CEST 2021
17th International Conference on Environmental Science and Technology
Athens, Greece, 1 to 4 September 2021

Data Visualization Techniques for Contaminated Site Investigation and Remediation

Toskos T.[1,*]

[1] Woodard & Curran, New Jersey, USA

*corresponding author:
e-mail: theodoros.toskos@gmail.com

Abstract Site investigations and remedial actions generate large amounts of data. While tabulating the data is necessary for the purpose of satisfying reporting requirements and comparisons to standards, it does not allow us to understand and explore relationships in the data to obtain important information on items such as contaminant source or fate and transport. Data visualization tools allow us to explore and understand these relationships. However, application of these tools requires planning and proper selection of assumptions, to avoid reaching improper conclusions. We are providing examples of how simple tools provide powerful information for site closure.

Keywords: visualization, contamination, investigation

1. Introduction

Most environmental investigations generate a substantial amount of data that must be understood and interpreted. While certain types of data, such as groundwater elevations, are almost automatically reduced and presented in a graphical manner that is easy to visualize and put into context, analytical data, which is the most voluminous data generated, is often cosigned to endless tables and mechanical recitation in the report text, without any effort to analyze and visualize in a holistic manner, to explain the behavior of the site and help the investigator develop appropriate solutions. Luckily, as the cost of computer hardware and software has declined over the years, we now have a wide of powerful tools that allow us to analyze and visualize investigation data, with minimal effort and cost. However, while these tools are cheap, fast, and easy to use, forethought in the application is required, to avoid arriving at incorrect conclusions.

2. Considerations and Planning

The process of data visualization must start with the clear definition of the question to be answered and the ultimate purpose of the analysis. General statements, such as, "show concentration trends over time" almost always lead to the use of the wrong tools and end-up with inappropriate conclusions. At a contaminated site, many things will happen "over time" which will affect contaminant concentrations. Failing to consider these actions, we can easily construct an inappropriate visualization. Consider the difference between "contaminant concentration trends from before remedial action to the present" and "contaminant concentrations following remedial action to the present". While both questions sound similar, they are not.

The ultimate use of the outcome of the data analysis must also be considered and stated clearly. While the objectives "show that the two well sets behave differently" and "show that the two well sets reflect different aquifer conditions" may sound the same, the regulatory implications and the selection of visualization tools is great.

3. Case Examples

At a site where groundwater was contaminated with benzene, a plot of the sampling data over time was prepared and a least-squares fit of a trend line was developed. The conclusion was that concentrations were steadily declining and compliance with regulatory limits would be attained soon (Figure 1).

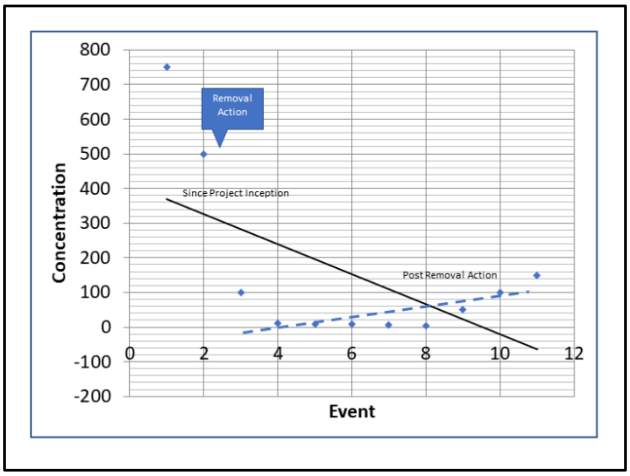

Figure 1- Effects of improper data aggregation

While the visualization shows that there was a net improvement of groundwater quality, the investigator did not consider the history of the data collection. Some data had been collected before a source removal action and some data had been collected after the source removal action. The investigator also did not consider the objective behind the analysis, which was do demonstrate that the removal action was successful and groundwater quality would be restored within a short period. When the points representing conditions prior to the removal action were removed from the data fitting, a different picture emerged (Figure 1). Concentrations initially declined but are ultimately increasing. While that does not imply failure of the removal action, it does create a very different conceptual model of the site behavior following removal. Processes, such as contaminant back-diffusion from clayey sediments may be operating, resulting in the upward trend. While the visualization does not provide the explanation for the observed trend, it prompts the investigator to reconsider their assumptions about site conditions and their ability to meet regulatory closure requirements.

