About the presenter

Alan Silva is a Statistician, has a Master and a PhD in Transportation and he is an Associate Professor of Statistics at University of Brasilia, Brazil. Working with SAS since 2002 and developing solutions using SAS/IML and SAS/AF.

SAS® GLOBAL FORUM 2016

IMAGINE. CREATE. INNOVATE.

Working with PROC SPP, PROC GMAP and PROC GINSIDE to Produce Nice Maps

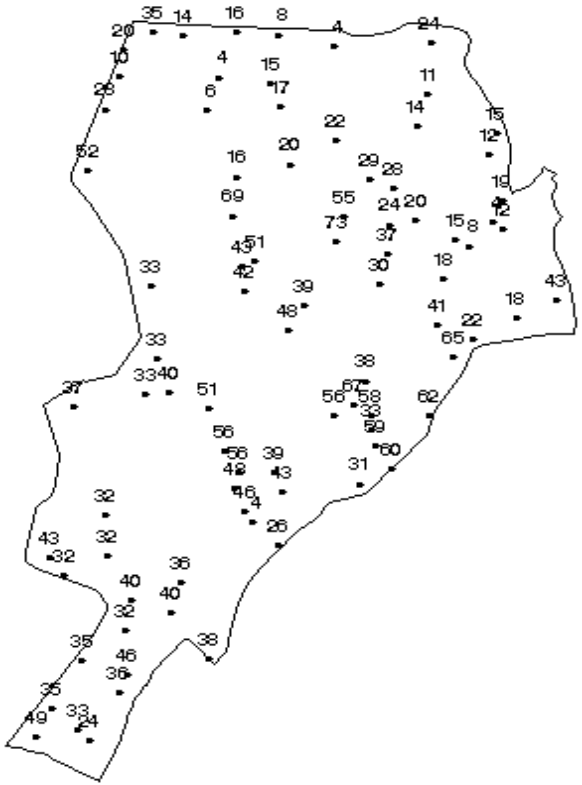
Alan Ricardo da Silva, PhD
University of Brasilia, Brazil

