

# Air Quality Data: A New Conceptual Approach



Spatial Data Analysis Workshop  
Alan J. Cimorelli, EPA Region 3  
December 3, 2001



# Outline

- Goal
- Present Concept
- New Concept
- Advantages
- Details of Approach
- Approach to Network Design
- Network Design Example
- Next Steps



# Basic Goal: Air Quality Data Collection

Produce as complete and accurate a picture (spatial field) of air quality in as cost effective a manner as possible.



# Present Concept

- Tenets:
  - Air Quality Data (AQD) are truth
  - If no monitors then no information
- Problems
  - Ignores substantial & relevant information (i.e., Interpolation)
  - Concept is too limited for planning purposes (we must estimate)
  - Therefore we use AQD to create an implicit spatial picture: For Designations - **AQD represent only the county in which they are taken**
  - Disincentive to monitor



# New Concept

- Tenets:
  - Measured or interpolated data are the same except for uncertainty – They are an estimate of the actual field
  - Define Air quality as an estimated field of concentrations with associated uncertainties
- Estimate Concentration Field:
  - Recognize that AQD are simply a sample of the “Actual” air quality FIELD
  - Then, estimate the complete field by interpolating (kriging) the AQD
- Estimate an associated uncertainty field based on area modeling



# Advantages to New Concept

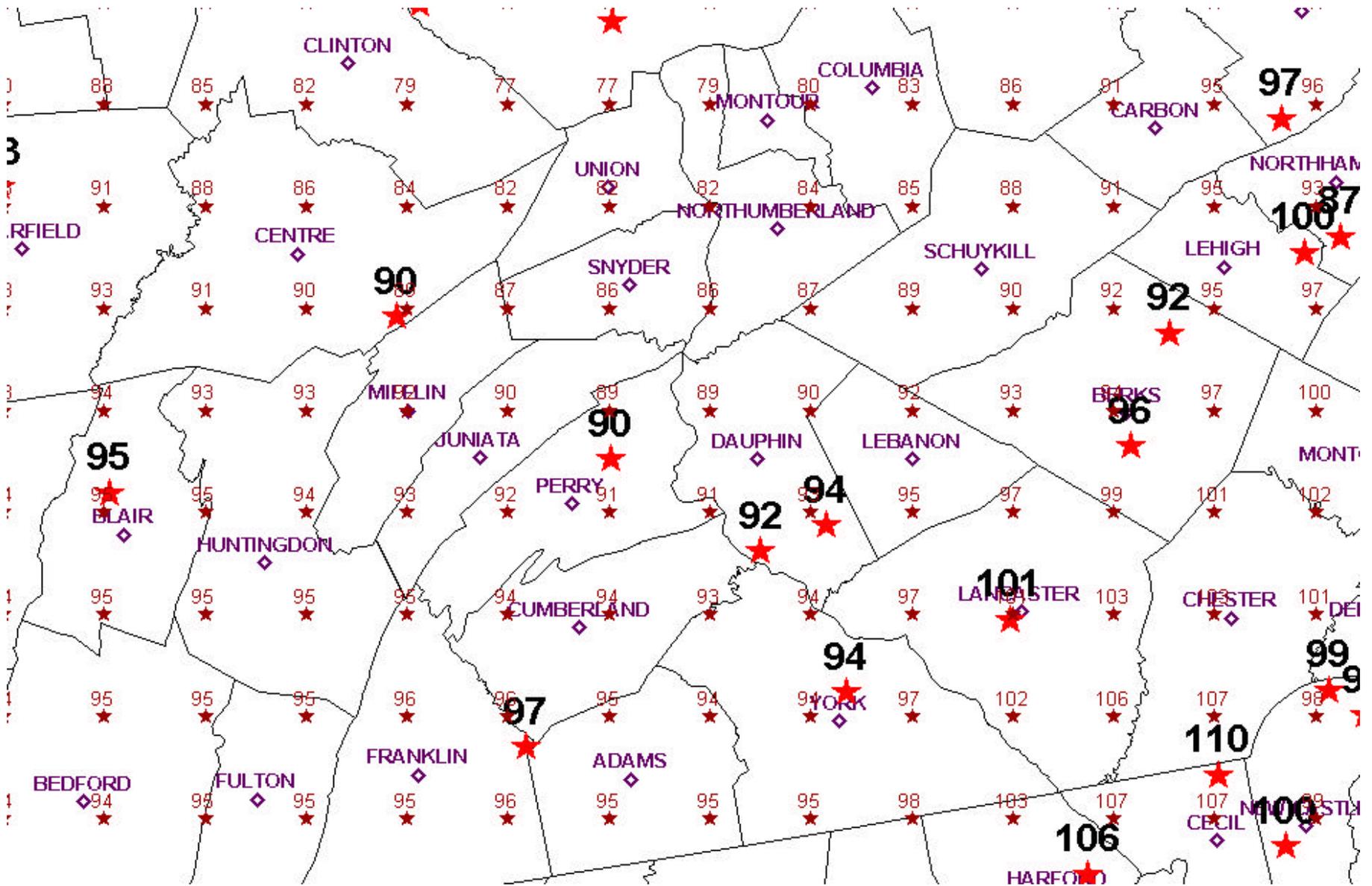
- A complete field of air quality is available for policy development, trends analysis, etc.
- Robust: Changes to an optimized network should not significantly affect the estimated field
- Removes monitoring disincentive
- Provides a direct blueprint for developing cost-effective networks



