Spatial temporal analysis of health effects associated with sources and speciation of fine PM

Montse Fuentes

Professor of Statistics NCSU

STAR award workshop, EPA, RTP. June 10, 2009.

This STAR/coop award team

- M. Fuentes (PI) – Statistics, NCSU
- C. Frey (co-PI)– Env. and Civil Eng. NCSU
- Yang Zhang (co-PI) – Atmospheric Sciences, NCSU
- M. Bell (co-PI) – Environmental Health, Yale University
- F. Dominici (consultant)- Biostatistics, Johns Hopkins
- EPA collaborators: Hubell, Ozkaynak, Neas, Norris, Luben, Holland, Burke.

This presentation has been prepared jointly by the team.

Objectives

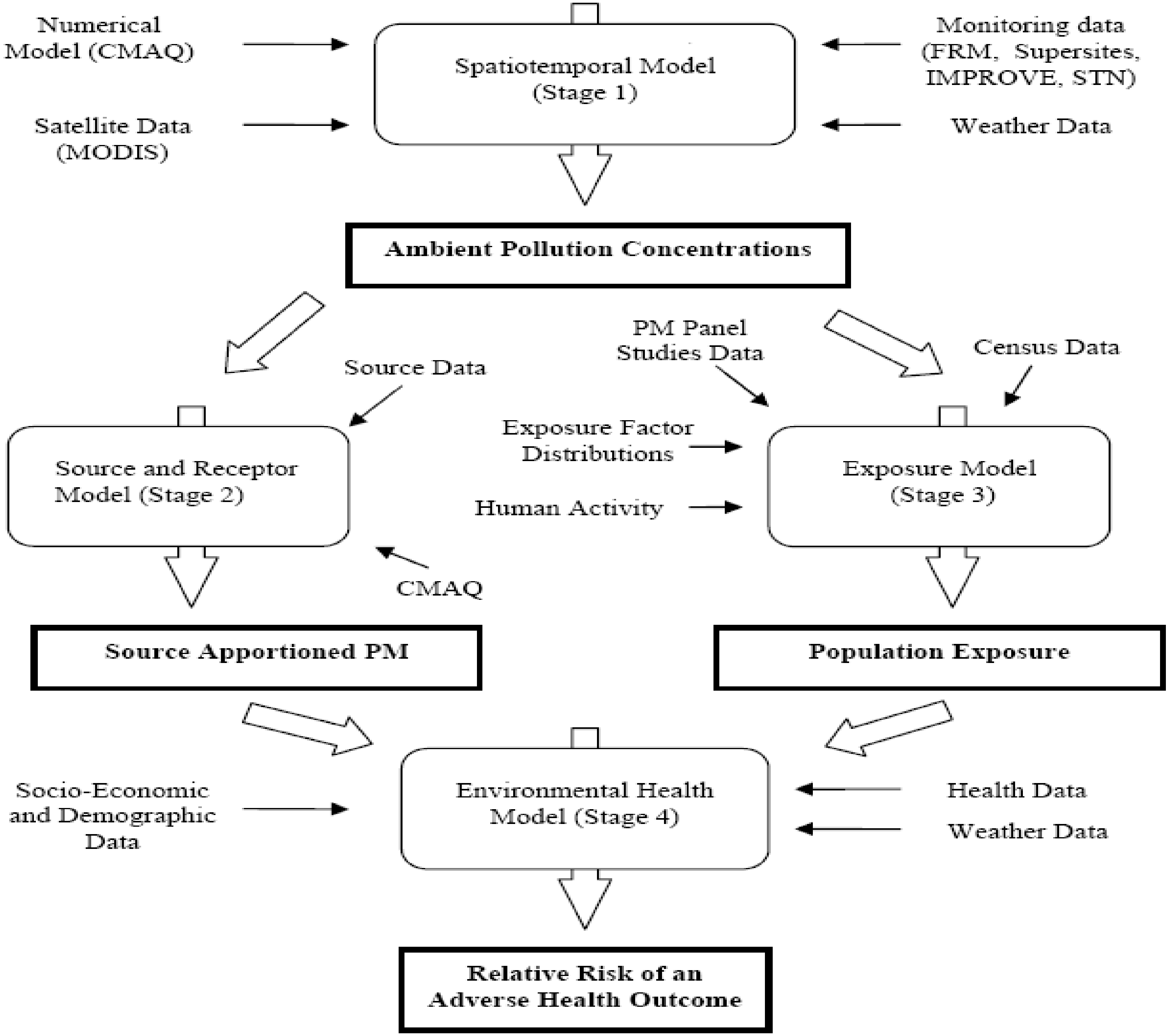
- Investigate the adverse health outcomes associated with population exposure to fine particulate matter (*PM2.5) and speciation, characterizing geographic differences, sources, and population heterogeneity.*

We aim to answer the following research questions:

- Can we improve the PM component-based epidemiologic studies by using atmospheric models (**CMAQ**) and exposure models (**SHEDS**)?
- How to use source apportionment approaches in national epidemiologic studies, while characterizing different sources of **uncertainty** in the models and the data?

Approach

We develop a statistical hierarchical Bayesian framework (with 3 stages) that provides a very broad, flexible approach to studying the spatiotemporal associations between adverse health outcomes and population exposure to daily PM2.5 mass and its components, while characterizing its potential sources.


