Sitewide Groundwater Monitoring Report Mound, Ohio, Site

Calendar Year 2016

June 2017

U.S. DEPARTMENT OF ENERGY | Legacy Management

This page intentionally left blank

Contents

Abbreviations iii
1.0 Introduction 1
1.1 Purpose 1
1.2 Project Description 1
1.2.1 Phase I 3
1.2.2 Parcels 6, 7, and 8 3
1.3 Geology and Hydrology 3
2.0 Monitoring Programs 5
2.1 Phase I 5
2.1.1 Monitored Natural Attenuation of TCE 5
2.1.2 Triggers 7
2.2 Parcels 6, 7, and 8 7
2.2.1 Monitored Natural Attenuation of TCE and Tritium 8
2.2.2 Triggers 10
2.3 Monitoring Network 10
2.4 Deviations from the Sitewide Operations and Maintenance Plan 10
2.5 Trend Analysis Methodology 11
3.0 Phase I MNA Remedy 13
3.1 Monitoring Results 13
3.2 Trend Analysis 15
3.3 Groundwater Elevations 17
3.4 Data Evaluation 17
3.5 Recommendations 17
4.0 Parcels 6, 7, and 8 MNA Remedy 19
4.1 Monitoring Results—VOCs 19
4.1.1 Seeps 19
4.1.2 Groundwater 21
4.1.3 Distribution of TCE 23
4.2 Monitoring Results—Tritium 25
4.3 Trend Analysis 27
4.3.1 VOCs 29
4.3.2 Tritium 30
4.4 Groundwater Elevations 30
4.5 Data Evaluation 32
4.5.1 VOCs 32
4.5.2 Tritium 32
4.6 Recommendations 35
5.0 Inspection of the Monitoring System 37
6.0 Data Validation 39
7.0 References 41

Figures

Figure 1. Mound, Ohio, Site 2
Figure 2. Generalized Cross Section Showing Flow from Bedrock to the BVA 4
Figure 3. Phase I MNA Remedy Monitoring Locations 6
Figure 4. Parcels 6, 7, and 8 Remedy Monitoring Locations 9
Figure 5. TCE Concentrations in Phase I—1999 Through 2016 14
Figure 6. cis-1,2-DCE Concentrations in Phase I—1999 Through 2016 15
Figure 7. 2016 Annual Averages for TCE and DCE in Phase I 16
Figure 8. 2016 Average Groundwater Elevations in Phase I 18
Figure 9. TCE Concentrations in Parcels 6, 7, and 8 Main Hill Seeps 20
Figure 10. PCE Concentrations in Parcels 6, 7, and 8 Seep 0601 21
Figure 11. TCE Concentrations in Parcels 6, 7, and 8 Groundwater 23
Figure 12. 2016 Annual Averages for TCE in Parcels 6, 7, and 8 Seeps and Groundwater 24
Figure 13. Tritium Activity in Parcels 6, 7, and 8 Main Hill Seeps 26
Figure 14. Tritium Activity in Parcels 6, 7, and 8 Wells 0138, 0346, 0347, and 0379 27
Figure 15. 2016 Annual Averages for Tritium in Parcels 6, 7, and 8 Seeps and Groundwater 28
Figure 16. 2016 Averages for Groundwater Elevations in Parcels 6, 7, and 8 31
Figure 17. Bedrock Topography in the Tributary Valley 33
Figure 18. Cross Section Through the Tributary Valley 34

Tables

Table 1. Remedy (MNA) Monitoring for Phase I 5
Table 2. Trigger Levels for Phase I MNA Remedy 7
Table 3. Monitoring for Parcels 6, 7, and 8 Area 8
Table 4. Trigger Levels for Parcels 6, 7, and 8 Monitoring Locations 10
Table 5. Summary of VOC Monitoring Results in Phase I for 2016 13
Table 6. Summary of Trend Analysis Results for TCE in Phase I for 2016 15
Table 7. Summary of VOC Results in the Main Hill Seeps for 2016 19
Table 8. Summary of VOC Results in Parcels 6, 7, and 8 Groundwater for 2016 22
Table 9. Summary of Tritium Results in the Main Hill Area for 2016 25
Table 10. Summary of Trend Analysis Results for VOCs in Parcels 6, 7, and 8 (2005–2016) ... 29
Table 11. Summary of Trend Analysis Results for VOCs in Parcels 6, 7, and 8 (2012–2016) ... 30
Table 12. Summary of Trend Analysis Results for Tritium in the Main Hill Seeps and Downgradient Wells (2005−2016) 30
Table 13. RINs for Calendar Year 2016 Sampling 39

Appendixes

Appendix A Well Construction Summary
Appendix B Mann-Kendall Trending Summaries
Appendix C 2016 Groundwater Elevations
Appendix D 2016 Groundwater and Seep Data Table
Appendix E Data Assessment Reports

Abbreviations

BVA	Buried Valley Aquifer
DCE	dichloroethene
DOE	U.S. Department of Energy
EPA	U.S. Environmental Protection Agency
MCL	maximum contaminant level
µg/L	micrograms per liter
MK	Mann-Kendall
MNA	monitored natural attenuation
nCi/L	nanocuries per liter
Ohio EPA	Ohio Environmental Protection Agency
OU	Operable Unit
PCE	tetrachloroethene (perchloroethene)
RAO	remedial action objective
RIN	report identification number
TCE	trichloroethene
VC	vinyl chloride
VOC	volatile organic compound
VSP	Visual Sample Plan

This page intentionally left blank

1.0 Introduction

1.1 Purpose

This report was prepared in support of the selected remedies for Phase I and Parcels 6, 7, and 8 of the Mound, Ohio, Site as outlined in the *Operations and Maintenance (O&M) Plan for the U.S. Department of Energy, Mound Site*, *Miamisburg, Ohio* (Sitewide Operations and Maintenance Plan) (DOE 2015). It summarizes the data collected in 2016 and documents the progress of the monitored natural attenuation (MNA) remedies for both areas of the Mound site. All sampling and data analyses were performed in accordance with the Sitewide Operations and Maintenance Plan, unless noted otherwise.

This report includes data collected during the groundwater sampling performed in 2016. Data are presented in both time-series and map-view plots. Trend analysis was performed on selected wells using the nonparametric Mann-Kendall (MK) test. This type of long-term trend analysis can be used to confirm trends in contaminant concentrations over time. The time-series plots will also be used to evaluate changes in data over time and to interpret the effectiveness of the MNA remedy.

This report also documents operational changes that occurred during the reporting period and identifies maintenance activities associated with the monitoring wells being sampled.

1.2 Project Description

The Mound site[1] is in Miamisburg, Ohio, approximately 10 miles southwest of Dayton. In 1995, the U.S. Department of Energy (DOE) Mound Plant, named after the Miamisburg Indian Mound adjacent to the site, comprised 120 buildings on 306 acres. The Great Miami River, west of the site, flows from northeast to southwest through Miamisburg and dominates the geography of the region surrounding the Mound site. Figure 1 shows the location of Phase I and Parcels 6, 7, and 8.

DOE remediated the Mound site to an "industrial use" standard consistent with the exposure assumptions provided in the *Mound 2000 Residual Risk Evaluation Methodology* (DOE 1997) and endorsed by the U.S. Environmental Protection Agency (EPA) and Ohio Environmental Protection Agency (Ohio EPA). The remedies for groundwater at the site combine groundwater monitoring and institutional controls in the form of deed restrictions on future land and groundwater use. These combined remedies will prevent current and future exposure of workers, the public, and the environment to contaminated groundwater from the Mound site.

[1] The Mound site has also been called the Mound Laboratory, Mound Laboratories, the Mound Plant (EPA ID OH6890008984), the USDOE Mound Plant, the Mound Facility, the USDOE Mound Facility, the Miamisburg Environmental Management Project (MEMP), and Miamisburg Closure Project (MCP). Currently, LM uses Mound, Ohio, Site as the formal name of the site.

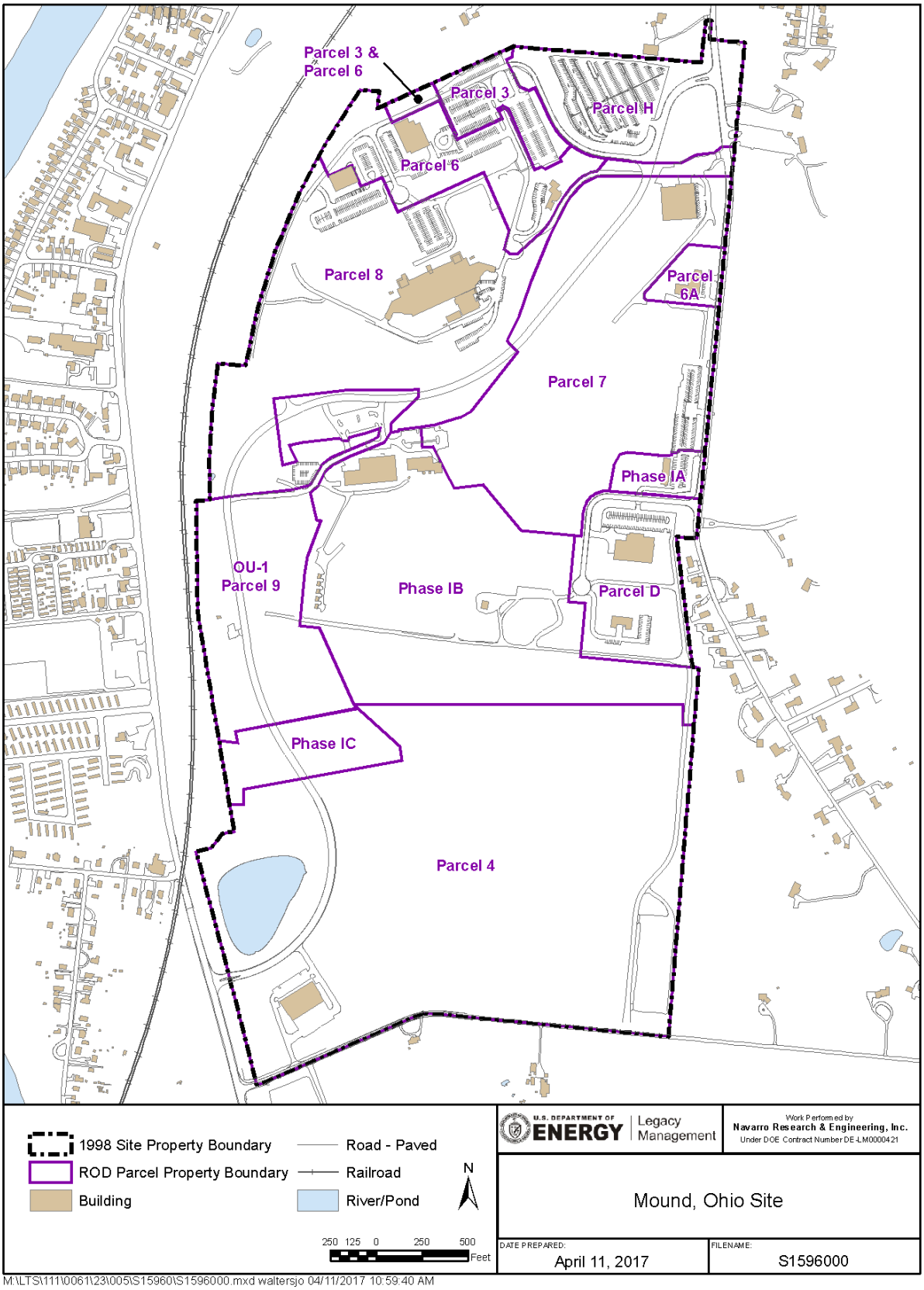


Figure 1. Mound, Ohio, Site

The long-term remedial action objective (RAO) for groundwater is to meet Safe Drinking Water Act maximum contaminant levels (MCLs) through MNA in the Phase I and Parcels 6, 7, and 8 areas. Until these goals are achieved, the near-term RAO is to prohibit the extraction and use of groundwater underlying the premises unless prior written approval is obtained from EPA, Ohio EPA, and Ohio Department of Health.

1.2.1 Phase I

Phase I is an approximately 52-acre area made up of three distinct sections. It lies on the southern border of the former production area of the Mound site. This area contains monitoring wells that are screened in both the Buried Valley Aquifer (BVA) and the upgradient bedrock aquifer system. MNA is being used as the remedy for a small, discrete section of the bedrock groundwater system contaminated with trichloroethene (TCE) to ensure that concentrations of TCE within the bedrock groundwater are decreasing to levels below the Safe Drinking Water Act MCL and do not impact the downgradient BVA.

1.2.2 Parcels 6, 7, and 8

Parcels 6, 7, and 8 occupy approximately 101 acres of the northern portion of the Mound Plant site. The main production facilities were located within Parcels 6 and 8, an area called the Main Hill. A tributary valley runs between these two parcels and Parcel 7; it contains a narrow tongue of glacial deposits that are in hydraulic communication with the BVA. Groundwater within the fractured bedrock beneath the Main Hill area, and in topographic highs within Parcel 7, flows along horizontal bedding planes and fractures and ultimately discharges to seeps or to the downgradient BVA.

Two monitoring wells on the western the edge of the BVA indicate volatile organic compound (VOC) impact, primarily TCE, that exceeds MCLs of the Safe Drinking Water Act. MNA is the remedy for the VOCs in groundwater associated with the Main Hill. Sampling is being performed to assess the contaminant concentrations and to verify that the BVA offsite and downgradient of these wells is not being adversely impacted.

Five seeps are also associated with this area and are located along the Main Hill of the plant property. Two seeps are within the plant property boundary, and the remaining three are offsite to the north. Several seeps have elevated levels of tritium and VOCs. These seeps, as well as several downgradient wells, are being monitored to verify that source removal (buildings and soil) on the Main Hill will result in decreasing concentrations over time.

1.3 Geology and Hydrology

The aquifer system at the Mound site consists of two distinct hydrogeologic environments: groundwater flow through the Ordovician shale and limestone bedrock beneath the hills, and groundwater flow within the unconsolidated glacial deposits and alluvium associated with the BVA in the Great Miami River valley. A thin tributary valley, which is located along the southern edge of the Main Hill, divides the two main portions of the Mound site and contains a narrow tongue of glacial deposits that are in hydraulic communication with the BVA.

The bedrock flow system is dominated by fracture flow and is not considered a highly productive aquifer. Groundwater flow in the bedrock typically mimics the topography, with groundwater discharging to the BVA or at seeps from the upper bedrock. The BVA is dominated by porous flow, with interbedded gravel deposits providing the major pathway for water movement. The unconsolidated deposits are Quaternary age sediments that consist of both glacial and fluvial deposits. The BVA is a highly productive aquifer capable of yielding a significant quantity of water. It is designated a sole-source aquifer. Groundwater flow in the BVA flows south, following the downstream course of the Great Miami River. The general structure and flow characteristics for these two interconnected systems are depicted on Figure 2.

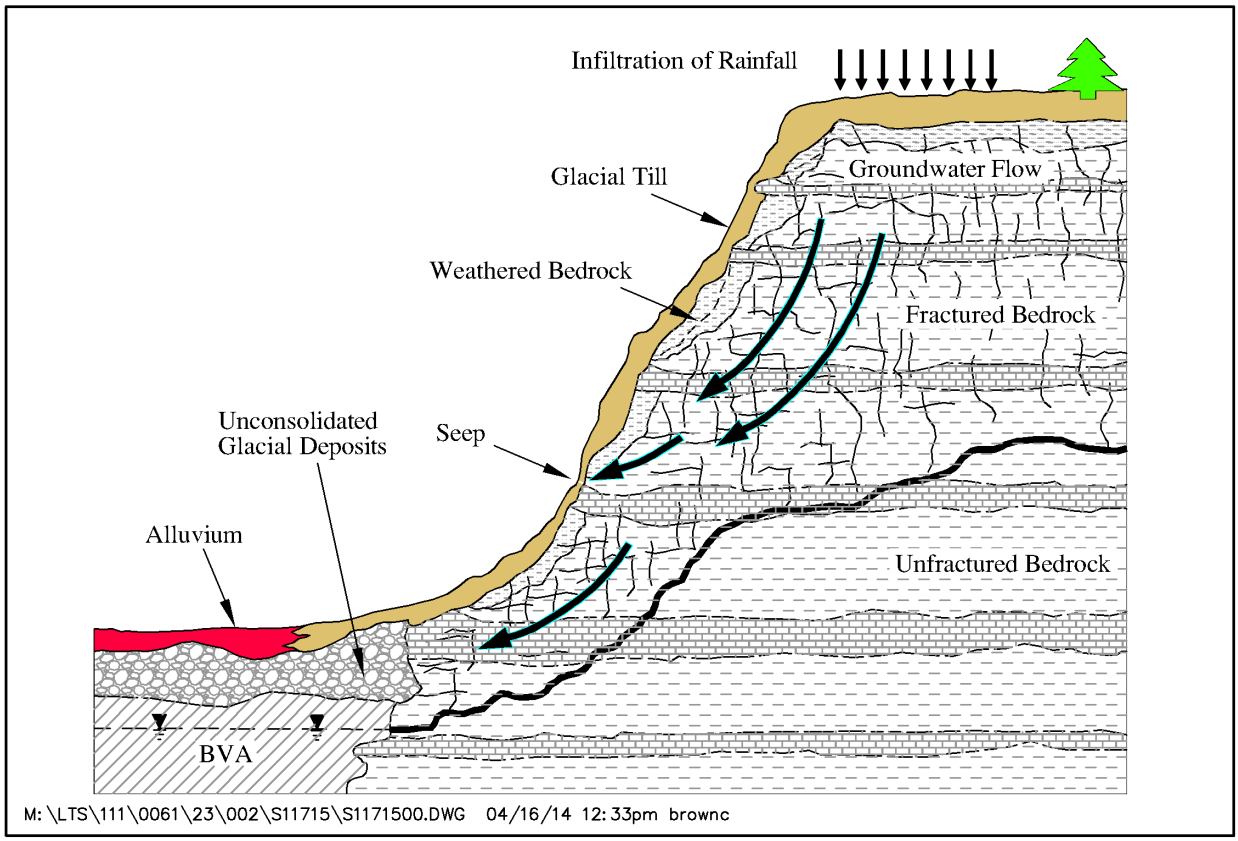


Figure 2. Generalized Cross Section Showing Flow from Bedrock to the BVA

For detailed descriptions of the geology, lithology, and groundwater flow regimes at the Mound site and specific hydrogeologic information for each area, refer to hydrogeologic investigation reports and work plans prepared for the site (DOE 1992, DOE 1994a, DOE 1994b, DOE 1995, and DOE 1999).

2.0 Monitoring Programs

2.1 Phase I

The Phase I groundwater monitoring program was established to verify that the BVA is not negatively affected by TCE-contaminated groundwater within the bedrock aquifer system. Groundwater in Phase I is monitored for TCE and its degradation products to verify that concentrations of TCE are decreasing by natural attenuation. The objective of this monitoring is to protect the BVA by verifying that the concentration of TCE near well 0411, well 0443, and seep 0617 is decreasing and to confirm that TCE is not adversely affecting the BVA.

2.1.1 Monitored Natural Attenuation of TCE

Under the Phase I MNA monitoring program, samples are collected semiannually from selected wells and a seep (Figure 3) and analyzed as outlined in Table 1. Sampling was performed in the first and third quarters of 2016.

Table 1. Remedy (MNA) Monitoring for Phase I

Monitoring Location	Area	Parameters
Well 0411	Well 0411 area	TCE Dichloroethene Vinyl chloride
Well 0443		
Well 0353	Bedrock monitoring	
Well 0444		
Well 0445		
Seep 0617		
Well 0400	BVA monitoring	
Well 0402		
Well P033		

Note:
All locations are sampled semiannually.

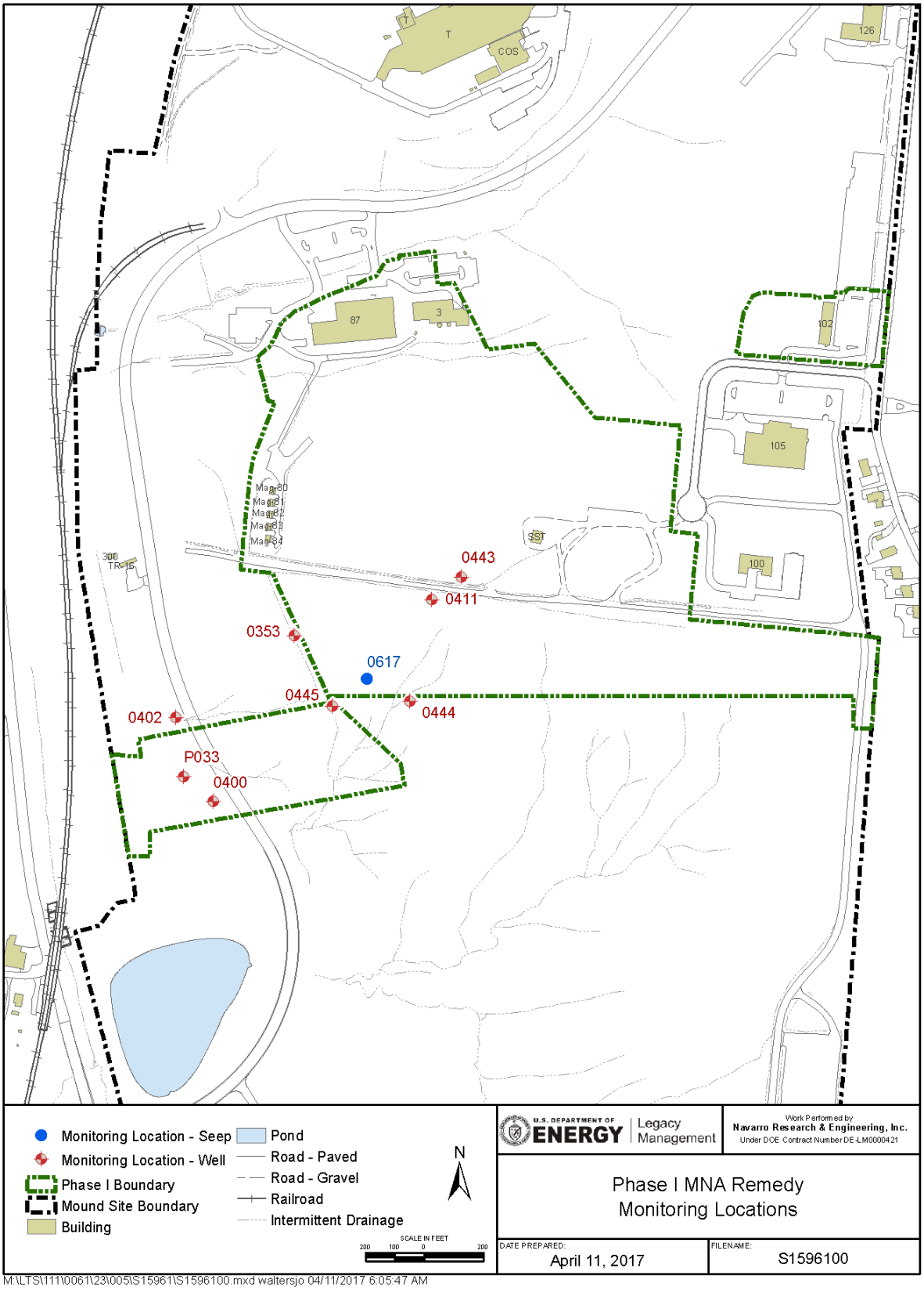


Figure 3. Phase I MNA Remedy Monitoring Locations

2.1.2 Triggers

The contaminant data are evaluated against previous data collected at each location to determine if MNA is adequately addressing groundwater impact and to monitor the geochemical conditions in the aquifer. Trigger levels and response actions have been established for each contaminant as presented in the Sitewide Operations and Maintenance Plan (DOE 2015). The triggers are summarized in Table 2.

Table 2. Trigger Levels for Phase I MNA Remedy

Location	TCE (µg/L)	DCE (µg/L)	Vinyl Chloride (µg/L)
0353	5	70	2
0400	5	70	2
0402	5	70	2
0411	30	70	2
0443	18	70	2
0444	5	70	2
0445	5	70	2
P033	5	70	2
0617 (seep)	16	70	2

Abbreviations:
DCE = dichloroethene
µg/L = micrograms per liter

EPA and Ohio EPA must be notified if trigger levels are exceeded. After notification, the Core Team (EPA, Ohio EPA, and DOE) will determine an appropriate course of action.

2.2 Parcels 6, 7, and 8

Groundwater in the Parcels 6, 7, and 8 area is monitored for TCE and its degradation products to verify that the downgradient BVA is not affected and that concentrations are decreasing. In addition, groundwater discharging from seeps is monitored for tritium and TCE and its degradation products to verify that source removal has resulted in decreasing concentrations over time.

The sampling program focuses on the following areas:

- **Well 0315/0347 Area:** Wells at the edge of the BVA on the southwestern corner of Parcel 8 that have elevated concentrations of VOCs. The program consists of wells that have TCE concentrations greater than the MCL and downgradient wells to the west. Wells 0315 and 0347 (source wells) and other selected downgradient BVA wells are monitored for VOCs—namely, tetrachloroethene (also known as perchloroethene or PCE), dichloroethene (DCE) isomers, TCE, and vinyl chloride (VC).
- **Main Hill Seeps:** Seeps on the northern and southern sides of the Main Hill that have elevated concentrations of VOCs and tritium. The program consists of seeps and downgradient wells to the west. Water from seeps 0601, 0602, 0605, 0606, and 0607 is

collected and analyzed for VOCs and tritium. Select wells within the BVA that are downgradient of the bedrock groundwater discharge area of the Main Hill are also sampled to monitor VOCs and tritium.

2.2.1 Monitored Natural Attenuation of TCE and Tritium

Under the Parcels 6, 7, and 8 MNA monitoring program, samples are collected quarterly for VOCs and semiannually for tritium in selected wells and seeps (Figure 4). Table 3 provides a summary of the monitoring locations as specified in the Sitewide Operations and Maintenance Plan.

Table 3. Monitoring for Parcels 6, 7, and 8 Area

Monitoring Location	Area	VOC	Tritium
Well 0315	Source wells	X	
Well 0347		X	
Well 0118	Downgradient BVA monitoring	X	X
Well 0124		X	
Well 0126		X	
Well 0138		X	X
Well 0346		X	X
Well 0379		X	X
Well 0386		X	
Well 0387		X	
Well 0389		X	
Well 0392		X	
Seep 0601	Main Hill seeps	X	X
Seep 0602		X	X
Seep 0605		X	X
Seep 0606		X	X
Seep 0607		X	X

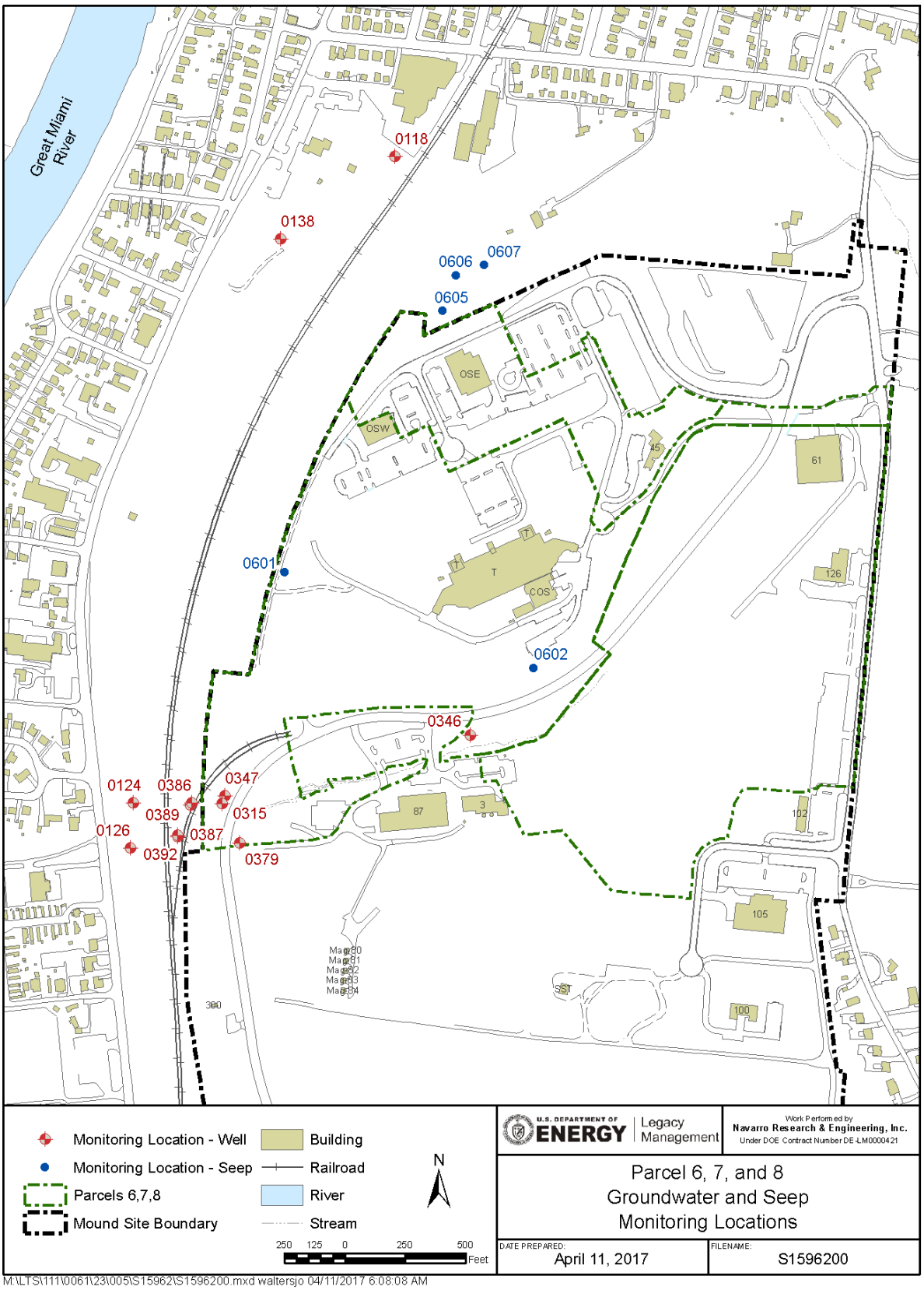


Figure 4. Parcels 6, 7, and 8 Remedy Monitoring Locations

2.2.2 Triggers

The contaminant data are evaluated against previous data collected at each location to determine if downward trends are occurring. Trigger levels and response actions have been established for each contaminant as presented in the Sitewide Operations and Maintenance Plan (DOE 2015). The triggers are summarized in Table 4.

Table 4. Trigger Levels for Parcels 6, 7, and 8 Monitoring Locations

Location	TCE (µg/L)	PCE (µg/L)	Tritium (nCi/L)
0315	30		
0347	30		
0124	5		
0126	5		
0386	5		
0387	5		
0389	5		
0392	5		
0601 (seep)		75	1500
0605 (seep)	150		

Abbreviations:
µg/L = micrograms per liter
nCi/L = nanocuries per liter

EPA and Ohio EPA must be notified if these trigger levels are exceeded. After notification, the Mound Core Team (EPA, Ohio EPA, and DOE) will determine an appropriate course of action.

2.3 Monitoring Network

The monitoring well and seep locations sampled under these programs were selected to provide data of sufficient quality to meet the objectives of the groundwater remedies for either Phase I or Parcels 6, 7, and 8. These wells were initially installed to support various site characterization activities and were designed and constructed to provide high-quality groundwater data. Appendix A contains construction information for each well used to support these remedies.

2.4 Deviations from the Sitewide Operations and Maintenance Plan

Sampling was performed as outlined in the Sitewide Operations and Maintenance Plan (DOE 2015), which compiles the sampling requirements outlined in previous plans for each area. Modifications to these monitoring programs (e.g., reduction in sampling frequency or discontinuation of monitoring locations) are also incorporated into the Sitewide Operations and Maintenance Plan.

Sampling was performed as follows:

- All required locations in Phase I were sampled in 2016.
- All required locations in Parcels 6, 7, and 8 were sampled in 2016 except seep 0602, which was dry during the third quarter. No sample was collected for this sampling event.
- Site-specific sampling methods for the Mound site were followed during these sampling events. These methods were developed by the Mound Groundwater Technical Team and approved by the Mound Core Team and are integrated into the Sitewide Operations and Maintenance Plan.

2.5 Trend Analysis Methodology

Groundwater data from select locations are evaluated for trends in contaminant concentrations over time. The computer program Visual Sample Plan (VSP), developed by Battelle Memorial Institute, was used to perform trend analysis; the method used was the Mann-Kendall test. Summary reports for each monitoring location are contained in Appendix B.

The Mann-Kendall test is a nonparametric statistical procedure that is appropriate for analyzing trends in data over time. There is no requirement that the data be normally distributed or that the trend, if present, is linear. The Mann-Kendall test can be used if there are missing values and values below the detection limit. The assumption of independence requires that the time between samples be sufficiently large so there is no correlation between measurements collected at different times. All locations were previously evaluated for seasonality as part of the annual review in 2014 (DOE 2015). Those results indicated there are no seasonal trends in contaminant data collected from any of the monitoring locations.

The Mann-Kendall procedure tests whether to reject the null hypothesis (H_0) and accept the alternative hypothesis (H_α), where:

- The null hypothesis (H_0) is that there is no monotonic trend in the series.
- The alternate hypothesis (H_α) is that a monotonic trend exists.

The initial assumption of the Mann-Kendall test is that the H_0 is true and that the data must be convincing beyond a reasonable doubt before H_0 is rejected and H_α is accepted. One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

The Mann-Kendall test analyzes differences in signs between earlier and later data points. The difference between the later-measured value and all earlier-measured values are computed and an integer value of 1, 0, or -1 (positive differences, no differences, and negative differences, respectively) is assigned. The Mann-Kendall statistic, S, is then computed as the sum of all of the integers and measures the trend in the data. Positive values indicate an increase in constituent concentrations over time, whereas negative values indicate a decrease in constituent concentrations over time. The strength of the trend is proportional to the magnitude of S (i.e., large magnitudes indicate a strong trend).

As part of this approach, a test statistic, Z_{MK}, was calculated using the Mann-Kendall statistic (S) and the variance of S. A positive value of Z would indicate that the data were skewed in an upward direction, and a negative value of Z would indicate that the data were skewed in a downward direction. The calculated Z_{MK} statistic was used to test whether to reject H_0 and accept H_α. The Z_{MK} statistic is compared to an established Z value that is based on user-specified false rejection rates (α). The α value used in these evaluations was 0.05; α is the tolerable probability that the Mann-Kendall test will falsely reject the null hypothesis. Then H_0 is rejected and H_α is accepted if $Z_{MK} \geq Z_{1-\alpha}$, where $Z_{1-\alpha}$, is the 100 $(1-\alpha)^{th}$ percentile of the standard normal distribution.

For this statistical evaluation, data sets were initially evaluated for downward trends. Those data sets that did not indicate a downward trend were then evaluated for upward trends. If neither of those evaluations rejected H_0, it was reported that no monotonic trend existed.

3.0 Phase I MNA Remedy

3.1 Monitoring Results

Monitoring results for 2016 (Table 5) continue to show low-level detections of TCE and *cis*-1,2-DCE, a TCE degradation product, in wells 0411 and 0443 and in seep 0617. All VOC concentrations were below the applicable trigger levels (Table 2). Concentrations of TCE in wells 0411 and 0443 and seep 0617 continue to exceed the MCL of 5 micrograms per liter (µg/L). No detectable concentrations of *trans*-1,2-DCE or VC were reported at these three monitoring locations. None of the downgradient wells indicated impact attributable to VOCs originating from the Phase I area. Concentrations of TCE and *cis*-1,2-DCE were reported in BVA wells 0400 and 0402 during the second half of 2016 and are attributable to VOC impact in Operable Unit (OU)-1, which is immediately upgradient of these wells. No detectable concentrations of *trans*-1,2-DCE or VC were reported in the BVA or bedrock wells.

Table 5. Summary of VOC Monitoring Results in Phase I for 2016

Well ID	Location	Parameter	First Semiannual Event	Second Semiannual Event
Source Area Wells and Seep				
0411	0411 Area	TCE (µg/L)	**11.9**	**11.7**
		cis-1,2-DCE (µg/L)	0.65 (J)	0.94 (J)
0443	0411 Area	TCE (µg/L)	**7.5**	**7.3**
		cis-1,2-DCE (µg/L)	0.23 (J)	0.33 (J)
0617	Seep/ Bedrock	TCE (µg/L)	**9.1**	3.5
		cis-1,2-DCE (µg/L)	1.8	1.1
Other Wells				
0353	Bedrock	TCE (µg/L)	ND (<1)	ND (<1)
		cis-1,2-DCE (µg/L)	ND (<1)	ND (<1)
0444	Bedrock	TCE (µg/L)	ND (<1)	ND (<1)
		cis-1,2-DCE (µg/L)	ND (<1)	ND (<1)
0445	Bedrock	TCE (µg/L)	ND (<1)	ND (<1)
		cis-1,2-DCE (µg/L)	ND (<1)	ND (<1)
0400	BVA	TCE (µg/L)	ND (<1)	ND (<1)
		cis-1,2-DCE (µg/L)	ND (<1)	1.1
0402	BVA	TCE (µg/L)	ND (<1)	1.0
		cis-1,2-DCE (µg/L)	ND (<1)	3.8
P033	BVA	TCE (µg/L)	ND (<1)	ND (<1)
		cis-1,2-DCE (µg/L)	ND (<1)	ND (<1)

Note:
Values in **bold** exceed the MCL of 5 µg/L for TCE.

Abbreviations:
J = estimated value less than the reporting limit
ND = not detected above reporting limit

TCE concentrations in well 0411 (Figure 5) decreased since monitoring began in 1999; however, since 2002, the concentrations of TCE in this well have ranged between 9 and 15 µg/L. Concentrations of TCE in well 0443 and seep 0617 have varied since monitoring of these locations started in 2002. Concentrations of TCE in well 0443 have been consistently greater

than the MCL since 2010. The time-concentration plots for well 0443 and seep 0617 indicate that concentrations vary and are less than those in well 0411.

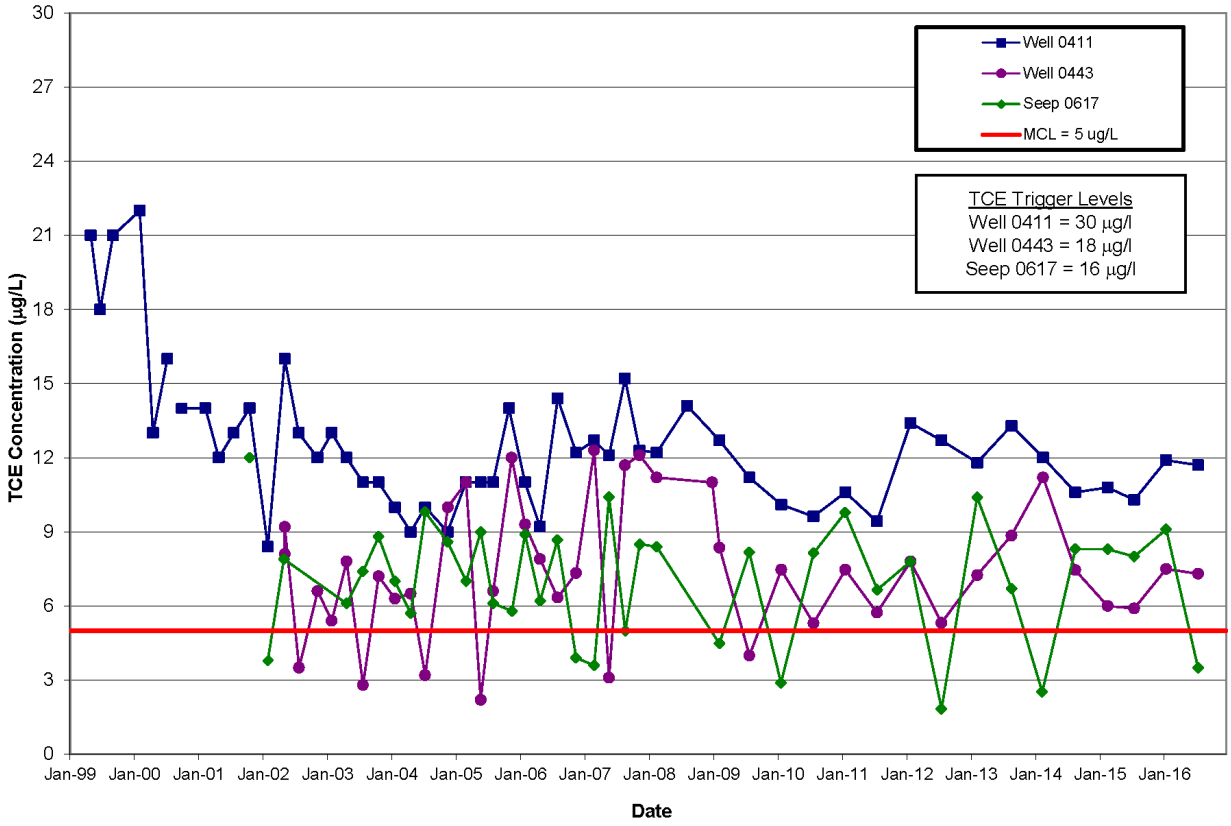


Figure 5. TCE Concentrations in Phase I—1999 Through 2016

The concentrations of *cis*-1,2-DCE in groundwater (Figure 6) have varied. Concentrations greater than the reporting limit of 1 µg/L have consistently been reported in well 0411 and seep 0617. Concentrations of *cis*-1,2-DCE in well 0411 are generally greater than those measured in seep 0617. Estimated detections less than 1 µg/L have been reported in well 0443 since 2009. None of the locations had concentrations of *cis*-1,2-DCE that exceeded the MCL of 70 µg/L.

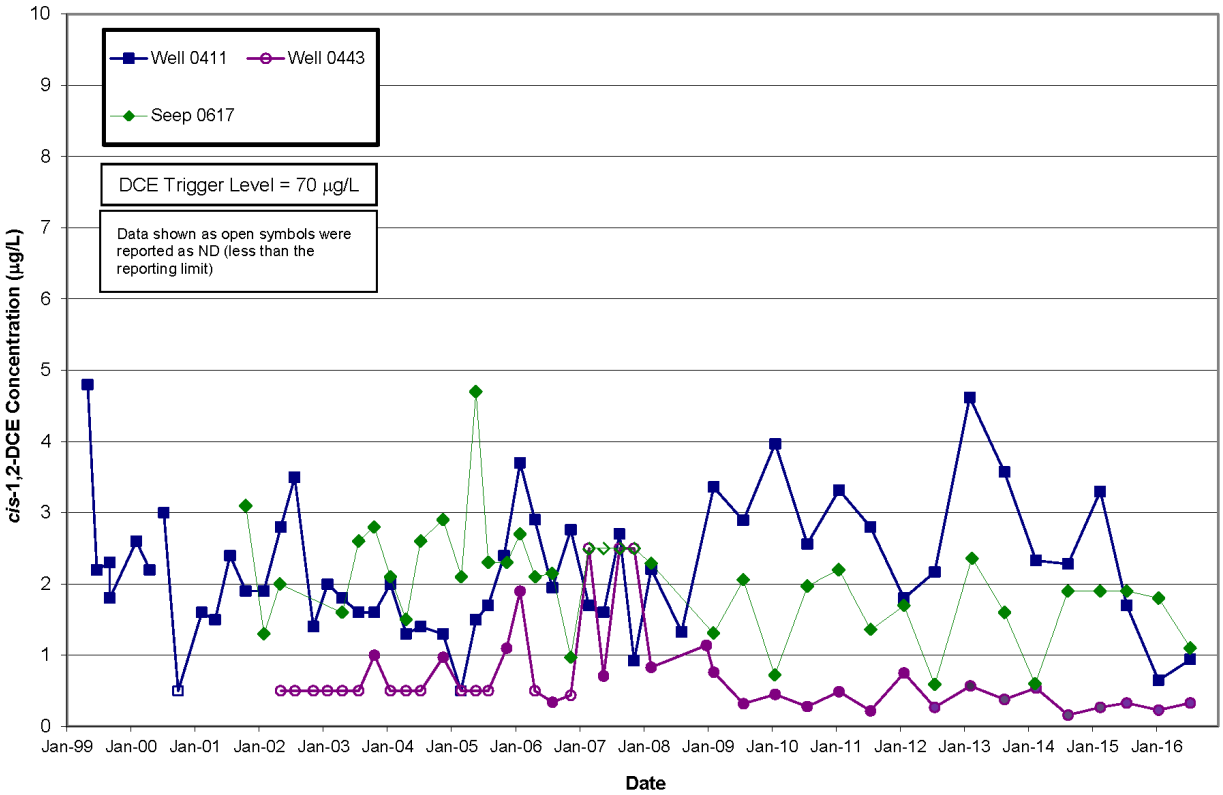


Figure 6. cis-1,2-DCE Concentrations in Phase I—1999 Through 2016

The distributions of TCE and DCE in groundwater (Figure 7) indicate that impact is localized in the bedrock groundwater near wells 0411 and 0443 and seep 0617. Wells screened in the bedrock and BVA do not have detectable concentrations of TCE or DCE, with the exception of wells 0400 and 0402. It has been determined that VOC impact in wells 0400 and 0402 is attributable to groundwater impact from OU-1, which is located immediately upgradient.

3.2 Trend Analysis

Trend analysis for TCE data collected since 1999 indicates downward trends for TCE in well 0411 and for *cis*-1,2-DCE in well 0443 and seep 0617 (Table 6). Summary reports providing details for each statistical evaluation for each monitoring location are contained in Appendix B.

Table 6. Summary of Trend Analysis Results for TCE in Phase I for 2016

Location	Analyte	No. of Samples	Trend
0411	TCE	53	**Down**
0443		41	None
0617		39	None
0411	*cis*-1,2-DCE	53	None
0443		41	**Down**
0617		39	**Down**

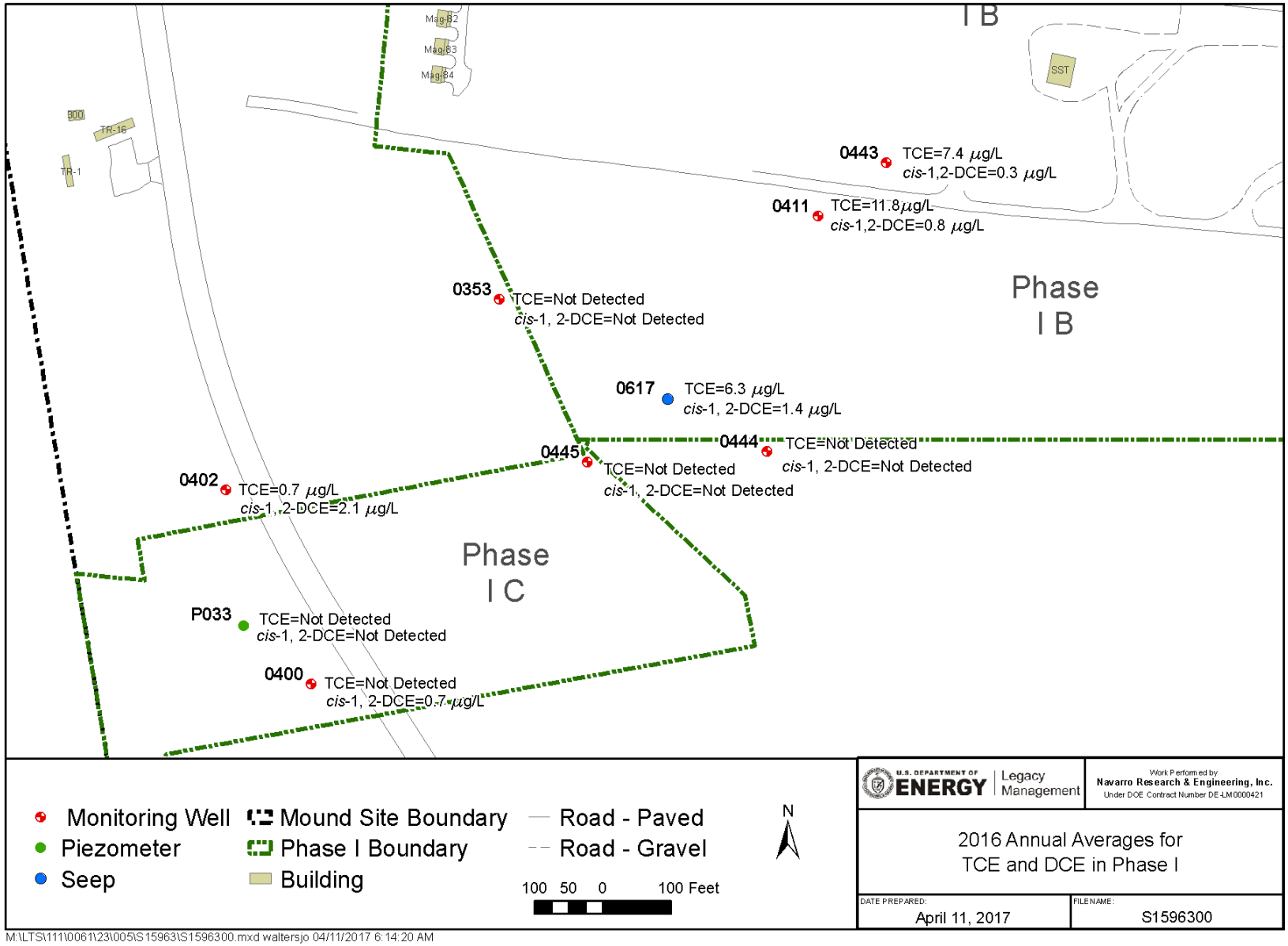


Figure 7. 2016 Annual Averages for TCE and DCE in Phase I

Evaluation of the downward trend in TCE concentrations in well 0411 may indicate the time frame when concentrations may approach the MCL of 5 µg/L. The nonparametric slope calculated for the trend analysis suggests that the MCL may be reached by 2041, which is similar to the estimated time frame from previous annual reports. The nonparametric analysis typically represents the decrease of contaminant concentrations in groundwater over time and provides good estimates of cleanup time frames.

3.3 Groundwater Elevations

A map of the average groundwater elevations measured in the Phase I area during 2016 (Figure 8) represents the two flow regimes present at the Mound site: bedrock and the unconsolidated materials of the BVA. The approximate location of contact of the BVA with the bedrock is indicated on this figure. Groundwater originating from the well 0411/0443 area flows southwest within the bedrock, following the bedrock topography. This groundwater enters the BVA along this contact. Flow within the BVA is generally to the south-southeast (parallel to the bedrock contact), although the groundwater elevations measured in the three wells screened in the BVA are similar. Appendix C presents a summary of the groundwater elevations measured during 2016.

3.4 Data Evaluation

The TCE and *cis*-1,2-DCE results for the groundwater sampling continues to show that VOC impact is localized in the bedrock groundwater near wells 0411 and 0443 and downgradient seep 0617. Concentrations of TCE at these three monitoring locations continue to exceed the MCL of 5 µg/L. TCE concentrations in well 0411, which has the highest concentrations in Phase I, have remained unchanged since 2009, ranging between 9 and 15 µg/L. Low levels of *cis*-1,2-DCE continue to be present at all three locations. Trend analysis continued to indicate a statistically significant downward trend in TCE concentrations in well 0411. Cleanup timeframe estimates based on the slope of the downward trend suggest that the MCL of 5 µg/L for TCE may be reached by 2041 in well 0411. Downward trends were calculated for *cis*-1,2-DCE in well 0443 and seep 0617. The wells located in the BVA do not show influence from localized TCE-impacted groundwater originating from the bedrock in Phase I.

3.5 Recommendations

No changes to the Phase I MNA monitoring program are warranted based on data collected in 2016. On the basis of no upward trends in TCE concentrations and the concentrations being considerably less than the trigger levels, monitoring frequency is recommended to remain semiannual for 2017. Sampling will continue during the first and third quarters of the year in an effort to bracket possible seasonal variations.

The Mound Core Team questioned whether adequate downgradient monitoring was being done based on groundwater flow maps developed for the Mound site, primarily the OU-1 area, which is immediately upgradient of the Phase I area. The Mound Core Team voiced concern that wells 0400, 0402, and P033, which are screened in the BVA, may be more cross gradient than downgradient of outflow from this area. The Mound Core Team requested that discrete groundwater samples be collected at locations along the interface of the bedrock and the BVA downgradient of wells 0411 and 0443 and seep 0617. This sampling will be performed and included in the monitoring report for calendar year 2017.

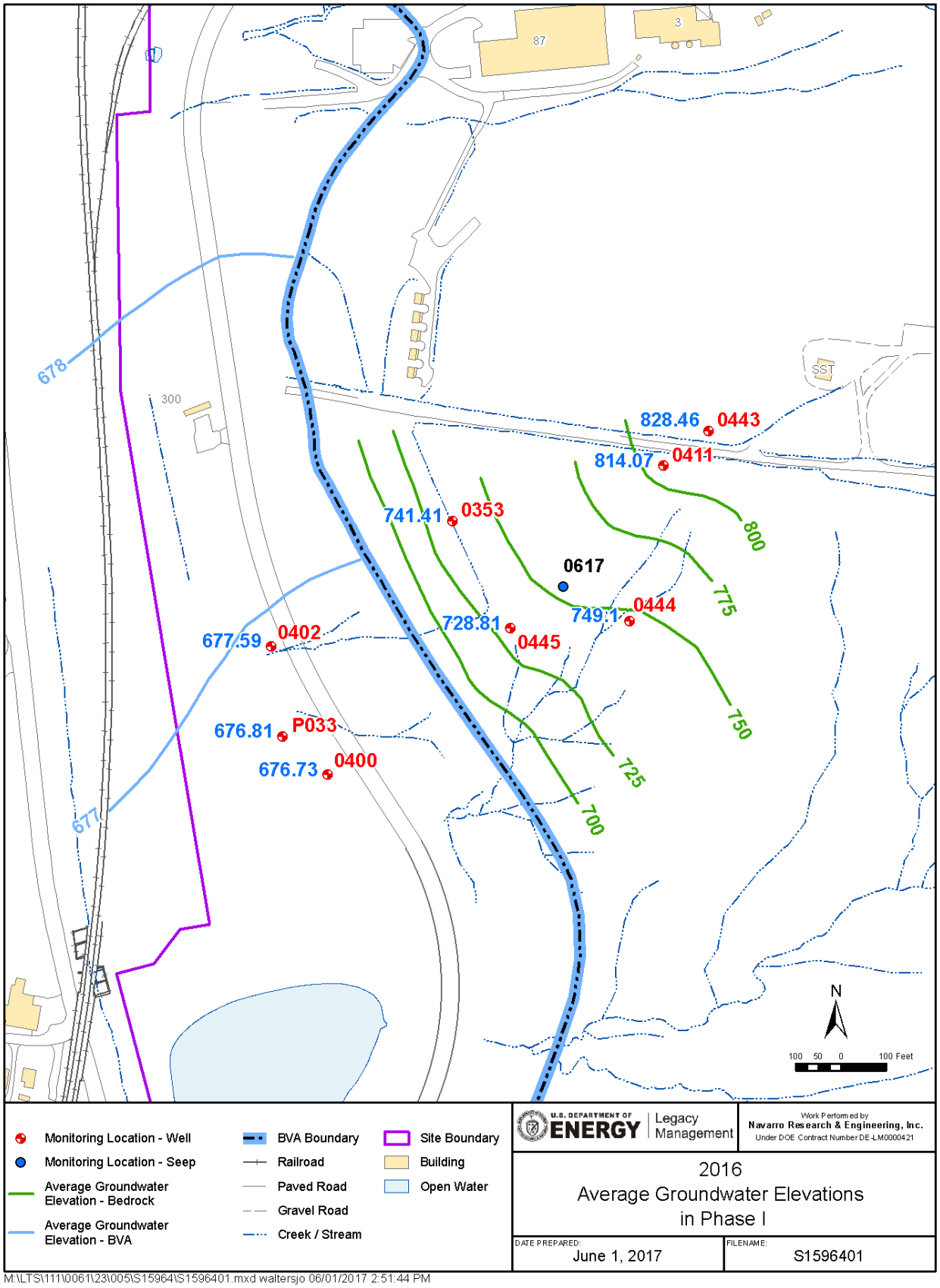


Figure 8. 2016 Average Groundwater Elevations in Phase I

4.0 Parcels 6, 7, and 8 MNA Remedy

4.1 Monitoring Results—VOCs

4.1.1 Seeps

Concentrations of TCE in all Main Hill seeps continued to exceed the MCL in 2016 (Table 7) with the exception of seep 0606. No locations had concentrations that exceeded the trigger level of 150 µg/L (established for seep 0605) (Table 4). The highest concentrations of TCE continued to be measured in seeps 0602 and 0605. PCE concentrations continued to exceed the MCL of 5 µg/L in seep 0601; however, concentrations at this location did not exceed the trigger level of 75 µg/L. Estimated detections of PCE were reported in seep 0605. *cis*-1,2-DCE was reported in seeps 0602 and 0605; seep 0602 had the highest concentrations. Estimated detections of *cis*-1,2-DCE (less than 1 µg/L) were reported in seeps 0601, 0606, and 0607. Estimated detections of *trans*-1,2-DCE (less than 1 µg/L) were reported in seep 0602 and 0605. No VC was detected in the seeps.

Table 7. Summary of VOC Results in the Main Hill Seeps for 2016

Location	Area	VOC Concentrations				
		VOC	Q1	Q2	Q3	Q4
Seeps						
0601	Onsite	PCE (µg/L)	**8.7**	**12.2**	**5.5**	**19.1**
		TCE (µg/L)	**6.5**	**7.3**	**5.3**	**7.8**
		cis-1,2-DCE (µg/L)	0.45 (J)	0.59 (J)	0.84 (J)	1.2
		trans-1,2-DCE (µg/L)	ND (<1)	ND (<1)	ND (<1)	ND (<1)
0602	Onsite	PCE (µg/L)	ND (<1)	ND (<1)	Dry	ND (<1)
		TCE (µg/L)	**12.8**	**6.7**		**10.0**
		cis-1,2-DCE (µg/L)	7.4	4.5		14.8
		trans-1,2-DCE (µg/L)	0.30 (J)	0.29 (J)		ND (<1)
0605	Offsite	PCE (µg/L)	0.21 (J)	0.33 (J)	ND (< 1)	ND (<1)
		TCE (µg/L)	**10.2**	**14.2**	**7.3**	**9.8**
		cis-1,2-DCE (µg/L)	1.21	2.0	1.0	**5.9**
		trans-1,2-DCE (µg/L)	0.18 (J)	0.21 (J)	ND (<1)	ND (<1)
0606	Offsite	PCE (µg/L)	ND (<1)	ND (<1)	ND (<1)	ND (<1)
		TCE (µg/L)	1.9	1.8	0.99 (J)	1.6
		cis-1,2-DCE (µg/L)	0.23 (J)	ND (<1)	0.30 (J)	1.3
		trans-1,2-DCE (µg/L)	ND (<1)	ND (<1)	ND (<1)	ND (<1)
0607	Offsite	PCE (µg/L)	ND (<1)	ND (<1)	ND (<1)	ND (<1)
		TCE (µg/L)	4.9	**7.3**	**6.9**	**5.3**
		cis-1,2-DCE (µg/L)	0.51 (J)	0.92 (J)	0.97 (J)	0.96 (J)
		trans-1,2-DCE (µg/L)	ND (<1)	ND (<1)	ND (<1)	ND (<1)

Notes:
"Dry" indicates no sample was collected because there was no flow during the scheduled sampling event.
PCE trigger level at 0601 = 75 µg/L.
TCE trigger level at the seeps = 150 µg/L.
Values in **bold** exceed the MCL.

Abbreviations:
J = estimated value that is less than the reporting limit
ND = not detected
Q = quarter

A graph of TCE concentrations measured in the seeps since the remediation of contaminated buildings and soil on the Main Hill (completed in mid-2006) (Figure 9) shows that the highest concentrations of TCE have been measured in seeps 0602 and 0605. In 2016, however, TCE concentrations decreased to about 10 µg/L. Concentrations of TCE have been variable in seep 0602, ranging from 15 µg/L to 139 µg/L. Concentrations of TCE in seep 0605 are relatively stable, and the remainder of the seeps follow a similar fluctuation. TCE concentrations in seep 0605 have been consistently less than 20 µg/L, and the concentrations in the remainder of the seeps have been less than 10 µg/L.

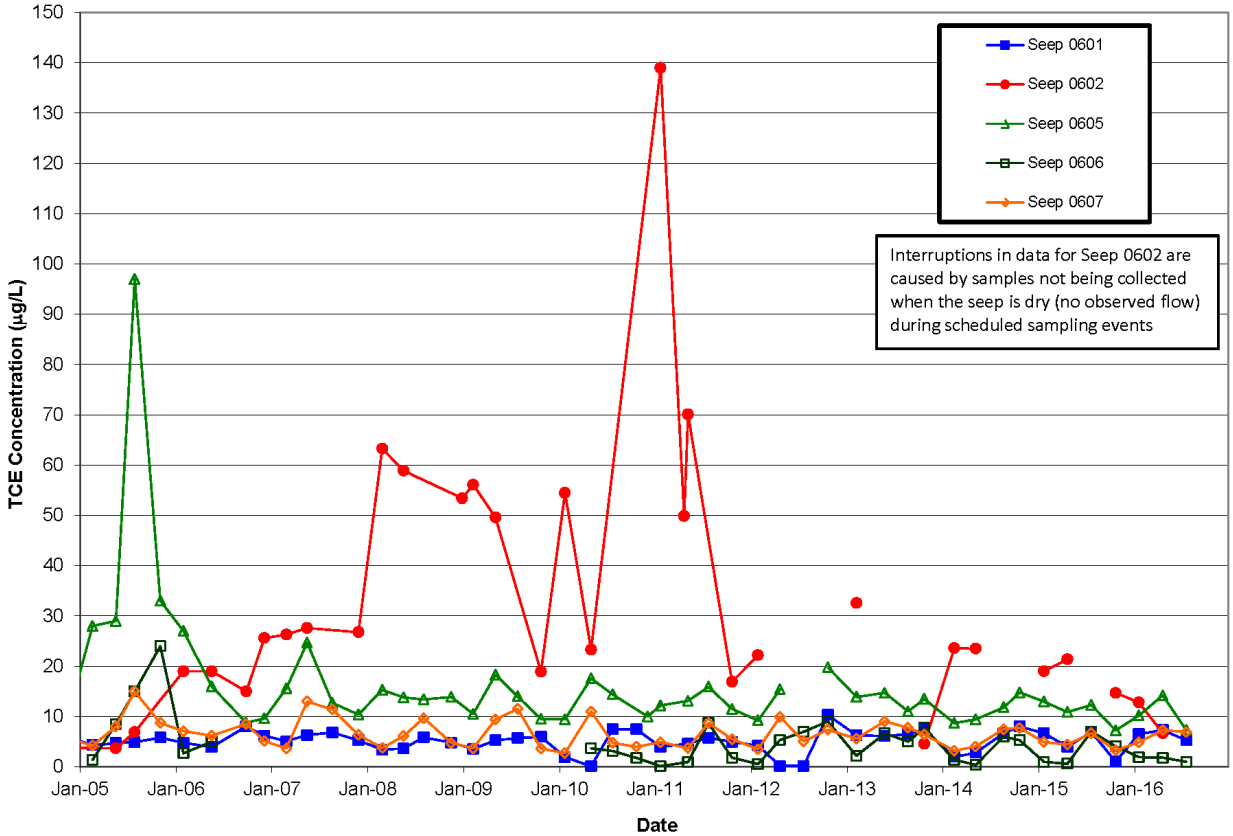


Figure 9. TCE Concentrations in Parcels 6, 7, and 8 Main Hill Seeps

Seep 0601 is the only location where detectable concentrations of PCE were reported. PCE concentrations in this seep (Figure 10) are similar to those measured before remediation on the Main Hill. Estimated detections of PCE (less than 1 µg/L) were reported in seep 0605.

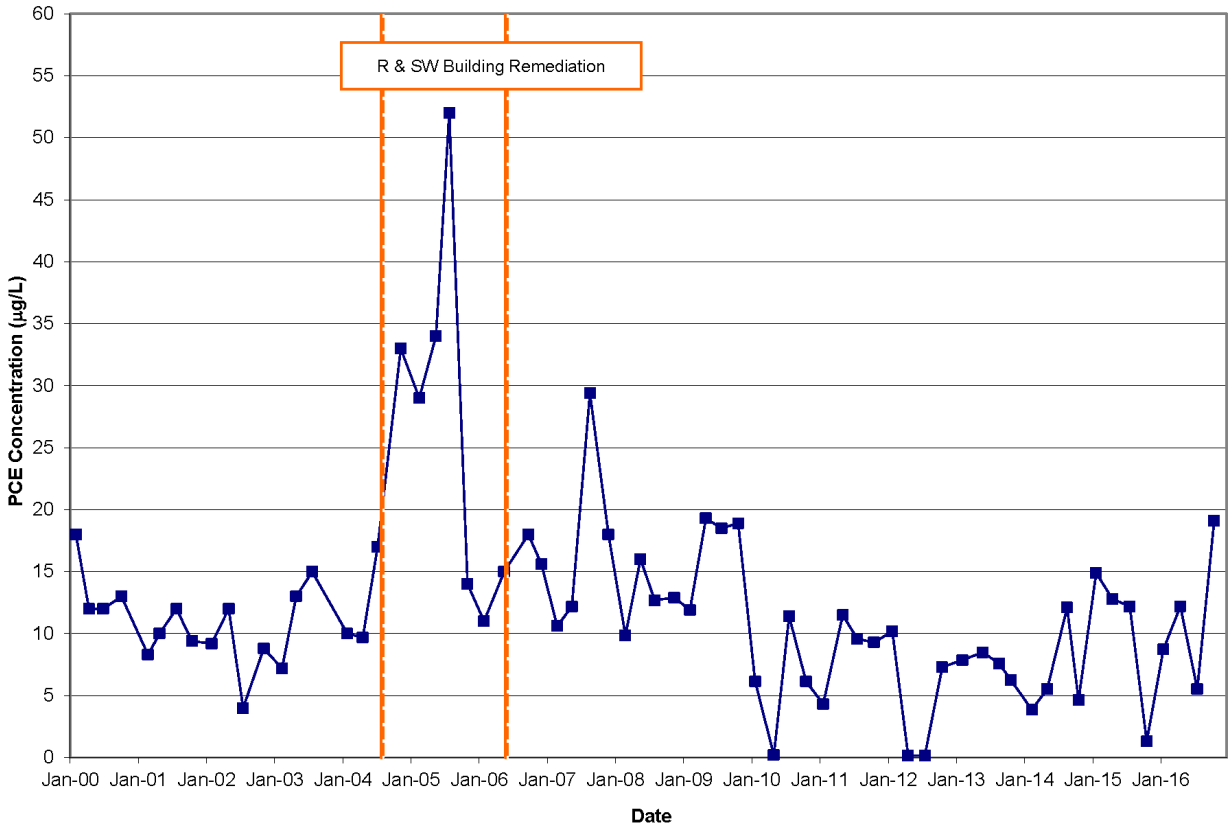


Figure 10. PCE Concentrations in Parcels 6, 7, and 8 Seep 0601

4.1.2 Groundwater

Monitoring results for 2016 (Table 8) continue to show TCE in wells 0315, 0347, 0379, and 0386; the highest concentrations are detected in wells 0315 and 0347 (source area wells), where concentrations also exceed the MCL. The concentrations of TCE reported in wells 0315 and 0347 were less than the trigger level of 30 µg/L established for these source area wells (Table 4). Well 0386 is downgradient of wells 0315 and 0347 just outside the Mound site boundary (Figure 4). Well 0379 is onsite within the tributary valley, where wells 0315 and 0347 are also located. Estimated detections of TCE were reported in well 0389. No detectable concentrations of TCE were measured in the other wells. All TCE concentrations were below applicable trigger levels.

Estimated detections of PCE less than 1 µg/L were reported in wells 0124, 0379, 0387, and 0392. A detectable concentration of 1 µg/L was reported in 0126 during the fourth quarter. All of these wells are located where the tributary valley enters into the BVA. No trigger levels for PCE have been set for these locations. No detectable concentrations of *cis*-1,2-DCE, *trans*-1,2-DCE, or VC were reported in any of the wells monitored as part of this program.

Table 8. Summary of VOC Results in Parcels 6, 7, and 8 Groundwater for 2016

Location	Area	VOC Concentrations				
		VOC	Q1	Q2	Q3	Q4
Onsite Wells						
0315	Source Area	PCE (µg/L)	ND (<1)	ND (<1)	ND (<1)	ND (<1)
		TCE (µg/L)	**3.8**	**5.7**	**6.8**	**7.5**
0347	Source Area	PCE (µg/L)	ND (<1)	ND (<1)	ND (<1)	ND (<1)
		TCE (µg/L)	**16.6**	**13.9**	**18.2**	**24.6**
0346	Onsite	PCE (µg/L)	ND (<1)	ND (<1)	ND (<1)	ND (<1)
		TCE (µg/L)	ND (<1)	ND (<1)	ND (<1)	ND (<1)
0379	Onsite	PCE (µg/L)	0.31 (J)	0.31 (J)	0.39 (J)	0.50 (J)
		TCE (µg/L)	1.6	1.5	1.7	1.7
Downgradient Wells—Near						
0386	BVA	PCE (µg/L)	ND (<1)	ND (<1)	ND (<1)	ND (<1)
		TCE (µg/L)	2.4	2.6	2.4	2.5
0387	BVA	PCE (µg/L)	0.20 (J)	0.23 (J)	ND (<1)	ND (<1)
		TCE (µg/L)	ND (<1)	ND (<1)	ND (<1)	ND (<1)
0389	BVA	PCE (µg/L)	ND (<1)	ND (<1)	ND (<1)	ND (<1)
		TCE (µg/L)	0.35 (J)	ND (<1)	0.80 (J)	0.74 (J)
0392	BVA	PCE (µg/L)	0.25 (J)	0.22 (J)	ND (<1)	ND (<1)
		TCE (µg/L)	ND (<1)	ND (<1)	ND (<1)	ND (<1)
Downgradient Wells—Far						
0118	BVA	PCE (µg/L)	ND (<1)	ND (<1)	ND (<1)	ND (<1)
		TCE (µg/L)	ND (<1)	ND (<1)	ND (<1)	ND (<1)
0124	BVA	PCE (µg/L)	ND (<1)	0.19 (J)	ND (<1)	ND (<1)
		TCE (µg/L)	ND (<1)	0.20 (J)	ND (<1)	ND (<1)
0126	BVA	PCE (µg/L)	0.77 (J)	0.82 (J)	0.97 (J)	1.0
		TCE (µg/L)	ND (<1)	ND (<1)	ND (<1)	ND (<1)
0138	BVA	PCE (µg/L)	ND (<1)	ND (<1)	ND (<1)	ND (<1)
		TCE (µg/L)	ND (<1)	ND (<1)	ND (<1)	ND (<1)

Notes:
TCE trigger level for 0315 and 0347 = 30 µg/L.
TCE trigger level for other wells = 5 µg/L.
Values in **bold** exceed the MCL.

Abbreviations:
J = estimated value that is less than the reporting limit
ND = not detected
Q = quarter

A graph of TCE concentrations measured in select wells shows that concentrations in wells 0315 and 0347 have consistently been greater than the MCL of 5 µg/L (Figure 11). The concentrations of TCE in the downgradient wells have been less than the MCL since 2000. The pattern in TCE concentrations in wells 0315 and 0347 has been similar since 2012. The concentrations in well 0347 are always higher and have greater changes (increases and decreases). An overall decrease in TCE concentrations can be observed beginning in the same time frame. It is possible

that surface water influences noted in previous reports (DOE 2014a; DOE 2014b) have been reduced and more recent data reflect concentrations of TCE in groundwater that are not influenced by external factors.

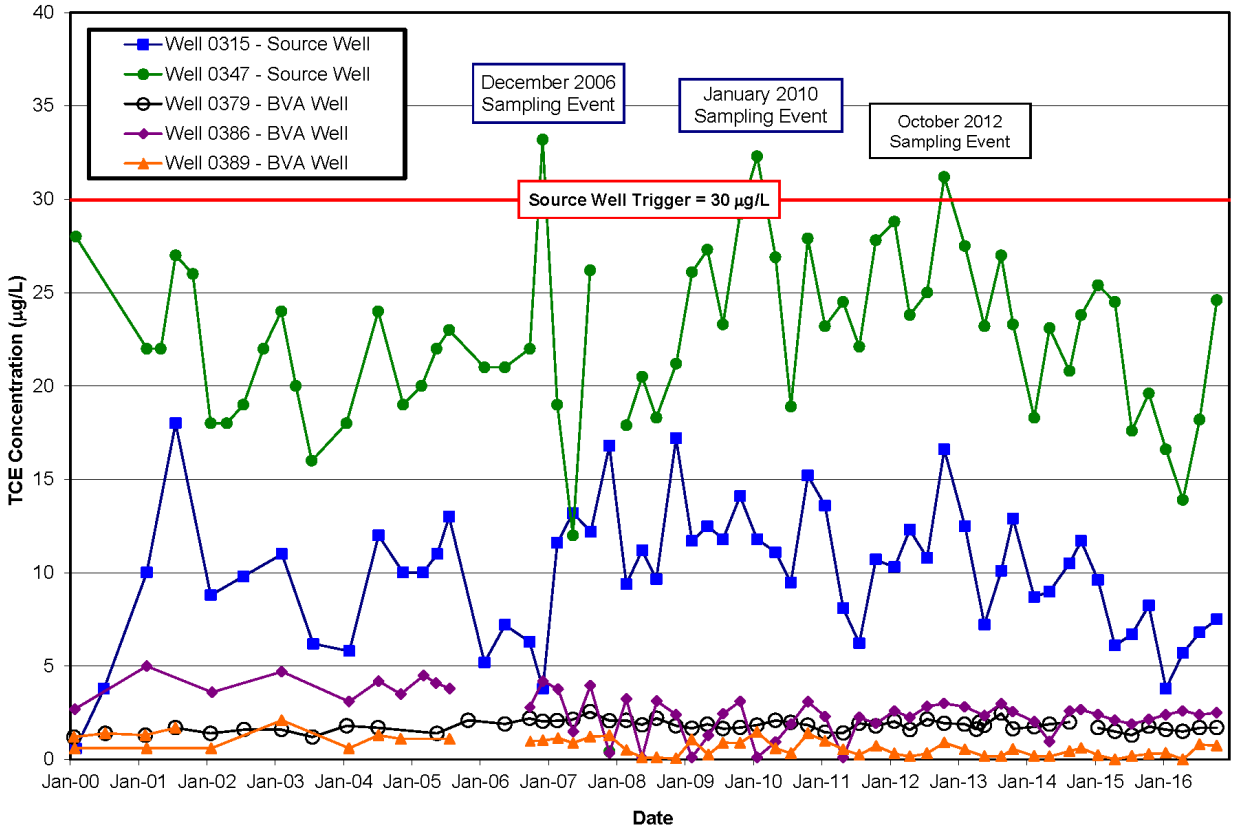


Figure 11. TCE Concentrations in Parcels 6, 7, and 8 Groundwater

4.1.3 Distribution of TCE

The distribution of TCE in groundwater (Figure 12) in the Main Hill area indicates that the highest area of impact is associated with wells 0315 and 0347. Concentrations of TCE in the seeps remain elevated. The BVA wells immediately downgradient of this area have TCE concentrations below the MCL. Figure 12 depicts the 2016 annual averages of TCE in the monitoring network.

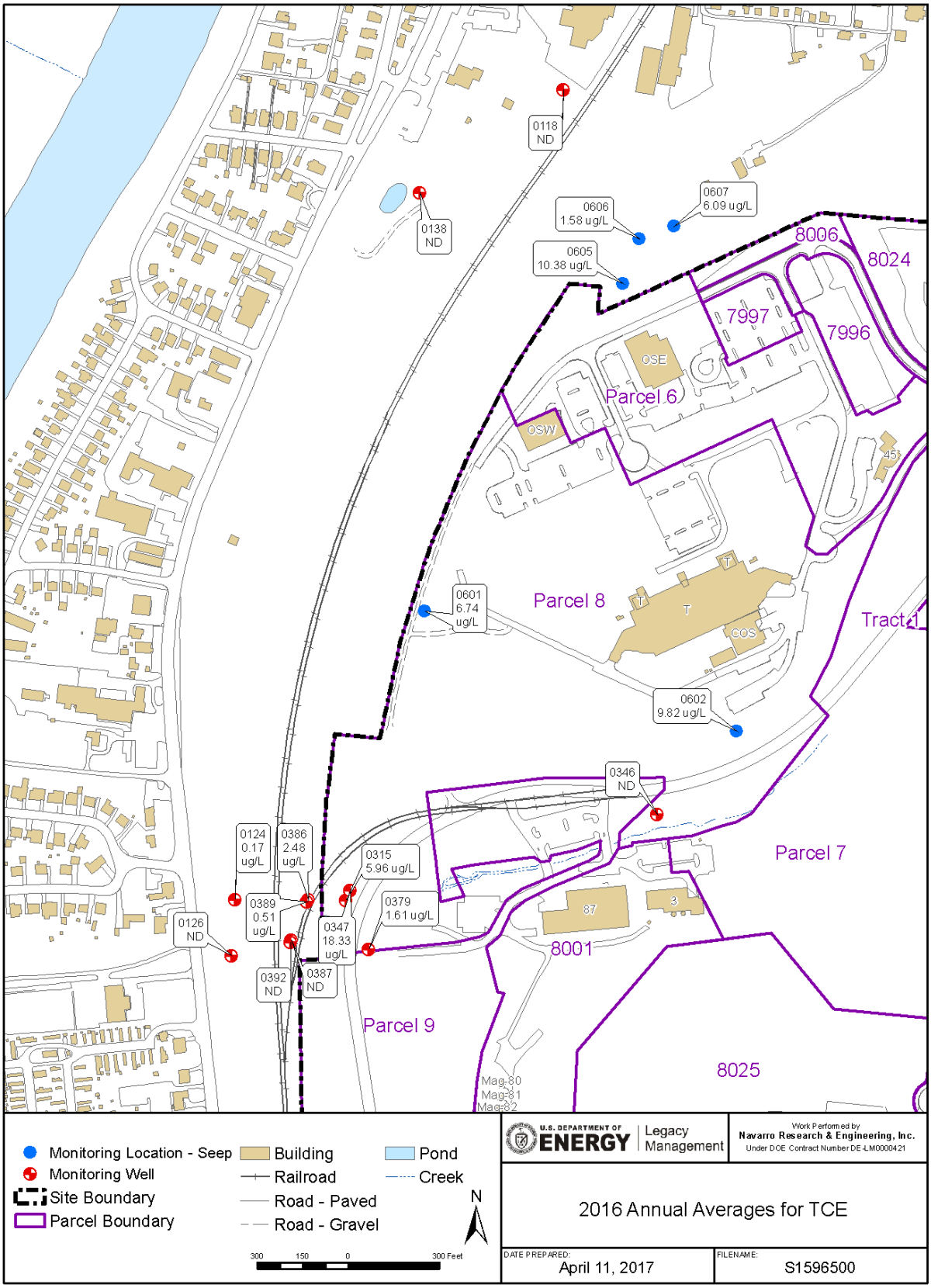


Figure 12. 2016 Annual Averages for TCE in Parcels 6, 7, and 8 Seeps and Groundwater

4.2 Monitoring Results—Tritium

Tritium levels in the Main Hill seeps continued to be elevated in 2016 and were higher than those in the downgradient groundwater wells (Table 9). The highest tritium activity continued to be measured in seep 0601, which is onsite. Seep 0601 is the only location that exceeded the MCL of 20 nanocuries per liter (nCi/L) in 2016. None of the seeps had tritium levels that exceeded the trigger level of 1500 nCi/L (Table 4). The wells had tritium levels similar to background (0.77 nCi/L [DOE 1996]).