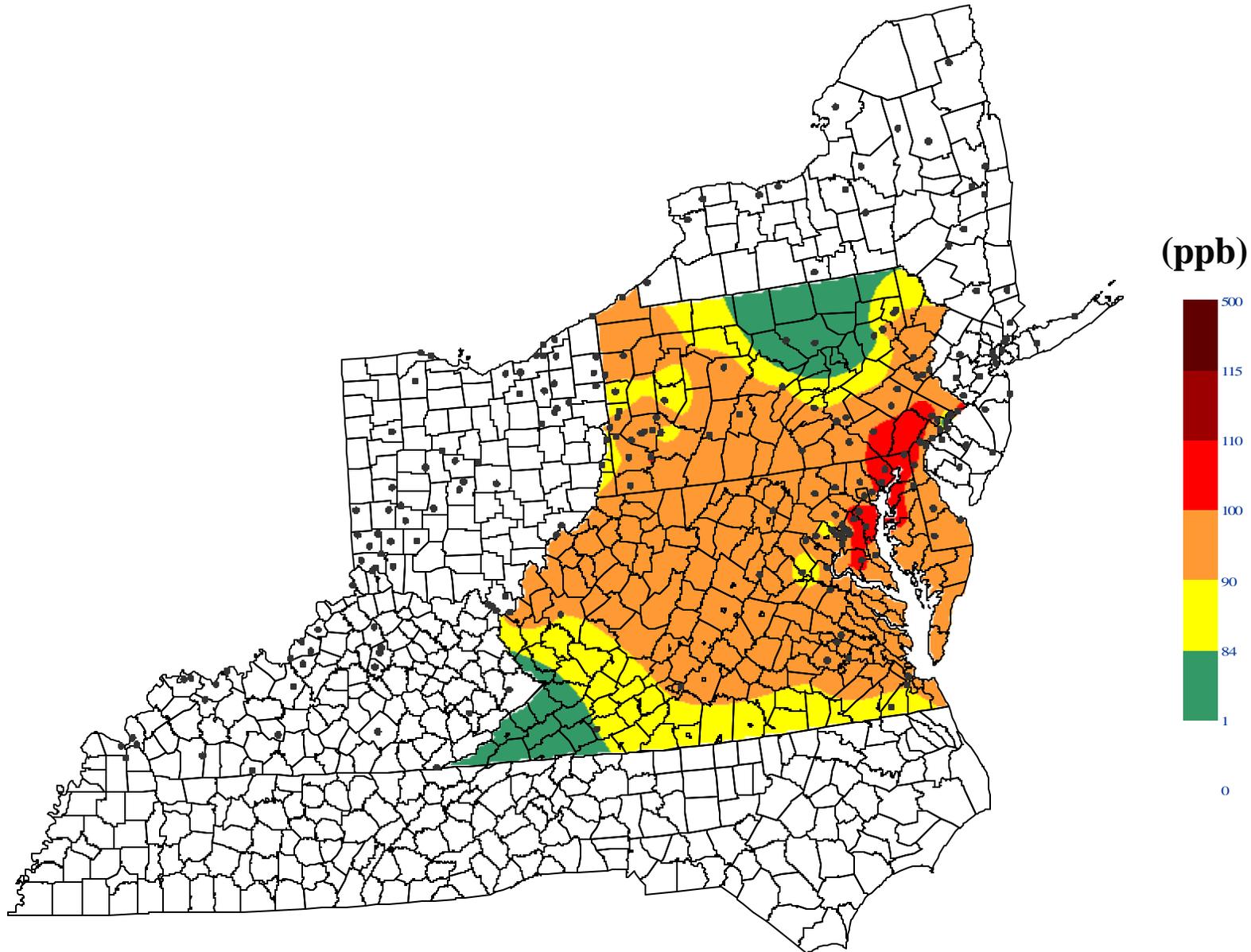
# Constructing Concentration Fields

- Establish the best variogram model for the concentration field base on:
  - Air quality data
  - BENCHMARK concentration field from area modeling (modeling data must adequately characterize important features of the field)
- Estimate, through kriging, the actual concentration field using:
  - The optimized variogram model
  - All available monitored air quality values both within and outside the area