Numerical
Model (CMAQ)
Spatiotemporal Model
(Stage 1)
Monitoring data
(FRM, Supersites,
IMPROVE, STN)
Satellite Data
(MODIS)
Weather Data
Ambient Pollution Concentrations
PM Panel
Studies Data
Census Data
Source Data
Exposure Factor
Distributions
Source and Receptor
Model (Stage 2)
Exposure Model
(Stage 3)
Human Activity
CMAQ
Source Apportioned PM
Population Exposure
Socio-Economic
and Demographic
Data
Environmental Health
Model (Stage 4)
Health Data
Weather Data
Relative Risk of an
Adverse Health Outcome

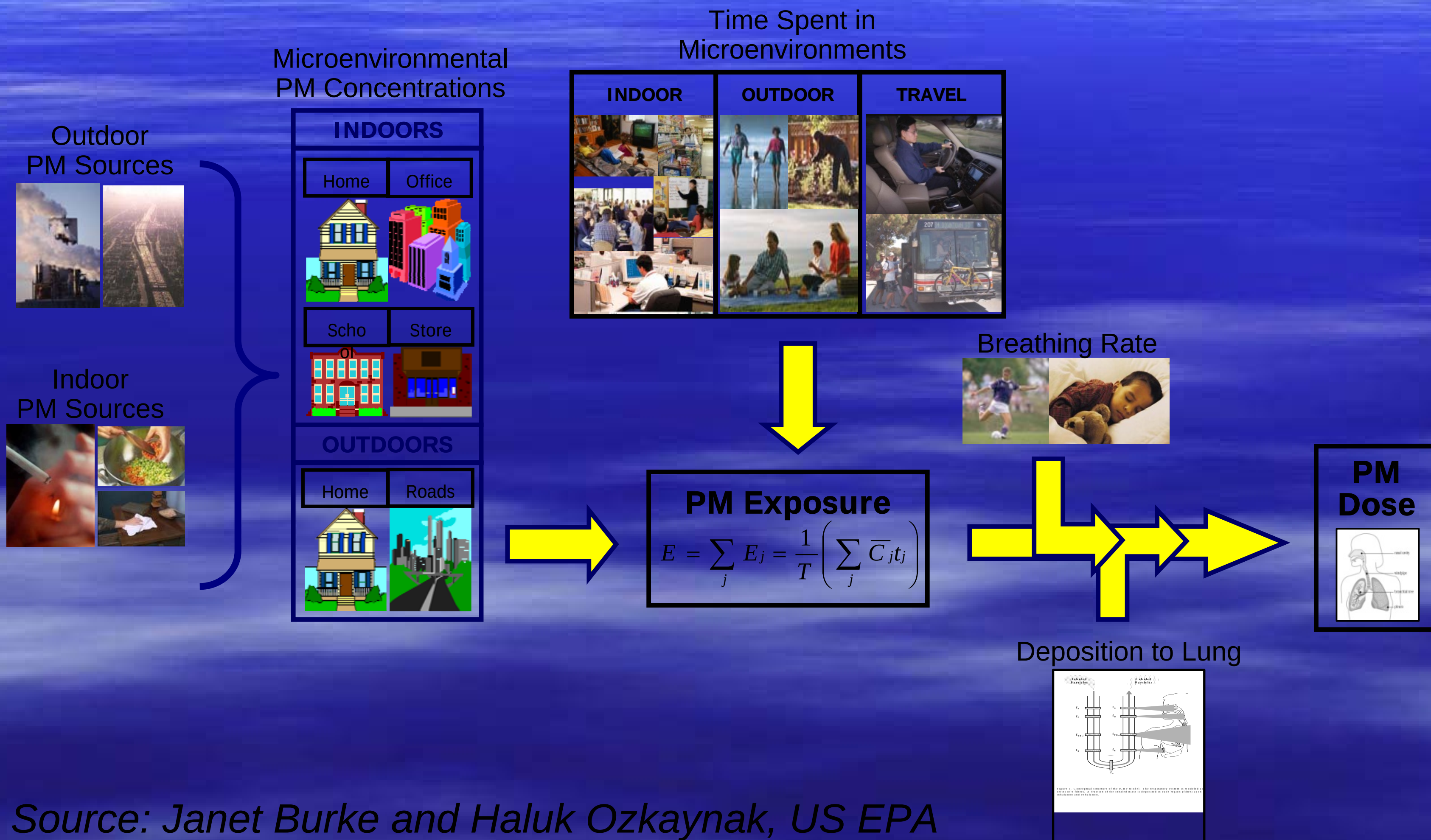
Expected benefits from this research

- Development of a new flexible spatiotemporal modeling framework for predicting fine PM mass and components, combining monitoring data with air quality numerical models (CMAQ), while accounting for different sources uncertainties.
- Improved characterization of the spatial temporal variation of PM sources by using atmospheric models, source and receptor spatial temporal analysis.
- Integration in the epidemiologic analysis of our results on PM composition and estimated sources.
- Better understanding of the changes in health effects estimates based on various methodologies for estimating exposure (e.g., monitoring, SHEDS, CMAQ).

Preliminary case studies

- Preliminary analysis with SHEDS to characterize human exposure:
  - **NYC and Houston case studies.**
- Source apportionment.
  - PMF approach:
    - **NYC case study.**
  - CMAQ Brute Force method:
    - **January and July 2002 case studies.**
- Human health data and analysis:
  - **Texas birth data case study.**

Stochastic Human Exposure and Dose Simulation (SHEDS) Model for Particulate Matter



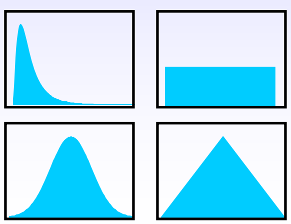
Source: Janet Burke and Haluk Ozkaynak, US EPA

SHEDS-PM Model Structure

Input Databases

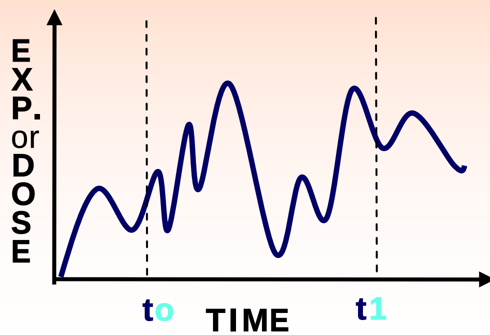
- US Census demographics
- Human activity patterns
- Air concentrations

