

Customizing ODS Graphics to Publish Visualizations

Paul A. Montagna, Harte Research Institute, Texas A&M University-Corpus Christi,
Corpus Christi, TX

ABSTRACT

The procedures provided in the ODS graphics package, such as SG PANEL, SG PLOT, and SG SCATTER make creating outstanding visualizations much easier than the original GPLOT procedure. They also provide great flexibility to create complex graphics with ease. However, it is a common desire to use something other than the default settings, sometimes we desire to tweak a plot, and often we need versions that are of publishable quality (i.e., a specific resolution, size, or restricted colors). All these requirements call for customization. This presentation describes several ways to customize ODS graphics plots including: using ODS options, altering styles using the %MODSTYLE macro, using PROC FORMAT, and some little know OPTION statements with the graphics procedures.

KEYWORDS

ARRAY statement, %MODSTYLE, ODS statement, PROC FORMAT, PPROC SG PLOT, PROC SG SCATTER

INTRODUCTION

A common task is the need to create publication quality graphics. The problem is that publishers have requirements that go beyond that provided by default settings. Many solutions are provided by SAS graphics software, such as the GPLOT procedure, but the ODS graphics procedures (SAS 2013) are especially well suited to solve the problem, easy to use, and easy to customize. While much visualization work is designed for web display, or other digital display, there is still a large print publishing industry, and the desire to create print publications, especially for technical applications.

Publishers often have an extensive list of requirements for including graphics in print publications, and do not allow for deviation from the requirements. Resolution is nearly always required to be a minimum of 300 dpi (dots per inch). Size of the graphic is optional, but it will depend on the finished published page sizes or column sizes. Fonts must be scaled to the graph size to ensure readability of text within the graphic. File formats are typically optional, but acceptable options include eps, tiff, jpg, png, or pdf. To use color, or not to use color, is always an important question. Color is tricky because it can enhance a graphic, but it can also be confusing or misdirect the reader from the main point of a graphic. Printing in monochrome is usually free, but there are usually charges for color production. Often it is useful to have monochrome and color versions of graphics because color is commonly free in the online and pdf versions of a publication. Therefore, I always start by creating a monochrome graphic, and then add color for emphasis or to draw the reader's attention to a specific point.

The solution to the problem presented here is to use the ODS Statistical Graphics (SG) procedures, such as SG PANEL, SG PLOT, and SG SCATTER (SAS 2016). Each of the ODS graphics procedures have common options, which can be used to change from the default settings. ODS statements can be used for setting specific output options. In addition, there is the %MODSTYLE macro, which can be used to modify the ODS Template directly. The first step is to know the default ODS settings, so that necessary changes are identified. The ODS default settings are listed in the text box to the right.

ODS GRAPHICS DEFAULTS

- Resolution = 96 dpi
- Graphic Image Format = jpg
- Frames/Borders = on
- Destination = HTML
- Style = HTMLBlue
- Markers = open circle
- Lines = solid
- Aspect Ratio = 480 X 640 pixels

DATA

The example data set is from a study to determine the role hydrology (i.e., river flow) plays in regulating the biogeochemistry (e.g., nutrient cycling) in Texas estuaries (Montagna et al. 2018). A climatic gradient exists along the Texas coast that provides a natural experiment for examining inflow effects because average salinity ranges from 16 to 35 practical salinity units (psu) in the four estuaries studied. There are two bays within estuaries, the primary bay connects to the Gulf of Mexico, and the secondary bay connects to the river. Thus, average salinities are lower in secondary bays than in primary bays. Measurements of dissolved inorganic nutrients, chlorophyll, dissolved organic matter, and carbonate chemistry were measured for three years to capture the range of year-to-year variability. Therefore, the experimental design contains two main effects: among estuaries (which tests for long-term climatic differences), and bays within estuaries (which tests for differences from rivers) (Table 1).

Table 1. Experimental Design. Bays within estuaries.