# 1999 8hr. Ozone Design Value: Kriged Grid (Linear Variogram) with Network Overlay



# 1999 8 hr. Ozone Design Values: Kriged Contour Map (Linear Variogram)

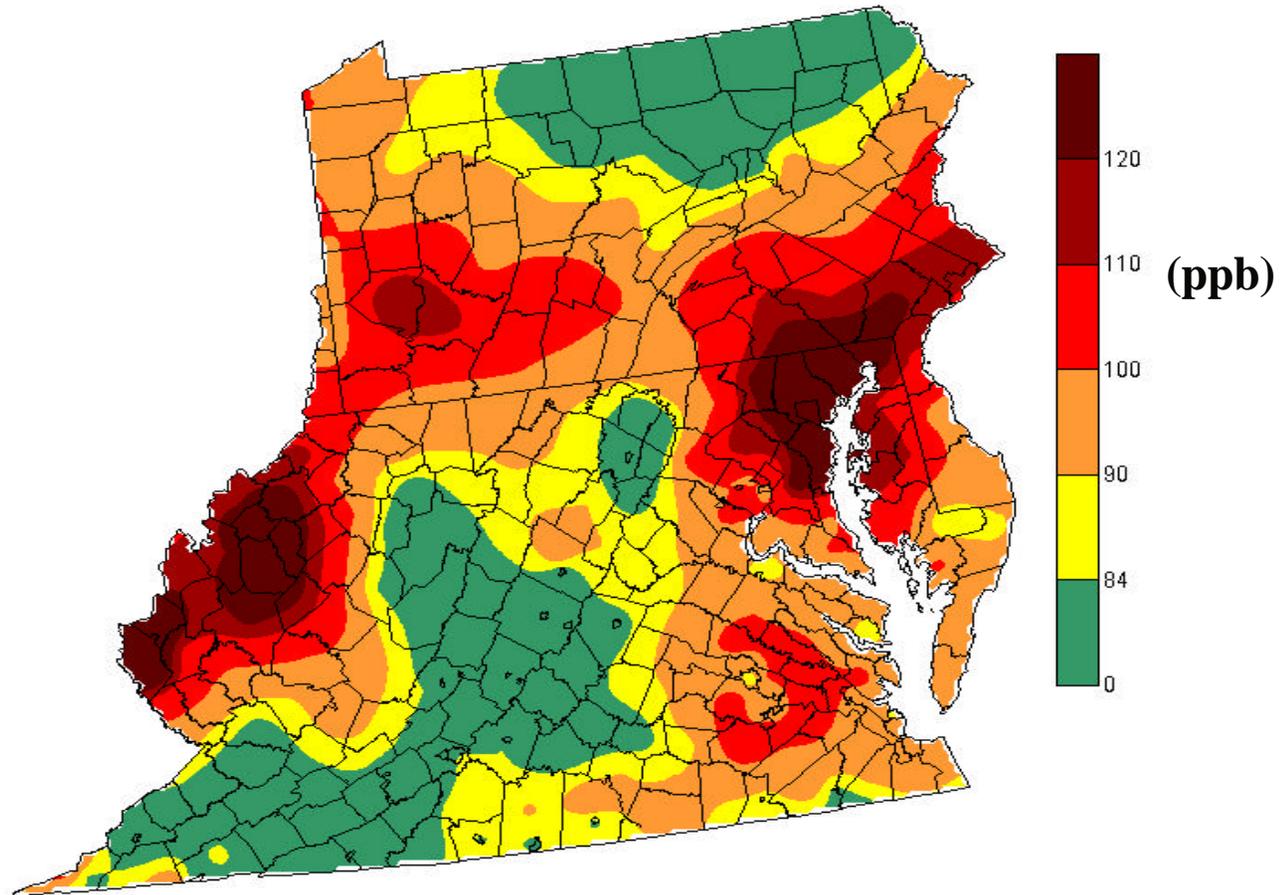




# Constructing Uncertainty Fields

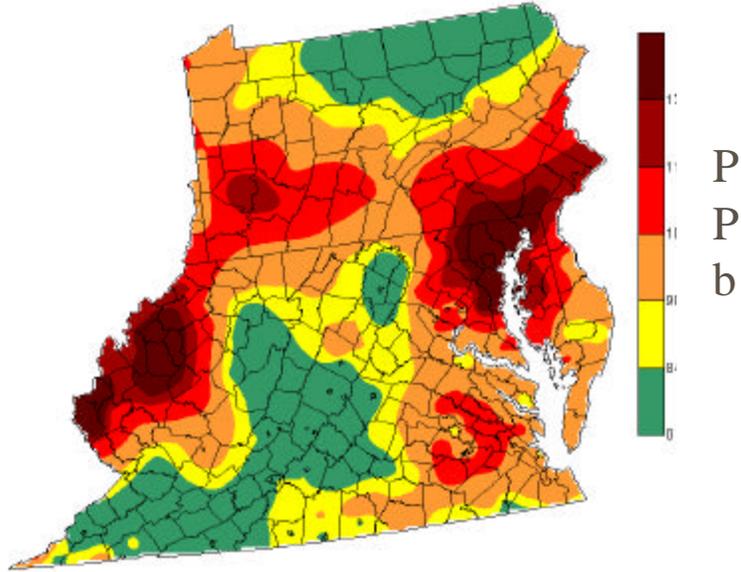
- Develop a subset of the benchmark (photochemical dispersion modeled) data based on network monitor locations only
- Estimate the full benchmark concentration field by kriging the benchmark data subset
- Compare the full benchmark field with the estimated field from the benchmark subset
- Construct a field of residuals (the uncertainty field)

**BENCHMARK Data Set**  
**4<sup>th</sup> High 8hr. Ozone: UAM-V Model Output**  
**1996 Emissions Inventory**  
**30 Days of 1995 Met**

