

Export Customized Graphs Directly into Microsoft Word Using Dynamic Data Exchange

Dennis J. Beal, Science Applications International Corporation, Oak Ridge, Tennessee

ABSTRACT

Highly customized graphs can be created with SAS® 9.1.3 and exported directly into Microsoft Word using dynamic data exchange (DDE) and SAS macros. SAS code is shown that produces side by side normal probability plots and box plots in both linear and log₁₀ scales for environmental data. The probability plots are customized using annotation to label potential outliers and create hollow symbols for censored results reported below the detection limit. A linear regression line is shown fitted to the environmental results on a log₁₀ scale. The box plots are also customized using annotation to show the mean, the 95% upper confidence limit on the mean, and special symbols for results outside the inner and outer fences of 1.5 and 3 times the interquartile range above the 75th percentile. SAS code displays the distribution (normal, lognormal, or nonparametric) on the graphs and determines the optimal number of tick marks for the vertical axes. DDE is used to open Word directly from within SAS and insert multiple pages of graphs into Word using SAS macros with unique figure titles and numbers. SAS code presented in this paper can be used for creating customized graphs for hundreds of constituents quickly. This paper is for intermediate SAS users of SAS/STAT and SAS/GRAPH.

Key words: dynamic data exchange, annotate, Microsoft Word, probability plots, box plots, macros

INTRODUCTION

Environmental scientists and risk assessors collect soil and water samples that are often analyzed for hundreds of chemicals and radionuclides in the analytical laboratory. The distributions of each of these constituents must be determined to properly calculate summary statistics, 95% upper confidence limits (UCLs) on the mean, and exposure point concentrations. The process for evaluating the distributions of each constituent involves both quantitative tests and graphical displays. This paper shows SAS code that can be used to create customized graphs and export them directly into Microsoft Word using dynamic data exchange (DDE) with unique figure titles for a large number of constituents. The SAS code presented in this paper uses the SAS System for personal computers version 9.1.3 running on a Windows XP Professional platform with Service Pack 3.

SAS CODE

The SAS code used to generate the graphs and export them directly into Microsoft Word using DDE is shown by topic as it appears in sequence in the SAS program.

INITIAL SAS MACROS

The following SAS macro variables that will be used later in the program are defined. These initial macros assume the data have already been stored in a SAS data set and reduced by removing duplicate samples in the INPDS macro variable. Furthermore, the summary statistics should have already been calculated in a SAS data set and are contained in the INPDS_SUM macro variable.

```
%let pathdat = C:\HHRA\Data\;      ** location of data;
%let pathtab = C:\HHRA\Tables\;    ** location of tables;
libname hhra "&pathdat";
%let pathplot = C:\HHRA\Plots\;    ** location of plots;
%let INPDS = hhra.soil_data;        ** soil data with individual results;
%let INPDS_SUM = hhra.sumstats_surfsoil; ** summarized surface soil data set;
%let WORDFN = HHRA plots for detected chemicals.doc; ** name of Word file;
%let NUMTICKS = 10;                ** desired number of major tick marks on the Y axis;
```

The macro OBSNVARs returns the number of variables and observations from a data set.

```
%macro obsnvars(ds);
  %global dset nvars nob;
  %let dset=&ds;
```

```

%let dsid = %sysfunc(open(&dset));
%if &dsid %then %do;
  %let nobs = %sysfunc(attrn(&dsid, NOBS));
  %let nvars = %sysfunc(attrn(&dsid, NVAR));
  %let rc = %sysfunc(close(&dsid));
%end;
%else %put Open for data set &dset failed - %sysfunc(sysmsg());
%mend obsnvars;

```

PREPARE DATA SETS

The following SAS code prepares the input data sets for creating the graphs by calculating the range of results that will be plotted on the graphs.

```

proc sort data=&INPDS out=plotdat; by med_type anagroup chemical units;
proc summary data=plotdat; ** find the largest and smallest individual result ;
  var results;
  by med_type anagroup chemical units;
  output out=minmaxout(drop=_type_ _freq_) min=MINRES max=MAXRES; run;
proc sort data=&INPDS_SUM out=sumstats; by med_type anagroup chemical units;

data minmaxout2;
  merge minmaxout
        sumstats(keep=med_type anagroup chemical units meanprox ucl95 sflag totn pdetc);
  by med_type anagroup chemical units;
  PLOTMIN = min(minres, meanprox); ** find smallest result or mean of proxy values;
  PLOTMAX = max(maxres, ucl95); ** find largest result or UCL95;
  PLOTNUM = _N_; run; ** plot number for each constituent to be plotted;

```

The variable PLOTMIN calculates the minimum of the individual results and the mean of the proxy values because for distributions with non-detected chemical results, half the non-detect is usually used as a proxy value for the censored result. Therefore, it is possible the mean could be smaller than the minimum detected result if there are many non-detects for a particular constituent. The UCL could exceed the maximum result for skewed distributions. The variable SFLAG is the distribution (normal, lognormal, or nonparametric), TOTN is the total number of results, and PDETC is the frequency of detection.

