Chapter 12
Ground Water Data Analysis

John R. Kasich, Governor
Mary Taylor, Lt. Governor
Craig W. Butler, Director

February 2018
Chapter 12

Ground Water Data Analysis

February 2018

Revision 1

Ohio Environmental Protection Agency
Division of Drinking and Ground Waters
P.O. Box 1049
50 West Town Street
Columbus, Ohio 43216-1049
Phone: (614) 644-2752
epa.ohio.gov/ddaqw/
Preface

This document is part of a series of chapters incorporated in Ohio EPA’s Technical Guidance Manual for Hydrogeologic Investigations and Ground Water Monitoring (TGM), which was originally published in 1995. DDAGW now maintains this technical guidance as a series of chapters rather than as an individual manual. These chapters can be obtained at epa.ohio.gov/ddagw/gw_support.aspx

The TGM identifies technical considerations for performing hydrogeologic investigations and ground water monitoring at potential or known ground water pollution sources. The purpose of the guidance is to enhance consistency within the Agency and inform the regulated community of the Agency’s technical recommendations and the basis for them.

Ohio EPA utilizes guidance to aid regulators and the regulated community in meeting laws, rules, regulations and policy. Guidance outlines recommended practices and explains their rationale. The methods and practices described in this guidance are not intended to be the only methods and practices available to an entity for complying with a specific rule. Unless following the guidance is specifically required within a rule, the Agency cannot require an entity to follow methods recommended within the guidance. The procedures used should be tailored to the specific needs and circumstances of the individual site, project and applicable regulatory program, and should not comprise a rigid step-by-step approach utilized in all situations.
Changes from the February 1995 TGM

Ohio EPA’s *Technical Guidance Manual for Hydrogeologic Investigations and Ground Water Monitoring* (TGM) was first finalized in 1995. Chapter 12 (Ground Water Quality Data Organization and Interpretation) was added as a new chapter in February 2005. This is the first revision to the chapter.

Section numbers were added to make the document easier to read.

References were updated, in particular, the references for ASTM standards and USEPA guidance documents.

Additional information has been added on:

- Data validation.

- Terminology changed throughout document to make concepts flow better and easier to understand.

- Multivariate Statistical Evaluation Section.
# Table of Contents

Preface ................................................................................................................................. 12-iii

Table of Contents .................................................................................................................. 12-v

1.0 Lab Analytical Methods and QA/QC ............................................................................. 12-2
   1.1 Analytical Methods ....................................................................................................... 12-2
   1.2 Lab QA/QC .................................................................................................................. 12-2

2.0 Validation of Data ............................................................................................................ 12-4
   2.1 U.S. EPA Guidance ..................................................................................................... 12-4
   2.1 Ohio EPA Guidance .................................................................................................. 12-5

3.0 Tools of Interpretation ..................................................................................................... 12-6
   3.1 Maps ............................................................................................................................. 12-6
   3.2 Graphs ........................................................................................................................ 12-6
      3.2.1 Bar Charts ........................................................................................................... 12-6
      3.2.2 XY Charts .......................................................................................................... 12-6
      3.2.3 Box Plots ............................................................................................................ 12-9
      3.2.4 Notched Boxplots ............................................................................................... 12-10
      3.2.5 Trilinear Diagrams ............................................................................................. 12-10
      3.2.6 Stiff Diagrams .................................................................................................... 12-11
   3.3 Statistics ....................................................................................................................... 12-12
      3.3.1 Principle Component Analysis - Multivariate Statistical Evaluation .................. 12.14
   3.4 Modeling ....................................................................................................................... 12-14
   3.5 Professional Judgment ................................................................................................ 12-14

4.0 Data Submittal ................................................................................................................ 12-16
   4.1 Data Validation Summary ............................................................................................ 12-16
   4.2 Statistical Data Submittals ......................................................................................... 12-16
   4.3 Tables .......................................................................................................................... 12-17
   4.4 Maps ............................................................................................................................ 12-17

5.0 References ....................................................................................................................... 12-18
Large amounts of ground water quality data can be generated during a hydrogeologic investigation and/or ground water monitoring sampling event. Proper data analysis and assessment, including validation, are necessary to enable sound decisions. Data Quality Assessment (DQA) is the scientific and statistical evaluation of environmental data to determine if they meet the project objectives, and thus are of the right type, quality, and quantity to support their intended use. DQA is meaningful only when it relates to the intended use of the data, answering the question: Can an environmental decision be made with the desired level of certainty, given the quality of the data?