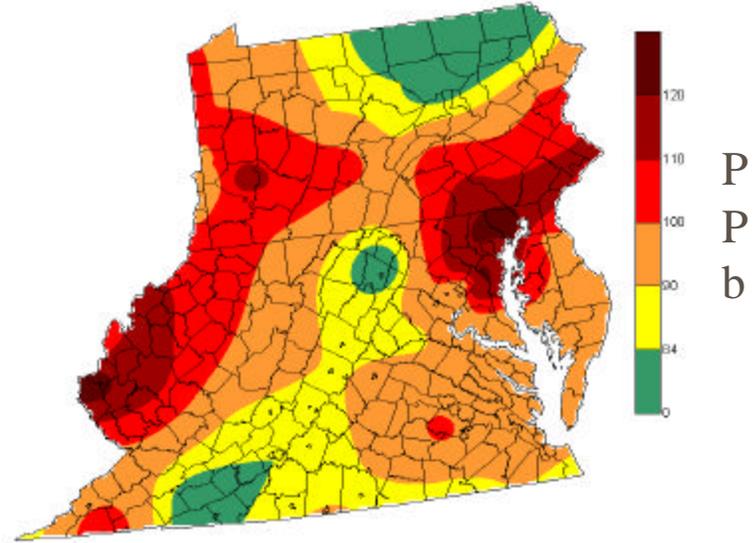




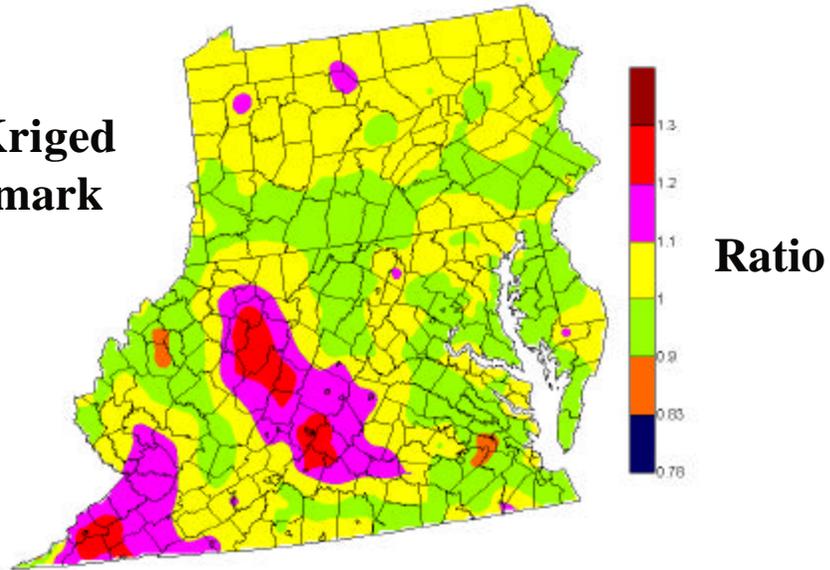
**Benchmark Data Set**



**Kriged Data Sub-Set  
(1999 O<sub>3</sub> Network)**

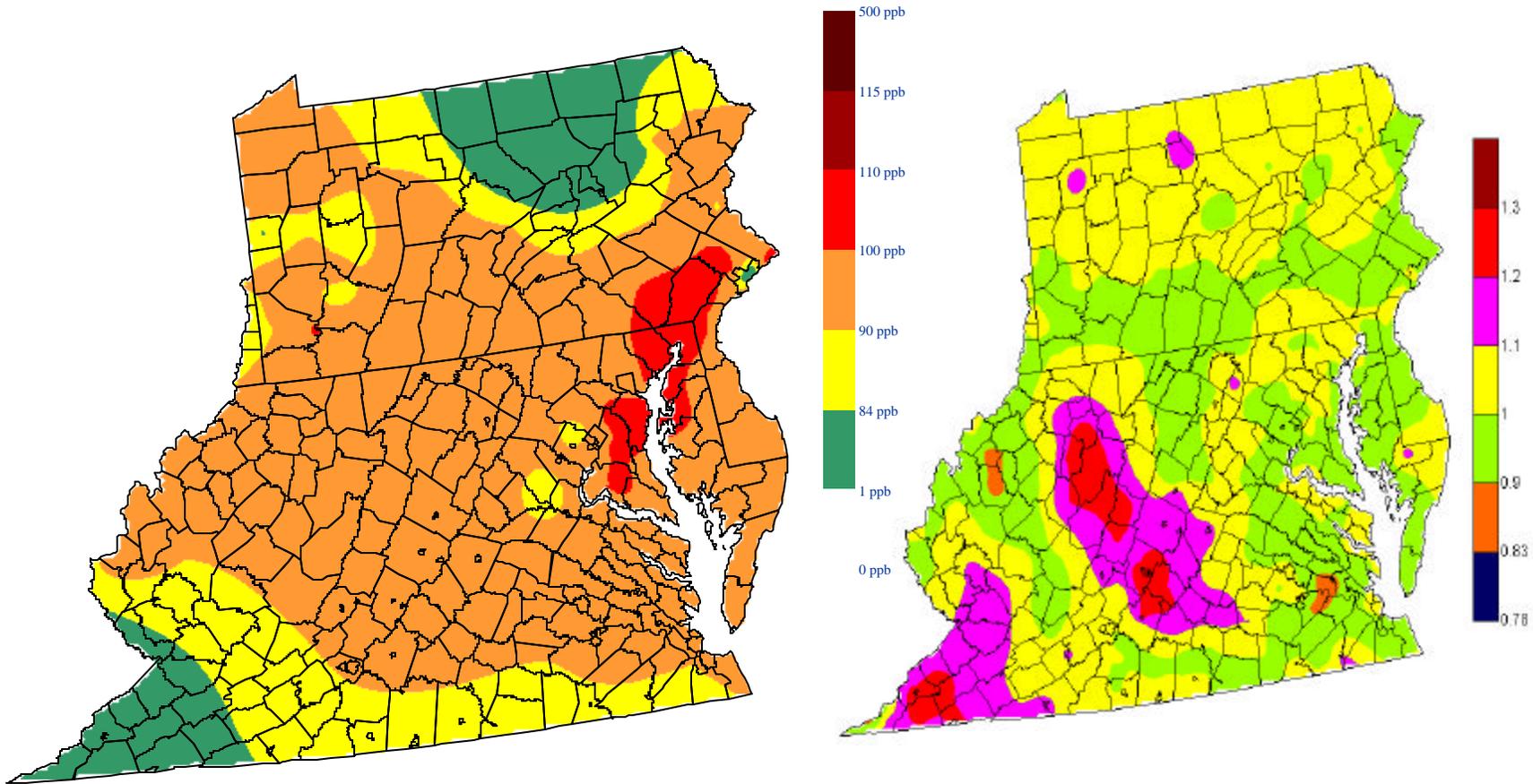


**Ratio of Kriged  
To Benchmark**



**Air Quality is defined as an interpolated concentration field plus an associated uncertainty field**

1999 8hr. DV O<sub>3</sub>





# Network Design: Premise

The primary purpose of a monitoring network is to estimate a variety of pollutant fields with the least uncertainty considering resource demands and other appropriate design criteria



# Network Design: Approach

- Develop appropriate benchmark data sets (modeled)
  - Multiple episodes
  - Multiple metrics
  - Multiple pollutants
- Construct potential new network designs (i.e., data subsets from the benchmark field)
- Estimate (krig) concentration fields for each network design
- Develop decision criteria (e.g.):
  - Compare each estimated field to the benchmark fields using statistics such as: Correlation Coefficient; Maximum residual; etc.
  - Multiple areas
  - Resource demand
  - State preference
  - Community preference
  - Other needs: e.g., statutory, trends, PSI, etc.
- Use MIRA to choose design