Exposure Factor Distributions



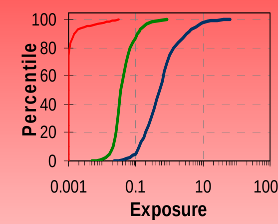
Algorithms

- Calculate Individual Exposure/Dose Profile

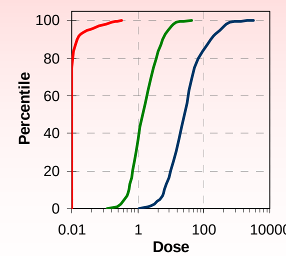


Output

- Population Exposure



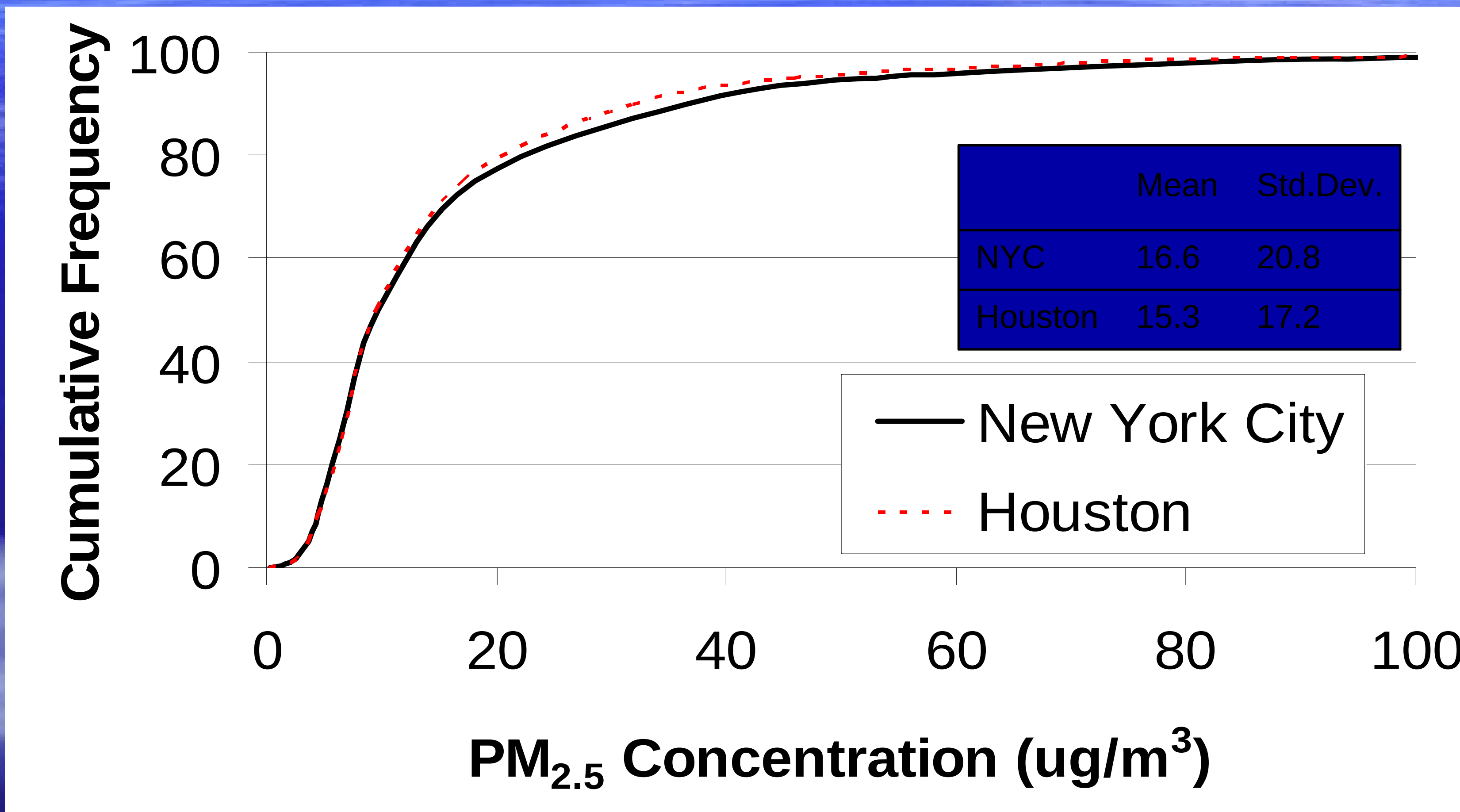
- Population Dose



- *Source: Janet Burke and Haluk Ozkaynak, US EPA*

- **Preliminary work with SHEDS:** we identify differences in estimated $PM_{2.5}$ exposure between New York City and Houston attributable to demographics and housing type for July 2002.
- **Methodology:**
  - Randomly chose 10 census tracts in each city.
  - Keep ambient $PM_{2.5}$ concentration constant for both New York City and Houston, to enable focus on differences in demographics and housing stock.
  - Use same microenvironmental algorithms and inputs for both cities.
  - Environmental Tobacco Smoke (ETS) is modeled in residential, restaurant, and bar microenvironments.
  - Compare distribution of inter-individual variability in daily average exposures using cumulative distribution functions.