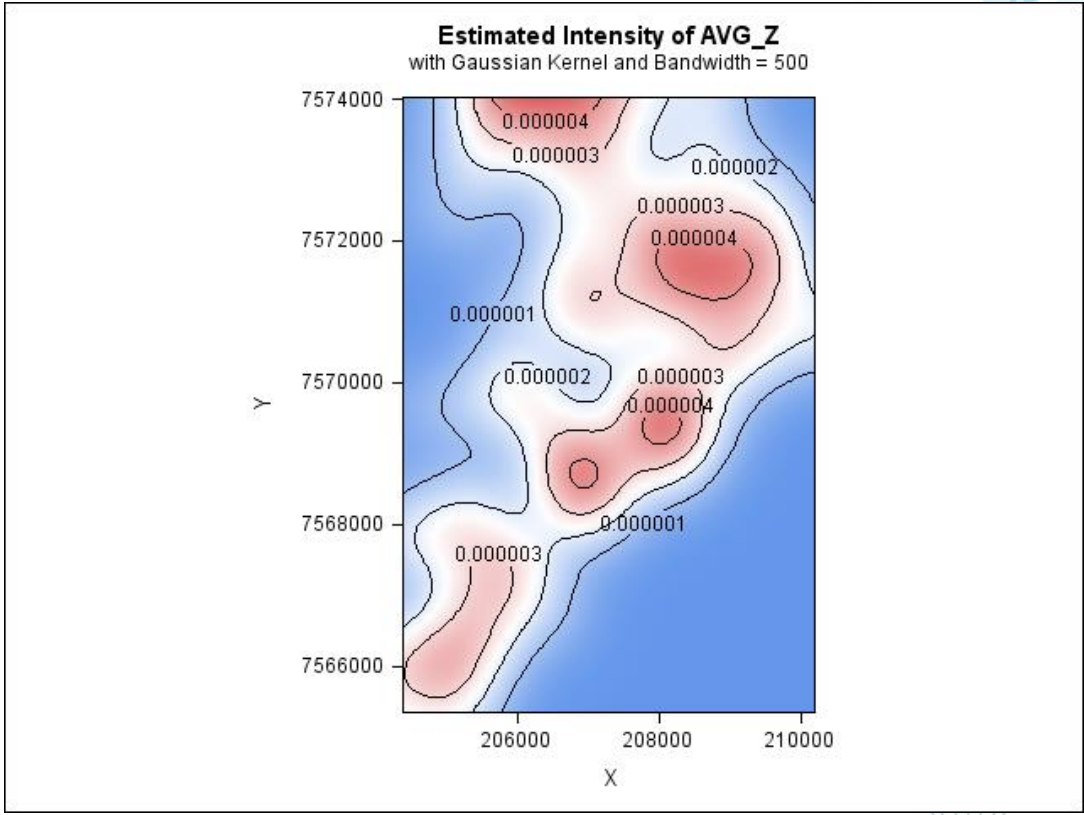
#SASGF

Spatial Point Pattern

- The new PROC SPP (Spatial Point Pattern) deal with spatial data, which are a collection of locations of single events of a spatial process (SAS, 2014);
- It is possible to use PROC SSP to create a surface of the intensity of the point pattern process;
- The problem is that PROC SPP generates data only for a squared area, even the data are bordered by an irregular area.

Spatial Point Pattern





Spatial Point Pattern

- The first analysis in order to characterizing the intensity of the data points in an area can be done by a kernel estimator of the intensity function. The general form of this kind of estimator is given by (Cressie, 1991):

$$\hat{\lambda}_h(\boldsymbol{s}) = \frac{1}{\rho_h(\boldsymbol{s})}\left\{\sum_{i=1}^{n} k_h(\boldsymbol{s}-\boldsymbol{s_i})\right\}$$

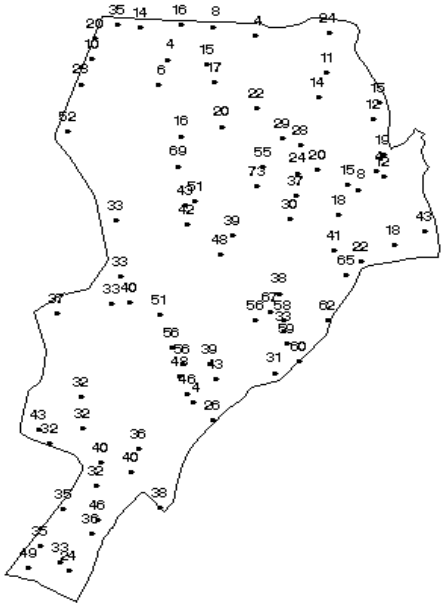
$$\hat{\lambda}_h(\boldsymbol{s}) = \frac{1}{\rho_h(\boldsymbol{s})}\left\{\sum_{i=1}^{n} h^{-2} k_h\left(\frac{\boldsymbol{s}-\boldsymbol{s_i}}{h}\right)\right\}$$ (PROC SPP)

$$K(d) = \frac{e^{\frac{-d^2}{2h^2}}}{\sqrt{2\pi}}$$

$$d = \sqrt{(x_i - x_j)^2 + (y_i + y_j)^2}$$

Illustration

- let us use an irregular shape from the Canchim farm (EMBRAPA) in São Carlos, São Paulo, Brazil. There are 85 data referring to the clay content.



The shape file (*.shp) can be imported by PROC MAPIMPORT.

Illustration

- First, one can use PROC SQL to select the borders of the area named MINX, MINY, MAXX, MAXY.

```
proc sql noprint;
 select min(x) into:minx from sao_carlos;
 select min(y) into:miny from sao_carlos;
 select max(x) into:maxx from sao_carlos;
 select max(y) into:maxy from sao_carlos;
quit;
%put minx=&minx maxx=&maxx miny=&miny maxy=&maxy;
```

Illustration

- After that, one can use that information about the borders of the area in the AREA= option of the PROCESS statement of PROC SPP. The b= option referred to the kernel bandwidth parameter of the kernel first-order intensity estimates and GRID= specifies a reference grid for computing the kernel estimate.

```
proc spp data=sao_carlos_pt plots(equate)=(trends observations);
   process   AVG_Z   =   (x,   y   /area=(&minx,&miny,&maxx,&maxy)
Event=AVG_Z) /
   kernel(type=gaussian b=500 out=kernel grid(90,90));
run;
```

Illustration

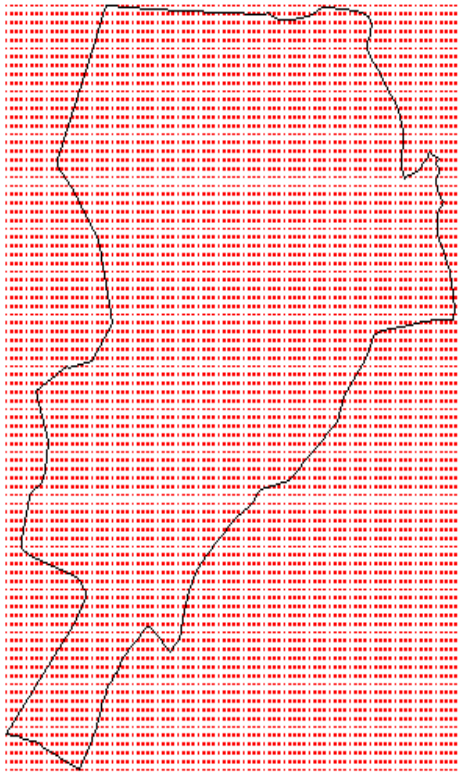
- To plot the results, one can use the ANNOTATE Facility from the dataset generated by KERNEL sub-option OUT= and PROC GMAP with ANNO= option in the CHORO statement.

