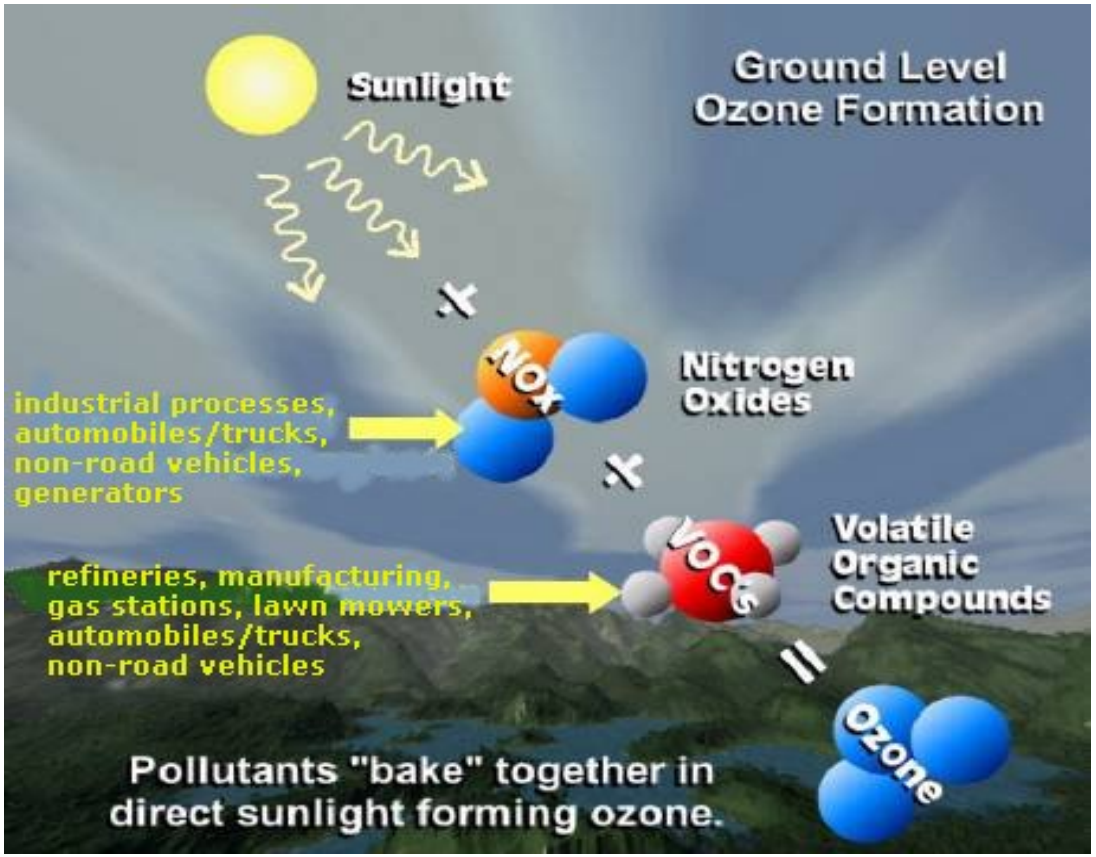
