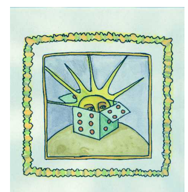
Measuring air quality standards and spatial designs

Douglas Nychka and Eric Gilleland
Geophysical Statistics Project,
National Center for Atmospheric Research
www.cgd.ucar.edu/stats

- The problem: A nonlinear, seasonal statistic
- A space-time model for ozone
- Sampling the posterior for the regulatory standard



The problem

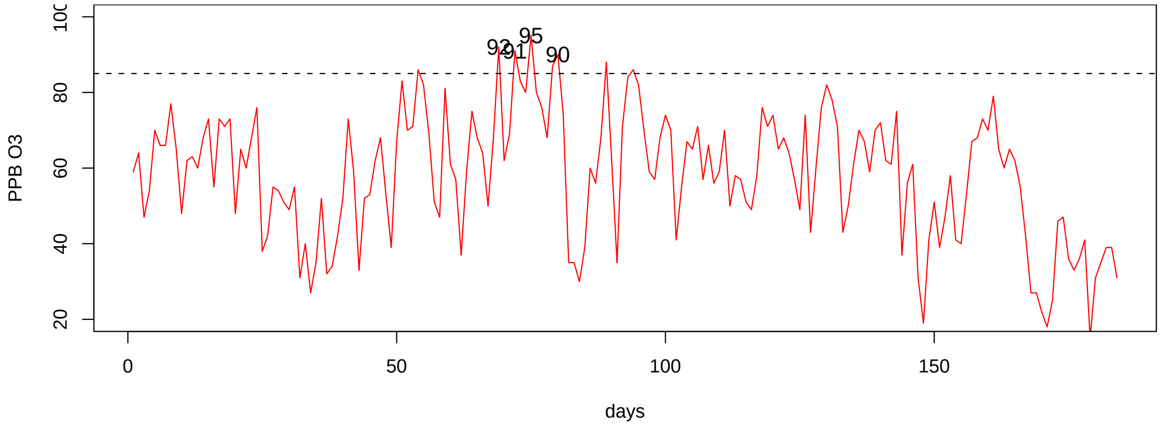
A suggested ozone pollutant standard is based on the fourth highest (max) 8-hour daily average (FHDA) recorded during the year. Compliance is related to a three year average being less than 85 PPB.

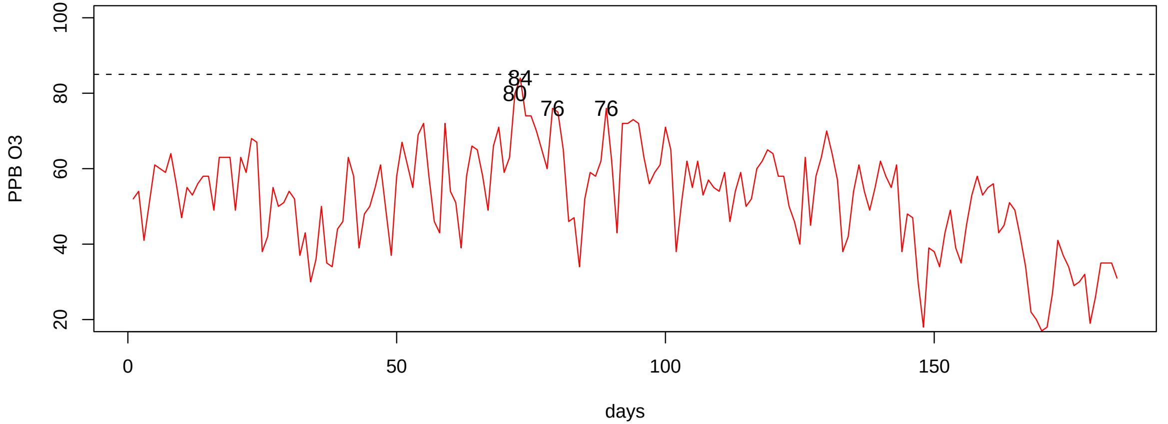
Spatial inference for the standard
What can be deduced about FHDA at unobserved locations?

Thinning monitoring networks
How should the existing network be reduced but still maintain the best performance?

Main idea:
Use a simpler daily model for ozone to sample the nonlinear seasonal statistic.

