Interactive Data Visualization Dashboard Project

Jeff Peterson

BACKGROUND

Dugway Proving Ground has recently conducted a test that will measure concentrations of chemicals within a closed environment. In the past, when we have conducted tests like this, we have produced hundreds of plots to show the changes in concentration versus time for the entire experiment. This type of data visualization is not only time consuming, but inefficient and ineffective. In order to really understand a concentration versus time experiment, one needs to use an interactive visualization dashboard. This kind of visualization will allow the user to drive the data story and customize their experience to answer many different questions both quickly and effectively. Building a visualization like this with the Tableau software is the main focus of this project.

A special test chamber has been constructed that will house the various instruments during this test. The instruments will include 32 RAE[1-2] instruments at two different heights, five different chemical release locations, and four separate GASMET[3] instruments.

The visualization will be structured in a quad chart format. A single Tableau Dashboard will be divided into four separate quadrants that will each highlight different data elements from the test. The quad chart format is the preferred method for Department of Defense visualizations.

CONSTRUCTING THE DATA FILES

There are five files within four separate Excel spreadsheets that I needed for the visualization. Since I am not including the actual data file for this test, I will include a sample of each file so that the user has a better understanding of the individual data elements. A description of each data file and the information required from each is described below.

1. **Elapsed time_Sync**—this is an xlsx file that I created so that I would have a common column with which I could do a join with the other Excel files. The file only contains one column.
    a. **Elapsed time**: this is an elapsed time column that I made that is in hh:mm:ss format.

Elapsed time
0:00:00
0:01:00
0:02:00
0:03:00
0:04:00
0:05:00

Figure 1. Sample of Elapsed time_Sync Data File

2. **T1A**—this is a xlsx file for the T1A experiment trial that contains the following variables:
    a. **Time**: a timestamp column of values collected each second.

b. **Point:** a label column for each of the four GASMET instruments. The possible values are GASMET 1, GASMET 2, GASMET 3, and GASMET 4.
c. **Diethyl Sulfide**: this is a parts per million (ppm) concentration column for diethyl sulfide chemical.
d. Iso-Valeric Acid: this is a ppm concentration column for the iso-valeric acid chemical.
t-Butyl Mercaptan: this is a ppm concentration column for the t-butyl mercaptan chemical.

Time	Point	Diethyl Sulfide	iso-Valeric Acid	t-Butyl Mercaptan
12:21:07 AM	GASMET 1	-0.12	0.06	-0.16
12:21:09 AM	GASMET 2	0.05	-0.01	0.03
12:21:10 AM	GASMET 4	0.21	-0.05	0.22
12:21:20 AM	GASMET 3	-0.03	0.00	0.07
12:21:28 AM	GASMET 1	-0.11	0.05	-0.19
12:21:30 AM	GASMET 2	0.05	-0.02	0.07
12:21:31 AM	GASMET 4	0.25	0.00	0.20
12:21:41 AM	GASMET 3	0.01	-0.01	-0.02

Figure 2. Sample of T1A Data File

3. **Test Plot for Color Changing Data Points_T1A**—this is an xlsx file for the T1A trial that contains the following variables:
    a. **Elapsed Time**: this is a column that I generated that has values from 0-60 to indicate the elapsed minutes for the trial.
    b. **X:** this is a column I made that assigns an x-location based on the relative instrument location within the building. This was needed in order to place the instruments in the correct locations for displaying within the visualization.
    c. **Y:** this is a column I made that assigns an y-location based on the relative instrument location within the building. This was needed in order to place the instruments in the correct locations for displaying within the visualization.
    d. **Point:** this is a column I made that lists all of the instruments by name as well as the names of the chemical evaporation locations.
    e. **Maximum concentration:** this is an instrument generated value that is simply the maximum concentration detected across all instruments at any given time.
    f. **ET**: this is an instrument generated file of elapsed time in hh:mm:ss format.

