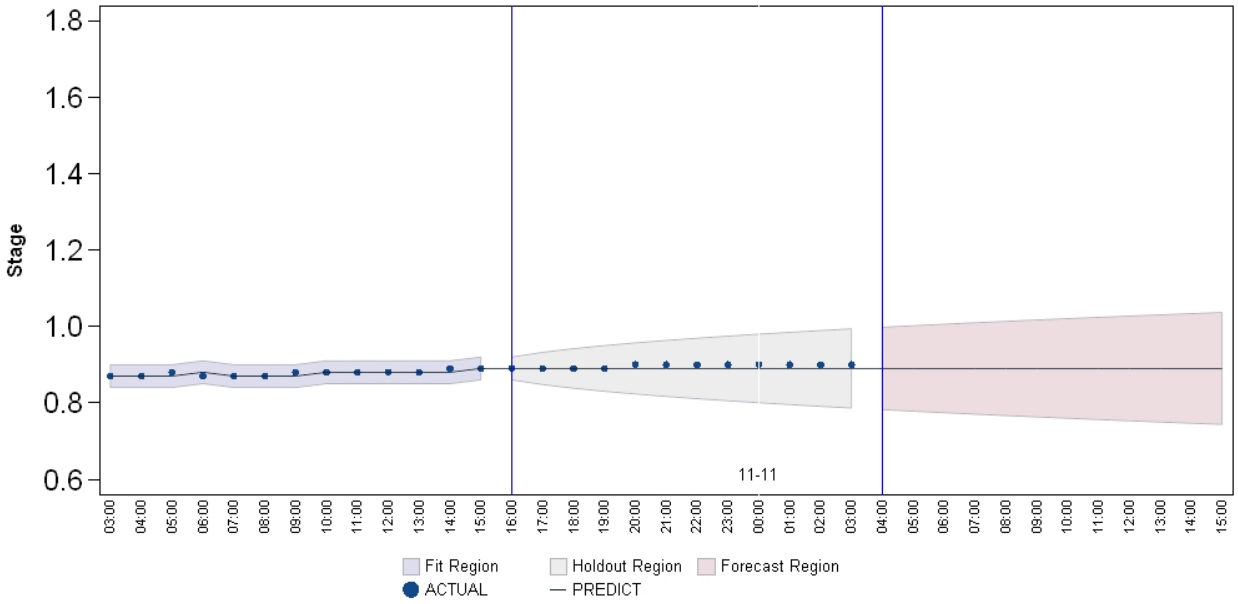
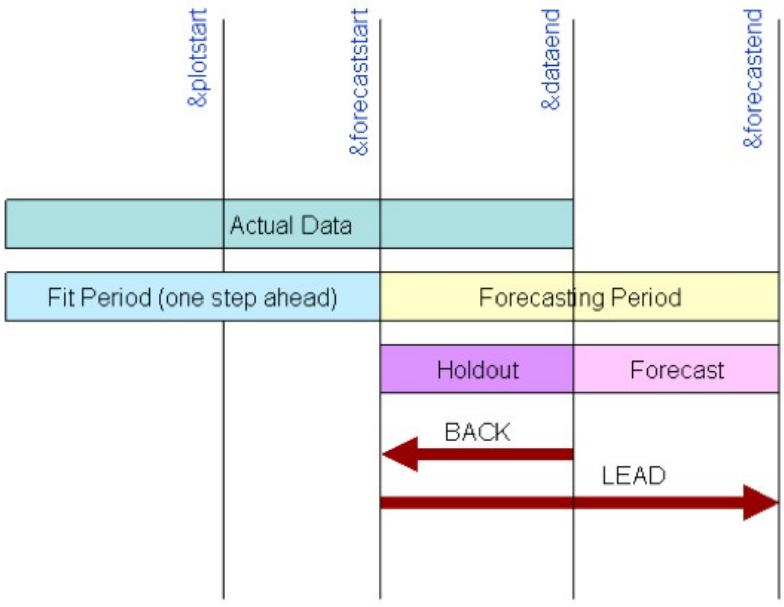
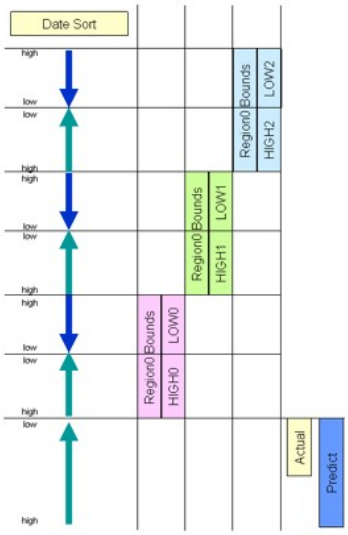
Handout For:
Effective Forecast Visualization With SAS SAS®
Detailed Instructions and Documentation