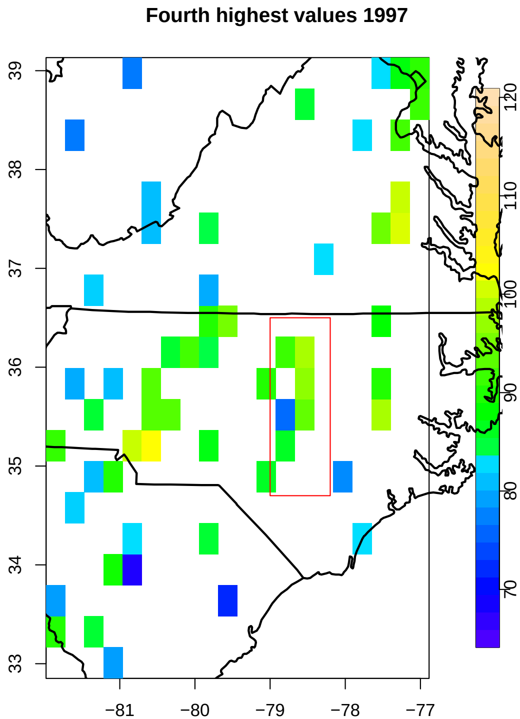
Two monitoring stations in Raleigh/Durham Area, North Carolina (1997)

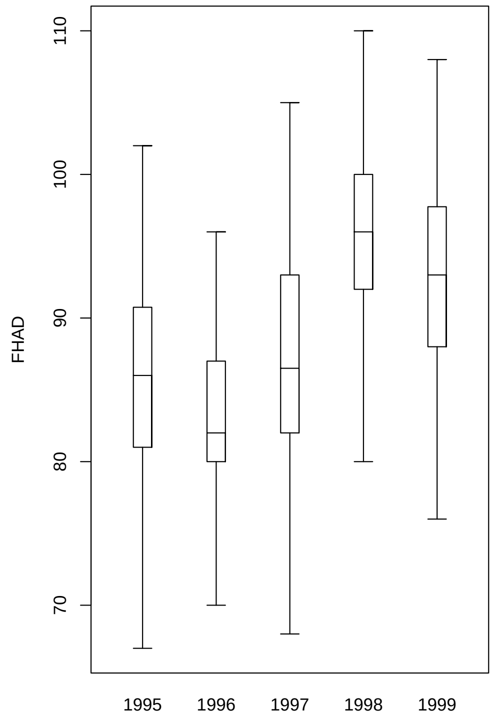




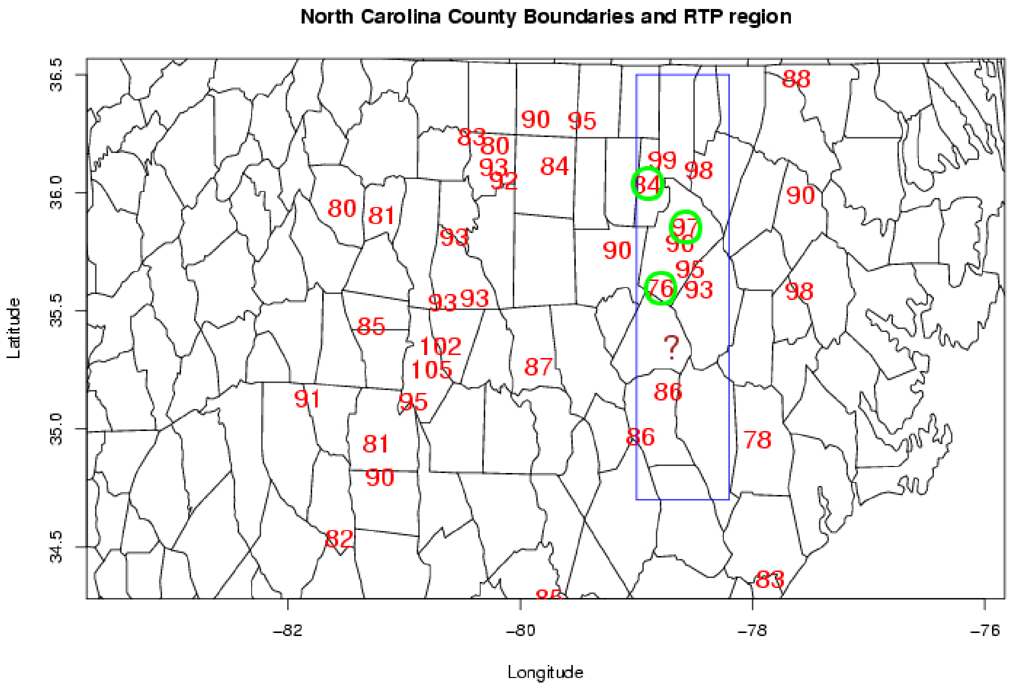
An example for North Carolina

Fourth highest daily average (FHDA) values (1997)





By counties (1997)



Spatial models for the annual fourth highest

Although it is straight forward to build spatial statistical models for the daily ozone field,

the extension to the fourth highest measurement is difficult ....

