

A Snazzy Graph Featuring Five Variables, Axis Breaks, and Good Resolution

Rachel A. Baker, Savannah River National Laboratory, Aiken, SC
Anthony T. Baker, Westinghouse Savannah River Company, Aiken, SC

ABSTRACT

Monitoring groundwater contamination is of great concern at operational sites in the United States Department of Energy (DOE) complex. The Savannah River Site (SRS) in Aiken, South Carolina is one such site. Elevated measurements of radium-226 and radium-228 were recently observed in numerous monitoring wells at the SRS Sanitary Landfill (SLF). The issue at hand was whether these elevated data are consistent with groundwater impacted by SRS operations or whether they are typical of natural conditions. Graphical data displays are important tools for gaining an understanding of the data and of relationships between various parameters. SAS[®] software was used to develop graphs of the radium isotopes versus other analytes, including field pH, for the SLF wells. Five variables were depicted in each well plot: analyte, time, measurement, lab, and detection. Two different result axes were required, in addition to breaks on the pH axis. Clarity is an issue for graphical output to be displayed in final reports. SAS software provided the means to create high resolution graphs of many dimensions with vertical axis breaks. The weight of evidence from these and other graphs, in addition to statistical and geochemical analyses, supported a conclusion of naturally-occurring radium in the SRS SLF.

INTRODUCTION

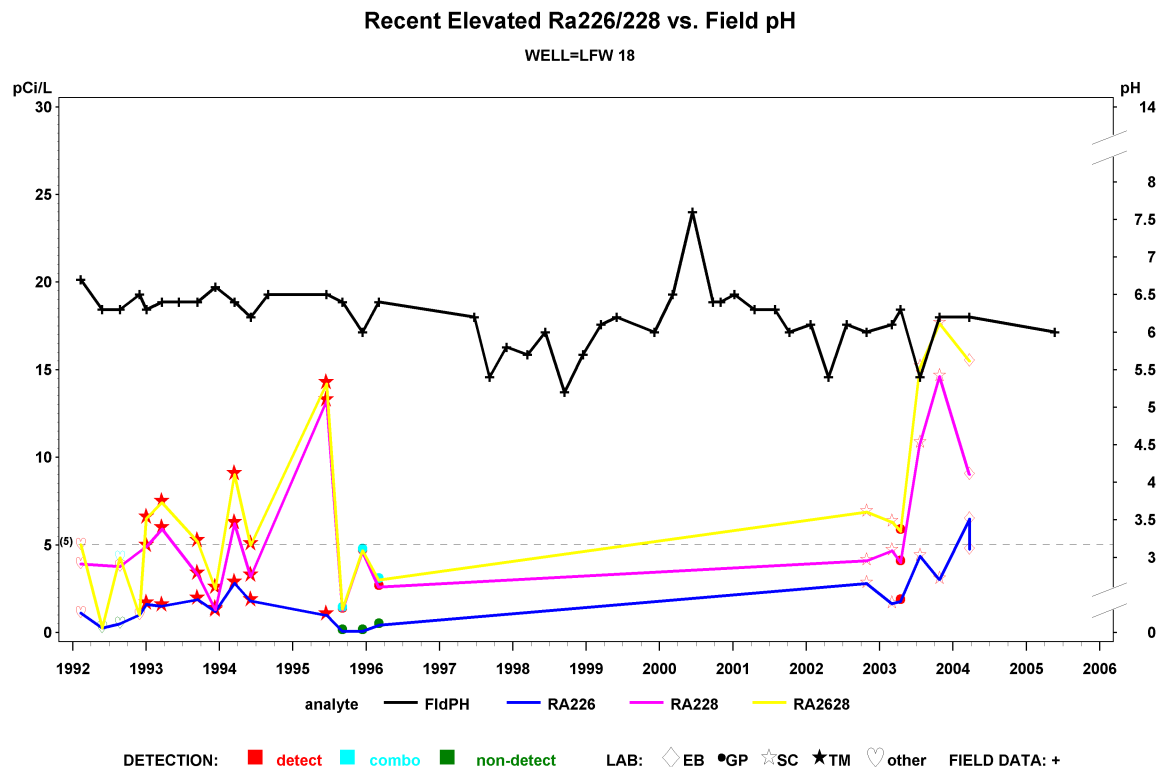
The Westinghouse Savannah River Company operates the Savannah River National Laboratory and the Savannah River Site (SRS) for the United States Department of Energy (DOE). Current SRS missions include defense programs, non-proliferation of nuclear weapons, environmental remediation and restoration, waste management, chemical processing, research and development, and technology transfer. Groundwater contamination arising from legacy weapons work is a significant environmental issue at DOE sites.