		Bay Distance from River Source	
		Near	Far
Estuary	Lavaca-Colorado	Lavaca	Matagorda
	Guadalupe	Upper Guadalupe	Lower Guadalupe
	Nueces	Nueces	Corpus Christi
	Baffin-Laguna Madre	Baffin	Laguna

The analysis approach is to create mixing curves, which are plots salinity versus a chemical parameter. Nutrients enter estuaries via river flow, so nutrient concentrations are highest in low salinity areas. A one to one relationship means the chemical is diluted conservatively as fresh water mixes with sea water. The concentration of a nutrient decreases rapidly as salinity increases when chemicals are consumed or transformed within estuaries along the gradient from the river to the sea. The goal of the project is to create graphs with the ranges represented, so the first step for analysis is to create a dataset of the means and standard errors for of the chemicals for each bay (Table 2).

Table 2. Data set (TXWQ_summary) of mean and standard error of chemical measurements.

Obs	Bay	_TYPE_	_FREQ_	Temp	Sal	DO	pH	PO4	SiO4	NH4	NOx	Chl	sTemp	sSal	sDO	spH	sPO4	sSiO4	sNH4	sNOx	sChl
1	Corpus Christi	0	39	23.36	33.64	7.38	8.11	0.500	33.138	0.662	0.7399	6.36	1.03	0.58	0.18	0.02	0.203	2.591	0.076	0.10257	0.55
2	Lavaca	0	26	23.84	19.16	8.26	8.23	1.453	70.204	1.087	3.2345	8.65	1.41	2.02	0.24	0.04	0.279	8.852	0.233	1.37586	1.34
3	Lower Baffin	0	13	23.38	40.13	6.06	8.21	0.248	68.170	3.599	0.4130	13.92	1.67	1.84	0.40	0.03	0.078	8.033	0.881	0.14508	2.36
4	Lower SA	0	26	23.26	22.17	8.08	8.12	1.500	82.166	1.114	1.4559	11.22	1.46	1.93	0.33	0.06	0.474	12.598	0.318	0.33379	1.47
5	Matagorda	0	52	23.09	24.39	8.26	8.18	0.962	35.538	1.077	4.9014	9.15	0.98	1.01	0.42	0.03	0.148	3.450	0.196	1.75468	0.89
6	Nueces	0	26	23.90	31.55	7.54	8.10	1.681	80.380	1.545	1.7429	9.25	1.27	1.24	0.18	0.03	0.325	8.600	0.492	0.52369	1.06
7	Upper Baffin	0	26	23.34	41.72	5.95	8.20	0.239	63.709	3.493	0.6965	18.43	1.05	2.39	0.27	0.03	0.026	6.870	0.636	0.22165	2.29
8	Upper SA	0	26	23.54	15.34	10.26	8.49	2.609	122.855	1.188	15.8234	27.36	1.41	1.84	0.82	0.05	0.715	15.367	0.294	7.12933	3.23

The second step is to create a dataset with the upper and lower bounds for each chemical parameter. This is easy to do using the %LET and the ARRAY statements. These statements create the dataset that will be graphed. For simplicity, only the salinity (Sal) and phosphate (PO4) variables are displayed (Table 3)

```

%let chem=Temp Sal DO pH PO4 SiO4 NH4 NOx Chl;
%let Schem=sTemp sSal sDO spH sPO4 sSiO4 sNH4 sNOx sChl;
%let lchem=lTemp lSal lDO lpH lPO4 lSiO4 lNH4 lNOx lChl;
%let uchem=uTemp uSal uDO upH uPO4 uSiO4 uNH4 uNOx uChl;
data TXWQ_g;
  set TXWQ_summary;
  array C {9} &chem;
  array S {9} &schem;
  array L {9} &lchem;
  array U {9} &uchem;
  do count=1 to 9;
    L(count) = C(count) - S(count);
    U(count) = C(count) + S(count);
  end;
run;

```

Table 3. Dataset to be graphed with mean, standard error, lower bounds, and upper bounds for salinity and phosphate.

