

Design Interface: Interactive tools for editing and evaluating spatial designs.

Douglas Nychka and Eric Gilleland
Nancy Saltzman and J. Andrew Royle



Geophysical Statistics Project,
National Center for Atmospheric Research
www.cgd.ucar.edu/stats

- The problem
- Spatial models for air quality
- Di Toolbar
- An Example: 8-hour average ozone in NC

National Center for Atmospheric Research



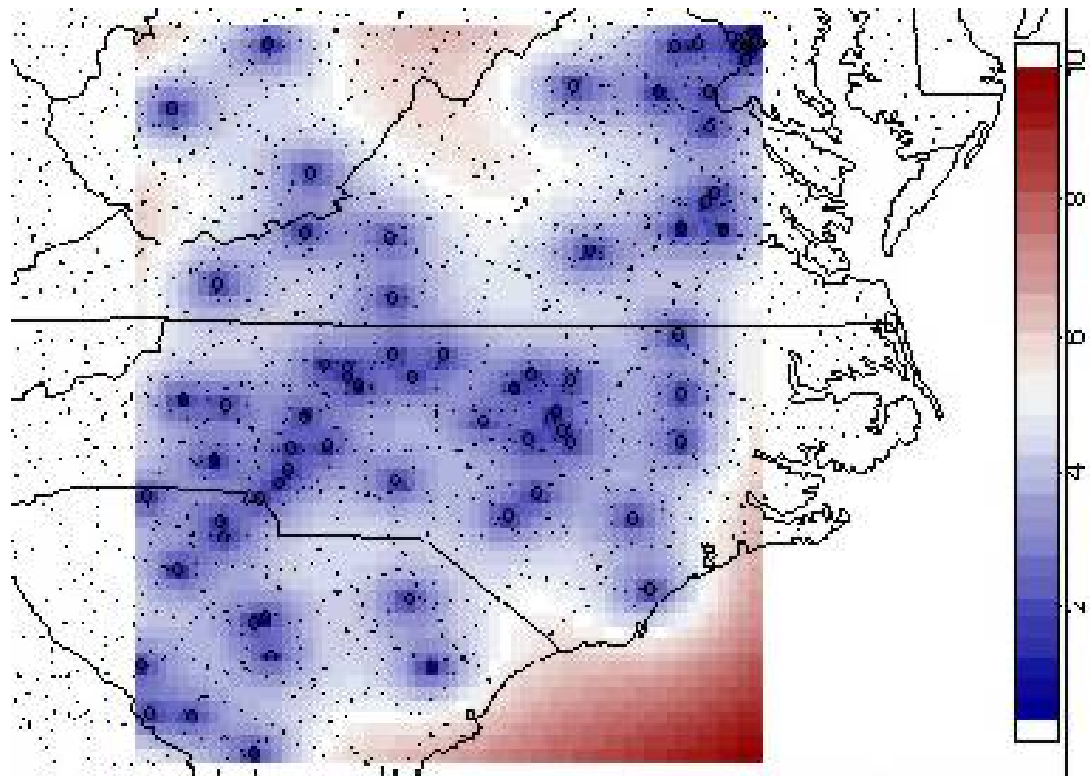
\approx 1000 people total, several hundred PH D (physical) scientists,
half the budget (\approx 60M) is a single grant from NSF-ATM

Research on nearly every aspect related to the atmosphere Climate, Weather, the Sun, Ocean/atmosphere, Ecosystems, Economic impacts, Air quality, Instrumentation, Scientific computing and ...

Statistical methods for the geosciences

Translating monitoring measurements into informative spatial information.

Spatial statistics can provide extensive information about the spatial characteristics of a pollutant and also the precision of a monitoring network.



Design Interface was proposed as a means to make this statistical information readily available to non-statisticians.

- What is a best estimate of a pollutant at an unmonitored location?
- What is the estimate of an air quality standard?
- What is the uncertainty in the spatial predictions?
- What is the effect of increasing or decreasing the monitoring locations on spatial prediction?

Basic premises:

- Measurements made at a point location have some relevance for an area surrounding the site.
- Correlations among locations are roughly consistent over time.

The first assumption suggests using spatial statistics to predict pollutants at unmonitored locations. The second allows for modeling the spatial structure using data at more than one time.

