CENTER FOR CLEAN AIR RESEARCH

UNIVERSITY *of* WASHINGTON

UW CCAR Scientific Advisory Committee Meeting

July 23rd & 24th 2013

SCHOOL OF PUBLIC HEALTH

UNIVERSITY *of* WASHINGTON

Lovelace

LOVELACE RESPIRATORY RESEARCH INSTITUTE

WASHINGTON STATE UNIVERSITY

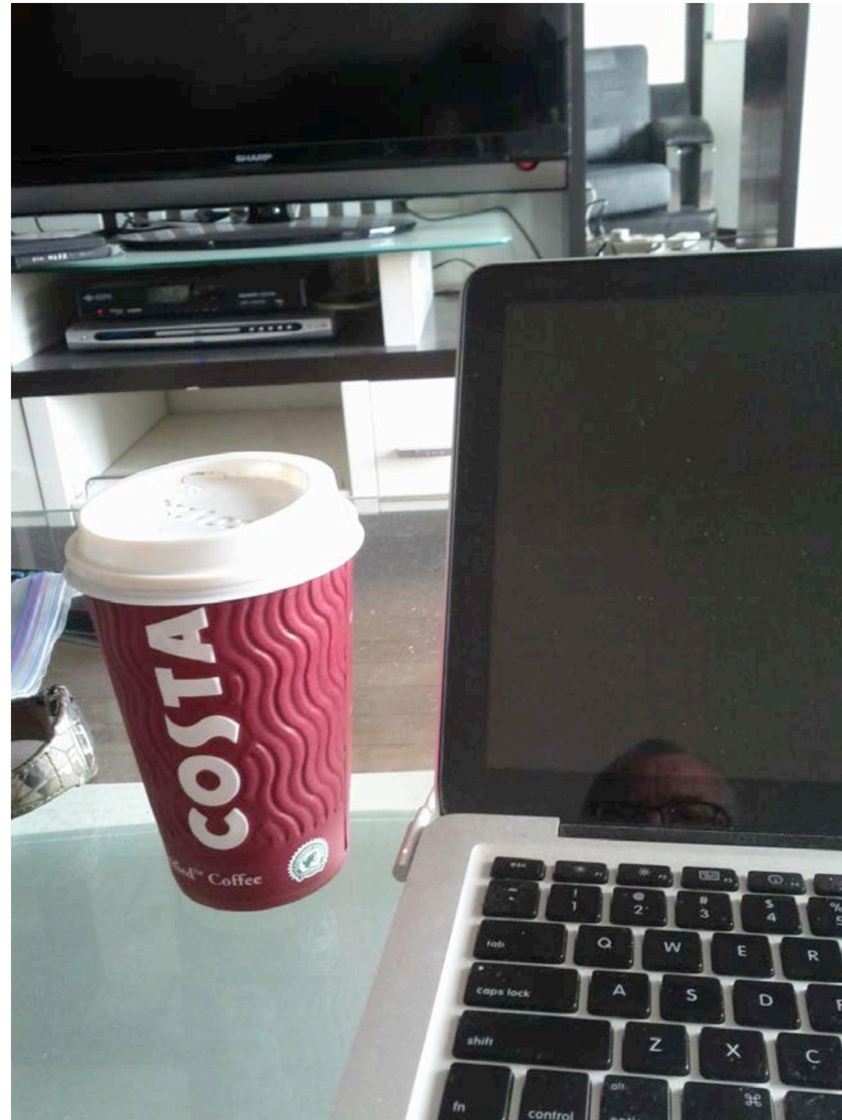
World Class. Face to Face.