In order to have adequate DQA, it is important that the regulatory basis for collecting the samples is understood. In addition, the following information is necessary for DQA:

1) the correct methods of analysis and quality assurance/quality control (QA/QC) documentation are requested from the lab;
2) the data from the lab are checked for technical soundness and documentation;
3) appropriate interpretive tools are used in evaluating the data; and
4) adequate and appropriate data are submitted in a manner that is easily understood to allow for credible decisions.
1.0 Lab Analytical Methods and QA/QC

1.1 Analytical Methods

Environmental data are generated from numerous analytical methods and different types of equipment. In order to select the appropriate analytical method(s), basic information is required including:

- The type of sample (water, leachate, soil, etc.)
- The analytes of interest
- The sensitivity or quantitation limits needed
- The data quality objective(s)
- Whether the purpose is detection, assessment or evaluating corrective measures

Ohio EPA's Division of Environmental Response and Revitalization (DERR) discusses the analytical methods most widely used for ground water monitoring by Ohio EPA in Chapter 2 of their Tier I Data Validation Manual.

1.1.1 Analytical Methods Approved for Ground Water Monitoring

The U.S. EPA publication SW-846, is the Office of Solid Waste's official compendium of analytical and sampling methods that have been evaluated and approved for use in complying with the Resource and Recovery Act (RCRA) regulations. SW-846 functions primarily as a guidance document setting forth recommended methods for the regulated and regulatory communities to use in responding to RCRA-related sampling and analysis requirements. Additional information on SW-846 can be found on the U.S. EPA website address at https://www.epa.gov/hw-sw846.

1.1.2 Analytical Methods Approved for Drinking Water Monitoring

The following website outlines documents listing all of the methods that are currently approved by U.S. EPA for the analysis of drinking water samples. Each document focuses on the methods that are applicable to a specific monitoring requirement: https://www.epa.gov/dwanalyticalmethods.

1.2 Lab QA/QC

The following is a list of common data quality parameters that should be requested from the lab to be reviewed during data validation:

- Chain of custody;
- Case narrative discussing any qualified or rejected data;
- Batch/field/lab identifications (IDs) and ID correspondence sheet;
- Holding times;
- Preservation and cooler receipt;
- Cooler temperatures on receipt;
- Analytical technique and method numbers (i.e., from SW-846, ASTM, etc.)
• Dilution factors;
• Date/time of sample collection;
• Date/time sample was received by the laboratory;
• Date/time of sample extraction (if applicable);
• Date/time analyses were performed, and individuals who performed them;
• Surrogate recoveries with control limits (for organics only);
• Laboratory blank data (method blanks, preparation blanks);
• Field blank data (duplicate, trip, field and equipment blanks);
• Spike data (including matrix spike/ matrix spike duplicate);
• Laboratory control samples (LCS);
• Method detection limit (MDL);
• Practical quantitation limit (PQL);
• Units of measurement;
• Data results; and
• Supporting field data (e.g., field sampling data sheets, field logbook notes, etc.)
2.0 Validation of Data

Data validation is the process of evaluating the completeness, correctness, consistency, and compliance of a data package before using it as a basis for making environmental decisions. The validity of analytical data is crucial because it serves as a basis for evaluating compliance with the Ohio EPA rules, permits and findings and orders. Data validation should identify laboratory and analytical errors that are associated with a data set.

Data must be systematically compared against a standard or against project-specific criteria to provide assurance that the data are adequate for the intended use. Validation consists of editing, screening, checking, auditing, verification, certification, and review (Canter et al., 1988). By working through the quality control information associated with a method, sample data may be either validated, qualified as estimated, or rejected.

2.1 U.S. EPA Guidance

U.S. EPA Guidance for validation of chemical analyses is available on the web at:


  This document is designed to offer guidance on Contract Lab Program (CLP) organic analytical data evaluation and review. It is intended to assist in the technical review of analytical data generated through the CLP Statement of Work (SOW) SOMO2.2 (September 2015).


  This document is designed to offer guidance on CLP inorganic analytical data evaluation and review. It is intended to assist in the technical review of analytical data generated through the CLP SOW ISMO2.3 (September 2015).