Results: Inter-Individual Variability in Daily Average Total Exposure



- 99th percentile is 104 ug/m^3 for NYC, 89.0 ug/m^3 for Houston

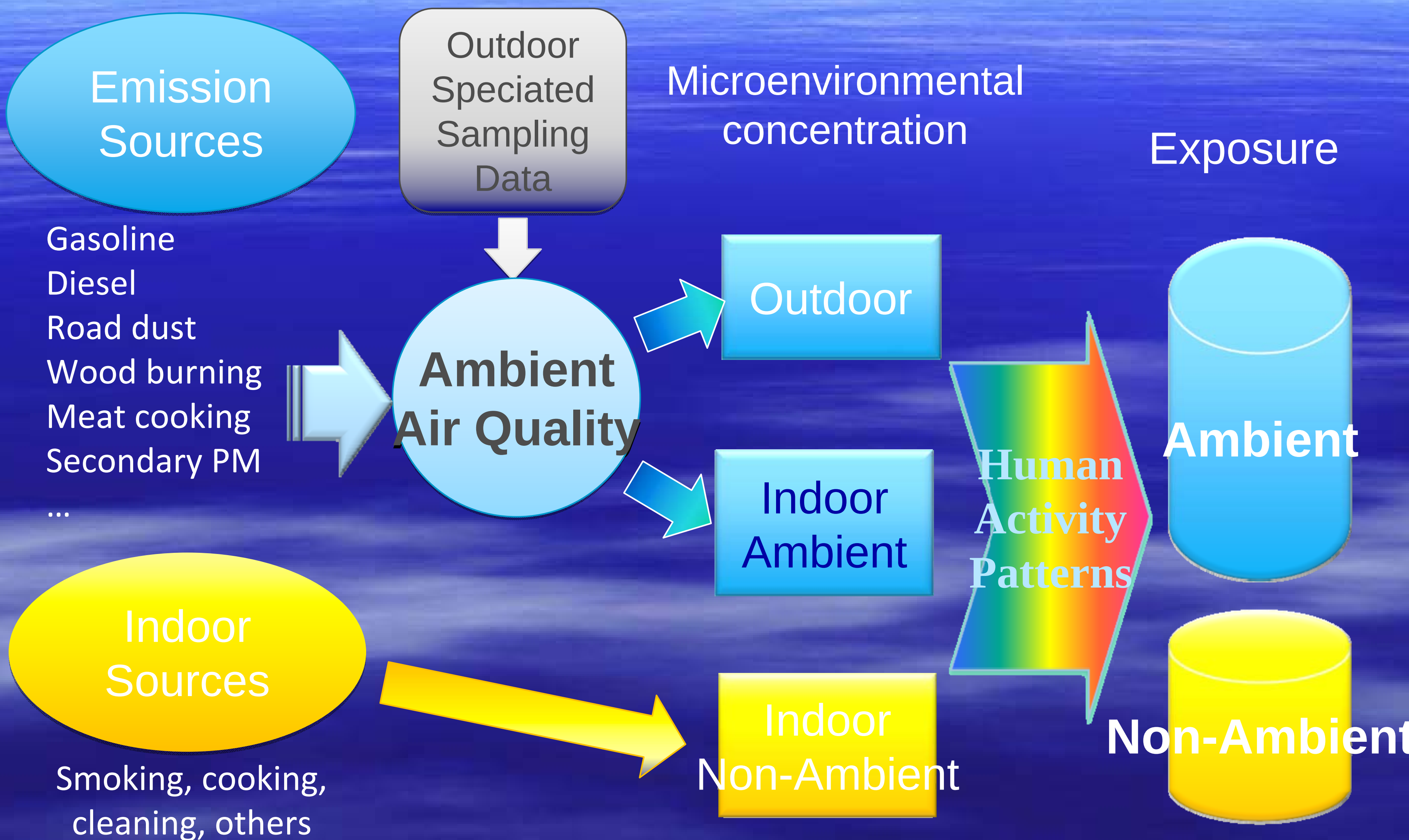
Conclusions from case study

- There is substantial inter-individual variability in exposure, even if ambient concentration is constant.
- Approximately 70 percent of exposure occurs at home, and approximately 94 percent occurs in indoor microenvironments.
- Houston has a somewhat younger, more male population, with a higher proportion of detached houses
- Differences in distributions of gender, age, and housing type are associated with slightly higher residential indoor exposure in New York City and slightly higher exposures in other microenvironments in Houston
- Activity data from CHAD are not specific to each city; hence, differences in transportation infrastructure and commuting patterns are not accounted for.
- In ongoing work, the implications of spatial and temporal variability in ambient $PM_{2.5}$ concentration specific to each area will be assessed.

Daily Average Exposures in NYC and Houston

	Ambient Exposure (ug/m^3)		Non-Ambient Exposure (ug/m^3)		Total Exposure (ug/m^3)	
	NYC	Houston	NYC	Houston	NYC	Houston
5th percentile	2.40	2.52	0.00	0.00	3.61	3.57
Median	5.44	5.57	3.41	3.30	9.59	9.57
Mean	5.49	5.59	10.9	10.0	16.4	15.6
95th percentile	8.68	8.67	47.1	43.1	52.6	48.9
Exposure-to-Ambient Concentration	0.55	0.56	1.09	1.00	1.64	1.56

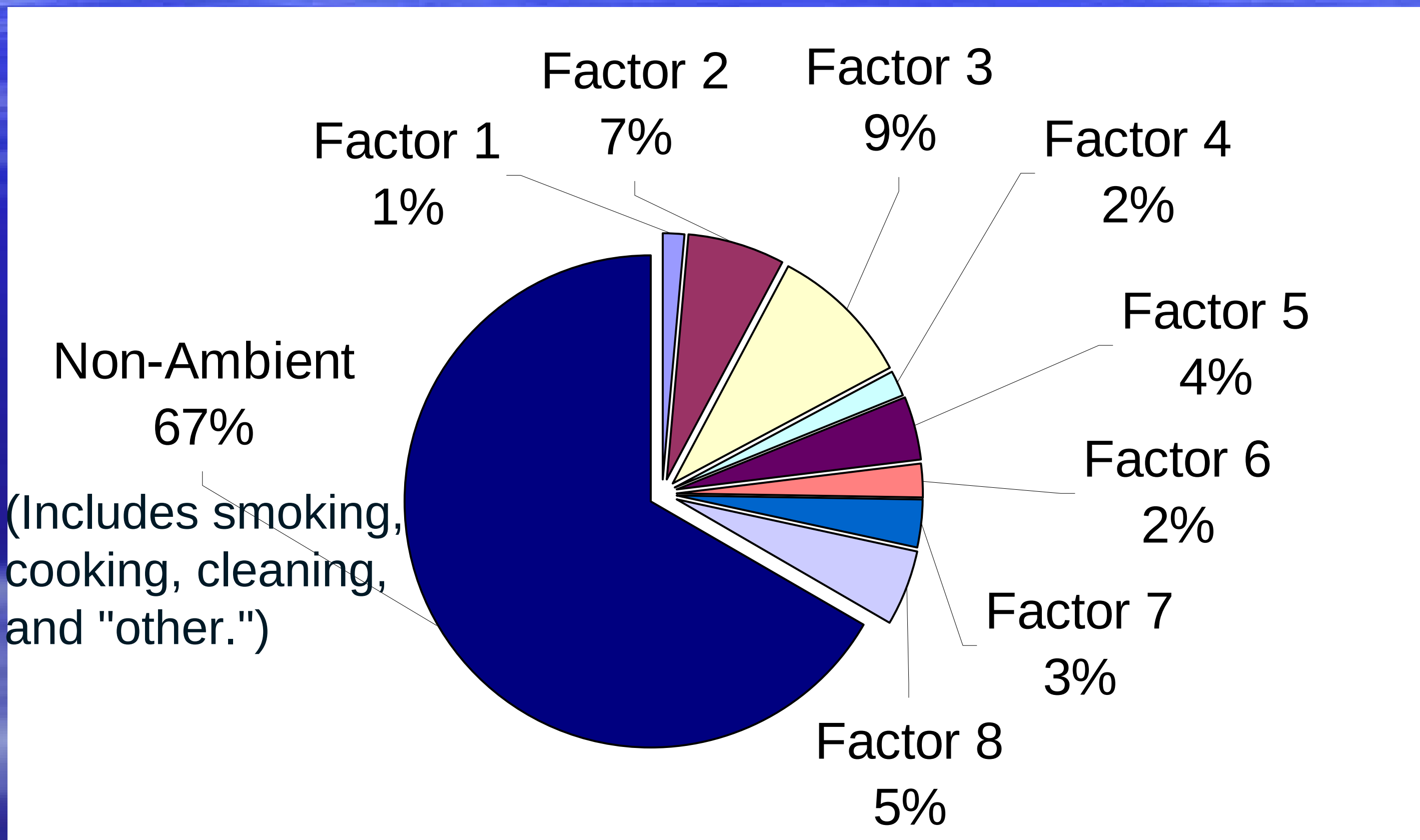
$PM_{2.5}$ Source Apportionment Framework