Elapsed Time	X	Y	Point	Maximum Concentration	ET
0	38	5	RAE 8H	10.3	0:00:00
1	38	5	RAE 8H	6.7	0:01:00
2	38	5	RAE 8H	0.6	0:02:00
3	38	5	RAE 8H	2.6	0:03:00
4	38	5	RAE 8H	7.1	0:04:00
5	38	5	RAE 8H	19.1	0:05:00
6	38	5	RAE 8H	28.5	0:06:00
7	38	5	RAE 8H	34.8	0:07:00
8	38	5	RAE 8H	29.8	0:08:00

Figure 3. Sample of Test Plot for Color Changing Data Points_T1A Data File

4. **MAX MIN AVERAGE_T1A**—this is an xlsx file that consists of two worksheets (**T1A High** and **T1A Low**). Each worksheet contains the following variables:
    a. **Time:** this is time column generated by the instruments at 1-minute intervals.
    b. **Concentration:** this is a column of the measured concentration at each timestamp for each of the high and low RAE instruments.
    c. **Point:** this is a column of each of the RAE instruments at the 66” height for the high instruments and 18” height for the low instruments. They are grouped by elapsed time.
    d. **Elapsed Time:** this is an elapsed time column that I generated that reset the starting Time column above to a zero value.

Time	Concentration	Point	Elapsed Time
0:21:00	0	RAE 1 H	0:00:00
0:21:00	0	RAE 1 H	0:00:00
0:21:00	0	RAE 1 H	0:00:00
0:21:00	0	RAE 1 H	0:00:00
0:21:00	0	RAE 1 H	0:00:00
0:21:00	0	RAE 1 H	0:00:00
0:21:00	0	RAE 1 H	0:00:00
0:21:00	0	RAE 1 H	0:00:00
0:21:00	0	RAE 1 H	0:00:00

Figure 4. Sample of MAX MIN AVERAGE_T1A Data File - T1A High Worksheet

The T1A Low Worksheet values are similar to what is shown in Figure 3.

The final data files were loaded into Tableau and joined on the Elapsed Time column from the Elapsed time_Sync file as shown in Figure 5.