Spatial Models

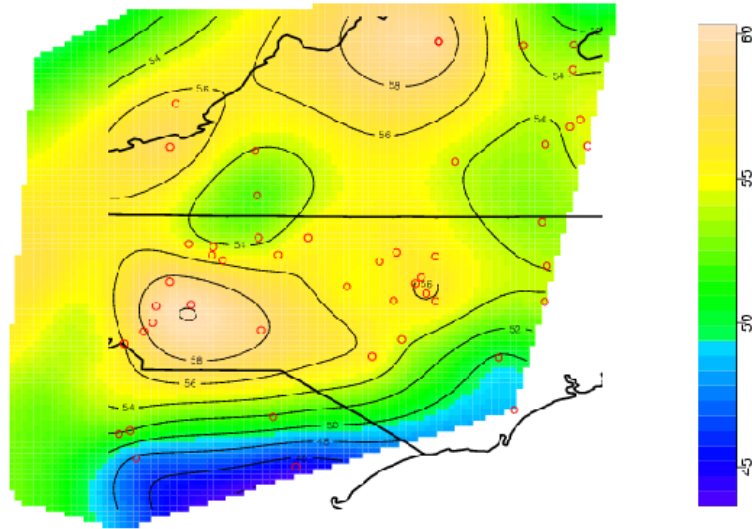
Starting point for spatial model

- Measurement is normally distributed or can be transformed (e.g. square root) to be normal $z(\mathbf{x})$.
- Each location has an associated mean level and variance
- The correlation among any two locations is specified.

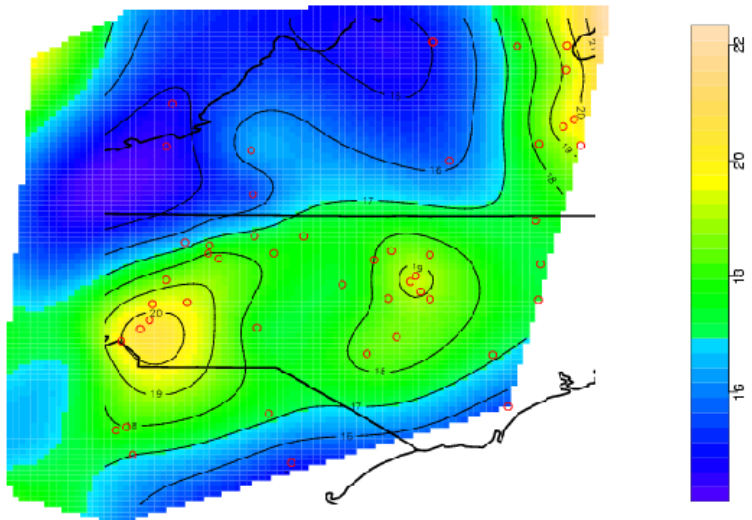
Each of these items can slowly vary over time or depend on seasonality.