The following SAS code merges the summary statistics with the individual results for probability plots. The macro variable UNITS stores the units associated with the results.

```

proc sort data=minmaxout2; by med_type anagroup chemical units;
proc sort data=plotdat; by med_type anagroup chemical units results;

data plotdat2;
  merge plotdat minmaxout2;
  by med_type anagroup chemical units;

proc sort data=plotdat2; by plotnum med_type anagroup chemical units;
proc sort data=plotdat2 out=unitsonly(keep=units) nodupkey; by units;

data _null_;
  set unitsonly; call symput('UNITS'||left(_N_), trim(left(units))); run;

```

VALUES FOR THE VERTICAL AXIS

Extensive SAS code is available from the author that calculates the optimal range of values to be displayed for the vertical axis for each constituent, the number of major tick marks closest to the desired number of major tick marks declared in the macro variable NUMTICKS, and the number of minor tick marks. These calculations are then stored in macro variables. Due to space limitations, the code is not included in this paper, but can be obtained from the author. The macro variables DIST, MED, CHEM, CHEMNUM, UNITS, UNITSFN, VM, MINTICK, MAXTICK, BYVAR, FORMAT_LOGRES, FORMAT_RES, and N_PLOTS are assigned during the execution of this code.

PROBABILITY PLOT CALCULATIONS

This SAS code merges the vertical axis information with the individual results and calculates the normal deviates from the normal cumulative distribution function (CDF). The variable HIT = 1 if the result is detected and 0 if the result is not detected.

```

proc sort data=plotdat2; by med_type anagroup chemical units;
proc sort data=aggs2; by med_type anagroup chemical units;

data plotdat5;
merge plotdat2 aggs2;
by med_type anagroup chemical units;
attrib _char_ format=$varying. informat=$varying.;

proc sort data=plotdat5; by plotnum med_type anagroup chemical units results hit;

data plotdat6;
set plotdat5;
by plotnum med_type anagroup chemical units;
if first.units then i=0; i+1;
phiinv_p = probit(i / (totn + 1)); run;

```

PROBABILITY PLOT REGRESSION LINE FOR LOG₁₀ SCALE

This SAS code calculates the linear regression line that will be plotted in log₁₀ scale.

```

proc sort data=plotdat6; by plotnum pdetc results;
proc reg data=plotdat6; ** used for plotting linear regression lines in log10 space;
model log10res = phiinv_p;
by plotnum pdetc;
output out=plotdat7 r=residual p=predicted; run; quit;

data plotdat7;
set plotdat6;
plot2res = 10**(predicted); run;

```

This SAS code ensures the scales of both vertical axes in the probability plots are the same in log₁₀ scale.

```

** add missing values for min and max so right Y axis matches left Y axis;
proc sort data=plotdat6; by plotnum results;
data minout;
set plotdat6; by plotnum; where results ^=.;
if first.plotnum;
plot2res = results;
phiinv_p = .;

proc sort data=plotdat6; by plotnum descending results;
data maxout;
set plotdat6; by plotnum; where results ^=.;
if first.plotnum;
plot2res = results;
phiinv_p = .;

proc sort data=plotdat7; by plotnum plot2res;
data minregout;
set plotdat7; by plotnum; where plot2res ^=.;
if first.plotnum;
results = plot2res;
phiinv_p = .;

proc sort data=plotdat7; by plotnum descending plot2res;
data maxregout;
set plotdat7; by plotnum; where plot2res ^=.;
if first.plotnum;
results = plot2res;
phiinv_p = .;

data plotdat8; set plotdat7 minout maxout minregout maxregout; run;

```

OPEN MICROSOFT WORD FROM WITHIN SAS

This SAS code opens Microsoft Word from within SAS.

```

options noxsync noxwait xmin;
filename sas2word dde 'winword|system';
data _null_;
  length fid rc start stop time 8;
  fid=fopen('sas2word','s');
  if (fid le 0) then do;
    rc=system('start winword');
    start=datetime();
    stop=start+100;
    do while (fid le 0);
      fid=fopen('sas2word','s');
      time=datetime();
      if (time ge stop) then fid=1;
    end;
  end;
  rc=fclose(fid); run;

data _null_;
  file sas2word;
  put '[AppMinimize]';
  put '[FileNew.Template="Normal.dot"]';
  put '[ChDefaultDir ""@; put "&pathplot"@; put "",0]';
  put '[EditBookmark.Name="HHRA",.Add]'; run;