Obs	Bay	Sal	sSal	lSal	uSal	PO4	sPO4	lPO4	uPO4
1	Corpus Christi	33.64	0.58	33.0577	34.2274	0.500	0.203	0.29660	0.70254
2	Lavaca	19.16	2.02	17.1477	21.1792	1.453	0.279	1.17408	1.73233
3	Lower Baffin	40.13	1.84	38.2901	41.9715	0.248	0.078	0.17056	0.32586
4	Lower SA	22.17	1.93	20.2442	24.0950	1.500	0.474	1.02620	1.97418
5	Matagorda	24.39	1.01	23.3817	25.4010	0.962	0.148	0.81403	1.11027
6	Nueces	31.55	1.24	30.3093	32.7961	1.681	0.325	1.35671	2.00586
7	Upper Baffin	41.72	2.39	39.3291	44.1173	0.239	0.026	0.21273	0.26479
8	Upper SA	15.34	1.84	13.5008	17.1877	2.609	0.715	1.89413	3.32407

PROC SGPLOT is used in the graphing examples below, but most of the statements are available in PROC SGPANEL and PROC SGSCATTER. The default settings creates the graph in Figure 1.

```
proc sgplot data=TXWQ_g;
scatter x=Sal y=PO4 / GROUP=Bay
        YERRORLOWER=lPO4 YERRORUPPER=uPO4
        XERRORLOWER=lSal XERRORUPPER=uSal;
xaxis values=(10 to 50 by 10) minor ;
run;quit;
```

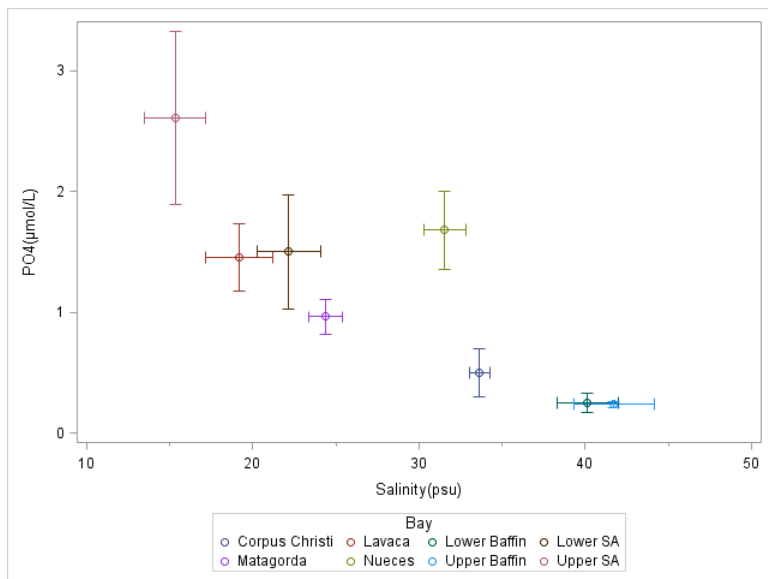


Figure 1. Default settings.

Notice the default settings creates a graph with different colors for different bays, and the symbols are all the same (Fig. 1). Also, notice the graph has a frame. The axis labels for the variables SAL and PO4 are already in the dataset.

I want to start using monochrome and other specific formats. This is accomplished with the ODS HTML and the ODS GRAPH statements. The ODS HTML statement is used to change the style, folder destination, and file name. The ODS GRAPH statement is used to change the image size, image resolution, output format, image name, and features (e.g., borders, data tips, scaling, anti-aliasing).

The HTML statement below create a monochrome graph using the JOURNAL style option, a resolution of 300 DPI, and designate folder location for the graph file. He GRAPHICS statement turns of the border, sets a specific size, change the output format to PNG, and names the image. The result is Figure 2.

```
ods html style=journal IMAGE_DPI=300 gpath="C:\Users\pmontagna\Documents\MySAS\Images";
ods graphics on / border=off height=3in width=4in outputfmt=png imagename="Salinity_v_PO4_bw";
proc sgplot data=TXWQ_g;
  scatter x=Sal y=PO4 / group=bay
          YERRORLOWER=lPO4 YERRORUPPER=uPO4
          XERRORLOWER=lSal XERRORUPPER=uSal;
  xaxis values=(10 to 50 by 10) minor ;
run;quit;
```