Ingredients for a spatial model

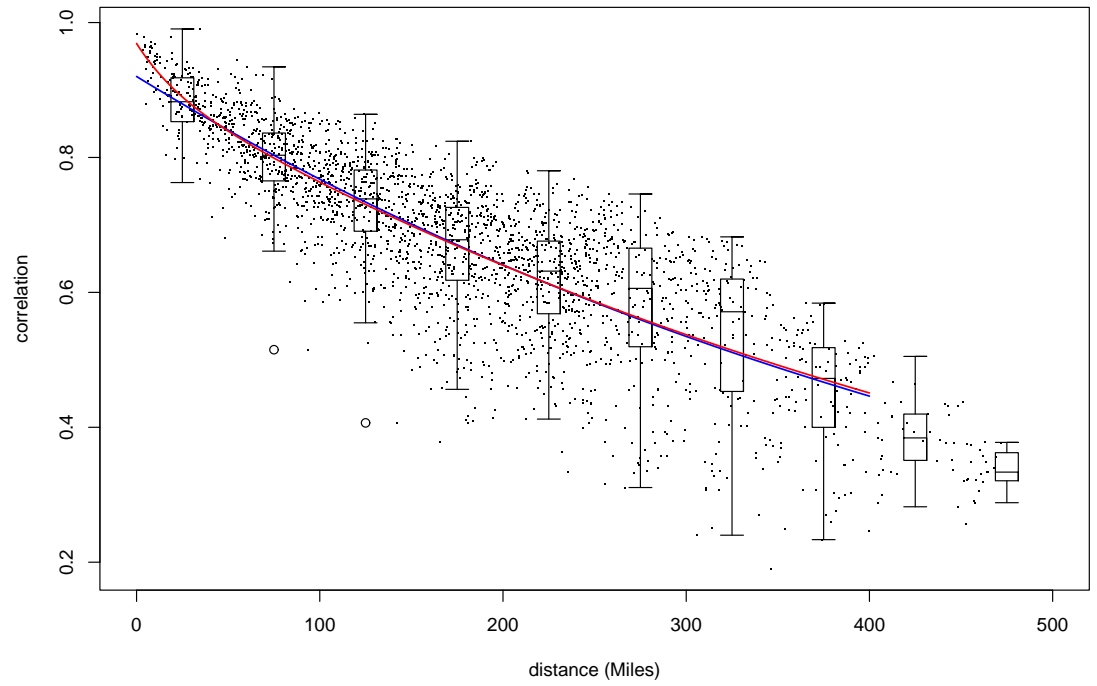
Mean 8-hour average O3



Standard Deviation 8-hour average O3



Spatial dependence



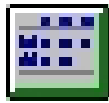
DI functions

How DI works: Creates, modifies and displays a network object using the **S** statistical environment. The initial version assumes that the basic data has been read into **S**.

The Tool Bar



- Define a covariance
- Create a network
 - Network locations
 - Spatial covariance
 - Data (optional)
- Network summary plot
- Spatial prediction plot
- Editing a network



Defining a covariance

Create a new covariance object

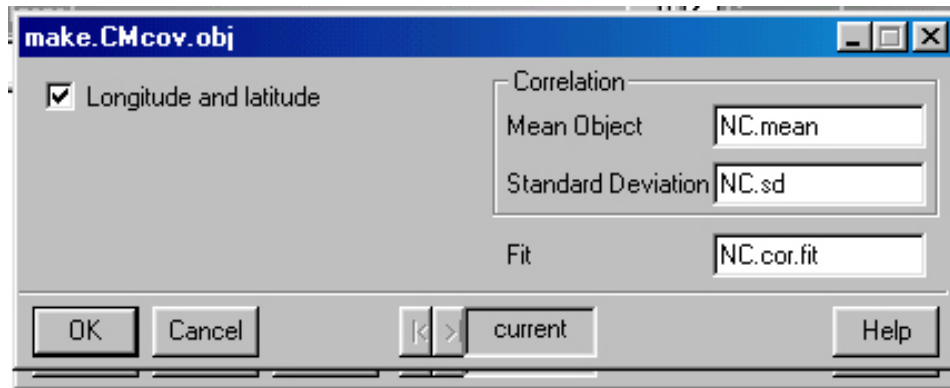
New Covariance Object Name:

Type:

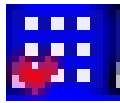
- CM.cov
- Exponential
- Other

CM.cov Correlation model

Adding some details:



NC.mean and NC.sd are the mean and standard deviation surfaces. NC.cor.fit is correlation as a function of distance.



Creating a network object

Create a new network object

Return Value
Network Object: NC.net

Covariance Object
Covariance Object: NC.cov.obj

Design
Locations: NC.ex\$loc
Measurements: NC.ozmax8

Grid Points
Grid Points for Prediction: [dropdown]
x lower: [input]
x upper: [input]
y lower: [input]
y upper: [input]
Number of x grid points: 50
Number of y grid points: 80

OK Cancel Apply < > current Help

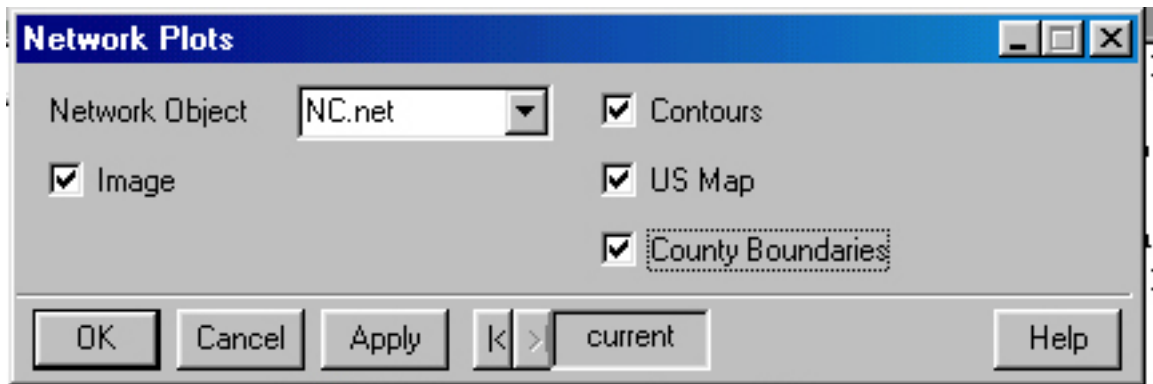
We used the covariance `NC.cov.obj` from previous step.