```

DEFINE MACRO FOR GENERATING PLOTS

This SAS code defines the macro for generating probability and box plots for one or many constituents. The interquartile range and possible outliers are calculated. The PROC GREPLAY deletes all previous plots from GSEG so the most current plots are used.

```

%macro plots(startno, endno);
  %do i = &startno %to &endno;
    data one&i.;
      set plotdat6; where plotnum=&i.; run;

    proc univariate data=one&i. noprint; var results; output out=uclout q1=p25 q3=p75;

    data merged;
      set one&i.;
      if _N_=1 then set uclout;
      IQR = p75 - p25;
      if 3*iqr + p75 > results >= 1.5*iqr + p75 or p25 - 3*iqr < results <= p25 - 1.5*iqr
        then OUTLIER1 = '#';
      if results >= 3*iqr + p75 or results <= p25 - 3*iqr then OUTLIER2 = '##';
    run;

    proc greplay nofs; igout work.gseg; delete _all_; run; quit;
  %enddo;
%macro end;

```

PROBABILITY PLOT IN LINEAR SCALE

This SAS code creates a normal probability in linear scale. The annotation data set ANNO1 identifies all the non-detects that will make the symbol appear hollow to distinguish them from detected concentrations. The annotation data set ANNO1b puts the distribution label in the top left corner of the plot. The data set ANNO1c labels the potential outliers on the plots.

```

filename plot "&pathplot.probability plot &&MED&i &&CHEM&i &&UNITSFN&i outlier.cgm";

goptions reset=all display gsfname=replace noprompt gsfname=plot device=CGMOFMP
  rotate=portrait htext=0.75 ftext="TimesRoman" horigin=0 in vorigin=0 in vsize=4.2
  in hsize=3.2 in cback=white;

data anno1; /* create annotate dataset;
  length function color style $ 8;
  set one&i.; where hit=0; ** keep only non-detected results;
  size=0.3; xsys='2'; ysys='2'; hsys='4'; x = phiinv_p; y = results; style='marker';
  text='W'; color='snow'; when='a'; position='5'; function='SYMBOL'; run;

```

```

%obsnvars(anno1); %let ND = &nobs;

data anno1b;
  length function color style $ 8 text $15;
  function='LABEL'; size=0.75; xsys='1'; ysys='1'; hsys='4'; x = 3; y = 97;
  style=' '; text="&&DIST&i."; color='black'; when='a'; position='6'; output; run;

data anno1c;
  length function color style $ 8 text $35;
  set merged; where outlier1^=' ' or outlier2^=' '; ** keep only potential outliers;
  function='LABEL'; size=0.7; xsys='2'; ysys='2'; hsys='4'; x = phiinv_p - 1;
  y = results; style=' '; text=station; color='black'; when='a'; position='6';
  output;

data anno; length text $50; set anno1 anno1b anno1c; run;

proc gplot data=one&i. annotate=anno;
  plot results*phiinv_p / vaxis=axis1 haxis=axis2 vm=&&VM&i. hm=1 name='plot1'
  nolegend;
  format results &&FORMAT_RES&i. phiinv_p 3.0;
  axis1 width=5 value=(h=0.75) order=(&&MINTICK&i. to &&MAXTICK&i. by &&BYVAR&i.)
  label=(a=90 r=0 h=0.75 j=c "&&CHEM&i. (&&UNITS&i.)");
  axis2 origin=(1 in) length=2.2 in width=5 value=(h=0.75) order=(-2 to 2 by 1)
  offset=(0.1 in) label=(h=0.75 'Deviates from the Normal CDF');
  label results='00'x phiinv_p='00'x project_pdetc='00'x;
  footnote1 j=l h=0.65 ' ';
  footnote2 j=l h=0.2 ' ';
  symbol1 f=marker v='W' i=rl h=0.7 c=blue w=5 l=1; run; quit;