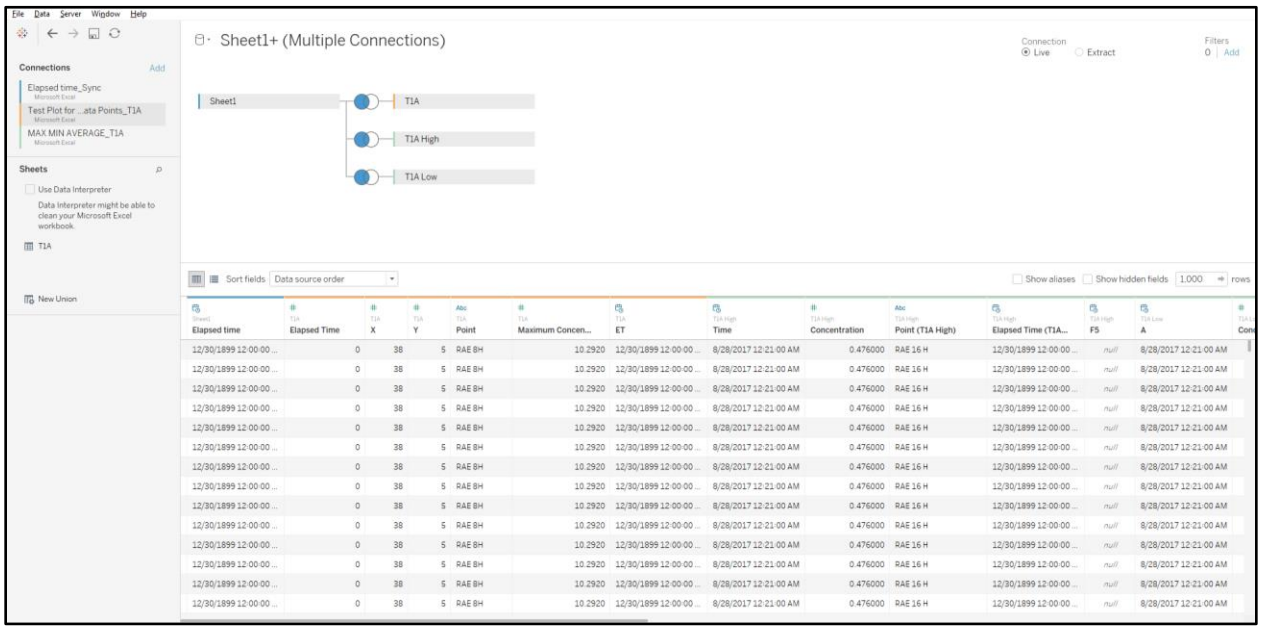

Figure 5. Screenshot of Data Loaded in Tableau

BUILDING QUADRANT 1

The first quadrant will consist of the top-down visualization of the 32 RAE instruments, the five release locations, and the four GASMET instrument locations. To make this visualization, I decided to make a coordinate grid system where I could assign an X and Y value to each of the instruments and locations. Doing this allowed me to place them in a position that was representative of their actual placement during the trial. One particular challenge occurred because I needed to find a way to represent two RAE instruments at the same location but at different heights. I ended up deciding to represent them as slightly offset icons with the instrument at the higher location overlapping the instrument at the lower location. Figure 6 shows how this configuration was visualized.

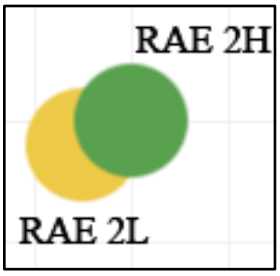

Figure 6. Overlapping Representation of Two RAE Instruments

The RAE 2H label is indicative of an RAE instrument at position 2 at the 66” height, while RAE 2L indicates an RAE instrument at position 2 at the 18” height.

The icon colors in Figure 6 are based on a color scale that I implemented so that I could monitor concentration changes in real-time throughout the trial. I made three concentration bins in ppm units and corresponding colors, which are:

1. <1 ppm = Green
2. 1-5 ppm = Yellow
3. 5+ ppm = Red

I specifically chose those colors because the concentration levels we were monitoring correspond to certain toxicities. Therefore, a level that indicates a “green” color would symbolize a concentration reading that is safe. Similarly, a “yellow” reading would indicate a somewhat hazardous concentration and a “red” reading would signify a dangerous concentration level.

In addition to the RAE icons, I also placed icons for the chemical release locations as well as the GASMET instrument locations. The four GASMET icons are also tracking concentration levels throughout the trial and have been assigned the color-coded values as well. The five release locations (labeled D1-D5) are shown in the visualization as grey icons. These will not change color during the trial and are representative of the possible chemical release location pans in the test chamber.

In order to have the visualization determine the correct icon color, I had to develop the following formulas:[4]

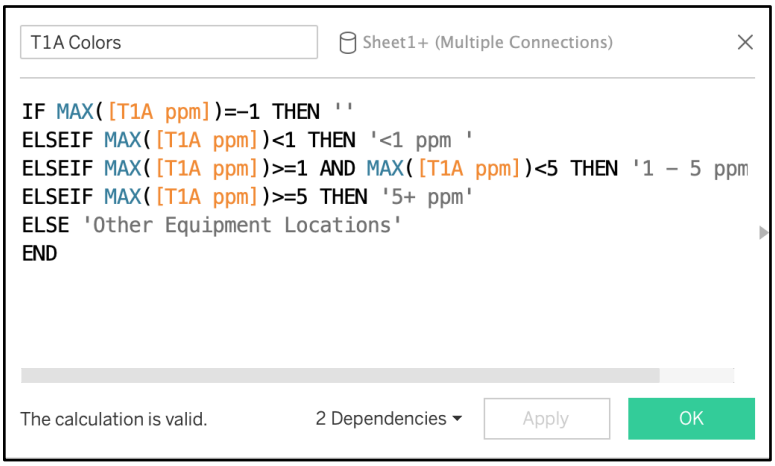

Figure 7. Tableau Formulas that Create Color Changing Icons

The bin sizes created above were then assigned their particular colors and the formulas in Figure 7 did the rest.

The final result of quadrant 1 is shown in Figure 8.