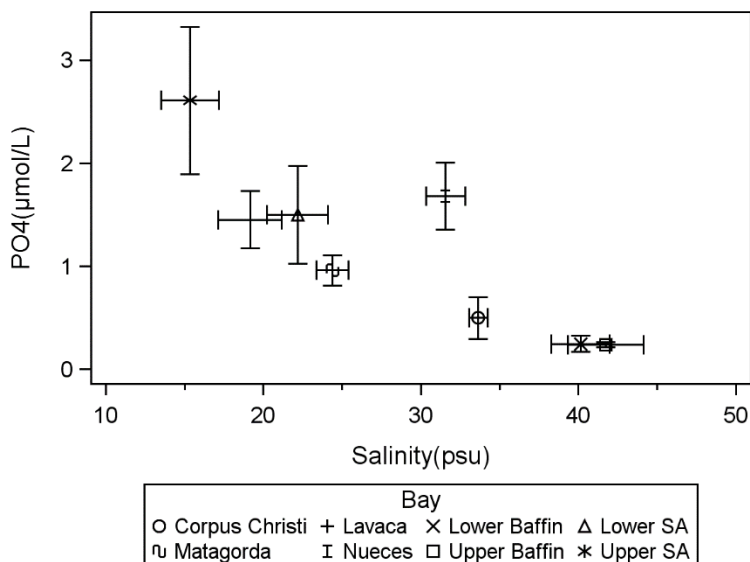


Figure 2. Graph using JOURNAL style, 300 dpi no border, and 4 by 3 inch graph size.

Notice that the fonts are now scaled to the final image size of 4 inches wide, which is a result of the size specification (Fig. 2). The frame is deleted. The JOURNAL style makes it monochrome, but also changes the symbol with each change in the bay group.

Designating specific symbol or line types is a common desire. The current example is essentially a hierarchical design with bays within estuaries (Table 1), so it would enhance the visualization if the symbols for the two bays within each estuary were the same, with one open and one closed. Because each bay is also near the river (i.e., a secondary bay), or far from the river (i.e., a primary bay). The bay types should have consistent symbols also. A new style with these characteristics can be created with the %MODSTYLE macro using the following statements. In many SAS statements, options follow a backslash (/); however in %MODSTYLE statement, the options are enclosed with parentheses. The symbol names are listed in the same sort order that the bay names are listed. Lines could be listed also. The symbols are a bit small, so the MARKERATTRS option will be used to increase the size.

```
%modstyle (parent=Journal, name=MyJournal, type=CLM,
          markers= triangle squarefilled diamond circle
                square trianglefilled diamondfilled circlefilled);

ods html style=MyJournal IMAGE_DPI=300 gpath="C:\Users\pmontagna\Documents\MySAS\Images";
ods graphics on / border=off height=3in width=4in outputfmt=png imagename="Salinty_v_P04_bw";
proc sgplot data=TXWQ_g;
  scatter x=Sal y=P04 / group=bay MARKERATTRS=(size=9)
          YERRORLOWER=lP04 YERRORUPPER=uP04
          XERRORLOWER=lSal XERRORUPPER=uSal;
  xaxis values=(10 to 50 by 10) minor;
run; quit;
```