Samuel T. Croker, Independent Consultant

CONFIDENCE INTERVALS AS REGIONS







SAS GRAPH MACRO

```
/*****************************************************************
*   Program Name     :   $RCSfile: threeregionforecast.sas,v $
*   REV/REV AUTH     :   $Revision: 1.2 $ $Author: scoyote $
*   REV DATE         :   $Date: 2007/03/05 21:27:31 $
*****************************************************************/
%macro threeregionforecast(
```

```
/* positional parameters (required) */
 ds                   /* dataset containing forecasts */
,dataend              /* ending date of the time series - not the forecast series */
,datastart            /* starting date of the time series */
,plotback             /* how many date values to go back from the end of the series in the plot */
,predictlag           /* how many date values to predict ahead */
/* optional parameters */
,order=                /* date axis ordering (should be order=(<<data>>) )*/
,lciname=L95           /* variable name of the lower confidence bound */
,uciname=U95           /* variable name of the upper confidence bound */
,rsname=residual       /* variable name of the residual values */
,fcname=forecast       /* variable name of the forecast values */
,varname=actual        /* variable name of the actual values */
,varlab=               /* label for the actual values */
,fclab=                /* label for the predict values */
,dtm=date              /* variable name of the date axis variable */
,dtdisplay=datetime28. /* default format for the datetime axis */
,dtformat=mmddyy8.     /* format for time axis values */
,gname=FCST            /* name for the SAS graph output */
,gcat=work.GSEG        /* ouptut catalog */
,gdesc=Forecast Plot   /* description and title of the graph */
,fontname=SWISS        /* fonts */
,htitle=1              /* title height */
,cback=white           /* background color */
,grtitle=              /* main title of the graph */
,vaxisvalh=1           /* vertical axis value height */
,haxisvalh=1           /* horizontal axis value height */
,xinterval=hour.       /* order interval for the x axis */
,xminorticks=2
,ymajnum=10            /* number of y tick marks */
,hatitle=              /* horizontal axis title */
,vatitle=              /* vertical axis title */
,cicol1=bwh            /* first confidence region color */
,cicol2=gwh            /* second confidence region color */
,cicol3=pkwh           /* third confidence region color */
/* the following apply to actual symbol values */
,actcol=vigb           /* color */
,acth=1                /* height */
,actw=1                /* width */
,actv=dot              /* value type */
,actl=1                /* line type */
,acti=none             /* interpolation type */
/* the following apply to forecast symbol values */
,fcstcol=degb          /* color */
,fcsth=1               /* height */
,fcstw=1               /* width */
,fcstl=1               /* line type */
,fcstv=none            /* value type */
,fcsti=j               /* interpolation type */
);
data _null_;
    format forecaststart forecastend 20.;
    forecaststart=intnx('dthour',&dataend,-&plotback);
    forecastend=intnx('dthour',forecaststart,&predictlag);
    plotstart=intnx('dthour',forecaststart,-&plotback);
    call symput ('forecaststart',forecaststart);
    call symput ('forecastend',forecastend);
    call symput ('plotstart',plotstart);
run;
/* rebuild the output data so that the cis plot as polygons */
data out(  drop=     sval0 sval1 sval2)
        low0( keep=&dtm sval0 sval1 sval2)
        high0(keep=&dtm sval0 sval1 sval2)