THE UNIVERSITY *of* NEW MEXICO

Campus-Wide Wireless Connection – NetID: **event0216** Password: **Q6p4-R5j3-R9x7**

the challenges of a sabbatical

Wangjing district, Beijing



Updates on the Scientific Advisory Committee

original members	current meeting	next meeting
John Balmes (chair)		John Balmes (chair)
Mike Brauer	Mike Brauer	Mike Brauer
Brent Coull	Brent Coull	Brent Coull
Ian Gilmour	Ian Gilmour	Ian Gilmour
Jake Lusis		Jesus Araujo
Nick Mills	Nick Mills	Nick Mills
Tom Peters	Tom Peters	Tom Peters
Arden Pope	Arden Pope	Arden Pope
Sanjay Rajagopalan	S Rajagopalan (chair)	Sanjay Rajagopalan
Barbara Turpin		Barbara Turpin

CCAR projects & cores

Project 1
roadway exposure characterization

M Yost (PI), T Larson, C Simpson, T Jobson, T VanReken

Project 2
exposure atmosphere generation

J McDonald (PI), T Larson

Project 3
toxicology

M Campen (PI), M Rosenfeld, J McDonald

Project 4
human clinical studies

J Kaufman (PI)

Project 5
epidemiology cohort study

J Kaufman (PI), S Vedal, C Curl

Project 6
multipollutant exposure modeling

L Sheppard (PI), A Szpiro, P Sampson

Biostats Core

Admin Core

Since we last met:

1. responses to SAC review

1. Clean Air Research Centers (CLARC) and CCAR meetings/seminars:
   - EPA center webinar, March 2013 – M Yost & T Larson (project 1)
   - Scott Fruin (UCLA) visit and seminar, Seattle May 2013
   - Planning annual meeting Seattle, July 2013 – highlights + collaborative projects

Since we last met:

3. Projects
   - P1 – Los Angeles x 2 seasons; South Seattle pilot study; LRRI chamber analyses (with P2)
   - P2 – irradiation chamber refinements
   - P3 – new endpoints and collaborations
   - P4 – to be discussed, again
   - P5 – in transit monitoring – pilot studies and heating season in Winston-Salem
   - Biostats Core – sparse PCA CSN data; measurement error correction methods applied to Sister Study

overview (selected) of SAC comments:

- overarching hypothesis?
- integration of mobile and chamber characterization data, and with experimental and observational exposures
- distinguishing roadway pollution from other sources
- more sensitive toxicologic endpoint(s)
- scripted study – prefer crossover study
- drop use of “eigenpollutant”; keep in mind spatial scales of contrasts

SAC input especially on:

1. reactions to early findings and approaches:
   - mobile and chamber monitoring (project 1)
   - tox models/endpoints (project 3)
   - in-transit exposures (project 5)
2. project 2 exposure atmospheres
3. project 4 – what now, again?
4. Biostats core – approaches to dimension reduction and measurement error correction
5. collaborative projects

<u>Outline of today's activities</u>

1. Individual project reviews, updates, discussions
   - Biostats Core I
   - Project 1

[LUNCH] with <u>posters</u> in R2228

- projects 2, 3 and 4

[BREAK]

- Project 5
- Biostats Core II

Today contin.

2. Cross-center collaborations
3. General discussion

[DINNER] at Foege Hall

Tomorrow's activities

1. SAC closed meeting
2. SAC report and discussion

UW CCAR - Project 1 work in progress: Developing multi-pollutant features of traffic emissions with mobile monitoring

CLARC/SAC meeting, July 23-26, 2013

Investigators: Michael Yost, Tim Larson, Chris Simpson, UW; Tom Jobson, Tim VanReken, WSU

University of Washington
Center for Clean Air Research (CCAR)

Project Aims

1. Near-roadway gradients exist in traffic-associated air pollution, and these gradients can be measured using an assortment of gas and particles sensors deployed in a moving vehicle.
2. Air pollutant measurements collected at “fuzzy points” on multiple occasions over short time periods using a mobile monitoring platform efficiently capture spatial variation in pollutant concentrations, and are well correlated with two-week integrated measurements collected at those same fuzzy points, and can be predicted using GIS spatial co-variates.
3. Measure spatial variation in concentrations of selected air pollutants at two-week average fixed sites