The importance of setting up a proper statement of the objective of the analysis is illustrated by the following example. The remedial investigation at a manufacturing site showed high concentrations of sodium in groundwater. The initial conclusion was this was due to historical site operations. Subsequent work provided indications that de-icing operation along the surrounding roadways, and natural conditions in the formation may be contributing to the observed concentrations. Data had been tabulated and examined for evidence that certain wells were reflecting site operations, while other wells were reflecting the impact of ambient conditions. No conclusion could be drawn from such review and attempts at statistical analysis, such as ion ratios. Ultimately, the problem formulation and the objective of the analysis were reformulated to be: "is there more than one source of contamination" and "identify the wells which reflect the impact of historic plant operations." With those consideration in mind, a classic hydrogeologic data visualization tool was selected. Figure 2 shows a Piper Diagram for the site data, which demonstrates the clustering of wells according to formation and installation details, which, in turn, reflects the origin of the observed contamination and has implications for the obligation to remediate.

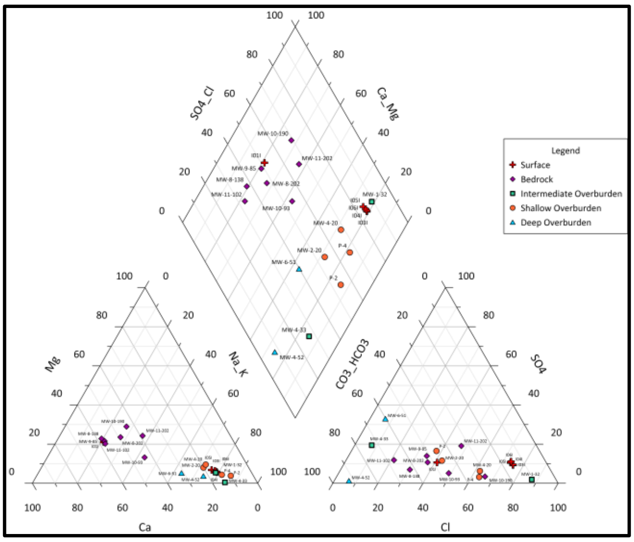

Figure 2- Piper Diagram showing well grouping

Such software can be obtained free of charge from government agency websites or can be purchased from specialty software vendors.

Finally, there are many more simple tools that the investigator can use to understand plume evolution in space and/or time, especially when more than one contaminant is involved. As an example, two disposal areas have resulted in two similar and overlapping plumes. To assess plume behavior and remedial action options, a "radar" or "spider" plot was constructed that helped the investigator visualize the two plumes and their evolution across the site (Figure 3).

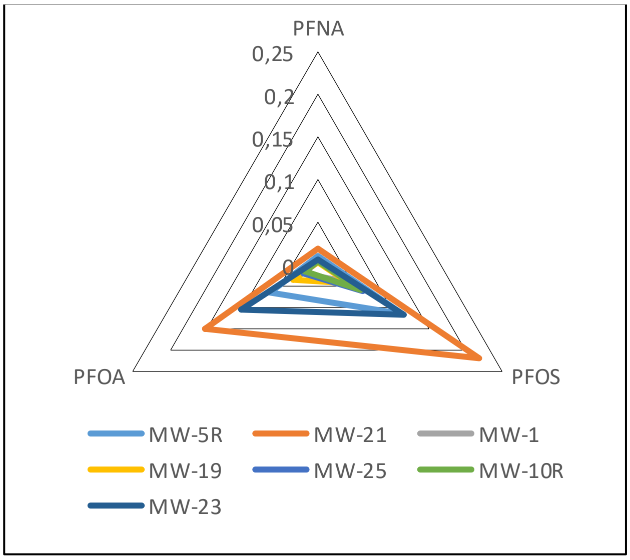

Figure 3- Radar plot of PFAS data

This tool allows the investigator to visualize the relative concentration of contaminants in each sample without the need for elaborate ratio calculations. The data generate two types of triangles, based on the relative composition of each source. This allowed the investigator to answer fundamental questions such as "which well reflects which source," and "how is each plume changing as it migrates across the site." The visualization reveals that both plumes attenuate rapidly over a short distance and that, perhaps, monitored natural attenuation is a viable alternative. This visualization also prompts the investigator to question what may be causing the observed rapid attenuation of contaminants that are considered persistent. Such tools are often included in commercial spreadsheet software (some offered under open-source licensing) or from specialty software vendors, and are broadly available.

4. Discussion

Investigators of contaminated sites need to not only compare sampling results to cleanup levels, but to also construct conceptual models of the site behavior. Many software tools are available, that enable the investigator to convert large volumes of data quickly and cheaply to images that illustrate site conditions and facilitate site closure. It is critical for the investigator to carefully craft the problem definition and project objective, so that they can select the proper data visualization tool and apply it to the proper data set