PMF Case Study: NYC

- Monitoring Station for PM2.5
  - Location: New York Botanical Garden, Bronx
  - ID: 360050083
- Study Period:Year of 2002
  - Dates excluded because of missing data: 01/01-02/01/2002, 09/26/2002
  - Date excluded becasue of Canadian Wildfires: 07/07/2002
- Species Categories
  - Strong: Br, Ca, Cu, Cl, Fe, Mn, Ni, Ti, V, Si, Zn, K, Na^+, NH_4^+, S, OC, NO_3^-, EC
  - Weak (down-weighted by factor of 3): As, Al, Ba, Cr, Pb, Mg, Se, Sn,Ta

Example Exposure Source Apportionment



Factor Contributions to $PM_{2.5}$ Daily Exposure (µg/m^3, percentage)

Conclusions from NYC case study

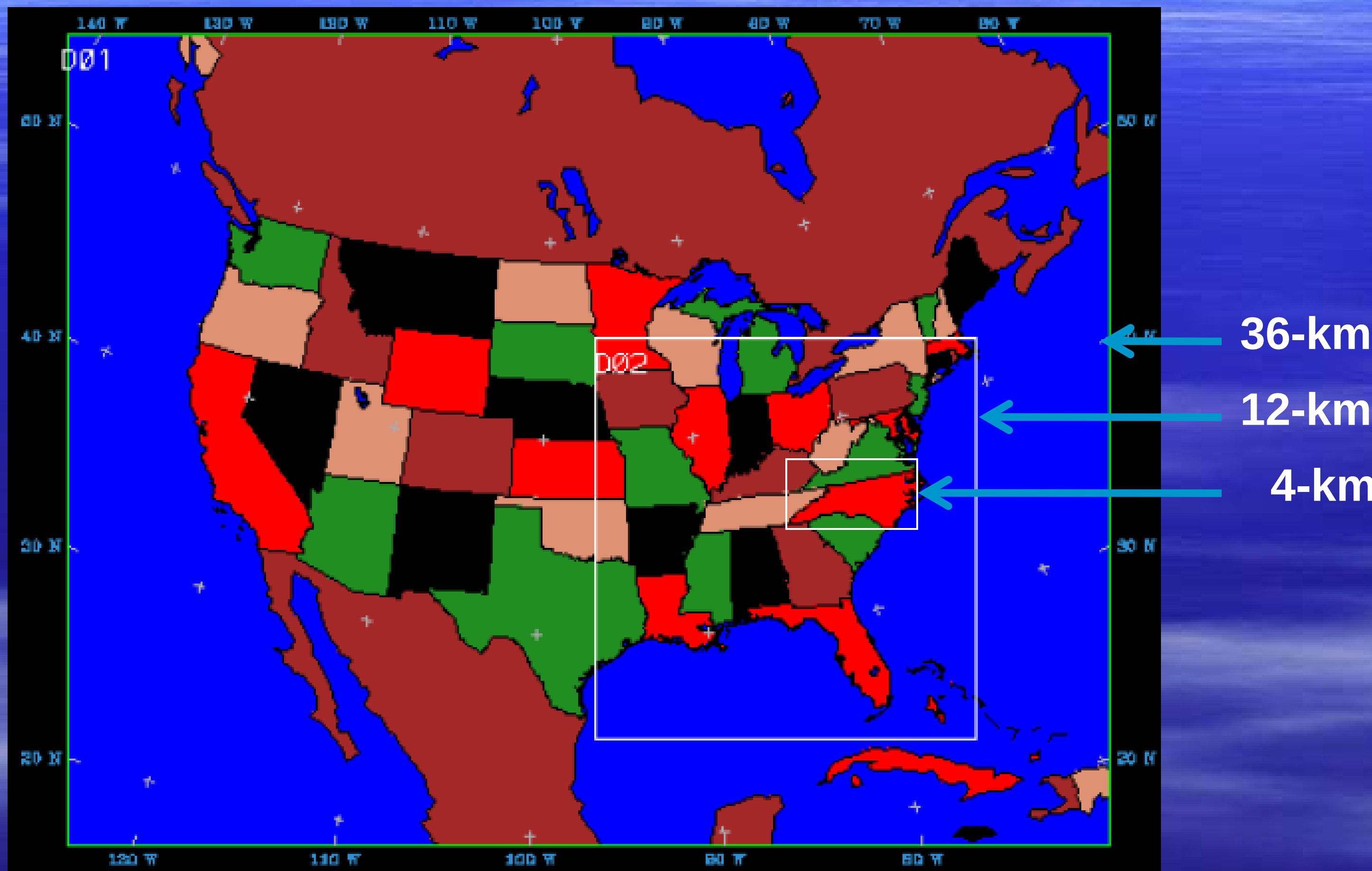
- Total exposure can be apportioned to sources by combining ambient exposure source apportionment and non-ambient exposure source apportionment
- Some of the factors are associated with vehicle emissions, oil combustion and secondary sulfates.
- Non-ambient sources are significant contributors to total exposure.