```

PROBABILITY PLOT IN LOG₁₀ SCALE

This SAS code creates a normal probability plot in log₁₀ scale. The same annotation data set ANNO used for the previous probability plot is used for the probability plot in log₁₀ scale. The footnote about hollow symbols is printed only when non-detects are present in the graph by referencing the macro variable ND.

```

data one8_&i.;
  set plotdat8; where plotnum=&i.; run;
  proc sort data=one8_&i.; by plotnum results; run;

filename plot "&pathplot.probability plot &&MED&i &&CHEM&i log10 &&UNITSFN&i. .cgm";

goptions reset=all display gsfname=replace noprompt gsfname=plot device=CGMOFMP
  rotate=portrait htext=0.75 ftext="TimesRoman" horigin=0 in vorigin=0 in vsize=4.2
  in hsize=3.2 in cback=white;

proc gplot data=one8_&i. annotate=anno;
  plot results*phiinv_p / vaxis=axis1 haxis=axis2 vm=9 hm=1 name='plot2' nolegend;
  plot2 plot2res*phiinv_p / vaxis=axis3 vm=9 nolegend noaxis;
  format results &&FORMAT_LOGRES&i. phiinv_p 3.0 plot2res 1.0;
  axis1 width=5 value=(h=0.75) logbase=10 logstyle=expand
  label=(a=90 r=0 h=0.75 j=c "&&CHEM&i. (&&UNITS&i.) in Log10 scale");
  axis2 origin=(1 in) length=2.2 in width=5 value=(h=0.75) order=(-2 to 2 by 1)
  offset=(0.1 in) label=(h=0.75 'Deviates from the Normal CDF');
  axis3 width=5 value=(h=0.75) logbase=10 label=none;
  label results='00'x phiinv_p='00'x project_pdetc='00'x plot2res='00'x;
  %if &ND > 0 %then %do;
    footnote1 j=l h=0.65 'Hollow symbols denote non-detects.'; %end;
  %else %if &ND = 0 %then %do;
    footnote1 j=l h=0.65 ' '; %end;
  footnote2 j=l h=0.2 ' ';
  symbol1 f=marker h=0.7 v='W' i=none l=1 w=5 c=blue; run; quit;

```

BOX PLOT IN LINEAR SCALE

This SAS code creates a box plot in linear scale. The annotation data set ANNO1 identifies the mean that will be displayed as a black star.

```

data merged;
  set one&i.;
  if _N_=1 then set &INPDS_SUM;
  IQR = p75 - p25;
  if 3*iqr + p75 > results >= 1.5*iqr + p75 or p25 - 3*iqr < results <= p25 - 1.5*iqr
    then OUTLIER1 = '#';
  if results >= 3*iqr + p75 or results <= p25 - 3*iqr then OUTLIER2 = '##'; run;

data anno1;          ** display mean symbol;
  length function color style $ 8 text XC $15 ;
  set merged;
  if _N_=1;
  xsys='2'; ysys='2'; function='SYMBOL'; xc=pdetc ; y=meanprox; hsys='4'; size=0.7;
  style='marker'; text='V'; color='black'; position='6'; when='a'; output; run;

```

The annotation data set ANNO2 identifies the 95% (UCL) on the mean that will be displayed as a green horizontal bar.

```

data anno2;          ** displays UCL95;
  length function color style $ 8 text XC $15 ;
  set merged;
  if _N_=1;
  xsys='2'; ysys='2'; function='SYMBOL'; xc=pdetc ; y=ucl95; hsys='4'; size=1.3;
  style='music'; text='J'; color='green'; position='6'; when='a'; output; run;

```

The annotation data set ANNO3 identifies the individual results outside the inner fence between 1.5 and 3 times the interquartile range above the 75th percentile that will be displayed as blue solid circles.

```

data anno3;          ** little outliers;
  length function color style $ 8 text XC $15 ;
  set merged; where outlier1='#';
  xsys='2'; ysys='2'; function='SYMBOL'; xc=pdetc ; y=results; hsys='4'; size=0.7;
  style='marker'; text='W'; color='blue'; position='6'; when='a'; output; run;

```

```
%obsnvars(anno3); %let OUT1 = &nobs;
```

The annotation data set ANNO4 identifies the individual results outside the outer fence beyond 3 times the interquartile range above the 75th percentile that will be displayed as red solid squares.

```

data anno4;          ** big outliers;
  length function color style $ 8 text XC $15 ;
  set merged; where outlier2='##';
  xsys='2'; ysys='2'; function='SYMBOL'; xc=pdetc ; y=results; hsys='4'; size=0.7;
  style='marker'; text='U'; color='red'; position='6'; when='a'; output; run;

```

```
%obsnvars(anno4); %let OUT2 = &nobs;
```

The annotation data set ANNO3a identifies the individual censored or non-detected results outside the inner fence between 1.5 and 3 times the interquartile range above the 75th percentile that will be displayed as blue hollow circles.

```

data anno3a;          ** little outliers non-detects;
  length function color style $ 8 text XC $15 ;
  set merged; where outlier1='#' and hit=0;
  xsys='2'; ysys='2'; function='SYMBOL'; xc=pdetc ; y=results; hsys='4'; size=0.3;
  style='marker'; text='W'; color='snow'; position='6'; when='a'; output; run;

```

The annotation data set ANNO4a identifies the individual censored or non-detected results outside the outer fence beyond 3 times the interquartile range above the 75th percentile that will be displayed as red hollow squares.

```

data anno4a;          ** big outliers non-detects;
  length function color style $ 8 text XC $15 ;
  set merged; where outlier2='##' and hit=0;
  xsys='2'; ysys='2'; function='SYMBOL'; xc=pdetc ; y=results; hsys='4'; size=0.3;
  style='marker'; text='U'; color='snow'; position='6'; when='a'; output; run;