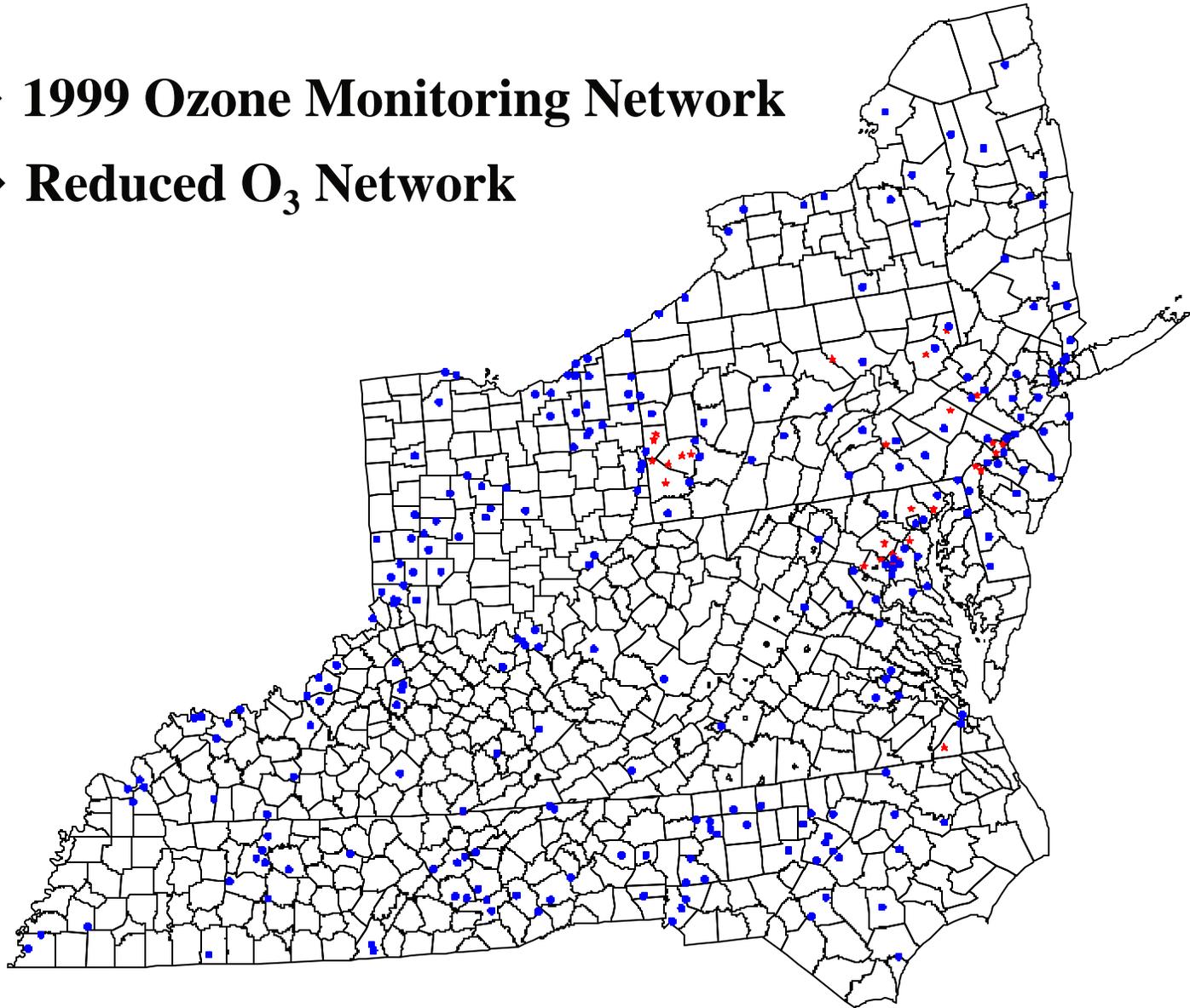


# Network Design: Example

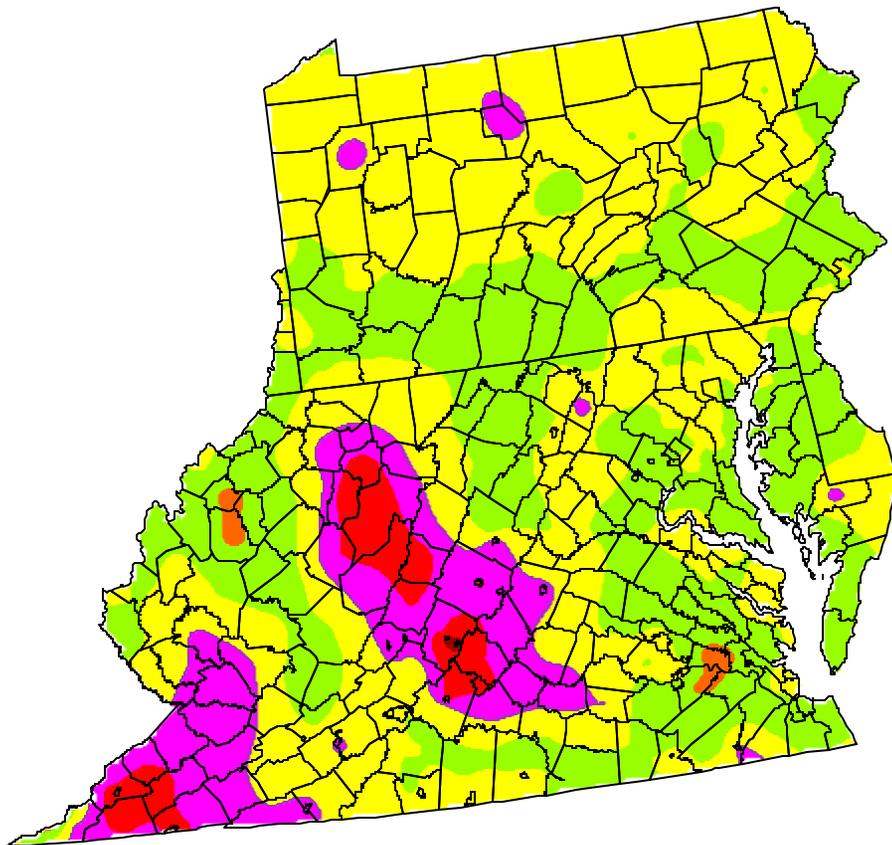
- Consider 4 networks:
  - 1999 O<sub>3</sub> network
  - Reduced network (28 Rg III monitors removed): For counties having multiple monitors, remove all but the one having the max 1999 8 hr. design value
  - Network A: Add 1 monitor to Albemarle, VA
  - Network AH: Add a monitor each to Albemarle, VA & Harrison, WV
- Consider 3 areas:
  - Region 3
  - Region 3's Non-Attainment counties
  - Region 3's Attainment counties
- Consider 1 decision criteria: correlation coefficient

★ + ● → 1999 Ozone Monitoring Network

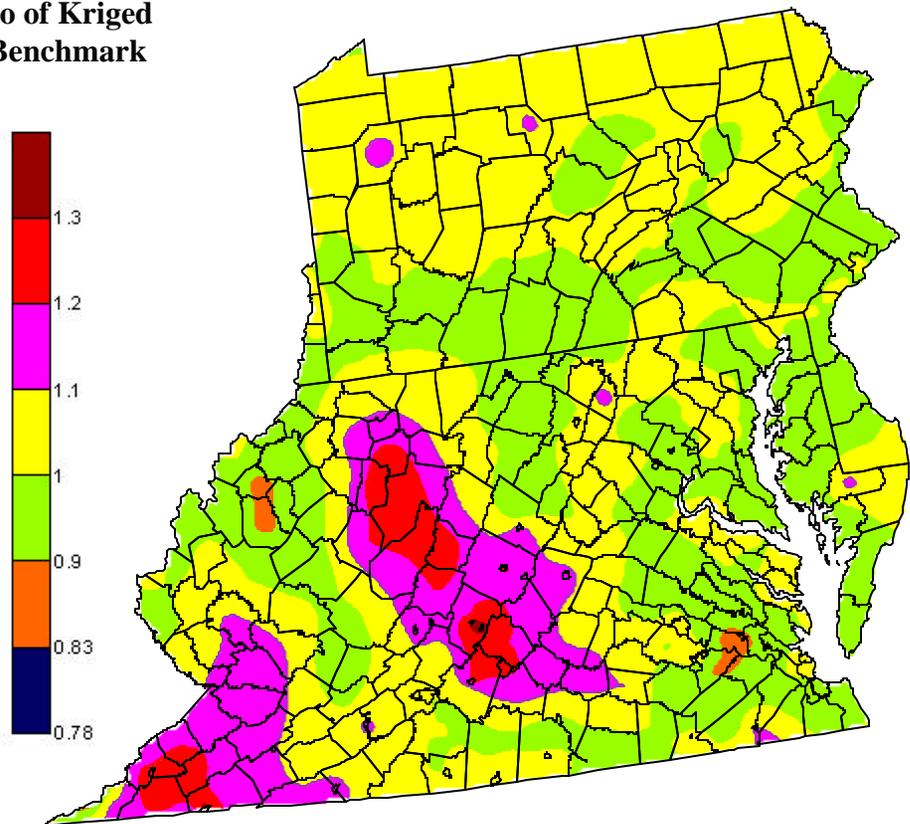
● → Reduced O<sub>3</sub> Network



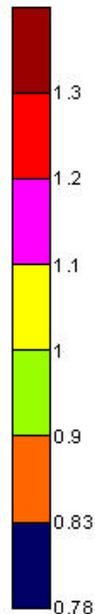
**1999 O<sub>3</sub> Network**  
(Corr Coeff = 0.89)