Table 9. Summary of Tritium Results in the Main Hill Area for 2016

Location	Tritium Activity (nCi/L)	
	S1	S2
Seeps		
0601	**23.5**	**29.2**
0602	5.7	Dry
0605	7.4	4.4
0606	3.6	4.8
0607	3.7	3.1
Downgradient Wells		
0118	ND (<0.31)	ND (<0.34)
0138	0.36	0.43
0346	0.86	0.57
0347	1.4	1.6
0379	1.2	0.59

Notes:
"Dry" indicates no sample was collected because there was no flow during the scheduled sampling event.
Tritium trigger level at the seeps = 1500 nCi/L.
Values in **bold** exceed the MCL of 20 nCi/L.

Abbreviations:
J = estimated value that is less than the reporting limit
ND = not detected
S = semiannual

Tritium levels in the seeps were highest during remediation activities on the Main Hill (2004−2006). Tritium data collected after building demolition and soil removal indicate decreasing levels in all of the seeps (Figure 13). The decrease in tritium levels in postremediation data continues to support that the majority of the source was removed from the Main Hill area and that, with continued flushing, levels should continue to decline. Starting in 2009, the tritium levels in all of the seeps except seep 0601 were less than the MCL of 20 nCi/L.

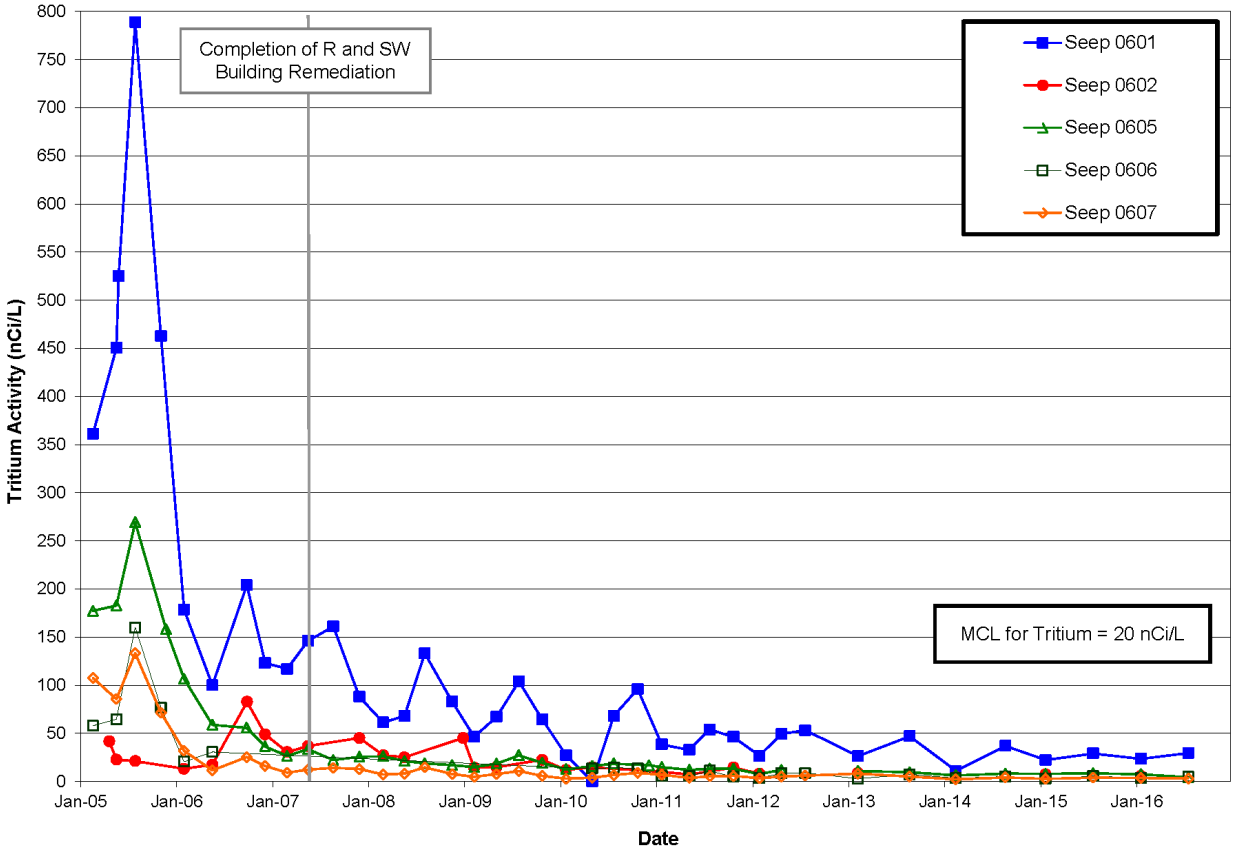


Figure 13. Tritium Activity in Parcels 6, 7, and 8 Main Hill Seeps

A graph of tritium levels in downgradient wells (Figure 14) illustrates that groundwater impact in the wells lagged behind impact expressed in the seeps. Groundwater impact increased near the end of remediation activities on the Main Hill, and impact in the seeps occurred as remediation activities were being performed and began to decrease as activities were completed. The tritium levels in the wells also responded quickly to remediation activities. Well 0347 has the highest levels of tritium. Tritium levels in wells 0138, 0346, and 0379 have leveled off and are similar to background. All tritium levels in the monitoring wells were below the MCL of 20 nCi/L.

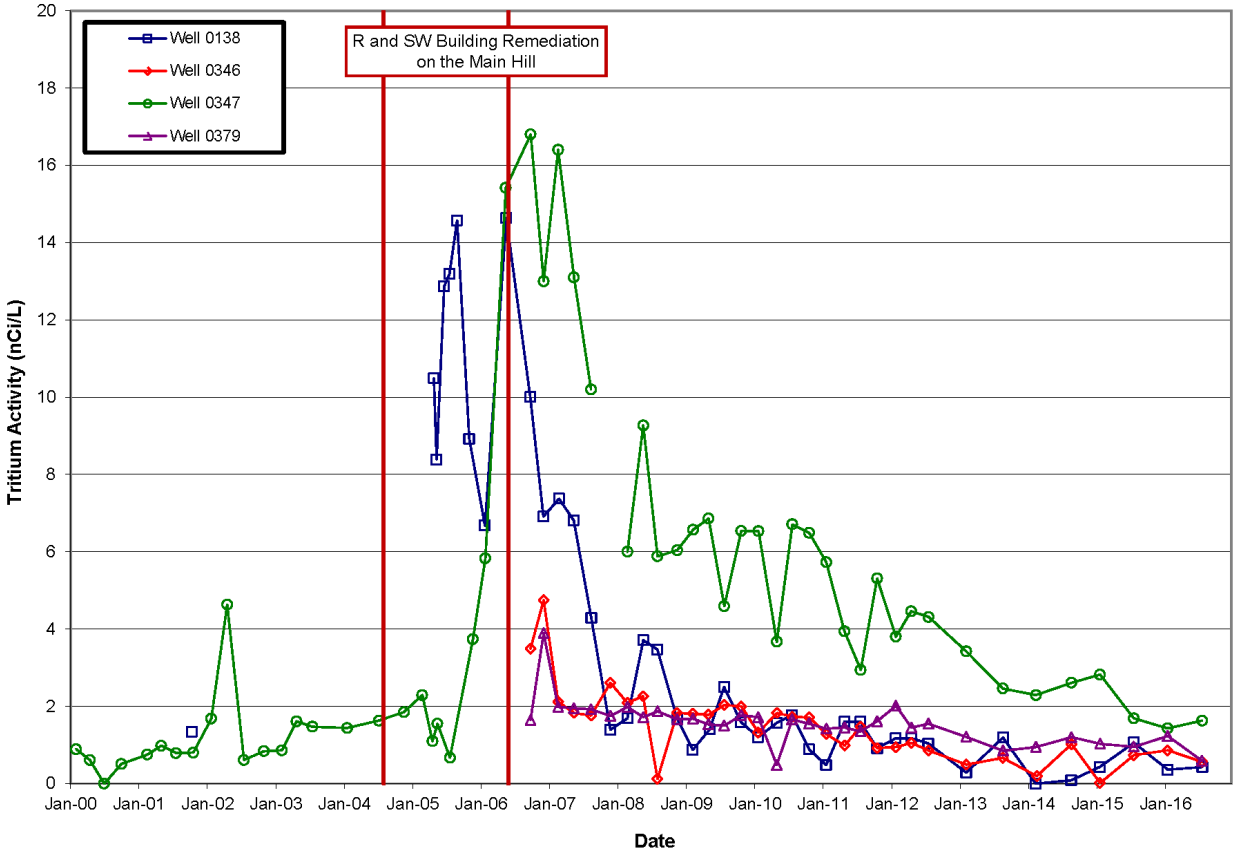


Figure 14. Tritium Activity in Parcels 6, 7, and 8 Wells 0138, 0346, 0347, and 0379

The distribution of tritium in groundwater (Figure 15) in the Main Hill area indicates that the greatest impact is still associated with the seeps, particularly seep 0601. Downgradient wells showed some elevated levels of tritium; however, levels are similar to background. Figure 15 depicts the 2016 annual averages of tritium in the monitoring network.

4.3 Trend Analysis

Trend analysis was performed on VOCs and tritium data using the nonparametric MK test. Trend analysis is reported for data collected since 2005. This period was selected to represent data collected since the completion of remediation activities on the Main Hill.

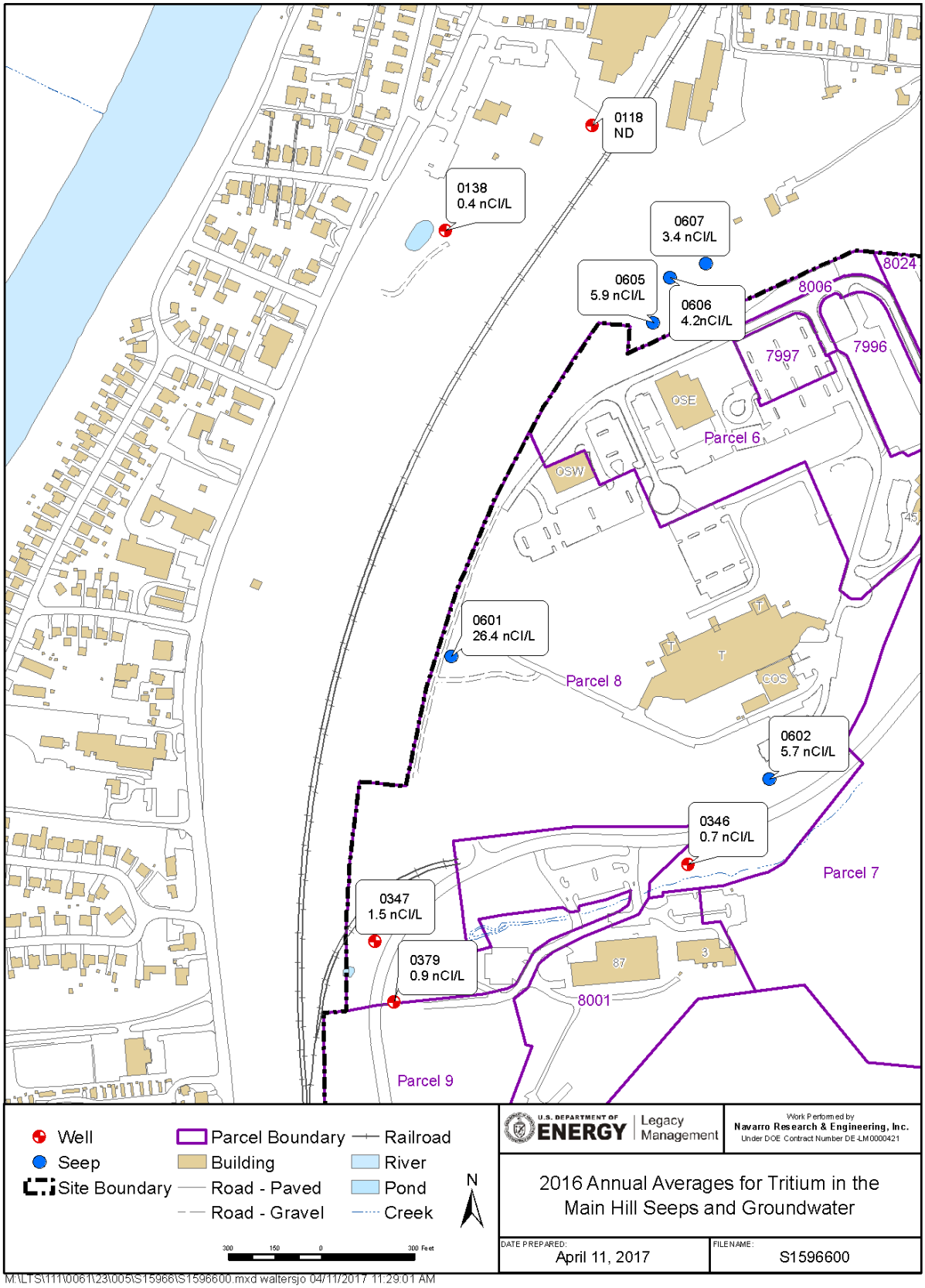


Figure 15. 2016 Annual Averages for Tritium in Parcels 6, 7, and 8 Seeps and Groundwater

4.3.1 VOCs

Trend analysis of TCE data collected since 2005 indicates statistically significant downward trends were calculated for seep 0605 and wells 0315, 0386, and 0389 (Table 10). Trend analysis was not performed on data from the remainder of the wells because results consistently showed nondetects or sporadic estimated detections. Summary reports providing details for each statistical evaluation for each monitoring location are in Appendix B.

Table 10. Summary of Trend Analysis Results for VOCs in Parcels 6, 7, and 8 (2005–2016)

Location	Number of Samples	Trend
TCE		
0601	48	None
0602	32	None
0605	47	**Down**
0606	33	None
0607	48	None
0315	47	**Down**
0347	47	None
0386	45	**Down**
0389	43	**Down**
PCE		
0601	48	**Down**
***cis*-1,2-DCE**		
0602	32	**Down**
0605	47	**Down**

Concentrations of PCE in seep 0601 were evaluated for a trend in PCE concentrations (Table 10). A statistical downward trend was indicated in the PCE data from this seep. Data from seeps 0602 and 0605 were evaluated for trends in *cis*-1,2-DCE concentrations (Table 10). Statistical downward trends were calculated for *cis*-1,2-DCE for both seeps.

A separate trend analysis of TCE data collected since 2012 was performed (Table 11). As previously noted, the influences of surface water entering the subsurface appear to have been reduced. Similar patterns in concentration changes have been observed in well 0315 and 0347 starting in 2012. Statistically significant downward trends were calculated for seeps 0602 and 0605 and wells 0315 and 0347.

Table 11. Summary of Trend Analysis Results for TCE in Parcels 6, 7, and 8 (2012–2016)

Location	Number of Samples	Trend
0601	20	None
0602	11	**Down**
0605	19	**Down**
0606	20	None
0607	20	None
0315	20	**Down**
0347	20	**Down**
0386	20	None
0389	20	None

4.3.2 Tritium

Trend analysis for tritium data collected since 2005 was performed for all of the seeps and wells where detectable levels have been consistently measured. Statistically significant downward trends in tritium were calculated in all of the seeps and wells (Table 12). Summary reports providing details for each statistical evaluation for each monitoring location are in Appendix B.

Table 12. Summary of Trend Analysis Results for Tritium in the Main Hill Seeps and Downgradient Wells (2005−2016)

Location	Number of Samples	Trend
0601	40	**Down**
0602	26	**Down**
0605	39	**Down**
0606	24	**Down**
0607	39	**Down**
0138	41	**Down**
0346	43	**Down**
0347	40	**Down**
0379	38	**Down**

4.4 Groundwater Elevations

A map of the average groundwater elevations measured in the Parcels 6, 7, and 8 area during 2016 (Figure 16) represents the two flow regimes present at the Mound site: bedrock and the unconsolidated materials of the BVA. The approximate location of contact of the BVA with the bedrock is indicated on this figure. Groundwater originating from the well Main Hill area flows within the bedrock, following the bedrock topography. This groundwater enters the BVA along this contact. Flow within the BVA is parallel to the bedrock contact. Appendix C presents a summary of the groundwater elevations measured during 2016.

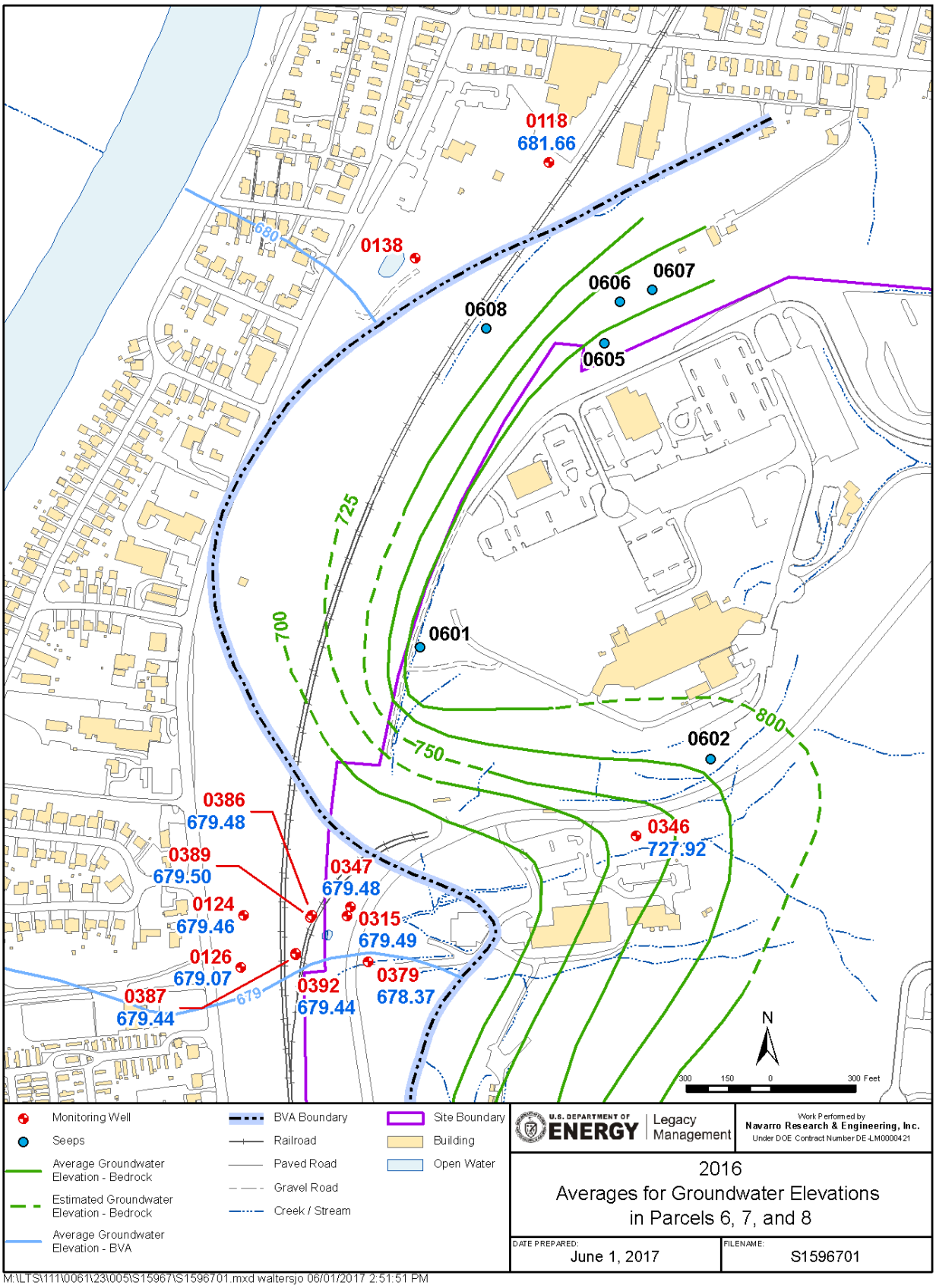


Figure 16. 2016 Averages for Groundwater Elevations in Parcels 6, 7, and 8

4.5 Data Evaluation

4.5.1 VOCs

Concentrations of TCE in the Main Hill seeps and source wells continued to exceed the MCL in 2016. The highest concentrations were measured in seeps 0602 and 0605 and wells 0315 and 0347. None of these values exceeded trigger levels. Detectable concentrations of the degradation products *cis*-1,2-DCE and *trans*-1,2-DCE were reported in seeps 0601, 0602, 0605, and 0606. Concentrations of VOCs were only detected in two of the far downgradient wells (see Table 8) at concentrations less than or equal to 1 µg/L.

Data collected over the past several years indicate variable concentrations of VOCs, primarily TCE, in the groundwater in Parcels 6, 7, and 8 as exhibited from the data from seep 0602 (Figure 9) and wells 0315 and 0347 (Figure 11). Seep 0602 and the downgradient wells 0315 and 0347 are in the tributary valley, which is along the southern edge of the Main Hill. As discussed in Section 1.3, the tributary valley is a narrow tongue of glacial deposits connected to the BVA that overlies the fractured bedrock at the site. Water infiltrating on the Main Hill moves through the fractured bedrock and ultimately discharges into the unconsolidated materials. Figure 17 depicts the bedrock topography beneath the tributary valley. Groundwater flow within the bedrock mimics the bedrock topography. TCE-impacted groundwater that originated on the Main Hill may move southward and discharge to seeps or the tributary valley. Seep 0602 is along the northern side of the tributary valley, and the wells are along the center of the valley. Figure 18 depicts the cross section along the transect from well pair 0315/0347 within the tributary valley to well 0126 in the BVA. Annual average TCE concentrations posted on the cross section show that the deep wells that are screened directly above the bedrock have the highest TCE concentrations. It is likely that these wells monitor the TCE-impacted groundwater discharging from the bedrock.

Trend analysis of TCE data collected since 2012 indicated downward trends in TCE concentrations in seeps 0602 and 0605 and wells 0315 and 0347. These decreases may be attributable to previous efforts that have been made to reduce the impact of surface water entering the subsurface (DOE 2014a; DOE 2014b).

4.5.2 Tritium

Tritium levels in the Main Hill seeps continued to be higher than those in the downgradient groundwater wells. The highest tritium activity was observed in seep 0601, which is onsite; this is the only location that exceeded the MCL of 20 nCi/L. Detectable levels of tritium were measured in four wells (0138, 0346, 0347, and 0379) downgradient of the seeps; however, most of the levels were similar to background. None of the groundwater wells had tritium levels that exceeded the MCL of 20 nCi/L.

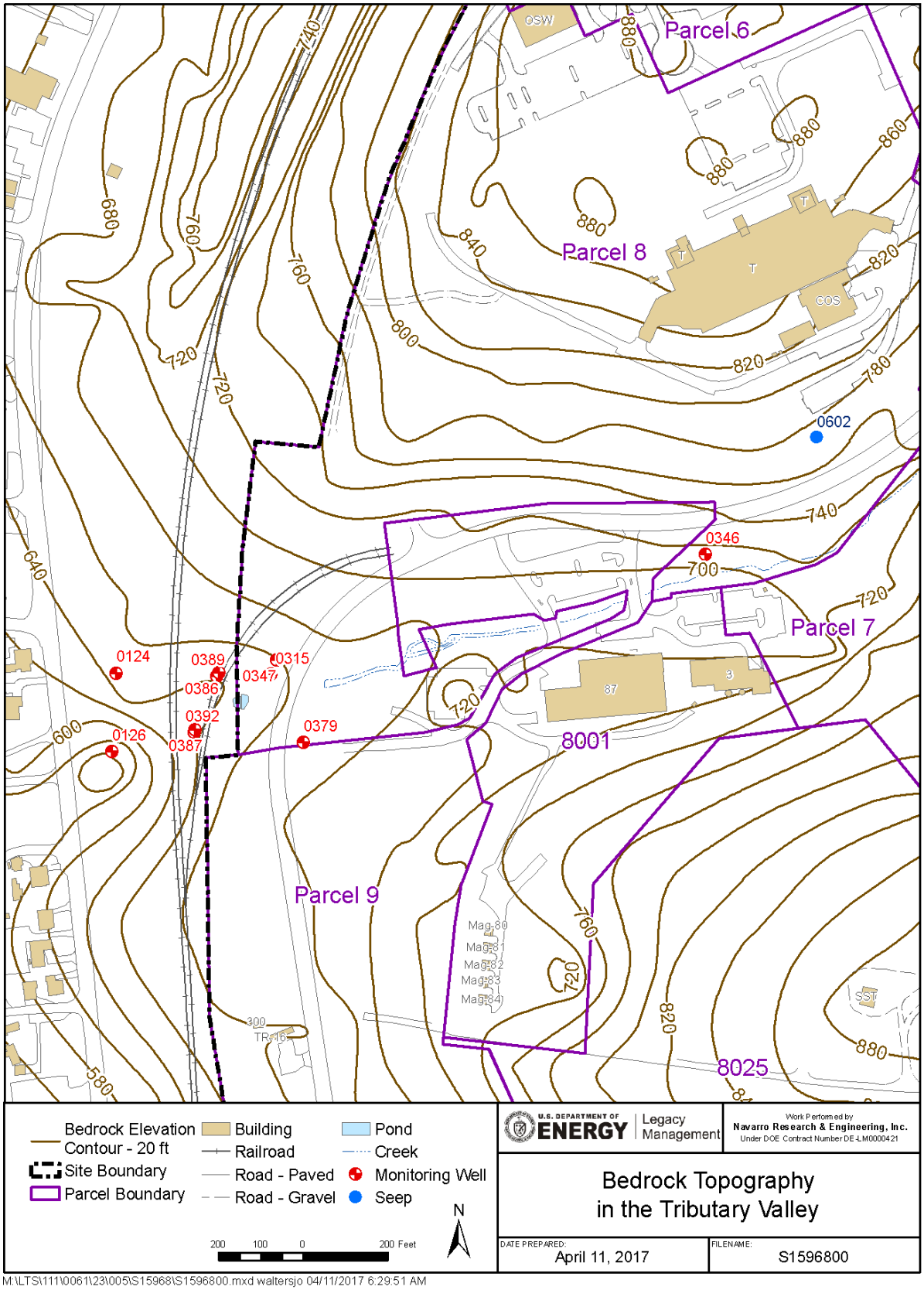


Figure 17. Bedrock Topography in the Tributary Valley

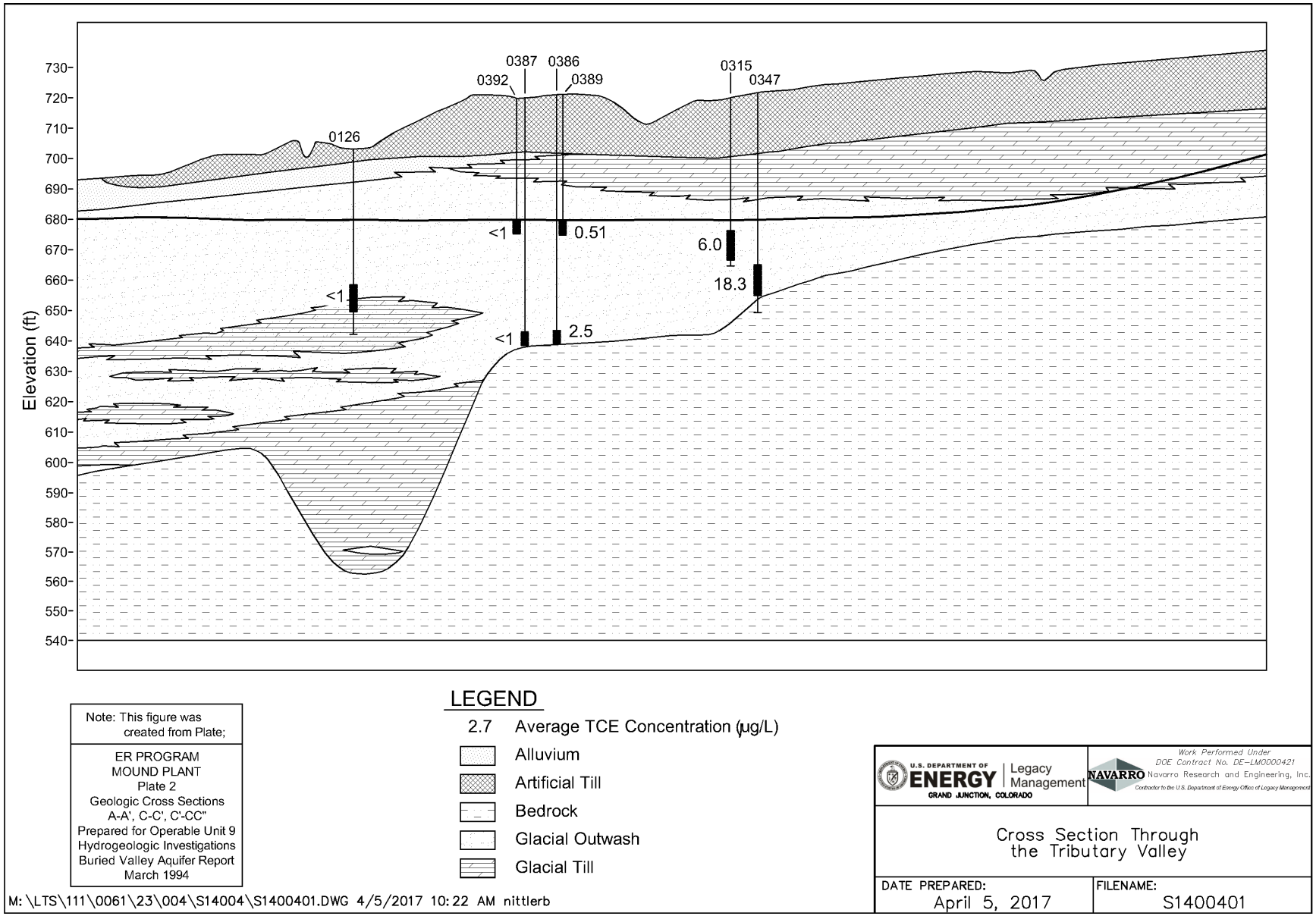


Figure 18. Cross Section Through the Tributary Valley

Trend analysis of tritium data collected since 2005 indicated statistically significant downward trends for all of the seeps and wells. The downward trends determined from postremediation data continue to support the interpretation that the majority of the source was removed from the Main Hill area during remediation and that flushing should continue to lower the levels.

4.6 Recommendations

The evaluation of the 2016 data does not indicate that the VOC monitoring program should be changed. TCE concentrations greater than the MCL continued to be measured in several seeps and in downgradient monitoring wells. The concentrations of VOCs continue to be variable at a few locations, although recent data (since 2012) indicate decreasing VOC concentrations at many locations. Quarterly sampling will continue at the seep and monitoring well locations in 2017 to determine if the system may be stabilizing since efforts were taken to reduce surface water infiltration into the subsurface (DOE 2014a; DOE 2014b).

No changes to the tritium monitoring program are warranted at this time; semiannual sampling for tritium will continue in 2017. Samples continue to be collected during the first and third quarters of the year to capture seasonal variation in tritium levels.

This page intentionally left blank

5.0 Inspection of the Monitoring System

A routine maintenance program has been established for the long-term groundwater monitoring locations at the Mound site. This program includes periodic inspections that focus on the integrity of each well and the condition of the protective casing and surface pad, the surrounding area, and the route of access. These inspections are usually performed during each sampling event.

Overall, the wells were in good condition and some routine repainting and vegetation removal were performed in 2016. Construction activities that started in 2015 near well 0138, which is in the city park, resulted in the need for riser pipe to be added to the top of the well on several occasions as additional soil was placed. Construction was still being performed (periodic placement of fill materials and grading) in 2016.

On December 7, 2016, it was discovered that the lock on well 0138 had been broken off and the well cap removed. It was not possible to determine if any vandalism had occurred. The well cap was replaced and locked at that time. It was decided to install a protective well casing to better secure the well; this work was completed on December 21, 2016. A downhole camera was used to determine if any objects, soil, or rocks were dropped into the well. None were observed. When construction activities are completed, a permanent protective casing and well pad will be installed and the top of the well casing resurveyed.

Well 0138 – A new protective casing was installed in December 2016

This page intentionally left blank

6.0 Data Validation

All data collected were validated in accordance with procedures specified in the Sitewide Operations and Maintenance Plan (DOE 2015). This procedure also fulfills the requirements of applicable procedures in the *Mound Methods Compendium* (MD 80045). Data validation was documented in reports prepared for each data package. All 2016 data, including data validation qualifiers, are summarized in Appendix D.

Laboratory performance is assessed by a review and evaluation of the following quality indicators:

- Sample shipping and receiving practices
- Chain of custody
- Laboratory blanks
- Preparation blanks
- Laboratory replicates
- Serial dilutions
- Detection limits
- Peak integrations
- Matrix spikes and matrix spike duplicates
- Holding times
- Instrument calibrations
- Interference check samples
- Radiochemical uncertainty
- Laboratory control samples
- Sample dilutions
- Surrogate recoveries
- Confirmation analyses
- Electronic data

A total of six report identification numbers (RINs) were established for the 2016 environmental sampling efforts at the Mound site. A RIN is a set of samples that are relinquished to the laboratory using a Chain of Custody form. Data Assessment Reports are prepared for each RIN and are presented in Appendix E.

The laboratory prepares an analytical package for each RIN that includes a summary of results, a complete set of supporting analytical data for every analysis reported, and an electronic data deliverable that is used to upload analytical data into databases for validation and qualification before the data are released. Every RIN received from the laboratory is thoroughly reviewed and evaluated before the data package is finalized and released to the public. Table 13 lists the RINs associated with this report.

Table 13. RINs for Calendar Year 2016 Sampling

RIN	Area	Sampling Date(s)
16017607	Parcels 6, 7, and 8	January 25–26, 2016
16017608	Phase I	January 25, 2016
16047763	Parcels 6, 7, and 8	April 25–26, 2016
16077934	Parcels 6, 7, and 8	July 25–28, 2016
16077935	Phase I	July 25–26, 2016
16108110	Parcels 6, 7, and 8	October 24–26, 2016

The Data Assessment Reports also summarize and assess the sampling quality control for each sampling event. The following items are included:

- Sampling protocol
- Trip blanks
- Outliers
- Equipment blanks
- Field duplicates

7.0 References

DOE (U.S. Department of Energy), 1992. *Remedial Investigation/Feasibility Study, Operable Unit 9, Site-Wide Work Plan*, *Mound Plant, Miamisburg, Ohio*, April.

DOE (U.S. Department of Energy), 1994a. *Operable Unit 9 Hydrogeologic Investigation: Bedrock Report*, *Mound Plant, Miamisburg, Ohio*, Technical Memorandum, January.

DOE (U.S. Department of Energy), 1994b. *Operable Unit 9 Hydrogeologic Investigation: Buried Valley Aquifer Report*, *Mound Plant, Miamisburg, Ohio*, Technical Memorandum (Revision 1), September.

DOE (U.S. Department of Energy), 1995. *Operable Unit 1 Record of Decision*, *Mound Plant, Miamisburg, Ohio*, June.

DOE (U.S. Department of Energy), 1996. *Operable Unit 9 Hydrogeologic Investigation: Groundwater Sweeps Report*, January.

DOE (U.S. Department of Energy), 1997. *Mound 2000 Residual Risk Evaluation Methodology*, *Mound Plant*, January 6.

DOE (U.S. Department of Energy), 1999. *Work Plan for Environmental Restoration of the DOE Mound Site, The Mound 2000 Approach*, February.

DOE (U.S. Department of Energy), 2014a. *Sitewide Groundwater Monitoring Report, Calendar Year 2012, Mound, Ohio*, LMS/MND/S10184, Office of Legacy Management, February.

DOE (U.S. Department of Energy), 2014b. *Sitewide Groundwater Monitoring Report Mound, Ohio, Site, Calendar Year 2013*, LMS/MND/S11737, Office of Legacy Management, May.

DOE (U.S. Department of Energy), 2015. *Operations and Maintenance Plan for the U.S. Department of Energy*, *Mound, Ohio, Site*, LMS/MND/S08406, January.

Mound Methods Compendium, Issue 2, MD-80045, prepared by BWXT of Ohio Inc.

This page intentionally left blank

Appendix A

Well Construction Summary

This page intentionally left blank

Location ID	Program	Northing	Easting	Ground Elevation	TOC Elevation	Well Depth	Top of Screen Elevation	Bottom of Screen Elevation	Screen Length	Well Material	Screened Formation
0118	Parcels 6, 7, 8	600464.95	1464737.80	705.36	704.86	40.1	674.73	664.73	10	4-inch SS	BVA
0124	Parcels 6, 7, 8	597789.14	1463654.10	704.18	705.12	55.9	659.18	649.18	10	4-inch SS	BVA
0126	Parcels 6, 7, 8	597603.58	1463643.30	704.61	705.54	54.8	660.78	650.78	10	4-inch SS	BVA
0138	Parcels 6, 7, 8	600124.02	1464263.30	698.59	697.76	40.2	667.59	657.59	10	4-inch SS	BVA
0315	Phase I	597786.28	1464020.40	722.57	723.99	54.8	679.17	669.17	10	4-inch SS	BVA
0346	Parcels 6, 7, 8	598070.11	1465048.90	743.50	742.97	45.5	702.50	697.50	5	4-inch SS	BVA
0347	Parcels 6, 7, 8	597819.31	1464034.10	723.76	725.20	68.4	666.76	656.76	10	4-inch SS	BVA
0353	Phase I	596686.11	1464609.40	744.04	745.33	19.3	731.04	726.04	5	4-inch SS	Bedrock
0379	Parcels 6, 7, 8	597624.41	1464095.90	715.24	716.11	40.9	685.24	675.24	10	4-inch SS	BVA
0386	Parcels 6, 7, 8	597789.23	1463896.00	725.16	724.79	86.6	648.16	638.16	10	4-inch SS	BVA
0387	Parcels 6, 7, 8	597654.63	1463839.50	721.26	720.89	81.6	644.26	639.26	5	4-inch SS	BVA
0389	Parcels 6, 7, 8	597781.29	1463891.90	724.96	724.65	51.7	682.96	672.96	10	4-inch SS	BVA
0392	Parcels 6, 7, 8	597648.77	1463838.30	721.18	720.84	44.7	681.18	676.18	5	4-inch SS	BVA
0400	Phase I	596122.80	1464333.10	703.22	705.11	34.4	680.72	670.72	10	2-inch SS	BVA
0402	Phase I	596407.78	1464208.00	702.48	704.02	32.3	681.74	671.74	10	2-inch SS	BVA
0411	Phase I	596808.81	1465077.10	834.83	836.57	39.7	806.89	796.89	10	2-inch SS	Bedrock
0443	Phase I	596886.22	1465177.11	856.89	858.78	39.6	829.20	819.20	10	2-inch PVC	Bedrock
0444	Phase I	596463.35	1465001.58	770.71	773.00	32.8	750.20	740.20	10	2-inch PVC	Bedrock
0445	Phase I	596448.12	1464738.54	741.29	743.43	42.5	710.93	700.93	10	2-inch PVC	Bedrock
P033	Phase I	596208.15	1464233.80	706.03	705.83	24.8	686.03	681.03	5	2-inch PVC	BVA
0601	Parcels 6, 7, 8	598743.22	1464280.80	817.52						Seep	Bedrock
0602	Parcels 6, 7, 8	598346.65	1465311.40	779.61						Seep	Bedrock
0605	Parcels 6, 7, 8	599824.63	1464935.40	817.70						Seep	Bedrock
0606	Parcels 6, 7, 8	699971.45	1464989.00	789.23						Seep	Bedrock
0607	Parcels 6, 7, 8	600015.30	1465105.70	797.00						Seep	Bedrock
0617	Phase I	596539.80	1464855.80	766.07						Seep	Bedrock

Abbreviations:
PVC = polyvinyl chloride
SS = stainless steel
TOC = top of casing

This page intentionally left blank

Appendix B

Mann-Kendall Trending Summaries

This page intentionally left blank

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$$x_2-x_1,\ x_3-x_1,\ \ldots,\ x_n-x_1,\ x_3-x_2,\ x_4-x_2,\ \ldots,\ x_n-x_{n-2},\ x_n-x_{n-1}$$

3. Let $\mathrm{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\mathrm{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\mathrm{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\mathrm{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \operatorname{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n > 10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \begin{cases} \dfrac{S-1}{\sqrt{VAR(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \dfrac{S+1}{\sqrt{VAR(S)}} & \text{if } S < 0 \end{cases} \qquad (3)$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Tritium

Analysis was performed on the data that was entered for the location labeled "0138." The following table lists the data points that were used in the analysis.

0138							
Time	**Tritium**		**Time**	**Tritium**		**Time**	**Tritium**
4/27/2005	10490		2/26/2008	1690		7/26/2011	1600
5/11/2005	8380		5/21/2008	3710		10/24/2011	914
6/21/2005	12870		8/4/2008	3470		1/31/2012	1170
7/20/2005	13190		11/18/2008	1680		4/25/2012	1170
8/29/2005	14570		2/9/2009	873		7/23/2012	1040
11/2/2005	8920		5/11/2009	1400		2/11/2013	281
1/25/2006	6680		7/27/2009	2500		8/27/2013	1190
5/19/2006	14630		10/26/2009	1590		2/18/2014	628
9/25/2006	10000		1/25/2010	1200		8/27/2014	862
12/4/2006	6910		5/5/2010	1570		1/27/2015	430
2/26/2007	7370		7/26/2010	1770		7/28/2015	1060
5/15/2007	6810		10/25/2010	880		1/25/2016	361
8/17/2007	4290		1/24/2011	479		7/28/2016	429
11/29/2007	1390		5/2/2011	1600			

SUMMARY STATISTICS for Tritium								
n				41				
Min				281				
Max				14630				
Range				14349				
Mean				3952.4				
Median				1600				
Variance				1.9094e+007				
StdDev				4369.6				
Std Error				682.42				
Skewness				1.2905				
Interquartile Range				5883				
Percentiles								
1%	5%	10%	25%	50%	75%	90%	95%	99%
281	367.8	439.8	977	1600	6860	1.239e+004	1.443e+004	1.463e+004

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. These values with low weights can be considered possible outliers and are listed in the table below.

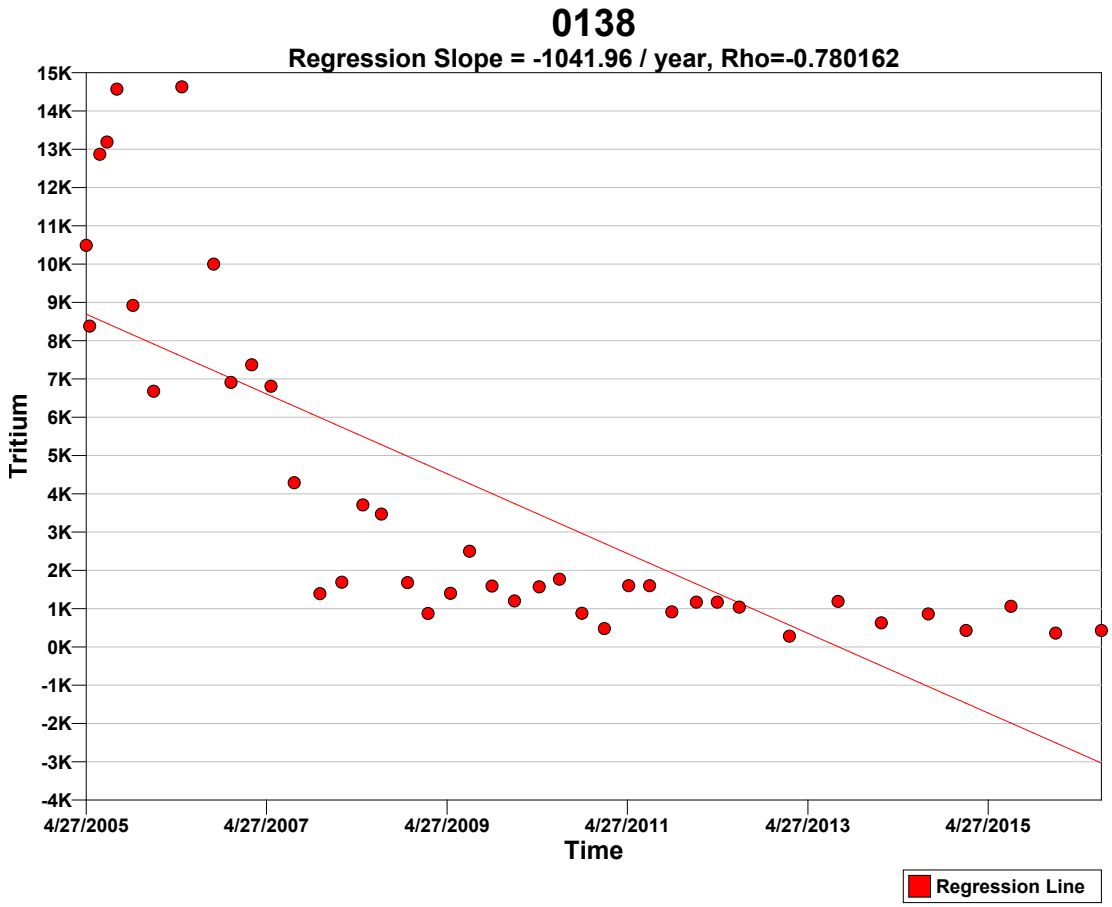
SUSPECTED OUTLIERS	
Date/Time	Value
5/11/2005	8380
8/29/2005	14570
1/25/2006	6680
5/19/2006	14630
11/29/2007	1390
2/26/2008	1690

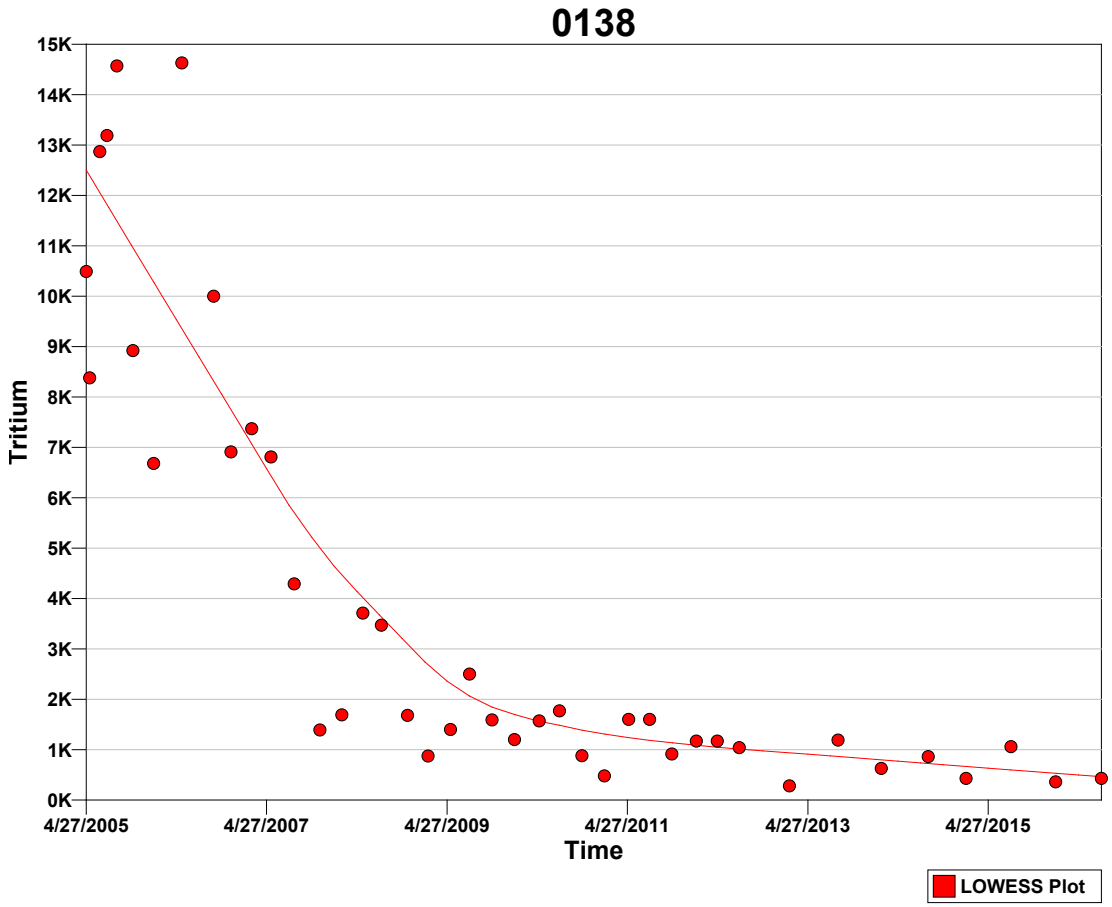
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots
The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.


0138
Regression Slope = -1041.96 / year, Rho=-0.780162
15K
14K
13K
12K
11K
10K
9K
8K
7K
6K
5K
4K
3K
2K
1K
0K
-1K
-2K
-3K
-4K
Tritium
4/27/2005
4/27/2007
4/27/2009
4/27/2011
4/27/2013
4/27/2015
Time
Regression Line



Tests

A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=41) is less than 50.

Normal Distribution Test for Tritium	
Shapiro-Wilk Test Statistic	0.75801
Shapiro-Wilk 5% Critical Value	0.941

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test

The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-596	
MK Test Statistic	Z_{MK}	-6.68385	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:27:51 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$$x_2-x_1,\ x_3-x_1,\ \ldots,\ x_n-x_1,\ x_3-x_2,\ x_4-x_2,\ \ldots,\ x_n-x_{n-2},\ x_n-x_{n-1}$$

3. Let $\mathrm{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\mathrm{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\mathrm{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\mathrm{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \operatorname{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n > 10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \begin{cases} \dfrac{S-1}{\sqrt{VAR(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \dfrac{S+1}{\sqrt{VAR(S)}} & \text{if } S < 0 \end{cases} \qquad (3)$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Trichloroethene

Analysis was performed on the data that was entered for the location labeled “0315.” The following table lists the data points that were used in the analysis.

0315							
Time	**Trichloroethene**		**Time**	**Trichloroethene**		**Time**	**Trichloroethene**
3/4/2005	10		5/5/2009	12.5		5/28/2013	7.23
5/19/2005	11		7/28/2009	11.8		8/26/2013	10.1
7/22/2005	13		10/27/2009	14.1		10/28/2013	12.9
1/27/2006	5.2		1/26/2010	11.8		2/19/2014	8.7
5/17/2006	7.2		5/5/2010	11.1		5/12/2014	8.99
9/26/2006	6.3		7/26/2010	9.48		8/28/2014	10.5
12/6/2006	3.79		10/25/2010	15.2		10/27/2014	11.7
2/21/2007	11.6		1/25/2011	13.6		1/28/2015	9.62
5/16/2007	13.2		5/2/2011	8.1		4/27/2015	6.07
8/20/2007	12.2		7/28/2011	6.23		7/27/2015	6.39
11/27/2007	16.8		10/24/2011	10.7		10/28/2015	8.24
2/26/2008	9.38		2/1/2012	10.3		1/25/2016	3.83
5/21/2008	11.2		4/23/2012	12.3		4/26/2016	5.67
8/5/2008	9.65		7/25/2012	10.8		7/27/2016	6.81
11/18/2008	17.2		10/23/2012	16.6		10/26/2016	7.53
2/12/2009	11.7		2/12/2013	12.5			

SUMMARY STATISTICS for Trichloroethene								
n				47				
Min				3.79				
Max				17.2				
Range				13.41				
Mean				10.23				
Median				10.5				
Variance				10.503				
StdDev				3.2408				
Std Error				0.47272				
Skewness				0.054322				
Interquartile Range				4.77				
Percentiles								
1%	5%	10%	25%	50%	75%	90%	95%	99%
3.79	4.378	5.99	7.53	10.5	12.3	14.32	16.72	17.2

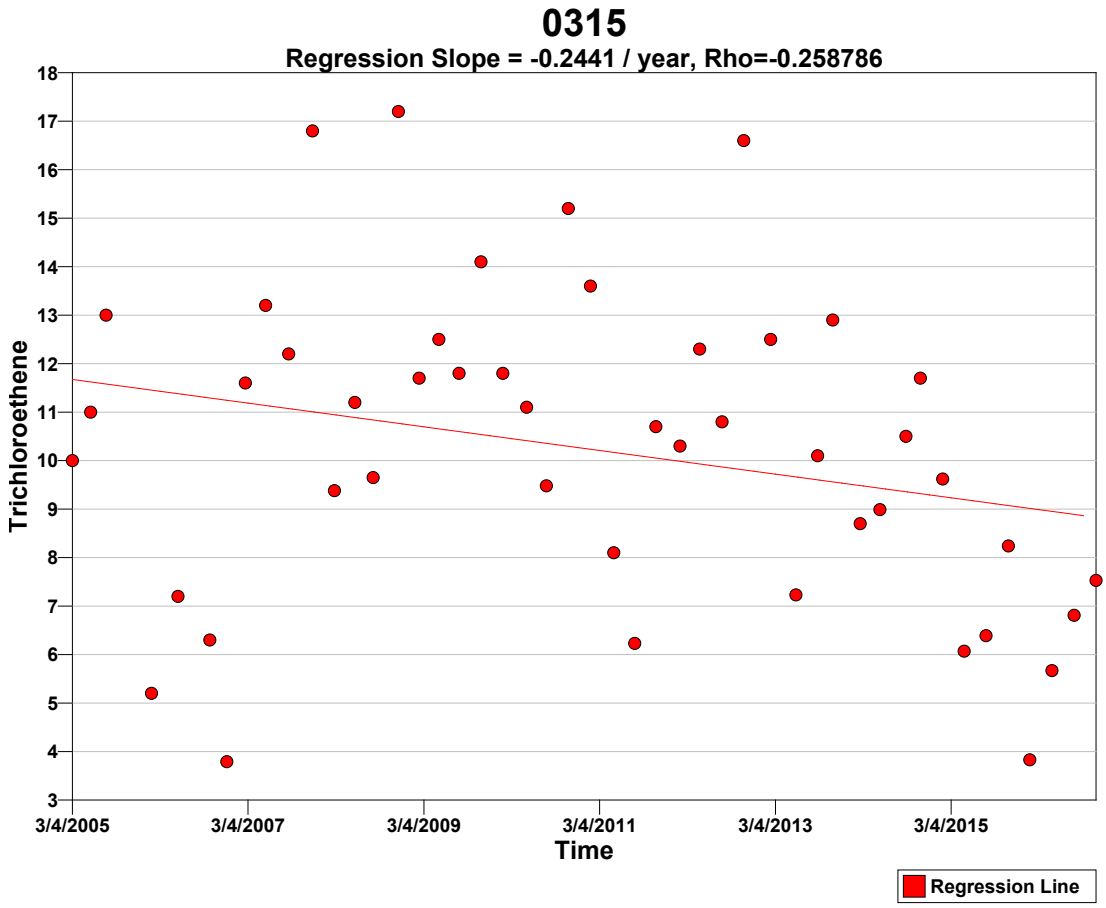
Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. The LOWESS method did not detect any outliers for the current data set.

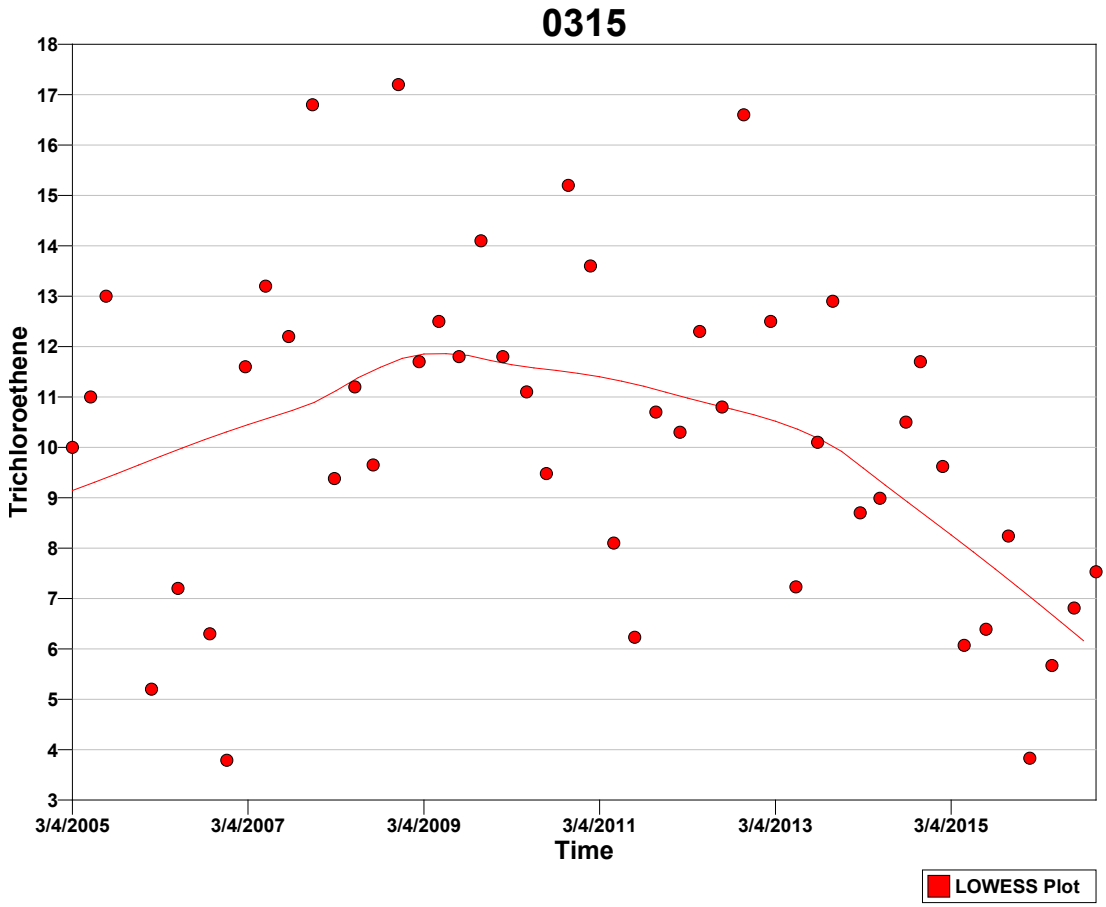
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots
The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.


0315
Regression Slope = -0.2441 / year, Rho=-0.258786
Trichloroethene
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
3/4/2005
3/4/2007
3/4/2009
3/4/2011
3/4/2013
3/4/2015
Time
Regression Line



Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=47) is less than 50.

Normal Distribution Test for Trichloroethene	
Shapiro-Wilk Test Statistic	0.97426
Shapiro-Wilk 5% Critical Value	0.946

The calculated test statistic exceeds the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-218	
MK Test Statistic	Z_{MK}	-1.99024	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:20:16 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$$x_2-x_1,\ x_3-x_1,\ \ldots,\ x_n-x_1,\ x_3-x_2,\ x_4-x_2,\ \ldots,\ x_n-x_{n-2},\ x_n-x_{n-1}$$

3. Let $\mathrm{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\mathrm{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\mathrm{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\mathrm{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \operatorname{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n > 10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \begin{cases} \dfrac{S-1}{\sqrt{VAR(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \dfrac{S+1}{\sqrt{VAR(S)}} & \text{if } S < 0 \end{cases} \qquad (3)$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Trichloroethene

Analysis was performed on the data that was entered for the location labeled "0315. The following table lists the data points that were used in the analysis.

0315							
Time	**Trichloroethene**		**Time**	**Trichloroethene**		**Time**	**Trichloroethene**
2/1/2012	10.3		10/28/2013	12.9		7/27/2015	6.39
4/23/2012	12.3		2/19/2014	8.7		10/28/2015	8.24
7/25/2012	10.8		5/12/2014	8.99		1/25/2016	3.83
10/23/2012	16.6		8/28/2014	10.5		4/26/2016	5.67
2/12/2013	12.5		10/27/2014	11.7		7/27/2016	6.81
5/28/2013	7.23		1/28/2015	9.62		10/26/2016	7.53
8/26/2013	10.1		4/27/2015	6.07			

SUMMARY STATISTICS for Trichloroethene								
n				20				
Min				3.83				
Max				16.6				
Range				12.77				
Mean				9.339				
Median				9.305				
Variance				9.1499				
StdDev				3.0249				
Std Error				0.67638				
Skewness				0.41083				
Interquartile Range				4.56				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
3.83	3.922	5.71	6.915	9.305	11.48	12.86	16.41	16.6

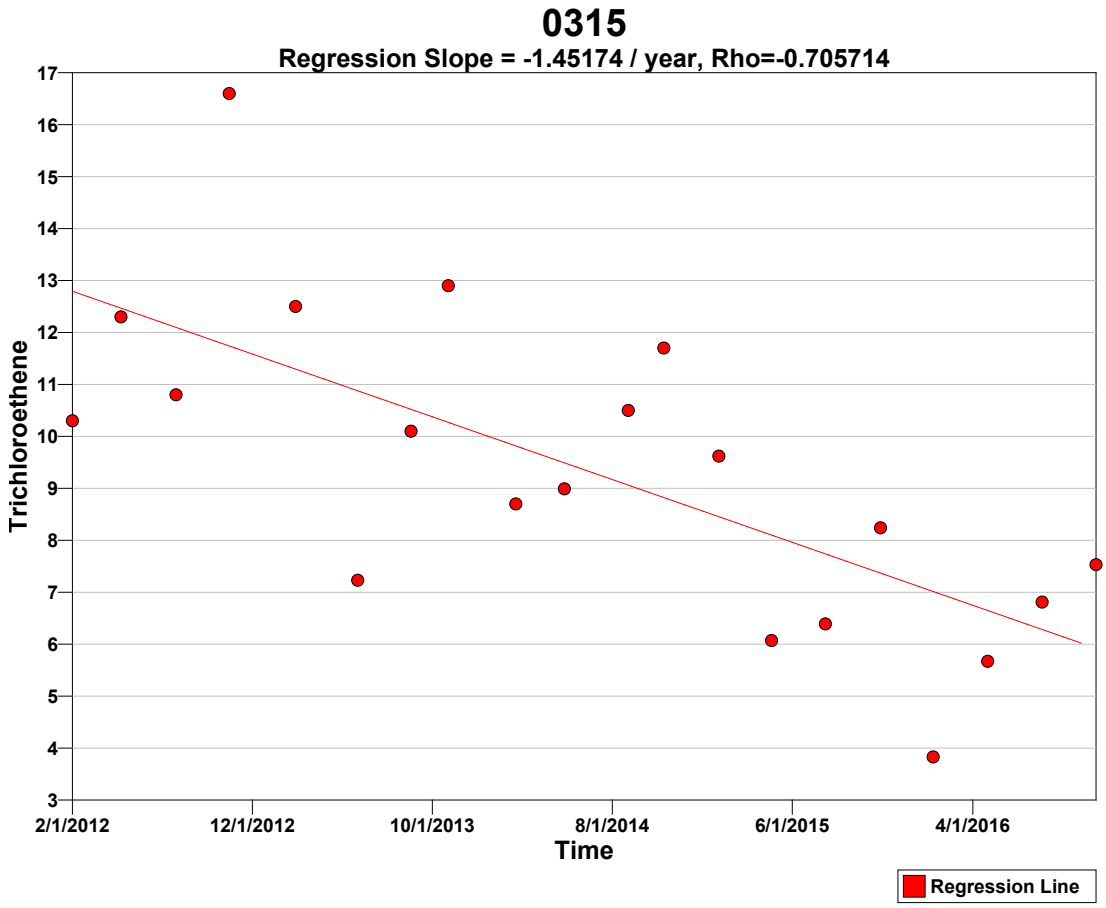
Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. The LOWESS method did not detect any outliers for the current data set.

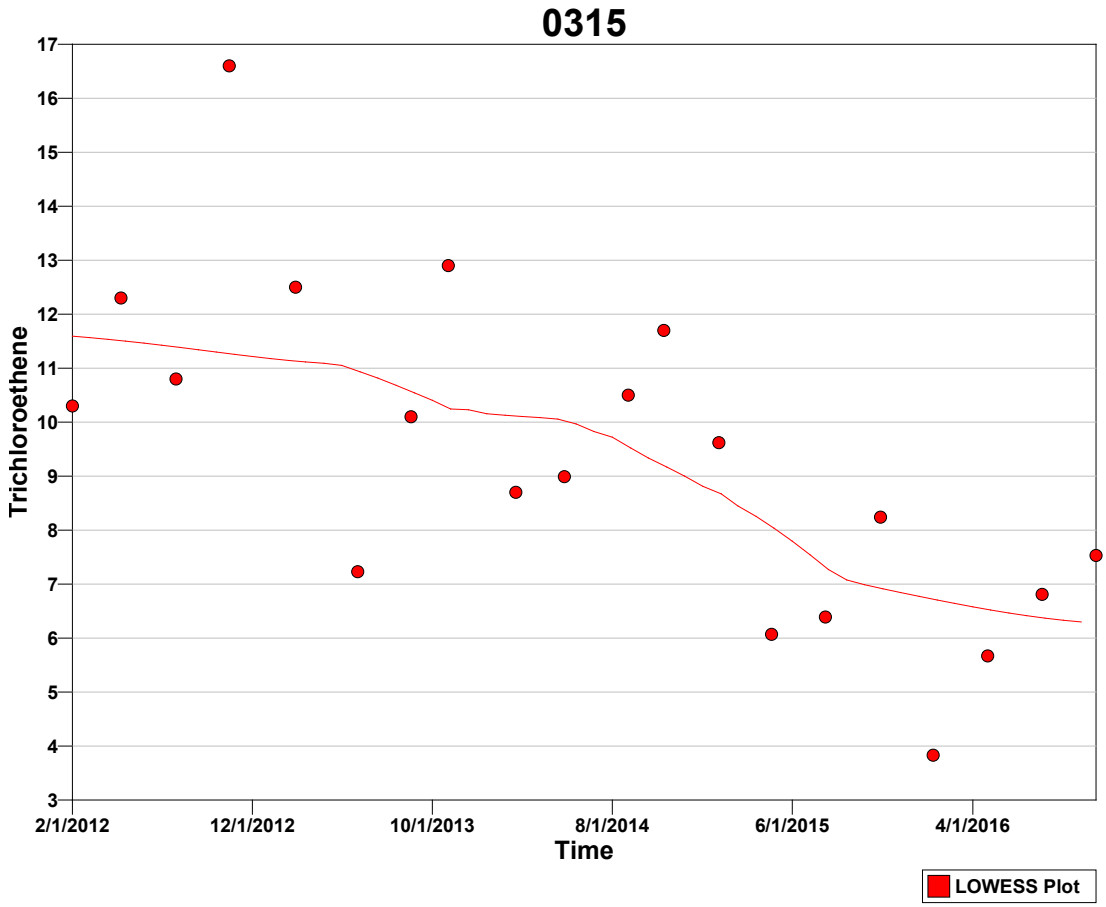
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots
The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.


0315
Regression Slope = -1.45174 / year, Rho=-0.705714
Trichloroethene
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2/1/2012
12/1/2012
10/1/2013
8/1/2014
6/1/2015
4/1/2016
Time
Regression Line



Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=20) is less than 50.

Normal Distribution Test for Trichloroethene	
Shapiro-Wilk Test Statistic	0.98175
Shapiro-Wilk 5% Critical Value	0.905

The calculated test statistic exceeds the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-94	
MK Test Statistic	Z_{MK}	-3.01732	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 4/10/2017 7:34:06 AM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$$x_2-x_1,\ x_3-x_1,\ \ldots,\ x_n-x_1,\ x_3-x_2,\ x_4-x_2,\ \ldots,\ x_n-x_{n-2},\ x_n-x_{n-1}$$

3. Let $\mathrm{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\mathrm{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\mathrm{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\mathrm{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \operatorname{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \begin{cases} \dfrac{S-1}{\sqrt{VAR(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \dfrac{S+1}{\sqrt{VAR(S)}} & \text{if } S < 0 \end{cases} \qquad (3)$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Tritium

Analysis was performed on the data that was entered for the location labeled "0346. The following table lists the data points that were used in the analysis.

0346							
Time	**Tritium**		**Time**	**Tritium**		**Time**	**Tritium**
3/2/2005	33770		11/28/2007	2610		7/27/2011	1470
4/19/2005	15100		2/26/2008	2090		10/24/2011	919
4/27/2005	14900		5/20/2008	2260		2/1/2012	943
5/16/2005	7850		8/5/2008	1200		4/24/2012	1050
6/21/2005	5220		11/18/2008	1830		7/25/2012	856
7/22/2005	2130		2/11/2009	1810		2/12/2013	494
11/3/2005	4460		5/5/2009	1780		8/26/2013	665
12/14/2005	4520		7/28/2009	2030		2/24/2014	196
1/25/2006	2320		10/27/2009	2000		8/28/2014	1020
5/22/2006	2050		1/26/2010	1320		1/28/2015	18.6
9/26/2006	3490		5/6/2010	1830		7/28/2015	736
12/5/2006	4750		7/27/2010	1730		1/25/2016	860
2/21/2007	2120		10/26/2010	1710		7/26/2016	566
5/16/2007	1830		1/25/2011	1280			
8/16/2007	1760		5/3/2011	987			

SUMMARY STATISTICS for Tritium								
n				43				
Min				18.6				
Max				33770				
Range				33751				
Mean				3314.7				
Median				1810				
Variance				3.2508e+007				
StdDev				5701.6				
Std Error				869.49				
Skewness				4.2079				
Interquartile Range				1333				
Percentiles								
1%	5%	10%	25%	50%	75%	90%	95%	99%
18.6	255.6	605.6	987	1810	2320	6798	1.506e+004	3.377e+004

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. These values with low weights can be considered possible outliers and are listed in the table below.

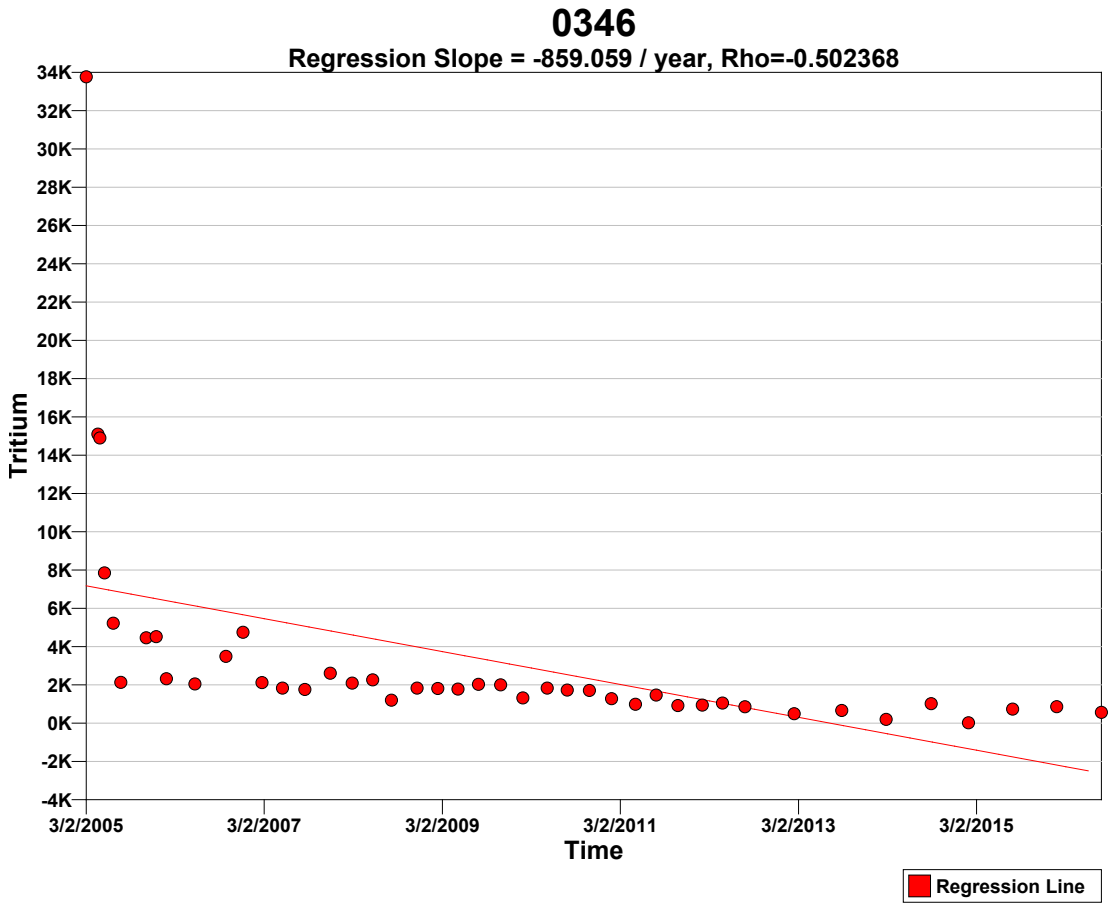
SUSPECTED OUTLIERS	
Date/Time	Value
3/2/2005	33770
4/19/2005	15100
4/27/2005	14900
5/16/2005	7850
7/22/2005	2130
1/25/2006	2320
5/22/2006	2050
12/5/2006	4750

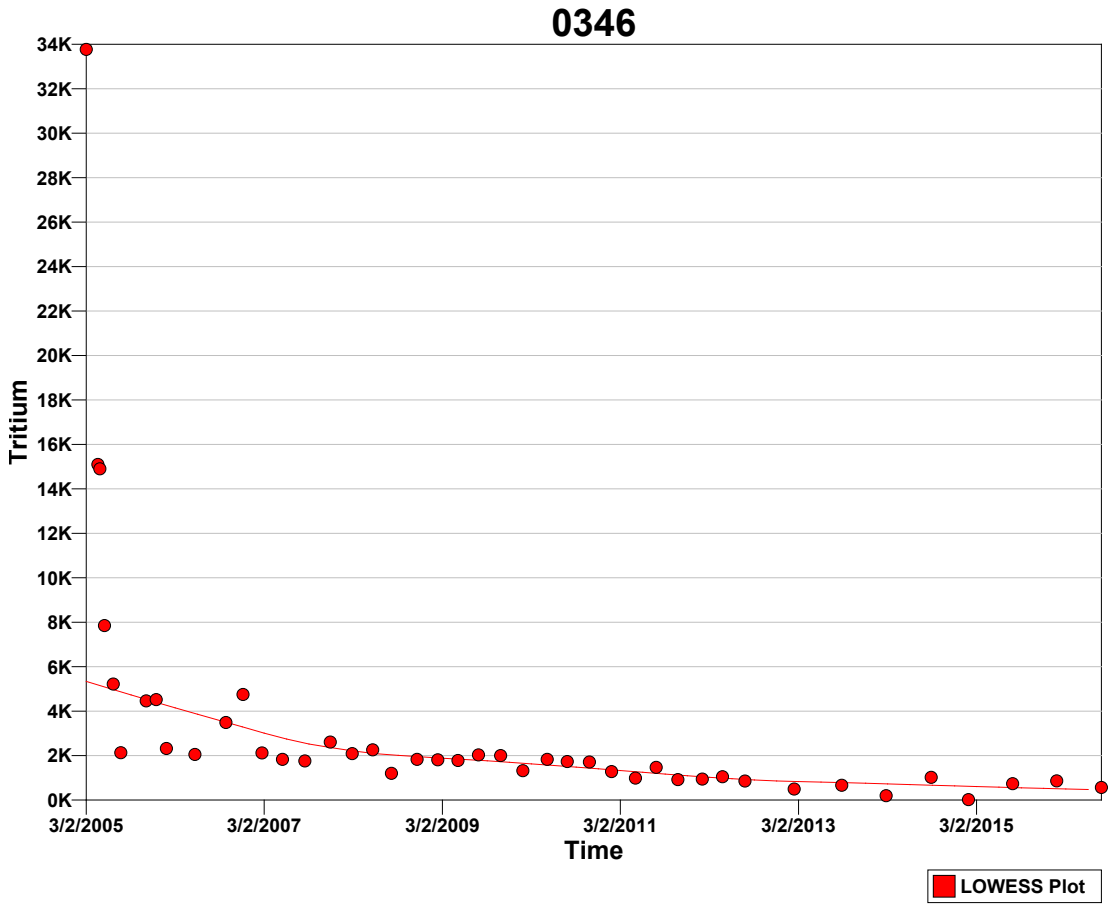
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots
The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.


0346
Regression Slope = -859.059 / year, Rho=-0.502368
Tritium
34K
32K
30K
28K
26K
24K
22K
20K
18K
16K
14K
12K
10K
8K
6K
4K
2K
0K
-2K
-4K
3/2/2005
3/2/2007
3/2/2009
3/2/2011
3/2/2013
3/2/2015
Time
Regression Line



Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=43) is less than 50.

Normal Distribution Test for Tritium	
Shapiro-Wilk Test Statistic	0.48299
Shapiro-Wilk 5% Critical Value	0.943

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-726	
MK Test Statistic	Z_{MK}	-7.58895	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:28:58 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$$x_2-x_1,\ x_3-x_1,\ \ldots,\ x_n-x_1,\ x_3-x_2,\ x_4-x_2,\ \ldots,\ x_n-x_{n-2},\ x_n-x_{n-1}$$

3. Let $\mathrm{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\mathrm{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\mathrm{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\mathrm{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \mathrm{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \begin{cases} \dfrac{S-1}{\sqrt{VAR(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \dfrac{S+1}{\sqrt{VAR(S)}} & \text{if } S < 0 \end{cases} \qquad (3)$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Trichloroethene

Analysis was performed on the data that was entered for the location labeled "0347. The following table lists the data points that were used in the analysis.

0347							
Time	**Trichloroethene**		**Time**	**Trichloroethene**		**Time**	**Trichloroethene**
2/24/2005	20		5/5/2009	27.3		5/28/2013	23.2
5/16/2005	22		7/28/2009	23.3		8/26/2013	27
7/22/2005	23		10/27/2009	29.2		10/28/2013	23.3
1/27/2006	21		1/27/2010	32.3		2/19/2014	18.3
5/17/2006	21		5/5/2010	26.9		5/12/2014	23.1
9/27/2006	22		7/27/2010	18.9		8/28/2014	20.8
12/5/2006	33.2		10/25/2010	27.9		10/27/2014	23.8
2/21/2007	19		1/25/2011	23.2		1/28/2015	25.4
5/16/2007	12		5/2/2011	24.5		4/27/2015	24.5
8/16/2007	26.2		7/28/2011	22.1		7/27/2015	17.6
11/27/2007	0.431		10/24/2011	27.8		10/28/2015	19.6
2/26/2008	17.9		2/1/2012	28.8		1/25/2016	16.6
5/21/2008	20.5		4/23/2012	23.8		4/26/2016	13.9
8/5/2008	18.3		7/25/2012	25		7/27/2016	18.2
11/19/2008	21.2		10/24/2012	31.2		10/26/2016	24.6
2/12/2009	26.1		2/12/2013	27.5			

SUMMARY STATISTICS for Trichloroethene								
n				47				
Min				0.431				
Max				33.2				
Range				32.769				
Mean				22.626				
Median				23.2				
Variance				31.164				
StdDev				5.5824				
Std Error				0.81428				
Skewness				-1.2558				
Interquartile Range				6.6				
Percentiles								
1%	5%	10%	25%	50%	75%	90%	95%	99%
0.431	12.76	17.4	19.6	23.2	26.2	28.88	31.86	33.2

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. These values with low weights can be considered possible outliers and are listed in the table below.