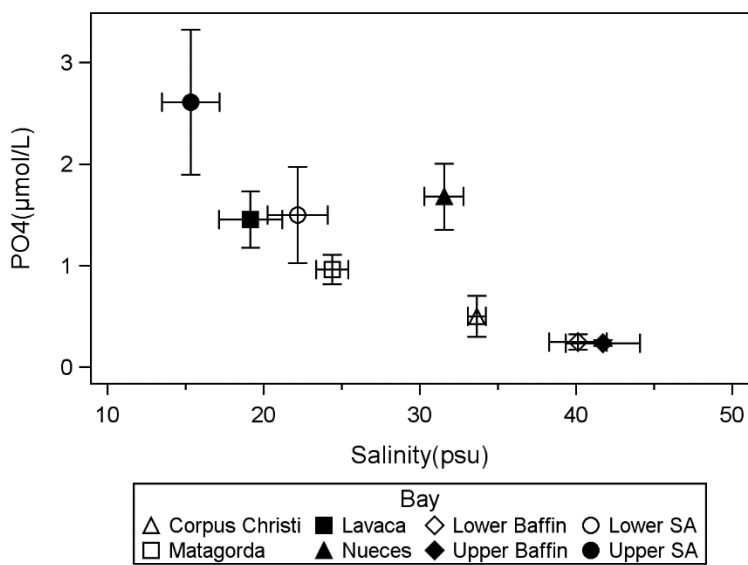


Figure 3. Graph using a modified style with designated symbols.

The relationships among and within estuaries are now easier to identify with the new symbols (Fig. 3). Secondary bays (closed symbols) always have higher nutrient concentrations than primary bays; except for Baffin Bay, which has the highest salinity. The larger symbol sizes make the treatment differences easier to see. All the bays are listed in alphabetical sort order, so two estuaries (Baffin Bay and San Antonio (SA) Bay) line up nicely, but the other don't.

The graph would be better if all the estuaries lined up with the secondary bays on the top line, and the primary bays on the bottom line. This can be accomplished using numbers as bay names and formats to print the desired labels. The bays can be renamed from 1 through 8 based on their geographic location (Fig. 4).

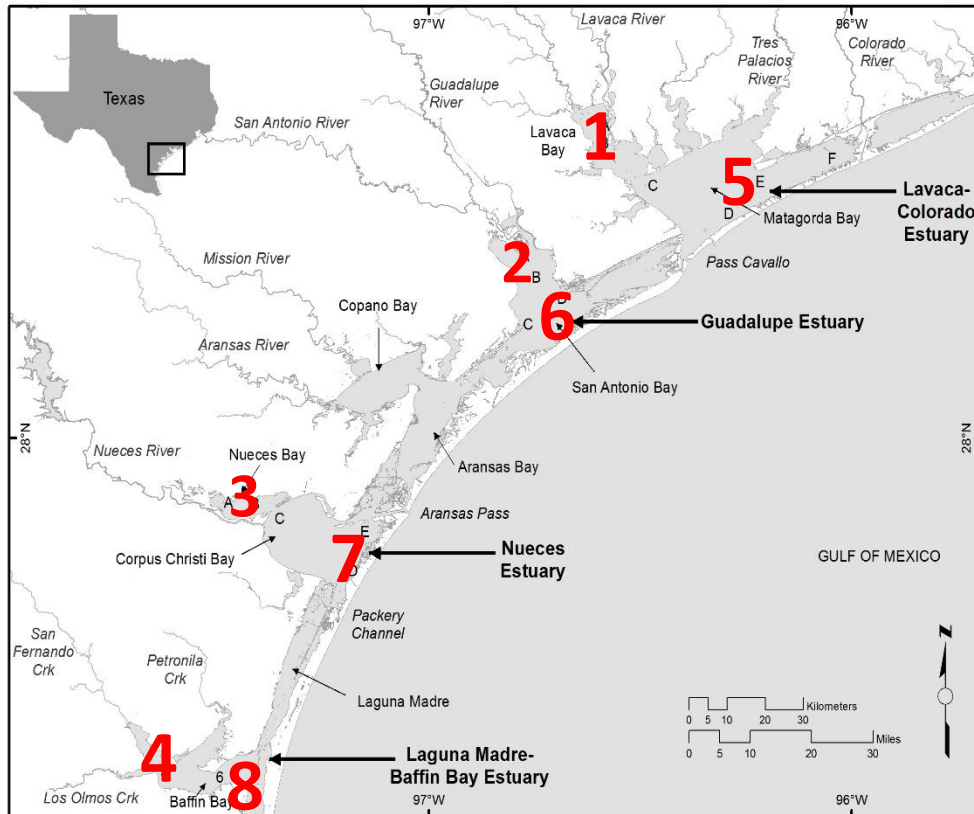


Figure 4. Map of study area with new bay number values.