- Time dependence
- Nongaussian statistic (extreme value)
- Covariance structure ???

The idea is to determine the distribution of the FHDA from simulating the daily ozone fields.

Model components

Transform: $O(\boldsymbol{x}, t) =$ 8-hour ozone at location $\boldsymbol{x}$ and time t.

$$u(\boldsymbol{x}, t) = \frac{O(\boldsymbol{x}, t) - \mu(\boldsymbol{x}, t)}{\sigma(\boldsymbol{x})}$$

Autoregression: $u(\boldsymbol{x}, t) = \rho(\boldsymbol{x}) u(\boldsymbol{x}, t) + e(\boldsymbol{x}, t)$

Spatial dependence: $e(\boldsymbol{x}, t)$ uncorrelated and stationary over time but correlated over space.

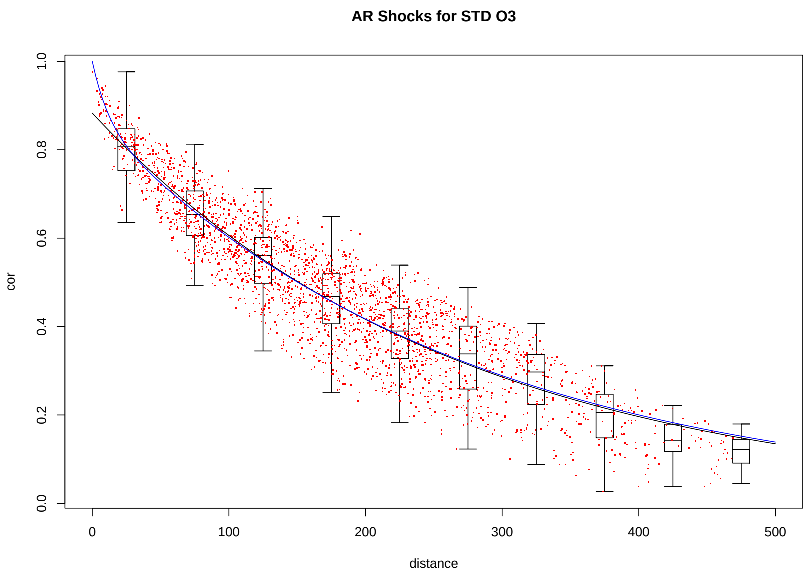
Under the assumption of multivariate normality one can generate fields of daily ozone.

Spatial dependence

Correlogram of shocks suggests an isotropic, mixture of exponential covariances

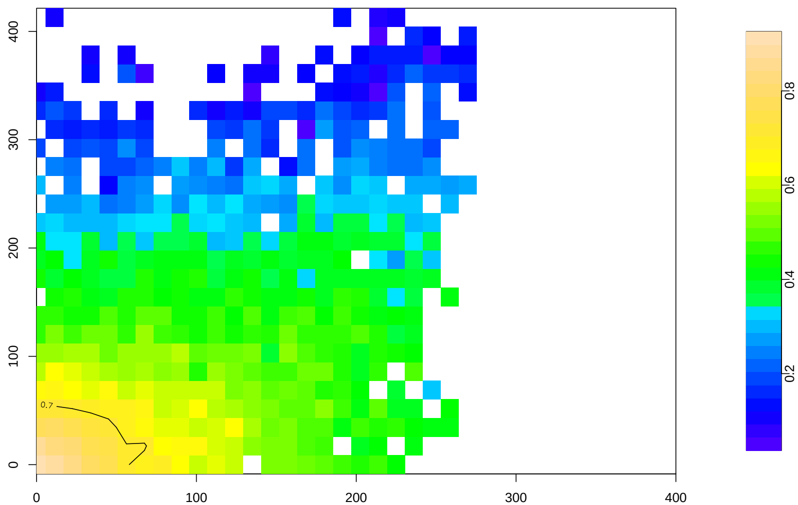
$$\alpha e^{-d/\theta_1} + (1-\alpha)e^{-d/\theta_2}$$

where d is the separation distance: $||\boldsymbol{x} - \boldsymbol{x}'||$.



Anisotropy?

Correlations as a function of both E-W and N-S differences.



Generating a conditional field.

The goal is to sample from $[FHDA(\boldsymbol{x})|dailydata]$. where $\boldsymbol{x}$ is an arbitrary location.