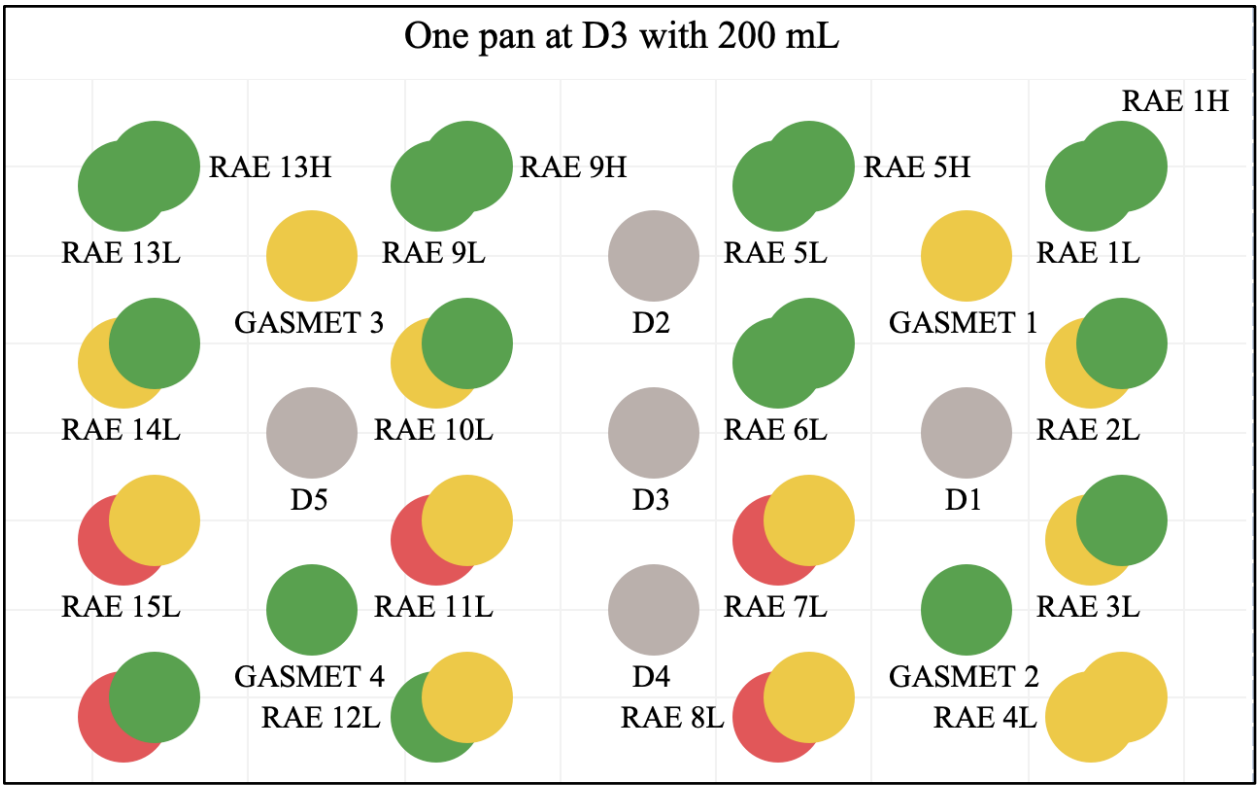

Figure 8. Quadrant 1 Visualization

The title of the quadrant indicates that one pan was being used for the chemical release location and that pan was positioned at D3 and was filled with 200 mL of chemical.

BUILDING QUADRANT 2

Quadrant 2 consists of the GASMET instrument data. Each of the GASMETs returned concentration values for three specific chemicals: diethyl sulfide, iso-valeric acid, and t-butyl mercaptan. The concentrations were measured roughly every 20 seconds during the hour-long trial. The visualization itself consists of three line charts overlaid on a uniform x-axis scale. The result is shown in Figure 9.