**Reduced Network**  
(Corr Coeff = 0.88)

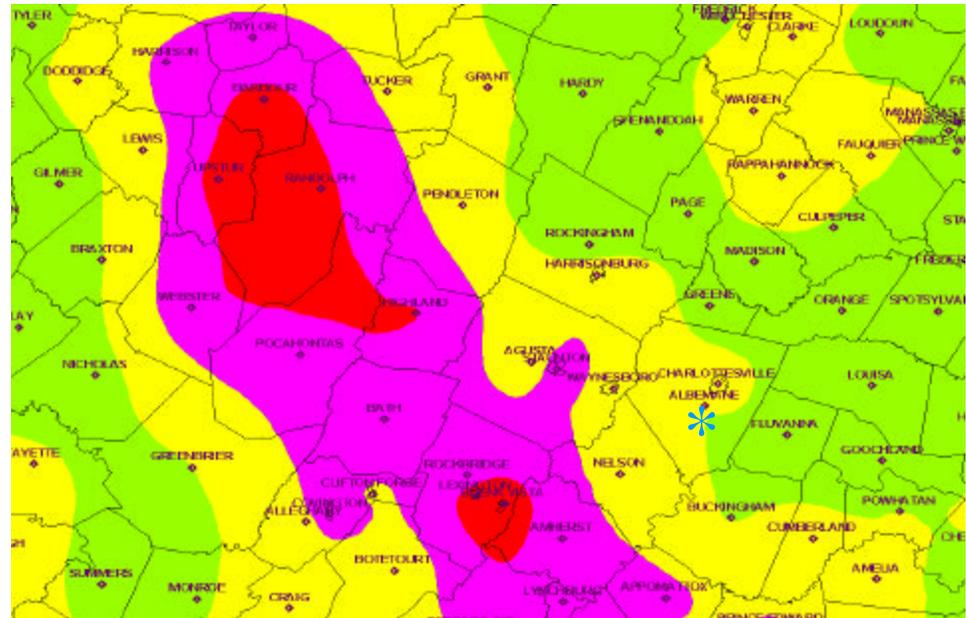
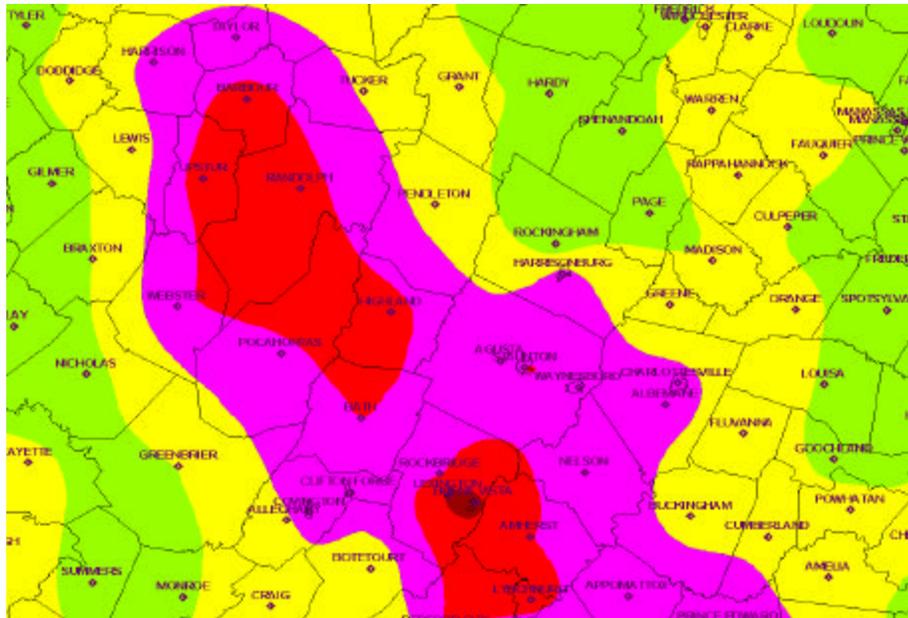


Ratio of Kriged  
To Benchmark

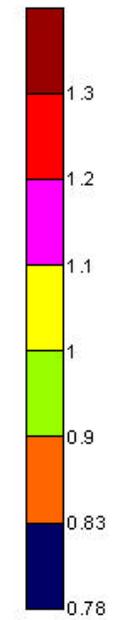
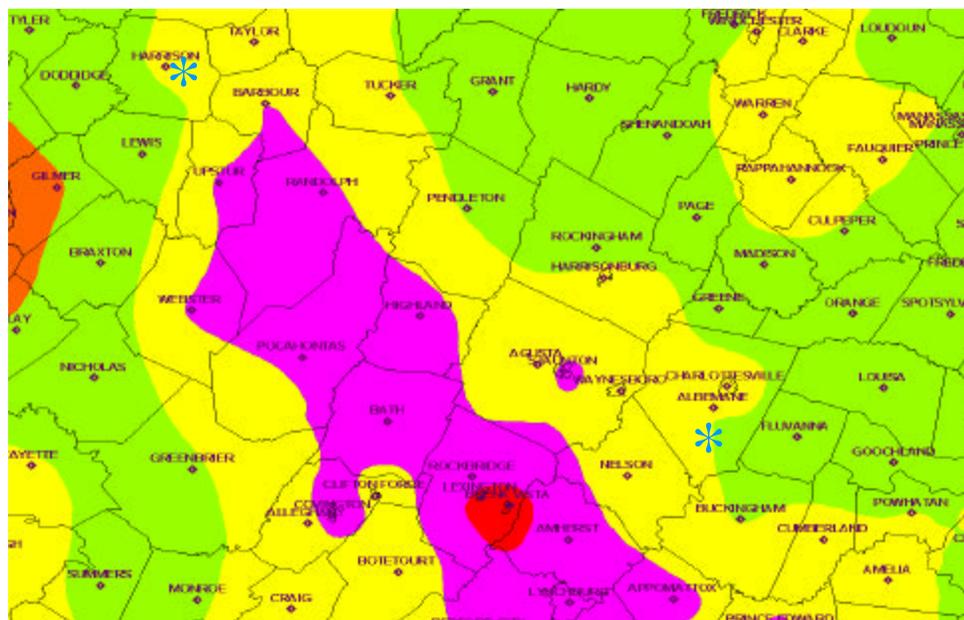