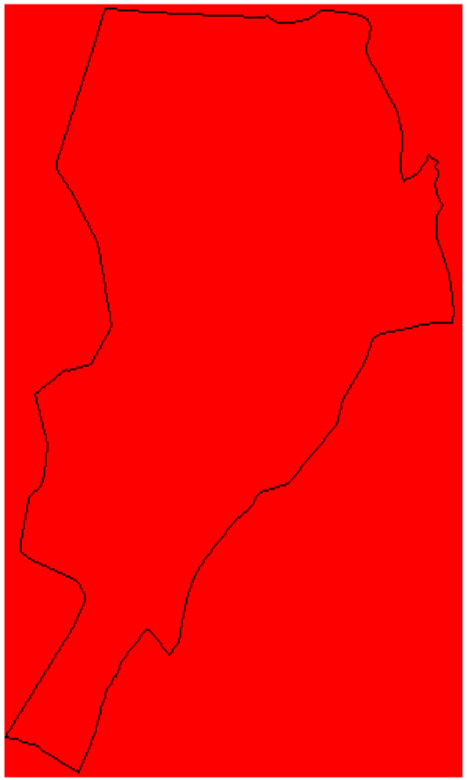
Just remember to rename the variables `GXC` and `GYC` to `X` and `Y`, respectively

```
data anno;set kernel(rename=(GXC=x GYC=y));
 length function style $10. color $8.;
 retain line 1 xsys ysys '2' hsys '3' color 'red';
 function='label';text='U';position='5';style='marker';
 size=1;
run;
proc gmap data=a map=sao_carlos all;
 id segment;
 choro v / anno=anno nolegend;
run;quit;
```

Illustration

- Using SIZE=0.5 (small squares) we can see how the coordinates are distributed on the area (left) an s) we can see thes ous way" on the a





Illustration

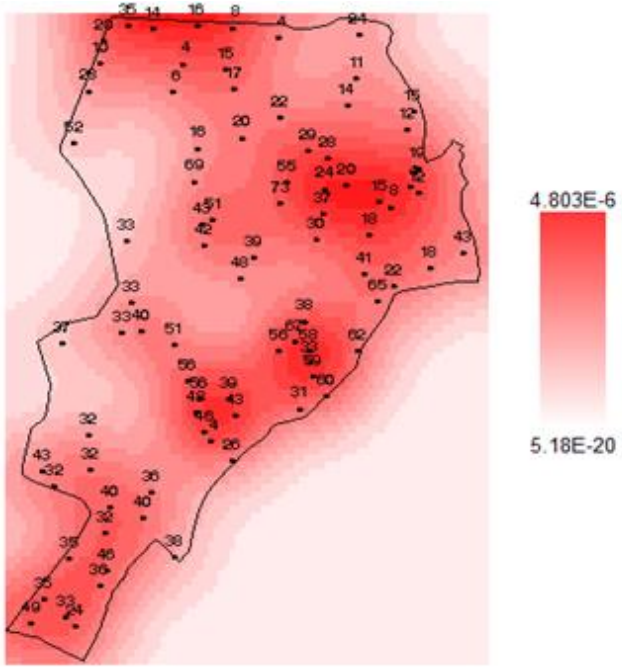
- Finally, to plot the continuous surface one can use the program described in the paper to color each coordinate (square created by the ANNOTATE Facility) and to create a continuous bar. This task can be done with %colorscale macro (SAS, 2003) with some adaptations. This macro is on Appendix I.

```
%colorscale(FFFFFF,,FF0000,&nc,clist,no);%patt;

%bar(FF3333,FFFFFF,&min,&max,vertical,y_i=44,x_i=80);
```