In 2002, elevated radium-226 and radium-228 measurements were noted in several groundwater monitoring wells at the SRS Sanitary Landfill (SLF). Of concern was the origin of the radium, whether the elevated radium was the byproduct of SRS operations or was naturally-occurring. A key aspect of the analysis was to gain an understanding of the data and of relationships between various parameters. For an individual well, it was of interest to examine the relationship between the two radium isotopes and field pH. The groundwater protection standard is 5 picocuries per liter for radium-226/228 (Ra2628), the combination of radium-226 and radium-228. The levels of the four analytes (Ra226, Ra228, Ra2628, and field pH) needed to be displayed through time for each well. Specialized graphs for the SLF wells were developed using SAS/GRAPH[®] software with the Annotate facility. On each graph, the horizontal axis depicts time. The primary vertical axis shows the concentration of the radioactive analytes, and the secondary vertical axis has pH data. Another wrinkle is that pH has a scale of 0 to 14, but the SLF pH data are between 3 and 8. Rather than compress the data or display only a section of the true pH scale, breaks were drawn on the pH axis using the Annotate facility to alert the viewer to the range of the data. Other factors such as analytical laboratory and detection are also of interest for groundwater data. Some data values are actual measurements, while others are detection limits reported by the lab for data that are found to be below the instrument detection levels. Typically, half of the detection limit is substituted for nondetect data. Since Ra2628 is derived from Ra226 and Ra228, some of the values are a combination of a detect and a nondetect. The Annotate facility was utilized to create specific symbols in which symbol color illustrates detection and symbol shape indicates the lab. The data for individual analytes are distinguished by connected lines. Since certain graphs were to be included in final reports, the clarity of the output was important. Graphics options control the production of graphs with a resolution of 600 dots per inch (dpi).

SAS software was utilized to construct specialized graphs to exhibit five dimensions of the data: results (1) for four analytes (2) through time (3) with lab (4) and detection (5) specified and to generate these graphs with high clarity and axis breaks. Information-rich graphs produced via SAS software, in conjunction with statistical and geochemical analyses, helped demonstrate that the levels of radium in the SRS Sanitary Landfill are not inconsistent with those of natural radium in the surrounding area.

SAMPLE GRAPH

The following sample graph is for a particular SLF monitoring well. Of note are the five dimensions of data, the two breaks in the pH axis, and the resolution of the output.



SNIPPETS OF SAS CODE

Graph of Five Dimensions of Data

The SAS Annotate facility was used to create plot points with color for detection and symbol for lab. The horizontal value is time, and the vertical component is the measurement. To differentiate analytes, the PLOT statement of type $x*y=z$ was employed in conjunction with a SYMBOL statement having the interpolation option of JOIN to draw lines. A PLOT2 statement added on the pH data, and regular symbols rather than annotated symbols were appropriate since detection and lab are not of concern for field data such as pH.

```
data symb_ra;
  length style color $7 text $3;
  retain function hsys xsys ysys
    ("label" "1" "2" "2")
    ;
  set recent_ra;
  position="+";
  x=sampdate;
  y=adj_ra;
  style='special'

  if detqual=1 then color='red';
  else if detqual=0 then color='green';
  else color='cyan';
```

```

    if lab='EB' then do;
        text='D';
        size=3;
    end; ...

symbol1 v=none i=join width=8;
symbol2 f="Arial/bold" v='+' h=1.25 i=join width=8;

proc gplot data=recent_ra_ph uniform;
    format sampdate year4. ph best.;
    where well='LFW 18' and sampdate>=&mindate and sampdate<=&maxdate;
    by well;
    plot adj_ra*sampdate=analyte/ anno=symb_ra frame vaxis=axis1 haxis=axis2
        lvref=2 vref=5 name="LFW_18";
    plot2 ph*sampdate=analyte/ anno=axisbrks vaxis=axis3 haxis=axis2
        nolegend;
run;
quit;