Location (lon/lat) information is `NC.ex$loc`

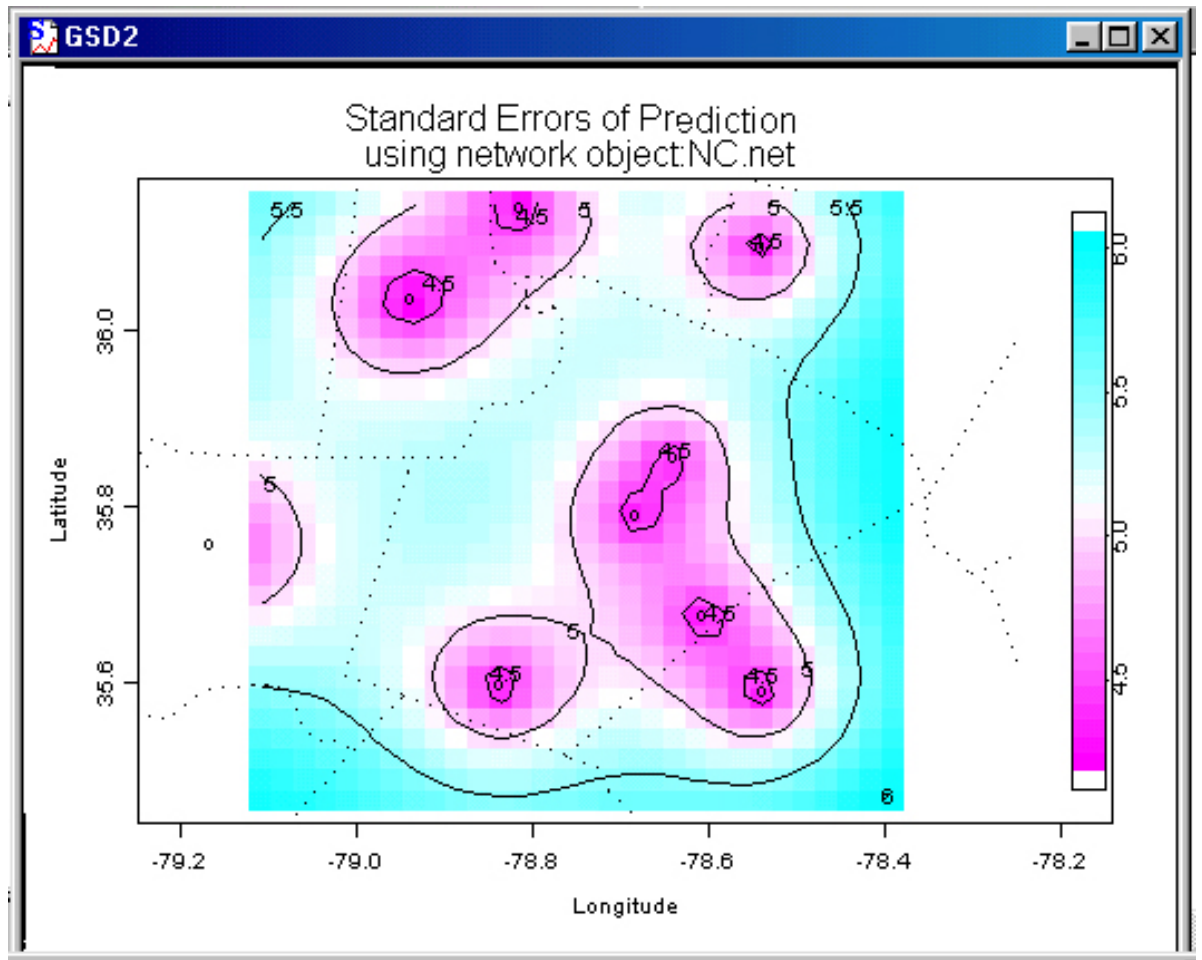
Data matrix is `NC.ozmax8`.



Graphical summary of network
Now it's easy!



Standard errors of prediction within RTP, NC.