Source Sensitivity Study using CMAQ

- Technical Approach
  - Identify major source categories based on raw emissions inventory
  - Use CMAQ with the Brute Force method (zero-out one source category at a time)
- Major Source Categories Examined
  - Coal Combustion (CC)
    - Includes electric generating, industrial, and commercial external combustion boilers
    - Includes electric utility, industrial, commercial, and residential stationary source coal combustion
  - Diesel Vehicles (DV)
    - Includes on-road and off-road diesel powered vehicles
    - Excludes emissions from marine vessels and aircrafts
  - Gasoline Vehicles (GV)
    - Includes on-road and off-road gasoline powered vehicles
    - Excludes emissions from marine vessels and aircrafts
  - Biomass Burning (BIO)
    - Includes wildfires, prescribed burning, agricultural burning, residential wood burning, open burning at landfills, and external combustion boilers

CMAQ Simulation Domains

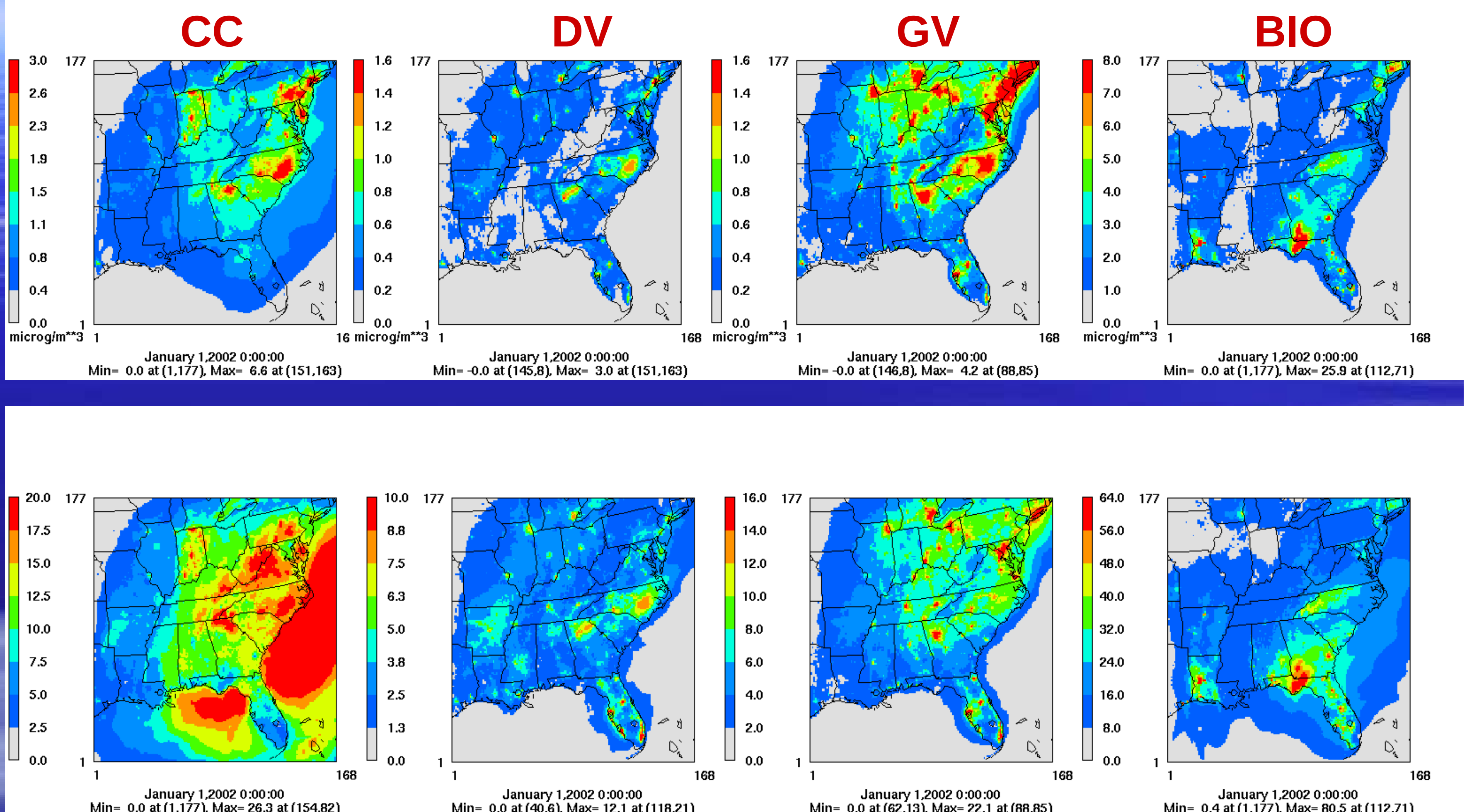
(Modified from Olerud and Sims, 2004)



Source Contributions (Jan)

Abs

%



Speciated % Concentrations Contributions to Total $PM_{2.5}$ Changes

January 2002

	NH_4^+	SO_4^{2-}	NO_3^-	EC	POA	SOA	OIN	$PM_{2.5}$
No CC (L1)	-1.58	-8.74	0.19	-0.03	-0.03	0.10	-1.15	-11.24
No CC (L1-19)	-0.89	-8.84	0.56	-0.02	-0.02	0.06	-0.81	-9.96
No Diesel	-0.19	0.31	-0.92	-0.90	-0.42	0.06	-0.02	-2.08
No Gasoline	-1.15	0.24	-2.78	-0.05	-0.29	-0.10	-0.15	-4.28
No Biomass	-0.44	-0.50	-0.81	-1.36	-8.30	-0.52	-3.10	-15.03