The following statements creates a new data set with bay values and formatted names for the values.

```

data Ordered;
attrib sb length=3 format=3.; /* sort geoc bay */
set TXWQ_g;
if bay='Lavaca'           then sb=1; /* Secondary*/
if bay='Upper SA'        then sb=2; /* Secondary*/
if bay='Nueces'          then sb=3; /* Secondary*/
if bay='Upper Baffin'    then sb=4; /* Secondary*/
if bay='Matagorda'       then sb=5; /* Primary */
if bay='Lower SA'        then sb=6; /* Primary */
if bay='Corpus Christi'  then sb=7; /* Primary */
if bay='Lower Baffin'    then sb=8; /* Primary */
run;

/* Create formats for sorted bays and estuaries */
proc format;
value Bays
1= 'Lavaca'
2= 'Upper SA'
3= 'Nueces'
4= 'Upper Baffin'
5= 'Matagorda'
6= 'Lower SA'
7= 'Corpus Christi'
8= 'Lower Baffin' ;
run;

```

Now, simply use the new dataset, and add custom formatting and labeling; however the %MODSTYLE macro must be changed also.

```

%modstyle (parent=journal, name=MyJournal, type=CLM,
          markers= triangle      square      diamond      circle
                trianglefilled squarefilled diamondfilled circlefilled);
ods html style=MyJournal IMAGE_DPI=300 gpath="C:\Users\pmontagna\Documents\MySAS\Images";
ods graphics on / border=off height=3in width=4in outputfmt=png imagename="Salinty_v_P04_bw";
proc sgplot data=Ordered;
  scatter x=Sal y=P04 / group=sb MARKERATTRS=(size=9)
          YERRORLOWER=lP04 YERRORUPPER=uP04
          XERRORLOWER=lSal XERRORUPPER=uSal;
  xaxis values=(10 to 50 by 10) minor ;
  format sb bays.;
  label sb='Bays';
run;quit;

```

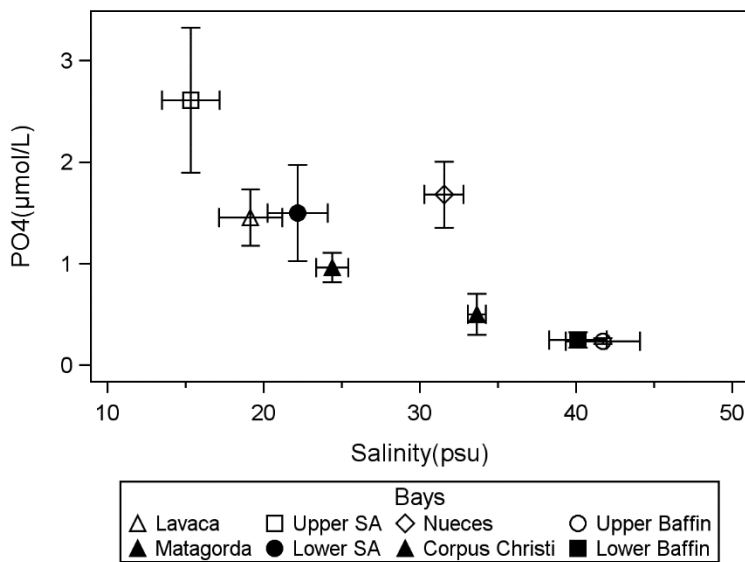


Figure 5. Graph using a customized ordered list symbols, and formatted labels.

The result is a graphic where the legend labels are ordered geographically (Fig. 5). Trends are easier to see, because salinity increases from left to right, and the labels are for estuaries that are aligned from northeast to southwest.