Analyzing observed data

The screenshot displays the S-PLUS software environment. The main window is titled "S-PLUS - Commands" and contains a menu bar (File, Edit, View, Insert, D), a toolbar, and a command window. The command window shows the following code and output:

```
3: get(nw.obj$Yname)[row.name, ]
2: plot.prob(nw.obj = NC.net2, row.name = "08/05/1998",
1:
>
```

The output of the command is a list of dates in quotes, arranged in two columns:

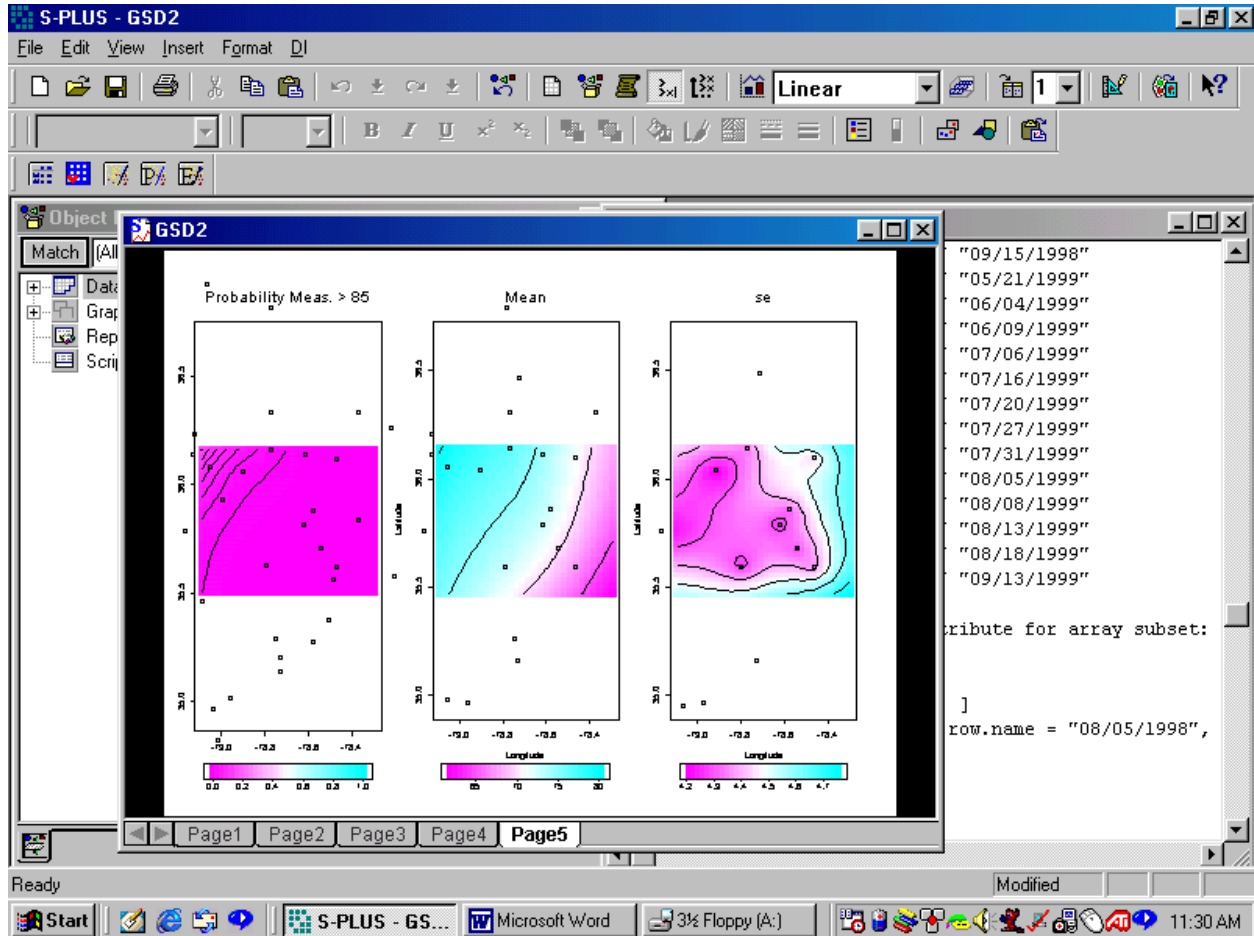
```
"09/14/1998" "09/15/1998"
"09/17/1998" "05/21/1999"
"05/30/1999" "06/04/1999"
"06/08/1999" "06/09/1999"
"06/11/1999" "07/06/1999"
"07/15/1999" "07/16/1999"
"07/19/1999" "07/20/1999"
"07/23/1999" "07/27/1999"
"07/30/1999" "07/31/1999"
"08/04/1999" "08/05/1999"
"08/07/1999" "08/08/1999"
"08/12/1999" "08/13/1999"
"08/17/1999" "08/18/1999"
"08/28/1999" "09/13/1999"
```