```

Breaks for pH Axis

On its Web site, SAS Institute, Inc. has a sample program (gplot-broken.html) that draws a lightning bolt type line for an axis break. That program was used as a starting point to create a more traditional axis break of angled parallel lines for the pH data. Either way, the approach is to draw a white box on top of the axis and to then draw lines to indicate the axis break. The following code sets up two breaks for the pH axis to show that even though pH values can be between 0 and 14, the SLF pH data are between 3 and 8. The axis values had to be carefully chosen for the breaks to appear in the right places. Note that it took some tweaking of the numbers to get it just right.

```

data annobrks;
    length function style color $8;
    retain xsys '5' ysys '2' when 'a' style 'solid';

    * solid white box to cover vaxis area for lower break;
    function='move'; x=93; y=.75; output;
    function='bar'; x=100; y=1.6; color='white'; output;

    * lines for lower vaxis break;
    color='black'; size=1.5;
    function='move'; xsys='1'; x=98; ysys='2'; y=.5; output;
    function='draw'; xsys='B'; ysys='B'; x=+3; y=+2; output;
    function='move'; xsys='1'; x=98; ysys='2'; y=1.2; output;
    function='draw'; xsys='B'; ysys='B'; x=+3; y=+2; output;

    * solid white box to cover vaxis area for upper break;
    function='move'; xsys='5'; x=93; ysys='2'; y=10.65; output;
    function='bar'; x=100; y=12.5; color='white'; output;

    * lines for upper vaxis break;
    color='black'; size=1.5;
    function='move'; xsys='1'; x=98; ysys='2'; y=9.85; output;
    function='draw'; xsys='B'; ysys='B'; x=+3; y=+2; output;
    function='move'; xsys='1'; x=98; ysys='2'; y=12; output;
    function='draw'; xsys='B'; ysys='B'; x=+3; y=+2; output;
run;
...
axis3 label=('pH') order=(0, 1.25, 3 to 8 by .5, 12, 14) ...;

```

Options for High Resolution Graphics

To produce output of high resolution, a PNG device with particular size and pixels options was specified. Note that the number of pixels is 600 times the x/y dimensions to produce a graph of 600 dpi. The thickness of the axis lines, values, and such also had to be increased.

```
goptions reset=all rotate=landscape colors=(black blue magenta yellow)
device=png target=png xmax=6in ymax=4in xpixels=3600 ypixels=2400
ftext="Arial/bold";

axis1 w=2 major=(w=3) ...;
```

CONCLUSION

Sometimes the only good way to get a handle on data is graphically. This paper explains how SAS software can be used to create highly informative, individualized graphs with excellent clarity.

- SAS/GRAPH software with graphics options and symbol statements and the Annotate facility can be employed to create specialized graphs of many dimensions.
- The SAS Annotate facility can be utilized to cover up part of an axis and then draw lines on the axis to indicate a break in the values. The SAS Institute, Inc. Web site has a sample program that draws lightning bolt type lines for the axis break, but angled parallel lines are more traditional.
- A SAS/GRAPH software GOPTIONS statement that specifies a DEVICE of PNG and appropriate values for XMAX, YMAX, XPIXELS, and YPIXELS will create high definition output.

Graphical displays developed using SAS software can be crucial in solving complex problems. In this case, such graphs helped address the issue of the source for elevated radium in the Sanitary Landfill at the U.S. Department of Energy's Savannah River Site in Aiken, South Carolina.

REFERENCES

SAS Institute, Inc. (1999), *SAS/GRAPH Software: Reference, Version 8, Volumes 1-3*, Cary, NC: SAS Institute, Inc.

SAS Institute, Inc. (2002), *Sample Graphs and Code Using PROC GPLOT: gplot-broken.html*.
<<http://ftp.sas.com/techsup/download/sample/graph/gplot-broken.html>>.

CONTACT INFORMATION

Please feel free to address questions or provide feedback to the authors.

Rachel A. Baker
Savannah River National Laboratory
Statistical Consulting Section
Building 773-42A, Room 162
Savannah River Site
Aiken, South Carolina 29808
(803)725-3199
rachel.baker@srnl.doe.gov

Anthony (Andy) T. Baker
Westinghouse Savannah River Company
Information Technology Department
Building 773-51A, Room 108
Savannah River Site
Aiken, South Carolina 29808
(803)725-2112
andy.baker@srs.gov

SAS and all other SAS Institute, Inc. product or service names are registered trademarks or trademarks of SAS Institute, Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.