Outline of talk

- Background on mobile sampling campaigns
- Measuring near-roadway gradients (Aim 1)
- Analysis of time-series multi-pollutant features
- Multi-pollutant analysis at fuzzy points (Aim 2)
- Next steps

Instrument Platform

Parameter	Instrument
Particle Light Scattering Coef.	Radiance Nephelometer M903
Particle-bound PAHs	EcoChem PAS 2000CE
Ultrafine Particle Counts	TSI P-Trak Model 8525
Black Carbon + Ultraviolet Abs. PM (UVPM)	Magee Scientific microAeth AE52
31 Sizes of Particle Counts Total Number Concentration Mean ultrafine Particle Diam.	GRIMM 1.109 spectrometer + GRIMM NanoCheck 1.320
Nitric Oxide (NO) Nitrogen Dioxide (NO_2) Oxides of Nitrogen (NO_x)	2B Technologies 410 -NO Aerodyne CAPS NO_2 Monitor 2B Technologies 410 w NO_x
Ozone (O_3)	Optec Chemiluminescent Analyzer
Carbon Monoxide (CO)	Langan CO Monitor T15N
Volatile Organic Compounds (VOCs)	Inficon 2020ppb Photoionization Detector
Carbon Dioxide (CO_2)	Senseair CO_2 K-30-FS Sensor
Visual Route Record	WebCam
Real-Time Positioning	Garmin GPS

Project 1: Sampling Schedule

Activity	Est. Begin Date	Est. End Date	Year of Study	Location - New
Pilot Testing of Mobile system	**8/15/11**	**11/15/11**	**1**	**Seattle, WA**
Field Sampling, City 1 (Heating)	**11/29/11**	**12/20/11**	**1**	**St. Paul, MN**
Field Sampling, City 2 (Heating)	**2/5/12**	**2/25/12**	**2**	**Baltimore, MD**
Characterization of LRRI Exposure Atmospheres	**4/16/12**	**5/17/12**	**2**	**Albuquerque, NM**
Field Sampling, City 2 (Non-Heating)	**6/8/12**	**6/30/12**	**2**	**Baltimore, MD**
Field Sampling, City 1 (Non-Heating)	**7/25/12**	**8/15/12**	**2**	**St. Paul, MN**
Field Sampling, City 3 (Heating) *	**1/3/13**	**1/23/13**	**3**	**Winston-Salem, NC**
Field Sampling, City 4 (Heating)	**2/5/13**	**2/25/13**	**3**	**Los Angeles, CA**
Field Sampling, City 4 (Non-Heating)	**6/15/13**	**6/30/13**	**3**	**Los Angeles, CA**
Field Sampling, City 3 (Non-Heating) *	**8/1/13**	**8/20/13**	**3**	**Winston-Salem, NC**
Characterization of UW Exposure Atmospheres	**10/1/13**	**10/15/13**	**3**	**Seattle, WA**
Sampling with GT CLARC Instrumentation	***9/1/13***	***9/20/13***	***3***	***Atlanta***