Finally, let's add color to enhance the graphic. Again, simply modify the %MODSTYLE macro. But, first we have to choose colors that would work well with the message of the graphic. Color ramps do this well. For the current example, a blue to brown color ramp would work to represent increasing salinity. Also, the Estuary-level treatment can be highlighted by using the same color for bays within estuary. Note, the filename is changed as well. Also, we may want to make a graphic for each nutrient, such that PO4 is subfigure A, so an inset is added.

```
%modstyle (parent=journal, name=MyJournal, type=CLM,
           markers= triangle      square      diamond      circle
                 trianglefilled squarefilled diamondfilled circlefilled,
           colors=Blue Green Orange Brown Blue Green Orange Brown);
ods html style=MyJournal IMAGE_DPI=300 gpath="C:\Users\pmontagna\Documents\MySAS\Images";
ods graphics on / border=off height=3in width=4in outputfmt=png imagename="Sal_v_PO4_color";
proc sgplot data=Ordered;
  scatter x=Sal y=PO4 / group=sb MARKERATTRS=(size=9)
          YERRORLOWER=lPO4 YERRORUPPER=uPO4
          XERRORLOWER=lSal XERRORUPPER=uSal;
  xaxis values=(10 to 50 by 10) minor ;
  format sb bays.;
  label sb='Bays';
  inset "A" / position=topright;
run;quit;
```

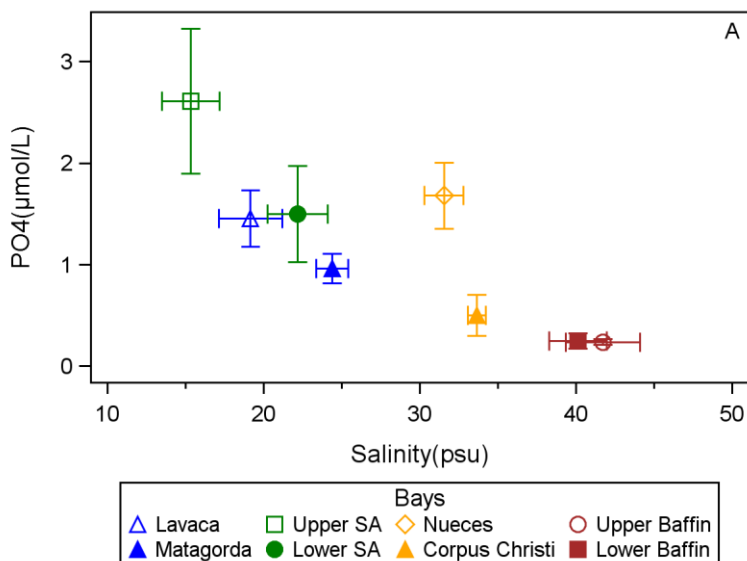


Figure 6. Final colored graphic with an inset to label a subfigure.

This example (Fig. 6) shows that color can be used to enhance the ability to easily see the trends in the data.

CONCLUSION

- You can make complex graphs easily using the statistical graphics (SG) procedures.
- You can control style, output, and file types using the ODS statements.
- You alter the style easily using %MODSTYLE macro.
- Start creating graphics with a monochrome version, and then add color for emphasis or clarity.

REFERENCES

Montagna, P.A., X. Hu, T.A. Palmer, and M. Wetz. 2018. Effect of hydrological variability on the biogeochemistry of estuaries across a regional climatic gradient. *Limnology and Oceanography* 63:2465-2478. [doi: 10.1002/lno.10953](https://doi.org/10.1002/lno.10953).

SAS Institute Inc. 2013. *SAS® 9.4 ODS Graphics: Getting Started with Business and Statistical Graphics*. Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2016. *SAS® 9.4 ODS Graphics: Procedures Guide, Sixth Edition*. Cary, NC: SAS Institute Inc.

ACKNOWLEDGMENTS

This publication was made possible, in part, by the National Oceanic and Atmospheric Administration award (NA15NOS4780185). Its contents are solely the responsibility of the award recipient Paul Montagna and do not necessarily represent the official views of the U.S. Department of Commerce, National Oceanic and Atmospheric Administration.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Paul A. Montagna
Harte Research Institute for Gulf of Mexico Studies
Texas A&M University-Corpus Christi
6300 Ocean Drive, Unit 5869
Corpus Christi, Texas 78712
paul.montagna@tamucc.edu
Office (361) 825-2040
<http://harteresearchinstitute.org/>

TRADEMARK INFORMATION

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 trademarks of their respective companies.