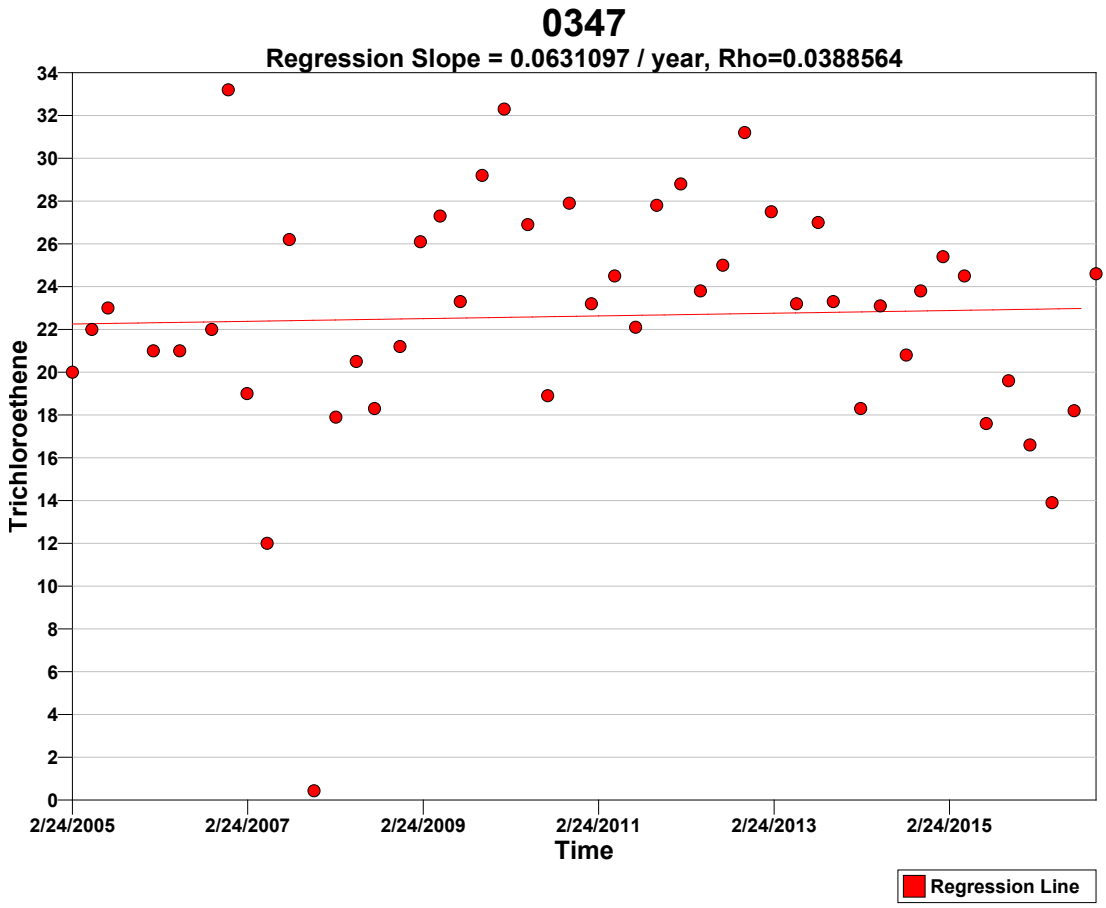
SUSPECTED OUTLIERS	
Date/Time	Value
11/27/2007	0.431

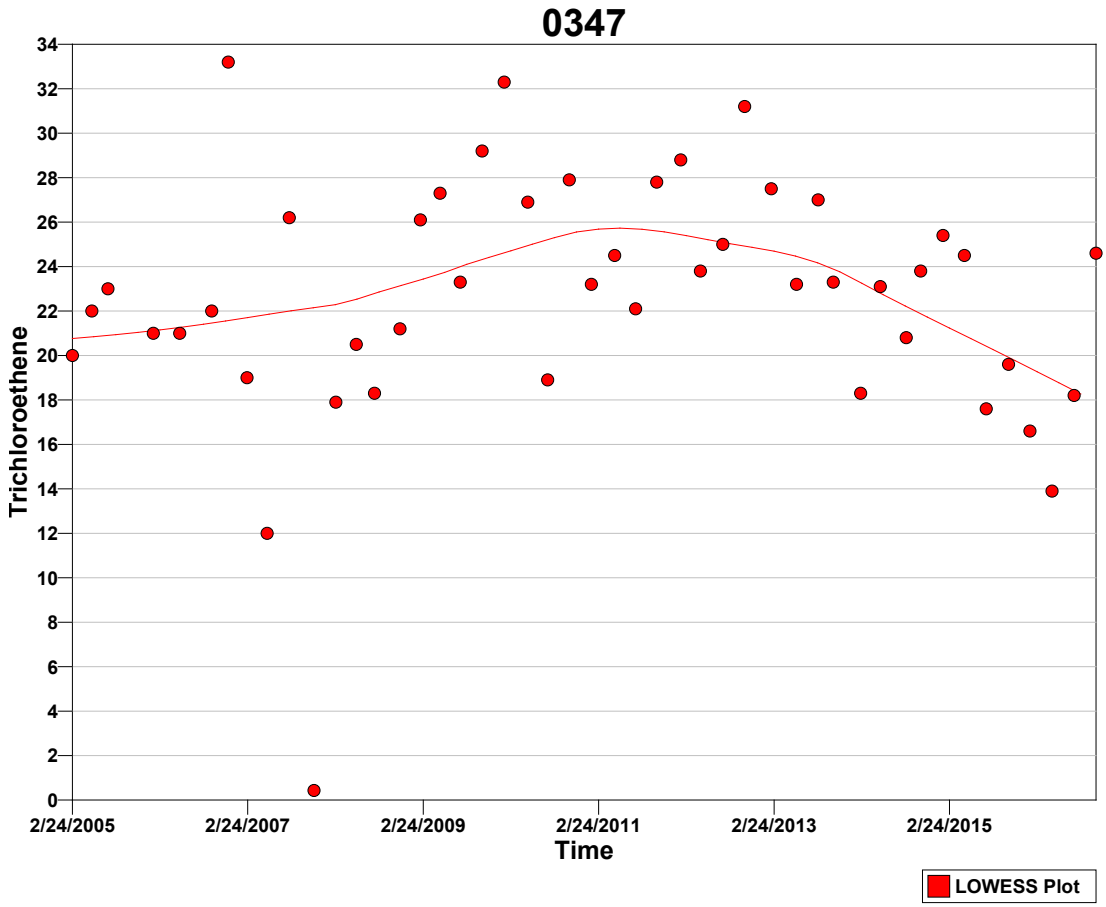
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots
The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.


0347
Regression Slope = 0.0631097 / year, Rho=0.0388564
Trichloroethene
34
32
30
28
26
24
22
20
18
16
14
12
10
8
6
4
2
0
2/24/2005
2/24/2007
2/24/2009
2/24/2011
2/24/2013
2/24/2015
Time
Regression Line



Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=47) is less than 50.

Normal Distribution Test for Trichloroethene	
Shapiro-Wilk Test Statistic	0.9305
Shapiro-Wilk 5% Critical Value	0.946

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	6	
MK Test Statistic	Z_{MK}	0.0458658	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	**Accept**
Alternative Hypothesis	H_a	A downward monotonic trend exists	Reject
Conclude with 95% confidence that a downward monotonic trend does not exist.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:24:26 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$$x_2-x_1,\ x_3-x_1,\ \ldots,\ x_n-x_1,\ x_3-x_2,\ x_4-x_2,\ \ldots,\ x_n-x_{n-2},\ x_n-x_{n-1}$$

3. Let $\mathrm{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\mathrm{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\mathrm{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\mathrm{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \operatorname{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n > 10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \begin{cases} \dfrac{S-1}{\sqrt{VAR(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \dfrac{S+1}{\sqrt{VAR(S)}} & \text{if } S < 0 \end{cases} \qquad (3)$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Trichloroethene

Analysis was performed on the data that was entered for the location labeled "0347. The following table lists the data points that were used in the analysis.

0347							
Time	**Trichloroethene**		**Time**	**Trichloroethene**		**Time**	**Trichloroethene**
2/1/2012	28.8		10/28/2013	23.3		7/27/2015	17.6
4/23/2012	23.8		2/19/2014	18.3		10/28/2015	19.6
7/25/2012	25		5/12/2014	23.1		1/25/2016	16.6
10/24/2012	31.2		8/28/2014	20.8		4/26/2016	13.9
2/12/2013	27.5		10/27/2014	23.8		7/27/2016	18.2
5/28/2013	23.2		1/28/2015	25.4		10/26/2016	24.6
8/26/2013	27		4/27/2015	24.5			

SUMMARY STATISTICS for Trichloroethene								
n				20				
Min				13.9				
Max				31.2				
Range				17.3				
Mean				22.81				
Median				23.55				
Variance				19.308				
StdDev				4.3941				
Std Error				0.98256				
Skewness				-0.20426				
Interquartile Range				6.675				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
13.9	14.04	16.7	18.63	23.55	25.3	28.67	31.08	31.2

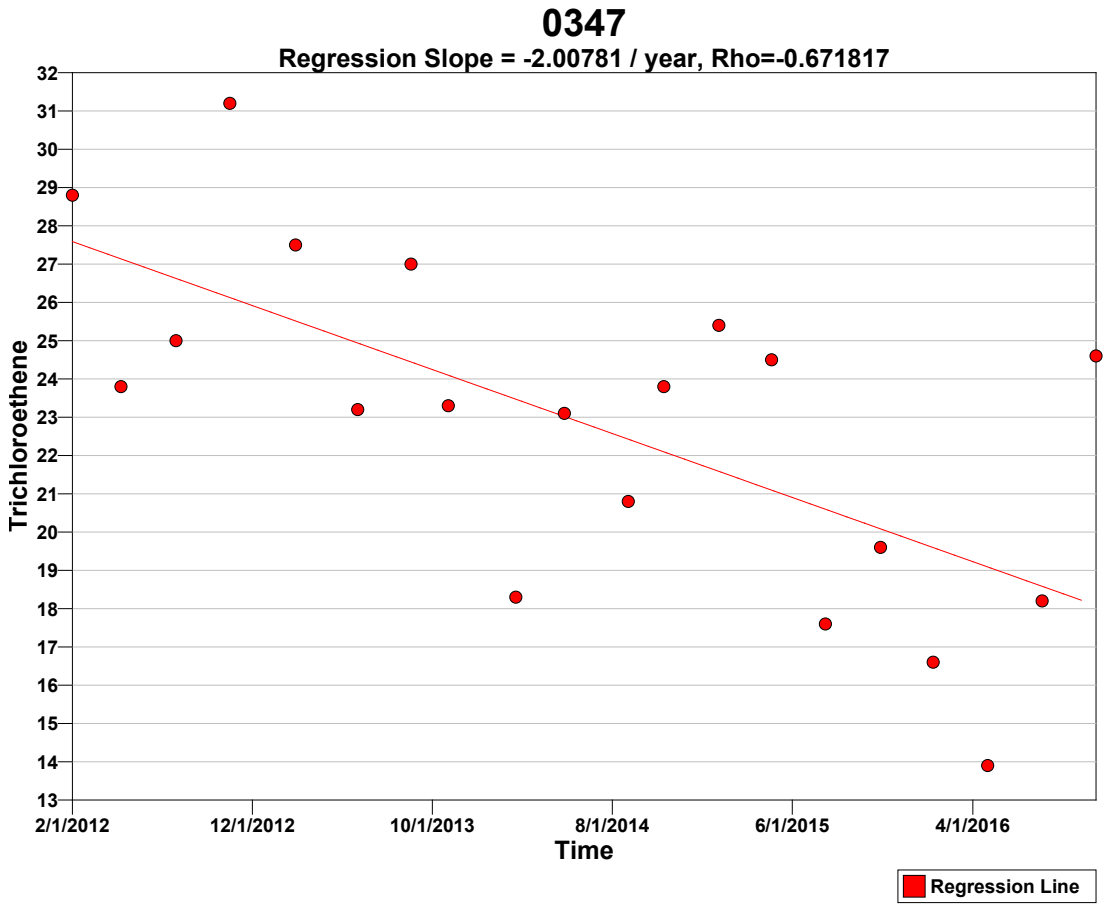
Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. The LOWESS method did not detect any outliers for the current data set.

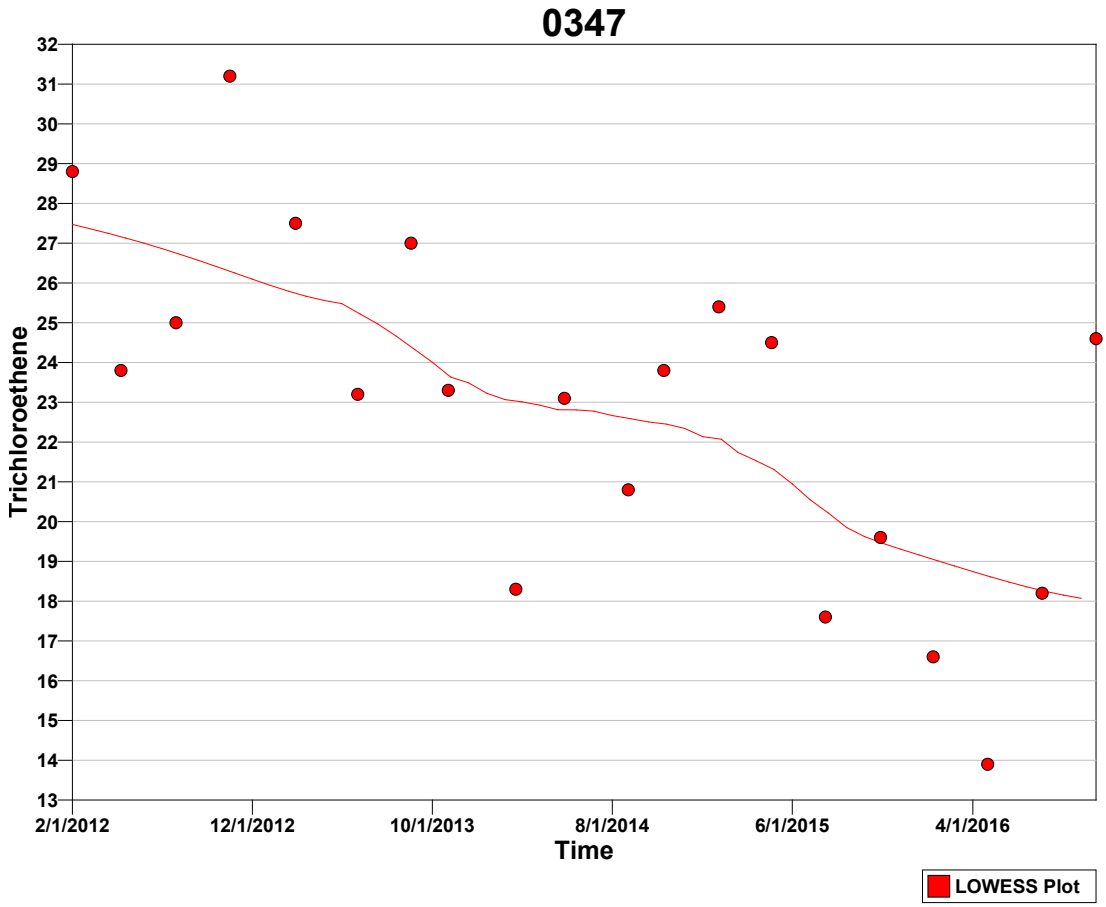
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots
The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.


0347
Regression Slope = -2.00781 / year, Rho=-0.671817
Trichloroethene
32
31
30
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
2/1/2012
12/1/2012
10/1/2013
8/1/2014
6/1/2015
4/1/2016
Time
Regression Line



Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=20) is less than 50.

Normal Distribution Test for Trichloroethene	
Shapiro-Wilk Test Statistic	0.97531
Shapiro-Wilk 5% Critical Value	0.905

The calculated test statistic exceeds the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-89	
MK Test Statistic	Z_{MK}	-2.8566	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 4/10/2017 7:35:24 AM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$$x_2-x_1,\ x_3-x_1,\ \ldots,\ x_n-x_1,\ x_3-x_2,\ x_4-x_2,\ \ldots,\ x_n-x_{n-2},\ x_n-x_{n-1}$$

3. Let $\mathrm{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\mathrm{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\mathrm{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\mathrm{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \operatorname{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n > 10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \begin{cases} \dfrac{S-1}{\sqrt{VAR(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \dfrac{S+1}{\sqrt{VAR(S)}} & \text{if } S < 0 \end{cases} \qquad (3)$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Tritium

Analysis was performed on the data that was entered for the location labeled "0347. The following table lists the data points that were used in the analysis.

0347							
Time	**Tritium**		**Time**	**Tritium**		**Time**	**Tritium**
2/24/2005	2290		5/21/2008	9270		10/24/2011	5310
4/20/2005	1100		8/5/2008	5880		2/1/2012	3800
5/16/2005	1550		11/19/2008	6040		4/23/2012	4460
7/22/2005	670		2/12/2009	6570		7/25/2012	4310
11/22/2005	3740		5/5/2009	6860		2/12/2013	3420
1/27/2006	5830		7/28/2009	4590		8/26/2013	2460
5/17/2006	15420		10/27/2009	6540		2/19/2014	2290
9/27/2006	16800		1/27/2010	6530		8/28/2014	2610
12/5/2006	13000		5/5/2010	3670		1/28/2015	2820
2/21/2007	16400		7/27/2010	6710		7/27/2015	1690
5/16/2007	13100		10/25/2010	6490		1/25/2016	1430
8/16/2007	10200		1/25/2011	5730		7/27/2016	1630
11/27/2007	157		5/2/2011	3940			
2/26/2008	6000		7/28/2011	2940			

SUMMARY STATISTICS for Tritium	
n	40
Min	157
Max	16800
Range	16643
Mean	5606.2
Median	4525
Variance	1.8141e+007
StdDev	4259.2
Std Error	673.44
Skewness	1.3051
Interquartile Range	4065

Percentiles								
1%	5%	10%	25%	50%	75%	90%	95%	99%
157	691.5	1442	2498	4525	6563	1.309e+004	1.635e+004	1.68e+004

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. These values with low weights can be considered possible outliers and are listed in the table below.

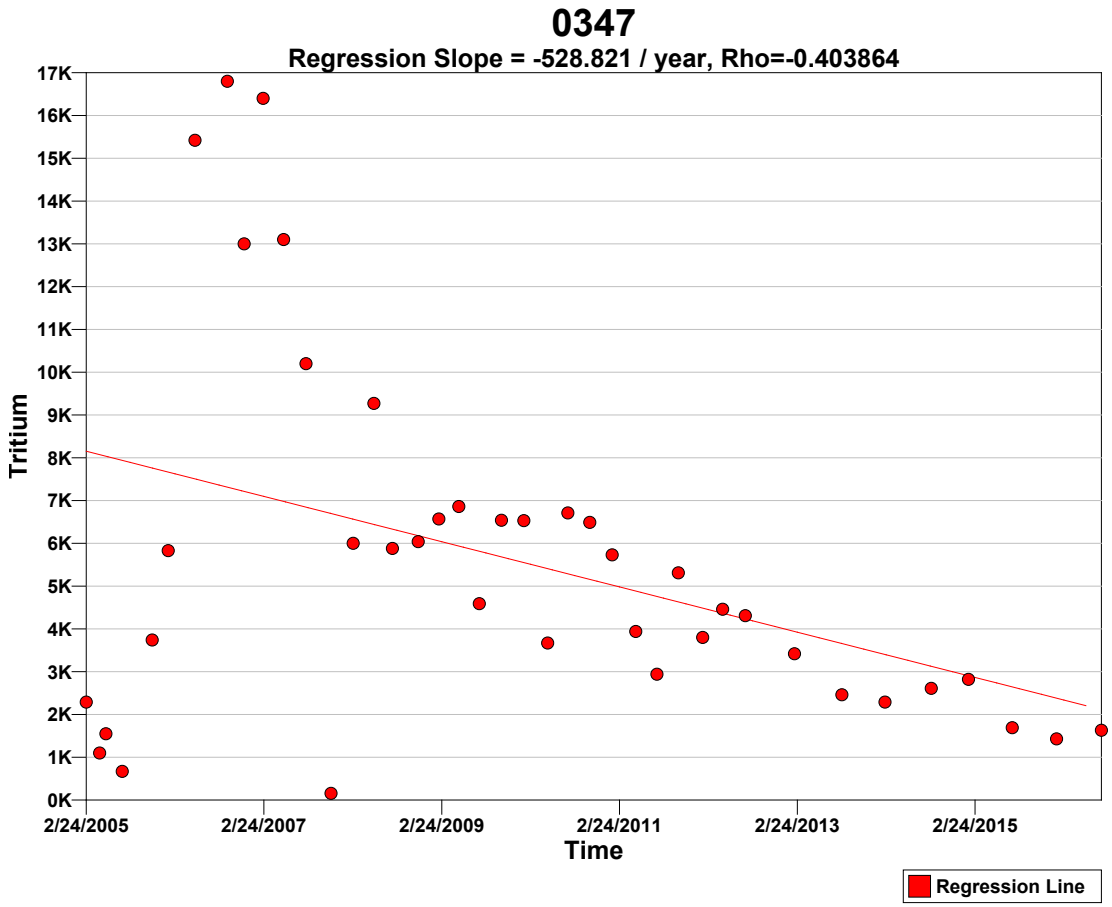
SUSPECTED OUTLIERS	
Date/Time	Value
5/17/2006	15420
9/27/2006	16800
12/5/2006	13000
2/21/2007	16400
5/16/2007	13100
11/27/2007	157

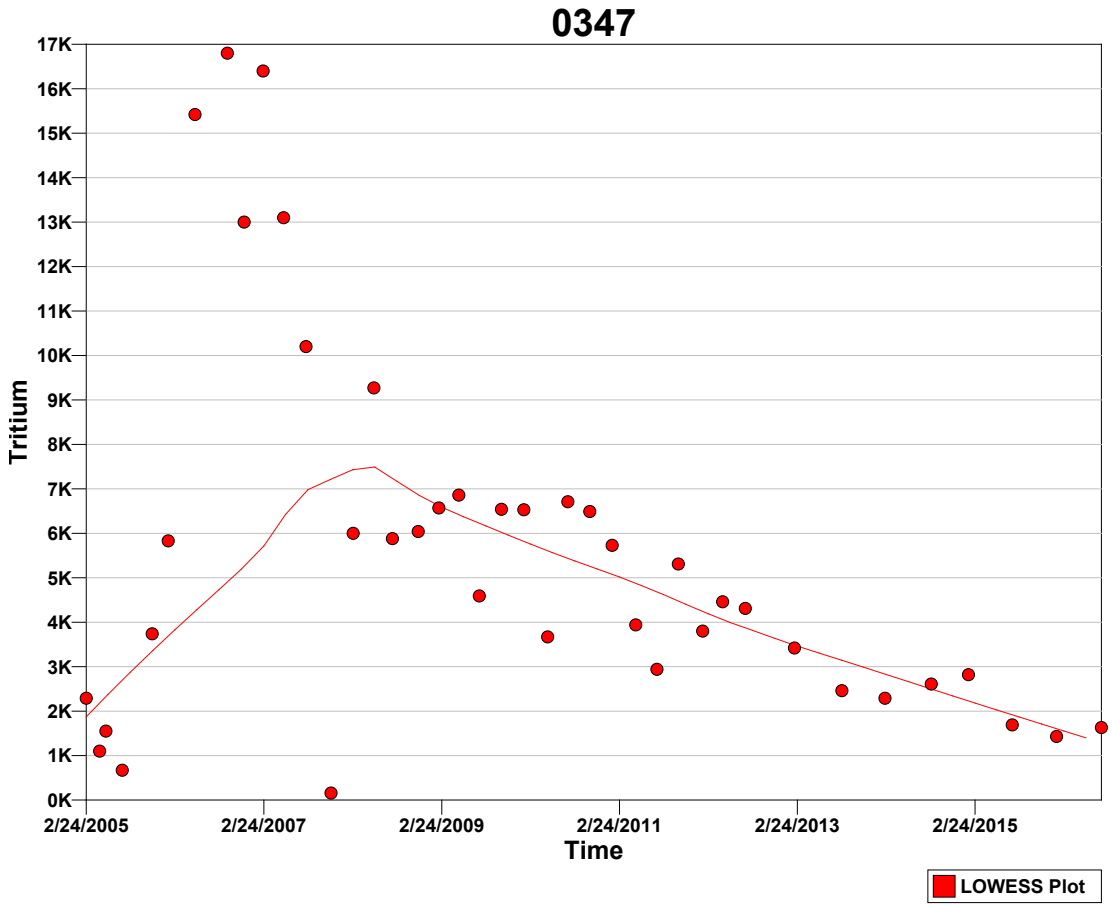
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots
The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.


0347
Regression Slope = -528.821 / year, Rho=-0.403864
Tritium
17K
16K
15K
14K
13K
12K
11K
10K
9K
8K
7K
6K
5K
4K
3K
2K
1K
0K
2/24/2005
2/24/2007
2/24/2009
2/24/2011
2/24/2013
2/24/2015
Time
Regression Line



Tests

A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=40) is less than 50.

Normal Distribution Test for Tritium	
Shapiro-Wilk Test Statistic	0.85746
Shapiro-Wilk 5% Critical Value	0.94

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test

The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-255	
MK Test Statistic	Z_{MK}	-2.95956	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:29:46 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$$x_2-x_1,\ x_3-x_1,\ \ldots,\ x_n-x_1,\ x_3-x_2,\ x_4-x_2,\ \ldots,\ x_n-x_{n-2},\ x_n-x_{n-1}$$

3. Let $\mathrm{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\mathrm{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\mathrm{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\mathrm{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \operatorname{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \begin{cases} \dfrac{S-1}{\sqrt{VAR(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \dfrac{S+1}{\sqrt{VAR(S)}} & \text{if } S < 0 \end{cases} \qquad (3)$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Tritium

Analysis was performed on the data that was entered for the location labeled "0379. The following table lists the data points that were used in the analysis.

0379							
Time	**Tritium**		**Time**	**Tritium**		**Time**	**Tritium**
4/19/2005	1410		8/6/2008	1870		10/24/2011	1610
5/18/2005	750		11/18/2008	1660		2/1/2012	2010
11/1/2005	1730		2/12/2009	1680		4/24/2012	1450
1/24/2006	600		5/5/2009	1530		7/25/2012	1550
5/17/2006	780		7/28/2009	1500		2/11/2013	1210
9/26/2006	1640		10/27/2009	1770		8/27/2013	852
12/5/2006	3890		1/26/2010	1720		2/19/2014	948
2/21/2007	1980		5/5/2010	479		8/28/2014	1200
5/16/2007	1950		7/26/2010	1660		1/29/2015	1030
8/16/2007	1920		10/25/2010	1550		7/28/2015	954
11/27/2007	1750		1/25/2011	1420		1/25/2016	1230
2/26/2008	1980		5/3/2011	1450		7/26/2016	590
5/21/2008	1720		7/27/2011	1350			

SUMMARY STATISTICS for Tritium								
n				38				
Min				479				
Max				3890				
Range				3411				
Mean				1483.5				
Median				1540				
Variance				3.4159e+005				
StdDev				584.46				
Std Error				94.811				
Skewness				1.5447				
Interquartile Range				577.5				
Percentiles								
1%	5%	10%	25%	50%	75%	90%	95%	99%
479	584.5	735	1158	1540	1735	1980	2104	3890

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. These values with low weights can be considered possible outliers and are listed in the table below.

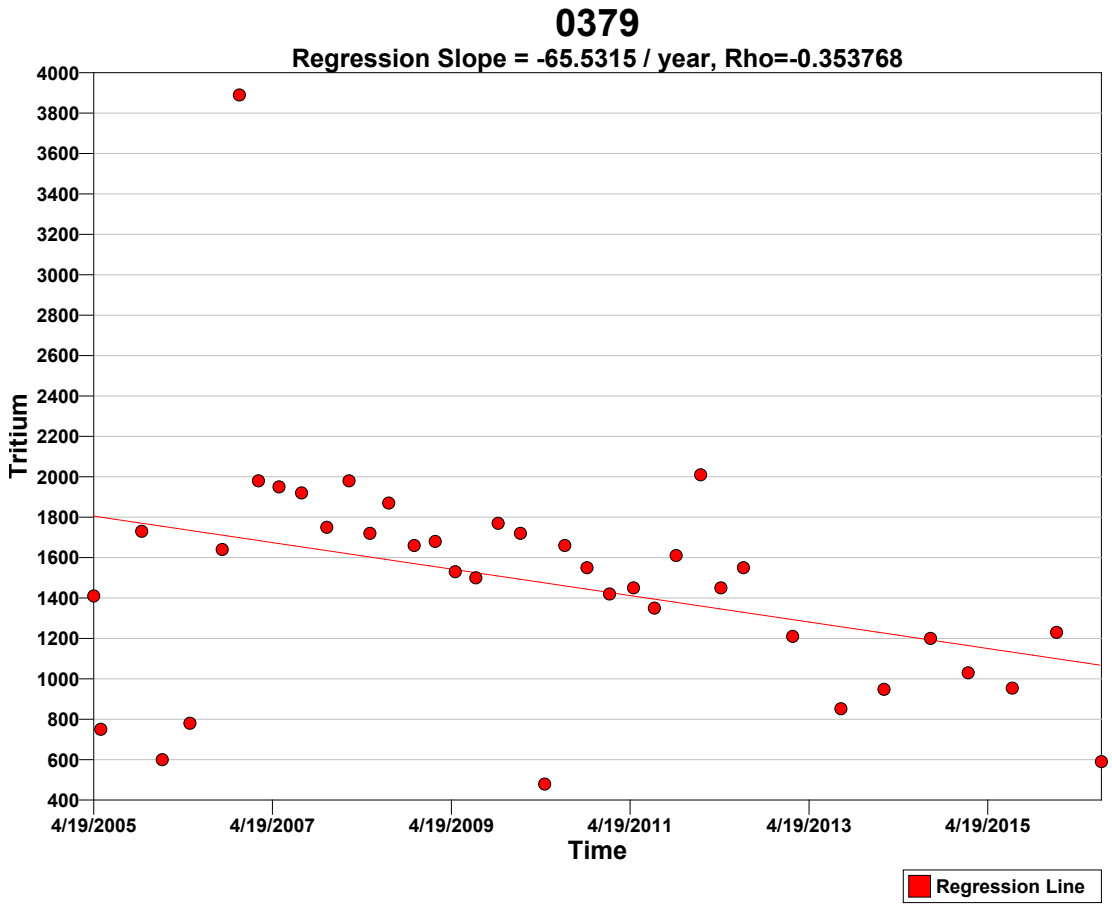
SUSPECTED OUTLIERS	
Date/Time	Value
1/24/2006	600
5/17/2006	780
12/5/2006	3890
5/5/2010	479

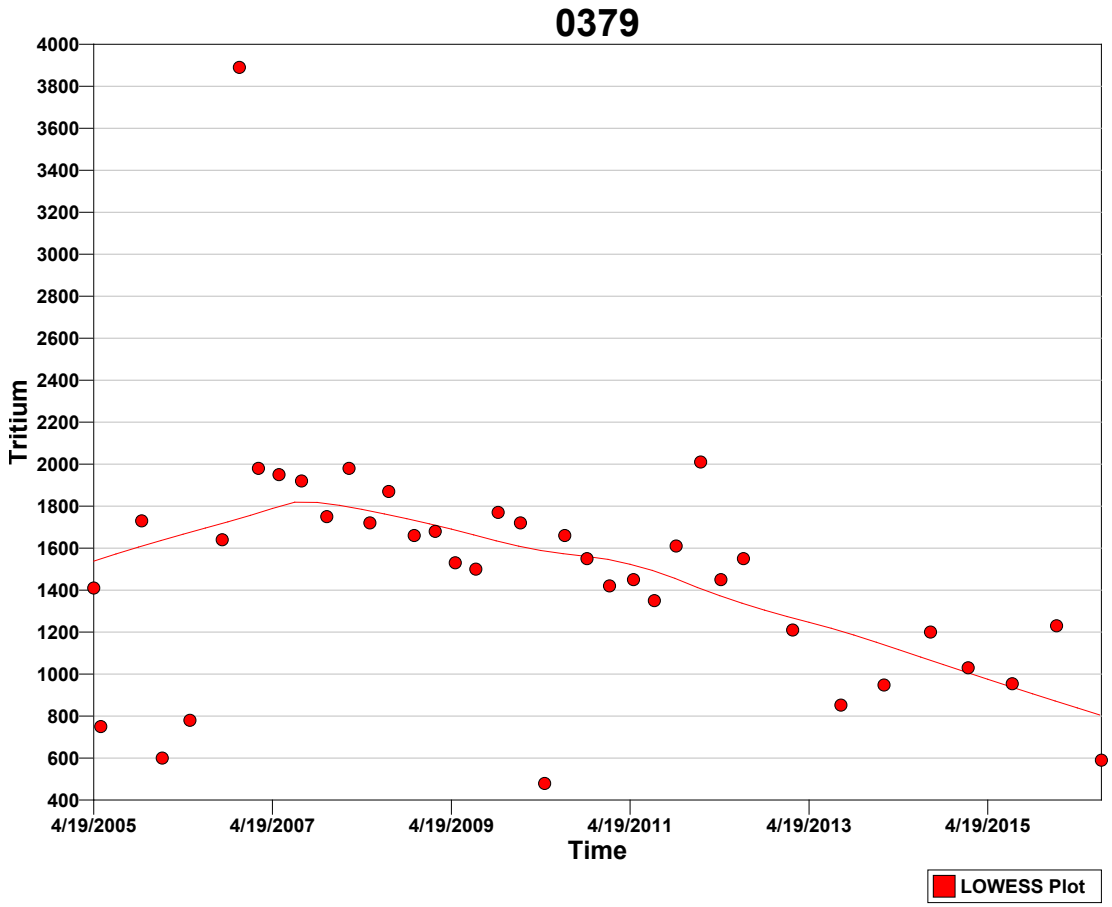
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots
The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.


0379
Regression Slope = -65.5315 / year, Rho=-0.353768
Tritium
4000
3800
3600
3400
3200
3000
2800
2600
2400
2200
2000
1800
1600
1400
1200
1000
800
600
400
4/19/2005
4/19/2007
4/19/2009
4/19/2011
4/19/2013
4/19/2015
Time
Regression Line



Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=38) is less than 50.

Normal Distribution Test for Tritium	
Shapiro-Wilk Test Statistic	0.86052
Shapiro-Wilk 5% Critical Value	0.938

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-244	
MK Test Statistic	Z_{MK}	-3.05618	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:30:34 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$$x_2-x_1,\ x_3-x_1,\ \ldots,\ x_n-x_1,\ x_3-x_2,\ x_4-x_2,\ \ldots,\ x_n-x_{n-2},\ x_n-x_{n-1}$$

3. Let $\mathrm{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\mathrm{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\mathrm{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\mathrm{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \operatorname{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n > 10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \begin{cases} \dfrac{S-1}{\sqrt{VAR(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \dfrac{S+1}{\sqrt{VAR(S)}} & \text{if } S < 0 \end{cases} \qquad (3)$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Trichloroethene

Analysis was performed on the data that was entered for the location labeled "0386. The following table lists the data points that were used in the analysis.

0386							
Time	**Trichloroethene**		**Time**	**Trichloroethene**		**Time**	**Trichloroethene**
3/7/2005	4.5		7/28/2009	2.45		5/28/2013	2.35
5/13/2005	4.1		10/27/2009	3.11		8/26/2013	2.99
7/25/2005	3.8		1/26/2010	0.11		10/28/2013	2.56
9/27/2006	2.8		5/6/2010	0.94		2/19/2014	2.03
12/6/2006	4.2		7/27/2010	1.9		5/12/2014	0.97
2/26/2007	3.76		10/26/2010	3.09		8/27/2014	2.56
5/17/2007	1.49		1/25/2011	2.3		10/27/2014	2.67
8/17/2007	3.96		5/3/2011	0.11		1/28/2015	2.35
11/27/2007	0.345		7/28/2011	2.25		4/27/2015	2.11
2/26/2008	3.24		10/26/2011	1.93		7/28/2015	1.94
5/20/2008	0.2		2/1/2012	2.65		10/27/2015	2.14
8/6/2008	3.13		4/24/2012	2.25		1/26/2016	2.42
11/19/2008	2.39		7/24/2012	2.84		4/25/2016	2.56
2/11/2009	0.11		10/24/2012	3.04		7/27/2016	2.38
5/11/2009	1.28		2/12/2013	2.82		10/25/2016	2.54

SUMMARY STATISTICS for Trichloroethene								
n				45				
Min				0.11				
Max				4.5				
Range				4.39				
Mean				2.3481				
Median				2.42				
Variance				1.1929				
StdDev				1.0922				
Std Error				0.16282				
Skewness				-0.44407				
Interquartile Range				1.08				
Percentiles								
1%	5%	10%	25%	50%	75%	90%	95%	99%
0.11	0.11	0.287	1.935	2.42	3.015	3.864	4.17	4.5

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. These values with low weights can be considered possible outliers and are listed in the table below.

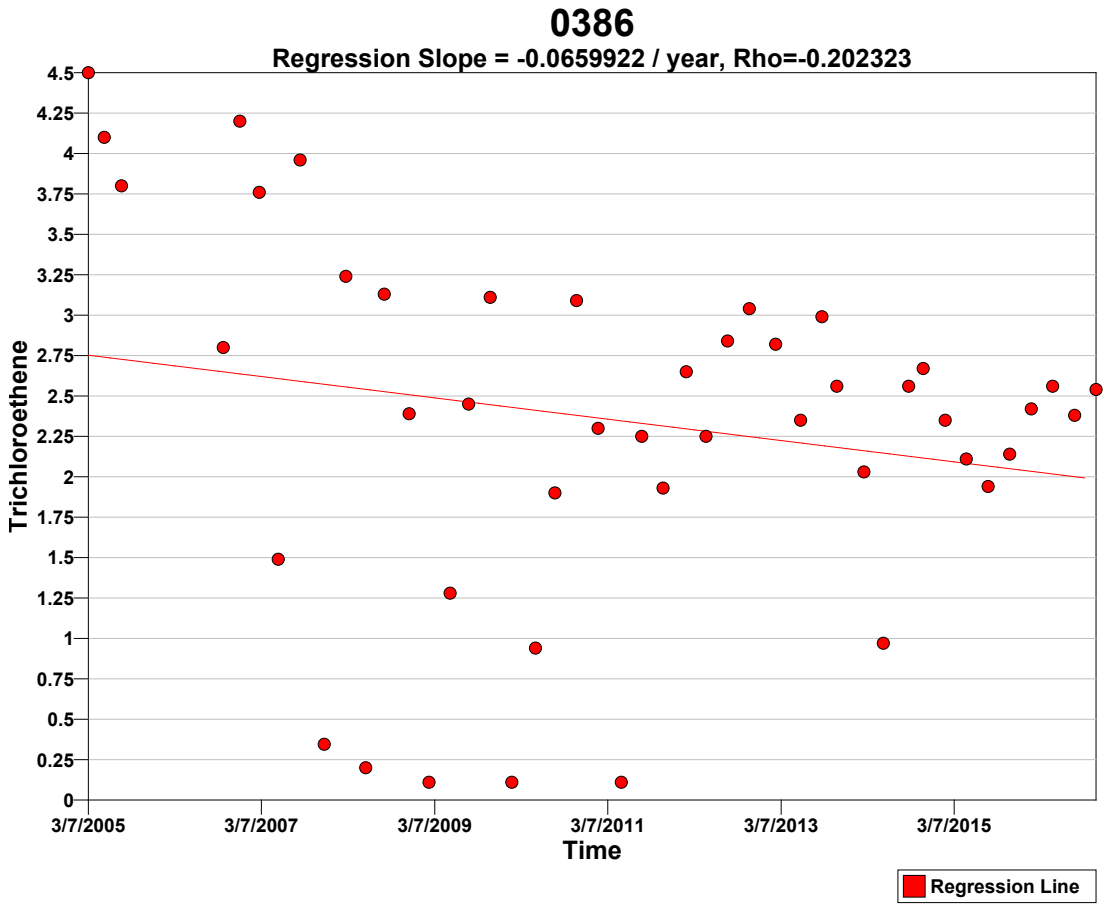
SUSPECTED OUTLIERS	
Date/Time	Value
11/27/2007	0.345
5/20/2008	0.2
2/11/2009	0.11
5/3/2011	0.11

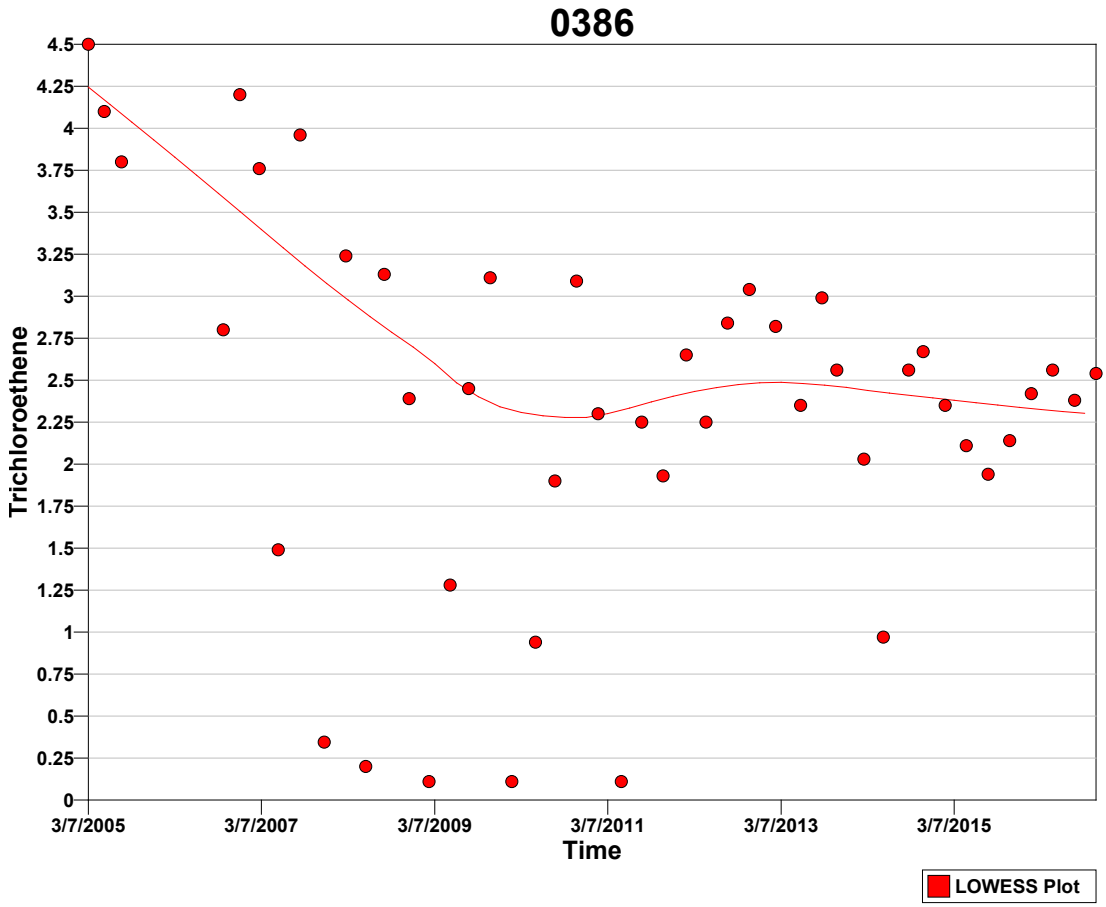
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots
The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.


0386
Regression Slope = -0.0659922 / year, Rho=-0.202323
Trichloroethene
4.5
4.25
4
3.75
3.5
3.25
3
2.75
2.5
2.25
2
1.75
1.5
1.25
1
0.75
0.5
0.25
0
3/7/2005
3/7/2007
3/7/2009
3/7/2011
3/7/2013
3/7/2015
Time
Regression Line



Tests

A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=45) is less than 50.

Normal Distribution Test for Trichloroethene	
Shapiro-Wilk Test Statistic	0.93685
Shapiro-Wilk 5% Critical Value	0.945

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test

The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-176	
MK Test Statistic	Z_{MK}	-1.71267	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:25:14 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$$x_2-x_1,\ x_3-x_1,\ \ldots,\ x_n-x_1,\ x_3-x_2,\ x_4-x_2,\ \ldots,\ x_n-x_{n-2},\ x_n-x_{n-1}$$

3. Let $\mathrm{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\mathrm{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\mathrm{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\mathrm{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \mathrm{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n > 10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \begin{cases} \dfrac{S-1}{\sqrt{VAR(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \dfrac{S+1}{\sqrt{VAR(S)}} & \text{if } S < 0 \end{cases} \qquad (3)$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Trichloroethene

Analysis was performed on the data that was entered for the location labeled "0386. The following table lists the data points that were used in the analysis.

0386							
Time	**Trichloroethene**		**Time**	**Trichloroethene**		**Time**	**Trichloroethene**
2/1/2012	2.65		10/28/2013	2.56		7/28/2015	1.94
4/24/2012	2.25		2/19/2014	2.03		10/27/2015	2.14
7/24/2012	2.84		5/12/2014	0.97		1/26/2016	2.42
10/24/2012	3.04		8/27/2014	2.56		4/25/2016	2.56
2/12/2013	2.82		10/27/2014	2.67		7/27/2016	2.38
5/28/2013	2.35		1/28/2015	2.35		10/25/2016	2.54
8/26/2013	2.99		4/27/2015	2.11			

SUMMARY STATISTICS for Trichloroethene								
n				20				
Min				0.97				
Max				3.04				
Range				2.07				
Mean				2.4085				
Median				2.48				
Variance				0.20692				
StdDev				0.45488				
Std Error				0.10171				
Skewness				-1.5787				
Interquartile Range				0.4975				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0.97	1.019	1.949	2.168	2.48	2.665	2.975	3.038	3.04

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. These values with low weights can be considered possible outliers and are listed in the table below.

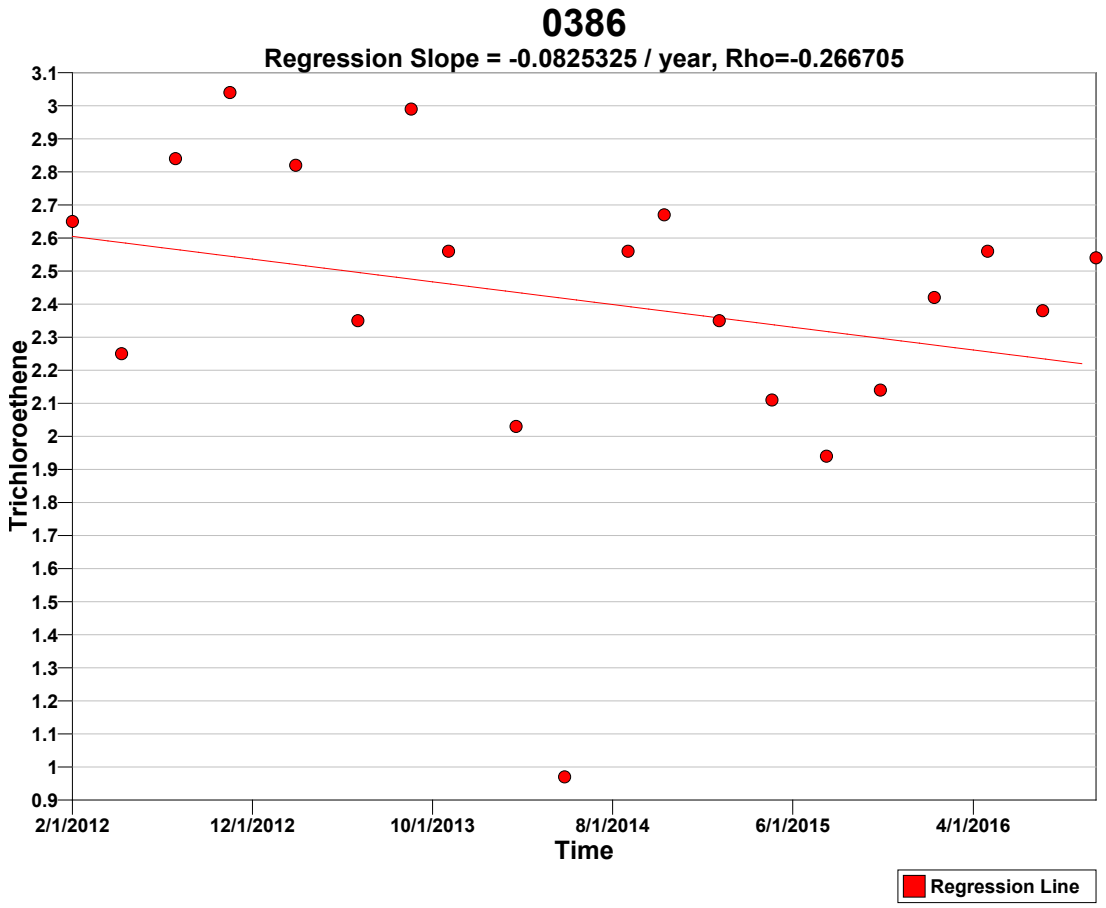
SUSPECTED OUTLIERS	
Date/Time	**Value**
5/12/2014	0.97

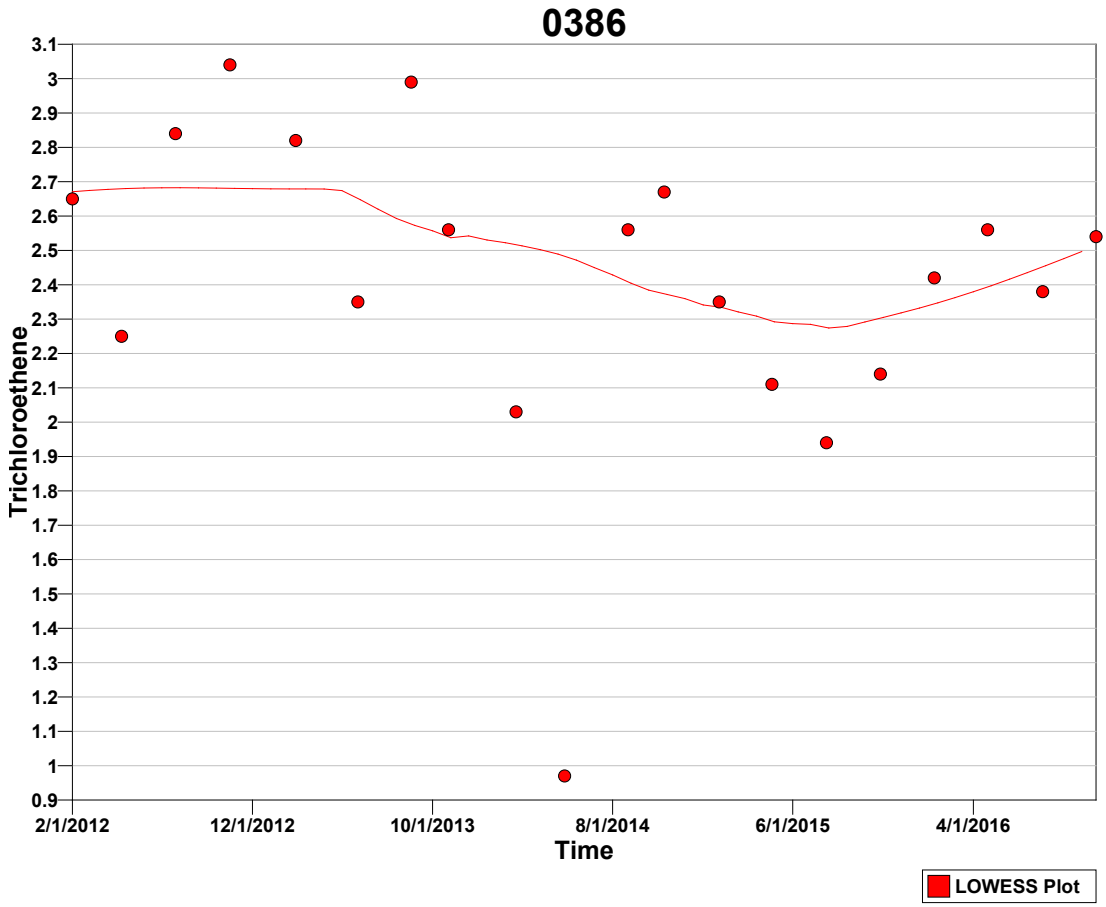
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots
The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.


0386
Regression Slope = -0.0825325 / year, Rho=-0.266705
Trichloroethene
3.1
3
2.9
2.8
2.7
2.6
2.5
2.4
2.3
2.2
2.1
2
1.9
1.8
1.7
1.6
1.5
1.4
1.3
1.2
1.1
1
0.9
2/1/2012
12/1/2012
10/1/2013
8/1/2014
6/1/2015
4/1/2016
Time
Regression Line



Tests

A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=20) is less than 50.

Normal Distribution Test for Trichloroethene	
Shapiro-Wilk Test Statistic	0.88101
Shapiro-Wilk 5% Critical Value	0.905

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test

The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-44	
MK Test Statistic	Z_{MK}	-1.39854	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	**Accept**
Alternative Hypothesis	H_a	A downward monotonic trend exists	Reject
Conclude with 95% confidence that a downward monotonic trend does not exist.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 4/10/2017 7:36:09 AM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$$x_2-x_1,\ x_3-x_1,\ \ldots,\ x_n-x_1,\ x_3-x_2,\ x_4-x_2,\ \ldots,\ x_n-x_{n-2},\ x_n-x_{n-1}$$

3. Let $\mathrm{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\mathrm{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\mathrm{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\mathrm{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \operatorname{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n > 10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \begin{cases} \dfrac{S-1}{\sqrt{VAR(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \dfrac{S+1}{\sqrt{VAR(S)}} & \text{if } S < 0 \end{cases} \qquad (3)$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Trichloroethene

Analysis was performed on the data that was entered for the location labeled "0389. The following table lists the data points that were used in the analysis.

0389							
Time	**Trichloroethene**		**Time**	**Trichloroethene**		**Time**	**Trichloroethene**
7/25/2005	1.1		1/26/2010	1.48		10/28/2013	0.55
9/27/2006	0.98		5/6/2010	0.59		2/19/2014	0.19
12/6/2006	1.01		7/27/2010	0.33		5/12/2014	0.16
2/26/2007	1.14		10/26/2010	1.39		8/27/2014	0.44
5/17/2007	0.881		1/25/2011	0.99		10/27/2014	0.61
8/17/2007	1.22		5/3/2011	0.54		1/28/2015	0.23
11/27/2007	1.28		7/28/2011	0.24		4/27/2015	0.16
2/26/2008	0.398		10/26/2011	0.72		7/28/2015	0.18
5/20/2008	0.52		2/1/2012	0.31		10/27/2015	0.29
8/6/2008	0.408		4/24/2012	0.16		1/26/2016	0.35
11/19/2008	1.19		7/24/2012	0.32		4/25/2016	0.16
2/11/2009	1.07		10/24/2012	0.9		7/27/2016	0.8
5/11/2009	0.256		2/12/2013	0.53		10/25/2016	0.74
7/28/2009	0.894		5/28/2013	0.16			
10/27/2009	0.887		8/26/2013	0.16			

SUMMARY STATISTICS for Trichloroethene								
n				43				
Min				0.16				
Max				1.48				
Range				1.32				
Mean				0.62591				
Median				0.54				
Variance				0.15736				
StdDev				0.39669				
Std Error				0.060494				
Skewness				0.46969				
Interquartile Range				0.724				
Percentiles								
1%	5%	10%	25%	50%	75%	90%	95%	99%
0.16	0.16	0.16	0.256	0.54	0.98	1.208	1.368	1.48

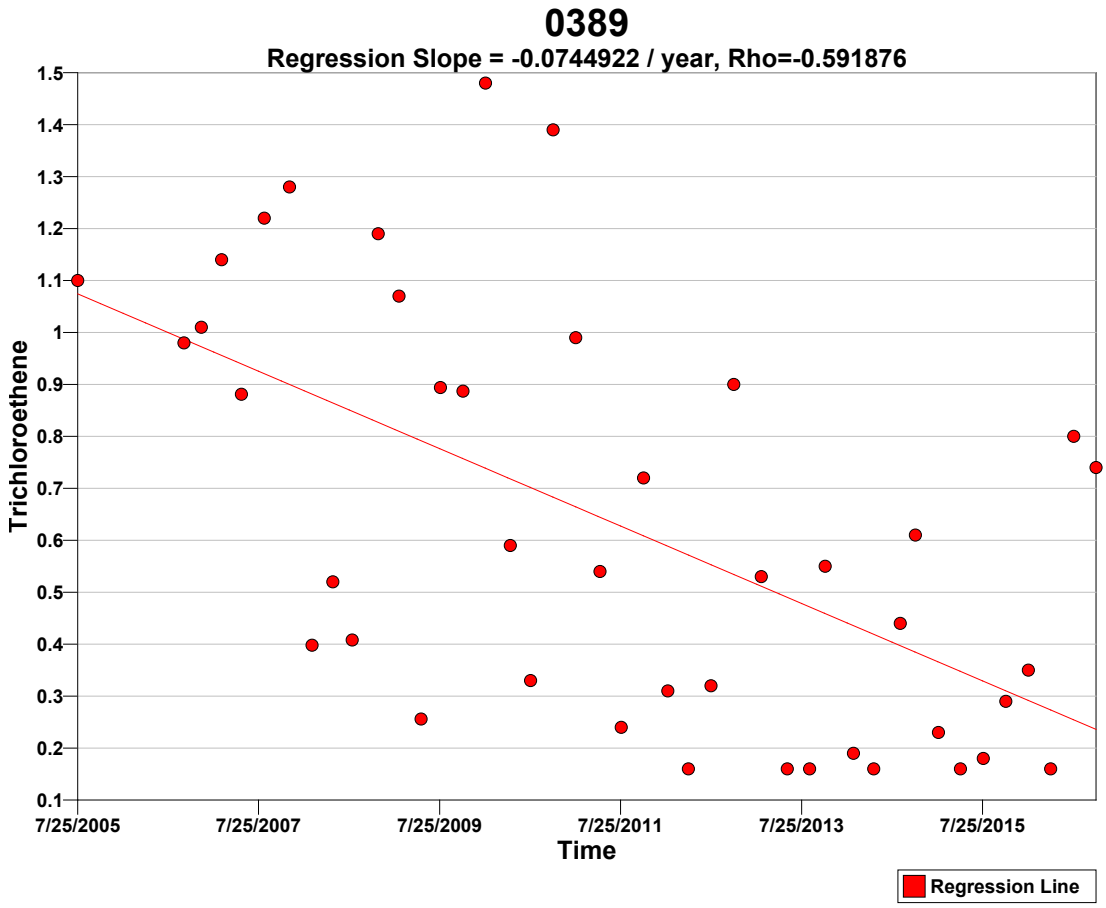
Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. The LOWESS method did not detect any outliers for the current data set.

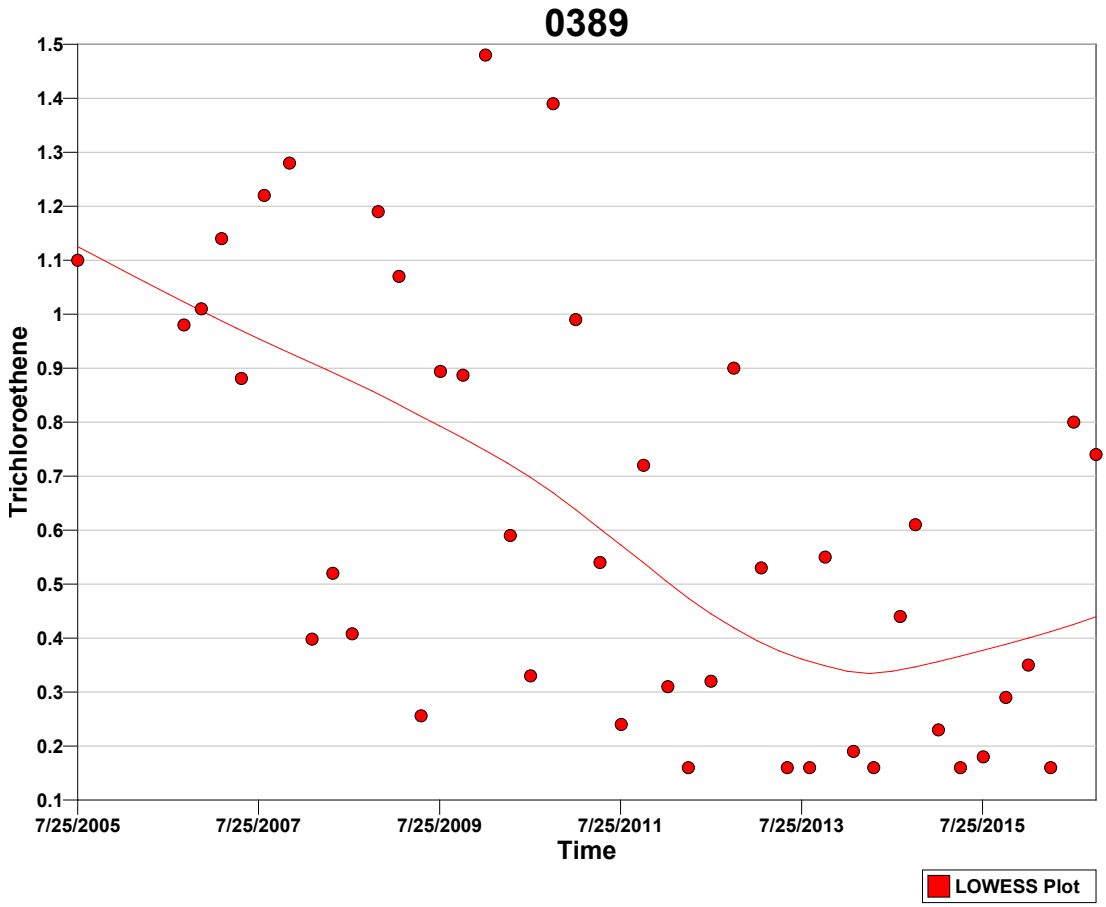
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots
The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.


0389
Regression Slope = -0.0744922 / year, Rho=-0.591876
Trichloroethene
1.5
1.4
1.3
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
7/25/2005
7/25/2007
7/25/2009
7/25/2011
7/25/2013
7/25/2015
Time
Regression Line



Tests

A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=43) is less than 50.

Normal Distribution Test for Trichloroethene	
Shapiro-Wilk Test Statistic	0.90702
Shapiro-Wilk 5% Critical Value	0.943

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test

The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-362	
MK Test Statistic	Z_{MK}	-3.78389	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:26:13 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$$x_2-x_1,\ x_3-x_1,\ \ldots,\ x_n-x_1,\ x_3-x_2,\ x_4-x_2,\ \ldots,\ x_n-x_{n-2},\ x_n-x_{n-1}$$

3. Let $\mathrm{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\mathrm{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\mathrm{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\mathrm{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \operatorname{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n > 10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \begin{cases} \dfrac{S-1}{\sqrt{VAR(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \dfrac{S+1}{\sqrt{VAR(S)}} & \text{if } S < 0 \end{cases} \qquad (3)$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Trichloroethene

Analysis was performed on the data that was entered for the location labeled "0389. The following table lists the data points that were used in the analysis.

0389							
Time	**Trichloroethene**		**Time**	**Trichloroethene**		**Time**	**Trichloroethene**
2/1/2012	0.31		10/28/2013	0.55		7/28/2015	0.18
4/24/2012	0.16		2/19/2014	0.19		10/27/2015	0.29
7/24/2012	0.32		5/12/2014	0.16		1/26/2016	0.35
10/24/2012	0.9		8/27/2014	0.44		4/25/2016	0.16
2/12/2013	0.53		10/27/2014	0.61		7/27/2016	0.8
5/28/2013	0.16		1/28/2015	0.23		10/25/2016	0.74
8/26/2013	0.16		4/27/2015	0.16			

SUMMARY STATISTICS for Trichloroethene								
n				20				
Min				0.16				
Max				0.9				
Range				0.74				
Mean				0.37				
Median				0.3				
Variance				0.057305				
StdDev				0.23939				
Std Error				0.053528				
Skewness				0.9859				
Interquartile Range				0.385				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0.16	0.16	0.16	0.16	0.3	0.545	0.794	0.895	0.9

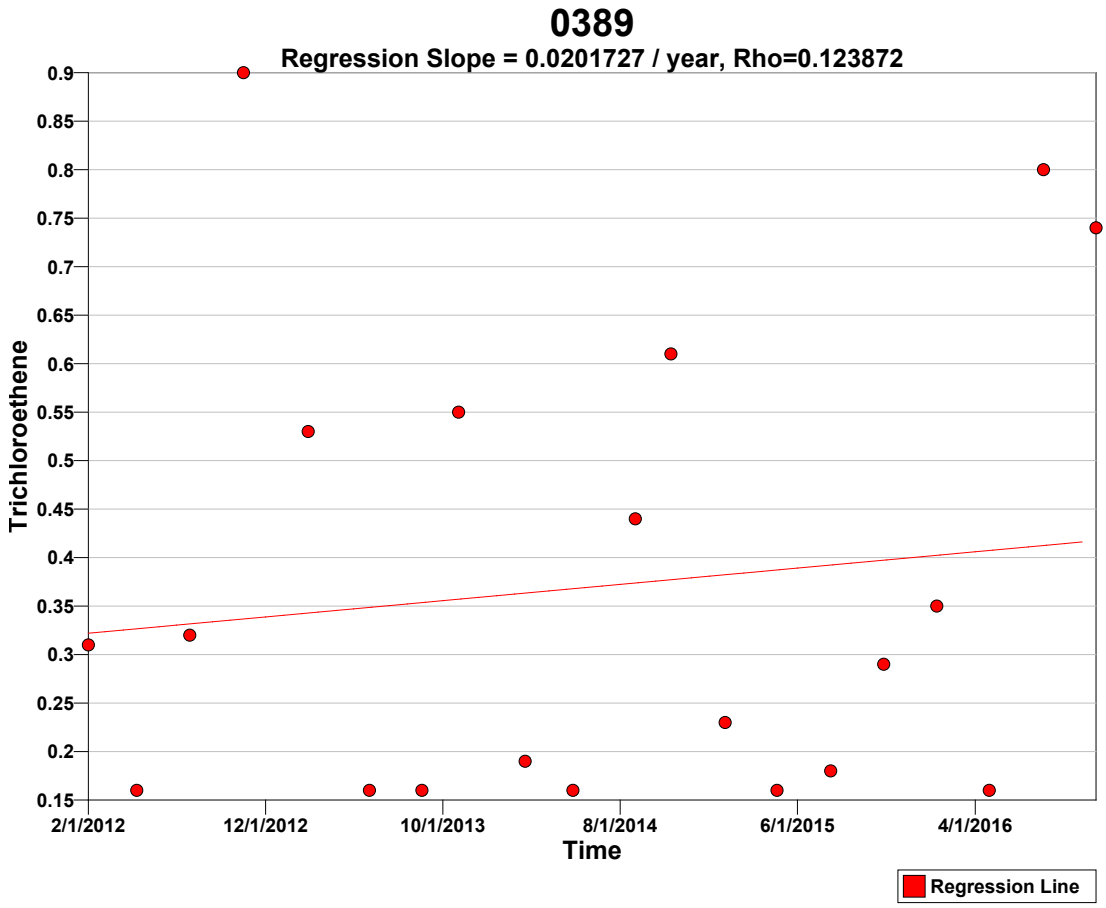
Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. The LOWESS method did not detect any outliers for the current data set.

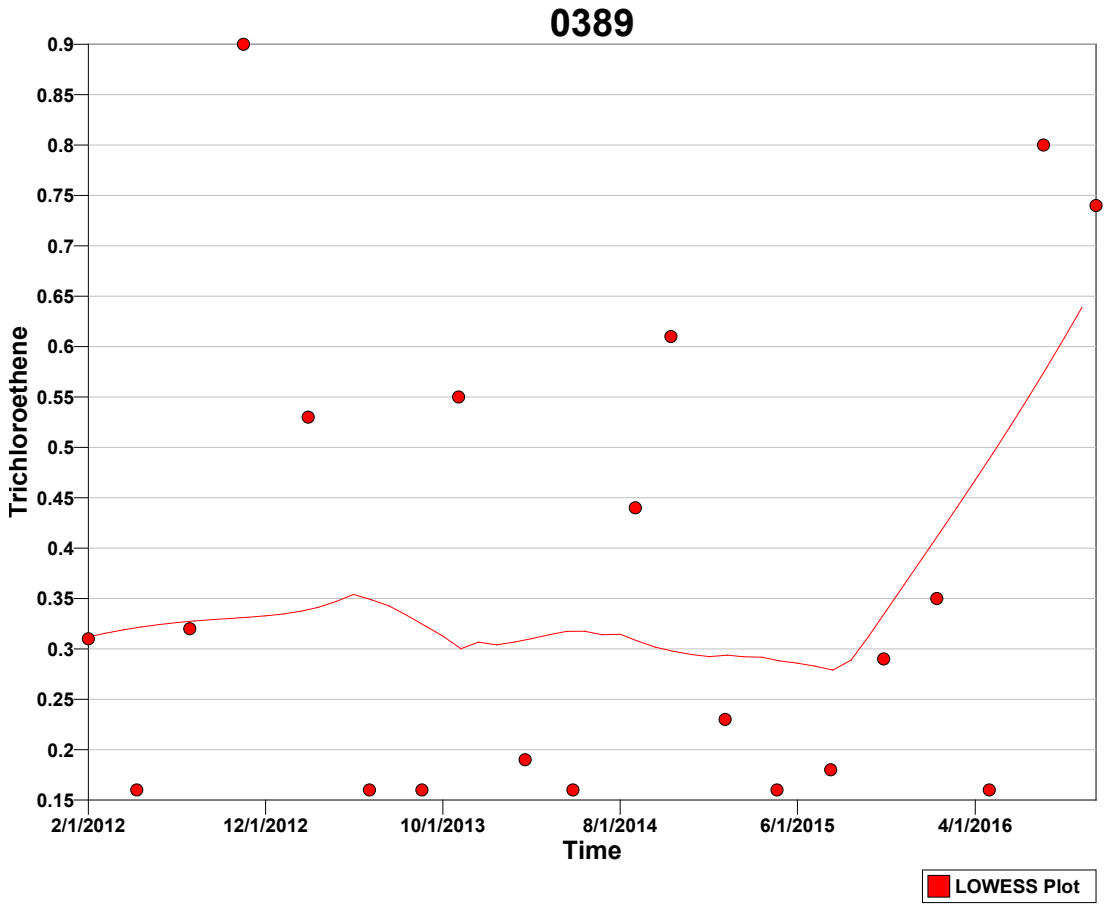
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots
The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.


0389
Regression Slope = 0.0201727 / year, Rho=0.123872
Trichloroethene
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
2/1/2012
12/1/2012
10/1/2013
8/1/2014
6/1/2015
4/1/2016
Time
Regression Line



Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=20) is less than 50.

Normal Distribution Test for Trichloroethene	
Shapiro-Wilk Test Statistic	0.8382
Shapiro-Wilk 5% Critical Value	0.905

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	19	
MK Test Statistic	Z_{MK}	0.592906	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	**Accept**
Alternative Hypothesis	H_a	A downward monotonic trend exists	Reject
Conclude with 95% confidence that a downward monotonic trend does not exist.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 4/10/2017 7:36:47 AM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$$x_2-x_1,\ x_3-x_1,\ \ldots,\ x_n-x_1,\ x_3-x_2,\ x_4-x_2,\ \ldots,\ x_n-x_{n-2},\ x_n-x_{n-1}$$

3. Let $\mathrm{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\mathrm{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\mathrm{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\mathrm{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \operatorname{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \begin{cases} \dfrac{S-1}{\sqrt{VAR(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \dfrac{S+1}{\sqrt{VAR(S)}} & \text{if } S < 0 \end{cases} \qquad (3)$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for cis-1x2-Dichloroethene

Analysis was performed on the data that was entered for the location labeled "0411. The following table lists the data points that were used in the analysis.

0411							
Time	**cis-1x2-Dichloroethene**		**Time**	**cis-1x2-Dichloroethene**		**Time**	**cis-1x2-Dichloroethene**
4/29/1999	4.8		10/22/2003	1.6		8/11/2008	1.33
6/21/1999	2.2		1/22/2004	2		2/10/2009	3.36
9/2/1999	1.8		4/21/2004	1.3		7/29/2009	2.89
2/1/2000	2.6		7/12/2004	1.4		1/27/2010	3.97
4/17/2000	2.2		11/17/2004	1.3		7/29/2010	2.56
7/7/2000	3		2/28/2005	1		1/26/2011	3.32
9/27/2000	1		5/24/2005	1.5		7/25/2011	2.8
2/9/2001	1.6		8/2/2005	1.7		1/30/2012	1.81
4/26/2001	1.5		11/1/2005	2.4		7/26/2012	2.17
7/20/2001	2.4		1/31/2006	3.7		2/14/2013	4.71
10/18/2001	1.9		4/26/2006	2.9		8/28/2013	3.57
1/29/2002	1.9		8/4/2006	1.95		2/24/2014	2.33
5/6/2002	2.8		11/17/2006	2.76		8/26/2014	2.28
7/24/2002	3.5		2/28/2007	1.7		1/27/2015	3.34
11/8/2002	1.4		5/23/2007	1.56		7/29/2015	1.69
1/28/2003	2		8/22/2007	2.73		1/25/2016	0.65
4/22/2003	1.8		11/13/2007	2.74		7/25/2016	0.94
7/25/2003	1.6		2/19/2008	2.21			

SUMMARY STATISTICS for cis-1x2-Dichloroethene								
n				53				
Min				0.65				
Max				4.8				
Range				4.15				
Mean				2.2674				
Median				2.17				
Variance				0.84282				
StdDev				0.91805				
Std Error				0.1261				
Skewness				0.78279				
Interquartile Range				1.2				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0.65	0.982	1.3	1.6	2.17	2.8	3.542	4.192	4.8

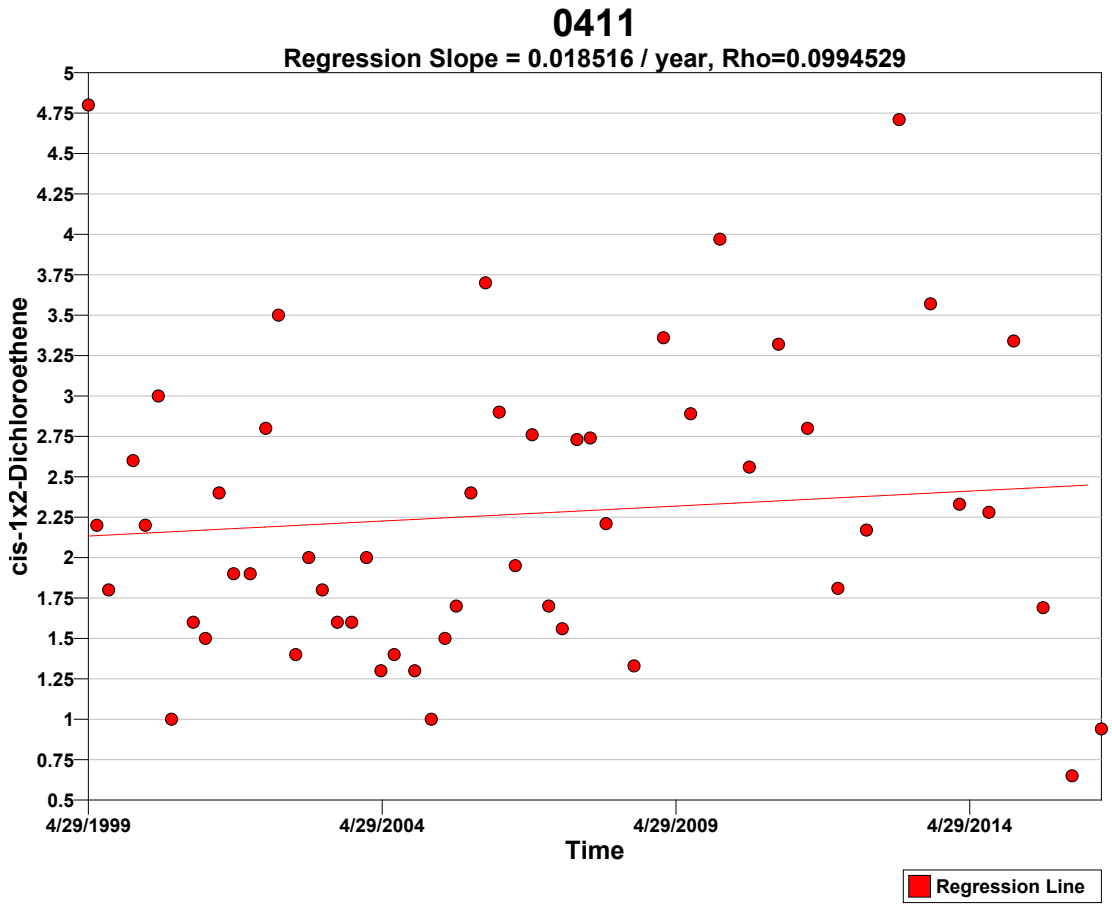
Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. The LOWESS method did not detect any outliers for the current data set.

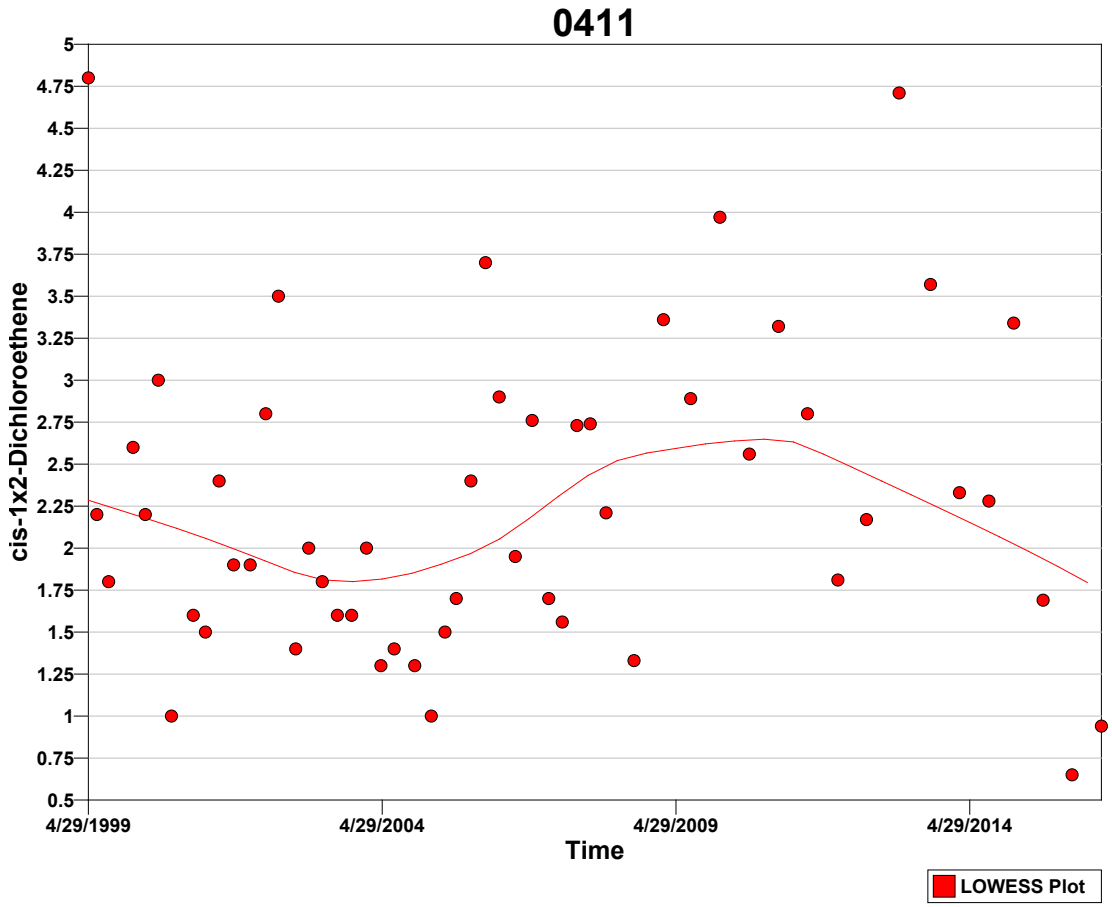
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots
The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.