```

The annotation data set ANNO5 displays the legend for the mean, UCL, and outliers. The legend will include symbols for outliers only if outliers are identified on each plot using the OUT1 and OUT2 macro variables.

```
data anno5;
  length function color style $ 8 text $25 ;
  xsys='1'; ysys='1'; function='SYMBOL'; x=3; y=5; hsys='4'; size=1.2; style='music';
  text='J'; color='green'; position='6'; when='a'; output;
  xsys='1'; ysys='1'; function='LABEL'; x=7; y=5; hsys='4'; size=0.68; style=' ';
  text="UCL95"; color='black'; position='6'; when='a'; output;
  xsys='1'; ysys='1'; function='SYMBOL'; x=3; y=9; hsys='4'; size=0.68;
  style='marker'; text='V'; color='black'; position='6'; when='a'; output;
  xsys='1'; ysys='1'; function='LABEL'; x=7; y=9; hsys='4'; size=0.68; style=' ';
  text="mean"; color='black'; position='6'; when='a'; output;
  %if &OUT1 > 0 and &OUT2 > 0 %then %do;
    xsys='1'; ysys='1'; function='SYMBOL'; x=3; y=13; hsys='4'; size=0.65;
    style='marker'; text='W'; color='blue'; position='6'; when='a'; output;
    xsys='1'; ysys='1'; function='LABEL'; x=7; y=13; hsys='4'; size=0.68; style=' ';
    text="> 1.5xIQR"; color='black'; position='6'; when='a'; output;
    xsys='1'; ysys='1'; function='SYMBOL'; x=3; y=17; hsys='4'; size=0.65;
    style='marker'; text='U'; color='red'; position='6'; when='a'; output;
    xsys='1'; ysys='1'; function='LABEL'; x=7; y=17; hsys='4'; size=0.68; style=' ';
    text="> 3xIQR"; color='black'; position='6'; when='a'; output;
  %end;
  %if &OUT1 = 0 and &OUT2 > 0 %then %do;
    xsys='1'; ysys='1'; function='SYMBOL'; x=3; y=13; hsys='4'; size=0.65;
    style='marker'; text='U'; color='red'; position='6'; when='a'; output;
    xsys='1'; ysys='1'; function='LABEL'; x=7; y=13; hsys='4'; size=0.68; style=' ';
    text="> 3xIQR"; color='black'; position='6'; when='a'; output;
  %end;
  %if &OUT1 > 0 and &OUT2 = 0 %then %do;
    xsys='1'; ysys='1'; function='SYMBOL'; x=3; y=13; hsys='4'; size=0.65;
    style='marker'; text='W'; color='blue'; position='6'; when='a'; output;
    xsys='1'; ysys='1'; function='LABEL'; x=7; y=13; hsys='4'; size=0.68; style=' ';
    text="> 1.5xIQR"; color='black'; position='6'; when='a'; output;
  %end; run;
```

The data set ANNO combines the annotation data sets together.

```
data anno;
  length text XC $25;
  set anno1 anno1b anno2 anno3 anno4 anno3a anno4a anno5; run;
proc sort data=merged; by results; run;

filename plot "&pathplot.box plots &&MED&i &&CHEM&i &&UNITSFN&i linear.cgm";

goptions reset=all display gsfname=replace noprompt gsfname=plot device=CGM0FMP
  rotate=portrait htext=0.75 ftext="TimesRoman" horigin=0 in vorigin=0 in vsize=4.2
  in hsize=3.2 in cback=white;
```

The box plot is created on a linear scale.

```
proc gplot data=merged annotate=anno;
  plot results*pdetc / vaxis=axis1 haxis=axis2 vm=&&VM&i. hm=0 name='plot3';
  format results &&FORMAT_RES&i.;
  axis1 width=5 value=(h=0.75) label=(a=90 r=0 h=0.75 j=c "&&CHEM&i. (&&UNITS&i.)")
  order=(&&MINTICK&i. to &&MAXTICK&i. by &&BYVAR&i.);
  axis2 width=5 value=(h=0.75) label=(h=0.75 'Frequency of Detection');
  footnote1 j=l h=0.65 ' '; footnote2 j=l h=0.2 ' ';
  label results='00'x; symbol1 h=0.75 bwidth=11 i=box00tf l=1 w=5 c=blue cv=yellow;
run; quit;
```