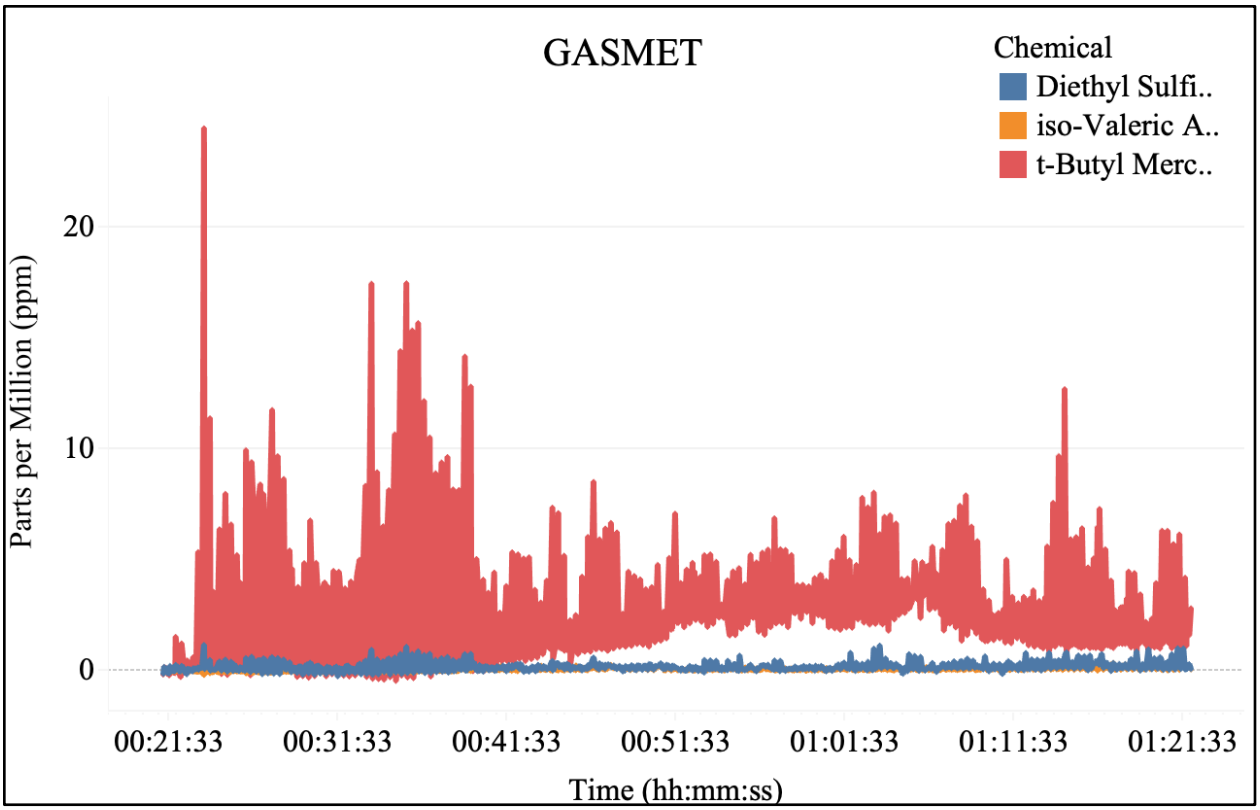

Figure 9. Quadrant 2 Visualization

Each of the chemicals that were detected by the GASMET instruments are shown with color-specific line plots. Since we are using four different GASMET instruments, each chemical line plot is the summed value from each instrument.

BUILDING QUADRANTS 3 AND 4

Each of these quadrants will represent the RAE instruments at the different height levels. Quadrant 3 will be for the 66" high instruments and quadrant 4 will be for the 18" high instruments. I will also separate these quadrants into maximum and minimum concentration values that were detected at each timestamp. This configuration was done in order to provide a three-dimensional cross section view of the different concentration levels that were measured during the trial.

As was described for Quadrant 1, Quadrants 3 and 4 will also be color-coded in response to detected concentration ranges. Figure 10 shows the layout of the visualization quadrants.