0411
Regression Slope = 0.018516 / year, Rho=0.0994529
cis-1x2-Dichloroethene
5
4.75
4.5
4.25
4
3.75
3.5
3.25
3
2.75
2.5
2.25
2
1.75
1.5
1.25
1
0.75
0.5
4/29/1999
4/29/2004
4/29/2009
4/29/2014
Time
Regression Line



Tests

A goodness-of-fit test was performed to test whether the values are normally distributed. The Lilliefors test was determined most appropriate for this VSP application because the size of the data set (n=53) is greater than 50.

Normal Distribution Test for cis-1x2-Dichloroethene	
Lilliefors Test Statistic	0.10513
Lilliefors 5% Critical Value	0.1217

The calculated test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test

The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	112	
MK Test Statistic	Z_{MK}	0.851816	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	**Accept**
Alternative Hypothesis	H_a	A downward monotonic trend exists	Reject
Conclude with 95% confidence that a downward monotonic trend does not exist.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:42:26 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$$x_2-x_1,\ x_3-x_1,\ \ldots,\ x_n-x_1,\ x_3-x_2,\ x_4-x_2,\ \ldots,\ x_n-x_{n-2},\ x_n-x_{n-1}$$

3. Let $\mathrm{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\mathrm{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\mathrm{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\mathrm{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \operatorname{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n > 10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \begin{cases} \dfrac{S-1}{\sqrt{VAR(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \dfrac{S+1}{\sqrt{VAR(S)}} & \text{if } S < 0 \end{cases} \qquad (3)$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Trichloroethene

Analysis was performed on the data that was entered for the location labeled "0411. The following table lists the data points that were used in the analysis.

0411							
Time	**Trichloroethene**		**Time**	**Trichloroethene**		**Time**	**Trichloroethene**
4/29/1999	21		10/22/2003	11		8/11/2008	14.1
6/21/1999	18		1/22/2004	10		2/10/2009	12.7
9/2/1999	21		4/21/2004	9		7/29/2009	11.2
2/1/2000	22		7/12/2004	10		1/27/2010	10.1
4/17/2000	13		11/17/2004	9		7/29/2010	9.62
7/7/2000	16		2/28/2005	11		1/26/2011	10.6
9/27/2000	14		5/24/2005	11		7/25/2011	9.42
2/9/2001	14		8/2/2005	11		1/30/2012	13.4
4/26/2001	12		11/1/2005	14		7/26/2012	12.7
7/20/2001	13		1/31/2006	11		2/14/2013	12.5
10/18/2001	14		4/26/2006	9.2		8/28/2013	13.3
1/29/2002	8.4		8/4/2006	14.4		2/24/2014	12
5/6/2002	16		11/17/2006	12.2		8/26/2014	10.6
7/24/2002	13		2/28/2007	12.7		1/27/2015	10.8
11/8/2002	12		5/23/2007	12.1		7/29/2015	10.3
1/28/2003	13		8/22/2007	15.2		1/25/2016	11.9
4/22/2003	12		11/13/2007	12.3		7/25/2016	11.7
7/25/2003	11		2/19/2008	12.2			

SUMMARY STATISTICS for Trichloroethene								
n				53				
Min				8.4				
Max				22				
Range				13.6				
Mean				12.597				
Median				12.1				
Variance				8.3614				
StdDev				2.8916				
Std Error				0.39719				
Skewness				1.5831				
Interquartile Range				2.8				
Percentiles								
1%	5%	10%	25%	50%	75%	90%	95%	99%
8.4	9	9.5	10.9	12.1	13.7	16	21	22

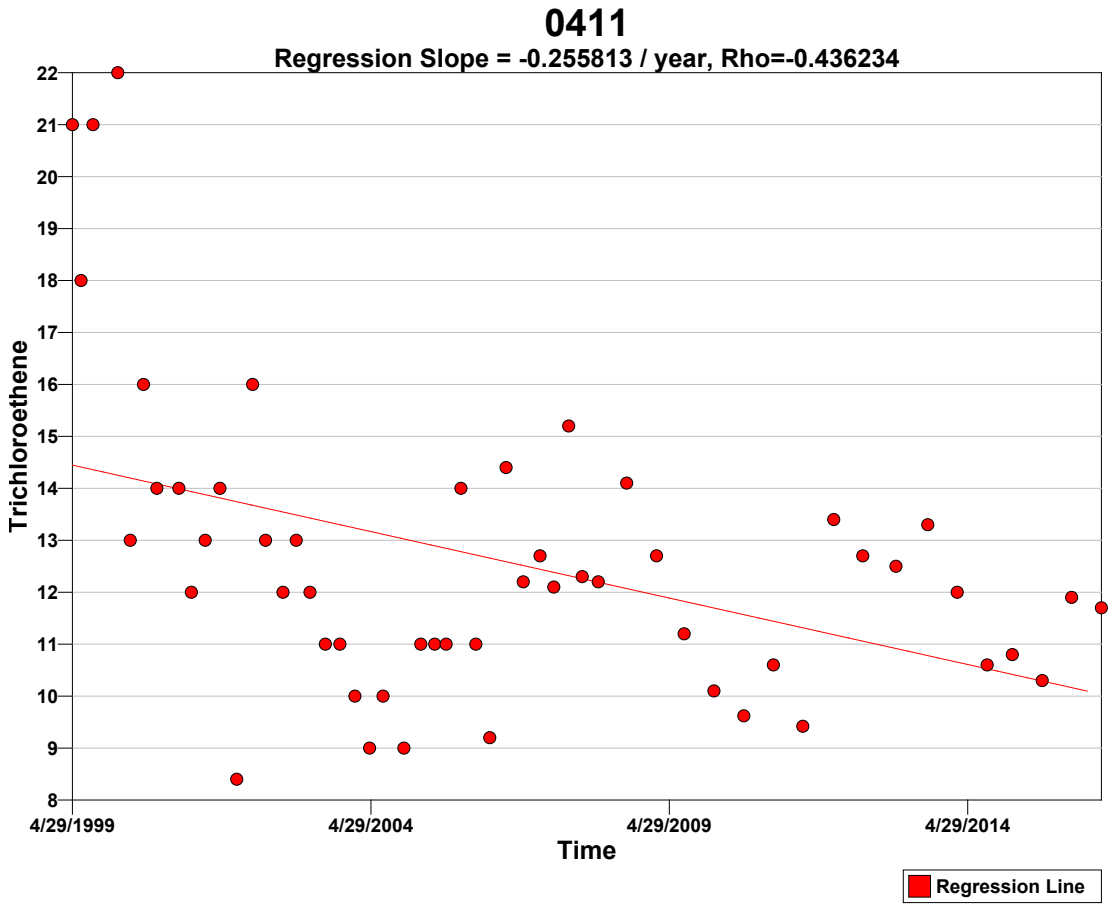
Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. The LOWESS method did not detect any outliers for the current data set.

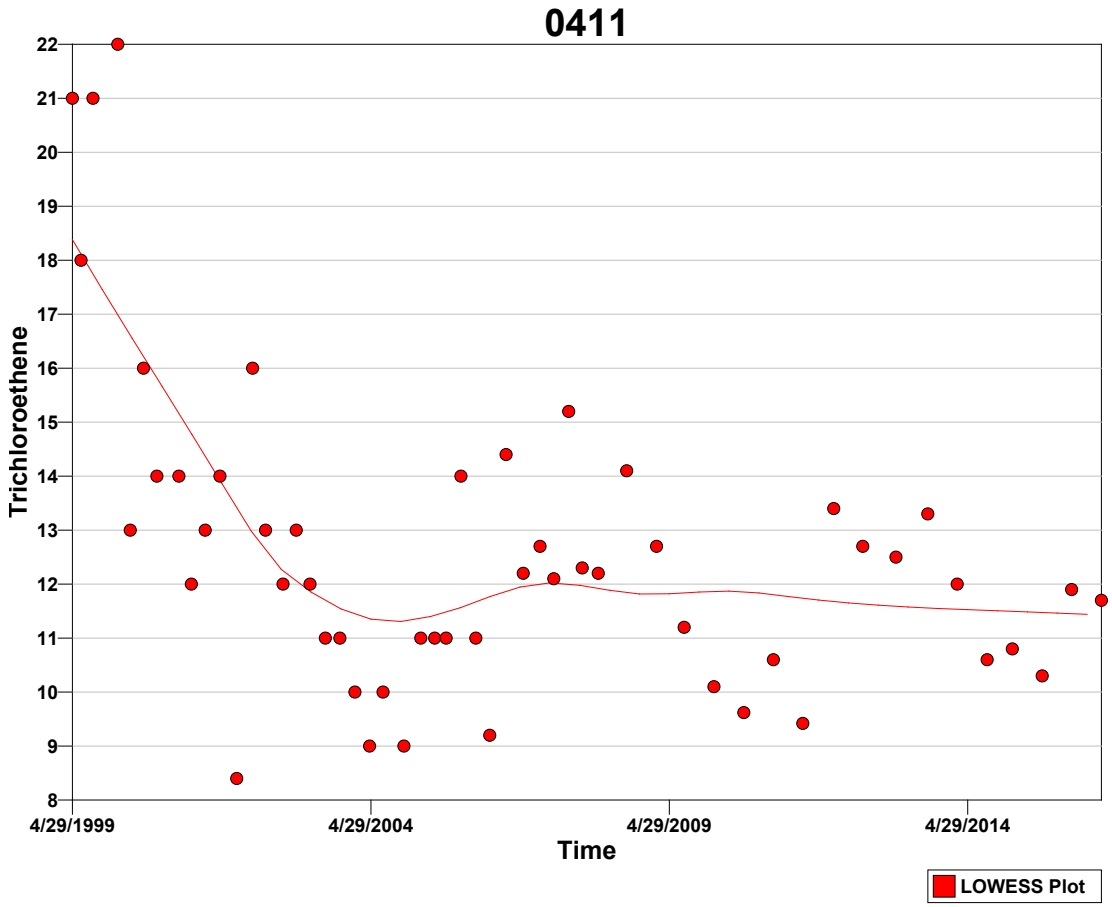
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots
The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.


0411
Regression Slope = -0.255813 / year, Rho=-0.436234
Trichloroethene
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
4/29/1999
4/29/2004
4/29/2009
4/29/2014
Time
Regression Line



Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Lilliefors test was determined most appropriate for this VSP application because the size of the data set (n=53) is greater than 50.

Normal Distribution Test for Trichloroethene	
Lilliefors Test Statistic	0.16156
Lilliefors 5% Critical Value	0.1217

The calculated test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-380	
MK Test Statistic	Z_{MK}	-2.91269	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:44:40 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$$x_2-x_1,\ x_3-x_1,\ \ldots,\ x_n-x_1,\ x_3-x_2,\ x_4-x_2,\ \ldots,\ x_n-x_{n-2},\ x_n-x_{n-1}$$

3. Let $\mathrm{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\mathrm{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\mathrm{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\mathrm{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \operatorname{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n > 10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \begin{cases} \dfrac{S-1}{\sqrt{VAR(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \dfrac{S+1}{\sqrt{VAR(S)}} & \text{if } S < 0 \end{cases} \qquad (3)$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for cis-1x2-Dichloroethene

Analysis was performed on the data that was entered for the location labeled "0443. The following table lists the data points that were used in the analysis.

0443							
Time	**cis-1x2-Dichloroethene**		**Time**	**cis-1x2-Dichloroethene**		**Time**	**cis-1x2-Dichloroethene**
5/6/2002	1		11/16/2005	1.1		7/29/2010	0.28
7/26/2002	1		1/31/2006	1.9		1/26/2011	0.49
11/8/2002	1		4/26/2006	1		7/25/2011	0.22
1/28/2003	1		8/4/2006	0.341		1/30/2012	0.75
4/22/2003	1		11/17/2006	0.438		7/26/2012	0.27
7/25/2003	1		2/28/2007	0.919		2/14/2013	0.57
10/22/2003	1		5/23/2007	5		8/28/2013	0.38
1/22/2004	1		8/22/2007	0.708		2/24/2014	0.54
4/21/2004	1		11/13/2007	0.872		8/26/2014	0.36
7/12/2004	1		2/19/2008	0.83		1/27/2015	0.27
11/18/2004	0.97		12/30/2008	1.04		7/29/2015	0.33
3/2/2005	1		2/10/2009	0.762		1/25/2016	0.23
5/24/2005	1		7/30/2009	0.32		7/25/2016	0.33
8/3/2005	1		1/27/2010	0.451			

SUMMARY STATISTICS for cis-1x2-Dichloroethene								
n				41				
Min				0.22				
Max				5				
Range				4.78				
Mean				0.84563				
Median				0.872				
Variance				0.56874				
StdDev				0.75415				
Std Error				0.11778				
Skewness				4.3455				
Interquartile Range				0.63				
Percentiles								
1%	5%	10%	25%	50%	75%	90%	95%	99%
0.22	0.234	0.272	0.37	0.872	1	1.032	1.82	5

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. These values with low weights can be considered possible outliers and are listed in the table below.

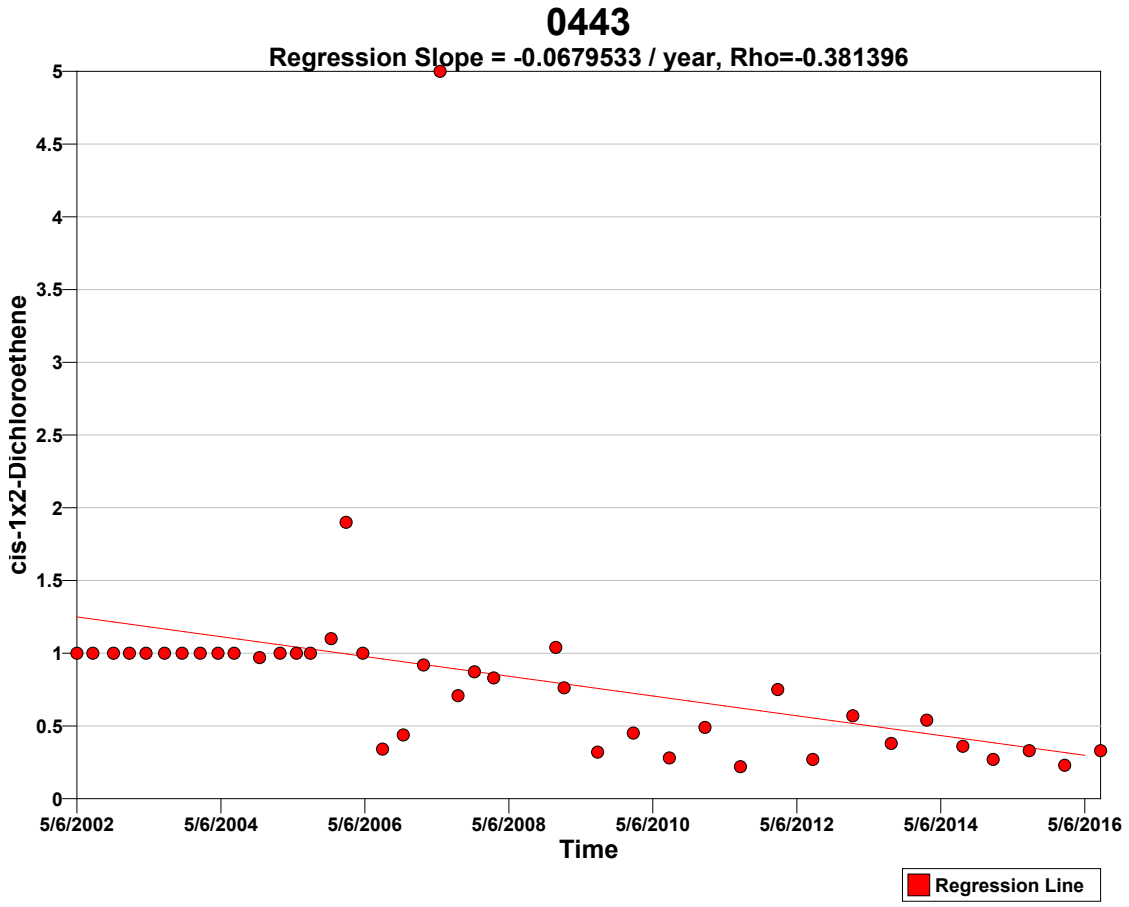
SUSPECTED OUTLIERS	
Date/Time	Value
1/31/2006	1.9
8/4/2006	0.341
11/17/2006	0.438
5/23/2007	5
8/22/2007	0.708
12/30/2008	1.04
7/30/2009	0.32
7/29/2010	0.28
7/25/2011	0.22
1/30/2012	0.75
7/26/2012	0.27

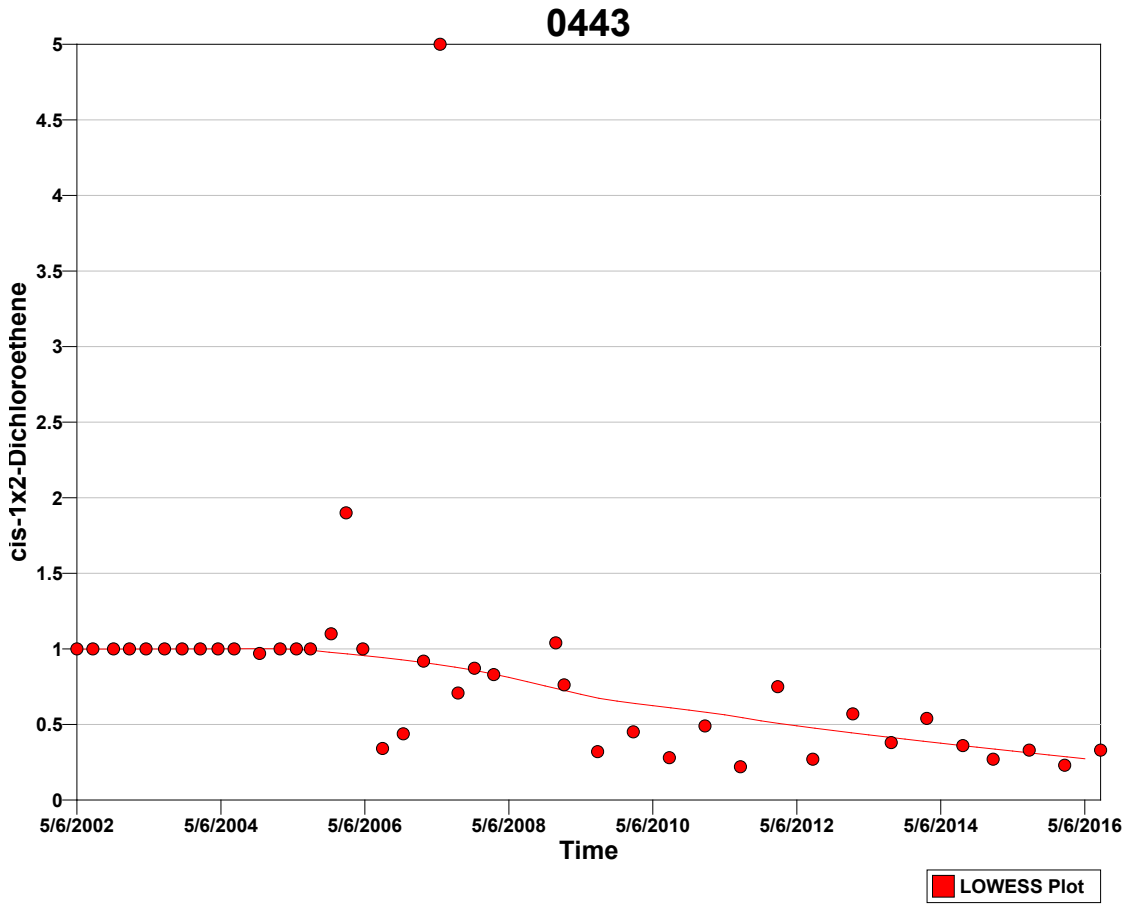
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots
The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.





Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=41) is less than 50.

Normal Distribution Test for cis-1x2-Dichloroethene	
Shapiro-Wilk Test Statistic	0.555
Shapiro-Wilk 5% Critical Value	0.941

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-437	
MK Test Statistic	Z_{MK}	-5.00423	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:43:36 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$$x_2-x_1,\ x_3-x_1,\ \ldots,\ x_n-x_1,\ x_3-x_2,\ x_4-x_2,\ \ldots,\ x_n-x_{n-2},\ x_n-x_{n-1}$$

3. Let $\mathrm{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\mathrm{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\mathrm{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\mathrm{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \operatorname{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n > 10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \begin{cases} \dfrac{S-1}{\sqrt{VAR(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \dfrac{S+1}{\sqrt{VAR(S)}} & \text{if } S < 0 \end{cases} \qquad (3)$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Trichloroethene
Analysis was performed on the data that was entered for the location labeled "0443. The following table lists the data points that were used in the analysis.

0443							
Time	**Trichloroethene**		**Time**	**Trichloroethene**		**Time**	**Trichloroethene**
5/6/2002	9.2		11/16/2005	12		7/29/2010	5.3
7/26/2002	3.5		1/31/2006	9.3		1/26/2011	7.47
11/8/2002	6.6		4/26/2006	7.9		7/25/2011	5.73
1/28/2003	5.4		8/4/2006	6.34		1/30/2012	14
4/22/2003	7.8		11/17/2006	7.33		7/26/2012	5.32
7/25/2003	2.8		2/28/2007	12.3		2/14/2013	7.25
10/22/2003	7.2		5/23/2007	3.06		8/28/2013	8.85
1/22/2004	6.3		8/22/2007	11.7		2/24/2014	11.2
4/21/2004	6.5		11/13/2007	12.1		8/26/2014	7.45
7/12/2004	3.2		2/19/2008	11.2		1/27/2015	6.03
11/18/2004	10		12/30/2008	11		7/29/2015	5.87
3/2/2005	11		2/10/2009	8.36		1/25/2016	7.53
5/24/2005	2.2		7/30/2009	3.99		7/25/2016	7.27
8/3/2005	6.6		1/27/2010	7.47			

SUMMARY STATISTICS for Trichloroethene								
n				41				
Min				2.2				
Max				14				
Range				11.8				
Mean				7.6005				
Median				7.33				
Variance				8.3825				
StdDev				2.8953				
Std Error				0.45216				
Skewness				0.20267				
Interquartile Range				3.85				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
2.2	2.826	3.26	5.8	7.33	9.65	11.94	12.28	14

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. The LOWESS method did not detect any outliers for the current data set.

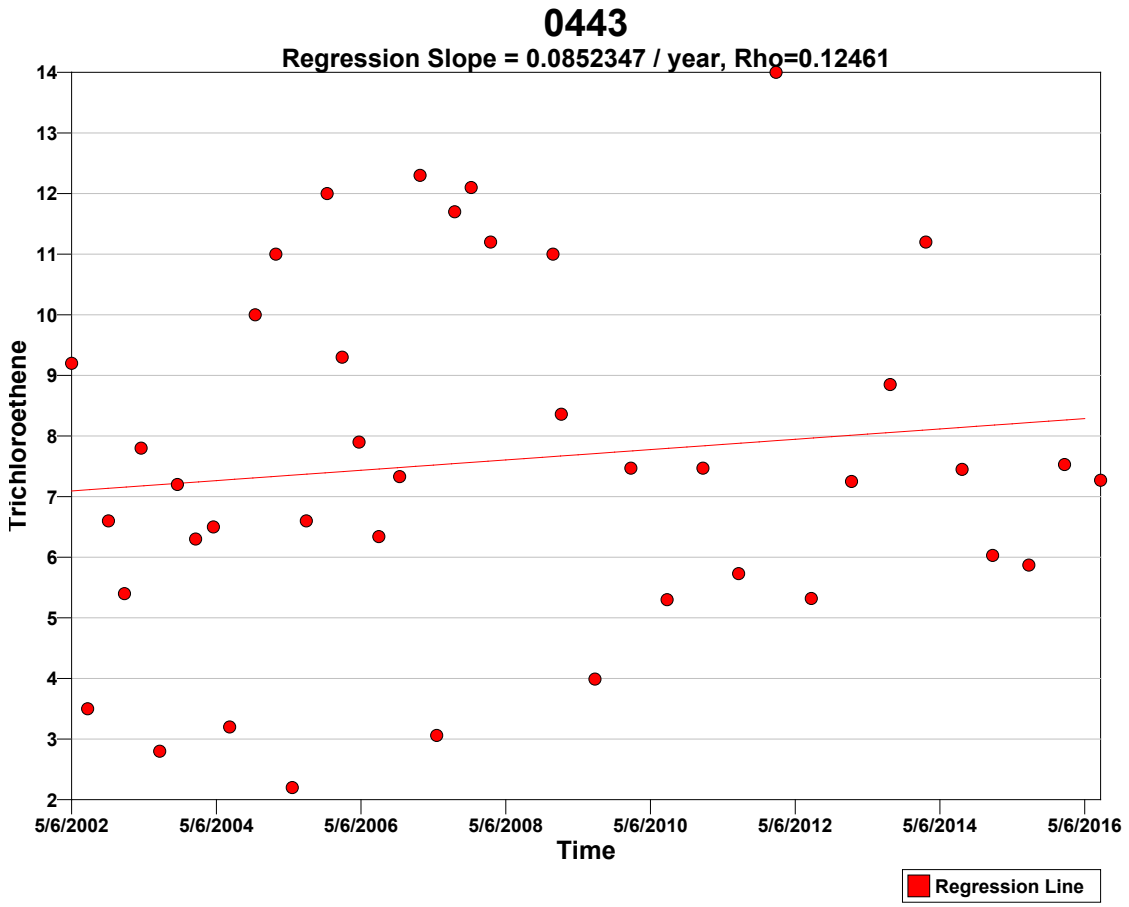
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

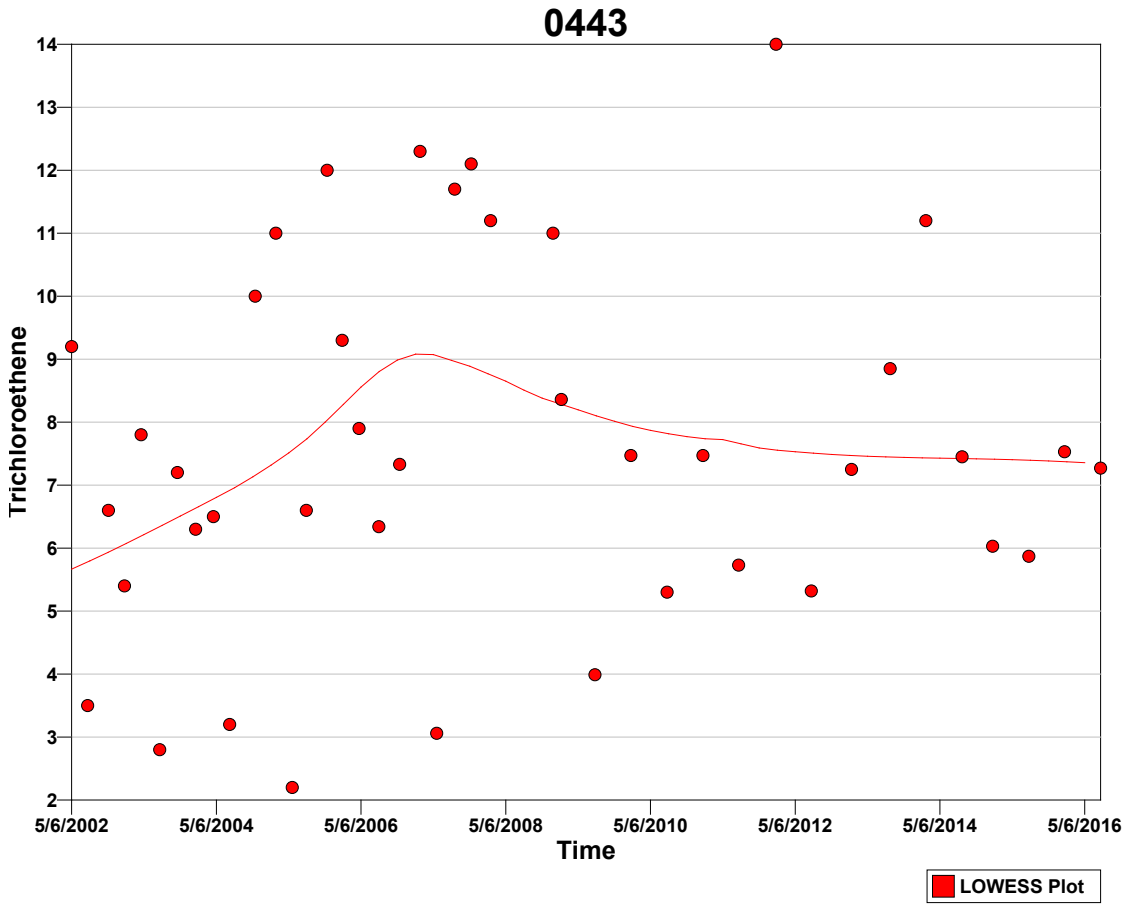
It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots

The Time vs. Measured Values Plot shows the best fitting “least squares” linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated “least squares” slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.





Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=41) is less than 50.

Normal Distribution Test for Trichloroethene	
Shapiro-Wilk Test Statistic	0.96412
Shapiro-Wilk 5% Critical Value	0.941

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	76	
MK Test Statistic	Z_{MK}	0.842608	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	**Accept**
Alternative Hypothesis	H_a	A downward monotonic trend exists	Reject
Conclude with 95% confidence that a downward monotonic trend does not exist.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:45:24 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$$x_2-x_1,\ x_3-x_1,\ \ldots,\ x_n-x_1,\ x_3-x_2,\ x_4-x_2,\ \ldots,\ x_n-x_{n-2},\ x_n-x_{n-1}$$

3. Let $\mathrm{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\mathrm{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\mathrm{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\mathrm{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \operatorname{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n > 10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \begin{cases} \dfrac{S-1}{\sqrt{VAR(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \dfrac{S+1}{\sqrt{VAR(S)}} & \text{if } S < 0 \end{cases} \qquad (3)$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Tetrachloroethene
Analysis was performed on the data that was entered for the location labeled "0601. The following table lists the data points that were used in the analysis.

0601							
Time	**Tetrachloroethene**		**Time**	**Tetrachloroethene**		**Time**	**Tetrachloroethene**
2/17/2005	29		2/9/2009	11.9		2/11/2013	7.88
5/17/2005	34		5/5/2009	19.3		5/30/2013	8.49
7/28/2005	52		7/29/2009	18.5		8/27/2013	7.59
11/3/2005	14		10/26/2009	18.9		10/28/2013	6.26
1/30/2006	11		1/25/2010	6.15		2/19/2014	3.86
5/18/2006	15		5/5/2010	0.2		5/12/2014	5.53
9/26/2006	18		7/26/2010	11.4		8/26/2014	12.1
12/5/2006	15.6		10/25/2010	6.13		10/27/2014	4.64
2/27/2007	10.6		1/25/2011	4.33		1/29/2015	14.9
5/17/2007	12.2		5/10/2011	11.5		4/27/2015	12.8
8/22/2007	29.4		7/27/2011	9.59		7/29/2015	12.2
11/29/2007	18		10/25/2011	9.29		10/28/2015	1.31
2/28/2008	9.86		1/31/2012	10.2		1/25/2016	8.73
5/19/2008	16		4/25/2012	0.16		4/25/2016	12.2
8/4/2008	12.7		7/24/2012	0.16		7/25/2016	5.54
11/17/2008	12.9		10/25/2012	7.29		10/24/2016	19.1

SUMMARY STATISTICS for Tetrachloroethene								
n	48							
Min	0.16							
Max	52							
Range	51.84							
Mean	12.466							
Median	11.45							
Variance	84.848							
StdDev	9.2113							
Std Error	1.3295							
Skewness	2.0481							
Interquartile Range	8.9325							
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0.16	0.178	3.605	6.518	11.45	15.45	20.27	31.93	52

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. These values with low weights can be considered possible outliers and are listed in the table below.

SUSPECTED OUTLIERS	
Date/Time	**Value**
7/28/2005	52

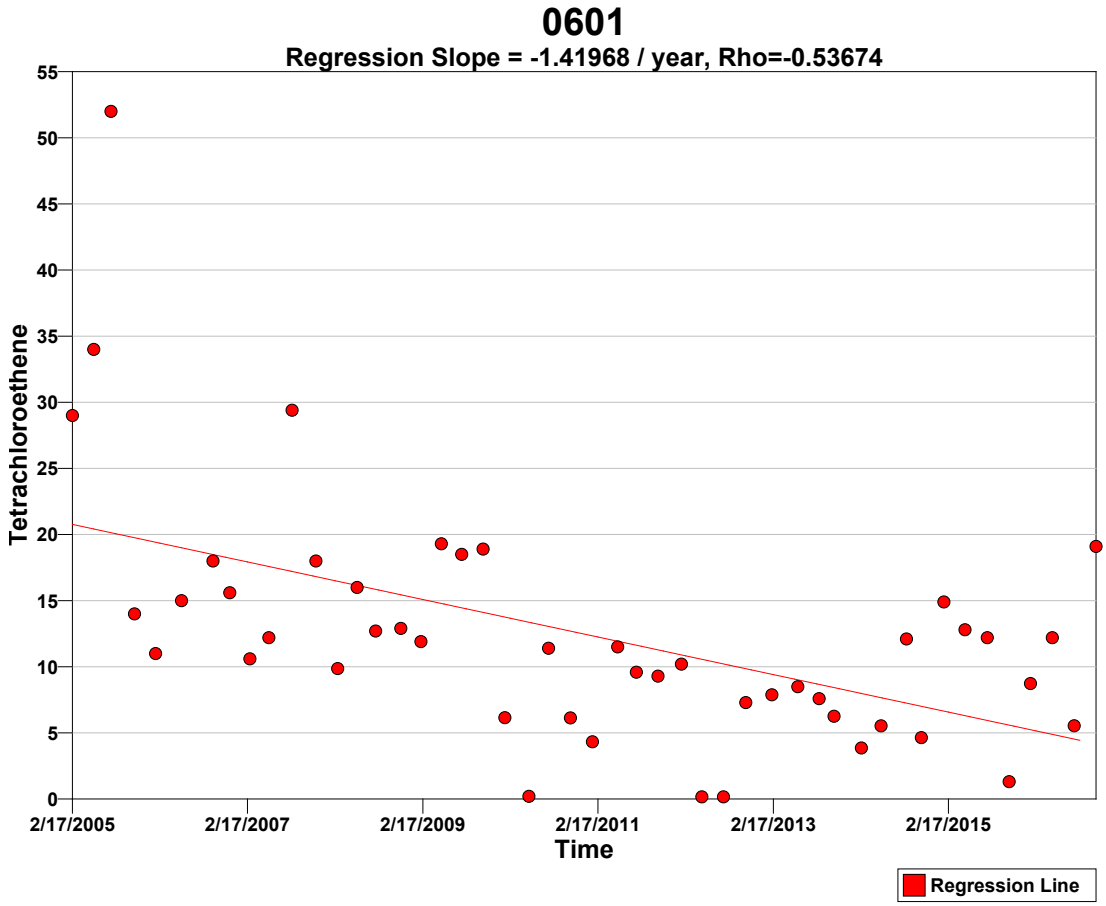
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

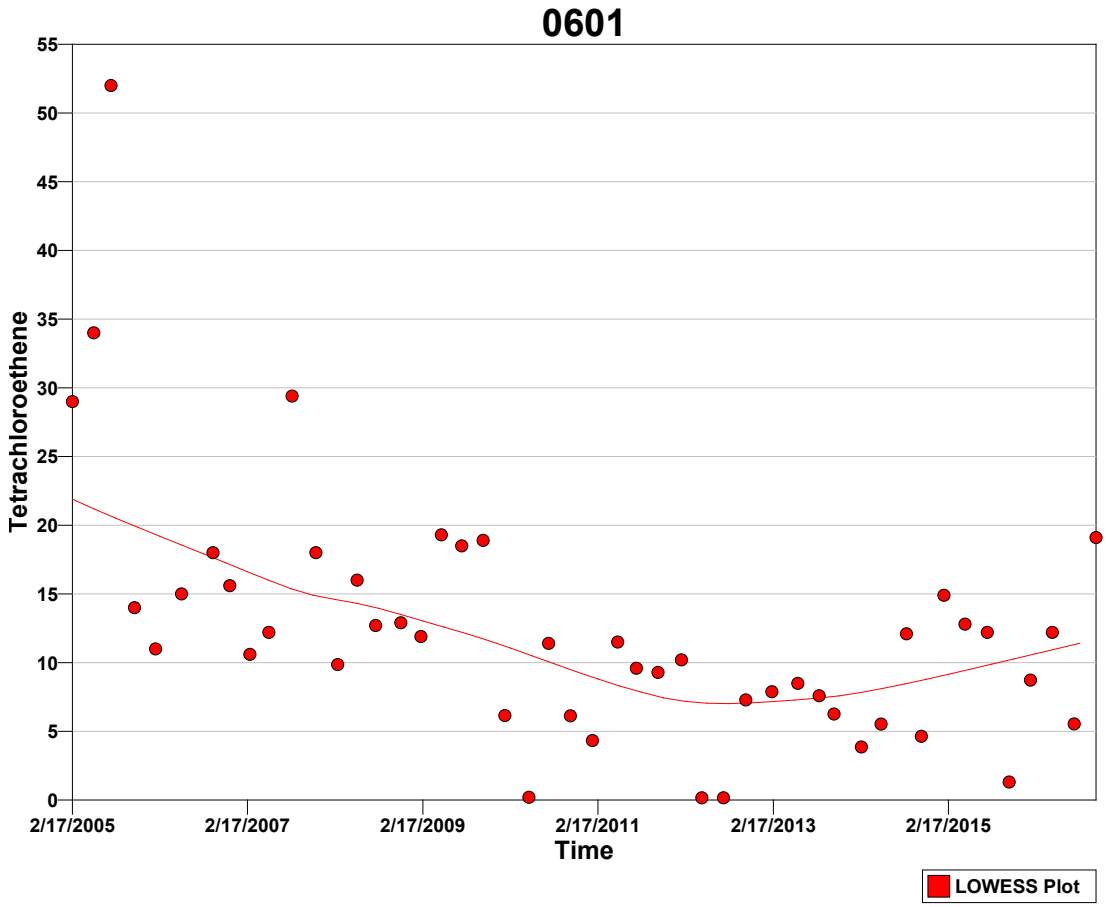
It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots

The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.





Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=48) is less than 50.

Normal Distribution Test for Tetrachloroethene	
Shapiro-Wilk Test Statistic	0.84241
Shapiro-Wilk 5% Critical Value	0.947

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-409	
MK Test Statistic	Z_{MK}	-3.62713	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:37:12 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$x_2-x_1, x_3-x_1, \ldots, x_n-x_1, x_3-x_2, x_4-x_2, \ldots, x_n-x_{n-2}, x_n-x_{n-1}$

3. Let $\text{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\text{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\text{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\text{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \text{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \frac{S-1}{\sqrt{VAR(S)}} \quad \text{if} \quad S > 0$$

$$= 0 \quad \text{if} \quad S = 0 \qquad (3)$$

$$= \frac{S+1}{\sqrt{VAR(S)}} \quad \text{if} \quad S < 0$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Trichloroethene
Analysis was performed on the data that was entered for the location labeled "0601. The following table lists the data points that were used in the analysis.

0601							
Time	**Trichloroethene**		**Time**	**Trichloroethene**		**Time**	**Trichloroethene**
2/17/2005	4.3		2/9/2009	3.55		2/11/2013	6.3
5/17/2005	4.7		5/5/2009	5.28		5/30/2013	6.1
7/28/2005	4.9		7/29/2009	5.71		8/27/2013	6.52
11/3/2005	5.8		10/26/2009	5.97		10/28/2013	7.34
1/30/2006	4.8		1/25/2010	1.95		2/19/2014	1.98
5/18/2006	3.9		5/5/2010	0.11		5/12/2014	2.8
9/26/2006	8		7/26/2010	7.46		8/26/2014	6.73
12/5/2006	6.18		10/25/2010	7.48		10/27/2014	8.03
2/27/2007	5.08		1/25/2011	3.94		1/29/2015	6.68
5/17/2007	6.32		5/10/2011	4.62		4/27/2015	3.98
8/22/2007	6.84		7/27/2011	5.78		7/29/2015	6.94
11/29/2007	5.25		10/25/2011	4.95		10/28/2015	1.04
2/28/2008	3.41		1/31/2012	4.16		1/25/2016	6.52
5/19/2008	3.63		4/25/2012	0.16		4/25/2016	7.3
8/4/2008	5.83		7/24/2012	0.16		7/25/2016	5.28
11/17/2008	4.82		10/25/2012	10.3		10/24/2016	7.85

SUMMARY STATISTICS for Trichloroethene								
n				48				
Min				0.11				
Max				10.3				
Range				10.19				
Mean				5.1402				
Median				5.28				
Variance				4.7819				
StdDev				2.1868				
Std Error				0.31563				
Skewness				-0.54879				
Interquartile Range				2.69				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0.11	0.16	1.859	3.95	5.28	6.64	7.517	8.016	10.3

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. The LOWESS method did not detect any outliers for the current data set.

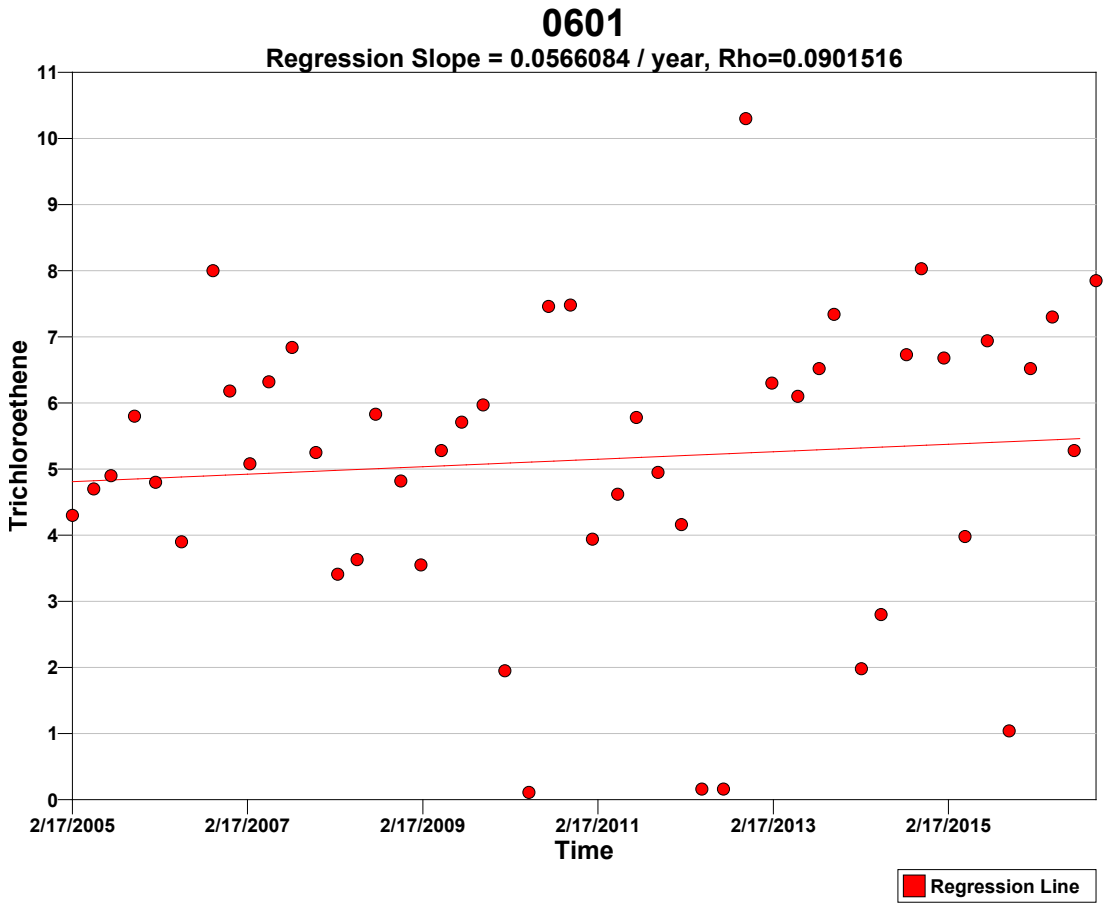
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

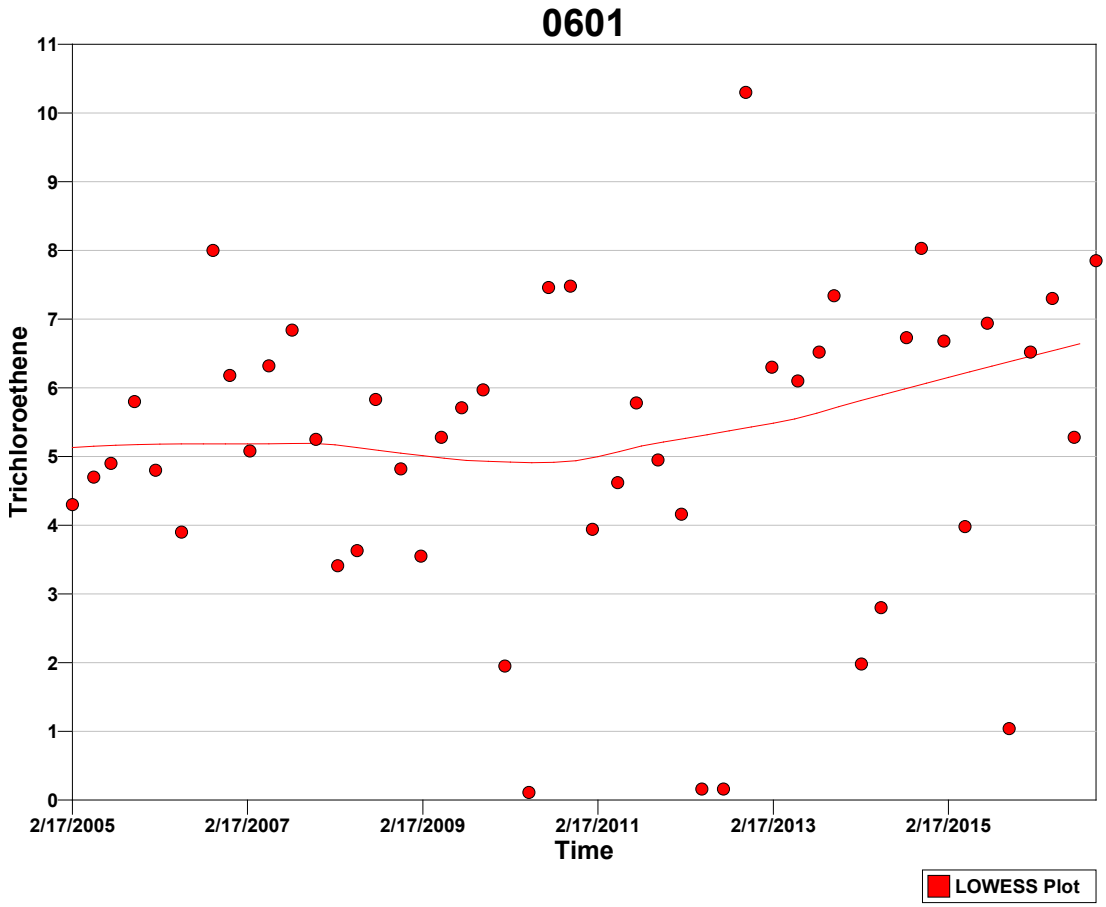
It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots

The Time vs. Measured Values Plot shows the best fitting “least squares” linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated “least squares” slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.





Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=48) is less than 50.

Normal Distribution Test for Trichloroethene	
Shapiro-Wilk Test Statistic	0.95604
Shapiro-Wilk 5% Critical Value	0.947

The calculated test statistic exceeds the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	159	
MK Test Statistic	Z_{MK}	1.40448	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	**Accept**
Alternative Hypothesis	H_a	A downward monotonic trend exists	Reject
Conclude with 95% confidence that a downward monotonic trend does not exist.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:32:43 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$x_2-x_1, x_3-x_1, \ldots, x_n-x_1, x_3-x_2, x_4-x_2, \ldots, x_n-x_{n-2}, x_n-x_{n-1}$

3. Let $\text{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\text{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\text{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\text{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \text{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \frac{S-1}{\sqrt{VAR(S)}} \quad \text{if} \quad S > 0$$

$$= 0 \quad \text{if} \quad S = 0 \qquad (3)$$

$$= \frac{S+1}{\sqrt{VAR(S)}} \quad \text{if} \quad S < 0$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Trichloroethene
Analysis was performed on the data that was entered for the location labeled "0601. The following table lists the data points that were used in the analysis.

0601							
Time	**Trichloroethene**		**Time**	**Trichloroethene**		**Time**	**Trichloroethene**
1/31/2012	4.16		10/28/2013	7.34		7/29/2015	6.94
4/25/2012	0.16		2/19/2014	1.98		10/28/2015	1.04
7/24/2012	0.16		5/12/2014	2.8		1/25/2016	6.52
10/25/2012	10.3		8/26/2014	6.73		4/25/2016	7.3
2/11/2013	6.3		10/27/2014	8.03		7/25/2016	5.28
5/30/2013	6.1		1/29/2015	6.68		10/24/2016	7.85
8/27/2013	6.52		4/27/2015	3.98			

SUMMARY STATISTICS for Trichloroethene								
n				20				
Min				0.16				
Max				10.3				
Range				10.14				
Mean				5.3085				
Median				6.41				
Variance				7.8775				
StdDev				2.8067				
Std Error				0.6276				
Skewness				-0.56918				
Interquartile Range				4.115				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0.16	0.16	0.248	3.095	6.41	7.21	8.012	10.19	10.3

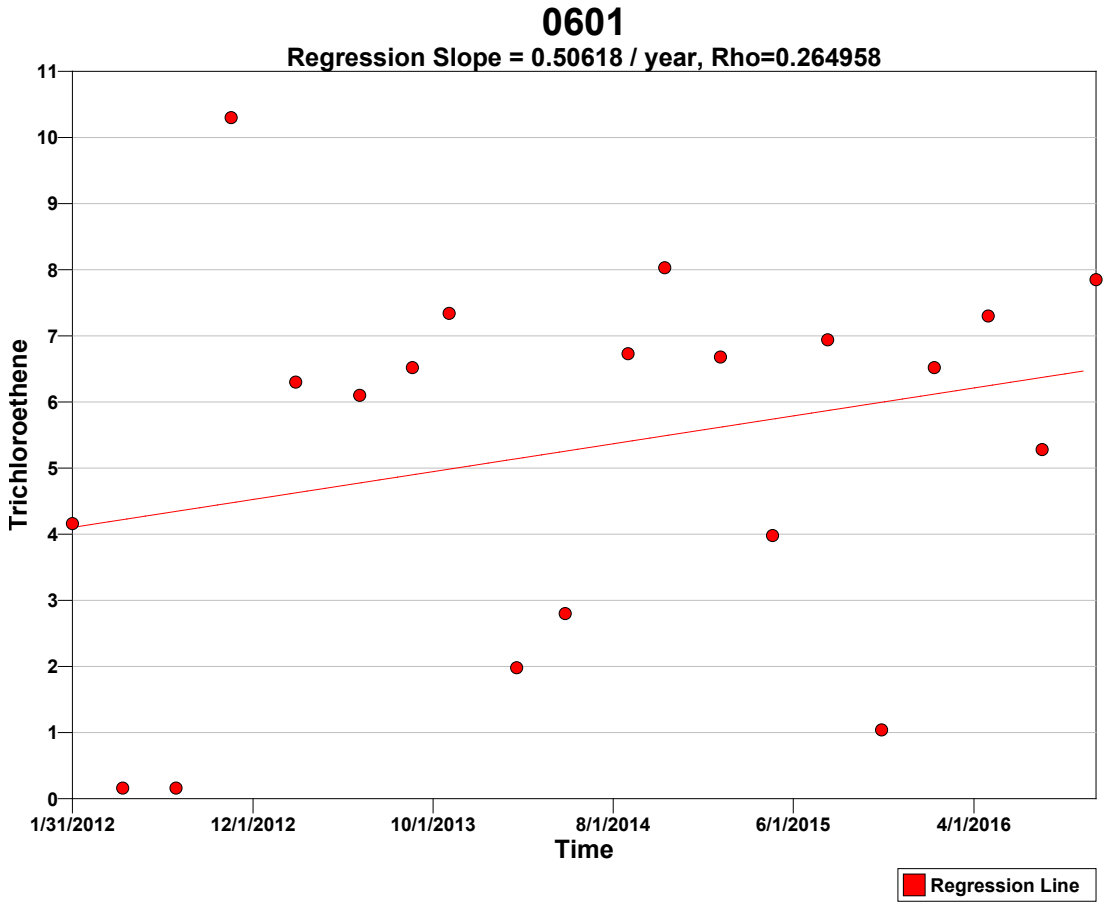
Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. The LOWESS method did not detect any outliers for the current data set.

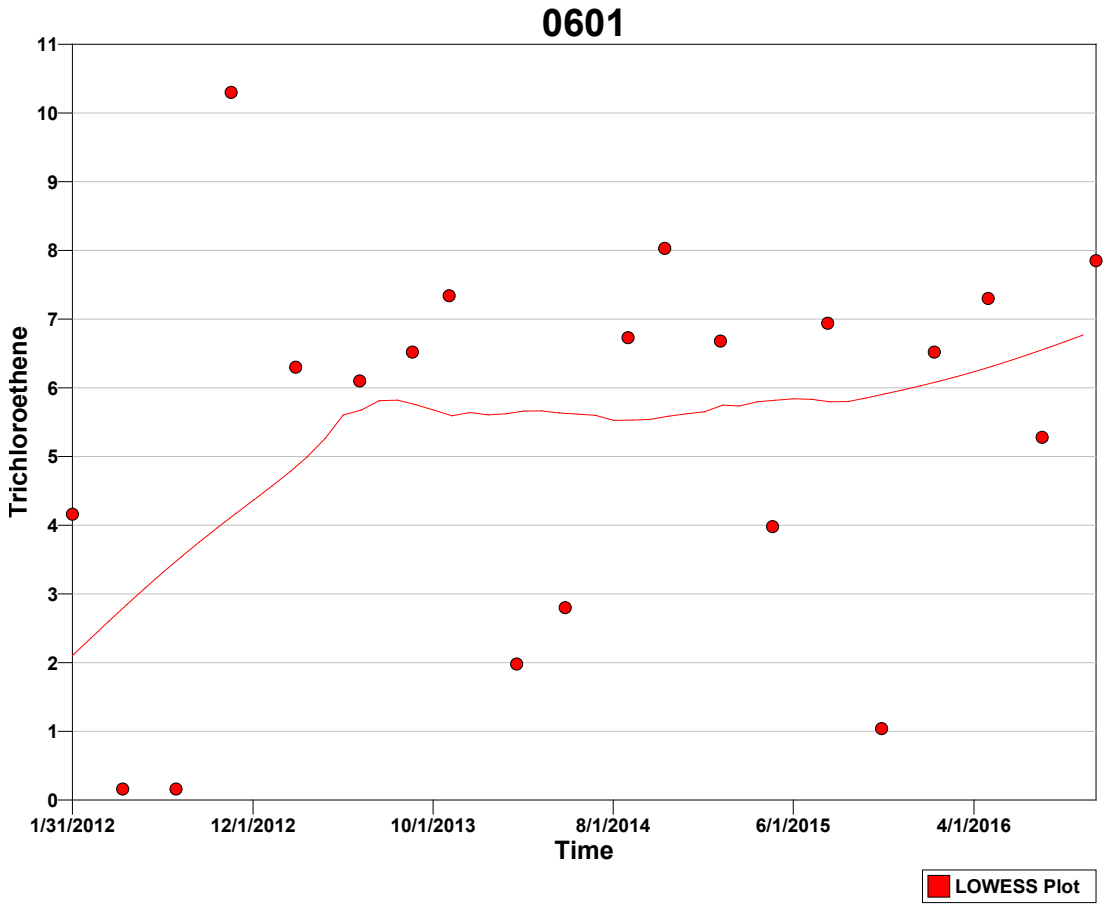
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots
The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.





Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=20) is less than 50.

Normal Distribution Test for Trichloroethene	
Shapiro-Wilk Test Statistic	0.91865
Shapiro-Wilk 5% Critical Value	0.905

The calculated test statistic exceeds the 5% Shapiro-Wilk critical value, so we cannot reject the hypothesis that the values are normally distributed, or in other words the values appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	42	
MK Test Statistic	Z_{MK}	1.33162	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	**Accept**
Alternative Hypothesis	H_a	A downward monotonic trend exists	Reject
Conclude with 95% confidence that a downward monotonic trend does not exist.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 4/10/2017 7:11:12 AM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$x_2-x_1, x_3-x_1, \ldots, x_n-x_1, x_3-x_2, x_4-x_2, \ldots, x_n-x_{n-2}, x_n-x_{n-1}$

3. Let $\text{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\text{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\text{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\text{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \text{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \frac{S-1}{\sqrt{VAR(S)}} \quad \text{if} \quad S > 0$$

$$= 0 \quad \text{if} \quad S = 0 \qquad (3)$$

$$= \frac{S+1}{\sqrt{VAR(S)}} \quad \text{if} \quad S < 0$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Tritium
Analysis was performed on the data that was entered for the location labeled "0601. The following table lists the data points that were used in the analysis.

0601							
Time	**Tritium**		**Time**	**Tritium**		**Time**	**Tritium**
2/17/2005	361200		5/19/2008	68100		10/25/2011	46200
5/17/2005	450460		8/4/2008	133000		1/31/2012	26800
5/26/2005	525340		11/17/2008	82900		4/25/2012	49200
7/28/2005	789040		2/9/2009	46300		7/24/2012	53200
11/3/2005	463100		5/5/2009	67100		2/11/2013	26300
1/30/2006	178480		7/29/2009	104000		8/27/2013	46900
5/18/2006	99900		10/26/2009	64500		2/19/2014	11000
9/26/2006	204000		1/25/2010	27400		8/26/2014	36900
12/5/2006	123000		5/5/2010	89.3		1/29/2015	22300
2/27/2007	117000		7/26/2010	68100		7/29/2015	29000
5/17/2007	146000		10/25/2010	96100		1/25/2016	23500
8/22/2007	161000		1/25/2011	38300		7/25/2016	29200
11/29/2007	88000		5/10/2011	32700			
2/28/2008	61200		7/27/2011	54000			

SUMMARY STATISTICS for Tritium								
n				40				
Min				89.3				
Max				789040				
Range				7.8895e+005				
Mean				1.2627e+005				
Median				65800				
Variance				2.7327e+010				
StdDev				1.6531e+005				
Std Error				26138				
Skewness				2.5321				
Interquartile Range				96750				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
89.3	1.156e+004	2.378e+004	3.375e+004	6.58e+004	1.305e+005	4.415e+005	5.222e+005	7.89e+005

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. These values with low weights can be considered possible outliers and are listed in the table below.

SUSPECTED OUTLIERS	
Date/Time	**Value**
2/17/2005	361200
7/28/2005	789040
1/30/2006	178480
5/18/2006	99900
12/5/2006	123000

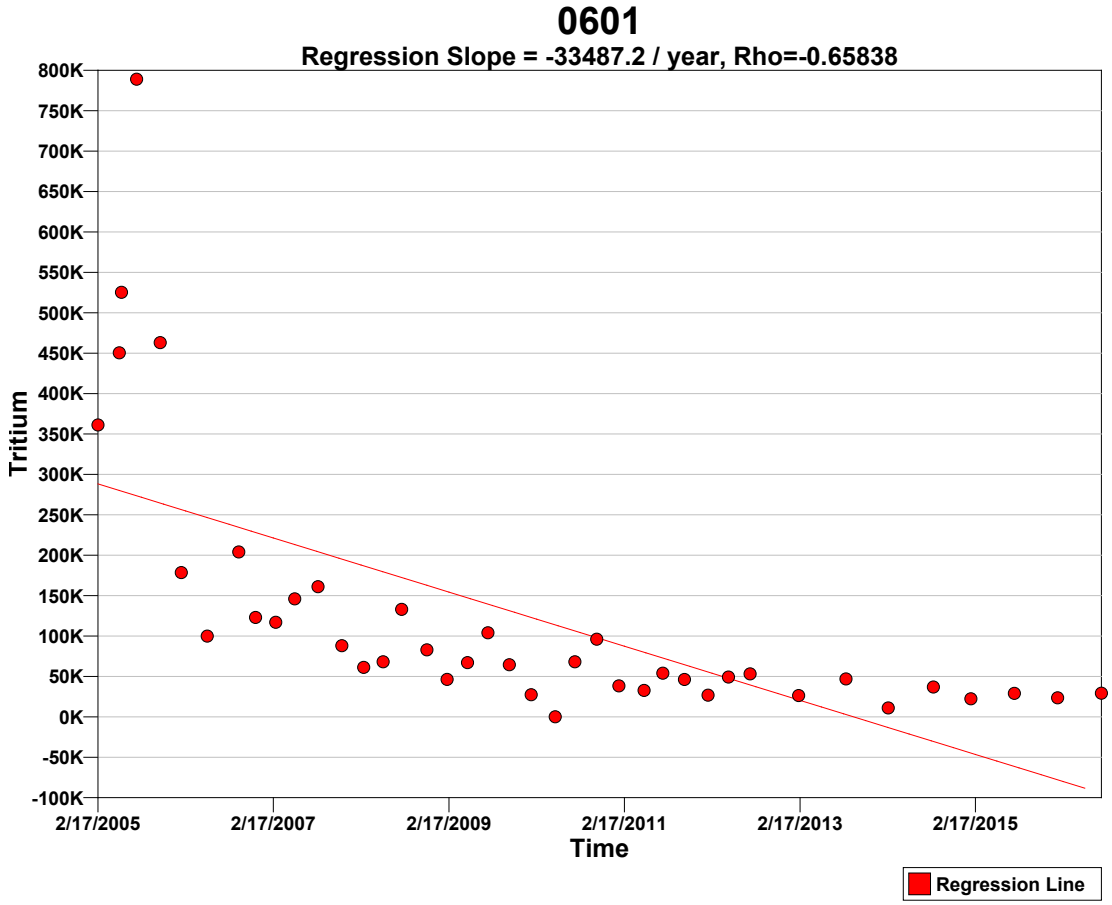
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

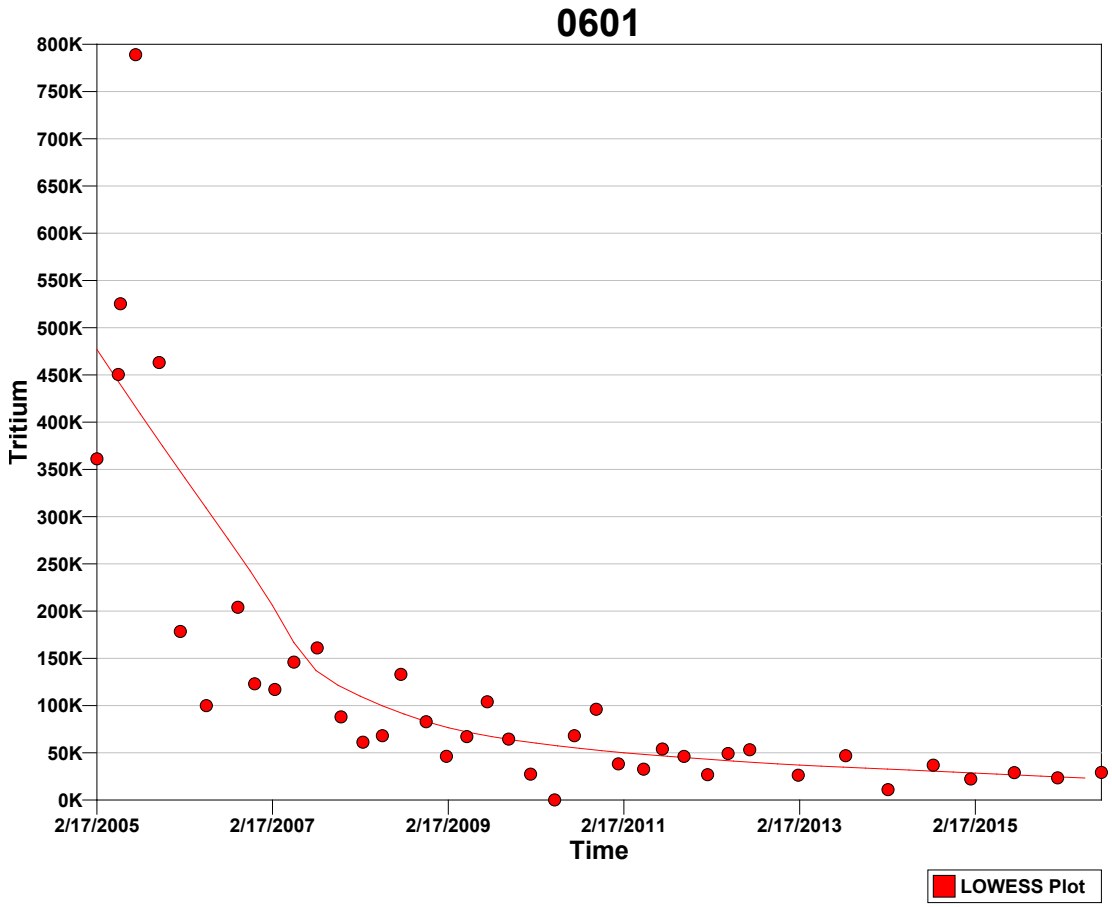
It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots

The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.





Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=40) is less than 50.

Normal Distribution Test for Tritium	
Shapiro-Wilk Test Statistic	0.6561
Shapiro-Wilk 5% Critical Value	0.94

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-541	
MK Test Statistic	Z_{MK}	-6.29199	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:51:48 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, …, n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$x_2-x_1, x_3-x_1, \ldots, x_n-x_1, x_3-x_2, x_4-x_2, \ldots, x_n-x_{n-2}, x_n-x_{n-1}$

3. Let $\text{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\text{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\text{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\text{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \text{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \frac{S-1}{\sqrt{VAR(S)}} \quad \text{if} \quad S > 0$$

$$= 0 \quad \text{if} \quad S = 0 \qquad (3)$$

$$= \frac{S+1}{\sqrt{VAR(S)}} \quad \text{if} \quad S < 0$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for cis-1x2-Dichloroethene
Analysis was performed on the data that was entered for the location labeled "0602. The following table lists the data points that were used in the analysis.

0602							
Time	**cis-1x2-Dichloroethene**		**Time**	**cis-1x2-Dichloroethene**		**Time**	**cis-1x2-Dichloroethene**
5/17/2005	2.7		12/29/2008	17.9		2/11/2013	42.3
7/28/2005	8.3		2/9/2009	19.1		10/28/2013	12.2
1/30/2006	21		5/5/2009	17.8		2/19/2014	19.4
5/18/2006	19		10/26/2009	18.1		5/12/2014	15.2
9/27/2006	16		1/25/2010	34.7		1/29/2015	9.89
12/5/2006	17.8		5/5/2010	12.2		4/27/2015	7.5
2/27/2007	23.9		1/25/2011	38		10/28/2015	4.56
5/17/2007	19.7		4/25/2011	8.11		1/25/2016	7.35
11/29/2007	18.2		5/10/2011	30.5		4/25/2016	4.54
2/28/2008	30.6		10/25/2011	22.6		10/24/2016	14.8
5/19/2008	26.3		1/31/2012	14.9			

SUMMARY STATISTICS for cis-1x2-Dichloroethene								
n				32				
Min				2.7				
Max				42.3				
Range				39.6				
Mean				17.973				
Median				17.85				
Variance				93.601				
StdDev				9.6748				
Std Error				1.7103				
Skewness				0.69887				
Interquartile Range				11.733				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
2.7	3.896	5.397	10.47	17.85	22.2	33.47	39.51	42.3

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. These values with low weights can be considered possible outliers and are listed in the table below.

SUSPECTED OUTLIERS	
Date/Time	**Value**
2/11/2013	42.3

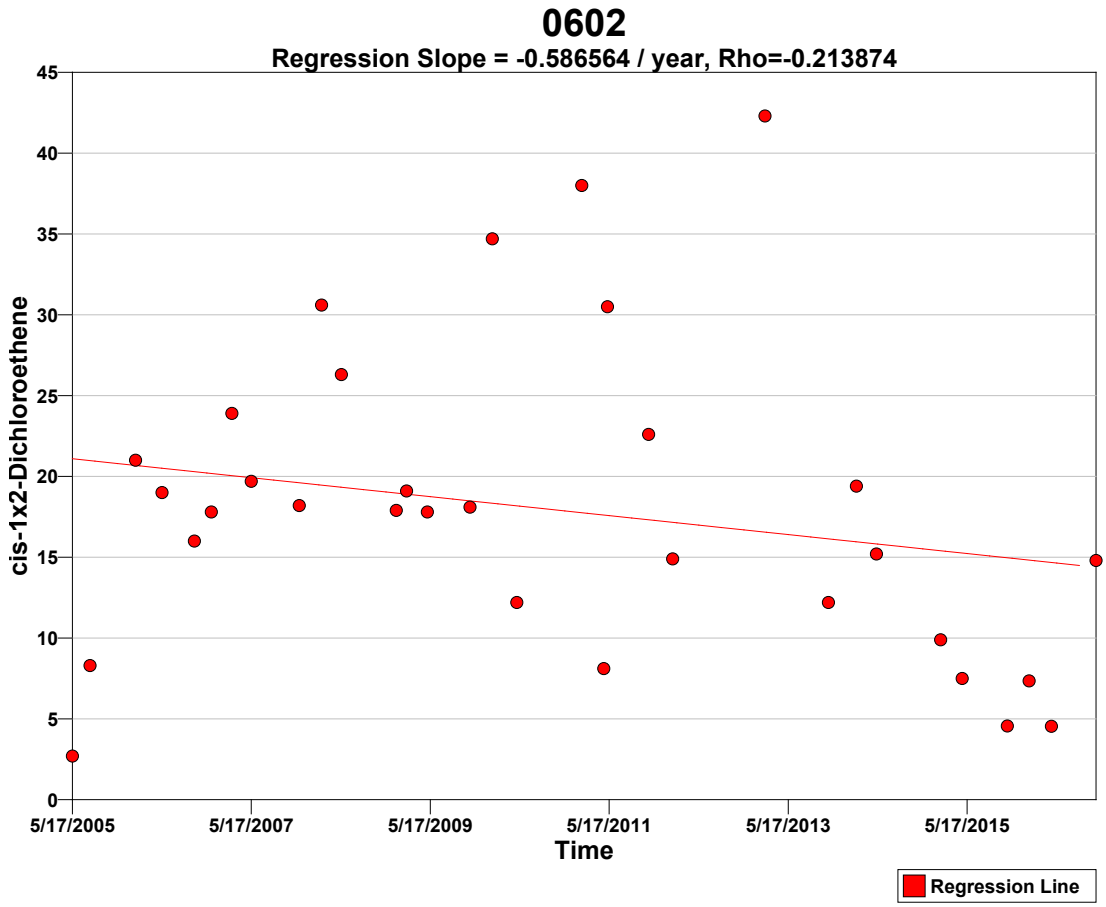
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

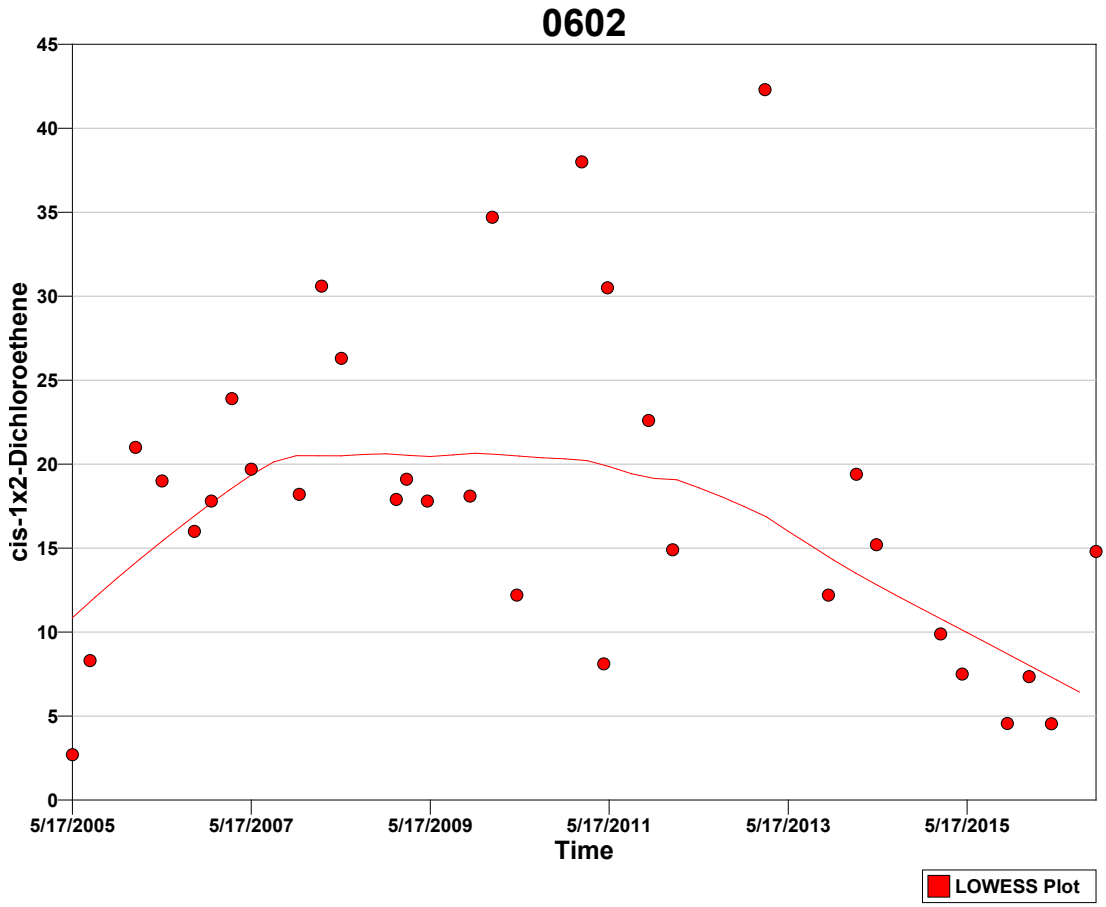
It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots

The Time vs. Measured Values Plot shows the best fitting “least squares” linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated “least squares” slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.





Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=32) is less than 50.

Normal Distribution Test for cis-1x2-Dichloroethene	
Shapiro-Wilk Test Statistic	0.94696
Shapiro-Wilk 5% Critical Value	0.93

The calculated test statistic exceeds the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-104	
MK Test Statistic	Z_{MK}	-1.67073	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:39:12 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$x_2-x_1, x_3-x_1, \ldots, x_n-x_1, x_3-x_2, x_4-x_2, \ldots, x_n-x_{n-2}, x_n-x_{n-1}$

3. Let $\text{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\text{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\text{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\text{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \text{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \frac{S-1}{\sqrt{VAR(S)}} \quad \text{if} \quad S > 0$$

$$= 0 \quad \text{if} \quad S = 0 \qquad (3)$$

$$= \frac{S+1}{\sqrt{VAR(S)}} \quad \text{if} \quad S < 0$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Trichloroethene
Analysis was performed on the data that was entered for the location labeled "0602. The following table lists the data points that were used in the analysis.

0602							
Time	**Trichloroethene**		**Time**	**Trichloroethene**		**Time**	**Trichloroethene**
5/17/2005	3.7		12/29/2008	53.4		2/11/2013	32.6
7/28/2005	6.9		2/9/2009	56.1		10/28/2013	4.58
1/30/2006	19		5/5/2009	49.6		2/19/2014	23.6
5/18/2006	19		10/26/2009	18.9		5/12/2014	23.5
9/27/2006	15		1/25/2010	54.5		1/29/2015	19
12/5/2006	25.6		5/5/2010	23.3		4/27/2015	21.4
2/27/2007	26.3		1/25/2011	139		10/28/2015	14.7
5/17/2007	27.6		4/25/2011	49.9		1/25/2016	12.8
11/29/2007	26.8		5/10/2011	70.1		4/25/2016	6.66
2/28/2008	63.3		10/25/2011	16.9		10/24/2016	10
5/19/2008	58.9		1/31/2012	22.2			

SUMMARY STATISTICS for Trichloroethene								
n				32				
Min				3.7				
Max				139				
Range				135.3				
Mean				31.714				
Median				23.4				
Variance				725.68				
StdDev				26.938				
Std Error				4.7621				
Skewness				2.246				
Interquartile Range				34.35				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
3.7	4.272	6.732	15.48	23.4	49.83	61.98	94.22	139

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. These values with low weights can be considered possible outliers and are listed in the table below.