* Passive only

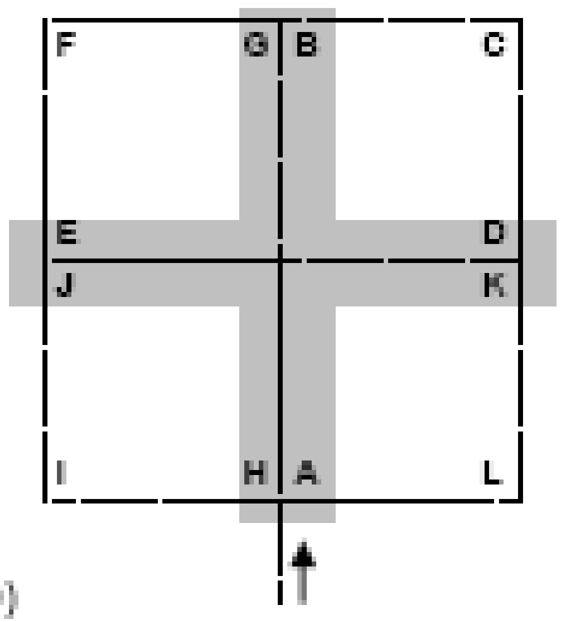
Monitoring Campaign Data

- Two Data sources:
  - PASSIVE – Passive samplers (2-week averages)
  - MOBILE – Mobile data (30s & ~15 min.; time-corrected)
- 3 Mobile Routes
  - One fixed route, 2-7 pm (**evening commute**)
  - All routes time adjusted to central fixed site
  - 15 Fuzzy points per route (43 total)
  - Fuzzy points coincident with passive samplers

Mobile Platform Collection

Traffic Intersections as "Fuzzy Points"

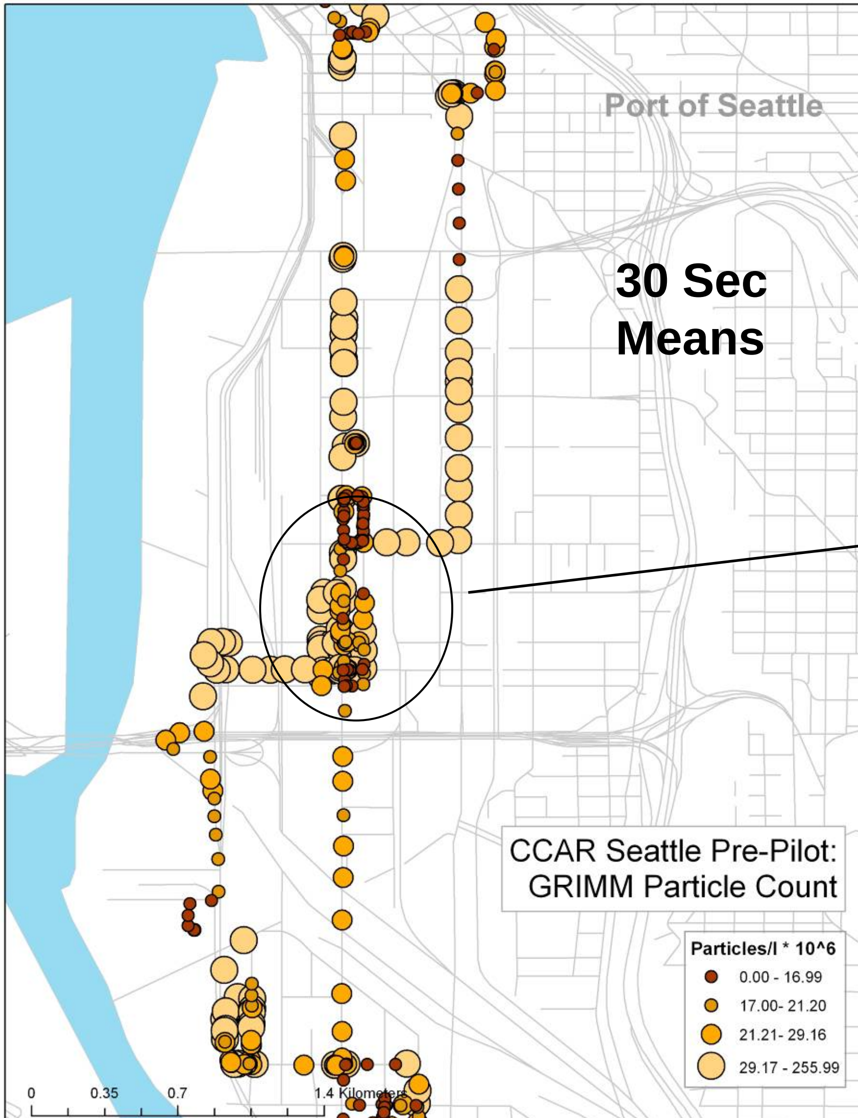
- Measure pollutant marker (e.g. σ_{ap}) at selected traffic intersections during peak afternoon traffic period
- Trace a cloverleaf / figure 8 at each intersection (~5-8 minutes); repeat
- Adjust the observed readings using fixed site data
- Calculate the median of the adjusted readings for each pass through a fuzzy point campaign