        low1( keep=&dtm sval0 sval1 sval2)
        high1(keep=&dtm sval0 sval1 sval2)
        low2( keep=&dtm sval0 sval1 sval2)
        high2(keep=&dtm sval0 sval1 sval2);
    set &ds;
    where &dtm>=&plotstart;
    output out;
    if &dtm <= &forecaststart then do;
        sval0=&lciname; output low0;
        sval0=&uciname; output high0;
    end;
    if &dtm > &forecaststart and &dtm <= &dataend then do;
        sval1=&lciname; output low1;
        sval1=&uciname; output high1;
    end;
    if &dtm > &dataend then do;
        sval2=&lciname; output low2;
        sval2=&uciname; output high2;
    end;
run;
/* sort the lower bound datasets so that the polygons will be drawn correctly */
```

```
    proc sort data=low0; by descending &dtm; run;
    proc sort data=low1; by descending &dtm; run;
    proc sort data=low2; by descending &dtm; run;

    /* stack the low and high datasets in this way so that the graphs will be drawn correctly */
    data forecast;
        set
            low2 high2
            low1 high1
            low0 high0
            out;
        if &dtm=. then delete;
    run;

    /* generate vertical lines to denote the date, and highlight the start of the different regions */
    data DayLines; set forecast(keep=&dtm );
        length color function $8 text $25;
        retain xsys '2' ysys '1' when 'a';
        if hour(&dtm)=0 and minute(&dtm)=0 and &dtm>=intnx('dthour',&plotstart,-1) then do;
            wdate=compress(day(datepart(&dtm))||'-'||month(datepart(&dtm)));
            function='move'; x=&dtm; y=0;
                output;
            function='draw'; x=&dtm;
                y=100; color='white'; size=1; output;
            function='label';x=&dtm;
                y=3; size=1; position='2';
                angle=90;color='black'; text=wdate; output;
        end;
        if &dtm=intnx('dthour',&forecaststart,1)
            or &dtm=intnx('dthour',&dataend,1) then do;
            function='move';x=&dtm; y=0; output;
            function='draw';x=&dtm; y=100; color='blue'; size=1; output;
        end;
    run;


/* draw the graph */
    goptions reset=all
            device=activex
            xpixels=1000
            ypixels=500
            ftext="&fontname"
            htitle=&htitle
            cback=&cback  ;
    title &grtitle;
    symbol1 i=ms                                    c=&cicol1  co=libgr;
    symbol2 i=ms                                    c=&cicol2  co=libgr;
    symbol3 i=ms                                    c=&cicol3  co=libgr;
    symbol4 i=&acti  v=&actv    l=&actl  h=&acth    w=&actw    c=&actcol;
    symbol5 i=&fcsti v=&fcstv   l=&fcstl h=&fcsth   w=&fcstw   c=&fcstcol;
    legend1 across=10;
    title &grtitle;
    axis1 label=(&hatitle )
        value=(f="&fontname" h=&haxisvalh angle=90 rotate=0)
        major=none
        minor=none
        order=(&plotstart to &forecastend by &xinterval);
    axis2 label=(&vatitle angle=90 rotate=0)
        value=(h=&vaxisvalh)
        minor=none;
    proc gplot data=forecast gout=work.gseg annotate=daylines;
        label sval0='Fit Region';
        label sval1='Holdout Region';
        label sval2='Forecast Region';
        label &varname=&varlab;
        label &fcname=&fclab;
        plot  sval0*&dtm=1
        sval1*&dtm=2
        sval2*&dtm=3
        &varname*&dtm=4
        &fcname*&dtm=5
            /   name="&gname" des="&gdesc "
                grid
                haxis=axis1
                vaxis=axis2
                legend=legend1
                overlay
                chref=palg;
        format &dtm &dtdisplay;
    run; quit;
%mend threeregionforecast;
```