Generating one sample of the FHDA

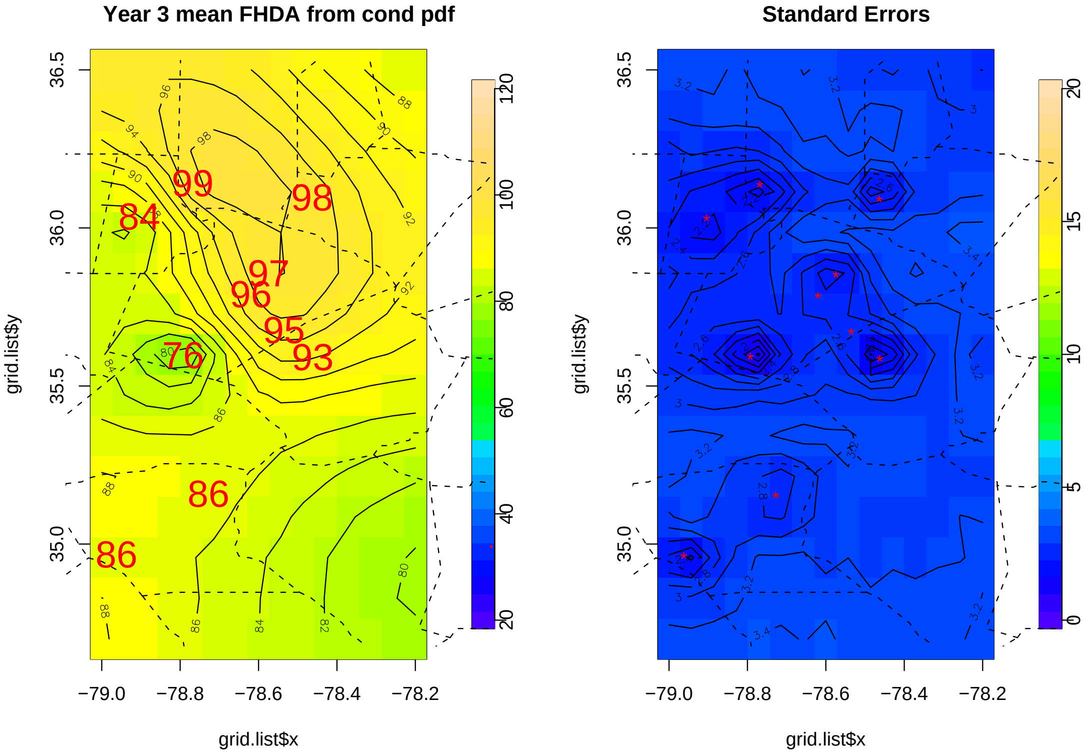
Start with initial value: $O_0(\boldsymbol{x})$, $u_0 = (O_0(\boldsymbol{x}) - \mu(\boldsymbol{x}))/\sigma(\boldsymbol{x})$ Loop through entire summer season

1. *conditional shock* sample from $[e_t(\boldsymbol{x})|\boldsymbol{e}_t^s]$
   where $\boldsymbol{e}_t^s$ are the shocks at the observed locations.
2. *propagate* $u_t = \rho(\boldsymbol{x})u_{t-1} +$ conditional shock
3. *back transform* $O_t(\boldsymbol{x}) = u_t\sigma(\boldsymbol{x}) + \mu(\boldsymbol{x})$

Calculate FHDA($\boldsymbol{x}$) based on conditional daily ozone values.

Repeat this process to accumulate a Monte Carlo sample of $[FHDA(\boldsymbol{x})|dailydata]$.

Results for 1997 RTP



Shortcuts

This seems like a lot of work just because we don't know the covariance!