  This document is designed to offer guidance on CLP high resolution analytical data evaluation and review. It is intended to assist in the technical review of analytical data generated through the CLP SOW HRSM01.2 (October 2014).

  This document is also designed to offer guidance on CLP dioxin and furan analytical data evaluation and review. It is intended to assist in the technical review of analytical data generated through the CLP SOW HRSM01.2 (October 2014).
2.2 Ohio EPA Guidance

Ohio EPA’s DERR has developed their Tier I Data Validation Manual and tiered checklist system available on the internet. The Tier I Data Validation Checklist addresses basic QA/QC parameters run by the lab, while the Tier II Data Validation Checklist includes all of the parameters assessed during the Tier I Data Validation at a more detailed level. In addition, instrument calibration and analysis sensitivity are addressed during Tier II validation.

Ohio EPA’s Division of Environmental Response and Revitalization (DERR) uses:

- Laboratory and Field Screening Data Review, August 2005.
- Data Quality Objectives Process Summary, January 2002
3.0 Tools of Interpretation

Ground water quality data should be compiled in a manner convenient for interpretation. Techniques include mapping, graphical displays, statistics, and modeling. The appropriate tools to be used depend on the goals of the sampling event or monitoring program.

3.1 Maps

Isopleth (isoconcentration) maps are contour maps constructed by drawing lines representing equal concentrations of dissolved constituents or single ions (Figure 12.1). These maps, when combined with site-specific geologic/hydrogeologic characteristics (see TGM Chapter 3), are useful in the identification, delineation and monitoring of contamination plumes. However, their applicability and accuracy depends on the homogeneity of ground water quality with depth and the concentration gradient among measuring points. Decisions on sampling point density, both horizontally and vertically, will need to be made based upon the intended use of the data. Potentiometric surface and isopleth maps should use data from the same hydrologic unit when multiple aquifers are present. Dashed contour lines should be used to indicate areas of uncertainty due to lack of data or low-quality data.

3.2 Graphs

Graphical presentation can be helpful in visualizing the temporal and spatial distribution of contaminants, identifying changes in water quality with time, spatially within the aquifer, and comparing waters of different compositions. Typical methods include, but are not limited to, bar charts, XY charts, box plots, trilinear diagrams, and stiff diagrams.

3.2.1 Bar Charts

Bar charts display analyte concentrations of ground water contamination with data typically reported in units of milligrams per liter (mg/l, equivalent to parts-per-million or ppm) or micrograms per liter (μg/l, equivalent to parts-per-billion or ppb). For such studies, bar charts can be constructed to display concentrations of constituents for single or multiple monitoring wells and/or sampling events. Figure 12.2 presents several examples of bar charts that may be useful.

3.2.2 XY Charts

Plotting changes in dissolved constituents with time constitutes one example of an XY chart that is extremely useful when evaluating contaminant migration or ground water remediation progress. Even with a relatively slow rate of ground water migration, long-term monitoring can detect gradual water quality changes. Time-series graphs can be used to compare individual parameters for a single well with time, multiple parameters for a single well with time (Figure 12.3), or illustrate changes with time for a common parameter in multiple wells (Sara and Gibbons, 1991). Caution should be used when evaluating data of variable quality, i.e., data with different levels of quality assurance/quality control. Variations in data quality could be misinterpreted as variations in ground water quality.
Figure 12.1 Example of an isopleth map of total VOC concentrations (ppb) at the Chem-Dyne site in Hamilton, Ohio for shallow well data from December 1985. U.S. EPA (1989b).
Figure 12.2 Example of bar charts. A) Shows concentrations of lead and chromium for one sampling event. B) Shows concentrations of several constituents at one well over multiple sampling events.
3.2.3 Box Plots

Box plots (also called box and whisker plots) can be used to compare ground water quality data (generally for the same parameter) between or among wells. The plots are constructed using the median value and the inter-quartile range (i.e., 25 and 75 cumulative frequencies as measured central tendency and variability) (U.S. EPA 2009, Part II, Chapter 9, Section 9.2, Page 9-5) (Figure 12.4). They are a quick and convenient way to visualize the spread of data.
3.2.4 Notched Boxplots