ODS GRAPHICS

With ODS GRAPHICS, the plotting of confidence intervals as regions is much easier. There is documentation for this type of work at http://support.sas.com/rnd/base/topics/statgraph/proctemplate/a002774500.htm.
ODS Graphics can be accessed through many procedures, but can also be used (experimentally in 9.1.3) via the data step. First a template must be declared. There are many templates associated with procedures also, and these can be copied. In the following example a template was created from scratch.

```
ODS PATH work.templat(update) sashelp.templat(read) sashelp.tmplmst(read);
proc template;
    define statgraph mygraphs.stcfor;
    dynamic graphtit;
    layout lattice /width=900 height=200 border=false;
        layout overlay /border=false
            xaxisopts=(display=(values TICKS) )
            yaxisopts=(display=all label="Stage" )
;
        entrytitle graphtit/
                fontsize=12
                fontweight=bold
                halign=left
                padtop=0
                padbottom=0
                valign=top;

            Band
                ylimitlower=fit_lower
                ylimitupper=fit_upper
                x=datestamp /
                    fill=true
                    lines=false
                    fillcolor=ywh
                    legendlabel="Fit CI"
                    name="Conf1";
            Band
                ylimitlower=holdout_lower
                ylimitupper=holdout_upper
                x=datestamp /
                    fill=true
                    lines=false
                    fillcolor=bwh
                    legendlabel="Holdout CI"
                    name="Conf2";
            Band
                ylimitlower=fcst_lower
                ylimitupper=fcst_upper
                x=datestamp /
                    fill=true
                    lines=false
                    fillcolor=pkwh
                    legendlabel="Forecast CI"
                    name="Conf3";
            scatter X=datestamp Y=actual  /name="act"  legendlabel="Actual Stage"  markers=true
             markersymbol=circlefilled markercolor=black ;
            SERIES X=datestamp Y=predict /name="pred" legendlabel="Predicted Stage" markers=false  linecolor=blue;
        endlayout;
    endlayout;
    end;
run;
```

A simple data step is used to plot the graph from ODS GRAPHICS. The SQL statement shows how the limits for each forecast region are retrieved from the data. The dataset MARYLAND contains the pre-forecasted data and OUTFOR contains the forecasts as generated by HPFENGINE.

```
goptions reset=all;
ods html gpath='C:\Documents and Settings\scoyote\Desktop\Output';
ods graphics on /reset imagename="forecastplot" imagefmt=jpeg border=off;

    proc sql noprint;
        select
               max(datestamp) format=30.
             , min(datestamp) format=30.
            into
                :maxdt
                ,:mindt
            from maryland
            where site_no="03078000";
    quit;
    data _null_;
        merge outfor(
                    where=(site_no='03078000')
                    rename=(lower=low upper=up)
```

```
                in=outfor)
            sitenames;
        by site_no;
        call symput('sitename',station_nm);
        if outfor;
        if datestamp<=intnx('hour',&maxdt,-&back) then do;
                fit_lower=low; fit_upper=up;
        end;
        else if datestamp<=&maxdt then do;
                holdout_lower=low; holdout_upper=up;
        end;
        else if datestamp>&maxdt then do;
                fcst_lower=low; fcst_upper=up;
        end;
        file print ods=( template='mygraphs.stcfor'
            objectlabel='Forecast Plot' dynamic=(graphtit=station_nm) );
        put _ods_ ;
    run;
ods html close;
ods graphics off;
```