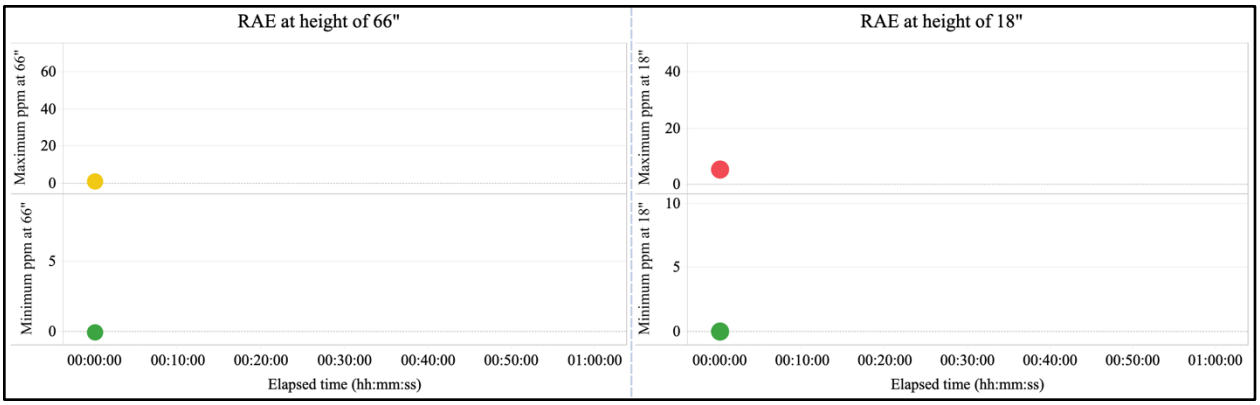

Figure 10. Quadrants 3 and 4 Visualizations

BUILDING THE INTERACTIVE ELEMENTS

The main feature of the visualization dashboard is to make it interactive, which allows the user to drive the data story and drill-down on individual data elements. To accomplish this, I had to use a Tableau feature called the Tableau Pages Shelf.

In essence, the Pages shelf lets you break a view into a series of pages so you can better analyze how a specific field affects the rest of the data in a view. When you place a dimension on the Pages shelf you are adding a new row for each member in the dimension. The Pages shelf creates a set of pages, with a different view on each page. Each view is based on a member of the field you placed on the Pages shelf. You can easily flip through the views and compare them on a common axis, using the controls that get added to the view when you move a field to the Pages shelf.[5]

By dragging the elapsed time column (from the Elapsed time_sync data file) to the Pages shelf, Tableau created a playback capability that would step through each timestamp element in a movie-like fashion.[6-7] The user can select from three different playback speeds and observe the changes in concentration for each of the RAE and GASMET sensors. Separate playback controls were generated for Quadrants 1, 3, and 4. I was able to combine all three controls into a master control that would start the playback for all three quadrants simultaneously. Double clicking the embedded movie below will demonstrate this capability.