BOX PLOT IN LOG₁₀ SCALE

This SAS code creates a box plot in log₁₀ scale using the same annotations created for the box plot in linear scale.

```
filename plot "&pathplot.box plots &&MED&i &&CHEM&i &&UNITSFN&i log10.cgm";
```

```

proc gplot data=merged annotate=anno;
  plot results*pdetc / vaxis=axis1 haxis=axis2 vm=9 hm=0 name='plot4';
  format results &&FORMAT_LOGRES&i.;
  axis1 width=5 value=(h=0.75) label=(a=90 r=0 h=0.75 j=c "&&CHEM&i. (&&UNITS&i.) in
    Log10 scale") logbase=10 logstyle=expand;
  axis2 width=5 value=(h=0.75) label=(h=0.75 'Frequency of Detection');
  footnote1 j=l h=0.65 ' '; footnote2 j=l h=0.2 ' ';
  label results='00'x;
  symbol1 h=0.75 bwidth=11 i=box00tf l=1 w=5 c=blue cv=yellow; run; quit;

```

PLACE GRAPHS TOGETHER USING PROC GREPLAY

This SAS code places the four individual plots together onto a single graph using PROC GREPLAY.

```

goptions reset=all display gsfname=replace noprompt gsfname=plot device=CGM0FMP
  rotate=portrait htext=0.8 ftext="TimesRoman" horigin=0 in vorigin=0 in vsize=8.5
  in hsize=6.5 in cback=white;

filename plot "&pathplot.4 plots &&MED&i &&CHEM&i .cgm";

```

```

proc greplay tc=tempcat nofs igout=work.gseg;
  tdef ucf des='UCF'
  1/llx = 0 lly = 51
    ulx = 0 uly = 100
    urx = 49 ury = 100
    lrx = 49 lry = 51
  2/llx = 51 lly = 51
    ulx = 51 uly = 100
    urx =100 ury = 100
    lrx =100 lry = 51
  3/llx = 0 lly = 0
    ulx = 0 uly = 49
    urx = 49 ury = 49
    lrx = 49 lry = 0
  4/llx = 51 lly = 0
    ulx = 51 uly = 49
    urx =100 ury = 49
    lrx =100 lry = 0
;

template = ucf;
treplay 1:plot1 2:plot3 3:plot2 4:plot4; run; quit;
%end; ** end do i;

```

SEND PLOTS TO MICROSOFT WORD

This SAS code sends the composite graph to Microsoft Word with a unique figure number and appropriate title.

```

data _null_;
  file sas2word;
  %if &NEWFILE = Y %then %do; put '[FileNew ]'; %end;
  %else %do; put '[FileOpen .Name = "&pathplot.&WORDFN"@; put '"']'; %end;
  put '[FilePageSetup.LeftMargin = 72]'; ** 1 inch margins;
  put '[FilePageSetup.RightMargin = 72]'; ** 1 inch margins;
  %local i;
  %do i=&startno %to &endno;
    put '[EditBookmark.Name= "&pathplot.&MED&i.&CHEM&i."@; Put '" ,.Add]';
    put '[InsertPicture.Name = "&pathplot.4 plots &&MED&i. &&CHEM&i. .cgm"@;
      put '"']';
    put '[InsertPara]';
    put '[FormatFont .Points = "11", .Font = "Time New Roman"]';
    put '[FormatParagraph.Alignment=1]'; ** center caption;
    put '[Insert '"@; put '"']';
    put '[Insert "Fig. '@; put '"&I"@; put '. Background probability and box plots of

```

```

        '@; put "&&CHEM&i. in &&MED&i."@; put '["]';
        %if &i. < &endno %then %do;    put '[InsertPageBreak]'; %end;
        %end;
        put '[FileSaveAs .Name = "'@; put "&pathplot.&WORDFN"@; put '["]'; run;
    %mend plots;    ** end plots macro;

    %let NEWFILE = Y; ** Y if a new file will be created, N otherwise;

```

Call the macro PLOTS for any number of plots.

```
%plots(1, &N_PLOTS)
```

GRAPHS DISPLAYED IN WORD

The final graphs as shown in Word are displayed in Figures 1 and 2 for the chemicals lead and dimethylbenzene, respectively, from 21 surface soil samples. The vertical scales are identical, as are the lengths of both axes for all four individual plots.

COMPARISON TO ODS

The use of DDE for generating these plots was compared to the Output Delivery System (ODS) in SAS. Currently, there are only four options for creating graphics output in rich text format (RTF) for Word: PNG, JPEG, SASEMF, and ActiveX. The image quality of the PNG, JPEG, and SASEMF graphs was poor compared to the CGM file created for Word using DDE. The image quality of the ActiveX graph was better, but Haworth (2001) states some limitations of ActiveX. ODS works only with the GCHART, GCONTOUR, GMAP, GPLOT, and G3D procedures, and it requires that SAS/GRAPH ActiveX control be installed. Since ActiveX is a Microsoft technology, it may only work with Microsoft Office products. The technology has proven to be a security risk since it can be used by hackers to insert viruses into the system. As a result, many sites choose to disable ActiveX.