**1999 O<sub>3</sub> Network (Corr Coeff = 0.89)**

**Network A (Corr Coeff = 0.90)**



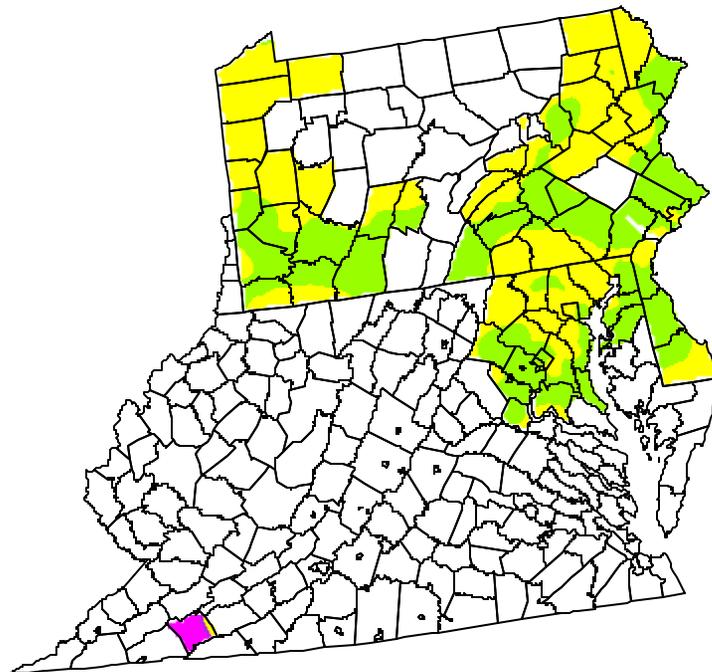
**Network AH  
(Corr Coeff = 0.91)**



**Ratio of Kriged  
To Benchmark**

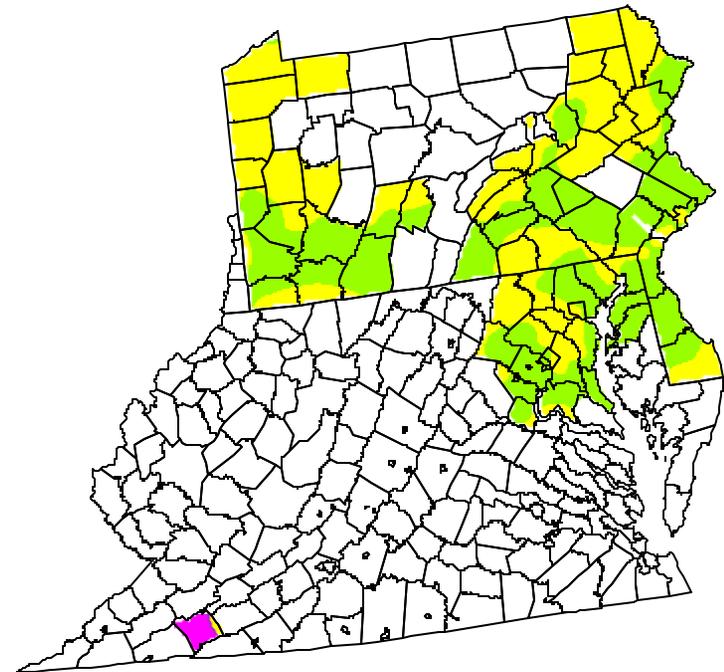
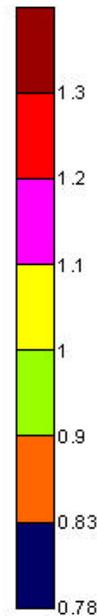
# 1999 O<sub>3</sub> Network 1 hr. Non-Attn. Areas

(Corr Coeff = 0.97)