Notched boxplots contain a “notch” (indentation) that represents a 95% confidence interval about the median for a particular constituent at a particular well. Comparison of notch locations (concentration ranges within notch areas) from different plots (well/constituent pairs) provides a good estimate of approximate lateral ground water quality spatial variability. For example, if chloride boxplots for three different wells have “notches” that all overlap, then the spatial variability among the chloride concentrations in the wells is statistically similar. This similarity indicates that the measured concentrations are likely from the same population (Figure 12.5). If none of the notches overlap, then considerable spatial differences in the constituent’s concentrations in ground water exist calling stationarity into question. Go to http://www.wessa.net/rwasp_notchedbox1.wasp to perform notched boxplot calculations using R website “Free Statistics and Forecasting Software.”

![Chloride Notched Boxplots, Landfill X](image)

**Figure 12.5** Example of a “notched” box plot

3.2.5 Trilinear Diagrams

Trilinear diagrams may be used in water chemistry studies to classify natural waters (Sara and Gibbons, 1991). They can show the percentage composition of three groups of ions and usually are in the form of two triangles bracketing a diamond-shaped plotting field (Figure 12.5).

These diagrams are useful in determining the similarities and/or differences in the general composition of ground water from specific hydrogeologic units, and are convenient for displaying a large number of analyses. The diagrams may help show whether particular units are hydraulically separate or interconnected and if ground water quality has been affected by dissolution or precipitation of a salt.
The value of trilinear diagrams may be limited for some investigations because ionic composition is represented as a percentage. Therefore, waters with very different total concentrations (or ionic strengths) may appear identical on the diagram. Because non-ionic solutes (e.g., silica and organics) are not represented (Hem, 1985), trilinear diagrams are generally not used when evaluating the presence or absence of contaminants.

Figure 12.5 Example of a trilinear diagram

3.2.6 Stiff Diagrams

Stiff diagrams are another graphical representation of the general chemistry of ground water. A polygonal shape is created from four parallel horizontal axes extending on either side of a vertical axis. Cations are plotted on the left of the vertical axis and anions are plotted on the right (Fetter, 1994). The relatively distinctive diagrams can be used to illustrate ground water composition differences or similarities. The width of the pattern approximates total ionic strength (Hem, 1985). One feature is the tendency of a pattern to maintain its characteristic shape as the ground water becomes diluted. It may be possible to trace the same types of ground water contamination from a source by studying the patterns. In the case presented in Figure 12.6, seepage of salt water from a brine disposal pit was suspected. Samples analyzed from the pit and the wells demonstrated the same pattern, showing evidence of contamination (Stiff, 1951).
3.3 Statistics

Whether a release has occurred, a plume is expanding and/or corrective actions are effective may also be evaluated by statistical analysis, if adequate data are available. The U.S. EPA March 2009 Unified Guidance, “Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities”, updates and replaces the earlier 1989 Interim Final Guidance and the July 1992 Addendum. The Unified Guidance provides...
The statistical evaluation of data resulting from the chemical analysis of ground water is required by Ohio’s regulations for RCRA Subtitle C and D facilities including Part B permitted hazardous waste land disposal facilities [OAC Rule 3745-54-97(G)], Part A permitted hazardous waste land disposal facilities [OAC Rule 3745-65-93], permitted municipal solid waste [OAC Rule 3745-27-10(C)], residual and industrial waste landfills [OAC Rule 3745-30-08(C)], and also construction and demolition debris sites [OAC Rule 3745-400-10] regulated by DMWM and/or DERR. Additionally, statistics may be employed for the evaluation of ground water quality at wastewater facilities (including non-toxic flyash, bottom ash, foundry sand and coal storage piles) regulated by DSW and to support corrective actions at remedial response [CERCLA (Superfund), 42 U.S.C. §§ 9601-9675 et seq.], voluntary action program [OAC Rule 3745-300-07] and brownfield [Small Business Liability Relief and Brownfields Revitalization Act, P.L. 107.118] sites regulated by DERR.

Regarding the methods mandated by OAC Rule 3745-65-93 (i.e., the Student’s T-test and the Average Replicate Test), the reader is referred to the U.S. EPA’s Technical Enforcement Guidance Document (U.S. EPA, 1986).