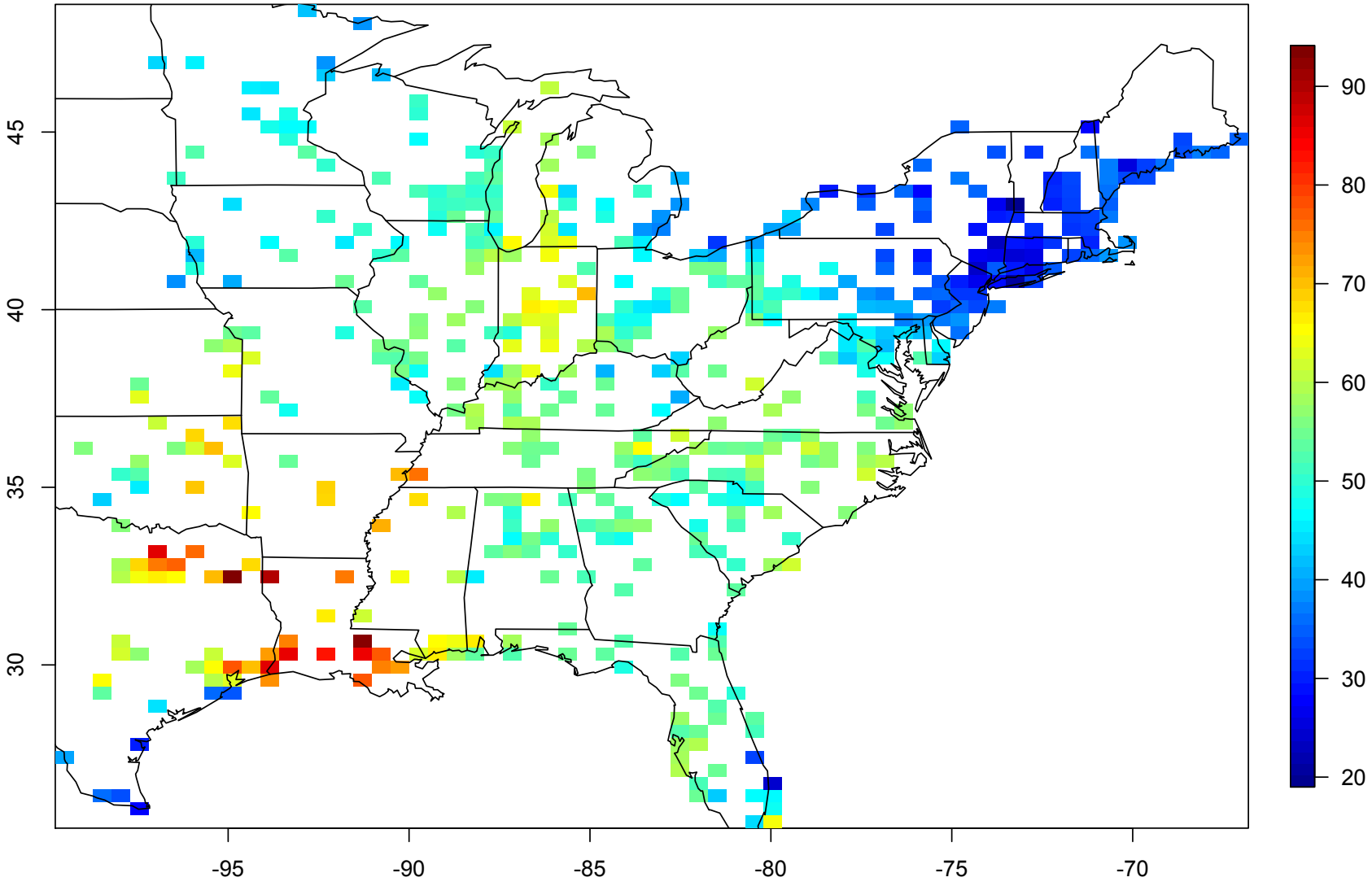
Midterm:
Ozone Prediction