CONCLUSION

Dynamic data exchange can be used to open Microsoft Word from within SAS and quickly send many plots to Word with the appropriate figure numbers and titles using SAS macros. Environmental data were presented using both probability and box plots in linear and $\log_{10}$ scales to examine the data distributions. Annotation is used on the graphs to identify possible outliers, label the distribution, create hollow symbols for non-detects, and show the mean, UCL, and possible outliers on box plots. A linear regression line is added to plots with a vertical axis on a $\log_{10}$ scale. The number of major and minor tick marks on the vertical axis of each plot is optimized for the range of data for each chemical. SAS code is presented to produce customized plots for a large number of chemicals directly into Word. DDE allows you more control over the graph in Word compared to ODS. Furthermore, the CGM driver produces higher quality resolution images over PNG, JPEG, and SASEMF drivers for graphs in Word.

REFERENCES

Haworth, Lauren E. 2001. *Output Delivery System: The Basics*. Cary, NC: SAS Institute Inc., p. 258.

CONTACT INFORMATION

The author welcomes and encourages any questions, corrections, feedback, and remarks. Contact the author at:

Dennis J. Beal, Ph.D. candidate
 Senior Statistician/Risk Scientist
 Science Applications International Corporation (SAIC)
 P.O. Box 2501
 151 Lafayette Drive
 Oak Ridge, Tennessee 37831

phone: 865-481-8736
 fax: 865-481-4757
 e-mail: beald@saic.com

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.