While statistics may be useful to determine if a release has occurred, is expanding, or is being remediated, professional judgment needs to be exercised to ensure that the results are representative of actual conditions. For example, the results may show either a “false positive” or “false negative” due to naturally occurring variations such as geologic heterogeneity and/or seasonal variability (Gibbons et al., 2009).

Statistical analyses are used to compare the chemical ground water quality of a monitored zone downgradient of the unit/area in question with either: 1) a set standard, such as an MCL or risk-based standard, 2) the chemical ground water quality from a background (or upgradient) well(s) screened in the same monitored zone and unaffected by facility operations (interwell statistical comparisons), or 3) historic concentrations from the same well (intrawell comparisons). These comparisons provide reliable determinations as to how and whether a waste disposal unit, area of contamination release and migration or remediation have influenced the quality of the ground water. Statistics can also be used to define the extent of ground water contamination. If a statistically significant difference is not observed between a downgradient well and background, the downgradient well is statistically considered uncontaminated.

Whether building a new database or adding to an established database, the initial steps in determining if a statistically significant difference has occurred are the same. First, before choosing the statistical method, an evaluation must be made as to whether the available site-specific data meet the basic assumptions necessary for a particular test to perform with appropriate accuracy and power. Power is the probability that the test will correctly identify contamination when it is present. The goal is to have the power be as high as reasonable, while keeping the accuracy (false-positive and false negative rates) under adequate control. If the basic assumptions of each test are met, the maximum power will be employed. The power to detect differences tends to increase as the concentration of a constituent within an individual well moves further from the mean or median of the background database. Additional assumptions that need to be verified for all methods include the independence of samples, number of samples, determination and distribution of the background data set, sampling frequency, and the need for seasonality corrections. Statistical methods employed by the owner/operator are required by rule to meet these basic, critical assumptions.

Once the basic assumptions have been met, or the data set has been mathematically transformed to meet
the basic assumptions, a statistical method may be chosen. The next step is to select the best options considering site-specific data and the specific assumptions that allow each method to perform with the greatest accuracy and power. Assumptions that change from method to method include minimum sample size, determination of distribution, homogeneity of variance, treatment of non-detect parameter levels, and experiment wise and comparison wise errors. Other factors to accounting for each assumption are the need to use a different statistical method as the site-specific database changes over time; and the need to use a different statistical method for some constituents.

3.3.1 Principle Component Analysis - Multivariate Statistical Evaluation

Principal component analysis (PCA) is one of several ordination techniques used to condense large data sets into underlying components (Shaw, 2003). PCA is most useful where a water quality data set has been influenced by only a few underlying components. The fundamental operation of PCA treats a set of data as a geometrical object, then finds the most informative orthogonal (angle) from which to view the shape. PCA is the default ordination tool on most statistics programs which perform ordinations, due to its ease of calculation, and robustness to deviations from multivariate normality in the data (Shaw, 2003). PCA is the primary ordination tool for the many environmental studies and has been used to differentiate component contribution in a comingled plume (Churchill, 2017).

3.4 Modeling

Ground water modeling can be used to assist in the determination of extent and rate of contaminant migration. Models can be conducted throughout the investigation and remedial processes. Information on modeling can be found in Chapter 14.

3.5 Professional Judgment

Professional judgment involves the use of education and experience. In some cases, a simple visual inspection of downgradient versus upgradient/background data can show obvious differences in chemical quality. The tabular and graphical presentations discussed earlier in this chapter can be used for this evaluation.

When evaluating potential ground water contamination, water quality data often are compared to primary and secondary drinking water standards. As important as it is to protect public health by identifying an exceedance, formulating a conclusion that ground water has been contaminated based solely on an exceedance is not appropriate. Certain inorganic constituents, such as iron, arsenic, and sulfate, can occur naturally in Ohio’s ground water at levels above standards; therefore, an exceedance for these or some other constituents may not necessarily imply contamination, but reflect natural variations in ground water quality. Conversely, values lower than a standard do not necessarily imply that contamination has not occurred. The background water quality could be impacted at a concentration below the standard. However, the mere presence of organics such as chlorinated solvents, petroleum distillates, pesticides/herbicides or PCBs generally indicates contamination, as these substances are rarely (if ever) naturally occurring in ground water.