July 2002

	NH_4^+	SO_4^{2-}	NO_3^-	EC	POA	SOA	OIN	$PM_{2.5}$
No CC (L1)	-3.98	-30.02	-0.18	-0.02	-0.03	-0.10	-0.65	-34.98
No CC (L1-19)	-2.33	-33.20	0.38	-0.02	-0.02	-0.01	-0.54	-35.74
No Diesel	-0.23	-1.06	-0.28	-0.89	-0.38	-0.23	-0.01	-3.08
No Gasoline	-0.70	-0.45	-0.28	-0.04	-0.27	-0.22	-0.12	-2.08
No Biomass	-0.07	-0.19	-0.01	-0.20	-1.27	-0.11	-0.66	-2.51

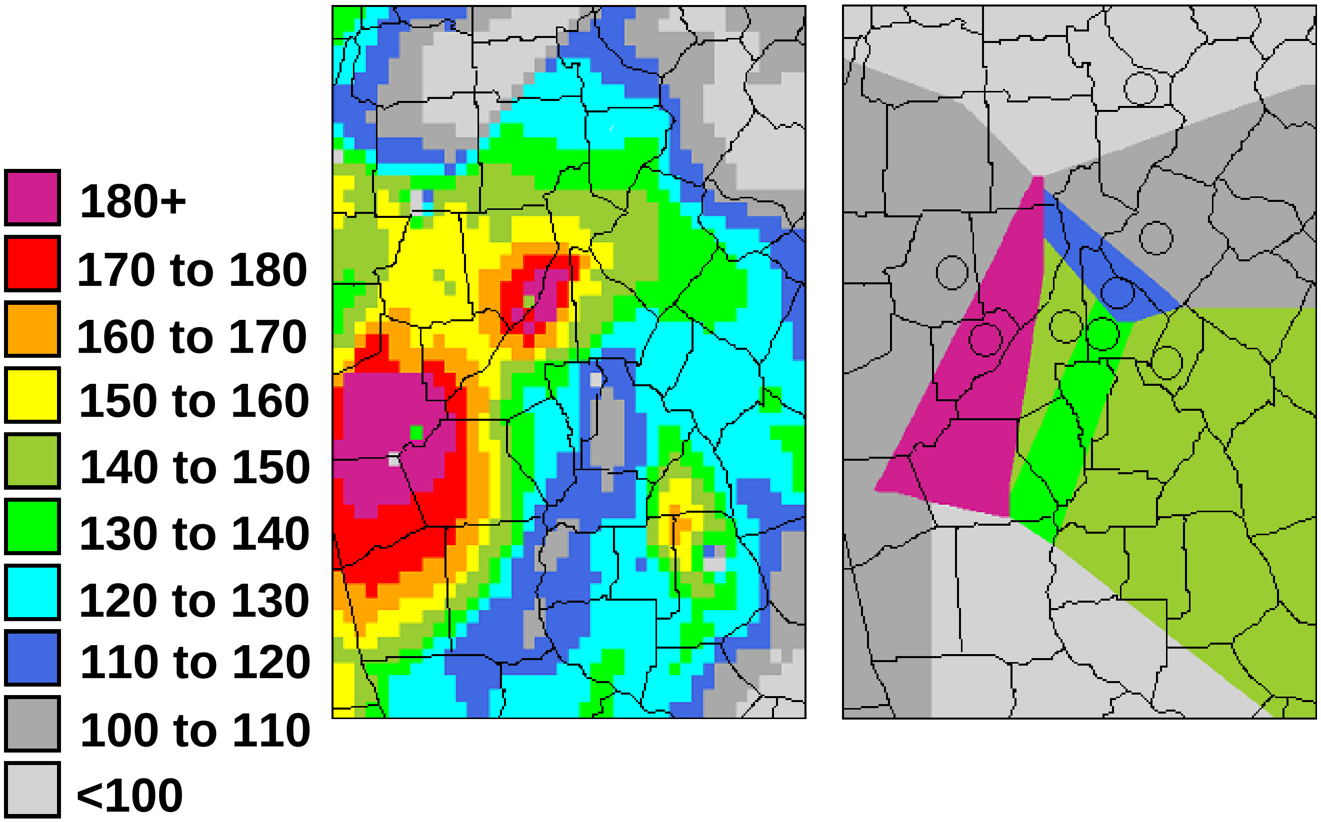
Summary of CMAQ source sensitivity study