Ground Level Ozone Analysis

Background on Ground-level Ozone (O3):



1. Main component of smog
2. Breathing high concentration of O3 triggers:
   - chest pain
   - bronchitis
   - emphysema
   - asthma
3. Monitored by EPA



Ozone Data:
801 station measurements of max.
8 hour O3 on May 22, 2005

```
library(maps)
library(LatticeKrig)
quilt.plot(Lon,Lat,o3)
map('state',add=TRUE)
```

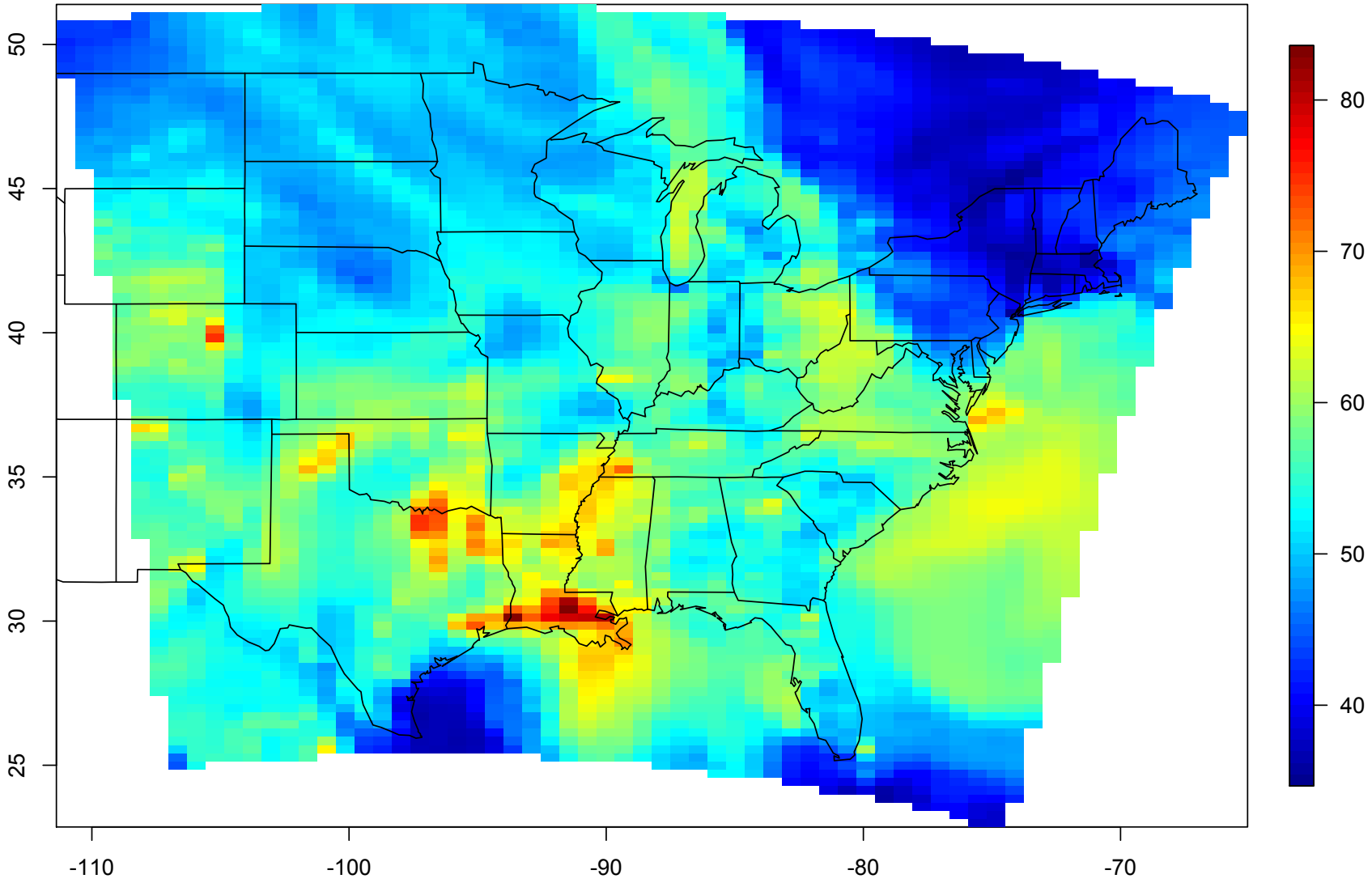
Ground Level Ozone Analysis

Community Multi-scale Air Quality Model (CMAQ):

- Mathematically simulates (on a fine spatial scale) formulation of ozone based on ground characteristics, temperatures, urban density, etc.

Issues with using CMAQ:

1. Doesn't reproduce observations exactly.
2. Returns an average O3 level, not at point level.



CMAQ Data:
66,960 Surrogate CMAQ O3 Measurements

```
quilt.plot(CLon,CLat,CMAQ)
map('state',add=TRUE)
```

Ground Level Ozone

Research Goals:

1. EPA scientists know that CMAQ is wrong:
   - Understand the relationship between CMAQ and Station Measurement.
2. Want ground-level O3 predictions at lots of locations
   - Prediction locations provided on website

Ground Level Ozone

Statistical Goals and Issues to consider:

1. Estimating relationship between ground level O3 and CMAQ:
   - Specifying a Model: CMAQ doesn't line up with station measurements. So, how are you going to define predictors? How many predictors are you going to use? Collinearity?
   - High Dimensions: You have 66960 potential "predictors" (the CMAQ values) that you could use to explain 801 observations.
   - No IID Errors: Station measurements are not independent but are spatially correlated.
   - Quantifying Relationship: How are you going to quantify the relationship between CMAQ and station measurements (HINT: intercepts are important here).
2. Predicting O3:
   - Nonlinearity: Ozone is dependent upon Lon/Lat but this relationship is highly non-linear.
   - Prediction Accuracy: How are you going to assess how accurate your predictions are?

Ground Level Ozone

Rules for the Midterm:

1. Written reports to be done individually – I want to see what you come up with.
2. I am available to answer questions related to HOW to do something not WHAT you should do.
3. Coding questions and "principle" questions are fine to ask me or your classmates.

Ground Level Ozone

A Few Expectations:

1. Justify your model: why did you do what you did? Does it answer the questions?
2. Justify your assumptions.
   - Justify linearity, normality, etc.
   - Justify assumptions you didn't make but might be a problem (e.g. collinearity, heteroskedasticity, normality). State why you're willing to overlook these issues.
3. I want to see your estimates/predictions.
4. I want to see uncertainties.
5. Interpretations in terms of the problem.