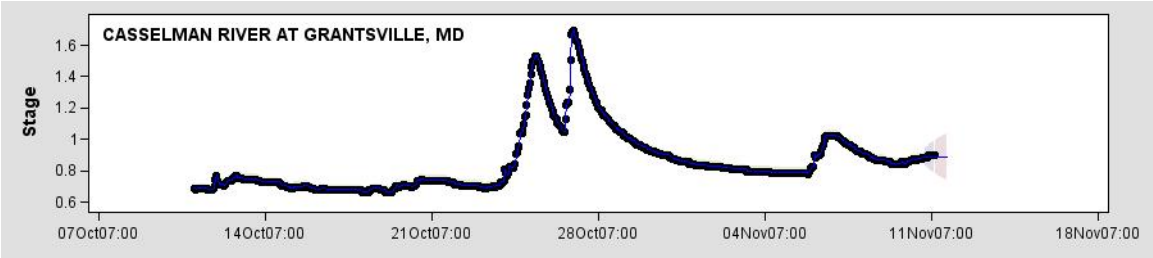


Figure 1: Cassleman River Near Grantsville, MD October 7th–November 11th, 2007

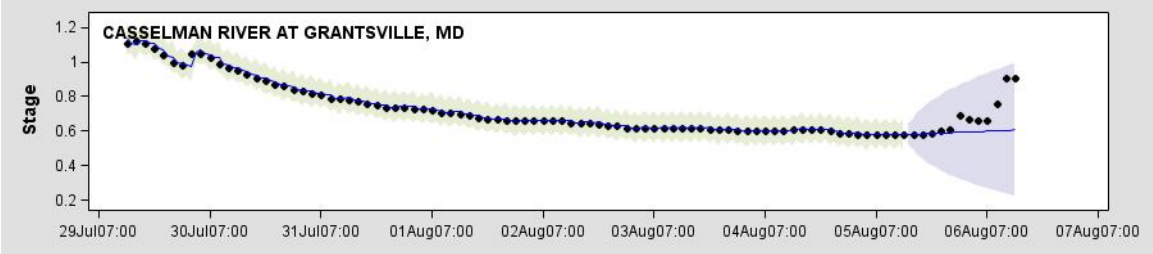


Figure 2: Cassleman River Near Grantsville, MD August 5th–6th, 2007

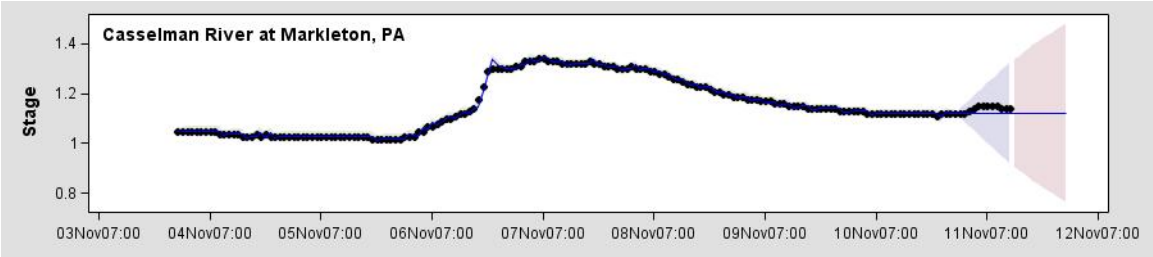
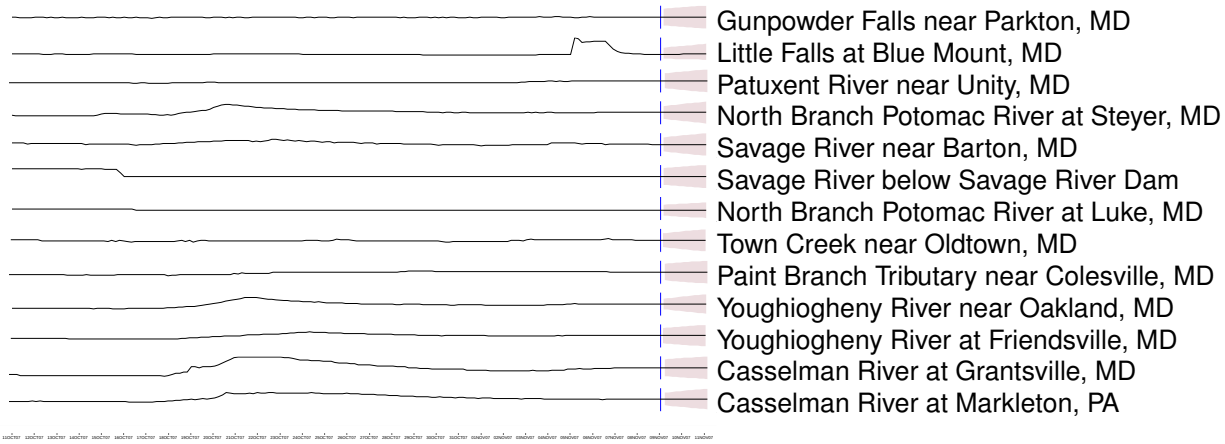


Figure 3: Cassleman River Near Markleton, PA October 7th–November 11th, 2007

SPARKLINES

SMALL STACKED GRAPHICS

The following graphs are easily comparable without too much detail information that could be included in many different ways, such as a drill down.