Illustration

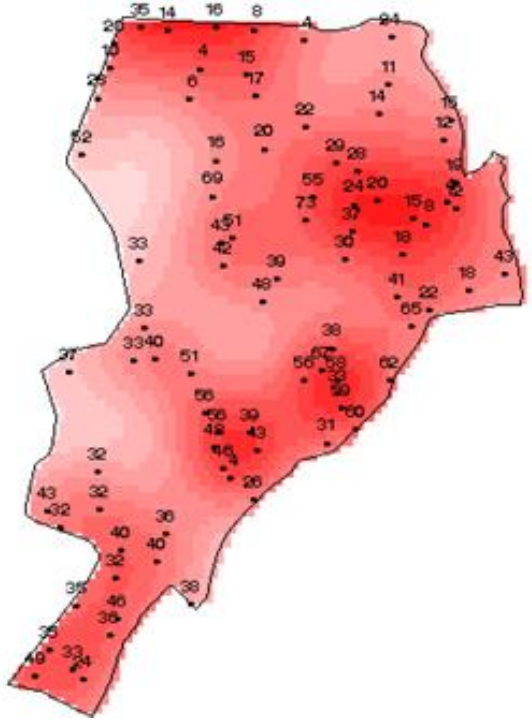
```
proc gmap data=a map=sao_carlos all anno=anno_points;
 id segment;
 choro v / anno=anno nolegend;
run;
quit;
```



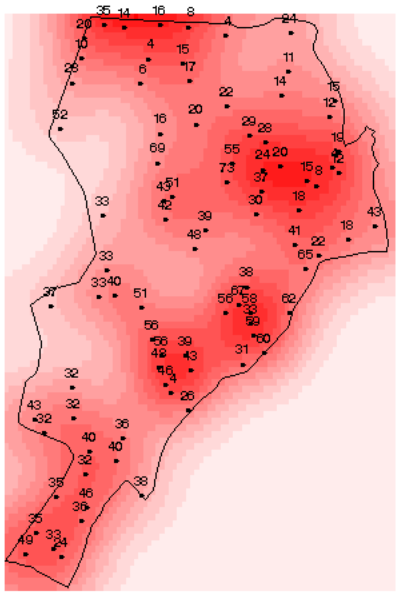
Illustration

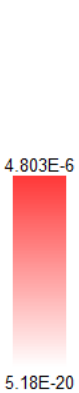
- Use PROC GINSIDE to show the intensity estimates only for the coordinates which are inside the polygon .

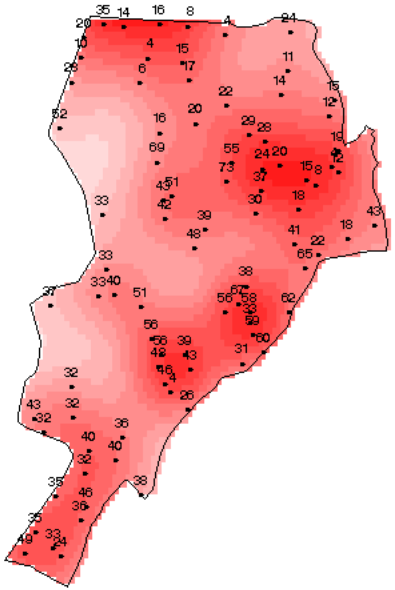
```
proc ginside data=anno map=Sao_carlos
out=anno2 insideonly;
 id segment;
 run;
data anno2;set _a_ anno2;
proc gmap data=a map=sao_carlos all anno=anno
 id segment;
 choro v / anno=anno2 nolegend;
run;quit;
```

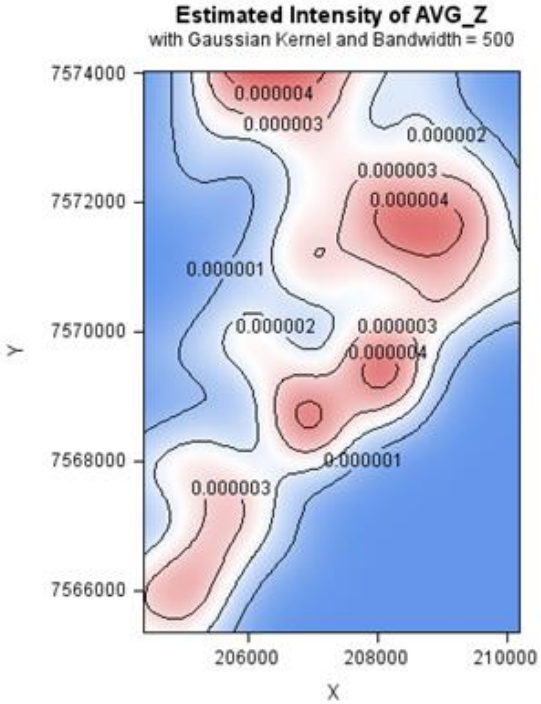


Illustration









Much Better!!

Acknowledgments: FAPDF, Brazil

Thank you!!

alansilva@unb.br