SUSPECTED OUTLIERS	
Date/Time	**Value**
2/28/2008	63.3
10/26/2009	18.9
5/5/2010	23.3
1/25/2011	139
5/10/2011	70.1
10/25/2011	16.9
10/28/2013	4.58

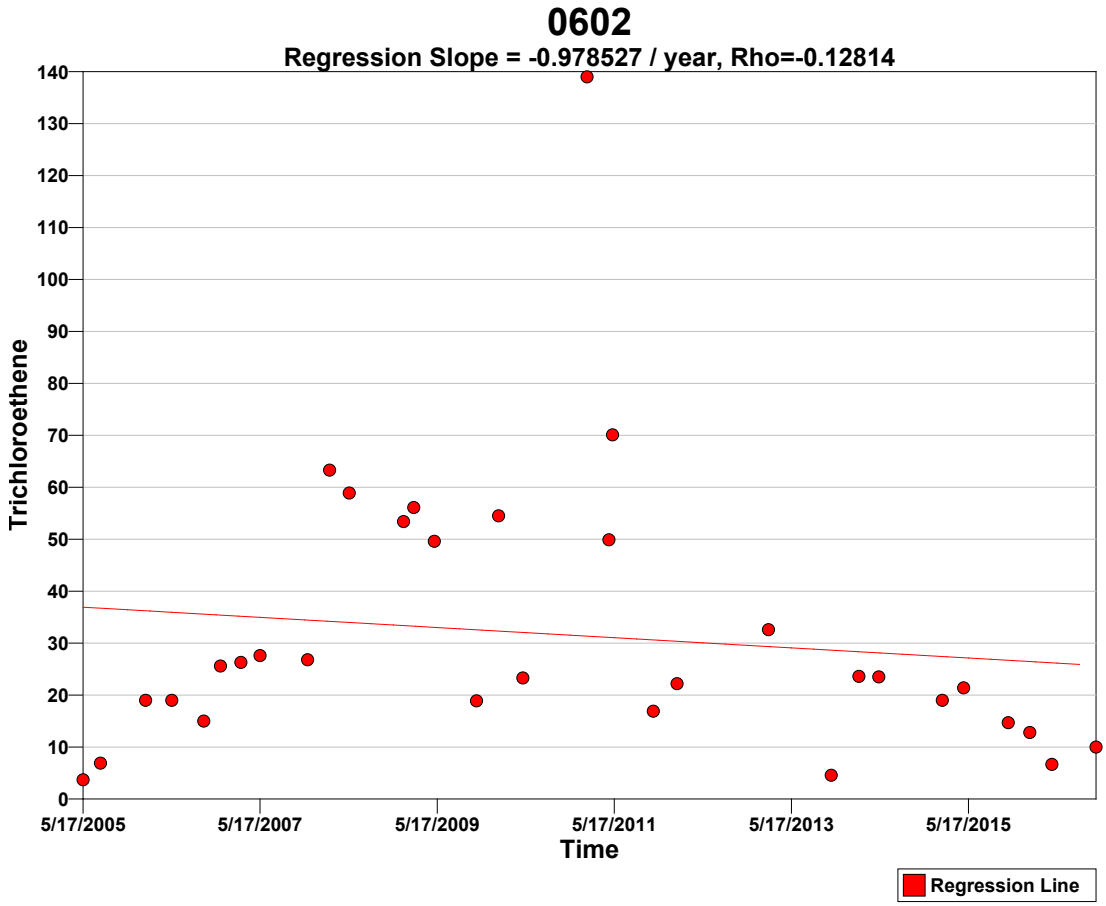
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

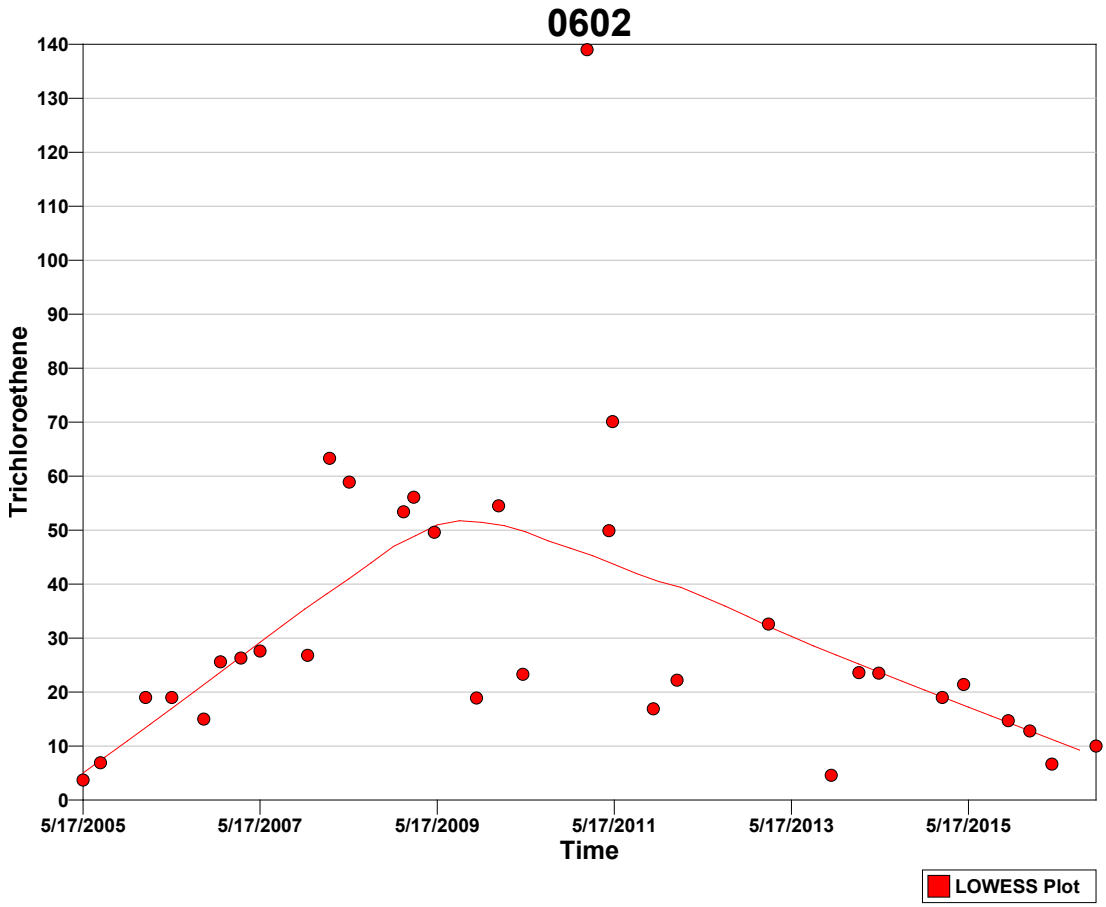
It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots

The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.





Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=32) is less than 50.

Normal Distribution Test for Trichloroethene	
Shapiro-Wilk Test Statistic	0.78432
Shapiro-Wilk 5% Critical Value	0.93

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-59	
MK Test Statistic	Z_{MK}	-0.941008	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	**Accept**
Alternative Hypothesis	H_a	A downward monotonic trend exists	Reject
Conclude with 95% confidence that a downward monotonic trend does not exist.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:33:42 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$x_2-x_1, x_3-x_1, \ldots, x_n-x_1, x_3-x_2, x_4-x_2, \ldots, x_n-x_{n-2}, x_n-x_{n-1}$

3. Let $\text{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\text{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\text{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\text{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \text{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \frac{S-1}{\sqrt{VAR(S)}} \quad \text{if} \quad S > 0$$

$$= 0 \quad \text{if} \quad S = 0 \qquad (3)$$

$$= \frac{S+1}{\sqrt{VAR(S)}} \quad \text{if} \quad S < 0$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Trichloroethene
Analysis was performed on the data that was entered for the location labeled "0602. The following table lists the data points that were used in the analysis.

0602							
Time	**Trichloroethene**		**Time**	**Trichloroethene**		**Time**	**Trichloroethene**
1/31/2012	22.2		5/12/2014	23.5		1/25/2016	12.8
2/11/2013	32.6		1/29/2015	19		4/25/2016	6.66
10/28/2013	4.58		4/27/2015	21.4		10/24/2016	10
2/19/2014	23.6		10/28/2015	14.7			

SUMMARY STATISTICS for Trichloroethene								
n	11							
Min	4.58							
Max	32.6							
Range	28.02							
Mean	17.367							
Median	19							
Variance	71.119							
StdDev	8.4332							
Std Error	2.5427							
Skewness	0.072865							
Interquartile Range	13.5							
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
4.58	4.58	4.996	10	19	23.5	30.8	32.6	32.6

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. These values with low weights can be considered possible outliers and are listed in the table below.

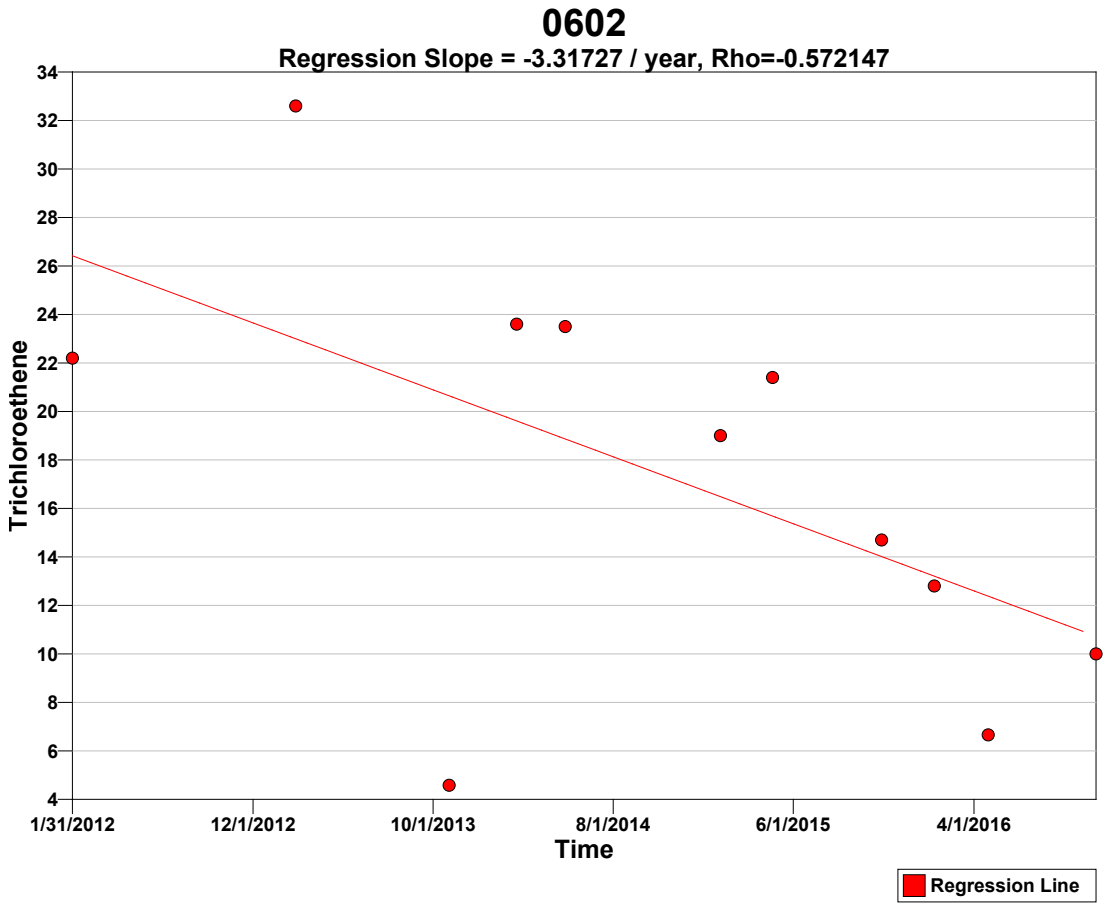
SUSPECTED OUTLIERS	
Date/Time	**Value**
10/28/2013	4.58

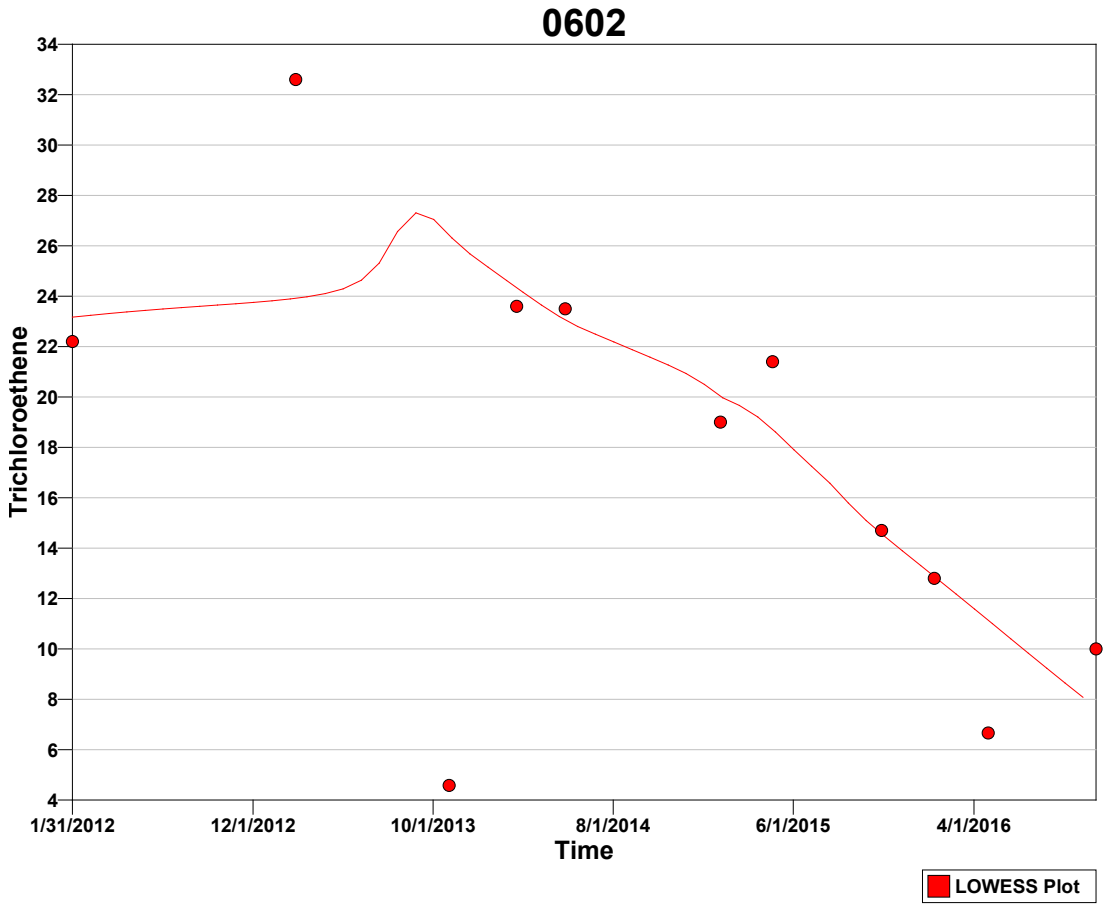
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots
The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.





Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=11) is less than 50.

Normal Distribution Test for Trichloroethene	
Shapiro-Wilk Test Statistic	0.96296
Shapiro-Wilk 5% Critical Value	0.85

The calculated test statistic exceeds the 5% Shapiro-Wilk critical value, so we cannot reject the hypothesis that the values are normally distributed, or in other words the values appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-29	
MK Test Statistic	Z_{MK}	-2.1798	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 4/10/2017 7:12:50 AM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$x_2-x_1, x_3-x_1, \ldots, x_n-x_1, x_3-x_2, x_4-x_2, \ldots, x_n-x_{n-2}, x_n-x_{n-1}$

3. Let $\text{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\text{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\text{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\text{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \text{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \frac{S-1}{\sqrt{VAR(S)}} \quad \text{if} \quad S > 0$$

$$= 0 \quad \text{if} \quad S = 0 \qquad (3)$$

$$= \frac{S+1}{\sqrt{VAR(S)}} \quad \text{if} \quad S < 0$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Tritium
Analysis was performed on the data that was entered for the location labeled "0602. The following table lists the data points that were used in the analysis.

0602							
Time	**Tritium**		**Time**	**Tritium**		**Time**	**Tritium**
4/20/2005	41770		11/29/2007	45100		1/25/2011	10100
5/17/2005	22740		2/28/2008	27200		5/10/2011	7240
7/28/2005	21090		5/19/2008	25100		10/25/2011	14500
1/30/2006	13160		12/29/2008	15400		1/31/2012	8120
5/18/2006	17380		2/9/2009	14200		2/11/2013	8810
9/27/2006	82700		5/5/2009	14800		2/19/2014	4570
12/5/2006	48900		10/26/2009	22500		1/29/2015	7280
2/27/2007	30500		1/25/2010	11900		1/25/2016	5740
5/17/2007	36800		5/5/2010	14700			

SUMMARY STATISTICS for Tritium								
n				26				
Min				4570				
Max				82700				
Range				78130				
Mean				22012				
Median				15100				
Variance				3.0561e+008				
StdDev				17482				
Std Error				3428.5				
Skewness				1.9507				
Interquartile Range				18248				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
4570	4980	6790	9778	1.51e+004	2.803e+004	4.624e+004	7.087e+004	8.27e+004

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. These values with low weights can be considered possible outliers and are listed in the table below.

SUSPECTED OUTLIERS	
Date/Time	**Value**
9/27/2006	82700

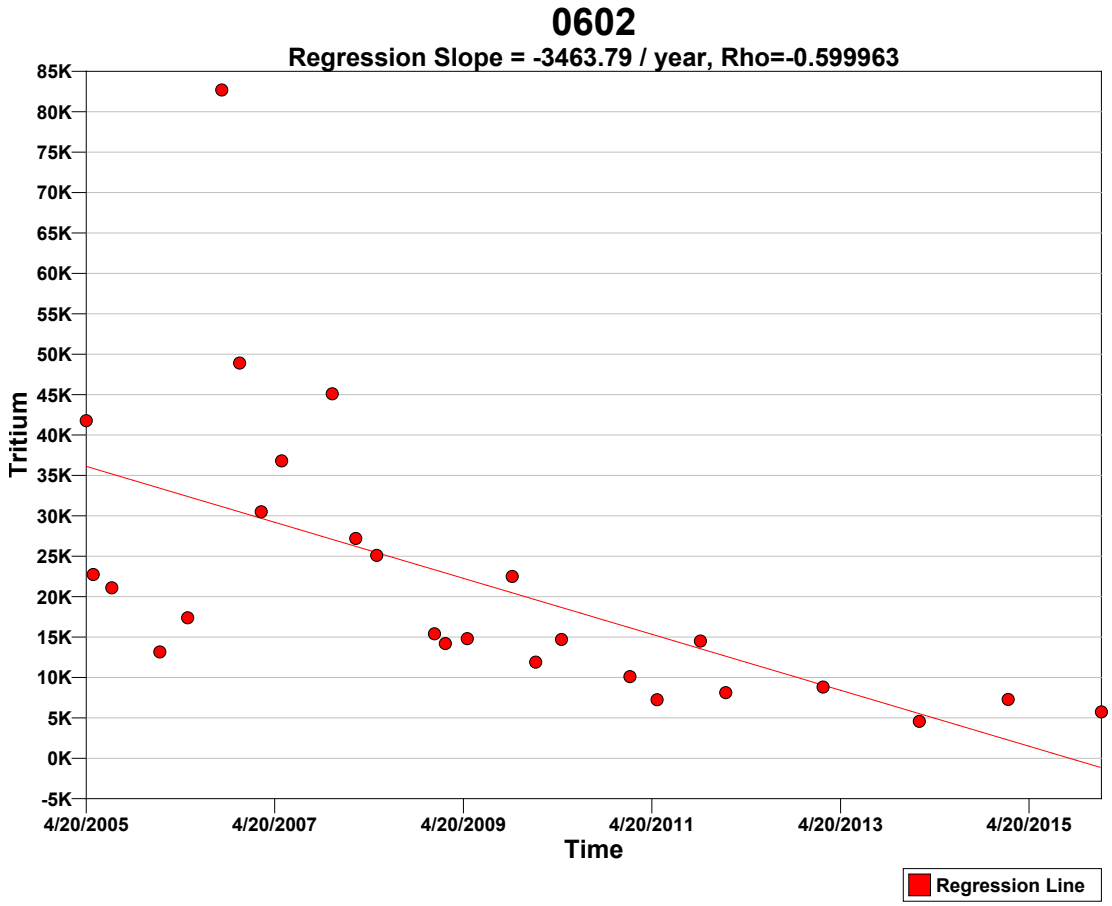
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

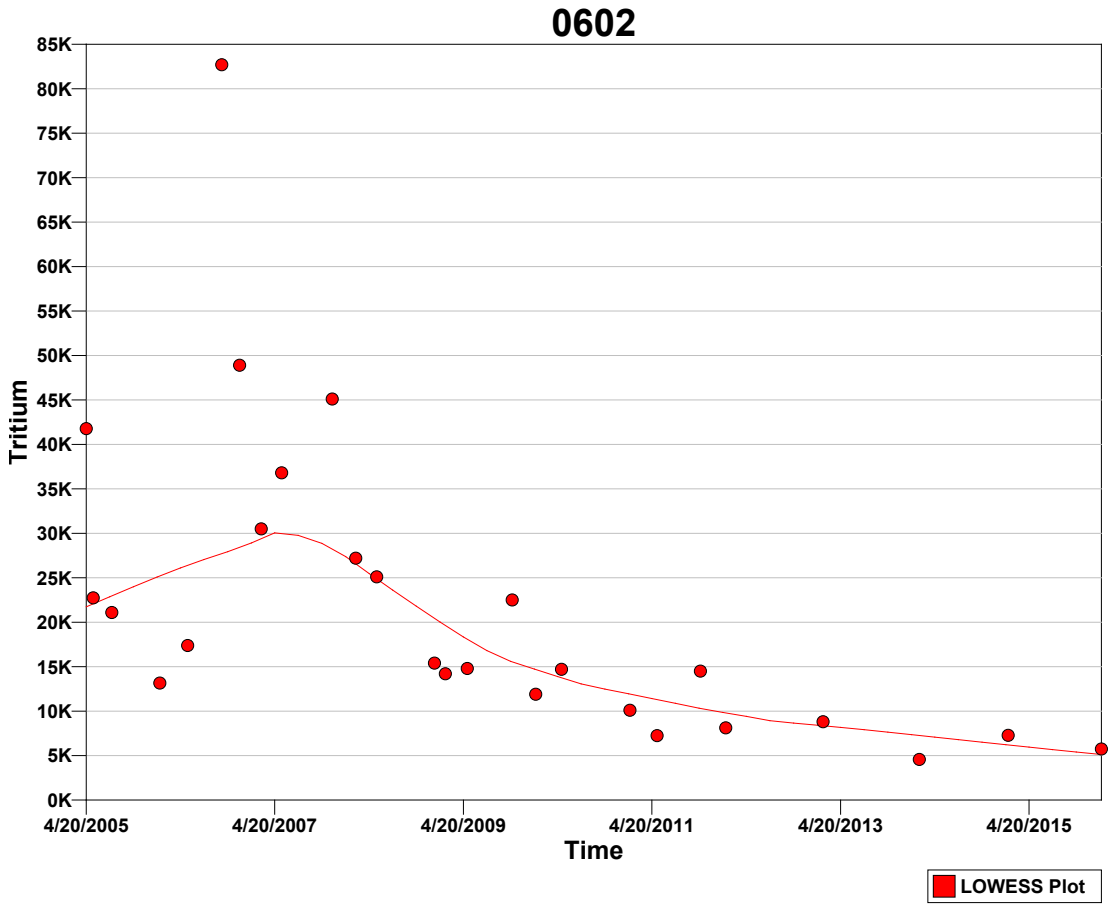
It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots
The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the

value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated “least squares” slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.





Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=26) is less than 50.

Normal Distribution Test for Tritium	
Shapiro-Wilk Test Statistic	0.81086
Shapiro-Wilk 5% Critical Value	0.92

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-207	
MK Test Statistic	Z_{MK}	-4.54056	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:52:37 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$x_2-x_1, x_3-x_1, \ldots, x_n-x_1, x_3-x_2, x_4-x_2, \ldots, x_n-x_{n-2}, x_n-x_{n-1}$

3. Let $\text{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\text{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\text{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\text{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \text{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \frac{S-1}{\sqrt{VAR(S)}} \quad \text{if} \quad S > 0$$

$$= 0 \quad \text{if} \quad S = 0 \qquad (3)$$

$$= \frac{S+1}{\sqrt{VAR(S)}} \quad \text{if} \quad S < 0$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for cis-1x2-Dichloroethene
Analysis was performed on the data that was entered for the location labeled "0605. The following table lists the data points that were used in the analysis.

0605							
Time	**cis-1x2-Dichloroethene**		**Time**	**cis-1x2-Dichloroethene**		**Time**	**cis-1x2-Dichloroethene**
2/17/2005	26		2/9/2009	7.92		5/30/2013	1.98
5/17/2005	30		5/5/2009	8.8		8/27/2013	1.49
7/28/2005	50		7/29/2009	6.83		10/28/2013	1.34
11/3/2005	17		10/26/2009	4.35		2/19/2014	2.9
1/30/2006	21		1/25/2010	4.52		5/12/2014	3.1
5/18/2006	16		5/5/2010	4.49		8/26/2014	9.93
9/26/2006	1.9		7/26/2010	1.91		10/27/2014	1.56
12/5/2006	9.91		12/7/2010	6.25		1/29/2015	3.56
2/27/2007	20.4		1/25/2011	4.48		4/27/2015	1.86
5/17/2007	32.2		5/10/2011	1.62		7/29/2015	1.09
8/22/2007	31.7		7/27/2011	7.08		10/28/2015	2.19
11/29/2007	9.54		10/25/2011	3.22		1/25/2016	1.18
2/28/2008	7.58		1/31/2012	2.39		4/25/2016	2.03
5/19/2008	5.15		4/25/2012	1.41		7/25/2016	1.02
8/4/2008	2.53		10/25/2012	2.47		10/24/2016	5.92
11/17/2008	12.4		2/11/2013	3.02			

SUMMARY STATISTICS for cis-1x2-Dichloroethene								
n				47				
Min				1.02				
Max				50				
Range				48.98				
Mean				8.6217				
Median				4.48				
Variance				109.76				
StdDev				10.477				
Std Error				1.5282				
Skewness				2.154				
Interquartile Range				7.93				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
1.02	1.126	1.396	1.98	4.48	9.91	26.8	32	50

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. These values with low weights can be considered possible outliers and are listed in the table below.

SUSPECTED OUTLIERS	
Date/Time	**Value**
7/28/2005	50
9/26/2006	1.9
12/5/2006	9.91
5/17/2007	32.2
8/22/2007	31.7
8/26/2014	9.93

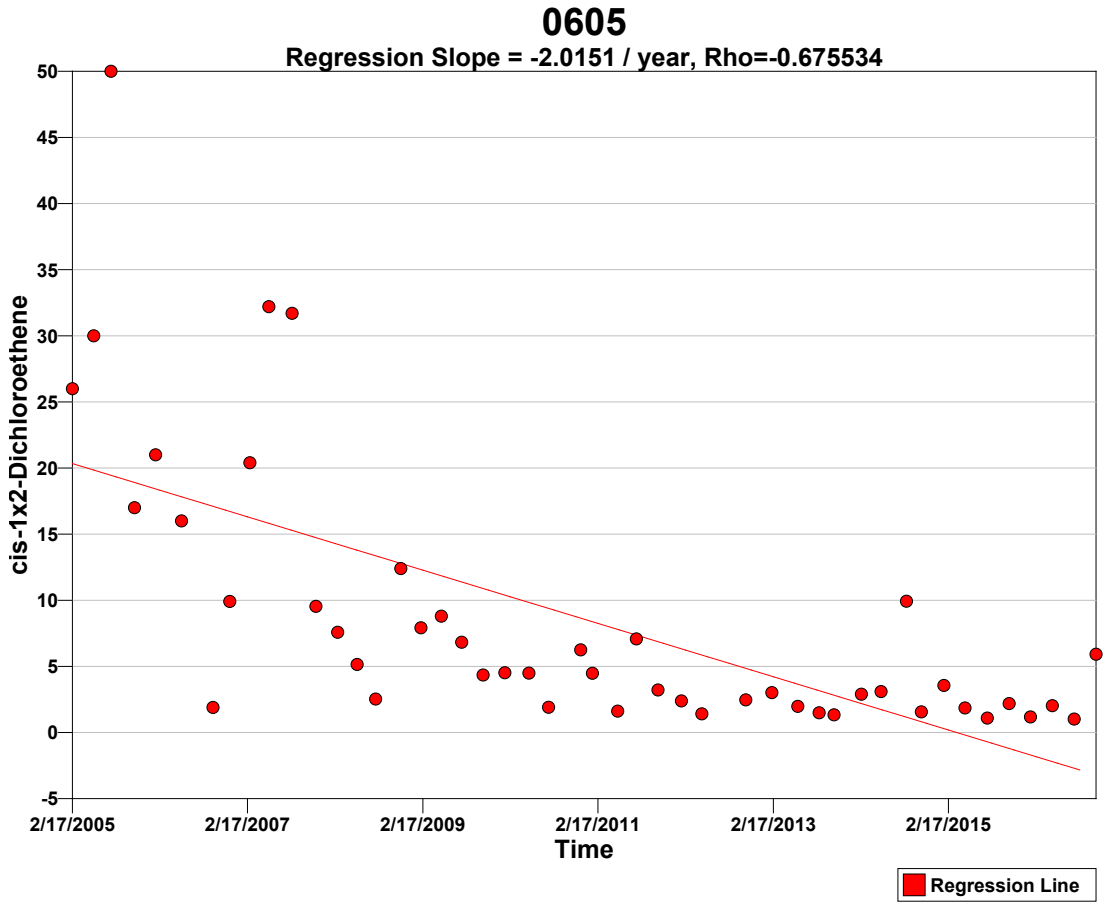
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

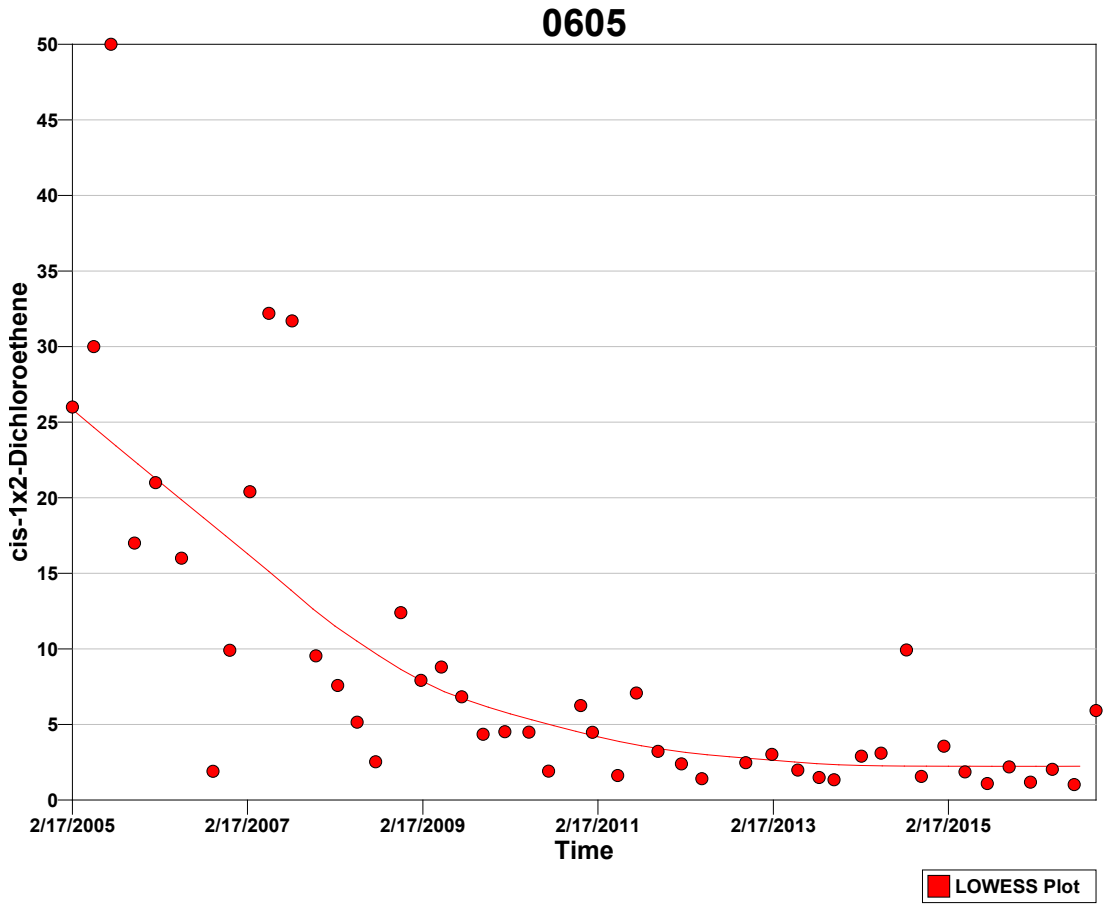
It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots

The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.





Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=47) is less than 50.

Normal Distribution Test for cis-1x2-Dichloroethene	
Shapiro-Wilk Test Statistic	0.71332
Shapiro-Wilk 5% Critical Value	0.946

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-635	
MK Test Statistic	Z_{MK}	-5.81407	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:40:30 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$x_2-x_1, x_3-x_1, \ldots, x_n-x_1, x_3-x_2, x_4-x_2, \ldots, x_n-x_{n-2}, x_n-x_{n-1}$

3. Let $\text{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\text{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\text{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\text{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \text{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \frac{S-1}{\sqrt{VAR(S)}} \quad \text{if} \quad S > 0$$

$$= 0 \quad \text{if} \quad S = 0 \qquad (3)$$

$$= \frac{S+1}{\sqrt{VAR(S)}} \quad \text{if} \quad S < 0$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Trichloroethene
Analysis was performed on the data that was entered for the location labeled "0605. The following table lists the data points that were used in the analysis.

0605							
Time	**Trichloroethene**		**Time**	**Trichloroethene**		**Time**	**Trichloroethene**
2/17/2005	28		2/9/2009	10.5		5/30/2013	14.7
5/17/2005	29		5/5/2009	18.3		8/27/2013	11
7/28/2005	97		7/29/2009	14		10/28/2013	13.5
11/3/2005	33		10/26/2009	9.54		2/19/2014	8.69
1/30/2006	27		1/25/2010	9.47		5/12/2014	9.43
5/18/2006	16		5/5/2010	17.6		8/26/2014	11.8
9/26/2006	8.8		7/26/2010	14.4		10/27/2014	14.8
12/5/2006	9.61		12/7/2010	10		1/29/2015	13
2/27/2007	15.6		1/25/2011	12.2		4/27/2015	10.9
5/17/2007	24.7		5/10/2011	13.1		7/29/2015	12.3
8/22/2007	12.7		7/27/2011	15.9		10/28/2015	7.23
11/29/2007	10.4		10/25/2011	11.5		1/25/2016	10.2
2/28/2008	15.3		1/31/2012	9.16		4/25/2016	14.2
5/19/2008	13.8		4/25/2012	15.4		7/25/2016	7.28
8/4/2008	13.4		10/25/2012	19.8		10/24/2016	9.82
11/17/2008	13.9		2/11/2013	13.9			

SUMMARY STATISTICS for Trichloroethene								
n				47				
Min				7.23				
Max				97				
Range				89.77				
Mean				15.996				
Median				13.4				
Variance				178.49				
StdDev				13.36				
Std Error				1.9488				
Skewness				5.115				
Interquartile Range				5.4				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
7.23	7.844	9.088	10.2	13.4	15.6	27.2	31.4	97

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. These values with low weights can be considered possible outliers and are listed in the table below.

SUSPECTED OUTLIERS	
Date/Time	**Value**
7/28/2005	97

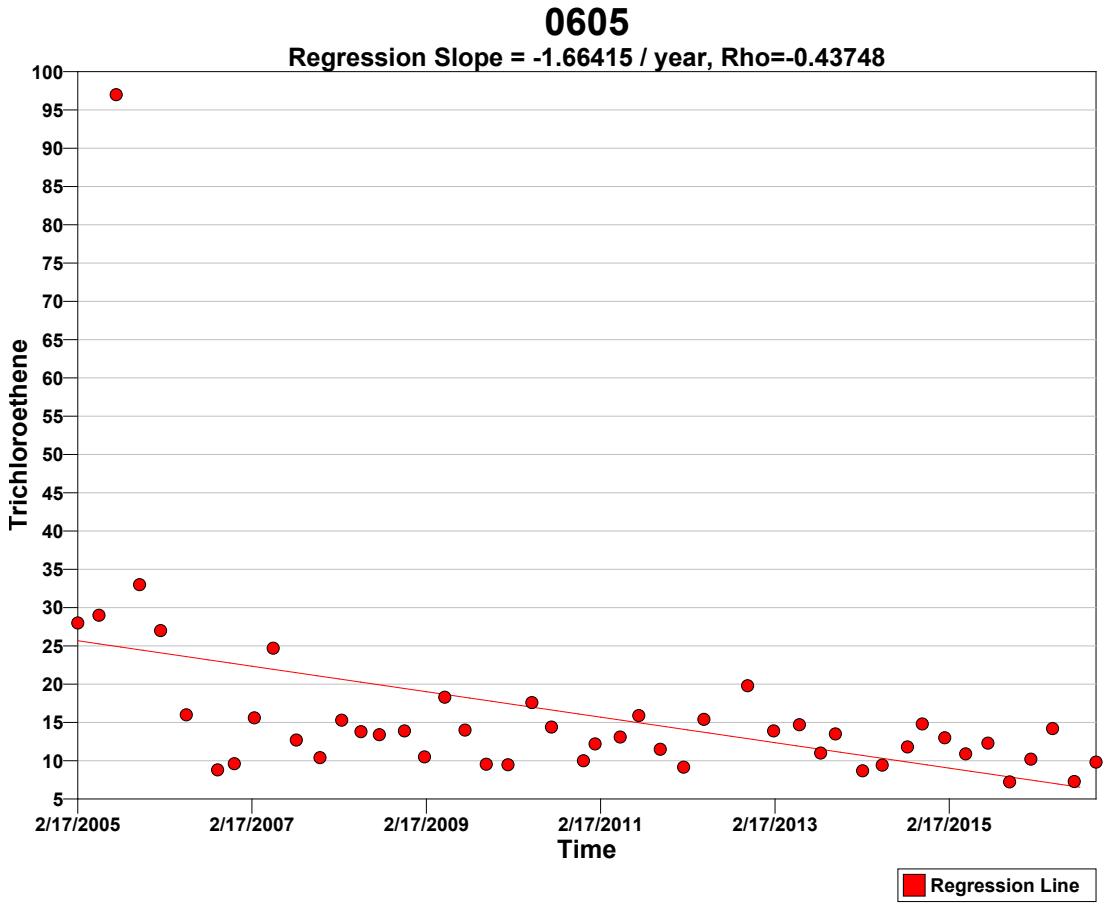
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

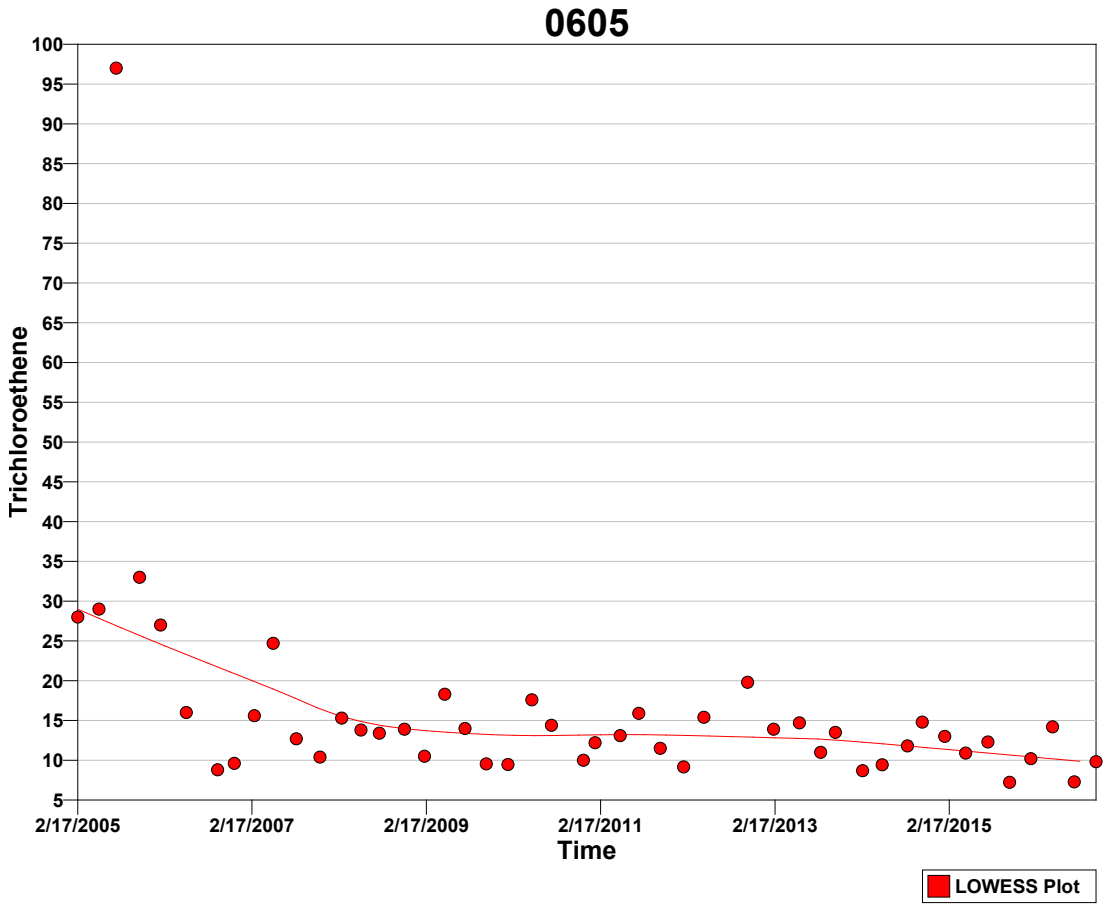
It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots

The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.





Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=47) is less than 50.

Normal Distribution Test for Trichloroethene	
Shapiro-Wilk Test Statistic	0.48314
Shapiro-Wilk 5% Critical Value	0.946

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-382	
MK Test Statistic	Z_{MK}	-3.49409	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:34:28 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$x_2-x_1, x_3-x_1, \ldots, x_n-x_1, x_3-x_2, x_4-x_2, \ldots, x_n-x_{n-2}, x_n-x_{n-1}$

3. Let $\text{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\text{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\text{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\text{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \text{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \frac{S-1}{\sqrt{VAR(S)}} \quad \text{if} \quad S > 0$$

$$= 0 \quad \text{if} \quad S = 0 \qquad (3)$$

$$= \frac{S+1}{\sqrt{VAR(S)}} \quad \text{if} \quad S < 0$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Trichloroethene
Analysis was performed on the data that was entered for the location labeled "0605. The following table lists the data points that were used in the analysis.

0605							
Time	**Trichloroethene**		**Time**	**Trichloroethene**		**Time**	**Trichloroethene**
1/31/2012	9.16		2/19/2014	8.69		10/28/2015	7.23
4/25/2012	15.4		5/12/2014	9.43		1/25/2016	10.2
10/25/2012	19.8		8/26/2014	11.8		4/25/2016	14.2
2/11/2013	13.9		10/27/2014	14.8		7/25/2016	7.28
5/30/2013	14.7		1/29/2015	13		10/24/2016	9.82
8/27/2013	11		4/27/2015	10.9			
10/28/2013	13.5		7/29/2015	12.3			

SUMMARY STATISTICS for Trichloroethene								
n				19				
Min				7.23				
Max				19.8				
Range				12.57				
Mean				11.953				
Median				11.8				
Variance				10.01				
StdDev				3.1638				
Std Error				0.72584				
Skewness				0.57712				
Interquartile Range				4.77				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
7.23	7.23	7.28	9.43	11.8	14.2	15.4	19.8	19.8

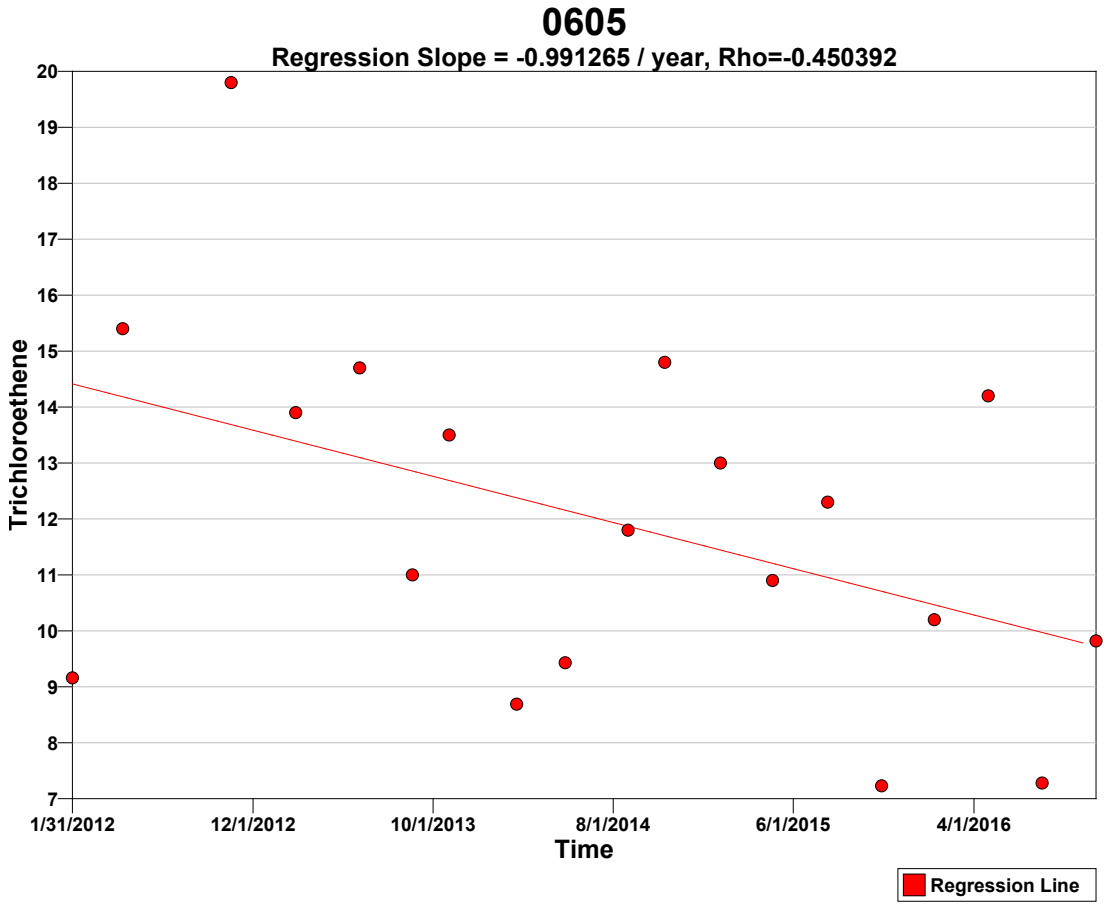
Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. The LOWESS method did not detect any outliers for the current data set.

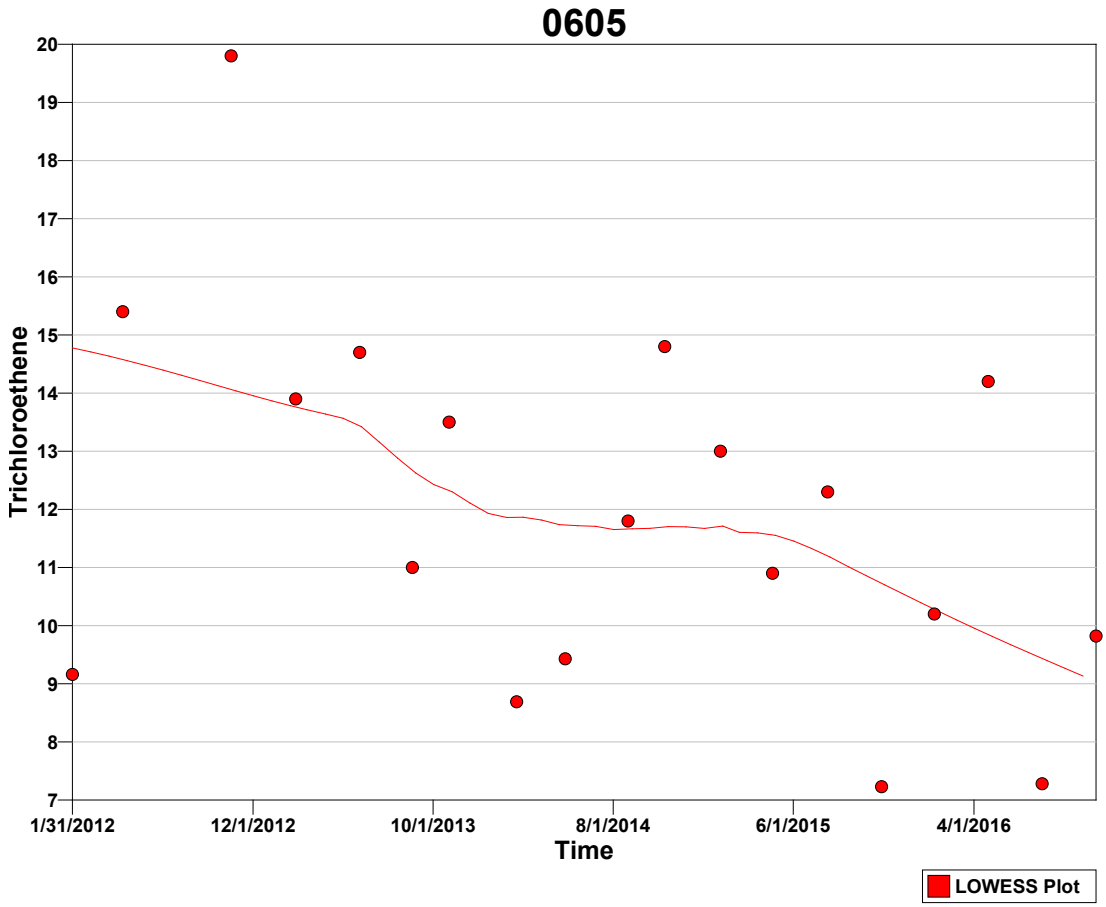
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots
The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.





Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=19) is less than 50.

Normal Distribution Test for Trichloroethene	
Shapiro-Wilk Test Statistic	0.95977
Shapiro-Wilk 5% Critical Value	0.901

The calculated test statistic exceeds the 5% Shapiro-Wilk critical value, so we cannot reject the hypothesis that the values are normally distributed, or in other words the values appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-53	
MK Test Statistic	Z_{MK}	-1.81925	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 4/10/2017 7:13:44 AM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$x_2-x_1, x_3-x_1, \ldots, x_n-x_1, x_3-x_2, x_4-x_2, \ldots, x_n-x_{n-2}, x_n-x_{n-1}$

3. Let $\text{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\text{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\text{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\text{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \text{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \frac{S-1}{\sqrt{VAR(S)}} \quad \text{if} \quad S > 0$$

$$= 0 \quad \text{if} \quad S = 0 \qquad (3)$$

$$= \frac{S+1}{\sqrt{VAR(S)}} \quad \text{if} \quad S < 0$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Tritium

Analysis was performed on the data that was entered for the location labeled "0605. The following table lists the data points that were used in the analysis.

0605							
Time	**Tritium**		**Time**	**Tritium**		**Time**	**Tritium**
2/17/2005	177210		2/28/2008	26100		5/10/2011	11900
5/17/2005	182750		5/19/2008	21300		7/27/2011	13200
7/28/2005	269300		8/4/2008	18600		10/25/2011	13100
11/3/2005	9390		11/17/2008	16900		1/31/2012	7820
11/22/2005	158300		2/9/2009	14800		4/25/2012	12000
1/30/2006	106520		5/5/2009	18400		2/11/2013	10700
5/18/2006	58880		7/29/2009	27100		8/27/2013	9610
9/26/2006	55800		10/26/2009	19300		2/19/2014	6300
12/5/2006	36200		1/25/2010	12500		8/26/2014	8060
2/27/2007	26500		5/5/2010	16000		1/29/2015	7910
5/17/2007	33400		7/26/2010	18500		7/29/2015	8340
8/22/2007	22200		12/7/2010	16700		1/25/2016	7460
11/29/2007	25600		1/25/2011	14900		7/25/2016	4410

SUMMARY STATISTICS for Tritium								
n				39				
Min				4410				
Max				269300				
Range				2.6489e+005				
Mean				39076				
Median				16900				
Variance				3.4479e+009				
StdDev				58719				
Std Error				9402.6				
Skewness				2.6451				
Interquartile Range				16400				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
4410	6300	7820	1.07e+004	1.69e+004	2.71e+004	1.583e+005	1.828e+005	2.693e+005

Outliers

One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. These values with low weights can be considered possible outliers and are listed in the table below.

SUSPECTED OUTLIERS	
Date/Time	**Value**
5/17/2005	182750
7/28/2005	269300
11/3/2005	9390
11/22/2005	158300
5/18/2006	58880
9/26/2006	55800
12/5/2006	36200
2/27/2007	26500
5/17/2007	33400
8/22/2007	22200

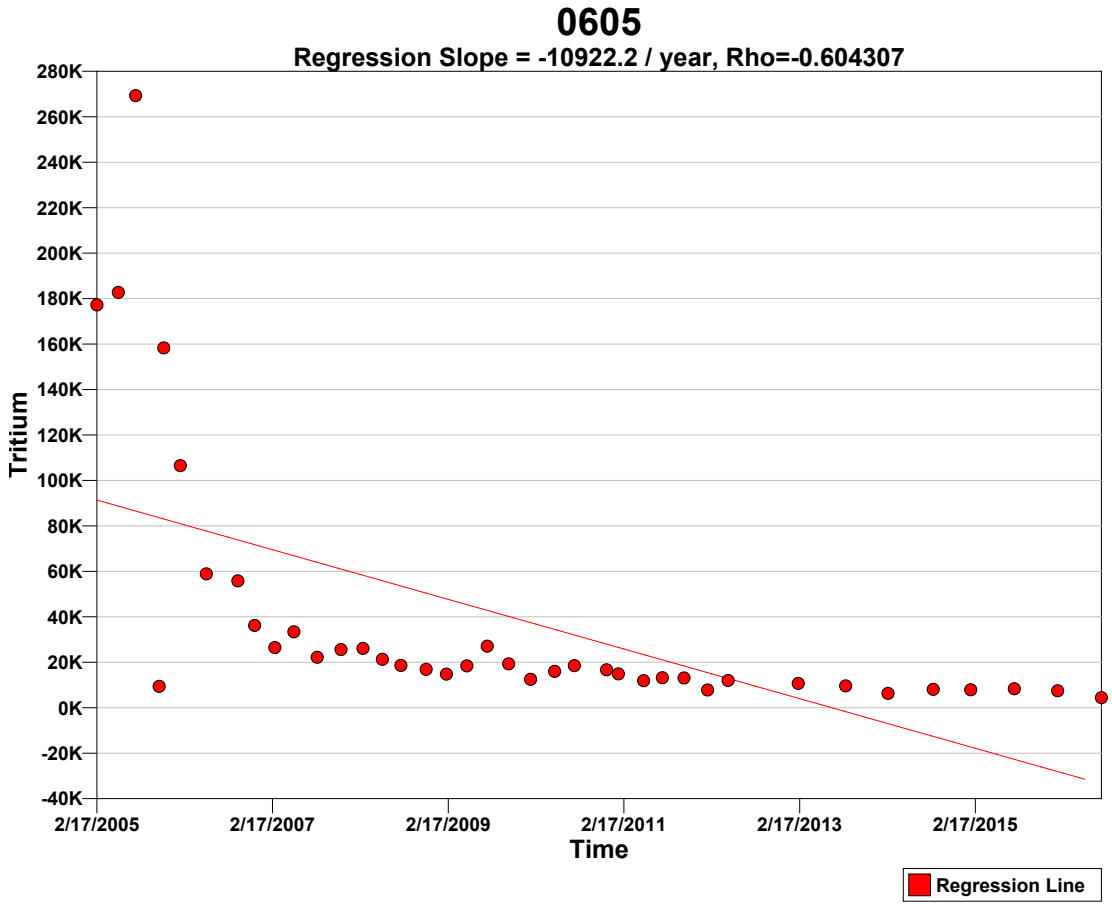
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

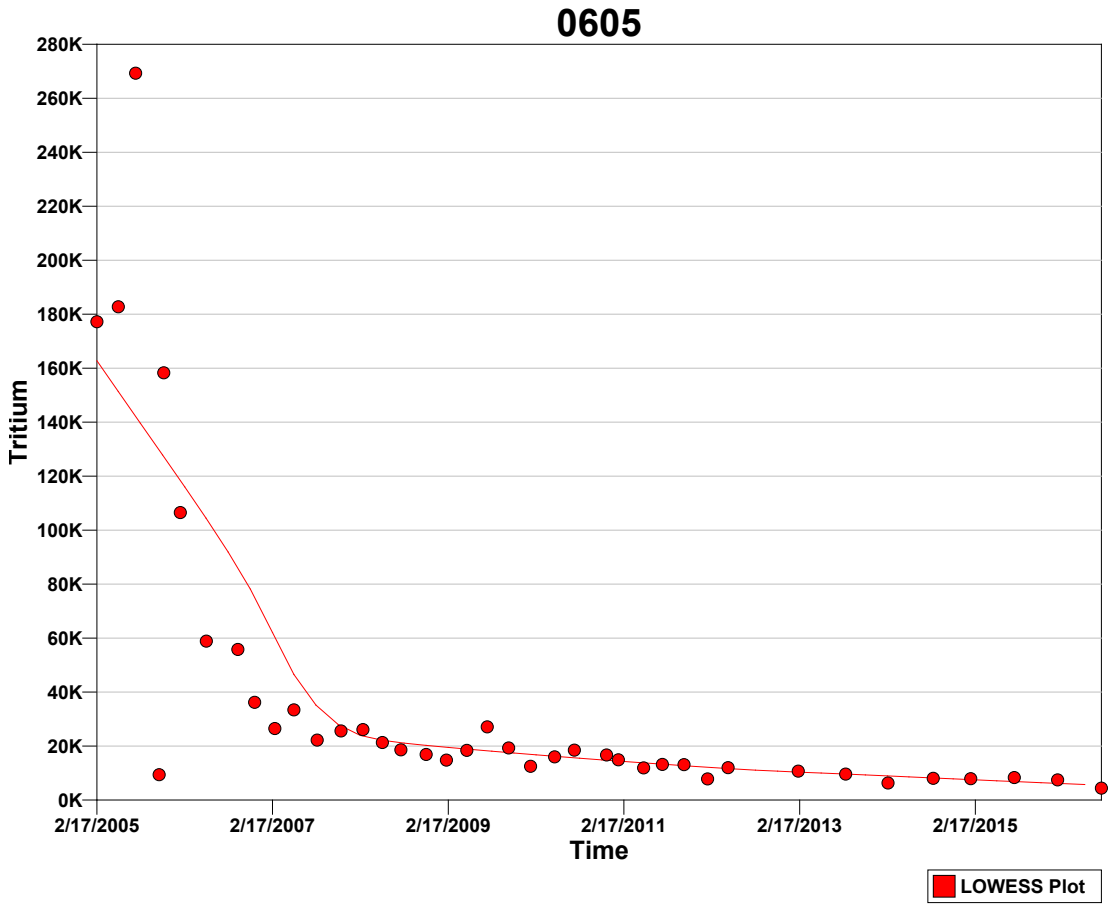
It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots

The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.





Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=39) is less than 50.

Normal Distribution Test for Tritium	
Shapiro-Wilk Test Statistic	0.57283
Shapiro-Wilk 5% Critical Value	0.939

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-583	
MK Test Statistic	Z_{MK}	-7.04038	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:53:15 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$x_2-x_1, x_3-x_1, \ldots, x_n-x_1, x_3-x_2, x_4-x_2, \ldots, x_n-x_{n-2}, x_n-x_{n-1}$

3. Let $\text{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\text{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\text{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\text{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \text{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \frac{S-1}{\sqrt{VAR(S)}} \quad \text{if} \quad S > 0$$

$$= 0 \quad \text{if} \quad S = 0 \qquad (3)$$

$$= \frac{S+1}{\sqrt{VAR(S)}} \quad \text{if} \quad S < 0$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Trichloroethene
Analysis was performed on the data that was entered for the location labeled "0606. The following table lists the data points that were used in the analysis.

0606							
Time	**Trichloroethene**		**Time**	**Trichloroethene**		**Time**	**Trichloroethene**
2/17/2005	1.4		7/27/2011	8.72		5/12/2014	0.36
5/17/2005	8.5		10/25/2011	1.8		8/26/2014	6.06
7/28/2005	15		1/31/2012	0.56		10/27/2014	5.29
11/3/2005	24		4/25/2012	5.34		1/29/2015	1
1/30/2006	2.7		7/24/2012	6.98		4/27/2015	0.61
5/18/2006	5		10/25/2012	9.01		7/29/2015	6.99
5/5/2010	3.69		2/11/2013	2.19		10/28/2015	4.08
7/26/2010	3.16		5/30/2013	6.66		1/25/2016	1.88
10/25/2010	1.78		8/27/2013	5		4/25/2016	1.81
1/25/2011	0.15		10/28/2013	7.72		7/25/2016	0.99
5/10/2011	0.95		2/19/2014	1.39		10/24/2016	1.64

SUMMARY STATISTICS for Trichloroethene								
n				33				
Min				0.15				
Max				24				
Range				23.85				
Mean				4.6185				
Median				3.16				
Variance				23.349				
StdDev				4.8321				
Std Error				0.84116				
Skewness				2.3679				
Interquartile Range				5.425				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0.15	0.297	0.58	1.395	3.16	6.82	8.894	17.7	24

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. These values with low weights can be considered possible outliers and are listed in the table below.

SUSPECTED OUTLIERS	
Date/Time	**Value**
11/3/2005	24

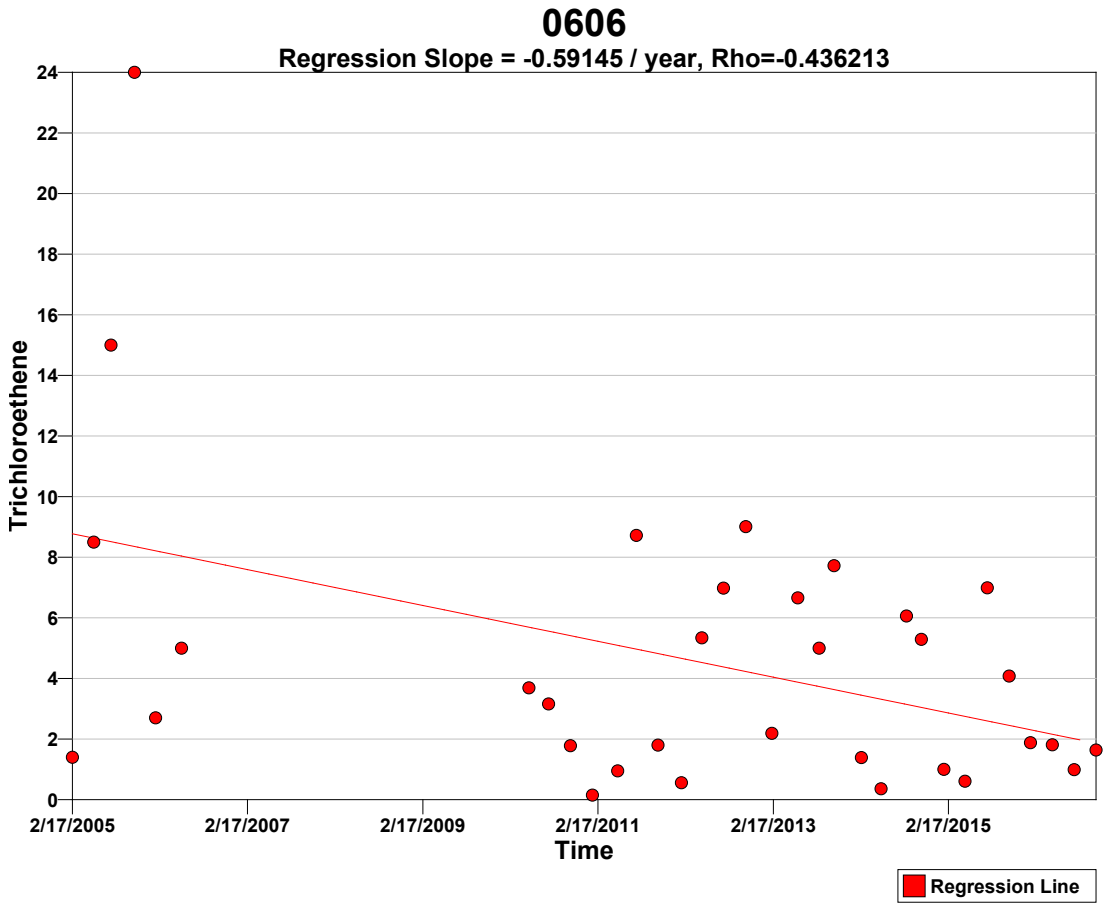
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

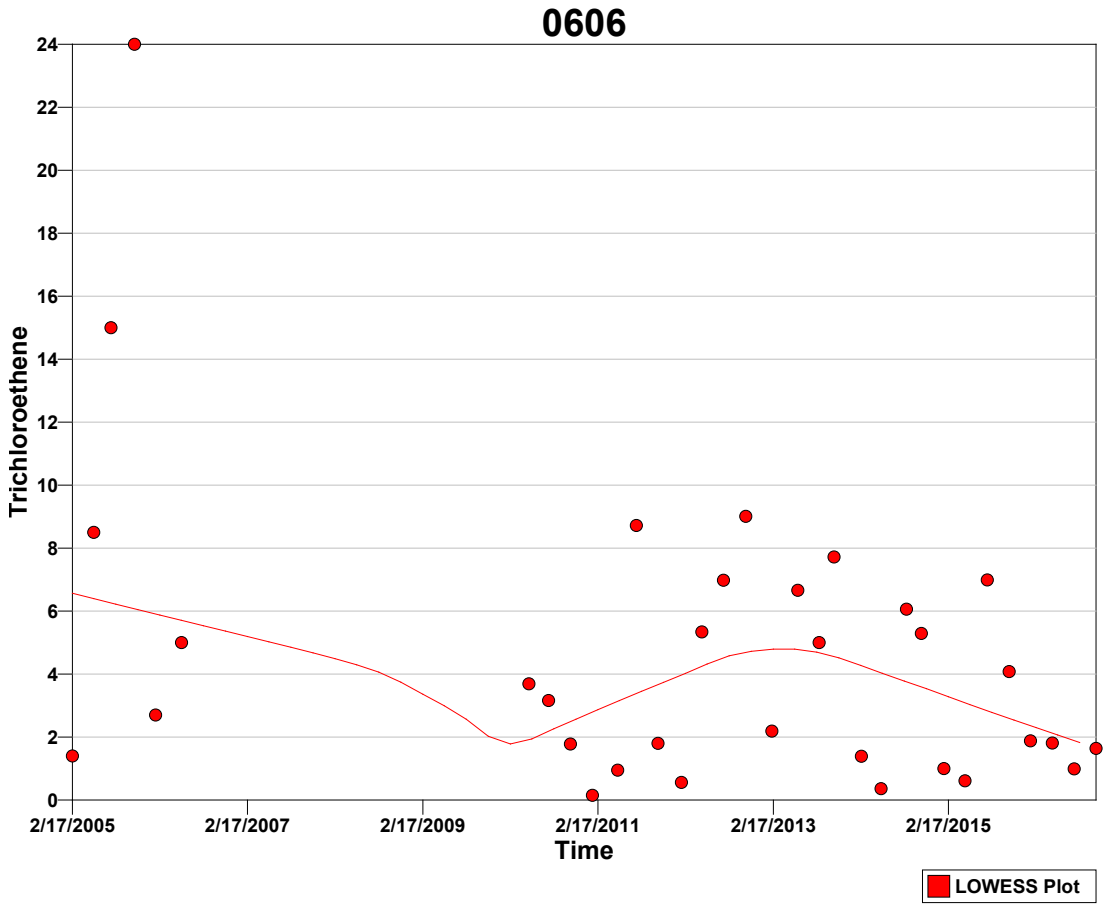
It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots

The Time vs. Measured Values Plot shows the best fitting “least squares” linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated “least squares” slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.





Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=33) is less than 50.

Normal Distribution Test for Trichloroethene	
Shapiro-Wilk Test Statistic	0.76747
Shapiro-Wilk 5% Critical Value	0.931

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-95	
MK Test Statistic	Z_{MK}	-1.45665	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	**Accept**
Alternative Hypothesis	H_a	A downward monotonic trend exists	Reject
Conclude with 95% confidence that a downward monotonic trend does not exist.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:35:09 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$x_2-x_1, x_3-x_1, \ldots, x_n-x_1, x_3-x_2, x_4-x_2, \ldots, x_n-x_{n-2}, x_n-x_{n-1}$

3. Let $\text{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\text{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\text{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\text{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \text{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \frac{S-1}{\sqrt{VAR(S)}} \quad \text{if} \quad S > 0$$

$$= 0 \quad \text{if} \quad S = 0 \qquad (3)$$

$$= \frac{S+1}{\sqrt{VAR(S)}} \quad \text{if} \quad S < 0$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Trichloroethene
Analysis was performed on the data that was entered for the location labeled "0606. The following table lists the data points that were used in the analysis.

0606							
Time	**Trichloroethene**		**Time**	**Trichloroethene**		**Time**	**Trichloroethene**
1/31/2012	0.56		10/28/2013	7.72		7/29/2015	6.99
4/25/2012	5.34		2/19/2014	1.39		10/28/2015	4.08
7/24/2012	6.98		5/12/2014	0.36		1/25/2016	1.88
10/25/2012	9.01		8/26/2014	6.06		4/25/2016	1.81
2/11/2013	2.19		10/27/2014	5.29		7/25/2016	0.99
5/30/2013	6.66		1/29/2015	1		10/24/2016	1.64
8/27/2013	5		4/27/2015	0.61			

SUMMARY STATISTICS for Trichloroethene								
n				20				
Min				0.36				
Max				9.01				
Range				8.65				
Mean				3.778				
Median				3.135				
Variance				7.9549				
StdDev				2.8204				
Std Error				0.63067				
Skewness				0.32678				
Interquartile Range				5.4125				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0.36	0.37	0.565	1.097	3.135	6.51	7.647	8.945	9.01

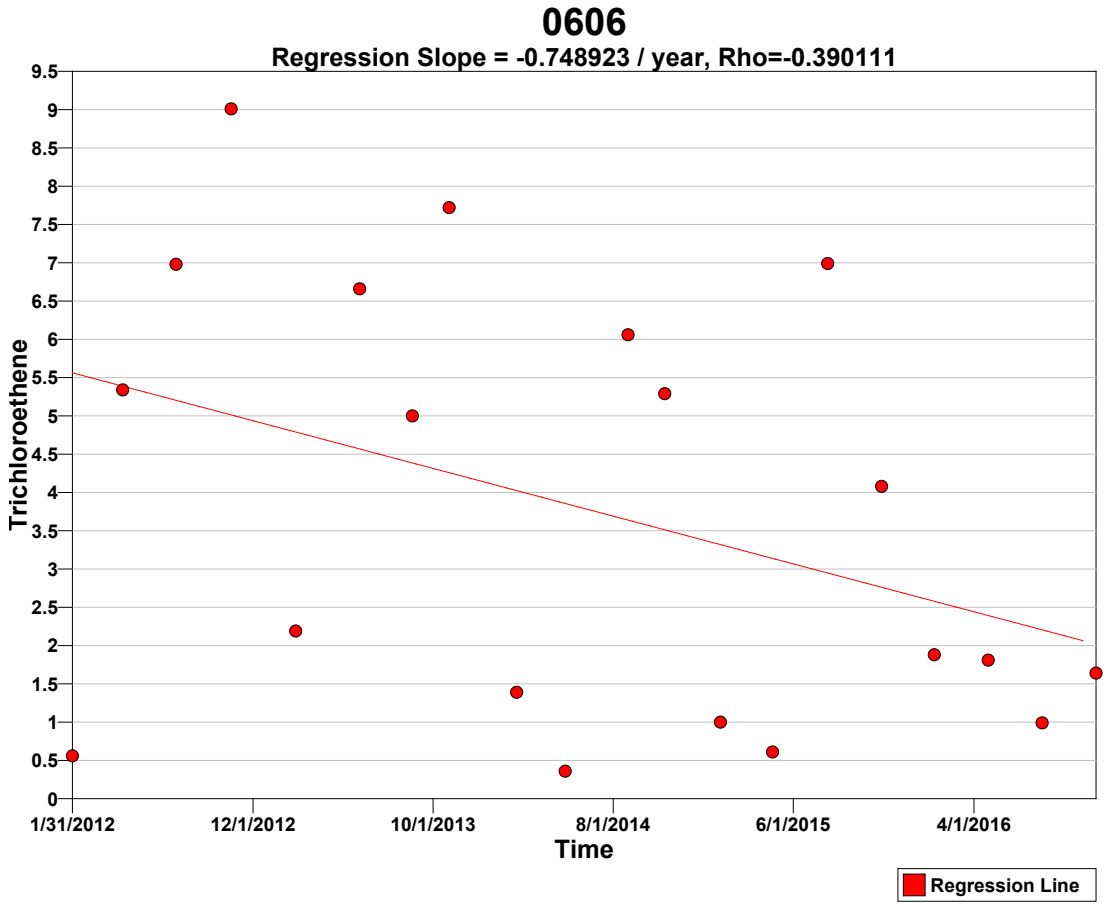
Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. The LOWESS method did not detect any outliers for the current data set.

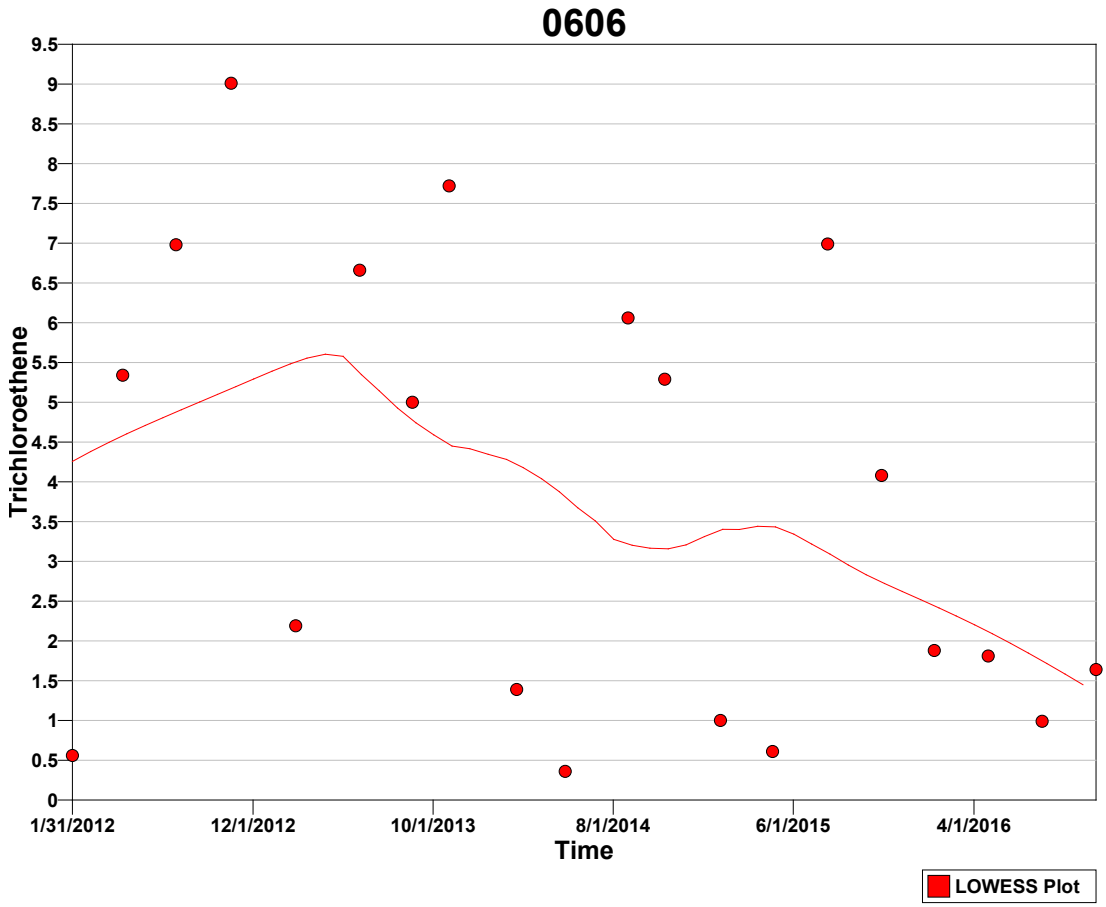
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots
The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.





Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=20) is less than 50.

Normal Distribution Test for Trichloroethene	
Shapiro-Wilk Test Statistic	0.89597
Shapiro-Wilk 5% Critical Value	0.905

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we cannot reject the hypothesis that the values are normally distributed, or in other words the values appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-48	
MK Test Statistic	Z_{MK}	-1.52488	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	**Accept**
Alternative Hypothesis	H_a	A downward monotonic trend exists	Reject
Conclude with 95% confidence that a downward monotonic trend does not exist.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 4/10/2017 7:14:32 AM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$x_2-x_1, x_3-x_1, \ldots, x_n-x_1, x_3-x_2, x_4-x_2, \ldots, x_n-x_{n-2}, x_n-x_{n-1}$

3. Let $\text{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\text{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\text{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\text{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \text{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \frac{S-1}{\sqrt{VAR(S)}} \quad \text{if} \quad S > 0$$

$$= 0 \quad \text{if} \quad S = 0 \qquad (3)$$

$$= \frac{S+1}{\sqrt{VAR(S)}} \quad \text{if} \quad S < 0$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Tritium
Analysis was performed on the data that was entered for the location labeled "0606. The following table lists the data points that were used in the analysis.

0606							
Time	**Tritium**		**Time**	**Tritium**		**Time**	**Tritium**
2/17/2005	57800		10/25/2010	13200		2/11/2013	3020
5/17/2005	64460		1/25/2011	5900		8/27/2013	7370
7/28/2005	159930		5/10/2011	5760		2/19/2014	3290
11/3/2005	76640		7/27/2011	11100		8/26/2014	5020
1/30/2006	20590		10/25/2011	4930		1/29/2015	2480
5/18/2006	30620		1/31/2012	3640		7/29/2015	5580
5/5/2010	13200		4/25/2012	8440		1/25/2016	3580
7/26/2010	14600		7/24/2012	8860		7/25/2016	4790

SUMMARY STATISTICS for Tritium								
n				24				
Min				2480				
Max				159930				
Range				1.5745e+005				
Mean				22283				
Median				7905				
Variance				1.2828e+009				
StdDev				35816				
Std Error				7310.9				
Skewness				2.9041				
Interquartile Range				14268				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
2480	2615	3155	4825	7905	1.909e+004	7.055e+004	1.391e+005	1.599e+005

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. These values with low weights can be considered possible outliers and are listed in the table below.