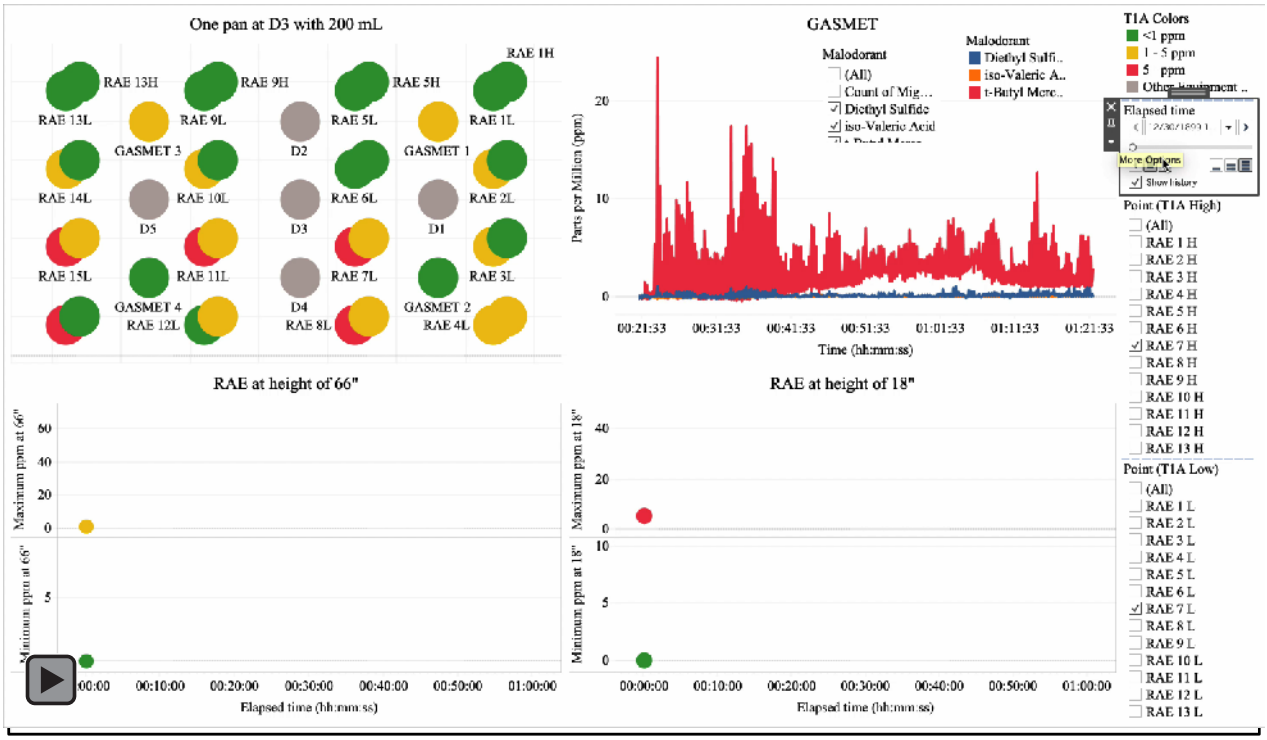

Figure 11. Movie Demonstrating Interactive Visualizations

The playback speed for this movie is set at the fastest level and it not the typical speed one would use for visualizing the data. The fastest speed was selected simply to keep the file size for the movie at a reasonable level. To fully experience the interactive nature of the visualization, the reader is directed to the accompanying Tableau Workbook.

For Quadrant 1 (upper left), each of the sensors will change color based on the concentration measured at each timestamp. For Quadrants 3 and 4 (bottom), the Show History box is checked in the playback control panel. This allows for a history trail to be displayed indicating the historical concentrations recorded throughout the trial.[8] This feature allows the user to get a better understanding of how the concentrations are changing over time as well as get an understanding of the cross-sectional concentration profile. The reader should note the individual scales on the y-axes for Quadrants 3 and 4. They will automatically change based on the selected instruments in the filters.

In the above movie, the user can select which RAE instrument concentrations they wish to monitor during the playback. For the included example, RAE 7 H and RAE 7 L are selected. The user can select any number of combinations of instruments they wish to observe. If two or more instruments are selected, the total summed concentration will be displayed.

Quadrant 2 (upper right) does not have any time dependent color changes. For this quadrant, I wanted to show the total GASMET concentration plots for the entire trial. The user can select/deselect the desired chemical traces they wish to see by the included check box filter. In the above example, all three chemicals are displayed.

VISUALIZATION DEPLOYMENT

Since I am the only one at my work that has a copy of the Tableau software, I had to find a way to share this visualization with the testing community and scientists that wanted access to the data. The U.S. Army has a Tableau Server[9] installed in Washington D.C. for use throughout the Department of Defense (DoD). I reached out to them to obtain access to this server site so that I could post this visualization. I was granted access and established a folder specific to my organization. One key element to this Tableau Server site is that it requires a secure login that is only available with a DoD issued Common Access Card (CAC). This security feature allows me to control who has access to our organizational folder and who has access to post material to it. It also helps ensure that our data is kept secure and only available to those with a need to know.

FUTURE WORK

I have posted this visualization to the Tableau Server and granted about 10 people access to it. It has been very well received and the testing community has expressed interest in me developing additional visualizations for them. To that end, my work has purchased three additional licenses for Tableau because of this project. I will continue to refine the visualization as requests for additional capabilities and/or views are made. This was an extremely rewarding project and one that was of great interest for my organization.

REFERNCES

[1] RAE Systems. (n.d.). Find Your Gas Detector Solution. Retrieved from https://www.raesystems.com/

[2] RAE Systems by Honeywell. (2019). *A Guideline For Pid Instrument Response*.

[3] Gasmet. (2020, April 24). Gasmet DX4015 is a portable FTIR gas analyzer for ambient air analysis. Retrieved from https://www.gasmet.com/products/category/portable-gas-analyzers/dx4015/

[4] Tableau. (2015). Conditional Formatting. Retrieved from https://community.tableau.com/thread/107750

[5] Tableau. (2020). Shelves and Cards Reference. Retrieved from https://help.tableau.com/current/pro/desktop/en-us/buildmanual_shelves.htm

[6] Sullins, B. (2019, April 5). Animate your Tableau dashboards on the Web. Retrieved from https://www.pluralsight.com/blog/data-professional/tableau-animation

[7] Nair, A., Gall, R., Lobo, S., & Davis, V. (2018, March 26). Building Motion Charts with Tableau for data storytelling. Retrieved from https://hub.packtpub.com/building-motion-charts-tableau/

[8] Page shelf - Playback and Show History functionality. (n.d.). Retrieved from https://community.tableau.com/ideas/3844

[9] Tableau. (2020). Get Started with Tableau Server on Windows. Retrieved from https://help.tableau.com/current/server/en-us/get_started_server.htm