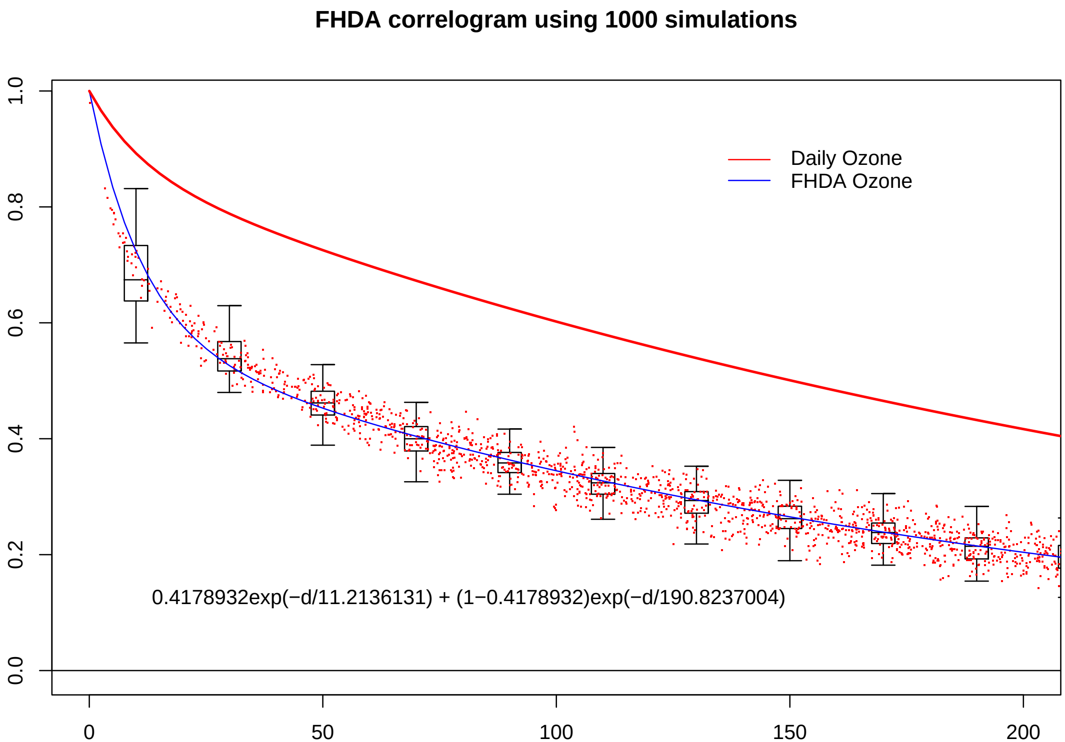
Especially to implement analysis of FHDA in an interactive framework.

What is the joint distribution of FHDA?

Is there any hope of it being simple?

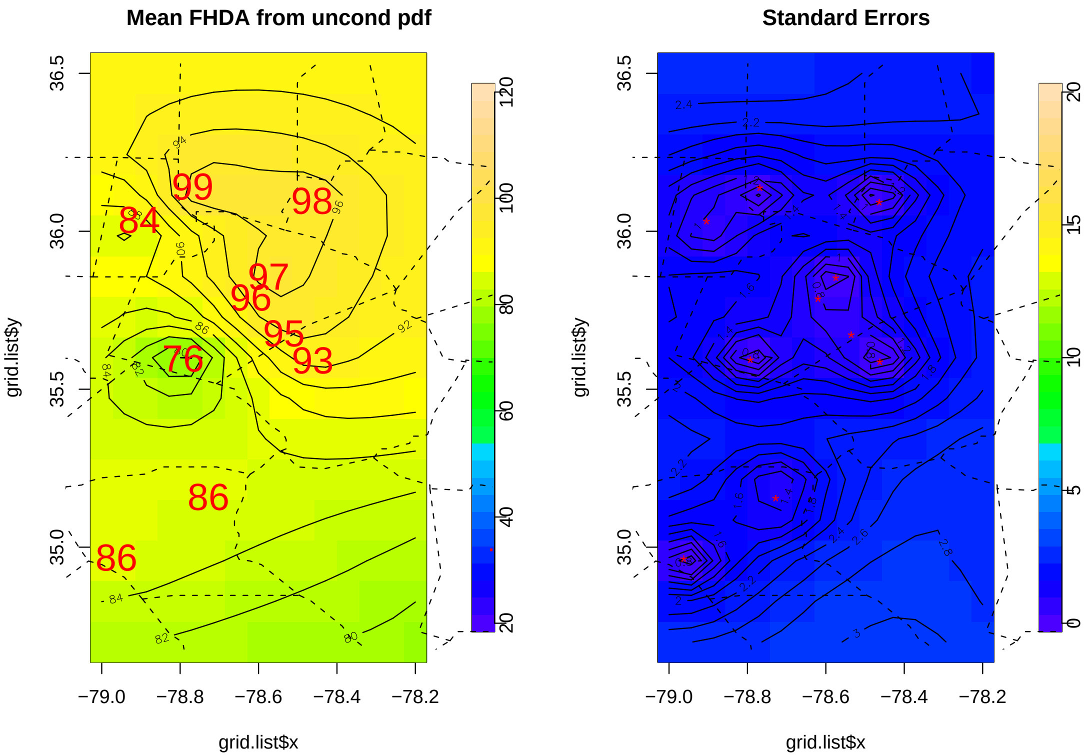
Estimated correlations for FHDA

Relationship among correlations



Also, bivariate scatterplots appear to be approximately normal.

Results for 1997 RTP, Approximate seasonal model



Discussion

- Examine effect of thinning O3 network using different design algorithms.
- Extend model to particles (PM).
- Extend model to more general tail behavior instead of just the fourth highest.