SUSPECTED OUTLIERS	
Date/Time	**Value**
2/17/2005	57800
7/28/2005	159930
1/30/2006	20590
5/18/2006	30620

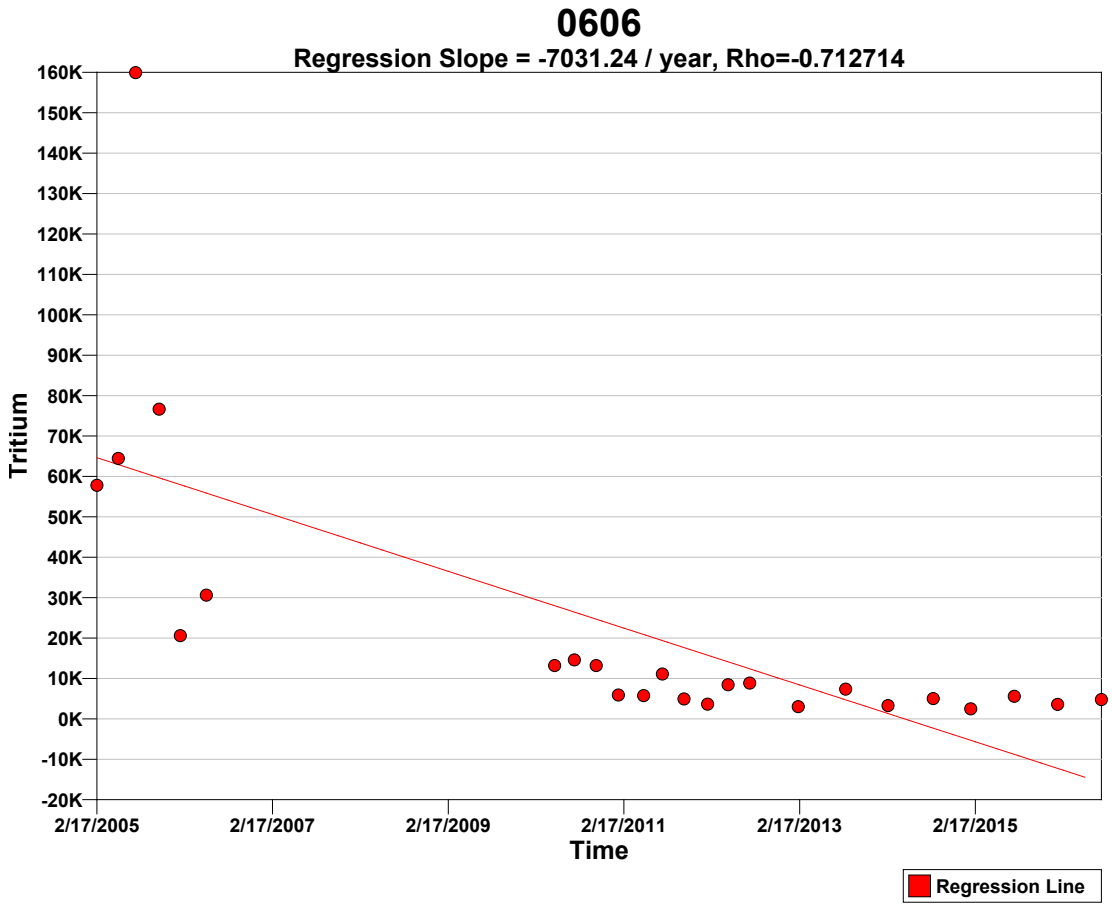
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

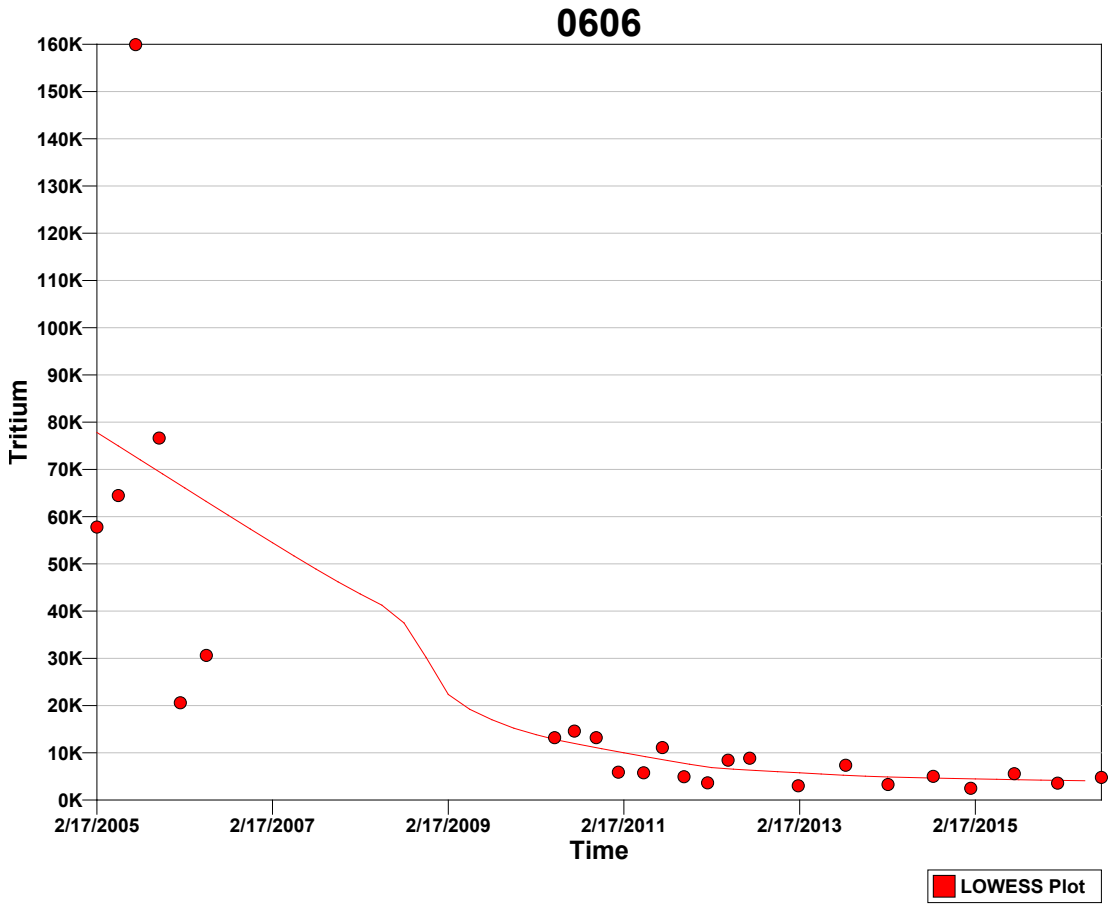
It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots

The Time vs. Measured Values Plot shows the best fitting “least squares” linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated “least squares” slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.





Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=24) is less than 50.

Normal Distribution Test for Tritium	
Shapiro-Wilk Test Statistic	0.5822
Shapiro-Wilk 5% Critical Value	0.916

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-191	
MK Test Statistic	Z_{MK}	-4.71429	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:53:55 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$x_2-x_1, x_3-x_1, \ldots, x_n-x_1, x_3-x_2, x_4-x_2, \ldots, x_n-x_{n-2}, x_n-x_{n-1}$

3. Let $\text{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\text{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\text{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\text{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \text{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \frac{S-1}{\sqrt{VAR(S)}} \quad \text{if} \quad S > 0$$

$$= 0 \quad \text{if} \quad S = 0 \qquad (3)$$

$$= \frac{S+1}{\sqrt{VAR(S)}} \quad \text{if} \quad S < 0$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Trichloroethene
Analysis was performed on the data that was entered for the location labeled "0607. The following table lists the data points that were used in the analysis.

0607							
Time	**Trichloroethene**		**Time**	**Trichloroethene**		**Time**	**Trichloroethene**
2/17/2005	4.2		2/9/2009	3.67		2/11/2013	5.61
5/17/2005	8		5/5/2009	9.43		5/30/2013	9.04
7/28/2005	15		7/29/2009	11.5		8/27/2013	7.75
11/3/2005	8.7		10/26/2009	3.63		10/28/2013	6.41
1/30/2006	7.1		1/25/2010	2.74		2/19/2014	3.18
5/18/2006	6.2		5/5/2010	11		5/12/2014	4.05
9/26/2006	8.5		7/26/2010	4.79		8/26/2014	7.53
12/5/2006	5.12		10/25/2010	3.98		10/27/2014	7.74
2/27/2007	3.67		1/25/2011	4.9		1/29/2015	4.86
5/17/2007	13		5/10/2011	3.7		4/27/2015	4.38
8/22/2007	11.4		7/27/2011	8.72		7/29/2015	6.67
11/29/2007	6.38		10/25/2011	5.51		10/28/2015	3.25
2/28/2008	3.75		1/31/2012	3.46		1/25/2016	4.87
5/19/2008	6.11		4/25/2012	9.95		4/25/2016	7.31
8/4/2008	9.69		7/24/2012	5.05		7/25/2016	6.92
11/17/2008	4.77		10/25/2012	7.45		10/24/2016	5.27

SUMMARY STATISTICS for Trichloroethene								
n	48							
Min	2.74							
Max	15							
Range	12.26							
Mean	6.5815							
Median	6.155							
Variance	7.8936							
StdDev	2.8096							
Std Error	0.40552							
Skewness	0.93891							
Interquartile Range	4.13							
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
2.74	3.212	3.613	4.245	6.155	8.375	11.04	12.32	15

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. The LOWESS method did not detect any outliers for the current data set.

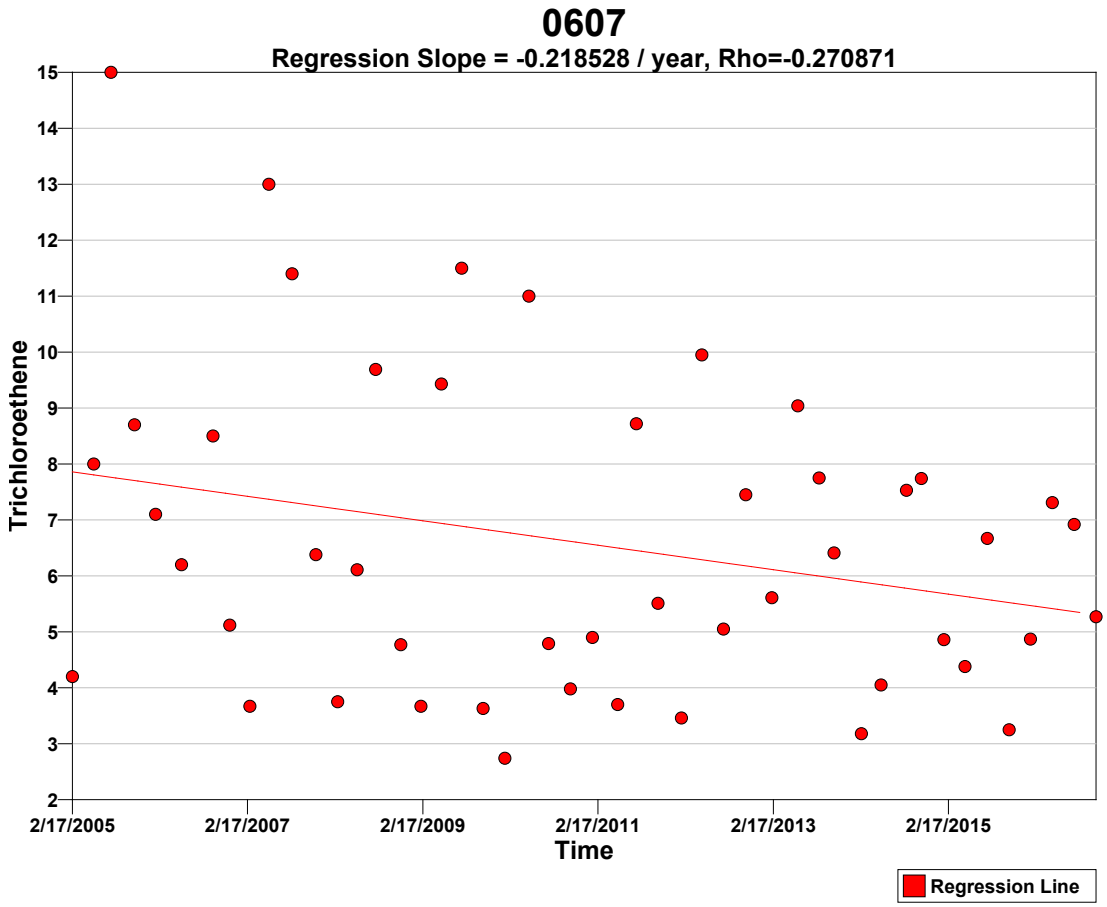
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

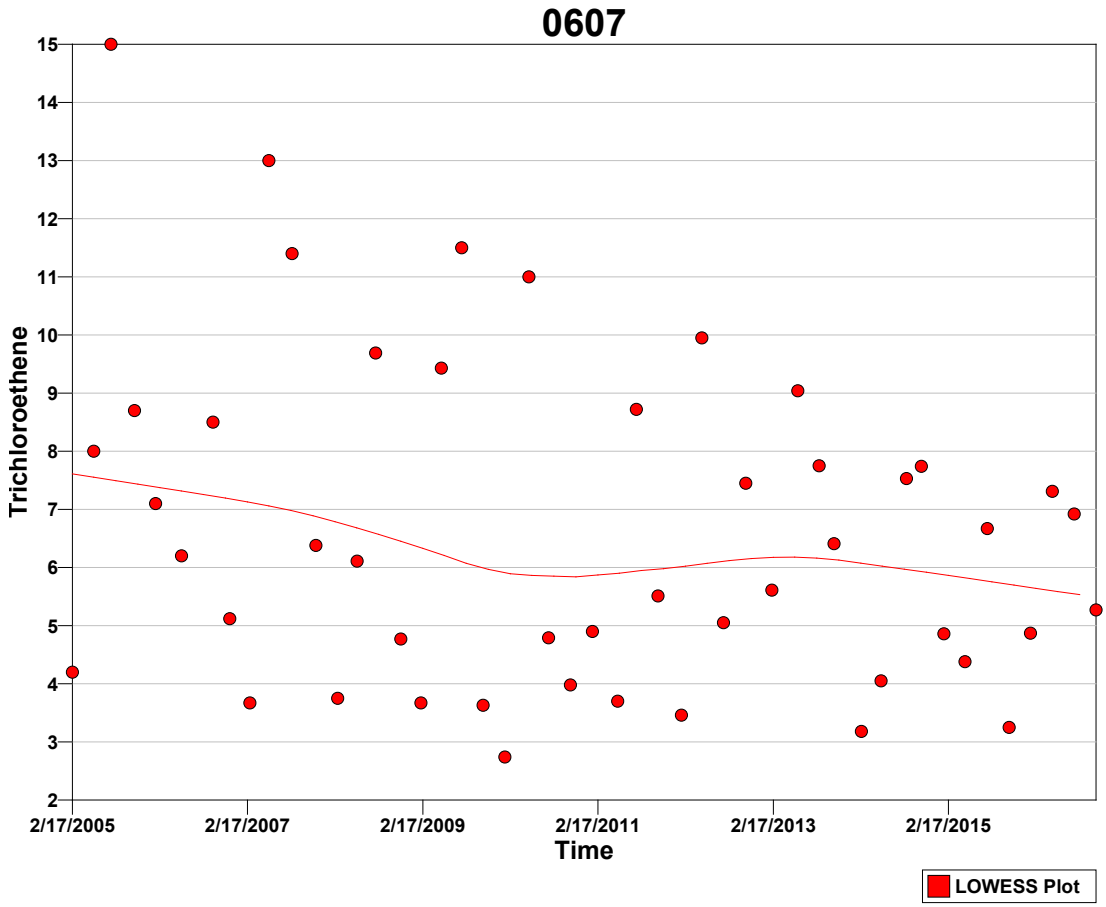
It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots

The Time vs. Measured Values Plot shows the best fitting “least squares” linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated “least squares” slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.





Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=48) is less than 50.

Normal Distribution Test for Trichloroethene	
Shapiro-Wilk Test Statistic	0.92399
Shapiro-Wilk 5% Critical Value	0.947

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-149	
MK Test Statistic	Z_{MK}	-1.31548	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	**Accept**
Alternative Hypothesis	H_a	A downward monotonic trend exists	Reject
Conclude with 95% confidence that a downward monotonic trend does not exist.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:35:44 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$x_2-x_1, x_3-x_1, \ldots, x_n-x_1, x_3-x_2, x_4-x_2, \ldots, x_n-x_{n-2}, x_n-x_{n-1}$

3. Let $\text{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\text{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\text{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\text{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \text{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \frac{S-1}{\sqrt{VAR(S)}} \quad \text{if} \quad S > 0$$

$$= 0 \quad \text{if} \quad S = 0 \qquad (3)$$

$$= \frac{S+1}{\sqrt{VAR(S)}} \quad \text{if} \quad S < 0$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Trichloroethene
Analysis was performed on the data that was entered for the location labeled "0607. The following table lists the data points that were used in the analysis.

0607							
Time	**Trichloroethene**		**Time**	**Trichloroethene**		**Time**	**Trichloroethene**
1/31/2012	3.46		10/28/2013	6.41		7/29/2015	6.67
4/25/2012	9.95		2/19/2014	3.18		10/28/2015	3.25
7/24/2012	5.05		5/12/2014	4.05		1/25/2016	4.87
10/25/2012	7.45		8/26/2014	7.53		4/25/2016	7.31
2/11/2013	5.61		10/27/2014	7.74		7/25/2016	6.92
5/30/2013	9.04		1/29/2015	4.86		10/24/2016	5.27
8/27/2013	7.75		4/27/2015	4.38			

SUMMARY STATISTICS for Trichloroethene								
n				20				
Min				3.18				
Max				9.95				
Range				6.77				
Mean				6.0375				
Median				6.01				
Variance				3.7414				
StdDev				1.9343				
Std Error				0.43252				
Skewness				0.21716				
Interquartile Range				3.01				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
3.18	3.184	3.271	4.5	6.01	7.51	8.911	9.904	9.95

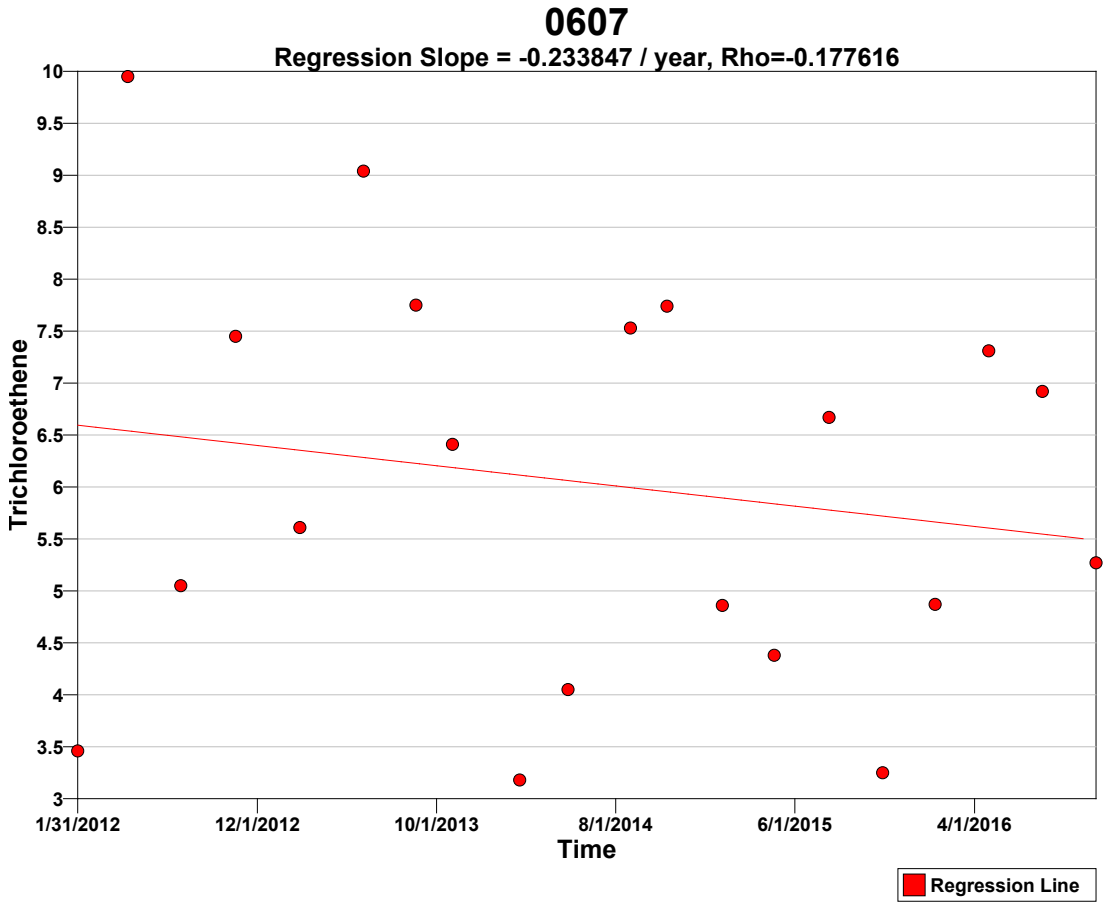
Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. The LOWESS method did not detect any outliers for the current data set.

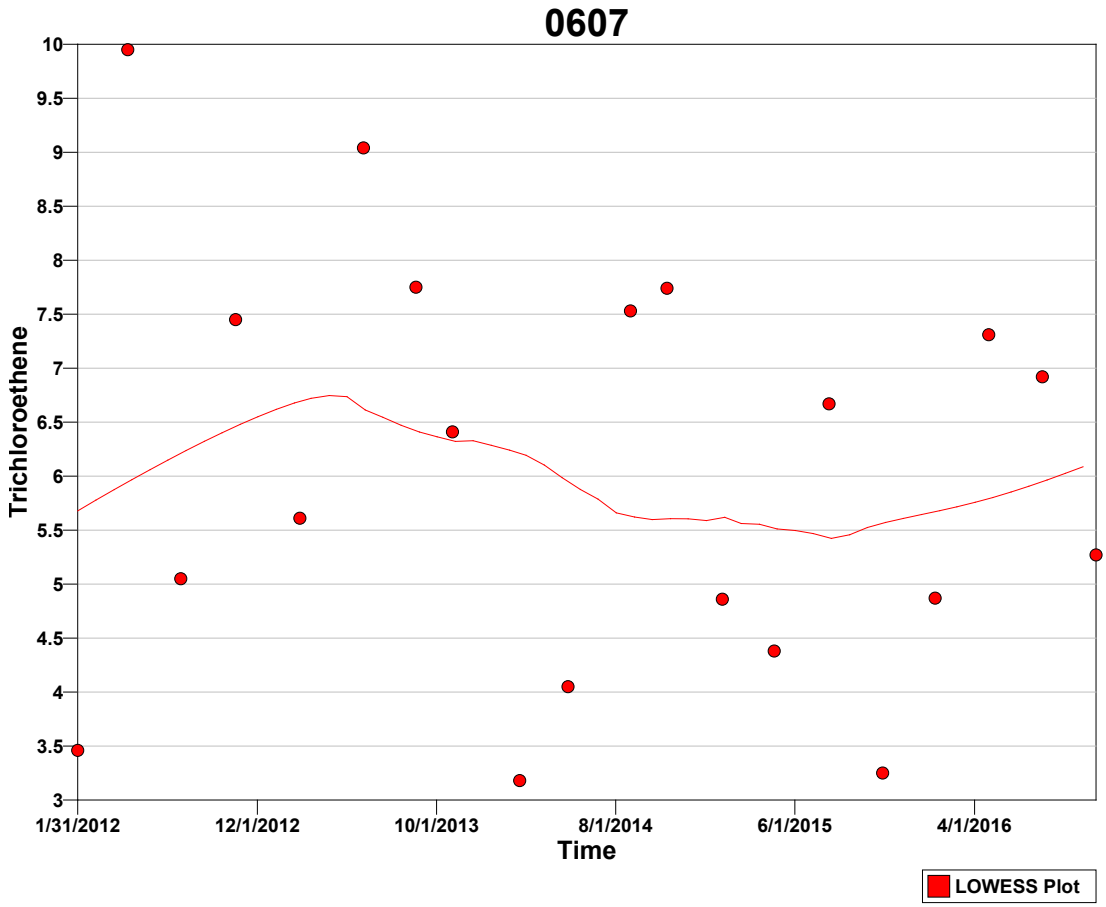
Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots
The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.





Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=20) is less than 50.

Normal Distribution Test for Trichloroethene	
Shapiro-Wilk Test Statistic	0.95998
Shapiro-Wilk 5% Critical Value	0.905

The calculated test statistic exceeds the 5% Shapiro-Wilk critical value, so we cannot reject the hypothesis that the values are normally distributed, or in other words the values appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-20	
MK Test Statistic	Z_{MK}	-0.616441	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	**Accept**
Alternative Hypothesis	H_a	A downward monotonic trend exists	Reject
Conclude with 95% confidence that a downward monotonic trend does not exist.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 4/10/2017 7:15:18 AM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$x_2-x_1, x_3-x_1, \ldots, x_n-x_1, x_3-x_2, x_4-x_2, \ldots, x_n-x_{n-2}, x_n-x_{n-1}$

3. Let $\text{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\text{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\text{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\text{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \text{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \frac{S-1}{\sqrt{VAR(S)}} \quad \text{if} \quad S > 0$$

$$= 0 \quad \text{if} \quad S = 0 \qquad (3)$$

$$= \frac{S+1}{\sqrt{VAR(S)}} \quad \text{if} \quad S < 0$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Tritium
Analysis was performed on the data that was entered for the location labeled "0607. The following table lists the data points that were used in the analysis.

0607							
Time	**Tritium**		**Time**	**Tritium**		**Time**	**Tritium**
2/17/2005	107450		5/19/2008	8060		7/27/2011	5370
5/17/2005	85560		8/4/2008	14700		10/25/2011	5690
7/28/2005	133130		11/17/2008	7680		1/31/2012	3430
11/3/2005	71410		2/9/2009	4760		4/25/2012	5170
1/30/2006	31830		5/5/2009	7660		7/24/2012	6120
5/18/2006	11430		7/29/2009	10700		2/11/2013	4510
9/26/2006	25300		10/26/2009	5750		8/27/2013	5100
12/5/2006	16000		1/25/2010	2910		2/19/2014	2170
2/27/2007	9030		5/5/2010	3880		8/26/2014	4160
5/17/2007	12200		7/26/2010	6630		1/29/2015	2760
8/22/2007	14100		10/25/2010	8840		7/29/2015	3900
11/29/2007	12900		1/25/2011	7040		1/25/2016	3710
2/28/2008	7190		5/10/2011	3710		7/25/2016	3080

SUMMARY STATISTICS for Tritium								
n	39							
Min	2170							
Max	133130							
Range	1.3096e+005							
Mean	17565							
Median	7040							
Variance	8.7825e+008							
StdDev	29635							
Std Error	4745.4							
Skewness	2.8683							
Interquartile Range	8740							
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
2170	2760	3080	4160	7040	1.29e+004	7.141e+004	1.075e+005	1.331e+005

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. These values with low weights can be considered possible outliers and are listed in the table below.

SUSPECTED OUTLIERS	
Date/Time	**Value**
2/17/2005	107450
7/28/2005	133130
1/30/2006	31830
5/18/2006	11430
9/26/2006	25300
12/5/2006	16000
2/27/2007	9030
5/17/2007	12200

Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing

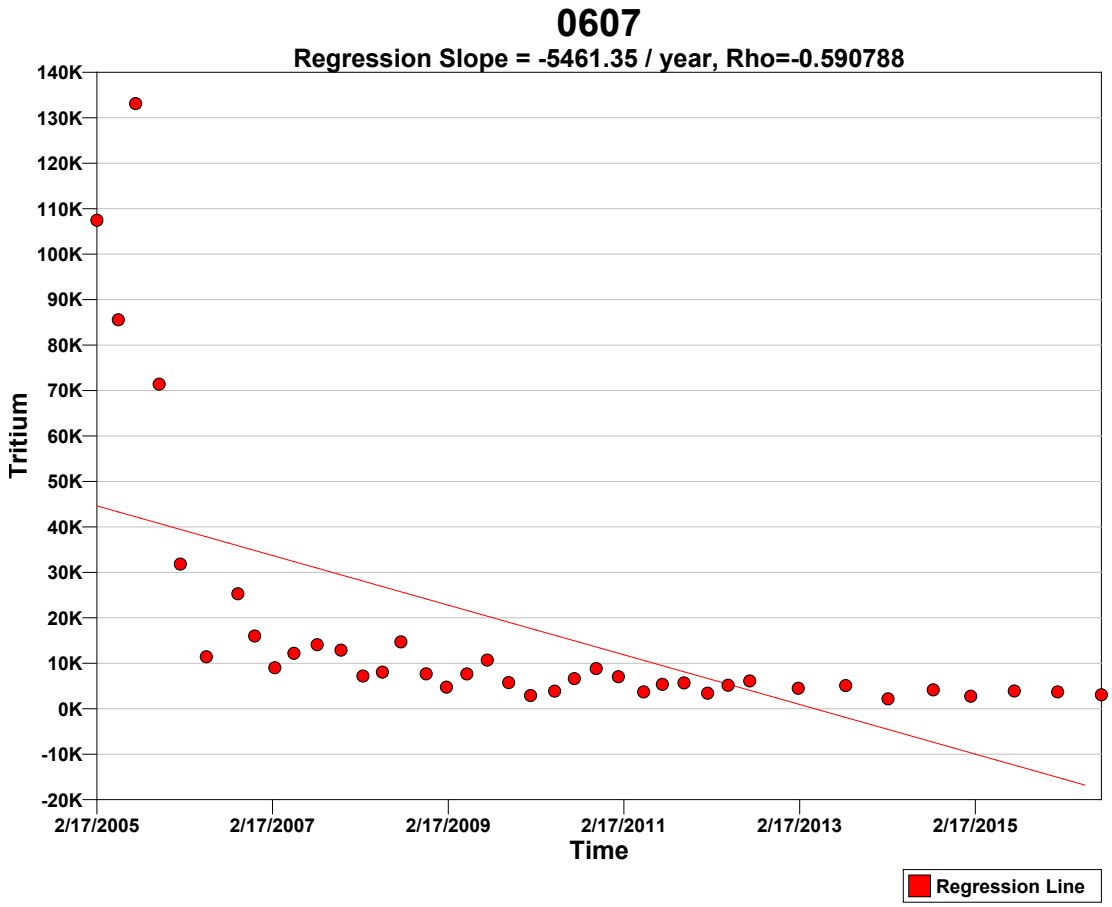
them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

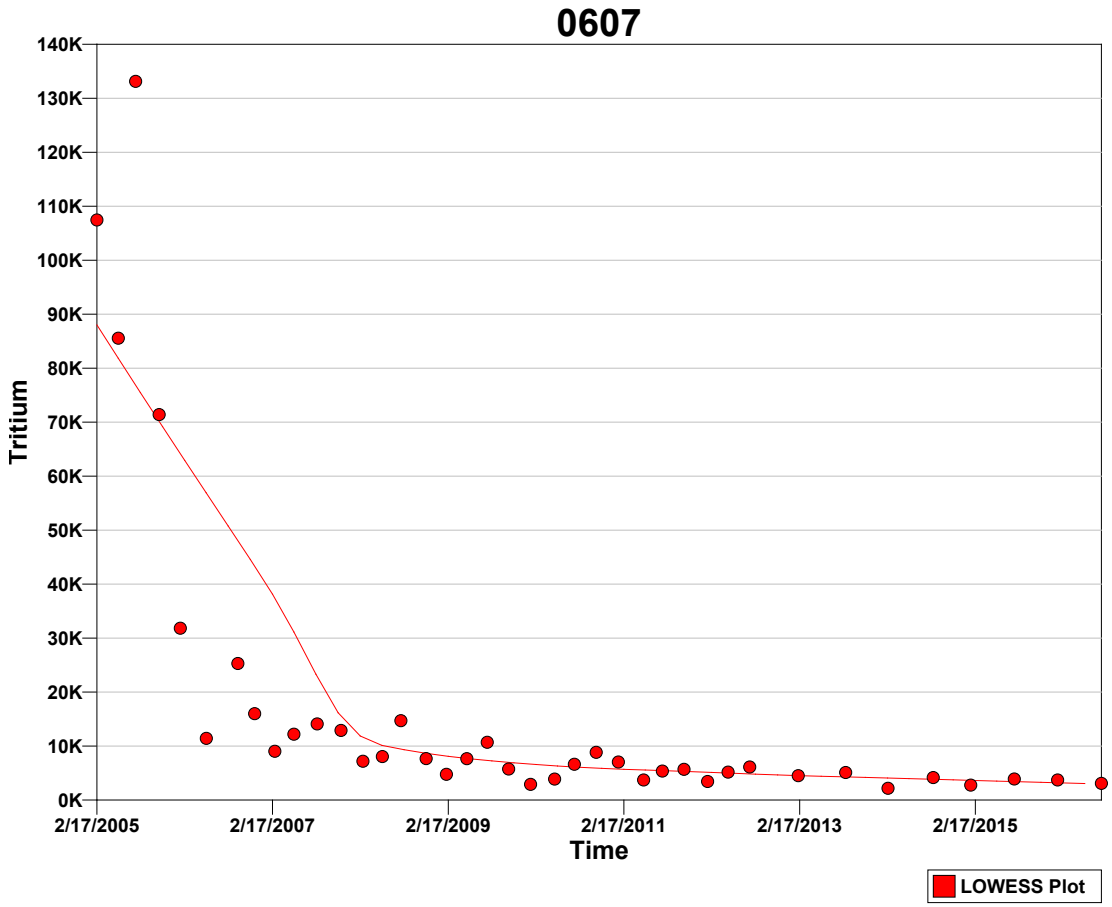
It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

Data Plots

The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated "least squares" slope obtained assuming that a linear change over time is occurring.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.





Tests
A goodness-of-fit test was performed to test whether the values are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=39) is less than 50.

Normal Distribution Test for Tritium	
Shapiro-Wilk Test Statistic	0.52206
Shapiro-Wilk 5% Critical Value	0.939

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the values are normally distributed, or in other words the values do not appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-534	
MK Test Statistic	Z_{MK}	-6.4481	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:54:35 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$x_2-x_1, x_3-x_1, \ldots, x_n-x_1, x_3-x_2, x_4-x_2, \ldots, x_n-x_{n-2}, x_n-x_{n-1}$

3. Let $\text{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\text{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\text{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\text{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \text{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \frac{S-1}{\sqrt{VAR(S)}} \quad \text{if} \quad S > 0$$

$$= 0 \quad \text{if} \quad S = 0 \qquad (3)$$

$$= \frac{S+1}{\sqrt{VAR(S)}} \quad \text{if} \quad S < 0$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for cis-1x2-Dichloroethene
Analysis was performed on the data that was entered for the location labeled "0617. The following table lists the data points that were used in the analysis.

0617							
Time	**cis-1x2-Dichloroethene**		**Time**	**cis-1x2-Dichloroethene**		**Time**	**cis-1x2-Dichloroethene**
10/18/2001	3.1		11/18/2005	2.3		7/30/2010	1.97
1/30/2002	1.3		1/31/2006	2.7		1/25/2011	2.2
5/1/2002	2		4/27/2006	2.1		7/27/2011	1.36
4/22/2003	1.6		8/3/2006	2.15		1/31/2012	1.71
7/25/2003	2.6		11/17/2006	0.969		7/26/2012	0.59
10/21/2003	2.8		2/28/2007	1.1		2/11/2013	2.36
1/22/2004	2.1		5/23/2007	2.82		8/27/2013	1.59
4/21/2004	1.5		8/24/2007	1.72		2/19/2014	0.6
7/12/2004	2.6		11/14/2007	2.57		12/8/2014	1.88
11/17/2004	2.9		2/19/2008	2.29		1/29/2015	1.94
3/2/2005	2.1		2/9/2009	1.31		7/29/2015	1.86
5/24/2005	4.7		7/29/2009	2.06		1/25/2016	1.8
8/3/2005	2.3		1/25/2010	0.724		7/25/2016	1.08

SUMMARY STATISTICS for cis-1x2-Dichloroethene								
n				39				
Min				-1.148031				
Max				2.49667				
Range				3.6447				
Mean				0				
Median				0.1110284				
Variance				0.50655				
StdDev				0.71172				
Std Error				0.11397				
Skewness				0.69096				
Interquartile Range				0.84532				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
-1.148	-1.134	-1.083	-0.4296	0.111	0.4157	0.6587	0.7641	2.497

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. These values with low weights can be considered possible outliers and are listed in the table below.

SUSPECTED OUTLIERS	
Date/Time	**Value**
5/24/2005	4.7

Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent

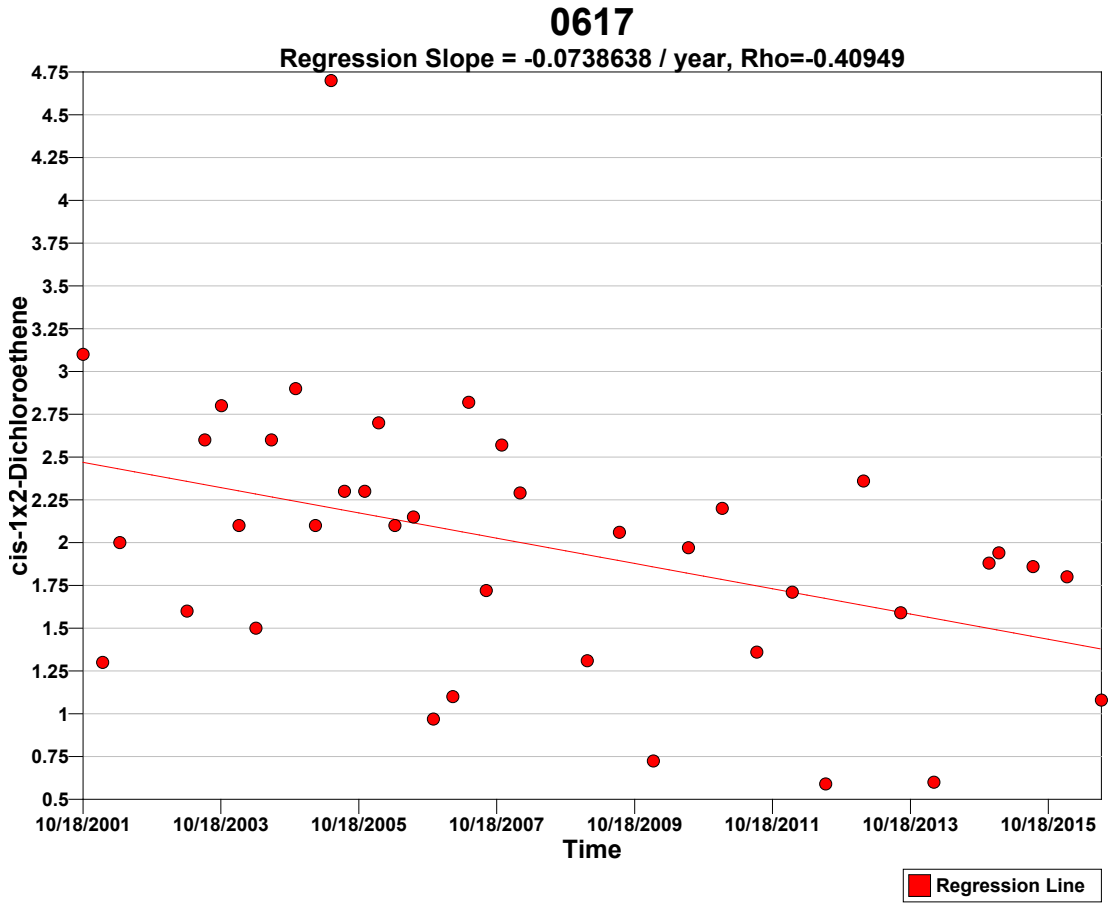
examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

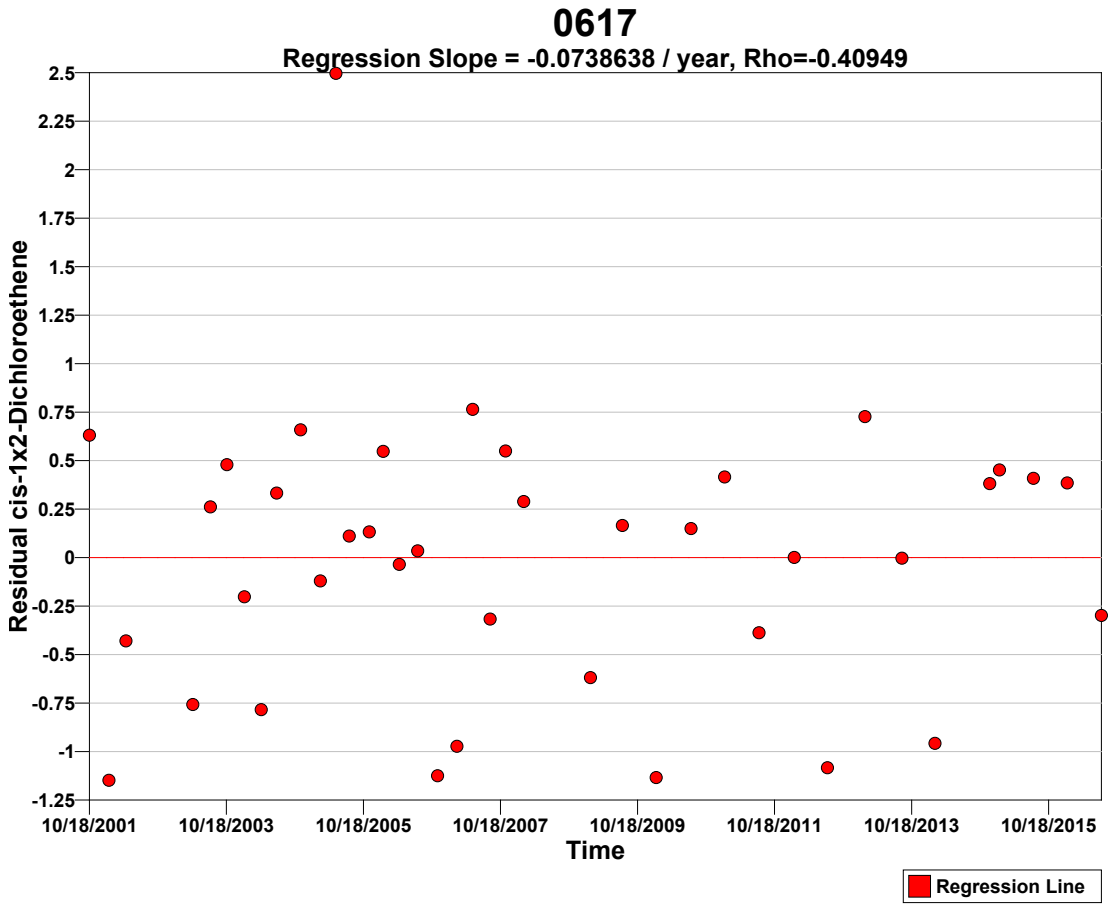
Data Plots

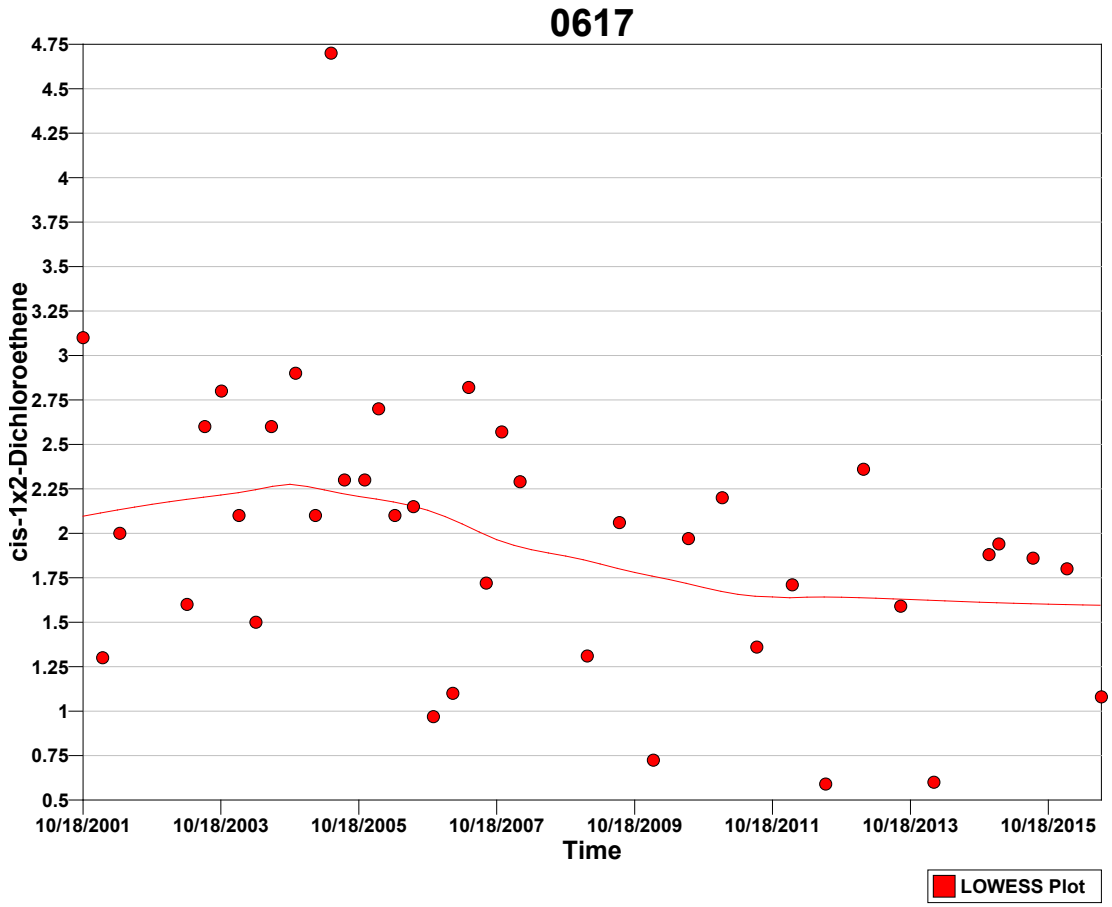
The Time vs. Measured Values Plot shows the best fitting “least squares” linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated “least squares” slope obtained assuming that a linear change over time is occurring.

The Time vs. Residual Values Plot shows the deviation (difference) of each observed value from the "least squares" linear line described above. The horizontal line at value zero represents a perfect fit (no difference) to the "least squares" linear line.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.




0617
Regression Slope = -0.0738638 / year, Rho=-0.40949
Residual cis-1x2-Dichloroethene
2.5
2.25
2
1.75
1.5
1.25
1
0.75
0.5
0.25
0
-0.25
-0.5
-0.75
-1
-1.25
10/18/2001
10/18/2003
10/18/2005
10/18/2007
10/18/2009
10/18/2011
10/18/2013
10/18/2015
Time
Regression Line



Tests

A goodness-of-fit test was performed to test whether the residuals (deviations from the linear regression line described and displayed above) are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=39) is less than 50.

Normal Distribution Test of Residuals for cis-1x2-Dichloroethene	
Shapiro-Wilk Test Statistic	0.91506
Shapiro-Wilk 5% Critical Value	0.939

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we cannot reject the hypothesis that the residuals are normally distributed, or in other words the residuals appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of residual values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test

The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-226	
MK Test Statistic	Z_{MK}	-2.72293	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	Reject
Alternative Hypothesis	H_a	A downward monotonic trend exists	**Accept**
Conclude with 95% confidence that a downward monotonic trend exists.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:49:21 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Mann-Kendall Test for Monotonic Trend

Primary Objective

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear.

Selected Statistical Testing Approach

The MK test can be used in place of a parametric linear regression analysis that is used to test if the slope of the estimated linear regression line is different from zero.The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test. Hence, the MK test is a nonparametric (distribution-free) test.

Hirsch, Slack and Smith (1982, page 107) indicate that the MK test is best viewed as an exploratory analysis and is most appropriately used to identify and quantify changes over time at sampling stations.

Calculations to Determine Whether a Trend Exists

The MK test is used to decide whether to reject the null hypothesis (H_o) that no monotonic trend exists in favor of the alternative hypothesis (H_a) that a monotonic trend exists.

One of three alternative hypotheses is chosen:

1. A monotonic downward trend exists
2. Either a monotonic upward or monotonic downward trend exists
3. A monotonic upward trend exists

For the current design, the 1st option (a downward trend) was chosen.

The data obtained over time must be convincing beyond a reasonable doubt before the MK test will reject the H_o and accept the H_a.

The MK test is conducted as follows (from Gilbert 1987, pp. 209-213):

1. List the data in the order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times 1, 2, ..., n, respectively. The data are not necessarily (and need not be) collected at equally spaced time intervals, although equally spaced sampling over time is often preferred.

2. Determine the sign of all *n(n-1)/2* possible differences $x_j - x_k$, where j > k. These differences are

$x_2-x_1, x_3-x_1, \ldots, x_n-x_1, x_3-x_2, x_4-x_2, \ldots, x_n-x_{n-2}, x_n-x_{n-1}$

3. Let $\text{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

$\text{sgn}(x_j - x_k) = 1$ if $x_j - x_k > 0$
$\text{sgn}(x_j - x_k) = 0$ if $x_j - x_k = 0$, or if the sign of $x_j - x_k$ cannot be determined due to nondetects
$\text{sgn}(x_j - x_k) = -1$ if $x_j - x_k < 0$

For example, if $x_j - x_k > 0$, then the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \text{sgn}(x_j - x_k) \qquad (1)$$

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be *larger* than observations made earlier. If S is a negative number, then observations made later in time tend to be *smaller* than observations made earlier.

5. If $n \leq 10$, follow the procedure described in Gilbert (1987, page 209, Section 16.4.1) by looking up S in a table of probabilities (Gilbert 1987, Table A18, page 272). If this probability is less than α (the probability of concluding a trend exists when there is none), then reject the null hypothesis and conclude the trend exists. If *n* cannot be found in the table of probabilities (which can happen if there are tied data values), the next value farther from zero in the table is used. For example, if S=12 and there is no value for S=12 in the table, it is handled the same as S=13.

If $n>10$, continue with steps 6 through 8 to determine whether a trend exists. This follows the procedure described in Gilbert (1987, page 211, Section 16.4.2).

6. Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5) \right] \qquad (2)$$

where g is the number of tied groups and t_p is the number of observations in the pth group. For example, in the sequence of measurements in time (23, 24, 29, 6, 29, 24, 24, 29, 23) we have g = 3 tied groups, for which t_1=2 for the tied value 23, t_2=3 for the tied value 24, and t_3=3 for the tied value 29.

7. Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \frac{S-1}{\sqrt{VAR(S)}} \quad \text{if} \quad S > 0$$

$$= 0 \quad \text{if} \quad S = 0 \qquad (3)$$

$$= \frac{S+1}{\sqrt{VAR(S)}} \quad \text{if} \quad S < 0$$

A positive value of Z_{MK} indicates that the data tend to increase with time; a negative value of Z_{MK} indicates that the data tend to decrease with time.

8. Finally, the hypothesis is tested. H_o is rejected and H_a is accepted if $Z_{MK} \leq -Z_{1-\alpha}$ where:

H_o (null hypothesis): no monotonic trend exists

H_a (alternative hypothesis): a downward monotonic trend exists

Alpha (α) is the Type I error rate, which is the user-specified small probability that can be tolerated that the MK test will falsely reject the null hypothesis (i.e. will conclude a trend exists when there is none)

$Z_{1-\alpha}$ is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution. For example, if $\alpha = 0.05$, then $Z_{1-\alpha} = 1.64485$. Values of $Z_{1-\alpha}$ are provided in many statistics books (for example Gilbert 1987, Table A1, page 254) and statistical software packages.

Assumptions

The following assumptions underlie the MK test:

1. When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
2. The observations obtained over time are representative of the true conditions at sampling times.
3. The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

The MK test does not require that the measurements or the residuals about a trend tine be normally distributed or that the trend, if present, be linear. However, when VSP computes the number of samples required to detect a linear trend, VSP assumes that the residuals about an assumed linear trend line are normally distributed as explained previously.

The MK test can be computed if there are missing values (no measurements for some sampling times), but the performance of the test will be adversely affected. The assumption of independence requires that the time between samples be sufficiently long so that there is no correlation between measurements collected at different times. Consult Helsel (2005, pages 209-215) for doing nonparametric regression (trend) analyses when data are less than one or more detection limits.

Data Analysis for Trichloroethene
Analysis was performed on the data that was entered for the location labeled "0617. The following table lists the data points that were used in the analysis.

0617							
Time	**Trichloroethene**		**Time**	**Trichloroethene**		**Time**	**Trichloroethene**
10/18/2001	12		11/18/2005	5.8		7/30/2010	8.14
1/30/2002	3.8		1/31/2006	8.9		1/25/2011	9.78
5/1/2002	7.9		4/27/2006	6.2		7/27/2011	6.66
4/22/2003	6.1		8/3/2006	8.67		1/31/2012	7.76
7/25/2003	7.4		11/17/2006	3.9		7/26/2012	1.84
10/21/2003	8.8		2/28/2007	3.57		2/11/2013	10.4
1/22/2004	7		5/23/2007	10.4		8/27/2013	6.7
4/21/2004	5.7		8/24/2007	4.95		2/19/2014	2.52
7/12/2004	9.8		11/14/2007	8.5		12/8/2014	8.31
11/17/2004	8.6		2/19/2008	8.4		1/29/2015	8.26
3/2/2005	7		2/9/2009	4.48		7/29/2015	8.05
5/24/2005	9		7/29/2009	8.17		1/25/2016	9.09
8/3/2005	6.1		1/25/2010	2.89		7/25/2016	3.53

SUMMARY STATISTICS for Trichloroethene								
n	39							
Min	-4.909218							
Max	4.470781							
Range	9.38							
Mean	0							
Median	0.4094443							
Variance	5.7011							
StdDev	2.3877							
Std Error	0.38234							
Skewness	-0.35627							
Interquartile Range	3.3289							
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
-4.909	-4.116	-3.709	-1.648	0.4094	1.681	2.922	3.69	4.471

Outliers
One of the trend data plots in VSP uses the LOWESS method to fit a nonlinear trend line through the data while giving little weight to values that appear to be out of place and that don't follow the trend. The LOWESS method did not detect any outliers for the current data set.

Data should not be excluded from analysis solely on the basis of the results of this method or any other outlier detection method. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them. Additionally, the default LOWESS parameters are robust for most datasets, but in some extreme cases some additional points besides the obvious outliers may be flagged as outliers while appearing to fit the fitted line rather well.

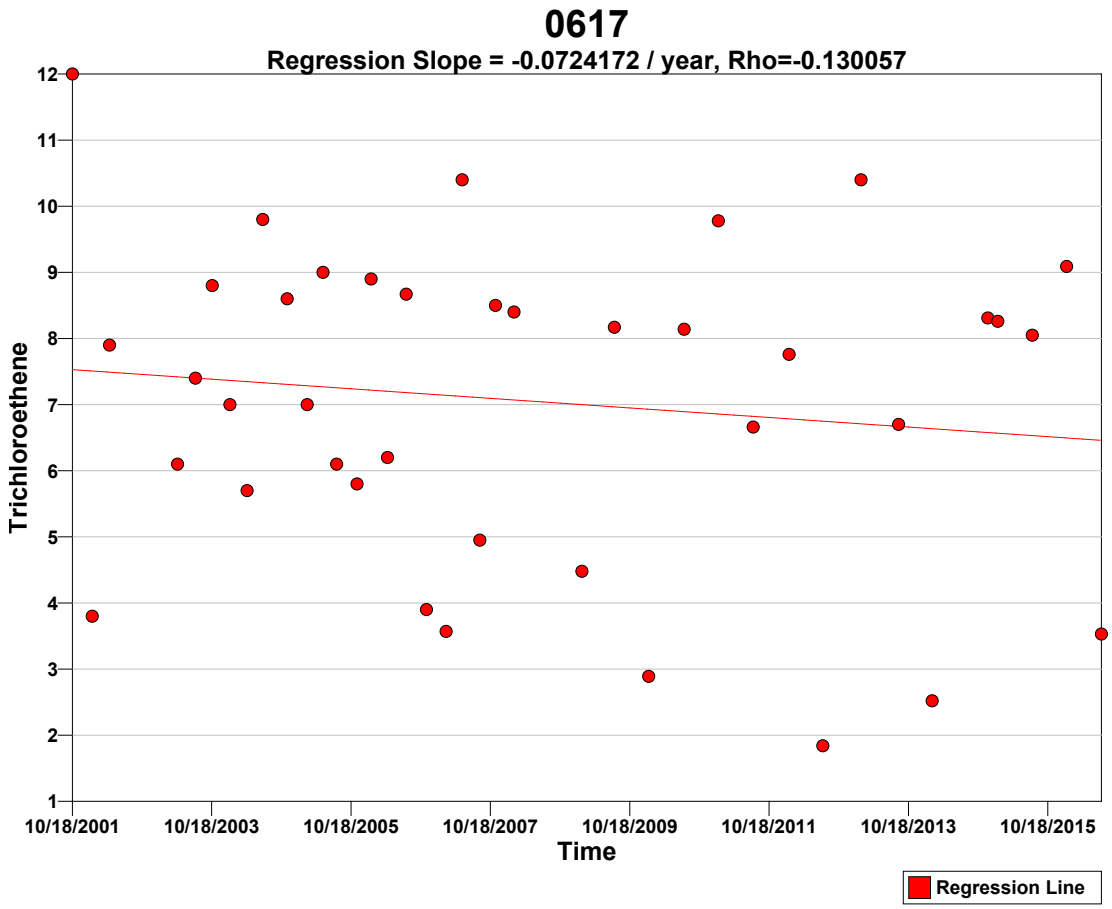
It is also recommended that the trend data plots be studied in depth to identify potential outliers and that the raw data be checked by an expert. No outlier detection method can replace a thorough and diligent examination by someone familiar with the data. Ideally, data plots are examined and the obvious erroneous values are addressed before relying on an outlier test.

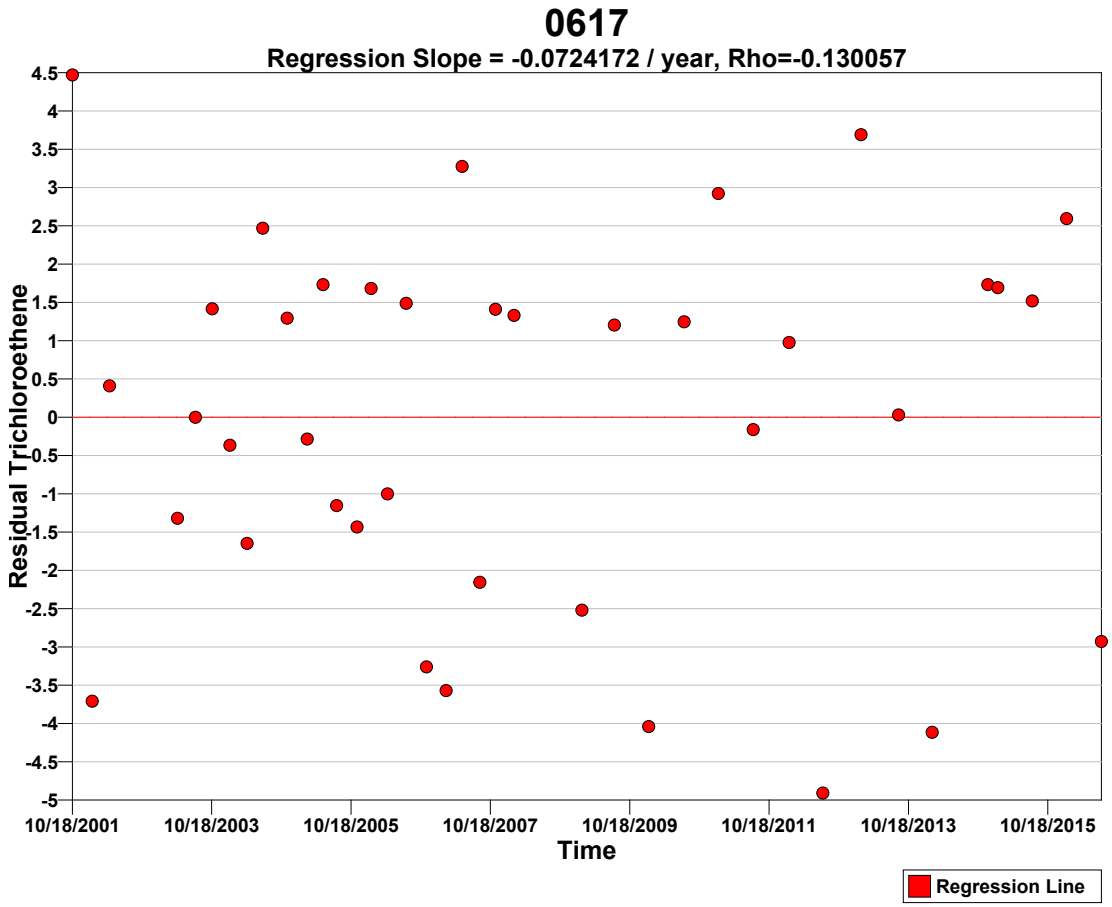
Data Plots
The Time vs. Measured Values Plot shows the best fitting "least squares" linear line to the observed n data values plotted against time. The x-axis is the time when data were collected and the y-axis is the

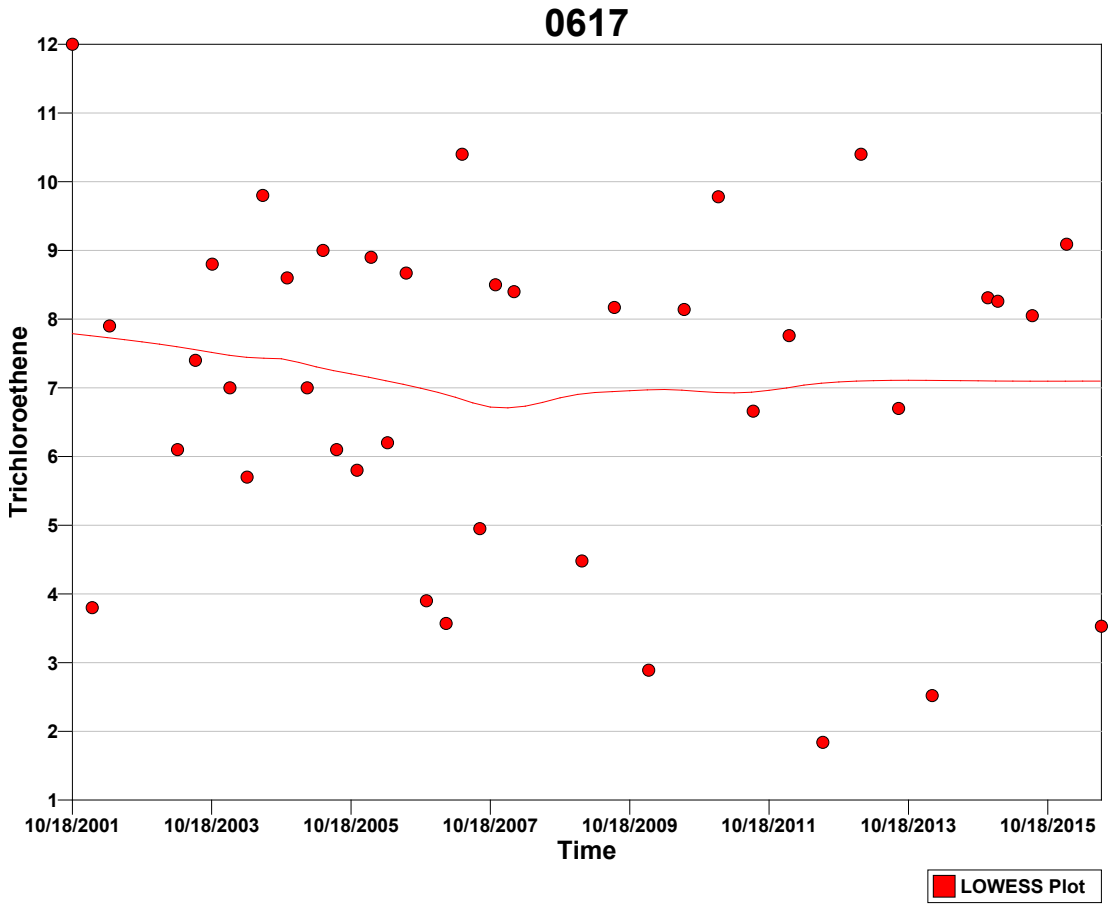
value of each datum. The slope of the linear line estimates the change in the data per unit time. If the linear line is a good fit to the data, then the slope of the line is an estimate of the linear trend over time in the data. If a linear line is not a good fit to the data, but an upward (or downward) tendency in the data is present, then a nonparametric estimator of monotonic change over time may be preferred to the estimated “least squares” slope obtained assuming that a linear change over time is occurring.

The Time vs. Residual Values Plot shows the deviation (difference) of each observed value from the "least squares" linear line described above. The horizontal line at value zero represents a perfect fit (no difference) to the "least squares" linear line.

The last plot is a Locally Weighted Smoothed Scatterplot (LOWESS Plot) developed by William S. Cleveland that shows a fitted line through the data that tends to eliminate the distortion that comes from deviant points.




0617
Regression Slope = -0.0724172 / year, Rho=-0.130057
Residual Trichloroethene
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
-3.5
-4
-4.5
-5
10/18/2001
10/18/2003
10/18/2005
10/18/2007
10/18/2009
10/18/2011
10/18/2013
10/18/2015
Time
Regression Line



Tests
A goodness-of-fit test was performed to test whether the residuals (deviations from the linear regression line described and displayed above) are normally distributed. The Shapiro-Wilk test was determined most appropriate for this VSP application because the size of the data set (n=39) is less than 50.

Normal Distribution Test of Residuals for Trichloroethene	
Shapiro-Wilk Test Statistic	0.95323
Shapiro-Wilk 5% Critical Value	0.939

The calculated test statistic does not exceed the 5% Shapiro-Wilk critical value, so we cannot reject the hypothesis that the residuals are normally distributed, or in other words the residuals appear to follow a normal distribution at the 5% level of significance. A Q-Q plot of residual values (displayed above) should be used to further assess the normality of the data.

Mann-Kendall Test
The Mann-Kendall Test for trend was performed on the data as outlined in the *Calculations to Determine Whether a Trend Exists* section above. The results are as follows:

Mann-Kendall Test			
Sum of Signs	S	-60	
MK Test Statistic	Z_{MK}	-0.713872	
Alpha	α	0.05	
Critical Value	$- Z_{1-\alpha}$	-1.64485	
Null Hypothesis	H_o	No trend exists	**Accept**
Alternative Hypothesis	H_a	A downward monotonic trend exists	Reject
Conclude with 95% confidence that a downward monotonic trend does not exist.			

References

Cleveland, William S., **Robust Locally Weighted Regression and Smoothing Scatterplots**, 1979, *Journal of the American Statistical Association*, Vol. 74, No. 368. p. 829-836.

Gilbert, R.O., 1987. **Statistical Methods for Environmental Pollution Monitoring**, Wiley & Sons, New York.

Helsel, D.R. 2005. **Nondetects and Data Analysis, Statistics for Censored Environmental Data**, Wiley & Sons, New York.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. **Techniques of trend analysis for monthly water quality data,** *Water Resources Research* 18(1):107-121.

Kendall, M.G. 1975. **Rank Correlation Methods,** 4th edition, Charles Griffin, London.

Mann, H.B. 1945. **Non-parametric tests against trend,** *Econometrica* 13:163-171.

This report was automatically produced* by Visual Sample Plan (VSP) software version 7.5.

This design was last modified 3/6/2017 3:48:00 PM.

Software and documentation available at http://vsp.pnnl.gov

Software copyright (c) 2017 Battelle Memorial Institute. All rights reserved.

* - The report contents may have been modified or reformatted by end-user of software.

Appendix C

2016 Groundwater Elevations

This page intentionally left blank

Table C-1. Phase I Groundwater Elevations

Location ID	Date	Elevation of Top of Casing	Depth from Top of Casing	Groundwater Elevation
0353	1/25/2016	745.33	2.25	743.08
	7/26/2016	745.33	5.60	739.73
0400	1/25/2016	705.11	23.22	681.89
	7/25/2016	705.11	29.58	675.53
0402	1/25/2016	704.02	22.08	681.94
	2/1/2016	704.02	22.74	681.28
	5/3/2016	704.02	26.12	677.90
	7/20/2016	704.02	28.60	675.42
	7/25/2016	704.02	28.20	675.82
	8/2/2016	704.02	28.48	675.54
	8/29/2016	704.02	27.68	676.34
	10/3/2016	704.02	25.60	678.42
	11/1/2016	704.02	24.75	679.27
0411	1/25/2016	836.57	17.08	819.49
	7/25/2016	836.57	27.93	808.64
0443	1/25/2016	858.78	30.32	828.46
	7/25/2016	858.78	Below top of pump	
0444	1/25/2016	773.00	22.44	750.56
	7/26/2016	773.00	25.37	747.63
0445	1/25/2016	743.43	13.83	729.60
	7/26/2016	743.43	15.41	728.02
P033	1/25/2016	705.83	23.97	681.86
	6/13/2016	705.83	30.99	674.84
	6/27/2016	705.83	30.99	674.84
	7/5/2016	705.83	30.83	675.00
	7/11/2016	705.83	30.70	675.13
	7/18/2016	705.83	30.66	675.17
	7/25/2016	705.83	30.38	675.45
	8/1/2016	705.83	30.58	675.25
	8/8/2016	705.83	30.74	675.09
	8/15/2016	705.83	30.79	675.04
	8/22/2016	705.83	30.51	675.32
	8/29/2016	705.83	29.81	676.02
	9/6/2016	705.83	29.77	676.06
	9/12/2016	705.83	29.73	676.10
	9/19/2016	705.83	29.45	676.38
	9/26/2016	705.83	28.26	677.57
	10/3/2016	705.83	27.56	678.27
	10/10/2016	705.83	27.21	678.62
	10/18/2016	705.83	27.23	678.60
	10/24/2016	705.83	26.83	679.00

Table C-1. Phase I Groundwater Elevations (continued)

Location ID	Date	Elevation of Top of Casing	Depth from Top of Casing	Groundwater Elevation
P033 (continued)	10/31/2016	705.83	26.65	679.18
	11/14/2016	705.83	26.81	679.02
	12/7/2016	705.83	26.94	678.89

Table C-2. Parcels 6, 7, and 8 Groundwater Elevations

Location ID	Date	Elevation of Top of Casing	Depth from Top of Casing	Groundwater Elevation
0118	1/25/2016	704.86	22.35	682.51
	4/25/2016	704.86	20.38	684.48
	7/28/2016	704.86	25.88	678.98
	10/24/2016	704.86	24.21	680.65
0124	1/26/2016	705.12	23.13	681.99
	4/26/2016	705.12	25.35	679.77
	7/28/2016	705.12	28.46	676.66
	10/24/2016	705.12	25.70	679.42
0126	1/26/2016	705.54	23.56	681.98
	4/26/2016	705.54	25.84	679.70
	7/28/2016	705.54	28.94	676.60
	10/24/2016	705.54	26.12	679.42
0138	1/25/2016	697.76	20.49	677.27
	4/26/2016	697.76	21.67	676.09
	7/28/2016	697.76	29.17	668.59
	10/26/2016	697.76	27.34	670.42
0315	1/25/2016	723.99	41.90	682.09
	4/26/2016	723.99	44.27	679.72
	7/27/2016	723.99	47.34	676.65
	10/26/2016	723.99	44.50	679.49
0346	1/25/2016	742.97	15.13	727.84
	4/25/2016	742.97	13.15	729.82
	7/26/2016	742.97	15.75	727.22
	10/25/2016	742.97	16.16	726.81
0347	1/25/2016	725.20	43.12	682.08
	4/26/2016	725.20	45.50	679.70
	7/27/2016	725.20	48.55	676.65
	10/26/2016	725.20	45.72	679.48
0379	1/25/2016	716.11	34.04	682.07
	2/1/2016	716.11	34.66	681.45
	4/26/2016	716.11	36.40	679.71
	5/2/2016	716.11	37.30	678.81
	7/26/2016	716.11	39.51	676.60
	8/1/2016	716.11	39.62	676.49
	8/29/2016	716.11	38.77	677.34
	10/3/2016	716.11	37.42	678.69
	10/25/2016	716.11	36.70	679.41
	10/31/2016	716.11	36.55	679.56
0386	1/26/2016	724.79	42.82	681.97
	4/25/2016	724.79	44.93	679.86
	7/27/2016	724.79	48.14	676.65
	10/25/2016	724.79	45.36	679.43

Table C-2. Parcels 6, 7, and 8 Groundwater Elevations (continued)

Location ID	Date	Elevation of Top of Casing	Depth from Top of Casing	Groundwater Elevation
0387	1/26/2016	720.89	38.96	681.93
	4/25/2016	720.89	41.05	679.84
	7/27/2016	720.89	44.29	676.60
	10/25/2016	720.89	41.50	679.39
0389	1/26/2016	724.65	42.70	681.95
	4/25/2016	724.65	44.70	679.95
	7/27/2016	724.65	47.97	676.68
	10/25/2016	724.65	45.22	679.43
0392	1/26/2016	720.84	38.77	682.07
	4/25/2016	720.84	40.90	679.94
	7/27/2016	720.84	44.11	676.73
	10/25/2016	720.84	41.30	679.54

Appendix D

2016 Groundwater and Seep Data Table

This page intentionally left blank

Table D-1. Phase I Groundwater and Seep Data

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0353	Dissolved oxygen	1/25/2016	8.32			QF	mg/L	F
		7/26/2016	2.42			QF	mg/L	F
0353	Oxidation reduction potential	1/25/2016	59			QF	mV	F
		7/26/2016	145.7			QF	mV	F
0353	Specific conductance	1/25/2016	1370			QF	µmhos/cm	F
		7/26/2016	1350			QF	µmhos/cm	F
0353	Tetrachloroethene	1/25/2016	0.16	0.16	U	QF	µg/L	F
		7/26/2016	0.16	0.16	U	QF	µg/L	F
0353	Trichloroethene	1/25/2016	0.16	0.16	U	QF	µg/L	F
		7/26/2016	0.16	0.16	U	QF	µg/L	F
0353	Turbidity	1/25/2016	37.9			QF	NTU	F
		7/26/2016	8.92			QF	NTU	F
0353	Vinyl chloride	1/25/2016	0.16	0.16	U	QF	µg/L	F
		7/26/2016	0.16	0.16	U	QF	µg/L	F
0353	*cis*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	QF	µg/L	F
		7/26/2016	0.16	0.16	U	QF	µg/L	F
0353	pH	1/25/2016	7.46			QF	s.u.	F
		7/26/2016	7.01			QF	s.u.	F
0353	*trans*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	QF	µg/L	F
		7/26/2016	0.16	0.16	U	QF	µg/L	F
0400	Dissolved oxygen	1/25/2016	1.57			F	mg/L	F
		7/25/2016	1.6			F	mg/L	F
0400	Oxidation reduction potential	1/25/2016	76.7			F	mV	F
		7/25/2016	33.3			F	mV	F
0400	Specific conductance	1/25/2016	1260			F	µmhos/cm	F
		7/25/2016	1350			F	µmhos/cm	F
0400	Tetrachloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		7/25/2016	0.16	0.16	U	F	µg/L	F
0400	Trichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		7/25/2016	0.16	0.16	U	F	µg/L	F

Table D-1. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0400	Turbidity	1/25/2016	90.9			F	NTU	F
		7/25/2016	177			F	NTU	F
0400	Vinyl chloride	1/25/2016	0.16	0.16	U	F	µg/L	F
		7/25/2016	0.16	0.16	U	F	µg/L	F
0400	*cis*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		7/25/2016	1.14	0.16		F	µg/L	F
0400	pH	1/25/2016	7			F	s.u.	F
		7/25/2016	7.04			F	s.u.	F
0400	*trans*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		7/25/2016	0.16	0.16	U	F	µg/L	F
0402	Dissolved oxygen	1/25/2016	4.23			F	mg/L	F
		2/1/2016	4.29			F	mg/L	F
		5/3/2016	1.93			F	mg/L	F
		6/30/2016	1.41				mg/L	F
		7/25/2016	1.53			F	mg/L	F
		8/2/2016	0.76			F	mg/L	F
		8/29/2016	1.49			F	mg/L	F
		10/3/2016	0.77			F	mg/L	F
		11/1/2016	0.38			F	mg/L	F
0402	Oxidation reduction potential	1/25/2016	147			F	mV	F
		2/1/2016	211.8			F	mV	F
		5/3/2016	158			F	mV	F
		6/30/2016	30.2				mV	F
		7/25/2016	59.5			F	mV	F
		8/2/2016	64.1			F	mV	F
		8/29/2016	91.4			F	mV	F
		10/3/2016	99.8			F	mV	F
		11/1/2016	102.4			F	mV	F

Table D-1. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0402	Specific conductance	1/25/2016	1360			F	µmhos/cm	F
		2/1/2016	1370			F	µmhos/cm	F
		5/3/2016	1120			F	µmhos/cm	F
		6/30/2016	1370				µmhos/cm	F
		7/25/2016	1410			F	µmhos/cm	F
		8/2/2016	1420			F	µmhos/cm	F
		8/29/2016	1460			F	µmhos/cm	F
		10/3/2016	1440			F	µmhos/cm	F
		11/1/2016	1430			F	µmhos/cm	F
0402	Tetrachloroethene	1/25/2016	0.27	0.16	J	F	µg/L	F
		2/1/2016	0.17	0.16	J	F	µg/L	F
		5/3/2016	0.16	0.16	U	F	µg/L	F
		6/30/2016	0.94	0.16	J		µg/L	F
		7/25/2016	1.22	0.16		F	µg/L	F
		8/2/2016	1.15	0.16		F	µg/L	F
		8/29/2016	1.33	0.16		F	µg/L	F
		10/3/2016	1.31	0.16		F	µg/L	F
		11/1/2016	1.68	0.16		F	µg/L	F
0402	Trichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		2/1/2016	0.16	0.16	U	F	µg/L	F
		5/3/2016	0.16	0.16	U	F	µg/L	F
		6/30/2016	0.91	0.16	J		µg/L	F
		7/25/2016	1.01	0.16		F	µg/L	F
		8/2/2016	0.94	0.16	J	F	µg/L	F
		8/29/2016	0.94	0.16	J	F	µg/L	F
		10/3/2016	0.92	0.16	J	F	µg/L	F
		11/1/2016	1.08	0.16		F	µg/L	F

Table D-1. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0402	Turbidity	1/25/2016	3.25			F	NTU	F
		2/1/2016	0.68			F	NTU	F
		5/3/2016	0.87			F	NTU	F
		6/30/2016	4.58				NTU	F
		7/25/2016	3.26			F	NTU	F
		8/2/2016	1.35			F	NTU	F
		8/29/2016	6.28			F	NTU	F
		10/3/2016	1.14			F	NTU	F
		11/1/2016	2.9			F	NTU	F
0402	Vinyl chloride	1/25/2016	0.16	0.16	U	F	µg/L	F
		2/1/2016	0.16	0.16	U	F	µg/L	F
		5/3/2016	0.16	0.16	U	F	µg/L	F
		6/30/2016	0.16	0.16	U		µg/L	F
		7/25/2016	0.31	0.16	J	UF	µg/L	F
		8/2/2016	0.31	0.16	J	F	µg/L	F
		8/29/2016	0.16	0.16	U	F	µg/L	F
		10/3/2016	0.16	0.16	U	F	µg/L	F
		11/1/2016	0.16	0.16	U	F	µg/L	F
0402	*cis*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		2/1/2016	0.16	0.16	U	F	µg/L	F
		5/3/2016	0.16	0.16	U	F	µg/L	F
		6/30/2016	3.29	0.16			µg/L	F
		7/25/2016	3.78	0.16		F	µg/L	F
		8/2/2016	3.92	0.16		F	µg/L	F
		8/29/2016	2.88	0.16		F	µg/L	F
		10/3/2016	2.37	0.16		F	µg/L	F
		11/1/2016	2.49	0.16		F	µg/L	F

Table D-1. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0402	pH	1/25/2016	6.81			F	s.u.	F
		2/1/2016	6.29			F	s.u.	F
		5/3/2016	6.09			F	s.u.	F
		6/30/2016	7.05				s.u.	F
		7/25/2016	6.68			F	s.u.	F
		8/2/2016	6.65			F	s.u.	F
		8/29/2016	6.51			F	s.u.	F
		10/3/2016	6.32			F	s.u.	F
		11/1/2016	6.85			F	s.u.	F
0402	*trans*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		2/1/2016	0.16	0.16	U	F	µg/L	F
		5/3/2016	0.16	0.16	U	F	µg/L	F
		6/30/2016	0.16	0.16	U		µg/L	F
		7/25/2016	0.16	0.16	U	F	µg/L	F
		8/2/2016	0.16	0.16	U	F	µg/L	F
		8/29/2016	0.16	0.16	U	F	µg/L	F
		10/3/2016	0.16	0.16	U	F	µg/L	F
		11/1/2016	0.16	0.16	U	F	µg/L	F
0411	Dissolved oxygen	1/25/2016	1.51			JF	mg/L	F
		7/25/2016	3.15			F	mg/L	F
0411	Oxidation reduction potential	1/25/2016	100.4			JF	mV	F
		7/25/2016	-32.9			F	mV	F
0411	Specific conductance	1/25/2016	1450			JF	µmhos/cm	F
		7/25/2016	1410			F	µmhos/cm	F
0411	Tetrachloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		1/25/2016	0.16	0.16	U	F	µg/L	D
		7/25/2016	0.16	0.16	U	F	µg/L	F
		7/25/2016	0.16	0.16	U	F	µg/L	D