# Reduced Network 1 hr. Non-Attn. Areas (Corr Coeff = 0.97)

Ratio of Kriged  
To Benchmark



# 1999 O<sub>3</sub> Network

## 1 hr. Attn. Areas

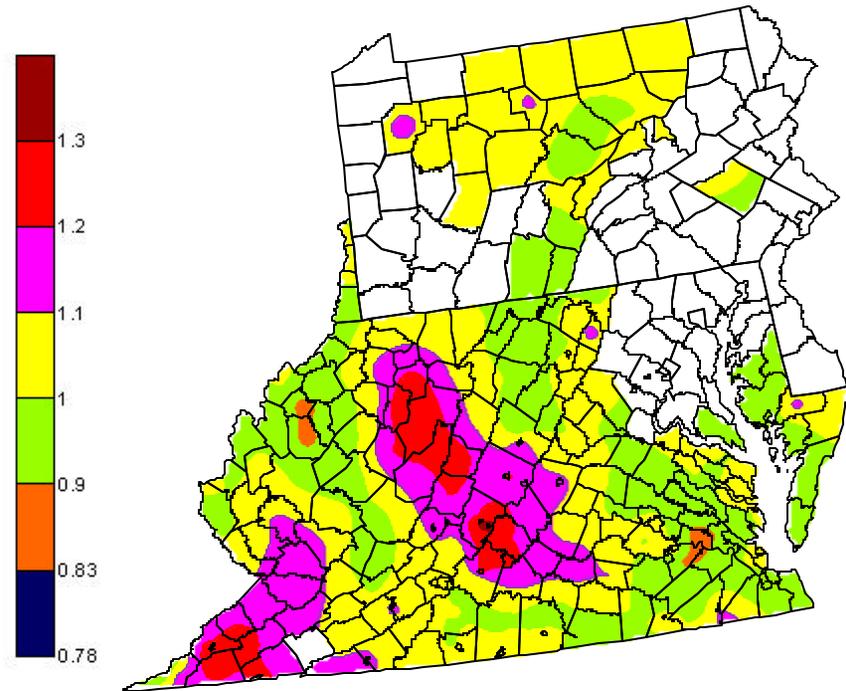
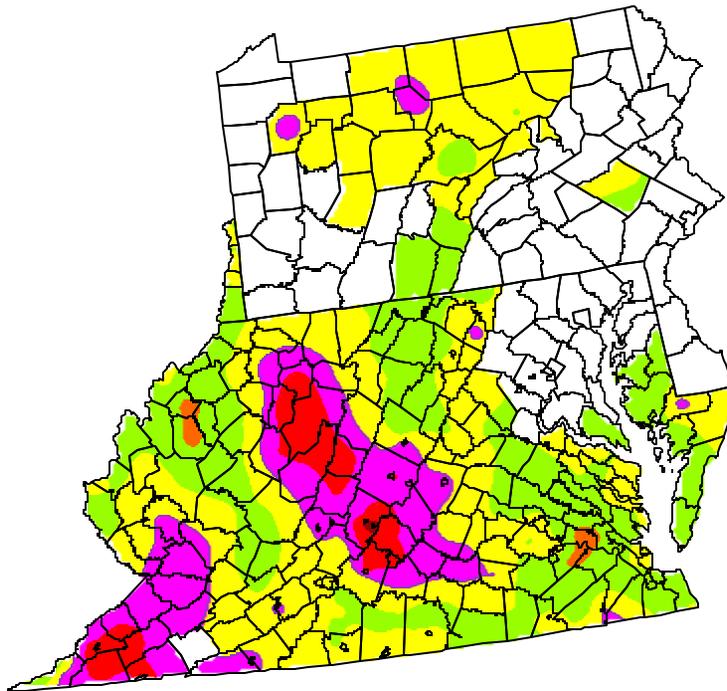
(Corr Coeff = 0.82)

# Reduced Network

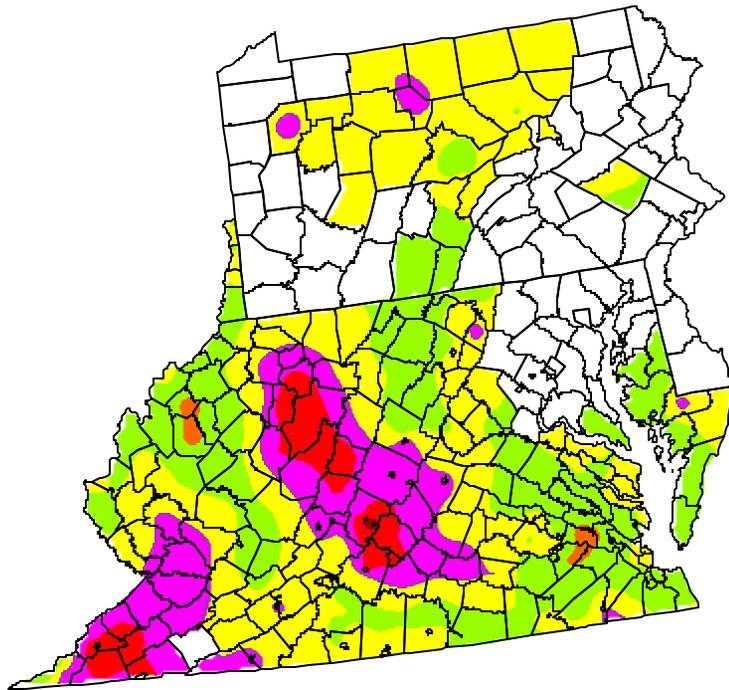
## 1 hr. Attn. Areas

(Corr Coeff = 0.82)

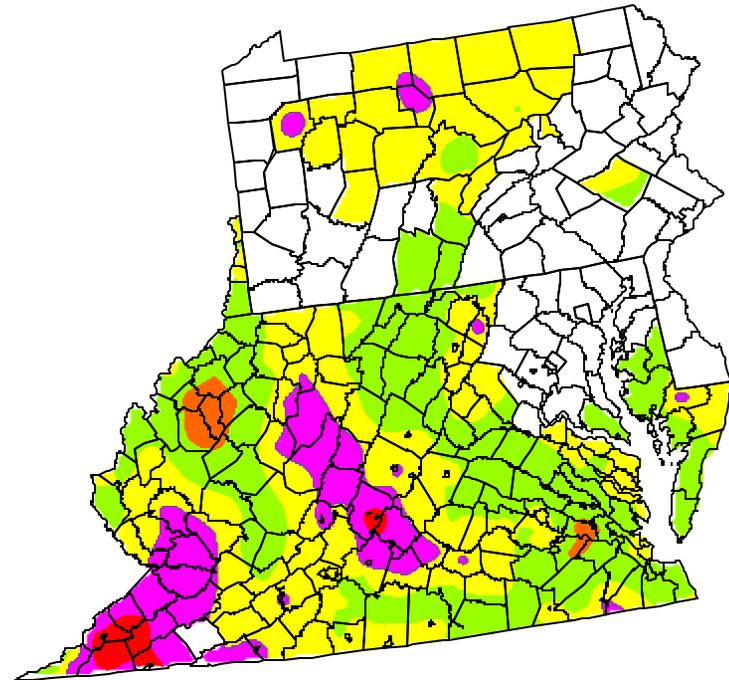
Ratio of Kriged  
To Benchmark



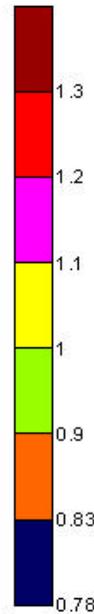
# 1999 O<sub>3</sub> Network 1 hr. Attn. Areas (Corr Coeff = 0.82)



# Network AH 1 hr. Attn. Areas (Corr Coeff = 0.86)



Ratio of Kriged  
To Benchmark





# Next Steps

- Optimized variogram model
- Proceed w/ network reassessment
  - Establish set of benchmark data sets:
    - Additional episodes
    - Additional metrics
  - Decide on comparison statistics
  - Establish other criteria
- Apply new concept to O<sub>3</sub> and PM<sub>2.5</sub> non-attainment designations – i.e., for counties wo/ monitors use interpolated data.
- Convince DECISION MAKERS to change their view