- Source Contributions for January and July 2002

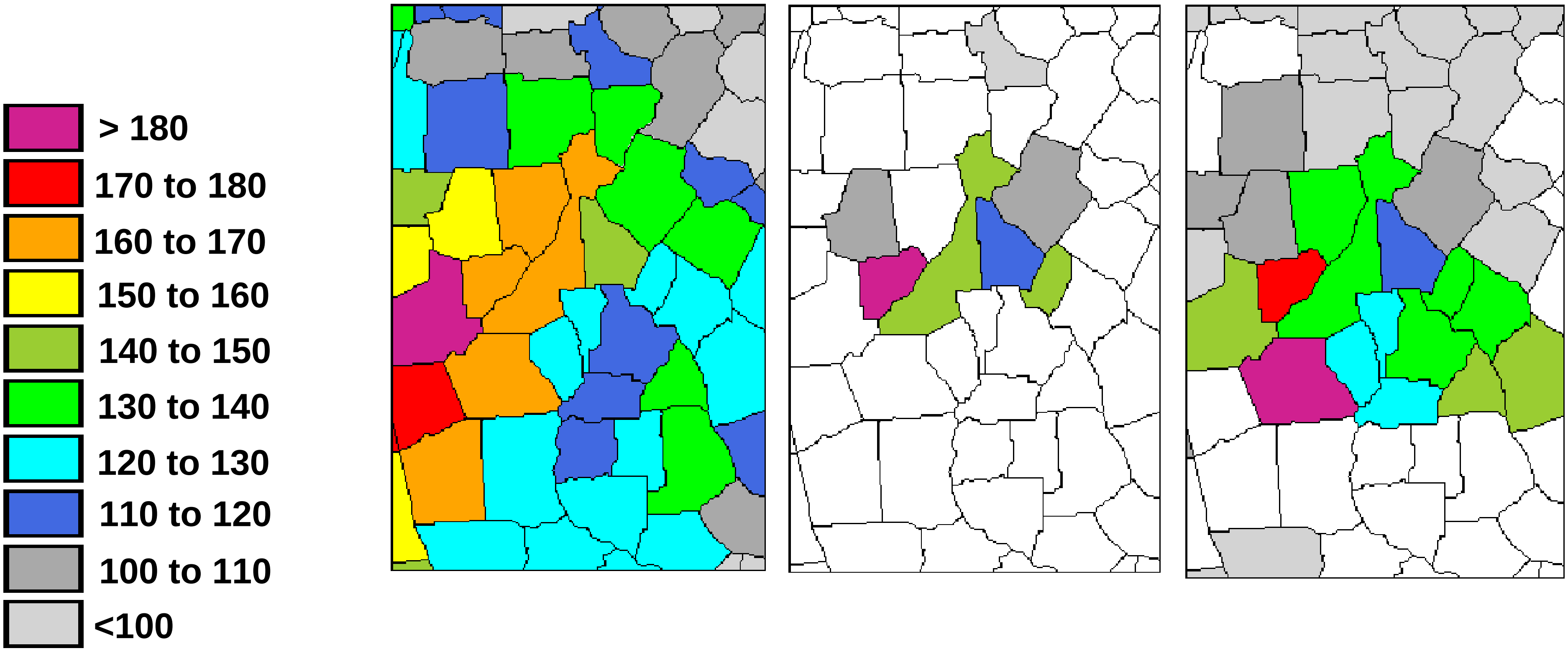
Source		Average $PM_{2.5}$ reduction (%)	Maximum $PM_{2.5}$ reduction (%)	Most affected region	Top 3 affected species
CC (L1)	Jan	11.2	26.3	ORV	SO_4^{2-}, NH_4^+, OIN
	Jul	35.0	67.8	Mid- Atlantic	SO_4^{2-}, NH_4^+, OIN
CC (L1-19)	Jan	10.0	30.1	ORV	SO_4^{2-}, NH_4^+, OIN
	Jul	35.7	69.8	Mid-Atlantic	SO_4^{2-}, NH_4^+, OIN
DV	Jan	2.1	12.1	Southern Fl.	NO_3^-, POA, EC
	Jul	3.1	27.4	Southern Fl.	SO_4^{2}, EC, POA
BIO	Jan	15.0	80.5	S.E. U.S., LA	POA, OIN, EC
	Jul	2.5	54.7	S.E. U.S., LA	POA, OIN, EC
GV	Jan	4.2	22.1	N.E. Corridor	NO_3^-, NH_4^+, POA
	Jul	2.1	22.0	Southern FL.	NH_4^+, SO_4^{2-}, NO_3^-

- Future Work
  - Source apportionment to 6 additional sources in Jan/Jul 2002
  - Inter-comparison with reactive tracer method (e.g., CAMx-PSAT, CMAQ-TSSA)

Health Analysis Example: Various Exposure Estimates



County-level Exposure Estimates



Health Datasets

MORTALITY

Timeframe: 1987-2005

Spatial scope: National (U.S.)

Spatial resolution: Community level for 108 urban communities

Other notes:

- Cause of death available (CVD, Respiratory, accident, non-accidental)
- Age categories (<65, 65 to 74, >75 years)

HOSPITAL ADMISSIONS

Timeframe: 1999-2005

Spatial scope: National (U.S.)

Spatial resolution: County-level for 204 U.S. urban counties

Other notes:

- **Cause of death available (CVD, Respiratory, accident, non-accidental)**
- **Age categories ($\geq$65 years)**

Other health outcomes: Birth weight and gestational age

Texas Birth Data:

- Texas Vital Record Information
- Dates: 1997-2004
- Geo-coded data with information on mother and father of baby and birth outcome

Case study in Harris County (2001-2002)

We simultaneously model birth weight and gestational age as a function of pollution and other covariates:

- Mother/Father Age Group
  - Ages 5-19,20-24,25-29,30-34,35-39,40-100
- Mother/Father Education
  - Less than High School, High School, More than High School
- Mother/Father Racial Group
  - Other, White, Black, Asian
- Average Temperature from Day of Birth
  - Fit using a polynomial spline with 4 df

Birth outcomes case study

- For an initial analysis in Harris County (2001-2002) we used the first trimester averaged concentration of a particular pollutant as the indicator of exposure to that pollutant
  - This average is based on the clinical estimate of gestational age and the date a woman gave birth
    - Different women had different exposures based on their window of time being pregnant
- For Ozone we used the AQS data
- For PM2.5 we used CMAQ

Results for Birth Weight

There is a significant association with PM2.5 and ozone.

Variable	Estimate (PM2.5 Model)	P-Value	Estimate (Ozone Model)	P-Value
Intercept	2686.9892	<.0001	2524.2717	<.0001
Female Baby	-114.9386	<.0001	-109.3604	0.0001233
Mother's Age >= 40	-481.3916	0.001114	-504.9132	0.0010263
Father's Education > High School	119.6293	0.009228	113.6548	0.0145684
First Trimester Pollution Average	-34.6582	<.0001	-6692.9765	0.0338547

Results for Gestational Age

There is a significant association with PM2.5 and ozone.

Variable	Estimate (PM2.5 Model)	P-Value	Estimate (Ozone Model)	P-Value
Intercept	37.9362417	<.0001	36.7545027	<.0001
Mother's Age >= 40	-2.5384447	0.001317	-2.5887878	0.001374
Black Mother	Not Significant at 0.05 Level	Not Significant at 0.05 Level	-2.7440289	0.042072
First Trimester Pollution Average	-0.2046888	<.0001	-34.5695212	0.038617

The integration in these epidemiologic analysis of our results on PM composition and sources, while accounting for uncertainty in the statistical and numerical models and the data, will determine the constituents and potential sources of the PM2.5 mixture that are most strongly associated with different adverse health effects.

We believe the knowledge and modeling based developed under this STAR award should be helpful in the assessment of the need for public policies aimed at managing fine PM air quality.

Acknowledgements

- U.S. EPA Science to Achieve Results (STAR) Grant # R833863
- Gregory Stella, Alpine Geophysics, Inc.
- Bebhinn Do and George Bridgers, N.C. DAQ
- Amy Herring, UNC
- ML Miranda, Duke