Table D-1. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0411	Trichloroethene	1/25/2016	11.9	0.16		F	µg/L	F
		1/25/2016	11.7	0.16		F	µg/L	D
		7/25/2016	11.7	0.16		F	µg/L	F
		7/25/2016	11.8	0.16		F	µg/L	D
0411	Turbidity	1/25/2016	1.48			JF	NTU	F
		7/25/2016	83			F	NTU	F
0411	Vinyl chloride	1/25/2016	0.16	0.16	U	F	µg/L	F
		1/25/2016	0.16	0.16	U	F	µg/L	D
		7/25/2016	0.16	0.16	U	F	µg/L	F
		7/25/2016	0.16	0.16	U	F	µg/L	D
0411	*cis*-1,2-Dichloroethene	1/25/2016	0.65	0.16	J	F	µg/L	F
		1/25/2016	0.65	0.16	J	F	µg/L	D
		7/25/2016	0.94	0.16	J	F	µg/L	F
		7/25/2016	0.88	0.16	J	F	µg/L	D
0411	pH	1/25/2016	7.11			JF	s.u.	F
		7/25/2016	7.01			F	s.u.	F
0411	*trans*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		1/25/2016	0.16	0.16	U	F	µg/L	D
		7/25/2016	0.16	0.16	U	F	µg/L	F
		7/25/2016	0.16	0.16	U	F	µg/L	D
0443	Dissolved oxygen	1/25/2016	5.06			F	mg/L	F
		7/25/2016	9.43			QF	mg/L	F
0443	Oxidation reduction potential	1/25/2016	77.3			F	mV	F
		7/25/2016	136.9			QF	mV	F
0443	Specific conductance	1/25/2016	1290			F	µmhos/cm	F
		7/25/2016	1440			QF	µmhos/cm	F
0443	Tetrachloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		7/25/2016	0.16	0.16	U	QF	µg/L	F
0443	Trichloroethene	1/25/2016	7.53	0.16		F	µg/L	F
		7/25/2016	7.27	0.16		QF	µg/L	F

Table D-1. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0443	Turbidity	1/25/2016	3.98			F	NTU	F
		7/25/2016	47.3			QF	NTU	F
0443	Vinyl chloride	1/25/2016	0.16	0.16	U	F	µg/L	F
		7/25/2016	0.16	0.16	U	QF	µg/L	F
0443	*cis*-1,2-Dichloroethene	1/25/2016	0.23	0.16	J	F	µg/L	F
		7/25/2016	0.33	0.16	J	QF	µg/L	F
0443	pH	1/25/2016	7.12			F	s.u.	F
		7/25/2016	6.98			QF	s.u.	F
0443	*trans*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		7/25/2016	0.16	0.16	U	QF	µg/L	F
0444	Dissolved oxygen	1/25/2016	4.97			QF	mg/L	F
		7/26/2016	2.17			QF	mg/L	F
0444	Oxidation reduction potential	1/25/2016	210.7			QF	mV	F
		7/26/2016	31.3			QF	mV	F
0444	Specific conductance	1/25/2016	1280			QF	µmhos/cm	F
		7/26/2016	1180			QF	µmhos/cm	F
0444	Tetrachloroethene	1/25/2016	0.16	0.16	U	QF	µg/L	F
		7/26/2016	0.16	0.16	U	QF	µg/L	F
0444	Trichloroethene	1/25/2016	0.16	0.16	U	QF	µg/L	F
		7/26/2016	0.16	0.16	U	QF	µg/L	F
0444	Turbidity	1/25/2016	5.44			QF	NTU	F
		7/26/2016	3.83			QF	NTU	F
0444	Vinyl chloride	1/25/2016	0.16	0.16	U	QF	µg/L	F
		7/26/2016	0.16	0.16	U	QF	µg/L	F
0444	*cis*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	QF	µg/L	F
		7/26/2016	0.16	0.16	U	QF	µg/L	F
0444	pH	1/25/2016	7.15			QF	s.u.	F
		7/26/2016	7.24			QF	s.u.	F
0444	*trans*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	QF	µg/L	F
		7/26/2016	0.16	0.16	U	QF	µg/L	F

Table D-1. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0445	Dissolved oxygen	1/25/2016	5.15			QF	mg/L	F
		7/26/2016	6.5			QF	mg/L	F
0445	Oxidation reduction potential	1/25/2016	-71.9			QF	mV	F
		7/26/2016	-91.9			QF	mV	F
0445	Specific conductance	1/25/2016	23,830			QF	µmhos/cm	F
		7/26/2016	17,370			QF	µmhos/cm	F
0445	Tetrachloroethene	1/25/2016	0.16	0.16	U	QF	µg/L	F
		7/26/2016	0.16	0.16	U	QF	µg/L	F
0445	Trichloroethene	1/25/2016	0.16	0.16	U	QF	µg/L	F
		7/26/2016	0.16	0.16	U	QF	µg/L	F
0445	Turbidity	1/25/2016	3.61			QF	NTU	F
0445	Vinyl chloride	1/25/2016	0.16	0.16	U	QF	µg/L	F
		7/26/2016	0.16	0.16	U	QF	µg/L	F
0445	*cis*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	QF	µg/L	F
		7/26/2016	0.16	0.16	U	QF	µg/L	F
0445	pH	1/25/2016	6.96			QF	s.u.	F
		7/26/2016	7.1			QF	s.u.	F
0445	*trans*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	QF	µg/L	F
		7/26/2016	0.16	0.16	U	QF	µg/L	F
P033	Dissolved oxygen	1/25/2016	3.29			F	mg/L	F
		7/25/2016	6.07			F	mg/L	F
P033	Oxidation reduction potential	1/25/2016	131.1			F	mV	F
		7/25/2016	76.1			F	mV	F
P033	Specific conductance	1/25/2016	1640			F	µmhos/cm	F
		7/25/2016	1530			F	µmhos/cm	F
P033	Tetrachloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		7/25/2016	0.16	0.16	U	F	µg/L	F
P033	Trichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		7/25/2016	0.16	0.16	U	F	µg/L	F
P033	Turbidity	1/25/2016	1.02			F	NTU	F
		7/25/2016	14.3			F	NTU	F

Table D-1. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
P033	Vinyl chloride	1/25/2016	0.16	0.16	U	F	µg/L	F
		7/25/2016	0.16	0.16	U	F	µg/L	F
P033	*cis*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		7/25/2016	0.16	0.16	U	F	µg/L	F
P033	pH	1/25/2016	6.9			F	s.u.	F
		7/25/2016	7.03			F	s.u.	F
P033	*trans*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		7/25/2016	0.16	0.16	U	F	µg/L	F
0617	Dissolved oxygen	1/25/2016	5.83				mg/L	F
		7/25/2016	5.76				mg/L	F
0617	Oxidation reduction potential	1/25/2016	87.5				mV	F
		7/25/2016	30.3				mV	F
0617	Specific conductance	1/25/2016	1670				µmhos/cm	F
		7/25/2016	1750				µmhos/cm	F
0617	Tetrachloroethene	1/25/2016	0.16	0.16	U		µg/L	F
		7/25/2016	0.16	0.16	U		µg/L	F
0617	Trichloroethene	1/25/2016	9.09	0.16			µg/L	F
		7/25/2016	3.53	0.16			µg/L	F
0617	Turbidity	1/25/2016	246				NTU	F
		7/25/2016	315				NTU	F
0617	Vinyl chloride	1/25/2016	0.16	0.16	U		µg/L	F
		7/25/2016	0.16	0.16	U		µg/L	F
0617	*cis*-1,2-Dichloroethene	1/25/2016	1.8	0.16			µg/L	F
		7/25/2016	1.08	0.16			µg/L	F
0617	pH	1/25/2016	7.74				s.u.	F
		7/25/2016	7.36				s.u.	F
0617	*trans*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U		µg/L	F
		7/25/2016	0.16	0.16	U		µg/L	F

Abbreviations:

D = analyte determined in diluted sample
F = low flow sampling method used
J = estimated value
µg/L = micrograms per liter
µmhos/cm = micromhos per centimeter
mg/L = milligrams per liter
mV = millivolts
pCi/L = picocuries per liter
Q = quantitative result due to sampling technique
s.u. = standard unit
U = analytical result below detection limit

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0118	Dissolved oxygen	1/25/2016	7.67			F	mg/L	F
		4/25/2016	5.87			F	mg/L	F
		7/28/2016	6.5			F	mg/L	F
		10/24/2016	5.69			F	mg/L	F
0118	Oxidation reduction potential	1/25/2016	117.6			F	mV	F
		4/25/2016	69.8			F	mV	F
		7/28/2016	173.7			F	mV	F
		10/24/2016	181.7			F	mV	F
0118	Specific conductance	1/25/2016	1220			F	µmhos/cm	F
		4/25/2016	1230			F	µmhos/cm	F
		7/28/2016	1220			F	µmhos/cm	F
		10/24/2016	1210			F	µmhos/cm	F
0118	Tetrachloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/28/2016	0.16	0.16	U	F	µg/L	F
		10/24/2016	0.16	0.16	U	F	µg/L	F
0118	Trichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/28/2016	0.16	0.16	U	F	µg/L	F
		10/24/2016	0.16	0.16	U	F	µg/L	F
0118	Tritium	1/25/2016	199	306	U	F	pCi/L	F
		7/28/2016	-142	345	U	F	pCi/L	F
0118	Turbidity	1/25/2016	17.4			F	NTU	F
		4/25/2016	15.2			F	NTU	F
		7/28/2016	12.9			F	NTU	F
		10/24/2016	10.1			F	NTU	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0118	Vinyl chloride	1/25/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/28/2016	0.16	0.16	U	F	µg/L	F
		10/24/2016	0.16	0.16	U	F	µg/L	F
0118	*cis*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/28/2016	0.16	0.16	U	F	µg/L	F
		10/24/2016	0.16	0.16	U	F	µg/L	F
0118	pH	1/25/2016	7.12			F	s.u.	F
		4/25/2016	6.94			F	s.u.	F
		7/28/2016	7.11			F	s.u.	F
		10/24/2016	6.97			F	s.u.	F
0118	*trans*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/28/2016	0.16	0.16	U	F	µg/L	F
		10/24/2016	0.16	0.16	U	F	µg/L	F
0124	Dissolved oxygen	1/26/2016	3.21			F	mg/L	F
		4/26/2016	3.92			F	mg/L	F
		7/28/2016	1.19			F	mg/L	F
		10/24/2016	1.66			F	mg/L	F
0124	Oxidation reduction potential	1/26/2016	163.6			F	mV	F
		4/26/2016	72.2			F	mV	F
		7/28/2016	132.9			F	mV	F
		10/24/2016	118.7			F	mV	F
0124	Specific conductance	1/26/2016	1170			F	µmhos/cm	F
		4/26/2016	1150			F	µmhos/cm	F
		7/28/2016	1300			F	µmhos/cm	F
		10/24/2016	1380			F	µmhos/cm	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0124	Tetrachloroethene	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/26/2016	0.19	0.16	JQ	JF	µg/L	F
		7/28/2016	0.16	0.16	U	F	µg/L	F
		10/24/2016	0.16	0.16	U	F	µg/L	F
0124	Trichloroethene	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/26/2016	0.2	0.16	JQ	JF	µg/L	F
		7/28/2016	0.16	0.16	U	F	µg/L	F
		10/24/2016	0.16	0.16	U	F	µg/L	F
0124	Turbidity	1/26/2016	3.2			F	NTU	F
		4/26/2016	2.11			F	NTU	F
		7/28/2016	1.74			F	NTU	F
		10/24/2016	7.34			F	NTU	F
0124	Vinyl chloride	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		7/28/2016	0.16	0.16	U	F	µg/L	F
		10/24/2016	0.16	0.16	U	F	µg/L	F
0124	*cis*-1,2-Dichloroethene	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		7/28/2016	0.16	0.16	U	F	µg/L	F
		10/24/2016	0.16	0.16	U	F	µg/L	F
0124	pH	1/26/2016	6.98			F	s.u.	F
		4/26/2016	6.91			F	s.u.	F
		7/28/2016	6.88			F	s.u.	F
		10/24/2016	6.86			F	s.u.	F
0124	*trans*-1,2-Dichloroethene	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		7/28/2016	0.16	0.16	U	F	µg/L	F
		10/24/2016	0.16	0.16	U	F	µg/L	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0126	Dissolved oxygen	1/26/2016	0.51			F	mg/L	F
		4/26/2016	1.33			F	mg/L	F
		7/28/2016	2.23			F	mg/L	F
		10/24/2016	0.66			F	mg/L	F
0126	Oxidation reduction potential	1/26/2016	162.8			F	mV	F
		4/26/2016	52.2			F	mV	F
		7/28/2016	143.7			F	mV	F
		10/24/2016	118.7			F	mV	F
0126	Specific conductance	1/26/2016	1250			F	µmhos/cm	F
		4/26/2016	1230			F	µmhos/cm	F
		7/28/2016	1220			F	µmhos/cm	F
		10/24/2016	1230			F	µmhos/cm	F
0126	Tetrachloroethene	1/26/2016	0.77	0.16	J	F	µg/L	F
		4/26/2016	0.82	0.16	JQ	JF	µg/L	F
		7/28/2016	0.97	0.16	J	F	µg/L	F
		10/24/2016	1.04	0.16		F	µg/L	F
0126	Trichloroethene	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		7/28/2016	0.16	0.16	U	F	µg/L	F
		10/24/2016	0.16	0.16	U	F	µg/L	F
0126	Turbidity	1/26/2016	0.89			F	NTU	F
		4/26/2016	0.86			F	NTU	F
		7/28/2016	0.43			F	NTU	F
		10/24/2016	2.08			F	NTU	F
0126	Vinyl chloride	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		7/28/2016	0.16	0.16	U	F	µg/L	F
		10/24/2016	0.16	0.16	U	F	µg/L	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0126	*cis*-1,2-Dichloroethene	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		7/28/2016	0.16	0.16	U	F	µg/L	F
		10/24/2016	0.16	0.16	U	F	µg/L	F
0126	pH	1/26/2016	7.04			F	s.u.	F
		4/26/2016	7.01			F	s.u.	F
		7/28/2016	6.87			F	s.u.	F
		10/24/2016	6.83			F	s.u.	F
0126	*trans*-1,2-Dichloroethene	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		7/28/2016	0.16	0.16	U	F	µg/L	F
		10/24/2016	0.16	0.16	U	F	µg/L	F
0138	Dissolved oxygen	1/25/2016	3.83			F	mg/L	F
		4/26/2016	3.74			F	mg/L	F
		7/28/2016	3.69			F	mg/L	F
		10/26/2016	4.14			F	mg/L	F
0138	Oxidation reduction potential	1/25/2016	83.9			F	mV	F
		4/26/2016	52			F	mV	F
		7/28/2016	64.8			F	mV	F
		10/26/2016	94.1			F	mV	F
0138	Specific conductance	1/25/2016	1260			F	µmhos/cm	F
		4/26/2016	1260			F	µmhos/cm	F
		7/28/2016	1240			F	µmhos/cm	F
		10/26/2016	1220			F	µmhos/cm	F
0138	Tetrachloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		7/28/2016	0.16	0.16	U	F	µg/L	F
		10/26/2016	0.16	0.16	U	F	µg/L	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0138 (continued)	Trichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		7/28/2016	0.16	0.16	U	F	µg/L	F
		10/26/2016	0.16	0.16	U	F	µg/L	F
0138	Tritium	1/25/2016	361	358		UF	pCi/L	F
		7/28/2016	429	347		JF	pCi/L	F
0138	Turbidity	1/25/2016	145			F	NTU	F
		4/26/2016	161			F	NTU	F
		7/28/2016	122			F	NTU	F
		10/26/2016	104			F	NTU	F
0138	Vinyl chloride	1/25/2016	0.16	0.16	U	F	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		7/28/2016	0.16	0.16	U	F	µg/L	F
		10/26/2016	0.16	0.16	U	F	µg/L	F
0138	*cis*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		7/28/2016	0.16	0.16	U	F	µg/L	F
		10/26/2016	0.16	0.16	U	F	µg/L	F
0138	pH	1/25/2016	6.94			F	s.u.	F
		4/26/2016	7.15			F	s.u.	F
		7/28/2016	7.09			F	s.u.	F
		10/26/2016	6.88			F	s.u.	F
0138	*trans*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		7/28/2016	0.16	0.16	U	F	µg/L	F
		10/26/2016	0.16	0.16	U	F	µg/L	F
0315	Dissolved oxygen	1/25/2016	1.91			F	mg/L	F
		4/26/2016	1.46			F	mg/L	F
		7/27/2016	1.3			F	mg/L	F
		10/26/2016	0.58			F	mg/L	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0315	Oxidation reduction potential	1/25/2016	11.2			F	mV	F
		4/26/2016	-17.7			F	mV	F
		7/27/2016	-13.7			F	mV	F
		10/26/2016	-19.2			F	mV	F
0315	Specific conductance	1/25/2016	1620			F	µmhos/cm	F
		4/26/2016	1680			F	µmhos/cm	F
		7/27/2016	1690			F	µmhos/cm	F
		10/26/2016	1670			F	µmhos/cm	F
0315	Tetrachloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	D
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/26/2016	0.16	0.16	U	F	µg/L	F
0315	Trichloroethene	1/25/2016	3.83	0.16		F	µg/L	F
		4/26/2016	5.67	0.16	Q	JF	µg/L	F
		4/26/2016	5.43	0.16	Q	JF	µg/L	D
		7/27/2016	6.81	0.16		F	µg/L	F
		10/26/2016	7.53	0.16		F	µg/L	F
0315	Turbidity	1/25/2016	223			F	NTU	F
		4/26/2016	65.5			F	NTU	F
		7/27/2016	31.4			F	NTU	F
		10/26/2016	46.2			F	NTU	F
0315	Vinyl chloride	1/25/2016	0.16	0.16	U	F	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	D
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/26/2016	0.16	0.16	U	F	µg/L	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0315	*cis*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	D
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/26/2016	0.16	0.16	U	F	µg/L	F
0315	pH	1/25/2016	7.1			F	s.u.	F
		4/26/2016	7.1			F	s.u.	F
		7/27/2016	7.15			F	s.u.	F
		10/26/2016	6.97			F	s.u.	F
0315	*trans*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	D
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/26/2016	0.16	0.16	U	F	µg/L	F
0346	Dissolved oxygen	1/25/2016	2.11			F	mg/L	F
		4/25/2016	3.69			F	mg/L	F
		7/26/2016	2.03			F	mg/L	F
		10/25/2016	5.42			F	mg/L	F
0346	Oxidation reduction potential	1/25/2016	-5			F	mV	F
		4/25/2016	-31.9			F	mV	F
		7/26/2016	3			F	mV	F
		10/25/2016	160.4			F	mV	F
0346	Specific conductance	1/25/2016	1580			F	µmhos/cm	F
		4/25/2016	1630			F	µmhos/cm	F
		7/26/2016	1730			F	µmhos/cm	F
		10/25/2016	403			F	µmhos/cm	F
0346	Tetrachloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/26/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0346	Trichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/26/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
0346	Tritium	1/25/2016	860	304		JF	pCi/L	F
		7/26/2016	566	331		JF	pCi/L	F
0346	Turbidity	1/25/2016	14.1			F	NTU	F
		4/25/2016	108			F	NTU	F
		7/26/2016	35.3			F	NTU	F
		10/25/2016	8.96			F	NTU	F
0346	Vinyl chloride	1/25/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/26/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
0346	*cis*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/26/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
0346	pH	1/25/2016	7.16			F	s.u.	F
		4/25/2016	7.22			F	s.u.	F
		7/26/2016	7.05			F	s.u.	F
		10/25/2016	7.24			F	s.u.	F
0346	*trans*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/26/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
0347	Dissolved oxygen	1/25/2016	0.72			F	mg/L	F
		4/26/2016	1.4			F	mg/L	F
		7/27/2016	1.36			F	mg/L	F
		10/26/2016	2.11			F	mg/L	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0347	Oxidation reduction potential	1/25/2016	-47			F	mV	F
		4/26/2016	-61.5			F	mV	F
		7/27/2016	-56.2			F	mV	F
		10/26/2016	-57.5			F	mV	F
0347	Specific conductance	1/25/2016	1620			F	µmhos/cm	F
		4/26/2016	1630			F	µmhos/cm	F
		7/27/2016	1660			F	µmhos/cm	F
		10/26/2016	1590			F	µmhos/cm	F
0347	Tetrachloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		1/25/2016	0.16	0.16	U	F	µg/L	D
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/26/2016	0.16	0.16	U	F	µg/L	F
		10/26/2016	0.16	0.16	U	F	µg/L	D
0347	Trichloroethene	1/25/2016	16.6	0.16		F	µg/L	F
		1/25/2016	16.1	0.16		F	µg/L	D
		4/26/2016	13.9	0.16	Q	JF	µg/L	F
		7/27/2016	18.2	0.16		F	µg/L	F
		10/26/2016	24.6	0.16		F	µg/L	F
		10/26/2016	24.6	0.16		F	µg/L	D
0347	Tritium	1/25/2016	1430	303		F	pCi/L	F
		1/25/2016	1670	297		F	pCi/L	D
		7/27/2016	1630	344		F	pCi/L	F
0347	Turbidity	1/25/2016	19.9			F	NTU	F
		4/26/2016	22.2			F	NTU	F
		7/27/2016	17.3			F	NTU	F
		10/26/2016	14.5			F	NTU	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0347	Vinyl chloride	1/25/2016	0.16	0.16	U	F	µg/L	F
		1/25/2016	0.16	0.16	U	F	µg/L	D
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/26/2016	0.16	0.16	U	F	µg/L	F
		10/26/2016	0.16	0.16	U	F	µg/L	D
0347	*cis*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		1/25/2016	0.16	0.16	U	F	µg/L	D
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/26/2016	0.16	0.16	U	F	µg/L	F
		10/26/2016	0.16	0.16	U	F	µg/L	D
0347	pH	1/25/2016	6.92			F	s.u.	F
		4/26/2016	6.96			F	s.u.	F
		7/27/2016	7.05			F	s.u.	F
		10/26/2016	6.75			F	s.u.	F
0347	*trans*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		1/25/2016	0.16	0.16	U	F	µg/L	D
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/26/2016	0.16	0.16	U	F	µg/L	F
		10/26/2016	0.16	0.16	U	F	µg/L	D
0379	Dissolved oxygen	1/25/2016	3.33			F	mg/L	F
		2/1/2016	2.17			F	mg/L	F
		4/26/2016	2.62			F	mg/L	F
		5/2/2016	2.69			F	mg/L	F
		7/26/2016	2.77			F	mg/L	F
		8/1/2016	2.4			F	mg/L	F
		8/29/2016	1.9			F	mg/L	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0379 (continued)	Dissolved oxygen	10/3/2016	1.4			F	mg/L	F
		10/25/2016	0.54			F	mg/L	F
		10/31/2016	0.58				mg/L	F
0379	Oxidation reduction potential	1/25/2016	35			F	mV	F
		2/1/2016	40.8			F	mV	F
		4/26/2016	21.8			F	mV	F
		5/2/2016	38			F	mV	F
		7/26/2016	85.4			F	mV	F
		8/1/2016	65.1			F	mV	F
		8/29/2016	-25.1			F	mV	F
		10/3/2016	-30.9			F	mV	F
		10/25/2016	11.7			F	mV	F
		10/31/2016	40.2				mV	F
0379	Specific conductance	1/25/2016	1680			F	µmhos/cm	F
		2/1/2016	1650			F	µmhos/cm	F
		4/26/2016	1520			F	µmhos/cm	F
		5/2/2016	1480			F	µmhos/cm	F
		7/26/2016	1650			F	µmhos/cm	F
		8/1/2016	1670			F	µmhos/cm	F
		8/29/2016	1810			F	µmhos/cm	F
		10/3/2016	1850			F	µmhos/cm	F
		10/25/2016	1850			F	µmhos/cm	F
		10/31/2016	1820				µmhos/cm	F
0379	Tetrachloroethene	1/25/2016	0.31	0.16	J	F	µg/L	F
		2/1/2016	0.29	0.16	J	F	µg/L	F
		4/26/2016	0.31	0.16	JQ	JF	µg/L	F
		5/2/2016	0.26	0.16	J	F	µg/L	F
		7/26/2016	0.39	0.16	J	F	µg/L	F
		7/26/2016	0.37	0.16	J	F	µg/L	D

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0379 (continued)	Tetrachloroethene	8/1/2016	0.32	0.16	J	F	µg/L	F
		8/29/2016	0.49	0.16	J	F	µg/L	F
		10/3/2016	0.46	0.16	J	F	µg/L	F
		10/25/2016	0.5	0.16	J	F	µg/L	F
		10/31/2016	0.52	0.16	J		µg/L	F
0379	Trichloroethene	1/25/2016	1.64	0.16		F	µg/L	F
		2/1/2016	1.65	0.16		F	µg/L	F
		4/26/2016	1.52	0.16	Q	JF	µg/L	F
		5/2/2016	1.27	0.16		F	µg/L	F
		7/26/2016	1.71	0.16		F	µg/L	F
		7/26/2016	1.7	0.16		F	µg/L	D
		8/1/2016	1.49	0.16		F	µg/L	F
		8/29/2016	1.72	0.16		F	µg/L	F
		10/3/2016	1.6	0.16		F	µg/L	F
		10/25/2016	1.73	0.16		F	µg/L	F
		10/31/2016	1.8	0.16			µg/L	F
0379	Tritium	1/25/2016	1230	308		F	pCi/L	F
		7/26/2016	590	336		JF	pCi/L	F
		7/26/2016	795	344		JF	pCi/L	D
0379	Turbidity	1/25/2016	26.7			F	NTU	F
		2/1/2016	9.98			F	NTU	F
		4/26/2016	34.6			F	NTU	F
		5/2/2016	85.8			F	NTU	F
		7/26/2016	9.98			F	NTU	F
		8/1/2016	14.2			F	NTU	F
		8/29/2016	150			F	NTU	F
		10/3/2016	35.2			F	NTU	F
		10/25/2016	41			F	NTU	F
		10/31/2016	13				NTU	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0379	Vinyl chloride	1/25/2016	0.16	0.16	U	F	µg/L	F
		2/1/2016	0.16	0.16	U	F	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		5/2/2016	0.16	0.16	U	F	µg/L	F
		7/26/2016	0.16	0.16	U	F	µg/L	F
		7/26/2016	0.16	0.16	U	F	µg/L	D
		8/1/2016	0.16	0.16	U	F	µg/L	F
		8/29/2016	0.16	0.16	U	F	µg/L	F
		10/3/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
		10/31/2016	0.16	0.16	U		µg/L	F
0379	*cis*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		2/1/2016	0.16	0.16	U	F	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		5/2/2016	0.16	0.16	U	F	µg/L	F
		7/26/2016	0.16	0.16	U	F	µg/L	F
		7/26/2016	0.16	0.16	U	F	µg/L	D
		8/1/2016	0.16	0.16	U	F	µg/L	F
		8/29/2016	0.16	0.16	U	F	µg/L	F
		10/3/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
		10/31/2016	0.16	0.16	U		µg/L	F
0379	pH	1/25/2016	7.09			F	s.u.	F
		2/1/2016	7.08			F	s.u.	F
		4/26/2016	7.15			F	s.u.	F
		5/2/2016	6.88			F	s.u.	F
		7/26/2016	7.07			F	s.u.	F
		8/1/2016	6.77			F	s.u.	F
		8/29/2016	6.97			F	s.u.	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0379 (continued)	pH	10/3/2016	7.18			F	s.u.	F
		10/25/2016	6.94			F	s.u.	F
		10/31/2016	7.14				s.u.	F
0379	*trans*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U	F	µg/L	F
		2/1/2016	0.16	0.16	U	F	µg/L	F
		4/26/2016	0.16	0.16	QU	JF	µg/L	F
		5/2/2016	0.16	0.16	U	F	µg/L	F
		7/26/2016	0.16	0.16	U	F	µg/L	F
		7/26/2016	0.16	0.16	U	F	µg/L	D
		8/1/2016	0.16	0.16	U	F	µg/L	F
		8/29/2016	0.16	0.16	U	F	µg/L	F
		10/3/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
		10/31/2016	0.16	0.16	U		µg/L	F
0386	Dissolved oxygen	1/26/2016	3.86			F	mg/L	F
		4/25/2016	4.4			F	mg/L	F
		7/27/2016	4.04			F	mg/L	F
		10/25/2016	3.87			F	mg/L	F
0386	Oxidation reduction potential	1/26/2016	168.5			F	mV	F
		4/25/2016	41.5			F	mV	F
		7/27/2016	90.8			F	mV	F
		10/25/2016	151.1			F	mV	F
0386	Specific conductance	1/26/2016	1190			F	µmhos/cm	F
		4/25/2016	1350			F	µmhos/cm	F
		7/27/2016	1240			F	µmhos/cm	F
		10/25/2016	1210			F	µmhos/cm	F
0386	Tetrachloroethene	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0386	Trichloroethene	1/26/2016	2.42	0.16		F	µg/L	F
		4/25/2016	2.56	0.16	Q	JF	µg/L	F
		7/27/2016	2.38	0.16		F	µg/L	F
		10/25/2016	2.54	0.16		F	µg/L	F
0386	Turbidity	1/26/2016	6.33			F	NTU	F
		4/25/2016	15.3			F	NTU	F
		7/27/2016	17.4			F	NTU	F
		10/25/2016	15.7			F	NTU	F
0386	Vinyl chloride	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
0386	*cis*-1,2-Dichloroethene	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
0386	pH	1/26/2016	6.81			F	s.u.	F
		4/25/2016	6.98			F	s.u.	F
		7/27/2016	6.95			F	s.u.	F
		10/25/2016	6.77			F	s.u.	F
0386	*trans*-1,2-Dichloroethene	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
0387	Dissolved oxygen	1/26/2016	0.72			F	mg/L	F
		4/25/2016	1.65			F	mg/L	F
		7/27/2016	1.78			F	mg/L	F
		10/25/2016	0.92			F	mg/L	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0387	Oxidation reduction potential	1/26/2016	175.5			F	mV	F
		4/25/2016	49.2			F	mV	F
		7/27/2016	52			F	mV	F
		10/25/2016	131.3			F	mV	F
0387	Specific conductance	1/26/2016	1260			F	µmhos/cm	F
		4/25/2016	1410			F	µmhos/cm	F
		7/27/2016	1350			F	µmhos/cm	F
		10/25/2016	1280			F	µmhos/cm	F
0387	Tetrachloroethene	1/26/2016	0.2	0.16	J	F	µg/L	F
		4/25/2016	0.23	0.16	JQ	JF	µg/L	F
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
0387	Trichloroethene	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
0387	Turbidity	1/26/2016	1.22			F	NTU	F
		4/25/2016	3.35			F	NTU	F
		7/27/2016	0.44			F	NTU	F
		10/25/2016	1.95			F	NTU	F
0387	Vinyl chloride	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
0387	*cis*-1,2-Dichloroethene	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0387	pH	1/26/2016	6.98			F	s.u.	F
		4/25/2016	6.99			F	s.u.	F
		7/27/2016	7.19			F	s.u.	F
		10/25/2016	6.9			F	s.u.	F
0387	*trans*-1,2-Dichloroethene	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
0389	Dissolved oxygen	1/26/2016	3.73			F	mg/L	F
		4/25/2016	2.64			F	mg/L	F
		7/27/2016	2.34			F	mg/L	F
		10/25/2016	2.1			F	mg/L	F
0389	Oxidation reduction potential	1/26/2016	171.7			F	mV	F
		4/25/2016	70.3			F	mV	F
		7/27/2016	91			F	mV	F
		10/25/2016	149.5			F	mV	F
0389	Specific conductance	1/26/2016	1330			F	µmhos/cm	F
		4/25/2016	1340			F	µmhos/cm	F
		7/27/2016	1210			F	µmhos/cm	F
		10/25/2016	1210			F	µmhos/cm	F
0389	Tetrachloroethene	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
0389	Trichloroethene	1/26/2016	0.35	0.16	J	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/27/2016	0.8	0.16	J	F	µg/L	F
		10/25/2016	0.74	0.16	J	F	µg/L	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0389	Turbidity	1/26/2016	21.8			F	NTU	F
		4/25/2016	17.8			F	NTU	F
		7/27/2016	19			F	NTU	F
		10/25/2016	49.8			F	NTU	F
0389	Vinyl chloride	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
0389	*cis*-1,2-Dichloroethene	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
0389	pH	1/26/2016	6.86			F	s.u.	F
		4/25/2016	6.92			F	s.u.	F
		7/27/2016	6.99			F	s.u.	F
		10/25/2016	6.86			F	s.u.	F
0389	*trans*-1,2-Dichloroethene	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
0392	Dissolved oxygen	1/26/2016	3.63			F	mg/L	F
		4/25/2016	4.92			F	mg/L	F
		7/27/2016	2.14			F	mg/L	F
		10/25/2016	4.56			F	mg/L	F
0392	Oxidation reduction potential	1/26/2016	216.8			F	mV	F
		4/25/2016	75.6			F	mV	F
		7/27/2016	145.6			F	mV	F
		10/25/2016	122.7			F	mV	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0392	Specific conductance	1/26/2016	1300			F	µmhos/cm	F
		4/25/2016	1230			F	µmhos/cm	F
		7/27/2016	1220			F	µmhos/cm	F
		10/25/2016	1170			F	µmhos/cm	F
0392	Tetrachloroethene	1/26/2016	0.25	0.16	J	F	µg/L	F
		4/25/2016	0.22	0.16	JQ	JF	µg/L	F
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
0392	Trichloroethene	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
0392	Turbidity	1/26/2016	5.39			F	NTU	F
		4/25/2016	1.77			F	NTU	F
		7/27/2016	5.2			F	NTU	F
		10/25/2016	20.5			F	NTU	F
0392	Vinyl chloride	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
0392	*cis*-1,2-Dichloroethene	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
0392	pH	1/26/2016	6.83			F	s.u.	F
		4/25/2016	6.87			F	s.u.	F
		7/27/2016	6.61			F	s.u.	F
		10/25/2016	6.78			F	s.u.	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0392	*trans*-1,2-Dichloroethene	1/26/2016	0.16	0.16	U	F	µg/L	F
		4/25/2016	0.16	0.16	QU	JF	µg/L	F
		7/27/2016	0.16	0.16	U	F	µg/L	F
		10/25/2016	0.16	0.16	U	F	µg/L	F
0601	Dissolved oxygen	1/25/2016	9.2				mg/L	F
		4/25/2016	9.11				mg/L	F
		7/25/2016	8.8				mg/L	F
		10/24/2016	3.72				mg/L	F
0601	Oxidation reduction potential	1/25/2016	226				mV	F
		4/25/2016	55				mV	F
		7/25/2016	94				mV	F
		10/24/2016	195.6				mV	F
0601	Specific conductance	1/25/2016	1500				µmhos/cm	F
		4/25/2016	1470				µmhos/cm	F
		7/25/2016	1210				µmhos/cm	F
		10/24/2016	1270				µmhos/cm	F
0601	Tetrachloroethene	1/25/2016	8.73	0.16			µg/L	F
		4/25/2016	12.2	0.16	Q	J	µg/L	F
		7/25/2016	5.54	0.16			µg/L	F
		10/24/2016	19.1	0.16			µg/L	F
0601	Trichloroethene	1/25/2016	6.52	0.16			µg/L	F
		4/25/2016	7.3	0.16	Q	J	µg/L	F
		7/25/2016	5.28	0.16			µg/L	F
		10/24/2016	7.85	0.16			µg/L	F
0601	Tritium	1/25/2016	23,500	333			pCi/L	F
		7/25/2016	29,200	342			pCi/L	F
0601	Turbidity	1/25/2016	10.1				NTU	F
		4/25/2016	9.08				NTU	F
		7/25/2016	998				NTU	F
		10/24/2016	2.96				NTU	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0601	Vinyl chloride	1/25/2016	0.16	0.16	U		µg/L	F
		4/25/2016	0.16	0.16	QU	J	µg/L	F
		7/25/2016	0.16	0.16	U		µg/L	F
		10/24/2016	0.16	0.16	U		µg/L	F
0601	*cis*-1,2-Dichloroethene	1/25/2016	0.48	0.16	J		µg/L	F
		4/25/2016	0.59	0.16	JQ	J	µg/L	F
		7/25/2016	0.84	0.16	J		µg/L	F
		10/24/2016	1.16	0.16			µg/L	F
0601	pH	1/25/2016	7.17				s.u.	F
		4/25/2016	7.28				s.u.	F
		7/25/2016	7.37				s.u.	F
		10/24/2016	7.18				s.u.	F
0601	*trans*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U		µg/L	F
		4/25/2016	0.16	0.16	QU	J	µg/L	F
		7/25/2016	0.16	0.16	U		µg/L	F
		10/24/2016	0.16	0.16	U		µg/L	F
0602	Dissolved oxygen	1/25/2016	6.42				mg/L	F
		4/25/2016	3.34				mg/L	F
		10/24/2016	3.84				mg/L	F
0602	Oxidation reduction potential	1/25/2016	56				mV	F
		4/25/2016	42				mV	F
		10/24/2016	128.6				mV	F
0602	Specific conductance	1/25/2016	1830				µmhos/cm	F
		4/25/2016	1540				µmhos/cm	F
		10/24/2016	1530				µmhos/cm	F
0602	Tetrachloroethene	1/25/2016	0.16	0.16	U		µg/L	F
		4/25/2016	0.16	0.16	QU	J	µg/L	F
		10/24/2016	0.16	0.16	U		µg/L	F
0602	Trichloroethene	1/25/2016	12.8	0.16			µg/L	F
		4/25/2016	6.66	0.16	Q	J	µg/L	F
		10/24/2016	10	0.16			µg/L	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0602	Tritium	1/25/2016	5740	572			pCi/L	F
0602	Turbidity	1/25/2016	1000		>		NTU	F
		4/25/2016	380				NTU	F
		10/24/2016	813				NTU	F
0602	Vinyl chloride	1/25/2016	0.16	0.16	U		µg/L	F
		4/25/2016	0.16	0.16	QU	J	µg/L	F
		10/24/2016	0.16	0.16	U		µg/L	F
0602	*cis*-1,2-Dichloroethene	1/25/2016	7.35	0.16			µg/L	F
		4/25/2016	4.54	0.16	Q	J	µg/L	F
		10/24/2016	14.8	0.16			µg/L	F
0602	pH	1/25/2016	7.36				s.u.	F
		4/25/2016	7				s.u.	F
		10/24/2016	7.32				s.u.	F
0602	*trans*-1,2-Dichloroethene	1/25/2016	0.3	0.16	J		µg/L	F
		4/25/2016	0.29	0.16	JQ	J	µg/L	F
		10/24/2016	0.16	0.16	U		µg/L	F
0605	Dissolved oxygen	1/25/2016	8.17				mg/L	F
		4/25/2016	2.45				mg/L	F
		7/25/2016	7.2				mg/L	F
		10/24/2016	5.7				mg/L	F
0605	Oxidation reduction potential	1/25/2016	44				mV	F
		4/25/2016	-69				mV	F
		7/25/2016	-63.5				mV	F
		10/24/2016	124.4				mV	F
0605	Specific conductance	1/25/2016	1980				µmhos/cm	F
		4/25/2016	2850				µmhos/cm	F
		7/25/2016	2090				µmhos/cm	F
		10/24/2016	1570				µmhos/cm	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0605	Tetrachloroethene	1/25/2016	0.21	0.16	J		µg/L	F
		4/25/2016	0.33	0.16	JQ	J	µg/L	F
		7/25/2016	0.16	0.16	U		µg/L	F
		10/24/2016	0.16	0.16	U		µg/L	F
0605	Trichloroethene	1/25/2016	10.2	0.16			µg/L	F
		4/25/2016	14.2	0.16	Q	J	µg/L	F
		7/25/2016	7.28	0.16			µg/L	F
		10/24/2016	9.82	0.16			µg/L	F
0605	Tritium	1/25/2016	7460	330			pCi/L	F
		7/25/2016	4410	395			pCi/L	F
0605	Turbidity	1/25/2016	949				NTU	F
		4/25/2016	804				NTU	F
		7/25/2016	936				NTU	F
		10/24/2016	5.62				NTU	F
0605	Vinyl chloride	1/25/2016	0.16	0.16	U		µg/L	F
		4/25/2016	0.16	0.16	QU	J	µg/L	F
		7/25/2016	0.16	0.16	U		µg/L	F
		10/24/2016	0.16	0.16	U		µg/L	F
0605	*cis*-1,2-Dichloroethene	1/25/2016	1.18	0.16			µg/L	F
		4/25/2016	2.03	0.16	Q	J	µg/L	F
		7/25/2016	1.02	0.16			µg/L	F
		10/24/2016	5.92	0.16			µg/L	F
0605	pH	1/25/2016	7.37				s.u.	F
		4/25/2016	7.24				s.u.	F
		7/25/2016	7.42				s.u.	F
		10/24/2016	7.49				s.u.	F
0605	*trans*-1,2-Dichloroethene	1/25/2016	0.18	0.16	J		µg/L	F
		4/25/2016	0.21	0.16	JQ	J	µg/L	F
		7/25/2016	0.16	0.16	U		µg/L	F
		10/24/2016	0.16	0.16	U		µg/L	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0606	Dissolved oxygen	1/25/2016	2.65				mg/L	F
		4/25/2016	8.56				mg/L	F
		10/24/2016	6.72				mg/L	F
0606	Oxidation reduction potential	1/25/2016	62				mV	F
		4/25/2016	-47				mV	F
		10/24/2016	147.5				mV	F
0606	Specific conductance	1/25/2016	1730				µmhos/cm	F
		4/25/2016	1770				µmhos/cm	F
		10/24/2016	1950				µmhos/cm	F
0606	Tetrachloroethene	1/25/2016	0.16	0.16	U		µg/L	F
		4/25/2016	0.16	0.16	QU	J	µg/L	F
		7/25/2016	0.16	0.16	U		µg/L	F
		10/24/2016	0.16	0.16	U		µg/L	F
0606	Trichloroethene	1/25/2016	1.88	0.16			µg/L	F
		4/25/2016	1.81	0.16	Q	J	µg/L	F
		7/25/2016	0.99	0.16	J		µg/L	F
		10/24/2016	1.64	0.16			µg/L	F
0606	Tritium	1/25/2016	3580	451			pCi/L	F
		7/25/2016	4790	357			pCi/L	F
0606	Turbidity	1/25/2016	1000		>		NTU	F
		4/25/2016	1000		>		NTU	F
		10/24/2016	739				NTU	F
0606	Vinyl chloride	1/25/2016	0.16	0.16	U		µg/L	F
		4/25/2016	0.16	0.16	QU	J	µg/L	F
		7/25/2016	0.16	0.16	U		µg/L	F
		10/24/2016	0.16	0.16	U		µg/L	F
0606	*cis*-1,2-Dichloroethene	1/25/2016	0.23	0.16	J		µg/L	F
		4/25/2016	0.16	0.16	QU	J	µg/L	F
		7/25/2016	0.3	0.16	J		µg/L	F
		10/24/2016	1.27	0.16			µg/L	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0606	pH	1/25/2016	7.56				s.u.	F
		4/25/2016	7.24				s.u.	F
		10/24/2016	7.48				s.u.	F
0606	*trans*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U		µg/L	F
		4/25/2016	0.16	0.16	QU	J	µg/L	F
		7/25/2016	0.16	0.16	U		µg/L	F
		10/24/2016	0.16	0.16	U		µg/L	F
0607	Dissolved oxygen	1/25/2016	9.48				mg/L	F
		4/25/2016	8.36				mg/L	F
		7/25/2016	9.93				mg/L	F
		10/24/2016	7.08				mg/L	F
0607	Oxidation reduction potential	1/25/2016	231				mV	F
		4/25/2016	51				mV	F
		7/25/2016	200.1				mV	F
		10/24/2016	162.9				mV	F
0607	Specific conductance	1/25/2016	1700				µmhos/cm	F
		4/25/2016	1640				µmhos/cm	F
		7/25/2016	1710				µmhos/cm	F
		10/24/2016	1400				µmhos/cm	F
0607	Tetrachloroethene	1/25/2016	0.16	0.16	U		µg/L	F
		4/25/2016	0.16	0.16	QU	J	µg/L	F
		7/25/2016	0.16	0.16	U		µg/L	F
		10/24/2016	0.16	0.16	U		µg/L	F
0607	Trichloroethene	1/25/2016	4.87	0.16			µg/L	F
		4/25/2016	7.31	0.16	Q	J	µg/L	F
		7/25/2016	6.92	0.16			µg/L	F
		10/24/2016	5.27	0.16			µg/L	F
0607	Tritium	1/25/2016	3710	314			pCi/L	F
		7/25/2016	3080	346			pCi/L	F

Table D-2. Parcels 6, 7, and 8 Groundwater and Seep Data (continued)

Location ID	Analyte	Sample Date	Value	Detection Limit	Laboratory Qualifier	Validation Qualifier	Units	Sample Type
0607	Turbidity	1/25/2016	932				NTU	F
		4/25/2016	195				NTU	F
		7/25/2016	263				NTU	F
		10/24/2016	249				NTU	F
0607	Vinyl chloride	1/25/2016	0.16	0.16	U		µg/L	F
		4/25/2016	0.16	0.16	QU	J	µg/L	F
		7/25/2016	0.16	0.16	U		µg/L	F
		10/24/2016	0.16	0.16	U		µg/L	F
0607	*cis*-1,2-Dichloroethene	1/25/2016	0.51	0.16	J		µg/L	F
		4/25/2016	0.92	0.16	JQ	J	µg/L	F
		7/25/2016	0.97	0.16	J		µg/L	F
		10/24/2016	0.96	0.16	J		µg/L	F
0607	pH	1/25/2016	7.32				s.u.	F
		4/25/2016	7.33				s.u.	F
		7/25/2016	7.38				s.u.	F
		10/24/2016	7.52				s.u.	F
0607	*trans*-1,2-Dichloroethene	1/25/2016	0.16	0.16	U		µg/L	F
		4/25/2016	0.16	0.16	QU	J	µg/L	F
		7/25/2016	0.16	0.16	U		µg/L	F
		10/24/2016	0.16	0.16	U		µg/L	F

Abbreviations:
D = analyte determined in diluted sample
F = low flow sampling method used
J = estimated value
µg/L = micrograms per liter
µmhos/cm = micromhos per centimeter
mg/L = milligrams per liter
mV = millivolts
pCi/L = picocuries per liter
Q = quantitative result due to sampling technique
s.u. = standard unit
U = analytical result below detection limit

Appendix E

Data Assessment Reports

This page intentionally left blank

Data Review and Validation Report

General Information

Requisition No. (RIN):	16017607
Sample Event:	January 25-26, 2016
Site(s):	Mound, Ohio; Groundwater
Laboratory:	GEL Laboratories, Charleston, South Carolina
Work Order No.:	390212
Analysis:	Organics and Radiochemistry
Validator:	Gretchen Baer
Review Date:	April 21, 2016

This validation was performed according to the *Environmental Procedures Catalog* (LMS/POL/S04325), *"Standard Practice for Validation of Environmental Data."* The procedure was applied at Level 2, Data Deliverables Verification. See attached Data Validation Worksheets for supporting documentation on the data review and validation. All analyses were successfully completed. The samples were prepared and analyzed using accepted procedures based on methods specified by line item code, which are listed in Table 1.

Table 1. Analytes and Methods

Analyte	Line Item Code	Prep Method	Analytical Method
Tritium	LSC-A-001	EPA 906.0 Mod	EPA 906.0 Mod
Volatile Organics, VOA	VOA-A-007	SW-846 5030B	SW-846 8260 LL

Data Qualifier Summary

Analytical results were qualified as listed in Table 2. Refer to the sections below for an explanation of the data qualifiers applied.

Table 2. Data Qualifier Summary

Sample Number	Location	Analyte(s)	Flag	Reason
390212-001	0138	Tritium	U	Less than the Decision Level Concentration
390212-008	0346	Methylene chloride	U	Less than 10 times the trip blank
390212-008	0346	Tritium	J	Less than the Determination Limit
390212-010	0347	Choroform	U	Less than 5 times the trip blank
390212-017	0605	Choroform	U	Less than 5 times the trip blank
390212-021	0347 Dup	Choroform	U	Less than 5 times the trip blank

Sample Shipping/Receiving

GEL Laboratories in Charleston, South Carolina, received 21 water samples on January 28, 2016, accompanied by Chain of Custody (COC) forms. The COC forms were checked to confirm that all of the samples were listed with sample collection dates and times, and that signatures and dates were present indicating sample relinquishment and receipt. The COC forms were complete with no errors or omissions. The air waybill numbers were listed on the receiving documentation.

Preservation and Holding Times

The sample shipment was received intact with the temperature inside the iced cooler at 4 °C, which complies with requirements. All samples were received in the correct container types and had been preserved correctly for the requested analyses. Sample analysis was completed within the applicable holding times.

Detection and Quantitation Limits

The method detection limit (MDL) was reported for all organic analytes as required. The MDL, as defined in 40 CFR 136, is the minimum concentration of an analyte that can be measured and reported with 99 percent confidence that the analyte concentration is greater than zero. The practical quantitation limit (PQL) for these analytes is the lowest concentration that can be reliably measured, and is defined as 5 times the MDL.

For radiochemical analytes (those measured by radiometric counting) the MDL and PQL are not applicable, and these results are evaluated using the minimum detectable concentration (MDC), Decision Level Concentration (DLC), and Determination Limit (DL). The MDC is a measure of radiochemical method performance and was calculated and reported as specified in the *Quality Systems Manual*. The DLC is the minimum concentration of an analyte that can be measured and reported with 99 percent confidence that the analyte concentration is greater than zero, and is estimated as 3 times the one-sigma total propagated uncertainty. Results that are greater than the MDC, but less than the DLC are qualified with a “U” flag (not detected). The DL for radiochemical results is the lowest concentration that can be reliably measured, and is defined as 3 times the MDC. Results not previously “U” qualified that are less than the DL are qualified with a “J” flag as estimated values.

The requested detection limit of 400 pCi/L for tritium was not met for three samples; however, all associated tritium results were well above the MDC. The reported MDLs for all organic

analytes and MDCs for radiochemical analytes demonstrate compliance with contractual requirements.

Laboratory Instrument Calibration

Compliance requirements for satisfactory instrument calibration are established to ensure that the instrument is capable of producing acceptable qualitative and quantitative data for all analytes. Initial calibration demonstrates that the instrument is capable of acceptable performance in the beginning of the analytical run. Compliance requirements for continuing calibration checks are established to ensure that the instrument continues to be capable of producing acceptable qualitative and quantitative data. All laboratory instrument calibrations were performed correctly in accordance with the cited methods. All calibration and laboratory spike standards were prepared from independent sources.

The Continuing Calibration Verification standards associated with the samples exhibited percent difference values within acceptance criteria for all compounds, with the following exceptions. 2-Butanone and 2-hexanone had percent drift values greater than 20 percent. There were no sample results greater than the MDL associated with these calibration verification compounds, so no qualification is necessary.

Volatiles Internal Standards and Surrogates

The volatile internal standard recoveries and surrogate recoveries were within the acceptance ranges for all samples.

Method Blanks

Method blanks are analyzed to assess any contamination that may have occurred during sample preparation. Methylene chloride (a common laboratory contaminant) was detected in some volatiles method blanks at concentrations above the MDL; methylene chloride was not detected in any sample associated with these blanks and no further qualification is required. The tritium method blank result was less than the DLC.

Trip Blank

Trip blanks were prepared and analyzed to document contamination attributable to shipping and field handling procedures. This type of blank is useful in documenting contamination of volatile organic samples. Acetone, 2-butanone, choroform, methylene chloride, toluene, and xylenes were detected in some trip blanks. Sample results for these compounds that are less than 5 times the trip blank concentration (and less than 10 times the blank concentration for common laboratory contaminants) are qualified with a “U” flag as not detected.

Matrix Spike Analysis

Matrix spike and matrix spike duplicate (MS/MSD) samples are used to measure method performance in the sample matrix. The spike recoveries met the acceptance criteria for all analytes evaluated.

Laboratory Replicate Analysis

Laboratory replicate analyses are used to determine laboratory precision for each sample matrix. The relative percent difference (RPD) for organic replicate results should be less than the laboratory-derived control limits. For radiochemical measurements, the relative error ratio (the ratio of the absolute difference between the sample and duplicate results and the sum of the 1-sigma uncertainties) is used to evaluate duplicate results and should be less than 3. All replicate results met these criteria, with the following exception. The RPD for trichlorofluoromethane exceeded the criteria in a matrix spike sample. Trichlorofluoromethane was not detected in the associated sample and no further qualification is necessary.

Laboratory Control Sample

Laboratory control samples were analyzed at the correct frequency to provide information on the accuracy of the analytical method and the overall laboratory performance, including sample preparation. All control sample results were acceptable.

Field Duplicate

Field duplicate samples are collected and analyzed as an indication of overall precision of the measurement process. The precision observed includes both field and laboratory precision and has more variability than laboratory duplicates, which measure only laboratory performance. Duplicate samples were collected from location 0347 (field duplicate ID 9347). For non-radiochemical measurements, the relative percent difference for duplicate results that are greater than 5 times the PQL should be less than 20 percent. For results less than 5 times the PQL, the range should be no greater than the PQL. For radiochemical measurements, the relative error ratio (the ratio of the absolute difference between the sample and duplicate results and the sum of the 1-sigma uncertainties) is used to evaluate duplicate results and should be less than 3. All duplicate results met these criteria, demonstrating acceptable precision.

Completeness

Results were reported in the correct units for all analytes requested using contract-required laboratory qualifiers.

Electronic Data Deliverable (EDD) File

The EDD file arrived on February 25, 2016. The Sample Management System EDD validation module was used to verify that the EDD file was complete and in compliance with requirements. The module compares the contents of the file to the requested analyses to ensure all and only the requested data are delivered. The contents of the EDD were manually examined to verify that the sample results accurately reflect the data contained in the sample data package.

Sampling Protocol

Sample results for all monitoring wells met the Category I low-flow sampling criteria and were qualified with an “F” flag in the database, indicating the wells were purged and sampled using

the low-flow sampling method. The turbidity values for locations 0138 and 0315 were greater than 50 NTU.

No field notes were provided for the seep samples.

Outliers Report

Potential outliers are measurements that are extremely large or small relative to the rest of the data and, therefore, are suspected of misrepresenting the population from which they were collected. Potential outliers may result from transcription errors, data-coding errors, or measurement system problems. However, outliers may also represent true extreme values of a distribution and indicate more variability in the population than was expected.

Statistical outlier tests give probabilistic evidence that an extreme value does not "fit" with the distribution of the remainder of the data and is therefore a statistical outlier. These tests should only be used to identify data points that require further investigation. The tests alone cannot determine whether a statistical outlier should be discarded or corrected within a data set.

There are three steps involved in identifying extreme values or outliers:

1. Identify extreme values that may be potential outliers by generating the Outliers Report using the Sample Management System from data in the environmental database. The application compares the new data set (in standard environmental database units) with historical data and lists the new data that fall outside the historical data range. A determination is also made if the data are normally distributed using the Shapiro-Wilk Test.
2. Apply the appropriate statistical test. Dixon's Extreme Value test is used to test for statistical outliers when the sample size is less than or equal to 25. This test considers both extreme values that are much smaller than the rest of the data (case 1) and extreme values that are much larger than the rest of the data (case 2). This test is valid only if the data without the suspected outlier are normally distributed. Rosner's Test is a parametric test that is used to detect outliers for sample sizes of 25 or more. This test also assumes that the data without the suspected outliers are normally distributed.
3. Scientifically review statistical outliers and decide on their disposition. The review should include an evaluation of any notable trends in the data that may indicate the outliers represent true extreme values.

No results from this sampling event were identified as potential outliers.

Report Prepared By: _____

Gretchen Baer
Data Validator

Data Validation Outliers Report - No Field Parameters
Comparison: All historical Data Beginning 1/1/2006
Laboratory: GEL Laboratories
RIN: 16017607
Report Date: 4/21/2016

					Current			Historical Maximum			Historical Minimum			Number of Data Points		Statistical Outlier
						Qualifiers			*Qualifiers*			*Qualifiers*				
Site Code	Location Code	Sample ID	Sample Date	Analyte	Result	*Lab*	*Data*	Result	*Lab*	*Data*	Result	*Lab*	*Data*	N	N Below Detect	
MND01	0608	N001	01/25/2016	Tritium	4580			54000			5790			29	0	No

STATISTICAL TESTS:
The distribution of the data is tested for normality or lognormality using the Shapiro-Wilk Test
Outliers are identified using Dixon's Test when there are 25 or fewer data points.
Outliers are identified using Rosner's Test when there are 26 or more data points.
See Data Quality Assessment: Statistical Methods for Practitioners, EPA QC/G-9S, February 2006.

SAMPLE MANAGEMENT SYSTEM
General Data Validation Report

RIN: 16017607 **Lab Code:** GEN **Validator:** Gretchen Baer **Validation Date:** 4/21/2016

Project: Mound LTS&M Ground Water **Analysis Type:** ☐ Metals ☐ General Chem ☑ Rad ☑ Organics

# of Samples: 21 **Matrix:** Water **Requested Analysis Completed:** Yes

Chain of Custody

Present: OK **Signed:** OK **Dated:** OK

Sample

Integrity: OK **Preservation:** OK **Temperature:** OK

Select Quality Parameters

☑ Holding Times	All analyses were completed within the applicable holding times.
☑ Detection Limits	There are 3 detection limit failures.
☑ Field/Trip Blanks	There were 2 trip/equipment blanks evaluated.
☑ Field Duplicates	There was 1 duplicate evaluated.

SAMPLE MANAGEMENT SYSTEM

Non-Compliance Report: Detection Limits

RIN: 16017607 **Lab Code:** GEN

Project: Mound LTS&M Ground Water

Validation Date: 4/21/2016

Ticket	Location	Lab Sample ID	Method Code	Lab Method	Analyte Name	Result	Qualifier	Reported Detection Limit	Required Detection Limit	Units
OCU 552	0602	390212014	LSC-A-001	EPA 906.0 Modified	Tritium	5740		572	400	pCi/L
OCU 553	0606	390212015	LSC-A-001	EPA 906.0 Modified	Tritium	3580		451	400	pCi/L
OCU 544	0608	390212006	LSC-A-001	EPA 906.0 Modified	Tritium	4580		416	400	pCi/L

SAMPLE MANAGEMENT SYSTEM

Validation Report: Equipment/Trip Blanks

RIN: 16017607 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 4/21/2016

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	390212019	EPA 8260 Low Level	Chloroform	1.73		0.160	ug/L

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
390212001	OCU 537	0138	0.160	1.00	U	
390212002	OCU 538	0386	0.160	1.00	U	
390212003	OCU 539	0387	0.160	1.00	U	
390212004	OCU 540	0389	0.160	1.00	U	
390212005	OCU 541	0392	0.160	1.00	U	
390212006	OCU 544	0608	0.160	1.00	U	
390212007	OCU 545	0118	0.160	1.00	U	
390212008	OCU 546	0346	0.160	1.00	U	
390212009	OCU 547	0379	0.160	1.00	U	
390212010	OCU 548	0347	0.240	1.00	J	U
390212011	OCU 549	0315	0.160	1.00	U	
390212012	OCU 550	0124	0.160	1.00	U	
390212013	OCU 551	0126	0.160	1.00	U	
390212014	OCU 552	0602	0.160	1.00	U	
390212015	OCU 553	0606	0.160	1.00	U	
390212016	OCU 554	0601	0.160	1.00	U	
390212017	OCU 555	0605	0.480	1.00	J	U
390212018	OCU 556	0607	0.160	1.00	U	
390212021	OCU 559	9347	0.270	1.00	J	U

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	390212019	EPA 8260 Low Level	Methylene Chloride (Lab Contaminant)	0.260	J	0.160	ug/L

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
390212001	OCU 537	0138	0.160	1.00	U	
390212002	OCU 538	0386	0.160	1.00	U	
390212003	OCU 539	0387	0.160	1.00	U	
390212004	OCU 540	0389	0.160	1.00	U	
390212005	OCU 541	0392	0.160	1.00	U	
390212006	OCU 544	0608	0.160	1.00	U	

SAMPLE MANAGEMENT SYSTEM

Validation Report: Equipment/Trip Blanks

RIN: 16017607 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 4/21/2016

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	390212019		Methylene Chloride (Lab Contaminant)				

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
390212007	OCU 545	0118	0.160	1.00	U	
390212008	OCU 546	0346	0.240	1.00	J	U
390212009	OCU 547	0379	0.160	1.00	U	
390212010	OCU 548	0347	0.160	1.00	U	
390212011	OCU 549	0315	0.160	1.00	U	
390212012	OCU 550	0124	0.160	1.00	U	
390212013	OCU 551	0126	0.160	1.00	U	
390212014	OCU 552	0602	0.160	1.00	U	
390212015	OCU 553	0606	0.160	1.00	U	
390212016	OCU 554	0601	0.160	1.00	U	
390212017	OCU 555	0605	0.160	1.00	U	
390212018	OCU 556	0607	0.160	1.00	U	
390212021	OCU 559	9347	0.160	1.00	U	

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	390212020	EPA 8260 Low Level	2-Butanone (Lab Contaminant)	2.07	J	0.500	ug/L

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
390212001	OCU 537	0138	0.500	1.00	U	
390212002	OCU 538	0386	0.500	1.00	U	
390212003	OCU 539	0387	0.500	1.00	U	
390212004	OCU 540	0389	0.500	1.00	U	
390212005	OCU 541	0392	0.500	1.00	U	
390212006	OCU 544	0608	0.500	1.00	U	
390212007	OCU 545	0118	0.500	1.00	U	
390212008	OCU 546	0346	0.500	1.00	U	
390212009	OCU 547	0379	0.500	1.00	U	
390212010	OCU 548	0347	0.500	1.00	U	
390212011	OCU 549	0315	0.500	1.00	U	
390212012	OCU 550	0124	0.500	1.00	U	

SAMPLE MANAGEMENT SYSTEM

Validation Report: Equipment/Trip Blanks

RIN: 16017607 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 4/21/2016

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	390212020		2-Butanone (Lab Contaminant)				

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
390212013	OCU 551	0126	0.500	1.00	U	
390212014	OCU 552	0602	0.500	1.00	U	
390212015	OCU 553	0606	0.500	1.00	U	
390212016	OCU 554	0601	0.500	1.00	U	
390212017	OCU 555	0605	0.500	1.00	U	
390212018	OCU 556	0607	0.500	1.00	U	
390212021	OCU 559	9347	0.500	1.00	U	

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	390212020	EPA 8260 Low Level	Acetone (Lab Contaminant)	14.6		0.500	ug/L

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
390212001	OCU 537	0138	0.500	1.00	U	
390212002	OCU 538	0386	0.500	1.00	U	
390212003	OCU 539	0387	0.500	1.00	U	
390212004	OCU 540	0389	0.500	1.00	U	
390212005	OCU 541	0392	0.500	1.00	U	
390212006	OCU 544	0608	0.500	1.00	U	
390212007	OCU 545	0118	0.500	1.00	U	
390212008	OCU 546	0346	0.500	1.00	U	
390212009	OCU 547	0379	0.500	1.00	U	
390212010	OCU 548	0347	0.500	1.00	U	
390212011	OCU 549	0315	0.500	1.00	U	
390212012	OCU 550	0124	0.500	1.00	U	
390212013	OCU 551	0126	0.500	1.00	U	
390212014	OCU 552	0602	0.500	1.00	U	
390212015	OCU 553	0606	0.500	1.00	U	
390212016	OCU 554	0601	0.500	1.00	U	
390212017	OCU 555	0605	0.500	1.00	U	
390212018	OCU 556	0607	0.500	1.00	U	

SAMPLE MANAGEMENT SYSTEM

Validation Report: Equipment/Trip Blanks

RIN: 16017607 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 4/21/2016

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	390212020		Acetone (Lab Contaminant)				

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
390212021	OCU 559	9347	0.500	1.00	U	

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	390212020	EPA 8260 Low Level	Methylene Chloride (Lab Contaminant)	0.290	J	0.160	ug/L

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
390212001	OCU 537	0138	0.160	1.00	U	
390212002	OCU 538	0386	0.160	1.00	U	
390212003	OCU 539	0387	0.160	1.00	U	
390212004	OCU 540	0389	0.160	1.00	U	
390212005	OCU 541	0392	0.160	1.00	U	
390212006	OCU 544	0608	0.160	1.00	U	
390212007	OCU 545	0118	0.160	1.00	U	
390212008	OCU 546	0346	0.240	1.00	J	U
390212009	OCU 547	0379	0.160	1.00	U	
390212010	OCU 548	0347	0.160	1.00	U	
390212011	OCU 549	0315	0.160	1.00	U	
390212012	OCU 550	0124	0.160	1.00	U	
390212013	OCU 551	0126	0.160	1.00	U	
390212014	OCU 552	0602	0.160	1.00	U	
390212015	OCU 553	0606	0.160	1.00	U	
390212016	OCU 554	0601	0.160	1.00	U	
390212017	OCU 555	0605	0.160	1.00	U	
390212018	OCU 556	0607	0.160	1.00	U	
390212021	OCU 559	9347	0.160	1.00	U	

SAMPLE MANAGEMENT SYSTEM

Validation Report: Equipment/Trip Blanks

RIN: 16017607 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 4/21/2016

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	390212020	EPA 8260 Low Level	Toluene (Lab Contaminant)	0.210	J	0.160	ug/L

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
390212001	OCU 537	0138	0.160	1.00	U	
390212002	OCU 538	0386	0.160	1.00	U	
390212003	OCU 539	0387	0.160	1.00	U	
390212004	OCU 540	0389	0.160	1.00	U	
390212005	OCU 541	0392	0.160	1.00	U	
390212006	OCU 544	0608	0.160	1.00	U	
390212007	OCU 545	0118	0.160	1.00	U	
390212008	OCU 546	0346	0.160	1.00	U	
390212009	OCU 547	0379	0.160	1.00	U	
390212010	OCU 548	0347	0.160	1.00	U	
390212011	OCU 549	0315	0.160	1.00	U	
390212012	OCU 550	0124	0.160	1.00	U	
390212013	OCU 551	0126	0.160	1.00	U	
390212014	OCU 552	0602	0.160	1.00	U	
390212015	OCU 553	0606	0.160	1.00	U	
390212016	OCU 554	0601	0.160	1.00	U	
390212017	OCU 555	0605	0.160	1.00	U	
390212018	OCU 556	0607	0.160	1.00	U	
390212021	OCU 559	9347	0.160	1.00	U	

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	390212020	EPA 8260 Low Level	Total Xylenes	0.540	J	0.160	ug/L

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
390212001	OCU 537	0138	0.160	1.00	U	
390212002	OCU 538	0386	0.160	1.00	U	
390212003	OCU 539	0387	0.160	1.00	U	
390212004	OCU 540	0389	0.160	1.00	U	
390212005	OCU 541	0392	0.160	1.00	U	
390212006	OCU 544	0608	0.160	1.00	U	

SAMPLE MANAGEMENT SYSTEM

Validation Report: Equipment/Trip Blanks

RIN: 16017607 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 4/21/2016

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	390212020		Total Xylenes				

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
390212007	OCU 545	0118	0.160	1.00	U	
390212008	OCU 546	0346	0.160	1.00	U	
390212009	OCU 547	0379	0.160	1.00	U	
390212010	OCU 548	0347	0.160	1.00	U	
390212011	OCU 549	0315	0.160	1.00	U	
390212012	OCU 550	0124	0.160	1.00	U	
390212013	OCU 551	0126	0.160	1.00	U	
390212014	OCU 552	0602	0.160	1.00	U	
390212015	OCU 553	0606	0.160	1.00	U	
390212016	OCU 554	0601	0.160	1.00	U	
390212017	OCU 555	0605	0.160	1.00	U	
390212018	OCU 556	0607	0.160	1.00	U	
390212021	OCU 559	9347	0.160	1.00	U	

SAMPLE MANAGEMENT SYSTEM

Validation Report: Field Duplicates

RIN: 16017607 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 4/21/2016

Duplicate: 9347 **Sample:** 0347

	Sample				Duplicate						
Analyte	Result	Flag	Error	Dilution	Result	Flag	Error	Dilution	RPD	RER	Units
1,1,1,2-Tetrachloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1,1-Trichloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1,2,2-Tetrachloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1,2-Trichloro-1,2,2-trifluoroethane	1.50	U		1.00	1.50	U		1.00			ug/L
1,1,2-Trichloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1-Dichloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1-Dichloroethene	0.160	U		1.00	0.160	U		1.00			ug/L
1,1-Dichloropropene	0.160	U		1.00	0.160	U		1.00			ug/L
1,2,3-Trichlorobenzene	0.200	U		1.00	0.200	U		1.00			ug/L
1,2,3-Trichloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
1,2,4-Trichlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,2,4-Trimethylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dibromo-3-chloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dibromoethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dichlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dichloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dichloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
1,3,5-Trimethylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,3-Dichlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,3-Dichloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
1,4-Dichlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
2,2-Dichloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
2-Butanone	0.500	U		1.00	0.500	U		1.00			ug/L
2-Chlorotoluene	0.160	U		1.00	0.160	U		1.00			ug/L
2-Hexanone	0.500	U		1.00	0.500	U		1.00			ug/L
4-Chlorotoluene	0.160	U		1.00	0.160	U		1.00			ug/L
4-Isopropyltoluene	0.160	U		1.00	0.160	U		1.00			ug/L
4-Methyl-2-pentanone	0.500	U		1.00	0.500	U		1.00			ug/L
Acetone	0.500	U		1.00	0.500	U		1.00			ug/L
Benzene	0.160	U		1.00	0.160	U		1.00			ug/L
Bromobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Bromochloromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Bromodichloromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Bromoform	0.160	U		1.00	0.160	U		1.00			ug/L
Bromomethane	0.160	U		1.00	0.160	U		1.00			ug/L
Carbon Disulfide	0.500	U		1.00	0.500	U		1.00			ug/L
Carbon Tetrachloride	1.44			1.00	1.35			1.00	6.45		ug/L
Chlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Chloroethane	0.160	U		1.00	0.160	U		1.00			ug/L

SAMPLE MANAGEMENT SYSTEM

Validation Report: Field Duplicates

RIN: 16017607 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 4/21/2016

Duplicate: 9347 **Sample:** 0347

	Sample				Duplicate						
Analyte	Result	Flag	Error	Dilution	Result	Flag	Error	Dilution	RPD	RER	Units
Chloroform	0.240	J		1.00	0.270	J		1.00			ug/L
Chloromethane	0.160	U		1.00	0.160	U		1.00			ug/L
cis-1,2-Dichloroethene	0.160	U		1.00	0.160	U		1.00			ug/L
cis-1,3-Dichloropropene	0.160	U		1.00	0.160	U		1.00			ug/L
Dibromochloromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Dibromomethane	0.160	U		1.00	0.160	U		1.00			ug/L
Dichlorodifluoromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Ethylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Hexachlorobutadiene	0.160	U		1.00	0.160	U		1.00			ug/L
Isopropylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Methylene Chloride	0.160	U		1.00	0.160	U		1.00			ug/L
Naphthalene	0.160	U		1.00	0.160	U		1.00			ug/L
n-Butylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
n-Propylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
sec-Butylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Styrene	0.160	U		1.00	0.160	U		1.00			ug/L
tert-Butylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Tetrachloroethene	0.160	U		1.00	0.160	U		1.00			ug/L
Toluene	0.160	U		1.00	0.160	U		1.00			ug/L
Total Xylenes	0.160	U		1.00	0.160	U		1.00			ug/L
trans-1,2-Dichloroethene	0.160	U		1.00	0.160	U		1.00			ug/L
trans-1,3-Dichloropropene	0.160	U		1.00	0.160	U		1.00			ug/L
Trichloroethene	16.6			1.00	16.1			1.00	3.06		ug/L
Trichlorofluoromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Tritium	1430		462	1.00	1670		506	1.00		0.7	pCi/L
Vinyl Chloride	0.160	U		1.00	0.160	U		1.00			ug/L

SAMPLE MANAGEMENT SYSTEM
Organics Data Validation Summary

RIN: 16017607 **Project:** Mound LTS&M Ground Water **Lab Code:** GEN **Validation Date:** 4/21/2016

LCS Recovery: All LCS recoveries were within the laboratory acceptance limits.

Method Blank(s): There were 2 method blank results above the MDL.

MS/MSD Recovery: There was 1 MS/MSD failure.

Surrogate Recovery: All surrogate recoveries were within the laboratory acceptance limits.

SAMPLE MANAGEMENT SYSTEM

Non-Compliance Report: MS/MSD Performance

RIN: 16017607 **Lab Code:** GEN

Project: Mound LTS&M Ground Water

Validation Date: 4/21/2016

MS/MSD	Date Analyzed	Method	Analyte	Recovery MS	Recovery MSD	Lower Limit	Upper Limit	MSD RPD	RPD Limit
OCU 599	02/03/2016	EPA 8260 Low Level	Trichlorofluoromethane	114.0	90.0	61.0	132.0	24.00	20.0

SAMPLE MANAGEMENT SYSTEM

Non-Compliance Report: Method Blanks

RIN: 16017607 **Lab Code:** GEN

Project: Mound LTS&M Ground Water

Validation Date: 4/21/2016

Method Blank	Date Analyzed	Method	Analyte	Result	Flag(s)	MDL
1203481259	2/3/2016	EPA 8260 Low Level	Methylene Chloride	0.26	J	0.16
1203481265	2/3/2016	EPA 8260 Low Level	Methylene Chloride	0.26	J	0.16

SAMPLE MANAGEMENT SYSTEM
Radiochemistry Data Validation Worksheet

RIN: 16017607 **Lab Code:** GEN **Date Due:** 2/25/2016

Matrix: Water **Site Code:** MND03 **Date Completed:** 2/25/2016

Sample	Analyte	Date Analyzed	Result	Flag	Tracer %R	LCS %R	MS %R	Duplicate RER
0347	Tritium	02/23/2016						0.49
Blank	Tritium	02/23/2016	0.0509	U				
0347	Tritium	02/23/2016					119.0	
Blank_Spike	Tritium	02/23/2016				104.00		

Data Review and Validation Report

General Information

Requisition No. (RIN):	16017608
Sample Event:	January 25, 2016
Site(s):	Mound, Ohio; LTS&M Groundwater
Laboratory:	GEL Laboratories, Charleston, South Carolina
Work Order No.:	390190
Analysis:	Organics
Validator:	Samantha Tigar
Review Date:	April 26, 2016

This validation was performed according to the *Environmental Procedures Catalog* (LMS/POL/S04325), *"Standard Practice for Validation of Environmental Data."* The procedure was applied at Level 2, Data Deliverables Verification. See attached Data Validation Worksheets for supporting documentation on the data review and validation. All analyses were successfully completed. The samples were prepared and analyzed using accepted procedures based on methods specified by line item code, which are listed in Table 1.

Table 1. Analytes and Methods

Analyte	Line Item Code	Prep Method	Analytical Method
Volatile Organics, VOA	VOA-A-007	SW-846 5030B	SW-846 8260 LL

Data Qualifier Summary

Analytical results were qualified as listed in Table 2. Refer to the sections below for an explanation of the data qualifiers applied.

Table 2. Data Qualifier Summary

Sample Number	Location	Analyte(s)	Flag	Reason
390190008	0999	Acetone	J	Percent drift > 20%
390190002	0400	Chloromethane	U	Less than 5x the trip blank

Sample Number	Location	Analyte(s)	Flag	Reason
390190011	0617	Chloromethane	U	Less than 5x the trip blank
390190001	0353	Methylene Chloride	U	Less than 10x the method blank
390190002	0400	Methylene Chloride	U	Less than 10x the method blank
390190007	0445	Methylene Chloride	U	Less than 10x the method blank
390190008	0999	Methylene Chloride	U	Less than 10x the method blank
390190010	P033	Methylene Chloride	U	Less than 10x the method blank

Sample Shipping/Receiving

GEL Laboratories in Charleston, South Carolina, received 11 water samples on January 28, 2016, accompanied by Chain of Custody (COC) forms. The COC forms were checked to confirm that all of the samples were listed with sample collection dates and times, and that signatures and dates were present indicating sample relinquishment and receipt. The COC forms were complete with no errors or omissions. The air waybill number was listed on the receiving documentation.

Preservation and Holding Times

The sample shipment was received intact with the temperature inside the iced cooler at 5 °C, which complies with requirements. All samples were received in the correct container types and had been preserved correctly for the requested analyses. Sample analysis was completed within the applicable holding times.

Detection and Quantitation Limits

The method detection limit (MDL) was reported for all analytes as required. The MDL, as defined in 40 CFR 136, is the minimum concentration of an analyte that can be measured and reported with 99 percent confidence that the analyte concentration is greater than zero. The practical quantitation limit (PQL) for these analytes is the lowest concentration that can be reliably measured, and is defined as 5 times the MDL. The reported MDLs for all analytes demonstrate compliance with contractual requirements.

Laboratory Instrument Calibration

Compliance requirements for satisfactory instrument calibration are established to ensure that the instrument is capable of producing acceptable qualitative and quantitative data for all analytes. Initial calibration demonstrates that the instrument is capable of acceptable performance in the beginning of the analytical run. Compliance requirements for continuing calibration checks are established to ensure that the instrument continues to be capable of producing acceptable qualitative and quantitative data. All laboratory instrument calibrations were performed correctly in accordance with the cited methods. All calibration and laboratory spike standards were prepared from independent sources.

The Continuing Calibration Verification standards associated with the samples exhibited percent difference values within acceptance criteria for all compounds, with the following exceptions.

Acetone and 2-butanone had percent drift values greater than 20 percent. Associated sample results greater than the MDL were qualified with a "J" as an estimated value.

Volatiles Internal Standards and Surrogates

The volatile internal standard recoveries and surrogate recoveries were within the acceptance ranges for all samples.

Method Blanks

Method blanks are analyzed to assess any contamination that may have occurred during sample preparation. Methylene chloride (a common laboratory contaminant) was detected in the volatiles method blanks at concentrations above the MDL. Associated samples with results less than 10 times the blank concentrations are qualified with a "U" flag (not detected).

Trip Blank

Trip blanks were prepared and analyzed to document contamination attributable to shipping and field handling procedures. This type of blank is useful in documenting contamination of volatile organic samples. Acetone, chloroform, chloromethane, methylene chloride, toluene, and vinyl chloride were detected in the trip blank. Sample results for these compounds that are less than 5 times the trip blank concentration (and less than 10 times the blank concentration for common laboratory contaminants) are qualified with a "U" flag as not detected.

Matrix Spike Analysis

Matrix spike and matrix spike duplicate (MS/MSD) samples are used to measure method performance in the sample matrix. The spike recoveries met the acceptance criteria for all analytes evaluated.

Laboratory Replicate Analysis

Laboratory replicate analyses are used to determine laboratory precision for each sample matrix. The relative percent difference (RPD) for organic replicate results should be less than the laboratory-derived control limits. All replicate results met these criteria, with the following exception. The RPD for one volatile compound exceeded the criteria in a matrix spike sample. The compound was not detected in the associated sample, no qualification is necessary.

Laboratory Control Sample

Laboratory control samples were analyzed at the correct frequency to provide information on the accuracy of the analytical method and the overall laboratory performance, including sample preparation. All control sample results were acceptable.

Field Duplicate

Field duplicate samples are collected and analyzed as an indication of overall precision of the measurement process. The precision observed includes both field and laboratory precision and has more variability than laboratory duplicates, which measure only laboratory performance. A duplicate sample was collected from location 0411. The relative percent difference for duplicate results that are greater than 5 times the PQL should be less than 20 percent. For results that are less than 5 times the PQL, the range should be no greater than the PQL. The duplicate results met the criteria for all analytes, demonstrating acceptable overall precision.