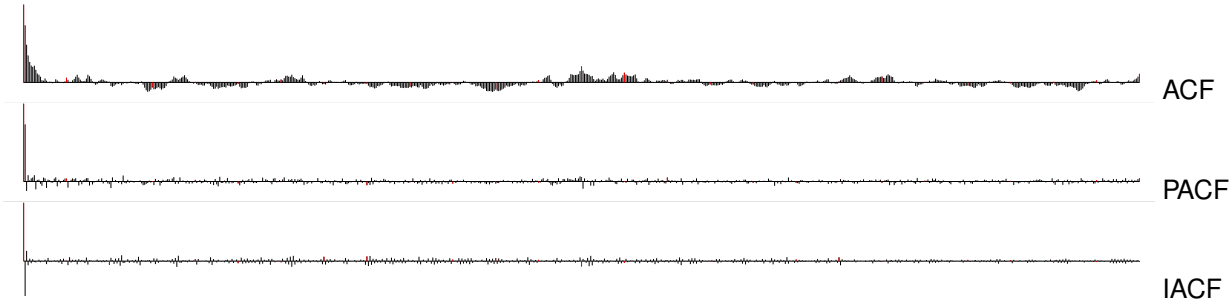
These graphics can also be used in the flow of the text if formatted correctly. The SASHELP.AIR series is shown  flowing with the text. The code to generate this sparkline is shown below, and is an example of how to use inline sizing options in the GOPTIONS statement:

```
filename outgraph "path.file";
goptions
     reset=all
     device=pdfc
     xmax=144pt  horigin=0.000pt  hsize=144pt xpixels=5000
     ymax=12pt   vorigin=0.000pt  vsize=12pt  ypixels=416
     cback=white
     noborder
     gsfname=outgraph
     gsfmode=replace;
     symbol1 v=none i=j c=black width=50;
     axis1 label=none value=none major=none minor=none
 offset=(0,0)
 style=0;
     axis2 label=none value=none major=none minor=none
 style=0;
proc sql noprint;
     select min(air)
          , max(air)
           into
                :minair
               , :maxair
     from sashelp.air;
quit;
data air;
     set sashelp.air;
     if air=&maxair then maxair=air;
     if air=&minair then minair=air;
run;
symbol2 v=dot h=40 c=red i=none;
symbol3 v=dot h=40 c=green i=none;

proc gplot data=air;
     plot (air maxair minair)*date / overlay   haxis=axis1 vaxis=axis2;run;
quit;
```