The "Network Probability Plot" dialog box is open, showing the following settings:

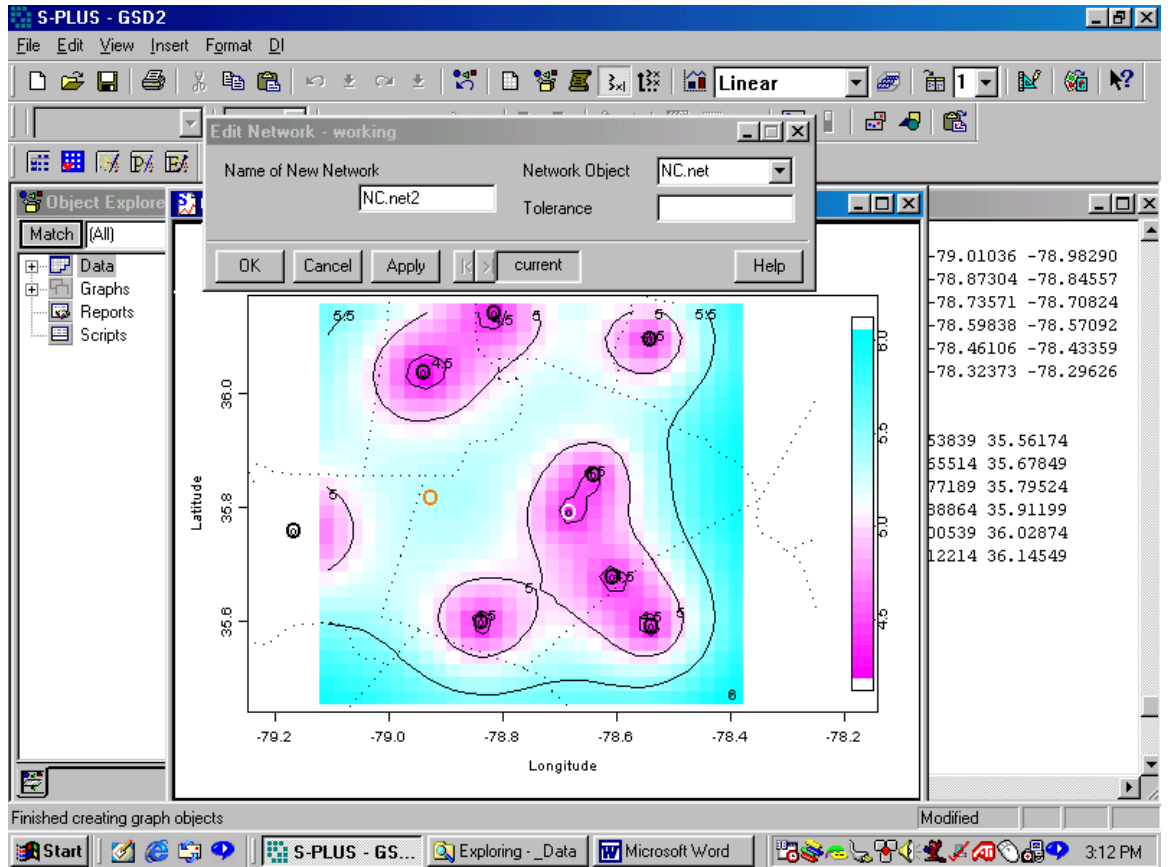
- Main Tab:**
 - Network Object: NC.net
 - Date: 08/05/1998
 - Exceedence: 85
- Advanced Tab:**
 - Krig Smoothing: GCV
 - df: 16 (range 3 to 50)
- Plots:**
 - Plot: Probabilities
 - US Map
 - County Boundaries

The dialog box also has buttons for OK, Cancel, Apply, and Help. The "Apply" button is labeled with "k" and "current".

Summary plot probability of exceedence, mean and standard error.



Editing a network



Future work

- GUI support for simple I/O to common formats.
- Tutorial on identifying spatial models.
- Diagnostics for adequacy of spatial model.
- Spatial models for the ozone standard (third highest yearly 8 hour average)