Completeness

Results were reported in the correct units for all analytes requested using contract-required laboratory qualifiers.

Electronic Data Deliverable (EDD) File

The EDD file arrived on February 25, 2016. The Sample Management System EDD validation module was used to verify that the EDD file was complete and in compliance with requirements. The module compares the contents of the file to the requested analyses to ensure all and only the requested data are delivered. The contents of the EDD were manually examined to verify that the sample results accurately reflect the data contained in the sample data package.

Sampling Protocol

Sample results for all monitoring wells met the Category I or II low-flow sampling criteria and were qualified with an “F” flag in the database, indicating the wells were purged and sampled using the low-flow sampling method. The groundwater sample results for the wells 0353, 0444, and 0445 were further qualified with a “Q” flag in the database indicating the data are considered qualitative because these are Category II wells.

Note that well 0444 was categorized as a catergory I, but was sampled using catergory II criteria.

The water level for location 0411 did not meet stability requirements and the field measurements were flagged with a “J” as an estimated value. The turbidity value for location 0400 was greater than 50 NTU.

Outliers Report

Potential outliers are measurements that are extremely large or small relative to the rest of the data and, therefore, are suspected of misrepresenting the population from which they were collected. Potential outliers may result from transcription errors, data-coding errors, or measurement system problems. However, outliers may also represent true extreme values of a distribution and indicate more variability in the population than was expected.

Statistical outlier tests give probabilistic evidence that an extreme value does not "fit" with the distribution of the remainder of the data and is therefore a statistical outlier. These tests should only be used to identify data points that require further investigation. The tests alone cannot determine whether a statistical outlier should be discarded or corrected within a data set.

There are three steps involved in identifying extreme values or outliers:

1. Identify extreme values that may be potential outliers by generating the Outliers Report using the Sample Management System from data in the environmental database. The application compares the new data set (in standard environmental database units) with historical data and lists the new data that fall outside the historical data range. A determination is also made if the data are normally distributed using the Shapiro-Wilk Test.
2. Apply the appropriate statistical test. Dixon's Extreme Value test is used to test for statistical outliers when the sample size is less than or equal to 25. This test considers both extreme values that are much smaller than the rest of the data (case 1) and extreme values that are much larger than the rest of the data (case 2). This test is valid only if the data without the suspected outlier are normally distributed. Rosner's Test is a parametric test that is used to detect outliers for sample sizes of 25 or more. This test also assumes that the data without the suspected outliers are normally distributed.
3. Scientifically review statistical outliers and decide on their disposition. The review should include an evaluation of any notable trends in the data that may indicate the outliers represent true extreme values.

There were no potential outliers identified, and the data for this RIN are acceptable as qualified.

Report Prepared By: _____

Samantha Tigar
Data Validator

Data Validation Outliers Report - No Field Parameters
Comparison: All historical Data Beginning 1/1/2006
Laboratory: GEL Laboratories
RIN: 16017608
Report Date: 4/26/2016

					Current			**Historical Maximum**			**Historical Minimum**			**Number of Data Points**		**Statistical Outlier**
						Qualifiers			*Qualifiers*			*Qualifiers*				
Site Code	Location Code	Sample ID	Sample Date	Analyte	Result	*Lab*	*Data*	Result	*Lab*	*Data*	Result	*Lab*	*Data*	N	N Below Detect	
MND01	0411	N002	01/25/2016	cis-1,2-Dichloroethene	0.650	J		4.71		F	1.33			24	0	No
MND01	0411	N001	01/25/2016	cis-1,2-Dichloroethene	0.650	J		4.71		F	1.33			24	0	No

STATISTICAL TESTS:
The distribution of the data is tested for normality or lognormality using the Shapiro-Wilk Test
Outliers are identified using Dixon's Test when there are 25 or fewer data points.
Outliers are identified using Rosner's Test when there are 26 or more data points.
See Data Quality Assessment: Statistical Methods for Practitioners, EPA QC/G-9S, February 2006.

SAMPLE MANAGEMENT SYSTEM

General Data Validation Report

RIN: 16017608 **Lab Code:** GEN **Validator:** Samantha Tigar **Validation Date:** 4/26/2016

Project: Mound LTS&M Ground Water **Analysis Type:** ☐ Metals ☐ General Chem ☐ Rad ☑ Organics

# of Samples: 11 **Matrix:** Water **Requested Analysis Completed:** Yes

Chain of Custody

Present: OK **Signed:** OK **Dated:** OK

Sample

Integrity: OK **Preservation:** OK **Temperature:** OK

Select Quality Parameters

☑ Holding Times	All analyses were completed within the applicable holding times.
☑ Detection Limits	The reported detection limits are equal to or below contract requirements.
☑ Field/Trip Blanks	There was 1 trip/equipment blank evaluated.
☑ Field Duplicates	There was 1 duplicate evaluated.

SAMPLE MANAGEMENT SYSTEM

Validation Report: Equipment/Trip Blanks

RIN: 16017608 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 4/26/2016

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	390190008	EPA 8260 Low Level	Acetone (Lab Contaminant)	3.51	J	0.500	ug/L

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
390190001	OCU 596	0353	0.500	1.00	U	
390190002	OCU 597	0400	0.500	1.00	U	
390190003	OCU 598	0402	0.500	1.00	U	
390190004	OCU 599	0411	0.500	1.00	U	
390190005	OCU 600	0443	0.500	1.00	U	
390190006	OCU 601	0444	0.500	1.00	U	
390190007	OCU 602	0445	0.500	1.00	U	
390190009	OCU 604	9411	0.500	1.00	U	
390190010	OCU 605	P033	0.500	1.00	U	
390190011	OCU 603	0617	0.500	1.00	U	

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	390190008	EPA 8260 Low Level	Chloroform	0.200	J	0.160	ug/L

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
390190001	OCU 596	0353	0.160	1.00	U	
390190002	OCU 597	0400	0.160	1.00	U	
390190003	OCU 598	0402	0.160	1.00	U	
390190004	OCU 599	0411	0.160	1.00	U	
390190005	OCU 600	0443	0.160	1.00	U	
390190006	OCU 601	0444	0.160	1.00	U	
390190007	OCU 602	0445	0.160	1.00	U	
390190009	OCU 604	9411	0.160	1.00	U	
390190010	OCU 605	P033	0.160	1.00	U	
390190011	OCU 603	0617	0.160	1.00	U	

SAMPLE MANAGEMENT SYSTEM

Validation Report: Equipment/Trip Blanks

RIN: 16017608 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 4/26/2016

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	390190008	EPA 8260 Low Level	Chloromethane	0.230	J	0.160	ug/L

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
390190001	OCU 596	0353	0.160	1.00	U	
390190002	OCU 597	0400	0.280	1.00	J	U
390190003	OCU 598	0402	0.160	1.00	U	
390190004	OCU 599	0411	0.160	1.00	U	
390190005	OCU 600	0443	0.160	1.00	U	
390190006	OCU 601	0444	0.160	1.00	U	
390190007	OCU 602	0445	0.160	1.00	U	
390190009	OCU 604	9411	0.160	1.00	U	
390190010	OCU 605	P033	0.160	1.00	U	
390190011	OCU 603	0617	0.220	1.00	J	U

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	390190008	EPA 8260 Low Level	Methylene Chloride (Lab Contaminant)	0.260	BJ	0.160	ug/L

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
390190001	OCU 596	0353	0.180	1.00	BJ	U
390190002	OCU 597	0400	0.170	1.00	BJ	U
390190003	OCU 598	0402	0.160	1.00	U	
390190004	OCU 599	0411	0.160	1.00	U	
390190005	OCU 600	0443	0.160	1.00	U	
390190006	OCU 601	0444	0.160	1.00	U	
390190007	OCU 602	0445	0.180	1.00	BJ	U
390190009	OCU 604	9411	0.160	1.00	U	
390190010	OCU 605	P033	0.170	1.00	BJ	U
390190011	OCU 603	0617	0.160	1.00	U	

SAMPLE MANAGEMENT SYSTEM

Validation Report: Equipment/Trip Blanks

RIN: 16017608 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 4/26/2016

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	390190008	EPA 8260 Low Level	Toluene (Lab Contaminant)	0.170	J	0.160	ug/L

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
390190001	OCU 596	0353	0.160	1.00	U	
390190002	OCU 597	0400	0.160	1.00	U	
390190003	OCU 598	0402	0.160	1.00	U	
390190004	OCU 599	0411	0.160	1.00	U	
390190005	OCU 600	0443	0.160	1.00	U	
390190006	OCU 601	0444	0.160	1.00	U	
390190007	OCU 602	0445	0.160	1.00	U	
390190009	OCU 604	9411	0.160	1.00	U	
390190010	OCU 605	P033	0.160	1.00	U	
390190011	OCU 603	0617	0.160	1.00	U	

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	390190008	EPA 8260 Low Level	Vinyl Chloride	0.300	J	0.160	ug/L

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
390190001	OCU 596	0353	0.160	1.00	U	
390190002	OCU 597	0400	0.160	1.00	U	
390190003	OCU 598	0402	0.160	1.00	U	
390190004	OCU 599	0411	0.160	1.00	U	
390190005	OCU 600	0443	0.160	1.00	U	
390190006	OCU 601	0444	0.160	1.00	U	
390190007	OCU 602	0445	0.160	1.00	U	
390190009	OCU 604	9411	0.160	1.00	U	
390190010	OCU 605	P033	0.160	1.00	U	

SAMPLE MANAGEMENT SYSTEM

Validation Report: Field Duplicates

RIN: 16017608 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 4/26/2016

Duplicate: 9411 **Sample:** 0411

	Sample				Duplicate						
Analyte	Result	Flag	Error	Dilution	Result	Flag	Error	Dilution	RPD	RER	Units
1,1,1,2-Tetrachloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1,1-Trichloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1,2,2-Tetrachloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1,2-Trichloro-1,2,2-trifluoroethane	1.50	U		1.00	1.50	U		1.00			ug/L
1,1,2-Trichloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1-Dichloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1-Dichloroethene	0.160	U		1.00	0.160	U		1.00			ug/L
1,1-Dichloropropene	0.160	U		1.00	0.160	U		1.00			ug/L
1,2,3-Trichlorobenzene	0.200	U		1.00	0.200	U		1.00			ug/L
1,2,3-Trichloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
1,2,4-Trichlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,2,4-Trimethylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dibromo-3-chloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dibromoethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dichlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dichloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dichloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
1,3,5-Trimethylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,3-Dichlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,3-Dichloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
1,4-Dichlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
2,2-Dichloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
2-Butanone	0.500	U		1.00	0.500	U		1.00			ug/L
2-Chlorotoluene	0.160	U		1.00	0.160	U		1.00			ug/L
2-Hexanone	0.500	U		1.00	0.500	U		1.00			ug/L
4-Chlorotoluene	0.160	U		1.00	0.160	U		1.00			ug/L
4-Isopropyltoluene	0.160	U		1.00	0.160	U		1.00			ug/L
4-Methyl-2-pentanone	0.500	U		1.00	0.500	U		1.00			ug/L
Acetone	0.500	U		1.00	0.500	U		1.00			ug/L
Benzene	0.160	U		1.00	0.160	U		1.00			ug/L
Bromobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Bromochloromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Bromodichloromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Bromoform	0.160	U		1.00	0.160	U		1.00			ug/L
Bromomethane	0.160	U		1.00	0.160	U		1.00			ug/L
Carbon Disulfide	0.500	U		1.00	0.500	U		1.00			ug/L
Carbon Tetrachloride	0.160	U		1.00	0.160	U		1.00			ug/L
Chlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Chloroethane	0.160	U		1.00	0.160	U		1.00			ug/L

SAMPLE MANAGEMENT SYSTEM

Validation Report: Field Duplicates

RIN: 16017608 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 4/26/2016

Duplicate: 9411 **Sample:** 0411

	Sample				Duplicate						
Analyte	Result	Flag	Error	Dilution	Result	Flag	Error	Dilution	RPD	RER	Units
Chloroform	0.160	U		1.00	0.160	U		1.00			ug/L
Chloromethane	0.160	U		1.00	0.160	U		1.00			ug/L
cis-1,2-Dichloroethene	0.650	J		1.00	0.650	J		1.00			ug/L
cis-1,3-Dichloropropene	0.160	U		1.00	0.160	U		1.00			ug/L
Dibromochloromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Dibromomethane	0.160	U		1.00	0.160	U		1.00			ug/L
Dichlorodifluoromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Ethylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Hexachlorobutadiene	0.160	U		1.00	0.160	U		1.00			ug/L
Isopropylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Methylene Chloride	0.160	U		1.00	0.160	U		1.00			ug/L
Naphthalene	0.160	U		1.00	0.160	U		1.00			ug/L
n-Butylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
n-Propylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
sec-Butylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Styrene	0.160	U		1.00	0.160	U		1.00			ug/L
tert-Butylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Tetrachloroethene	0.160	U		1.00	0.160	U		1.00			ug/L
Toluene	0.160	U		1.00	0.160	U		1.00			ug/L
Total Xylenes	0.160	U		1.00	0.160	U		1.00			ug/L
trans-1,2-Dichloroethene	0.160	U		1.00	0.160	U		1.00			ug/L
trans-1,3-Dichloropropene	0.160	U		1.00	0.160	U		1.00			ug/L
Trichloroethene	11.9			1.00	11.7			1.00	1.69		ug/L
Trichlorofluoromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Vinyl Chloride	0.160	U		1.00	0.160	U		1.00			ug/L

SAMPLE MANAGEMENT SYSTEM
Organics Data Validation Summary

RIN: 16017608 **Project:** Mound LTS&M Ground Water **Lab Code:** GEN **Validation Date:** 4/26/2016

LCS Recovery: All LCS recoveries were within the laboratory acceptance limits.

Method Blank(s): There was 1 method blank result above the MDL.

MS/MSD Recovery: There was 1 MS/MSD failure.

Surrogate Recovery: All surrogate recoveries were within the laboratory acceptance limits.

SAMPLE MANAGEMENT SYSTEM

Non-Compliance Report: MS/MSD Performance

RIN: 16017608 **Lab Code:** GEN

Project: Mound LTS&M Ground Water

Validation Date: 4/26/2016

MS/MSD	Date Analyzed	Method	Analyte	Recovery MS	Recovery MSD	Lower Limit	Upper Limit	MSD RPD	RPD Limit
OCU 599	02/03/2016	EPA 8260 Low Level	Trichlorofluoromethane	114.0	90.0	61.0	132.0	24.00	20.0

SAMPLE MANAGEMENT SYSTEM

Non-Compliance Report: Method Blanks

RIN: 16017608 **Lab Code:** GEN

Project: Mound LTS&M Ground Water

Validation Date: 4/26/2016

Method Blank	Date Analyzed	Method	Analyte	Result	Flag(s)	MDL
1203481265	2/3/2016	EPA 8260 Low Level	Methylene Chloride	0.26	J	0.16

Data Review and Validation Report

General Information

Requisition No. (RIN):	16047763
Sample Event:	April 25 – 26, 2016
Site(s):	Mound, Ohio; Groundwater
Laboratory:	GEL Laboratories, Charleston, South Carolina
Work Order No.:	396206
Analysis:	Organics
Validator:	Samantha Tigar
Review Date:	July 18, 2016

This validation was performed according to the *Environmental Procedures Catalog* (LMS/POL/S04325), *"Standard Practice for Validation of Environmental Data."* The procedure was applied at Level 2, Data Deliverables Verification. See attached Data Validation Worksheets for supporting documentation on the data review and validation. All analyses were successfully completed. The samples were prepared and analyzed using accepted procedures based on methods specified by line item code, which are listed in Table 1.

Table 1. Analytes and Methods

Analyte	Line Item Code	Prep Method	Analytical Method
Volatile Organics, VOA	VOA-A-007	SW-846 5030B	SW-846 8260 LL

Data Qualifier Summary

Analytical results were qualified as listed in Table 2. Refer to the sections below for an explanation of the data qualifiers applied.

Table 2. Data Qualifier Summary

Sample Number	Location	Analyte(s)	Flag	Reason
All	All	Volatiles	J	Received out of temperature range
396206013	0126	Acetone	U	Less than 10 times the trip blank
396206018	0315 duplicate	Acetone	U	Less than 10 times the trip blank

Sample Shipping/Receiving

GEL Laboratories in Charleston, South Carolina, received 20 water samples on April 28, 2016, accompanied by Chain of Custody (COC) forms. The COC forms were checked to confirm that all of the samples were listed with sample collection dates and times, and that signatures and dates were present indicating sample relinquishment and receipt. The COC forms were complete with no errors or omissions. The air waybill numbers were listed on the receiving documentation.

Preservation and Holding Times

The sample shipment was received intact with the temperature inside the iced cooler at 8 °C, which does not comply with requirements. All samples were qualified with a "J" flag as estimated. All samples were received in the correct container types and had been preserved correctly for the requested analyses. Sample analysis was completed within the applicable holding times.

Detection and Quantitation Limits

The method detection limit (MDL) was reported for all analytes as required. The MDL, as defined in 40 CFR 136, is the minimum concentration of an analyte that can be measured and reported with 99 percent confidence that the analyte concentration is greater than zero. The practical quantitation limit (PQL) for these analytes is the lowest concentration that can be reliably measured, and is defined as 5 times the MDL. The reported MDLs for all analytes demonstrate compliance with contractual requirements.

Laboratory Instrument Calibration

Compliance requirements for satisfactory instrument calibration are established to ensure that the instrument is capable of producing acceptable qualitative and quantitative data for all analytes. Initial calibration demonstrates that the instrument is capable of acceptable performance in the beginning of the analytical run. Compliance requirements for continuing calibration checks are established to ensure that the instrument continues to be capable of producing acceptable qualitative and quantitative data. All laboratory instrument calibrations were performed correctly in accordance with the cited methods. All calibration and laboratory spike standards were prepared from independent sources.

The Continuing Calibration Verification standards associated with the samples exhibited percent drift values greater than 20 percent for several target compounds. Acetone was the only compound, associated with these calibration verification compounds, with results greater than the MDL but was previously qualified.

Volatiles Internal Standards and Surrogates

The volatile internal standard recoveries and surrogate recoveries were within the acceptance ranges for all samples.

Method Blanks

Method blanks are analyzed to assess any contamination that may have occurred during sample preparation. All method blank results associated with the samples were below the PQL for all analytes. In cases where a blank concentration exceeds or equals the MDL, the associated sample results are qualified with a “U” flag (not detected) when the sample result is greater than the MDL but less than 5 times the blank concentration.

Trip Blank

Trip blanks were prepared and analyzed to document contamination attributable to shipping and field handling procedures. This type of blank is useful in documenting contamination of volatile organic samples. Acetone and vinyl chloride were detected in some trip blanks. Sample results for these compounds that are less than 5 times the trip blank concentration (and less than 10 times the blank concentration for common laboratory contaminants) are qualified with a “U” flag as not detected.

Matrix Spike Analysis

Matrix spike and matrix spike duplicate (MS/MSD) samples are used to measure method performance in the sample matrix. The spike recoveries met the acceptance criteria for all analytes evaluated.

Laboratory Replicate Analysis

Laboratory replicate analyses are used to determine laboratory precision for each sample matrix. The relative percent difference (RPD) for organic replicate results should be less than the laboratory-derived control limits. All replicate results met these criteria.

Laboratory Control Sample

Laboratory control samples were analyzed at the correct frequency to provide information on the accuracy of the analytical method and the overall laboratory performance, including sample preparation. All control sample results were acceptable with one exception. The recovery for acetone was above acceptance criteria but was not detected in the associated samples and no qualification was necessary.

Field Duplicate

Field duplicate samples are collected and analyzed as an indication of overall precision of the measurement process. The precision observed includes both field and laboratory precision and has more variability than laboratory duplicates, which measure only laboratory performance. A duplicate sample was collected from location 0315. The relative percent difference for duplicate results that are greater than 5 times the PQL should be less than 20 percent. For results that are less than 5 times the PQL, the range should be no greater than the PQL. The duplicate results met the criteria.

Completeness

Results were reported in the correct units for all analytes requested using contract-required laboratory qualifiers.

Electronic Data Deliverable (EDD) File

The EDD file arrived on May 25, 2016. The Sample Management System EDD validation module was used to verify that the EDD file was complete and in compliance with requirements. The module compares the contents of the file to the requested analyses to ensure all and only the requested data are delivered. The contents of the EDD were manually examined to verify that the sample results accurately reflect the data contained in the sample data package.

Sampling Protocol

Sample results for all monitoring wells met the Category I low-flow sampling criteria and were qualified with an “F” flag in the database, indicating the wells were purged and sampled using the low-flow sampling method.

The turbidity value for locations 0138, 0315, and 0346 were greater than 50 NTU.

Outliers Report

Potential outliers are measurements that are extremely large or small relative to the rest of the data and, therefore, are suspected of misrepresenting the population from which they were collected. Potential outliers may result from transcription errors, data-coding errors, or measurement system problems. However, outliers may also represent true extreme values of a distribution and indicate more variability in the population than was expected.

Statistical outlier tests give probabilistic evidence that an extreme value does not "fit" with the distribution of the remainder of the data and is therefore a statistical outlier. These tests should only be used to identify data points that require further investigation. The tests alone cannot determine whether a statistical outlier should be discarded or corrected within a data set.

There are three steps involved in identifying extreme values or outliers:

1. Identify extreme values that may be potential outliers by generating the Outliers Report using the Sample Management System from data in the environmental database. The application compares the new data set (in standard environmental database units) with historical data and lists the new data that fall outside the historical data range. A determination is also made if the data are normally distributed using the Shapiro-Wilk Test.
2. Apply the appropriate statistical test. Dixon's Extreme Value test is used to test for statistical outliers when the sample size is less than or equal to 25. This test considers both extreme values that are much smaller than the rest of the data (case 1) and extreme values that are much larger than the rest of the data (case 2). This test is valid only if the data without the suspected outlier are normally distributed. Rosner's Test is a parametric

test that is used to detect outliers for sample sizes of 25 or more. This test also assumes that the data without the suspected outliers are normally distributed.

3. Scientifically review statistical outliers and decide on their disposition. The review should include an evaluation of any notable trends in the data that may indicate the outliers represent true extreme values.

No results from this sampling event were identified as potential outliers.

Report Prepared By: _____

Samantha Tigar
Data Validator

Data Validation Outliers Report - No Field Parameters
Comparison: All historical Data Beginning 1/1/2006
Laboratory: GEL Laboratories
RIN: 16047763
Report Date: 7/18/2016

					Current			Historical Maximum			Historical Minimum			Number of Data Points		Statistical Outlier
						Qualifiers			*Qualifiers*			*Qualifiers*				
Site Code	Location Code	Sample ID	Sample Date	Analyte	Result	*Lab*	*Data*	Result	*Lab*	*Data*	Result	*Lab*	*Data*	N	N Below Detect	
MND01	0347	N001	04/26/2016	Chloroform	0.210	JQ		5.00	U		0.240	J	FQ	51	10	NA
MND01	0602	N001	04/25/2016	cis-1,2-Dichloroethene	4.54	Q		42.3			4.56			26	0	No

STATISTICAL TESTS:
The distribution of the data is tested for normality or lognormality using the Shapiro-Wilk Test
Outliers are identified using Dixon's Test when there are 25 or fewer data points.
Outliers are identified using Rosner's Test when there are 26 or more data points.
See Data Quality Assessment: Statistical Methods for Practitioners, EPA QC/G-9S, February 2006.

SAMPLE MANAGEMENT SYSTEM

General Data Validation Report

RIN: 16047763 **Lab Code:** GEN **Validator:** Samantha Tigar **Validation Date:** 7/18/2016

Project: Mound LTS&M Ground Water **Analysis Type:** ☐ Metals ☐ General Chem ☐ Rad ☑ Organics

# of Samples: 20 **Matrix:** Water **Requested Analysis Completed:** Yes

Chain of Custody

Present: OK **Signed:** OK **Dated:** OK

Sample

Integrity: OK **Preservation:** OK **Temperature:** OK

Select Quality Parameters

☑ Holding Times	All analyses were completed within the applicable holding times.
☑ Detection Limits	The reported detection limits are equal to or below contract requirements.
☑ Field/Trip Blanks	There were 2 trip/equipment blanks evaluated.
☑ Field Duplicates	There was 1 duplicate evaluated.

SAMPLE MANAGEMENT SYSTEM

Validation Report: Equipment/Trip Blanks

RIN: 16047763 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 7/18/2016

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	396206006	EPA 8260 Low Level	Vinyl Chloride	0.390	JQ	0.160	ug/L

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
396206002	OFR 882	0386	0.160	1.00	QU	
396206003	OFR 883	0387	0.160	1.00	QU	
396206004	OFR 884	0389	0.160	1.00	QU	
396206005	OFR 885	0392	0.160	1.00	QU	
396206007	OFR 889	0118	0.160	1.00	QU	
396206008	OFR 890	0346	0.160	1.00	QU	

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	396206020	EPA 8260 Low Level	Acetone (Lab Contaminant)	2.17	JQ	0.500	ug/L

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
396206001	OFR 881	0138	0.500	1.00	QU	
396206009	OFR 891	0379	0.500	1.00	QU	
396206010	OFR 892	0347	0.500	1.00	QU	
396206011	OFR 893	0315	0.500	1.00	QU	
396206012	OFR 894	0124	0.500	1.00	QU	
396206013	OFR 895	0126	2.43	1.00	JQ	U
396206018	OFR 901	9315	2.09	1.00	JQ	U

SAMPLE MANAGEMENT SYSTEM

Validation Report: Field Duplicates

RIN: 16047763 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 7/18/2016

Duplicate: 9315 **Sample:** 0315

	Sample				Duplicate						
Analyte	Result	Flag	Error	Dilution	Result	Flag	Error	Dilution	RPD	RER	Units
1,1,1,2-Tetrachloroethane	0.160	QU		1.00	0.160	QU		1.00			ug/L
1,1,1-Trichloroethane	0.160	QU		1.00	0.160	QU		1.00			ug/L
1,1,2,2-Tetrachloroethane	0.160	QU		1.00	0.160	QU		1.00			ug/L
1,1,2-Trichloro-1,2,2-trifluoroethane	1.50	QU		1.00	1.50	QU		1.00			ug/L
1,1,2-Trichloroethane	0.160	QU		1.00	0.160	QU		1.00			ug/L
1,1-Dichloroethane	0.160	QU		1.00	0.160	QU		1.00			ug/L
1,1-Dichloroethene	0.160	QU		1.00	0.160	QU		1.00			ug/L
1,1-Dichloropropene	0.160	QU		1.00	0.160	QU		1.00			ug/L
1,2,3-Trichlorobenzene	0.200	QU		1.00	0.200	QU		1.00			ug/L
1,2,3-Trichloropropane	0.160	QU		1.00	0.160	QU		1.00			ug/L
1,2,4-Trichlorobenzene	0.160	QU		1.00	0.160	QU		1.00			ug/L
1,2,4-Trimethylbenzene	0.160	QU		1.00	0.160	QU		1.00			ug/L
1,2-Dibromo-3-chloropropane	0.160	QU		1.00	0.160	QU		1.00			ug/L
1,2-Dibromoethane	0.160	QU		1.00	0.160	QU		1.00			ug/L
1,2-Dichlorobenzene	0.160	QU		1.00	0.160	QU		1.00			ug/L
1,2-Dichloroethane	0.160	QU		1.00	0.160	QU		1.00			ug/L
1,2-Dichloropropane	0.160	QU		1.00	0.160	QU		1.00			ug/L
1,3,5-Trimethylbenzene	0.160	QU		1.00	0.160	QU		1.00			ug/L
1,3-Dichlorobenzene	0.160	QU		1.00	0.160	QU		1.00			ug/L
1,3-Dichloropropane	0.160	QU		1.00	0.160	QU		1.00			ug/L
1,4-Dichlorobenzene	0.160	QU		1.00	0.160	QU		1.00			ug/L
2,2-Dichloropropane	0.160	QU		1.00	0.160	QU		1.00			ug/L
2-Butanone	0.500	QU		1.00	0.500	QU		1.00			ug/L
2-Chlorotoluene	0.160	QU		1.00	0.160	QU		1.00			ug/L
2-Hexanone	0.500	QU		1.00	0.500	QU		1.00			ug/L
4-Chlorotoluene	0.160	QU		1.00	0.160	QU		1.00			ug/L
4-Isopropyltoluene	0.160	QU		1.00	0.160	QU		1.00			ug/L
4-Methyl-2-pentanone	0.500	QU		1.00	0.500	QU		1.00			ug/L
Acetone	0.500	QU		1.00	2.09	JQ		1.00			ug/L
Benzene	0.160	QU		1.00	0.160	QU		1.00			ug/L
Bromobenzene	0.160	QU		1.00	0.160	QU		1.00			ug/L
Bromochloromethane	0.160	QU		1.00	0.160	QU		1.00			ug/L
Bromodichloromethane	0.160	QU		1.00	0.160	QU		1.00			ug/L
Bromoform	0.160	QU		1.00	0.160	QU		1.00			ug/L
Bromomethane	0.160	QU		1.00	0.160	QU		1.00			ug/L
Carbon Disulfide	0.500	QU		1.00	0.500	QU		1.00			ug/L
Carbon Tetrachloride	0.620	JQ		1.00	0.570	JQ		1.00			ug/L
Chlorobenzene	0.160	QU		1.00	0.160	QU		1.00			ug/L
Chloroethane	0.160	QU		1.00	0.160	QU		1.00			ug/L

SAMPLE MANAGEMENT SYSTEM

Validation Report: Field Duplicates

RIN: 16047763 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 7/18/2016

Duplicate: 9315 **Sample:** 0315

	Sample				Duplicate						
Analyte	Result	Flag	Error	Dilution	Result	Flag	Error	Dilution	RPD	RER	Units
Chloroform	0.160	QU		1.00	0.160	QU		1.00			ug/L
Chloromethane	0.160	QU		1.00	0.160	QU		1.00			ug/L
cis-1,2-Dichloroethene	0.160	QU		1.00	0.160	QU		1.00			ug/L
cis-1,3-Dichloropropene	0.160	QU		1.00	0.160	QU		1.00			ug/L
Dibromochloromethane	0.160	QU		1.00	0.160	QU		1.00			ug/L
Dibromomethane	0.160	QU		1.00	0.160	QU		1.00			ug/L
Dichlorodifluoromethane	0.160	QU		1.00	0.160	QU		1.00			ug/L
Ethylbenzene	0.160	QU		1.00	0.160	QU		1.00			ug/L
Hexachlorobutadiene	0.160	QU		1.00	0.160	QU		1.00			ug/L
Isopropylbenzene	0.160	QU		1.00	0.160	QU		1.00			ug/L
Methylene Chloride	0.160	QU		1.00	0.160	QU		1.00			ug/L
Naphthalene	0.160	QU		1.00	0.160	QU		1.00			ug/L
n-Butylbenzene	0.160	QU		1.00	0.160	QU		1.00			ug/L
n-Propylbenzene	0.160	QU		1.00	0.160	QU		1.00			ug/L
sec-Butylbenzene	0.160	QU		1.00	0.160	QU		1.00			ug/L
Styrene	0.160	QU		1.00	0.160	QU		1.00			ug/L
tert-Butylbenzene	0.160	QU		1.00	0.160	QU		1.00			ug/L
Tetrachloroethene	0.160	QU		1.00	0.160	QU		1.00			ug/L
Toluene	0.160	QU		1.00	0.160	QU		1.00			ug/L
Total Xylenes	0.160	QU		1.00	0.160	QU		1.00			ug/L
trans-1,2-Dichloroethene	0.160	QU		1.00	0.160	QU		1.00			ug/L
trans-1,3-Dichloropropene	0.160	QU		1.00	0.160	QU		1.00			ug/L
Trichloroethene	5.67	Q		1.00	5.43	Q		1.00	4.32		ug/L
Trichlorofluoromethane	0.160	QU		1.00	0.160	QU		1.00			ug/L
Vinyl Chloride	0.160	QU		1.00	0.160	QU		1.00			ug/L

SAMPLE MANAGEMENT SYSTEM
Organics Data Validation Summary

RIN: 16047763 **Project:** Mound LTS&M Ground Water **Lab Code:** GEN **Validation Date:** 7/18/2016

LCS Recovery: There was 1 LCS failure.

Method Blank(s): All method blanks results were below the method detection limit.

MS/MSD Recovery: All MS/MSD recoveries were within the laboratory acceptance limits.

Surrogate Recovery: All surrogate recoveries were within the laboratory acceptance limits.

SAMPLE MANAGEMENT SYSTEM

Non-Compliance Report: LCS Recovery

RIN: 16047763 **Lab Code:** GEN

Project: Mound LTS&M Ground Water

Validation Date: 7/18/2016

LCS/LCSD	Date Analyzed	Method	Analyte	Recovery	Lower Limit	Upper Limit
LCS	05/04/2016	EPA 8260 Low Level	Acetone	141.0	62.0	140.0

Data Review and Validation Report

General Information

Requisition No. (RIN):	16077934
Sample Event:	July 25-28, 2016
Site(s):	Mound, Ohio; Groundwater
Laboratory:	GEL Laboratories, Charleston, South Carolina
Work Order No.:	402684
Analysis:	Organics and Radiochemistry
Validator:	Gretchen Baer
Review Date:	December 19, 2016

This validation was performed according to the *Environmental Procedures Catalog* (LMS/POL/S04325), *"Standard Practice for Validation of Environmental Data."* The procedure was applied at Level 2, Data Deliverables Verification. See attached Data Validation Worksheets for supporting documentation on the data review and validation. All analyses were successfully completed. The samples were prepared and analyzed using accepted procedures based on methods specified by line item code, which are listed in Table 1.

Table 1. Analytes and Methods

Analyte	Line Item Code	Prep Method	Analytical Method
Tritium	LSC-A-001	EPA 906.0 Mod	EPA 906.0 Mod
Volatile Organics, VOA	VOA-A-007	SW-846 5030B	SW-846 8260 LL

Data Qualifier Summary

Analytical results were qualified as listed in Table 2. Refer to the sections below for an explanation of the data qualifiers applied.

Table 2. Data Qualifier Summary

Sample Number	Location	Analyte(s)	Flag	Reason
402684-001	0346	Tritium	J	Less than the Determination Limit
402684-002	0379	Tritium	J	Less than the Determination Limit
402684-004	0379 Dup	Tritium	J	Less than the Determination Limit
402684-019	0138	Tritium	J	Less than the Determination Limit

Sample Shipping/Receiving

GEL Laboratories in Charleston, South Carolina, received 20 water samples on July 29, 2016, accompanied by Chain of Custody (COC) forms. The COC forms were checked to confirm that all of the samples were listed with sample collection dates and times, and that signatures and dates were present indicating sample relinquishment and receipt. The COC forms were complete with no errors or omissions. The air waybill numbers were listed on the receiving documentation.

Preservation and Holding Times

The sample shipment was received intact with the temperature inside the iced cooler at 2 °C, which complies with requirements. All samples were received in the correct container types and had been preserved correctly for the requested analyses. Sample analysis was completed within the applicable holding times.

Detection and Quantitation Limits

The method detection limit (MDL) was reported for all organic analytes as required. The MDL, as defined in 40 CFR 136, is the minimum concentration of an analyte that can be measured and reported with 99 percent confidence that the analyte concentration is greater than zero. The practical quantitation limit (PQL) for these analytes is the lowest concentration that can be reliably measured, and is defined as 5 times the MDL.

For radiochemical analytes (those measured by radiometric counting) the MDL and PQL are not applicable, and these results are evaluated using the minimum detectable concentration (MDC), Decision Level Concentration (DLC), and Determination Limit (DL). The MDC is a measure of radiochemical method performance and was calculated and reported as specified in the *Quality Systems Manual*. The DLC is the minimum concentration of an analyte that can be measured and reported with 99 percent confidence that the analyte concentration is greater than zero, and is estimated as 3 times the one-sigma total propagated uncertainty. Results that are greater than the MDC, but less than the DLC are qualified with a “U” flag (not detected). The DL for radiochemical results is the lowest concentration that can be reliably measured, and is defined as 3 times the MDC. Results not previously “U” qualified that are less than the DL are qualified with a “J” flag as estimated values.

The reported MDLs for all organic analytes and MDCs for radiochemical analytes demonstrate compliance with contractual requirements.

Laboratory Instrument Calibration

Compliance requirements for satisfactory instrument calibration are established to ensure that the instrument is capable of producing acceptable qualitative and quantitative data for all analytes. Initial calibration demonstrates that the instrument is capable of acceptable performance in the beginning of the analytical run. Compliance requirements for continuing calibration checks are established to ensure that the instrument continues to be capable of producing acceptable qualitative and quantitative data. All laboratory instrument calibrations were performed correctly in accordance with the cited methods. All calibration and laboratory spike standards were prepared from independent sources.

The Continuing Calibration Verification standards associated with the samples exhibited percent difference values within acceptance criteria for all compounds.

Volatiles Internal Standards and Surrogates

The volatile internal standard recoveries and surrogate recoveries were within the acceptance ranges for all samples.

Method Blanks

Method blanks are analyzed to assess any contamination that may have occurred during sample preparation. The method blank results were below the MDL for all target organic compounds. The tritium method blank result was less than the DLC.

Trip Blank

Trip blanks were prepared and analyzed to document contamination attributable to shipping and field handling procedures. This type of blank is useful in documenting contamination of volatile organic samples. There were no target analytes detected in the trip blank at a concentration greater than the MDL.

Matrix Spike Analysis

Matrix spike and matrix spike duplicate (MS/MSD) samples are used to measure method performance in the sample matrix. The spike recoveries met the acceptance criteria for all analytes evaluated.

Laboratory Replicate Analysis

Laboratory replicate analyses are used to determine laboratory precision for each sample matrix. The relative percent difference (RPD) for organic replicate results should be less than the laboratory-derived control limits. For radiochemical measurements, the relative error ratio (the ratio of the absolute difference between the sample and duplicate results and the sum of the 1-sigma uncertainties) is used to evaluate duplicate results and should be less than 3. All replicate results met these criteria.

Laboratory Control Sample

Laboratory control samples were analyzed at the correct frequency to provide information on the accuracy of the analytical method and the overall laboratory performance, including sample preparation. All control sample results were acceptable.

Field Duplicate

Field duplicate samples are collected and analyzed as an indication of overall precision of the measurement process. The precision observed includes both field and laboratory precision and has more variability than laboratory duplicates, which measure only laboratory performance. Duplicate samples were collected from location 0379 (field duplicate ID 9379). For non-radiochemical measurements, the relative percent difference for duplicate results that are greater than 5 times the PQL should be less than 20 percent. For results less than 5 times the PQL, the range should be no greater than the PQL. For radiochemical measurements, the relative error ratio (the ratio of the absolute difference between the sample and duplicate results and the sum of the 1-sigma uncertainties) is used to evaluate duplicate results and should be less than 3. All duplicate results met these criteria, demonstrating acceptable precision.

Completeness

Results were reported in the correct units for all analytes requested using contract-required laboratory qualifiers.

Electronic Data Deliverable (EDD) File

The EDD file arrived on August 26, 2016. The Sample Management System EDD validation module was used to verify that the EDD file was complete and in compliance with requirements. The module compares the contents of the file to the requested analyses to ensure all and only the requested data are delivered. The contents of the EDD were manually examined to verify that the sample results accurately reflect the data contained in the sample data package.

Sampling Protocol

Sample results for all monitoring wells met the Category I low-flow sampling criteria and were qualified with an “F” flag in the database, indicating the wells were purged and sampled using the low-flow sampling method. The final turbidity value for location 0138 was greater than 50 NTU.

No field notes were provided for the seep samples.

Outliers Report

Potential outliers are measurements that are extremely large or small relative to the rest of the data and, therefore, are suspected of misrepresenting the population from which they were collected. Potential outliers may result from transcription errors, data-coding errors, or measurement system problems. However, outliers may also represent true extreme values of a distribution and indicate more variability in the population than was expected.

Statistical outlier tests give probabilistic evidence that an extreme value does not "fit" with the distribution of the remainder of the data and is therefore a statistical outlier. These tests should only be used to identify data points that require further investigation. The tests alone cannot determine whether a statistical outlier should be discarded or corrected within a data set.

There are three steps involved in identifying extreme values or outliers:

1. Identify extreme values that may be potential outliers by generating the Outliers Report using the Sample Management System from data in the environmental database. The application compares the new data set (in standard environmental database units) with historical data and lists the new data that fall outside the historical data range. A determination is also made if the data are normally distributed using the Shapiro-Wilk Test.
2. Apply the appropriate statistical test. Dixon's Extreme Value test is used to test for statistical outliers when the sample size is less than or equal to 25. This test considers both extreme values that are much smaller than the rest of the data (case 1) and extreme values that are much larger than the rest of the data (case 2). This test is valid only if the data without the suspected outlier are normally distributed. Rosner's Test is a parametric test that is used to detect outliers for sample sizes of 25 or more. This test also assumes that the data without the suspected outliers are normally distributed.
3. Scientifically review statistical outliers and decide on their disposition. The review should include an evaluation of any notable trends in the data that may indicate the outliers represent true extreme values.

No results from this sampling event were identified as potential outliers.

Report Prepared By: _____

Gretchen Baer
Data Validator

Data Validation Outliers Report - No Field Parameters
Comparison: All historical Data Beginning 1/1/2006
Laboratory: GEL Laboratories
RIN: 16077934
Report Date: 12/19/2016

					Current			Historical Maximum			Historical Minimum			Number of Data Points		Statistical Outlier
						Qualifiers			*Qualifiers*			*Qualifiers*				
Site Code	Location Code	Sample ID	Sample Date	Analyte	Result	*Lab*	*Data*	Result	*Lab*	*Data*	Result	*Lab*	*Data*	N	N Below Detect	
MND01	0347	N001	07/27/2016	Chloroform	0.160	U	F	5.00	U		0.210	JQ	JF	52	10	NA
MND01	0392	N001	07/27/2016	Tetrachloroethene	0.160	U	F	0.450	J		0.197	J		40	2	No
MND01	0605	N001	07/25/2016	cis-1,2-Dichloroethene	1.02			32.2			1.09			41	0	NA
MND01	0605	N001	07/25/2016	Tritium	4410			55800			6300			31	0	No

STATISTICAL TESTS:
The distribution of the data is tested for normality or lognormality using the Shapiro-Wilk Test
Outliers are identified using Dixon's Test when there are 25 or fewer data points.
Outliers are identified using Rosner's Test when there are 26 or more data points.
See Data Quality Assessment: Statistical Methods for Practitioners, EPA QC/G-9S, February 2006.

NA: Data are not normally or lognormally distributed.

SAMPLE MANAGEMENT SYSTEM
General Data Validation Report

RIN: 16077934 **Lab Code:** GEN **Validator:** Gretchen Baer **Validation Date:** 12/19/2016

Project: Mound LTS&M Ground Water **Analysis Type:** ☐ Metals ☐ General Chem ☑ Rad ☑ Organics

# of Samples: 20 **Matrix:** Water **Requested Analysis Completed:** Yes

Chain of Custody

Present: OK **Signed:** OK **Dated:** OK

Sample

Integrity: OK **Preservation:** OK **Temperature:** OK

Select Quality Parameters

☑ Holding Times	All analyses were completed within the applicable holding times.
☑ Detection Limits	The reported detection limits are equal to or below contract requirements.
☑ Field/Trip Blanks	There were 3 trip/equipment blanks evaluated.
☑ Field Duplicates	There was 1 duplicate evaluated.

SAMPLE MANAGEMENT SYSTEM

Validation Report: Field Duplicates

RIN: 16077934 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 12/19/2016

Duplicate: 9379 **Sample:** 0379

	Sample				Duplicate						
Analyte	Result	Flag	Error	Dilution	Result	Flag	Error	Dilution	RPD	RER	Units
1,1,1,2-Tetrachloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1,1-Trichloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1,2,2-Tetrachloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1,2-Trichloro-1,2,2-trifluoroethane	1.50	U		1.00	1.50	U		1.00			ug/L
1,1,2-Trichloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1-Dichloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1-Dichloroethene	0.160	U		1.00	0.160	U		1.00			ug/L
1,1-Dichloropropene	0.160	U		1.00	0.160	U		1.00			ug/L
1,2,3-Trichlorobenzene	0.200	U		1.00	0.200	U		1.00			ug/L
1,2,3-Trichloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
1,2,4-Trichlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,2,4-Trimethylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dibromo-3-chloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dibromoethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dichlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dichloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dichloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
1,3,5-Trimethylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,3-Dichlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,3-Dichloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
1,4-Dichlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
2,2-Dichloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
2-Butanone	0.500	U		1.00	0.500	U		1.00			ug/L
2-Chlorotoluene	0.160	U		1.00	0.160	U		1.00			ug/L
2-Hexanone	0.500	U		1.00	0.500	U		1.00			ug/L
4-Chlorotoluene	0.160	U		1.00	0.160	U		1.00			ug/L
4-Isopropyltoluene	0.160	U		1.00	0.160	U		1.00			ug/L
4-Methyl-2-pentanone	0.500	U		1.00	0.500	U		1.00			ug/L
Acetone	0.500	U		1.00	0.500	U		1.00			ug/L
Benzene	0.160	U		1.00	0.160	U		1.00			ug/L
Bromobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Bromochloromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Bromodichloromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Bromoform	0.160	U		1.00	0.160	U		1.00			ug/L
Bromomethane	0.160	U		1.00	0.160	U		1.00			ug/L
Carbon Disulfide	0.500	U		1.00	0.500	U		1.00			ug/L
Carbon Tetrachloride	0.790	J		1.00	0.820	J		1.00			ug/L
Chlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Chloroethane	0.160	U		1.00	0.160	U		1.00			ug/L

SAMPLE MANAGEMENT SYSTEM

Validation Report: Field Duplicates

RIN: 16077934 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 12/19/2016

Duplicate: 9379 **Sample:** 0379

	Sample				Duplicate						
Analyte	Result	Flag	Error	Dilution	Result	Flag	Error	Dilution	RPD	RER	Units
Chloroform	0.160	U		1.00	0.160	U		1.00			ug/L
Chloromethane	0.160	U		1.00	0.160	U		1.00			ug/L
cis-1,2-Dichloroethene	0.160	U		1.00	0.160	U		1.00			ug/L
cis-1,3-Dichloropropene	0.160	U		1.00	0.160	U		1.00			ug/L
Dibromochloromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Dibromomethane	0.160	U		1.00	0.160	U		1.00			ug/L
Dichlorodifluoromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Ethylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Hexachlorobutadiene	0.160	U		1.00	0.160	U		1.00			ug/L
Isopropylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Methylene Chloride	0.160	U		1.00	0.160	U		1.00			ug/L
Naphthalene	0.160	U		1.00	0.160	U		1.00			ug/L
n-Butylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
n-Propylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
sec-Butylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Styrene	0.160	U		1.00	0.160	U		1.00			ug/L
tert-Butylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Tetrachloroethene	0.390	J		1.00	0.370	J		1.00			ug/L
Toluene	0.160	U		1.00	0.160	U		1.00			ug/L
Total Xylenes	0.160	U		1.00	0.160	U		1.00			ug/L
trans-1,2-Dichloroethene	0.160	U		1.00	0.160	U		1.00			ug/L
trans-1,3-Dichloropropene	0.160	U		1.00	0.160	U		1.00			ug/L
Trichloroethene	1.71			1.00	1.70			1.00	0.59		ug/L
Trichlorofluoromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Tritium	590		247	1.00	795		279	1.00		1.1	pCi/L
Vinyl Chloride	0.160	U		1.00	0.160	U		1.00			ug/L

SAMPLE MANAGEMENT SYSTEM
Organics Data Validation Summary

RIN: 16077934 **Project:** Mound LTS&M Ground Water **Lab Code:** GEN **Validation Date:** 12/19/2016

LCS Recovery: All LCS recoveries were within the laboratory acceptance limits.

Method Blank(s): All method blanks results were below the method detection limit.

MS/MSD Recovery: All MS/MSD recoveries were within the laboratory acceptance limits.

Surrogate Recovery: All surrogate recoveries were within the laboratory acceptance limits.

SAMPLE MANAGEMENT SYSTEM
Radiochemistry Data Validation Worksheet

RIN: 16077934 **Lab Code:** GEN **Date Due:** 8/26/2016

Matrix: Water **Site Code:** MND03 **Date Completed:** 8/26/2016

Sample	Analyte	Date Analyzed	Result	Flag	Tracer %R	LCS %R	MS %R	Duplicate RER
0379	Tritium	08/21/2016						0.13
Blank_Spike	Tritium	08/21/2016				87.60		
0379	Tritium	08/21/2016					96.3	
Blank	Tritium	08/21/2016	-60.0000	U				

Data Review and Validation Report

General Information

Requisition No. (RIN):	16077935
Sample Event:	July 25 – 26, 2016
Site(s):	Mound, Ohio; Groundwater
Laboratory:	GEL Laboratories, Charleston, South Carolina
Work Order No.:	402672
Analysis:	Organics
Validator:	Samantha Tigar
Review Date:	September 1, 2016

This validation was performed according to the *Environmental Procedures Catalog* (LMS/POL/S04325), *"Standard Practice for Validation of Environmental Data."* The procedure was applied at Level 2, Data Deliverables Verification. See attached Data Validation Worksheets for supporting documentation on the data review and validation. All analyses were successfully completed. The samples were prepared and analyzed using accepted procedures based on methods specified by line item code, which are listed in Table 1.

Table 1. Analytes and Methods

Analyte	Line Item Code	Prep Method	Analytical Method
Volatile Organics, VOA	VOA-A-007	SW-846 5030B	SW-846 8260 LL

Data Qualifier Summary

Analytical results were qualified as listed in Table 2. Refer to the sections below for an explanation of the data qualifiers applied.

Table 2. Data Qualifier Summary

Sample Number	Location	Analyte(s)	Flag	Reason
402672002	0402	Vinyl Chloride	U	Less than 10 times the trip blank

Sample Shipping/Receiving

GEL Laboratories in Charleston, South Carolina, received 12 water samples on July 29, 2016, accompanied by Chain of Custody (COC) forms. The COC forms were checked to confirm that all of the samples were listed with sample collection dates and times, and that signatures and dates were present indicating sample relinquishment and receipt. The COC forms were complete with no errors or omissions. The air waybill numbers were listed on the receiving documentation.

Preservation and Holding Times

The sample shipment was received intact with the temperature inside the iced cooler at 2 °C, which complies with requirements. All samples were received in the correct container types and had been preserved correctly for the requested analyses. Sample analysis was completed within the applicable holding times.

Detection and Quantitation Limits

The method detection limit (MDL) was reported for all analytes as required. The MDL, as defined in 40 CFR 136, is the minimum concentration of an analyte that can be measured and reported with 99 percent confidence that the analyte concentration is greater than zero. The practical quantitation limit (PQL) for these analytes is the lowest concentration that can be reliably measured, and is defined as 5 times the MDL. The reported MDLs for all analytes demonstrate compliance with contractual requirements.

Laboratory Instrument Calibration

Compliance requirements for satisfactory instrument calibration are established to ensure that the instrument is capable of producing acceptable qualitative and quantitative data for all analytes. Initial calibration demonstrates that the instrument is capable of acceptable performance in the beginning of the analytical run. Compliance requirements for continuing calibration checks are established to ensure that the instrument continues to be capable of producing acceptable qualitative and quantitative data. All laboratory instrument calibrations were performed correctly in accordance with the cited methods. All calibration and laboratory spike standards were prepared from independent sources.

The Continuing Calibration Verification standards associated with the samples exhibited percent drift values less than 20 percent for all target compounds.

Volatiles Internal Standards and Surrogates

The volatile internal standard recoveries and surrogate recoveries were within the acceptance ranges for all samples.

Method Blanks

Method blanks are analyzed to assess any contamination that may have occurred during sample preparation. All method blank results associated with the samples were below the PQL for all analytes. In cases where a blank concentration exceeds or equals the MDL, the associated sample results are qualified with a “U” flag (not detected) when the sample result is greater than the MDL but less than 5 times the blank concentration.

Trip Blank

Trip blanks were prepared and analyzed to document contamination attributable to shipping and field handling procedures. This type of blank is useful in documenting contamination of volatile organic samples. Vinyl chloride was detected in a trip blank. Sample results for compounds that are less than 5 times the trip blank concentration (and less than 10 times the blank concentration for common laboratory contaminants) are qualified with a “U” flag as not detected.

Matrix Spike Analysis

Matrix spike and matrix spike duplicate (MS/MSD) samples are used to measure method performance in the sample matrix. The spike recoveries did not met the acceptance criteria for 1,1,2,2-tetrachloroethane and bromomethane. The compounds were not detected in the associated samples, no qualification required.

Laboratory Replicate Analysis

Laboratory replicate analyses are used to determine laboratory precision for each sample matrix. The relative percent difference (RPD) for organic replicate results should be less than the laboratory-derived control limits. All replicate results met these criteria.

Laboratory Control Sample

Laboratory control samples were analyzed at the correct frequency to provide information on the accuracy of the analytical method and the overall laboratory performance, including sample preparation. All control sample results were acceptable.

Field Duplicate

Field duplicate samples are collected and analyzed as an indication of overall precision of the measurement process. The precision observed includes both field and laboratory precision and has more variability than laboratory duplicates, which measure only laboratory performance. A duplicate sample was collected from location 0411. The relative percent difference for duplicate results that are greater than 5 times the PQL should be less than 20 percent. For results that are less than 5 times the PQL, the range should be no greater than the PQL. The duplicate results met the criteria.

Completeness

Results were reported in the correct units for all analytes requested using contract-required laboratory qualifiers.

Electronic Data Deliverable (EDD) File

The EDD file arrived on August 26, 2016. The Sample Management System EDD validation module was used to verify that the EDD file was complete and in compliance with requirements. The module compares the contents of the file to the requested analyses to ensure all and only the requested data are delivered. The contents of the EDD were manually examined to verify that the sample results accurately reflect the data contained in the sample data package.

Sampling Protocol

Sample results for all monitoring wells met the Category I or II low-flow sampling criteria and were qualified with an “F” flag in the database, indicating the wells were purged and sampled using the low-flow sampling method. The groundwater sample results for wells 0353, 0443, 0444, and 0445 were further qualified with a “Q” flag in the database indicating the data are considered qualitative because these are Category II wells.

The turbidity value for locations 0400 and 0411 were greater than 50 NTU.

Outliers Report

Potential outliers are measurements that are extremely large or small relative to the rest of the data and, therefore, are suspected of misrepresenting the population from which they were collected. Potential outliers may result from transcription errors, data-coding errors, or measurement system problems. However, outliers may also represent true extreme values of a distribution and indicate more variability in the population than was expected.

Statistical outlier tests give probabilistic evidence that an extreme value does not "fit" with the distribution of the remainder of the data and is therefore a statistical outlier. These tests should only be used to identify data points that require further investigation. The tests alone cannot determine whether a statistical outlier should be discarded or corrected within a data set.

There are three steps involved in identifying extreme values or outliers:

1. Identify extreme values that may be potential outliers by generating the Outliers Report using the Sample Management System from data in the environmental database. The application compares the new data set (in standard environmental database units) with historical data and lists the new data that fall outside the historical data range. A determination is also made if the data are normally distributed using the Shapiro-Wilk Test.
2. Apply the appropriate statistical test. Dixon's Extreme Value test is used to test for statistical outliers when the sample size is less than or equal to 25. This test considers both extreme values that are much smaller than the rest of the data (case 1) and extreme

values that are much larger than the rest of the data (case 2). This test is valid only if the data without the suspected outlier are normally distributed. Rosner's Test is a parametric test that is used to detect outliers for sample sizes of 25 or more. This test also assumes that the data without the suspected outliers are normally distributed.

3. Scientifically review statistical outliers and decide on their disposition. The review should include an evaluation of any notable trends in the data that may indicate the outliers represent true extreme values.

No results from this sampling event were identified as potential outliers.

Report Prepared By: _____

Samantha Tigar
Data Validator

SAMPLE MANAGEMENT SYSTEM

General Data Validation Report

RIN: 16077935 **Lab Code:** GEN **Validator:** Samantha Tigar **Validation Date:** 9/1/2016

Project: Mound LTS&M Ground Water **Analysis Type:** ☐ Metals ☐ General Chem ☐ Rad ☑ Organics

# of Samples: 12 **Matrix:** Water **Requested Analysis Completed:** Yes

Chain of Custody

Present: OK **Signed:** OK **Dated:** OK

Sample

Integrity: OK **Preservation:** OK **Temperature:** OK

Select Quality Parameters

☑ Holding Times	All analyses were completed within the applicable holding times.
☑ Detection Limits	The reported detection limits are equal to or below contract requirements.
☑ Field/Trip Blanks	There were 2 trip/equipment blanks evaluated.
☑ Field Duplicates	There was 1 duplicate evaluated.

SAMPLE MANAGEMENT SYSTEM

Validation Report: Equipment/Trip Blanks

RIN: 16077935 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 9/1/2016

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	402672005	EPA 8260 Low Level	Vinyl Chloride	0.400	J	0.160	ug/L

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
402672001	OIZ 752	0400	0.160	1.00	U	
402672002	OIZ 753	0402	0.310	1.00	J	U
402672003	OIZ 754	0411	0.160	1.00	U	
402672004	OIZ 755	0443	0.160	1.00	U	
402672006	OIZ 759	9411	0.160	1.00	U	
402672007	OIZ 760	P033	0.160	1.00	U	

SAMPLE MANAGEMENT SYSTEM

Validation Report: Field Duplicates

RIN: 16077935 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 9/1/2016

Duplicate: 9411 **Sample:** 0411

	Sample				Duplicate						
Analyte	Result	Flag	Error	Dilution	Result	Flag	Error	Dilution	RPD	RER	Units
1,1,1,2-Tetrachloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1,1-Trichloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1,2,2-Tetrachloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1,2-Trichloro-1,2,2-trifluoroethane	1.50	U		1.00	1.50	U		1.00			ug/L
1,1,2-Trichloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1-Dichloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1-Dichloroethene	0.160	U		1.00	0.160	U		1.00			ug/L
1,1-Dichloropropene	0.160	U		1.00	0.160	U		1.00			ug/L
1,2,3-Trichlorobenzene	0.200	U		1.00	0.200	U		1.00			ug/L
1,2,3-Trichloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
1,2,4-Trichlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,2,4-Trimethylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dibromo-3-chloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dibromoethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dichlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dichloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dichloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
1,3,5-Trimethylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,3-Dichlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,3-Dichloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
1,4-Dichlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
2,2-Dichloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
2-Butanone	0.500	U		1.00	0.500	U		1.00			ug/L
2-Chlorotoluene	0.160	U		1.00	0.160	U		1.00			ug/L
2-Hexanone	0.500	U		1.00	0.500	U		1.00			ug/L
4-Chlorotoluene	0.160	U		1.00	0.160	U		1.00			ug/L
4-Isopropyltoluene	0.160	U		1.00	0.160	U		1.00			ug/L
4-Methyl-2-pentanone	0.500	U		1.00	0.500	U		1.00			ug/L
Acetone	0.500	U		1.00	0.500	U		1.00			ug/L
Benzene	0.160	U		1.00	0.160	U		1.00			ug/L
Bromobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Bromochloromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Bromodichloromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Bromoform	0.160	U		1.00	0.160	U		1.00			ug/L
Bromomethane	0.160	U		1.00	0.160	U		1.00			ug/L
Carbon Disulfide	0.500	U		1.00	0.500	U		1.00			ug/L
Carbon Tetrachloride	0.160	U		1.00	0.160	U		1.00			ug/L
Chlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Chloroethane	0.160	U		1.00	0.160	U		1.00			ug/L

SAMPLE MANAGEMENT SYSTEM

Validation Report: Field Duplicates

RIN: 16077935 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 9/1/2016

Duplicate: 9411 **Sample:** 0411

	Sample				Duplicate						
Analyte	Result	Flag	Error	Dilution	Result	Flag	Error	Dilution	RPD	RER	Units
Chloroform	0.160	U		1.00	0.160	U		1.00			ug/L
Chloromethane	0.160	U		1.00	0.160	U		1.00			ug/L
cis-1,2-Dichloroethene	0.940	J		1.00	0.880	J		1.00	NA		ug/L
cis-1,3-Dichloropropene	0.160	U		1.00	0.160	U		1.00			ug/L
Dibromochloromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Dibromomethane	0.160	U		1.00	0.160	U		1.00			ug/L
Dichlorodifluoromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Ethylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Hexachlorobutadiene	0.160	U		1.00	0.160	U		1.00			ug/L
Isopropylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Methylene Chloride	0.160	U		1.00	0.160	U		1.00			ug/L
Naphthalene	0.160	U		1.00	0.160	U		1.00			ug/L
n-Butylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
n-Propylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
sec-Butylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Styrene	0.160	U		1.00	0.160	U		1.00			ug/L
tert-Butylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Tetrachloroethene	0.160	U		1.00	0.160	U		1.00			ug/L
Toluene	0.160	U		1.00	0.160	U		1.00			ug/L
Total Xylenes	0.160	U		1.00	0.160	U		1.00			ug/L
trans-1,2-Dichloroethene	0.160	U		1.00	0.160	U		1.00			ug/L
trans-1,3-Dichloropropene	0.160	U		1.00	0.160	U		1.00			ug/L
Trichloroethene	11.7			1.00	11.8			1.00	0.85		ug/L
Trichlorofluoromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Vinyl Chloride	0.160	U		1.00	0.160	U		1.00			ug/L

SAMPLE MANAGEMENT SYSTEM
Organics Data Validation Summary

RIN: 16077935 **Project:** Mound LTS&M Ground Water **Lab Code:** GEN **Validation Date:** 9/1/2016

LCS Recovery: All LCS recoveries were within the laboratory acceptance limits.

Method Blank(s): All method blanks results were below the method detection limit.

MS/MSD Recovery: There were 2 MS/MSD failures.

Surrogate Recovery: All surrogate recoveries were within the laboratory acceptance limits.

SAMPLE MANAGEMENT SYSTEM

Non-Compliance Report: MS/MSD Performance

RIN: 16077935 **Lab Code:** GEN

Project: Mound LTS&M Ground Water

Validation Date: 9/1/2016

MS/MSD	Date Analyzed	Method	Analyte	Recovery MS	Recovery MSD	Lower Limit	Upper Limit	MSD RPD	RPD Limit
OIZ 755	08/03/2016	EPA 8260 Low Level	1,1,2,2-Tetrachloroethane	121.0	110.0	73.0	120.0		
OIZ 755	08/03/2016	EPA 8260 Low Level	Bromomethane	130.0	121.0	66.0	129.0		

Data Review and Validation Report

General Information

Requisition No. (RIN):	16108110
Sample Event:	October 24 – 26, 2016
Site(s):	Mound, Ohio; Groundwater
Laboratory:	GEL Laboratories, Charleston, South Carolina
Work Order No.:	409186
Analysis:	Organics
Validator:	Samantha Tigar
Review Date:	January 3, 2017

This validation was performed according to the *Environmental Procedures Catalog* (LMS/POL/S04325), *"Standard Practice for Validation of Environmental Data."* The procedure was applied at Level 2, Data Deliverables Verification. See attached Data Validation Worksheets for supporting documentation on the data review and validation. All analyses were successfully completed. The samples were prepared and analyzed using accepted procedures based on methods specified by line item code, which are listed in Table 1.

Table 1. Analytes and Methods

Analyte	Line Item Code	Prep Method	Analytical Method
Volatile Organics, VOA	VOA-A-007	SW-846 5030B	SW-846 8260 LL

Data Qualifier Summary

Analytical results were qualified as listed in Table 2. Refer to the sections below for an explanation of the data qualifiers applied.

Table 2. Data Qualifier Summary

Sample Number	Location	Analyte(s)	Flag	Reason
409186007	0605	Acetone	U	Less than 10 times the trip blank
409186009	0607	Acetone	U	Less than 10 times the trip blank

Sample Shipping/Receiving

GEL Laboratories in Charleston, South Carolina, received 21 water samples on October 27, 2016, accompanied by Chain of Custody (COC) forms. The air waybill number was listed on the

receiving documentation. The COC forms were checked to confirm that all of the samples were listed with sample collection dates and times, and that signatures and dates were present indicating sample relinquishment and receipt. The COC forms were complete with no errors or omissions, with the following exception. The receiving date on one of the COC forms was listed as the 26th instead of the 27th.

Preservation and Holding Times

The sample shipment was received intact with the temperature inside the iced cooler at 1 °C, which complies with requirements. All samples were received in the correct container types and had been preserved correctly for the requested analyses. Sample analysis was completed within the applicable holding times.

Detection and Quantitation Limits

The method detection limit (MDL) was reported for all organic analytes as required. The MDL, as defined in 40 CFR 136, is the minimum concentration of an analyte that can be measured and reported with 99 percent confidence that the analyte concentration is greater than zero. The practical quantitation limit (PQL) for these analytes is the lowest concentration that can be reliably measured, and is defined as 5 times the MDL. The reported MDLs for all analytes demonstrate compliance with contractual requirements.

Laboratory Instrument Calibration

Compliance requirements for satisfactory instrument calibration are established to ensure that the instrument is capable of producing acceptable qualitative and quantitative data for all analytes. Initial calibration demonstrates that the instrument is capable of acceptable performance in the beginning of the analytical run. Compliance requirements for continuing calibration checks are established to ensure that the instrument continues to be capable of producing acceptable qualitative and quantitative data. All laboratory instrument calibrations were performed correctly in accordance with the cited methods. All calibration and laboratory spike standards were prepared from independent sources.

The Continuing Calibration Verification standards associated with the samples exhibited percent difference values outside the acceptance criteria for several target compounds. All associated results were below the MDL, no qualification necessary.

Volatiles Internal Standards and Surrogates

The volatile internal standard recoveries and surrogate recoveries were within the acceptance ranges for all samples.

Method Blanks

Method blanks are analyzed to assess any contamination that may have occurred during sample preparation. The method blank results were below the MDL for all target organic compounds.

Trip Blank

Trip blanks were prepared and analyzed to document contamination attributable to shipping and field handling procedures. This type of blank is useful in documenting contamination of volatile organic samples. Three trip blanks were submitted with these samples. Acetone and chlorobenzene were detected in the trip blanks. All associated results greater than the MDL and less than 5 times the blank concentration (10 times for common laboratory contaminants) were qualified with a “U” flag as not detected.

Matrix Spike Analysis

Matrix spike and matrix spike duplicate (MS/MSD) samples are used to measure method performance in the sample matrix. The spike recoveries met the acceptance criteria for all analytes evaluated.

Laboratory Replicate Analysis

Laboratory replicate analyses are used to determine laboratory precision for each sample matrix. The relative percent difference (RPD) for organic replicate results should be less than the laboratory-derived control limits. All replicate results met these criteria.

Laboratory Control Sample

Laboratory control samples were analyzed at the correct frequency to provide information on the accuracy of the analytical method and the overall laboratory performance, including sample preparation. All control sample results were acceptable.

Field Duplicate

Field duplicate samples are collected and analyzed as an indication of overall precision of the measurement process. The precision observed includes both field and laboratory precision and has more variability than laboratory duplicates, which measure only laboratory performance. A duplicate sample was collected from location 0347. The relative percent difference for duplicate results that are greater than 5 times the PQL should be less than 20 percent. For results less than 5 times the PQL, the range should be no greater than the PQL. All duplicate results met these criteria, demonstrating acceptable precision.

Completeness

Results were reported in the correct units for all analytes requested using contract-required laboratory qualifiers.

Electronic Data Deliverable (EDD) File

The EDD file arrived on November 18, 2016. The Sample Management System EDD validation module was used to verify that the EDD file was complete and in compliance with requirements. The module compares the contents of the file to the requested analyses to ensure all and only the requested data are delivered. The contents of the EDD were manually examined to verify that the sample results accurately reflect the data contained in the sample data package.

Sampling Protocol

Sample results for all monitoring wells met the Category I low-flow sampling criteria and were qualified with an “F” flag in the database, indicating the wells were purged and sampled using the low-flow sampling method. The final turbidity value for location 0138 was greater than 50 NTU.

Outliers Report

Potential outliers are measurements that are extremely large or small relative to the rest of the data and, therefore, are suspected of misrepresenting the population from which they were collected. Potential outliers may result from transcription errors, data-coding errors, or measurement system problems. However, outliers may also represent true extreme values of a distribution and indicate more variability in the population than was expected.

Statistical outlier tests give probabilistic evidence that an extreme value does not "fit" with the distribution of the remainder of the data and is therefore a statistical outlier. These tests should only be used to identify data points that require further investigation. The tests alone cannot determine whether a statistical outlier should be discarded or corrected within a data set.

There are three steps involved in identifying extreme values or outliers:

1. Identify extreme values that may be potential outliers by generating the Outliers Report using the Sample Management System from data in the environmental database. The application compares the new data set (in standard environmental database units) with historical data and lists the new data that fall outside the historical data range. A determination is also made if the data are normally distributed using the Shapiro-Wilk Test.
2. Apply the appropriate statistical test. Dixon's Extreme Value test is used to test for statistical outliers when the sample size is less than or equal to 25. This test considers both extreme values that are much smaller than the rest of the data (case 1) and extreme values that are much larger than the rest of the data (case 2). This test is valid only if the data without the suspected outlier are normally distributed. Rosner's Test is a parametric test that is used to detect outliers for sample sizes of 25 or more. This test also assumes that the data without the suspected outliers are normally distributed.
3. Scientifically review statistical outliers and decide on their disposition. The review should include an evaluation of any notable trends in the data that may indicate the outliers represent true extreme values.

No results from this sampling event were identified as potential outliers.

Report Prepared By: _____

Samantha Tigar
Data Validator

Data Validation Outliers Report - No Field Parameters
Comparison: All historical Data Beginning 1/1/2010
Laboratory: GEL Laboratories
RIN: 16108110
Report Date: 1/3/2017

					Current			**Historical Maximum**			**Historical Minimum**			**Number of Data Points**		**Statistical Outlier**
						Qualifiers			*Qualifiers*			*Qualifiers*				
Site Code	Location Code	Sample ID	Sample Date	Analyte	Result	*Lab*	*Data*	Result	*Lab*	*Data*	Result	*Lab*	*Data*	N	N Below Detect	
MND01	0601	N001	10/24/2016	cis-1,2-Dichloroethene	1.16			0.940	J		0.1000	U		33	5	No
MND01	0601	N001	10/24/2016	Tetrachloroethene	19.1			14.9			0.160	U		33	4	No
MND01	0602	N001	10/24/2016	1,1-Dichloroethane	0.450	J		0.422	J		0.160	U		16	6	No
MND01	0607	N001	10/24/2016	Acetone	2.88	J		2.59	BJ	U	0.500	QU	J	29	29	NA

STATISTICAL TESTS:
The distribution of the data is tested for normality or lognormality using the Shapiro-Wilk Test
Outliers are identified using Dixon's Test when there are 25 or fewer data points.
Outliers are identified using Rosner's Test when there are 26 or more data points.
See Data Quality Assessment: Statistical Methods for Practitioners, EPA QC/G-9S, February 2006.

NA: Data are not normally or lognormally distributed.

SAMPLE MANAGEMENT SYSTEM

General Data Validation Report

RIN: 16108110 **Lab Code:** GEN **Validator:** Samantha Tigar **Validation Date:** 1/3/2017

Project: Mound LTS&M Ground Water **Analysis Type:** ☐ Metals ☐ General Chem ☐ Rad ☑ Organics

# of Samples: 21 **Matrix:** Water **Requested Analysis Completed:** Yes

Chain of Custody

Present: OK **Signed:** OK **Dated:** OK

Sample

Integrity: OK **Preservation:** OK **Temperature:** OK

Select Quality Parameters

☑ Holding Times	All analyses were completed within the applicable holding times.
☑ Detection Limits	The reported detection limits are equal to or below contract requirements.
☑ Field/Trip Blanks	There were 3 trip/equipment blanks evaluated.
☑ Field Duplicates	There was 1 duplicate evaluated.

SAMPLE MANAGEMENT SYSTEM

Validation Report: Equipment/Trip Blanks

RIN: 16108110 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 1/3/2017

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	409186004	EPA 8260 Low Level	Acetone (Lab Contaminant)	5.51	J	0.500	ug/L

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
409186001	OLV 796	0118	0.500	1.00	U	
409186002	OLV 801	0124	0.500	1.00	U	
409186003	OLV 802	0126	0.500	1.00	U	
409186005	OLV 805	0601	0.500	1.00	U	
409186006	OLV 808	0602	0.500	1.00	U	
409186007	OLV 803	0605	2.90	1.00	J	U
409186008	OLV 806	0606	0.500	1.00	U	
409186009	OLV 804	0607	2.88	1.00	J	U

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	409186016	EPA 8260 Low Level	Acetone (Lab Contaminant)	5.88	J	0.500	ug/L

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
409186010	OLV 797	0346	0.500	1.00	U	
409186011	OLV 798	0379	0.500	1.00	U	
409186012	OLV 791	0386	0.500	1.00	U	
409186013	OLV 792	0387	0.500	1.00	U	
409186014	OLV 793	0389	0.500	1.00	U	

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	409186016	EPA 8260 Low Level	Chlorobenzene	0.320	J	0.160	ug/L

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
409186010	OLV 797	0346	0.160	1.00	U	
409186011	OLV 798	0379	0.160	1.00	U	
409186012	OLV 791	0386	0.160	1.00	U	
409186013	OLV 792	0387	0.160	1.00	U	
409186014	OLV 793	0389	0.160	1.00	U	

SAMPLE MANAGEMENT SYSTEM

Validation Report: Equipment/Trip Blanks

RIN: 16108110 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 1/3/2017

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	409186020	EPA 8260 Low Level	Acetone (Lab Contaminant)	6.65	J	0.500	ug/L

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
409186017	OLV 790	0138	0.500	1.00	U	
409186018	OLV 800	0315	0.500	1.00	U	
409186019	OLV 799	0347	0.500	1.00	U	
409186021	OLV 807	9347	0.500	1.00	U	

Blank Data

Blank Type	Lab Sample ID	Lab Method	Analyte Name	Result	Qualifier	MDL	Units
Trip Blank	409186020	EPA 8260 Low Level	Chlorobenzene	0.310	J	0.160	ug/L

Sample ID	Sample Ticket	Location	Result	Dilution Factor	Lab Qualifier	Validation Qualifier
409186017	OLV 790	0138	0.160	1.00	U	
409186018	OLV 800	0315	0.160	1.00	U	
409186019	OLV 799	0347	0.160	1.00	U	
409186021	OLV 807	9347	0.160	1.00	U	

SAMPLE MANAGEMENT SYSTEM

Validation Report: Field Duplicates

RIN: 16108110 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 1/3/2017

Duplicate: 9347 **Sample:** 0347

	Sample				Duplicate						
Analyte	Result	Flag	Error	Dilution	Result	Flag	Error	Dilution	RPD	RER	Units
1,1,1,2-Tetrachloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1,1-Trichloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1,2,2-Tetrachloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1,2-Trichloro-1,2,2-trifluoroethane	1.50	U		1.00	1.50	U		1.00			ug/L
1,1,2-Trichloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1-Dichloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,1-Dichloroethene	0.160	U		1.00	0.160	U		1.00			ug/L
1,1-Dichloropropene	0.160	U		1.00	0.160	U		1.00			ug/L
1,2,3-Trichlorobenzene	0.200	U		1.00	0.200	U		1.00			ug/L
1,2,3-Trichloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
1,2,4-Trichlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,2,4-Trimethylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dibromo-3-chloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dibromoethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dichlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dichloroethane	0.160	U		1.00	0.160	U		1.00			ug/L
1,2-Dichloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
1,3,5-Trimethylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,3-Dichlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
1,3-Dichloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
1,4-Dichlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
2,2-Dichloropropane	0.160	U		1.00	0.160	U		1.00			ug/L
2-Butanone	0.500	U		1.00	0.500	U		1.00			ug/L
2-Chlorotoluene	0.160	U		1.00	0.160	U		1.00			ug/L
2-Hexanone	0.500	U		1.00	0.500	U		1.00			ug/L
4-Chlorotoluene	0.160	U		1.00	0.160	U		1.00			ug/L
4-Isopropyltoluene	0.160	U		1.00	0.160	U		1.00			ug/L
4-Methyl-2-pentanone	0.500	U		1.00	0.500	U		1.00			ug/L
Acetone	0.500	U		1.00	0.500	U		1.00			ug/L
Benzene	0.160	U		1.00	0.160	U		1.00			ug/L
Bromobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Bromochloromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Bromodichloromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Bromoform	0.160	U		1.00	0.160	U		1.00			ug/L
Bromomethane	0.160	U		1.00	0.160	U		1.00			ug/L
Carbon Disulfide	0.500	U		1.00	0.500	U		1.00			ug/L
Carbon Tetrachloride	1.61			1.00	1.66			1.00	NA		ug/L
Chlorobenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Chloroethane	0.160	U		1.00	0.160	U		1.00			ug/L

SAMPLE MANAGEMENT SYSTEM

Validation Report: Field Duplicates

RIN: 16108110 **Lab Code:** GEN **Project:** Mound LTS&M Ground Water **Validation Date:** 1/3/2017

Duplicate: 9347 **Sample:** 0347

	Sample				Duplicate						
Analyte	Result	Flag	Error	Dilution	Result	Flag	Error	Dilution	RPD	RER	Units
Chloroform	0.480	J		1.00	0.480	J		1.00			ug/L
Chloromethane	0.160	U		1.00	0.160	U		1.00			ug/L
cis-1,2-Dichloroethene	0.160	U		1.00	0.160	U		1.00			ug/L
cis-1,3-Dichloropropene	0.160	U		1.00	0.160	U		1.00			ug/L
Dibromochloromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Dibromomethane	0.160	U		1.00	0.160	U		1.00			ug/L
Dichlorodifluoromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Ethylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Hexachlorobutadiene	0.160	U		1.00	0.160	U		1.00			ug/L
Isopropylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Methylene Chloride	0.160	U		1.00	0.160	U		1.00			ug/L
Naphthalene	0.160	U		1.00	0.160	U		1.00			ug/L
n-Butylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
n-Propylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
sec-Butylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Styrene	0.160	U		1.00	0.160	U		1.00			ug/L
tert-Butylbenzene	0.160	U		1.00	0.160	U		1.00			ug/L
Tetrachloroethene	0.160	U		1.00	0.160	U		1.00			ug/L
Toluene	0.160	U		1.00	0.160	U		1.00			ug/L
Total Xylenes	0.160	U		1.00	0.160	U		1.00			ug/L
trans-1,2-Dichloroethene	0.160	U		1.00	0.160	U		1.00			ug/L
trans-1,3-Dichloropropene	0.160	U		1.00	0.160	U		1.00			ug/L
Trichloroethene	24.6			1.00	24.6			1.00	0		ug/L
Trichlorofluoromethane	0.160	U		1.00	0.160	U		1.00			ug/L
Vinyl Chloride	0.160	U		1.00	0.160	U		1.00			ug/L

SAMPLE MANAGEMENT SYSTEM
Organics Data Validation Summary

RIN: 16108110 **Project:** Mound LTS&M Ground Water **Lab Code:** GEN **Validation Date:** 1/3/2017

LCS Recovery: All LCS recoveries were within the laboratory acceptance limits.

Method Blank(s): All method blanks results were below the method detection limit.

MS/MSD Recovery: All MS/MSD recoveries were within the laboratory acceptance limits.

Surrogate Recovery: All surrogate recoveries were within the laboratory acceptance limits.