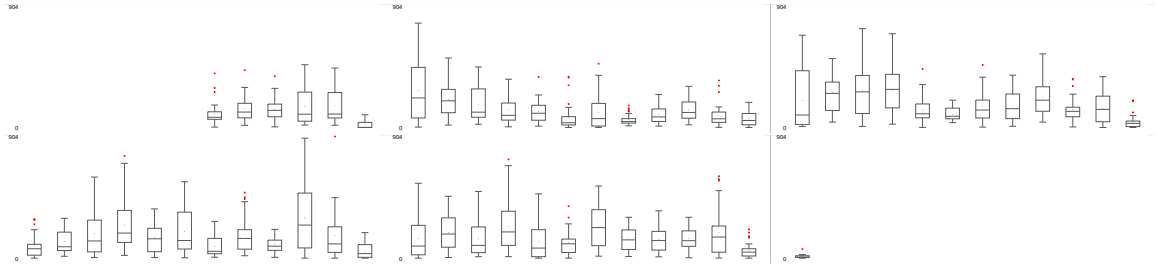
EXAMPLE - TIME SERIES DIAGNOSTICS

LONG PERIOD CORRELATIONS - 768 LAGS

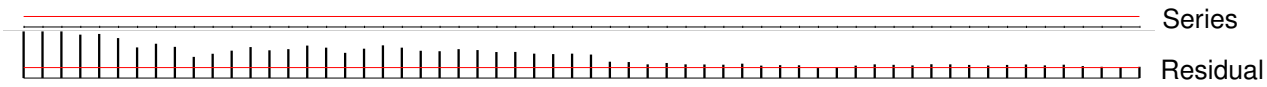


SHORT PERIOD CORRELATIONS

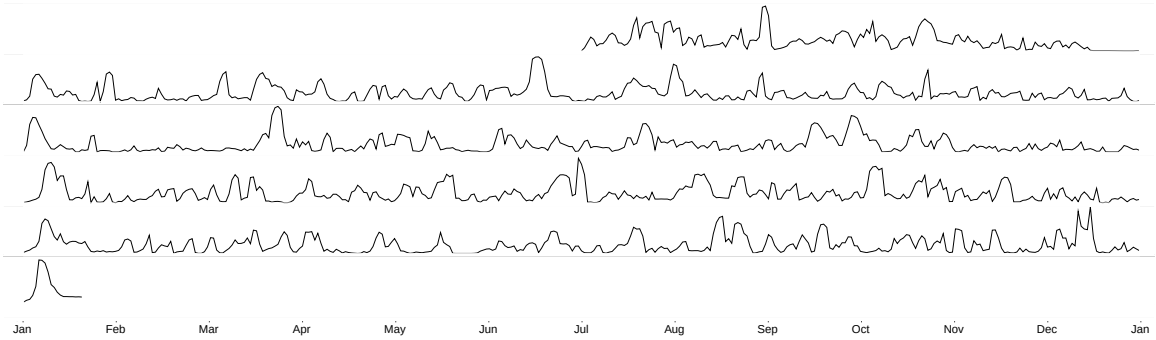
COMPARATIVE BOXPLOTS



χ^2 PROBABILITIES



TIME SERIES



REFERENCES

Tufte, Edward R. 2006. *Beautiful Evidence*, Cheshire, CT: Graphics Press LLC.

Gossens, Michael, Frank Mittlebach, Alexander Samarin *The LaTeXCompanion*, 1st Edition : Addison-Wesley Professional

Box, George E.P., Gwilym M. Jenkins and Gregory C. Reinsel. 1994. *Time Series Analysis: Forecasting and Control*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall.

Brocklebank, John and David A. Dickey. 2003. *SAS® for Forecasting Time Series*, 2nd ed. Cary, NC: SAS Institute Inc.

Gelso, Charlie, Larry Coburn. 2006. *Guide to Maryland Trout Fishing: The Catch and Release Streams* Carter, OK: Falling Star Publishing

Cartier, Jeff. "The Power of the Graphics Template Language." *Proceedings of the 30th Annual SAS® Users Group International Conference*. April 2004.
<http://support.sas.com/rnd/datavisualization/papers/sugi30/GTL.pdf>
(Accessed July 18, 2007).

Croker, Samuel T. "Effective Forecast Visualization with SAS/GRAPH." *SAS Global Forum 2007 Proceedings*. April 2007.
<http:
//www8.sas.com/scholars/Proceedings/2006/DataPresentation/DP01_06.PDF>

SAS Institute Inc. *Sample 1151: Filling the area between plot lines using SYMBOL statement*
<http://support.sas.com/ctx/samples/index.jsp?sid=115>

SAS Institute Inc.*ARIMA: Models for Series J from Box-Jenkins*
<http://ftp.sas.com/techsup/download/sample/samp_lib/etssampArima_Models_for_Series_J_from_B.html>

Shumway, Robert H. and David S. Stoffer. 2006. *Time Series Analysis and Its Applications with R Examples*, 2nd ed. New York: Springer Science+Business Media, LLC.

CONTACT INFORMATION

We value and encourage your comments and questions! You can find the latest version of the SAS code for this paper at: `http://www.scoyote.net/forecasting/`. Please note that we may update this code for use in other papers.

You can contact the authors at:

Name: Samuel T. Croker
E-Mail: `scoyote at scoyote.net`
Web: `http://www.scoyote.net/forecasting/`