Data for wells downgradient from a pollution source should be compared to data from an upgradient/background well that has not been affected by the source. If an upgradient/background well does not exist, then the results can initially be compared to known local or regional background values. However, utilization of regional values for evaluating potential contamination should be a part of initial
investigations only. Further evaluation should be based on site-specific background sampling. In any ground water contamination investigation, obtaining background concentrations for chemical constituents of concern is essential, particularly those that may be common to both the local ground water quality and the potential or known contaminant source.
4.0 Data Submittal

Ground water quality data should be presented in a manner convenient for interpretation. Adequate and appropriate data should be submitted in a manner that is easily understood to substantiate and support decisions.

4.1 Data Validation Summary

The data validation summary does not have a required format. However, it should contain key elements and a summary of the data validation findings, including the rational for collecting the data, the statement of the data quality objectives, the summary of findings, an analysis of whether the data quality objectives have been met, or if additional (higher level) data validation is necessary. Data qualifiers, if any, should be assigned to the data in the report with the problems encountered discussed. If there are other criteria that can have a significant bearing on the usability of the data they should be noted (e.g., missing data). Finally, a complete documented evaluation of whether the reported detection limits or quantitation limits met the regulatory or risk standards, whether any deviations of the methodology occurred, and whether any bias exists in the data must be made. The following common data quality parameters should be reviewed during data validation:

- Chain of custody;
- Case narrative discussing any missing, qualified or rejected data;
- Batch/field/lab identifications (IDs) and ID correspondence sheet;
- Holding times;
- Preservation and cooler receipt;
- Cooler temperature upon receipt;
- Analytical technique and method numbers (i.e., from SW-846, ASTM, etc.)
- Dilution factors;
- Date/time of sample collection;
- Date/time sample was received by the laboratory;
- Date/time of sample extraction (if applicable);
- Date/time analyses were performed, and individuals who performed them;
- Surrogate recoveries with control limits (for organics only);
- Laboratory blank data (method blanks, preparation blanks);
- Field blank data (duplicate, trip, field and equipment blanks);
- Spike data (including matrix spike/ matrix spike duplicate);
- Laboratory control samples (LCS);
- Method detection limit (MDL);
- Practical quantitation limit (PQL);
- Units of measurement;
- Data results; and
- Supporting field data (e.g., field sampling data sheets, field logbook notes, etc.)

4.2 Statistical Data Submittals

To facilitate correct interpretation of the statistical data by both Ohio EPA and the regulated entity, the Agency recommends that certain information be submitted with any statistical evaluations. Evidence should be provided that the statistical methods used are:

- as specified in the permit/plan, comparing the results to the appropriate standard;
• conducted separately for each constituent;

• completed within the time frame and at the frequency specified in the permit/plan;

• appropriate for the distribution of the constituents;

• performed at a Type 1 error level of 0.01 for each testing period for individual well comparisons; 0.05 for multiple comparisons (except for tolerance, confidence, prediction intervals and control charts); and 0.1 for the annual site-wide false positive rate;

• account for data between the method detection limit and the quantitation limit as specified in the permit/plan;

• include procedures to control or correct for seasonal, spatial and temporal variability, as necessary; and

• include example calculations of the statistical methods used.

4.3 Tables

Tables of data are the most common form in which the chemical analyses are reported. Tables generally are sorted by well, type of constituent, and/or time of sampling. For most constituents, data are expressed in milligrams per liter (mg/l) or micrograms per liter (μg/l). Data should be organized and presented in tabular form or as dictated by regulatory or program requirements. Reports from the laboratory also should be submitted. Some Ohio EPA programs (DMWM) require ground water quality data to be submitted in a set computer-based format and define which laboratory data must be submitted with required reports. Before submitting data in an electronic format, regulated entities should check with the appropriate program to determine the preferred media and format.

4.4 Maps

Isopleth maps are contour maps constructed by drawing lines representing equal concentrations of dissolved constituents or single ions (Figure 12.1). These maps, when combined with site-specific geologic/hydrogeologic characteristics (see Chapter 3), are useful in tracking plumes and documenting the effectiveness of remediation. Their applicability depends on the homogeneity of ground water quality with depth and the concentration gradient between/among measuring points. Restricting sampling points in either the vertical or horizontal direction limits usefulness (Sara and Gibbons, 1991). Dashed contour lines should be used to indicate areas of uncertainty due to lack of data or low-quality data.
5.0 References