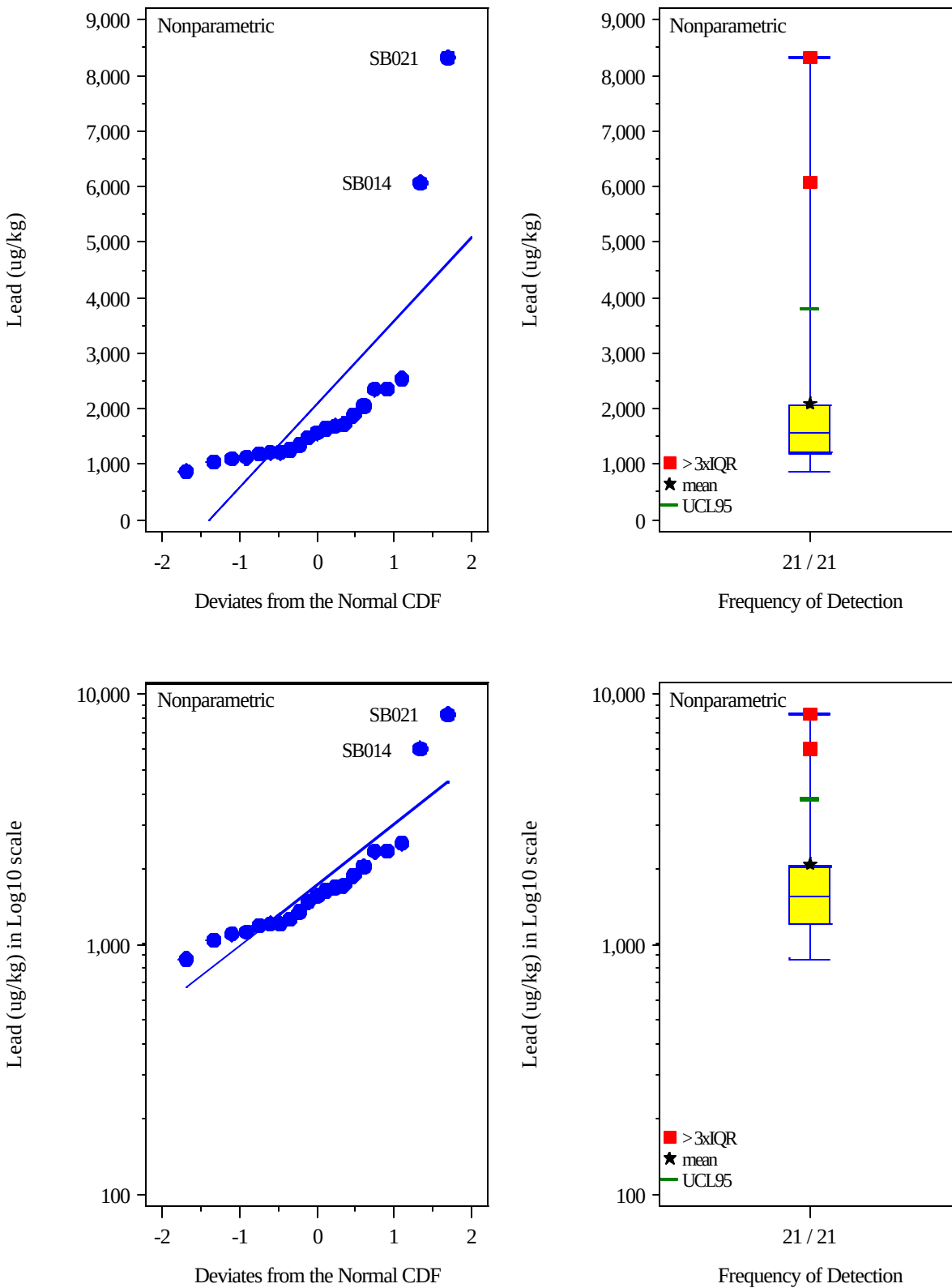
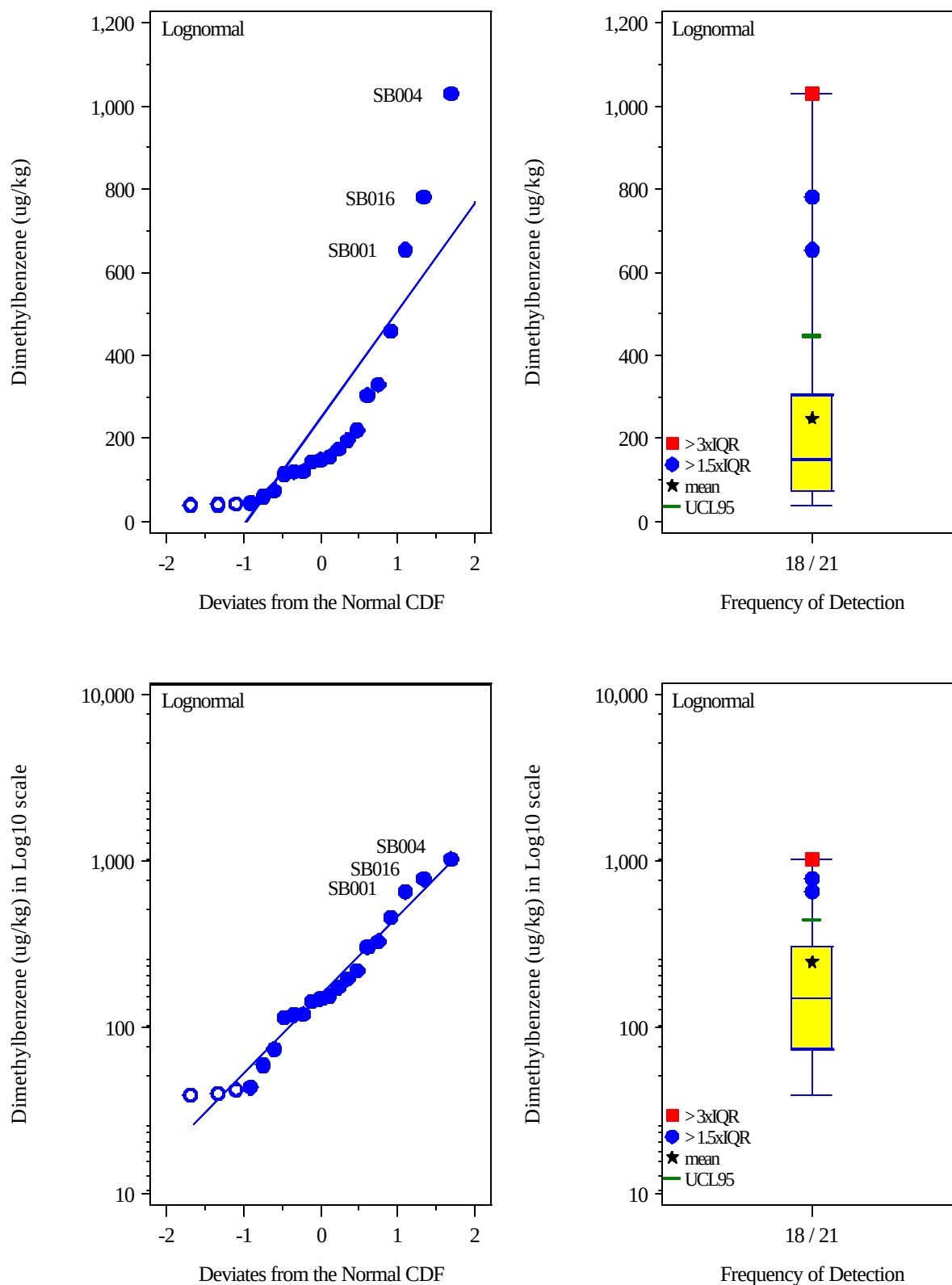


Figure 1. Background probability and box plots of lead in surface soil



Hollow symbols denote non-detects.

Figure 2. Background probability and box plots of dimethylbenzene in surface soil