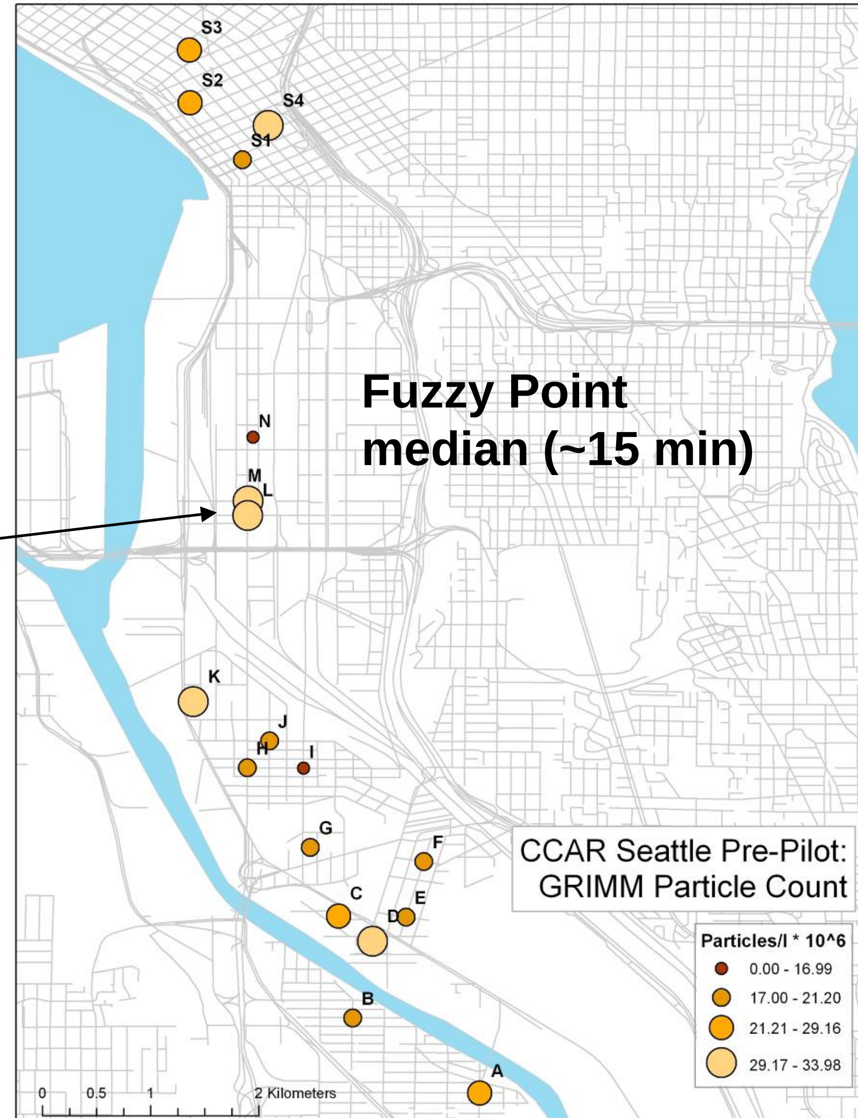


(b)

$$\text{Adjusted Reading} = \frac{\text{Observed 10-sec reading from mobile x Campaign median from fixed}}{\text{30-min moving median from fixed}}$$

Fuzzy Points - Detail Maps

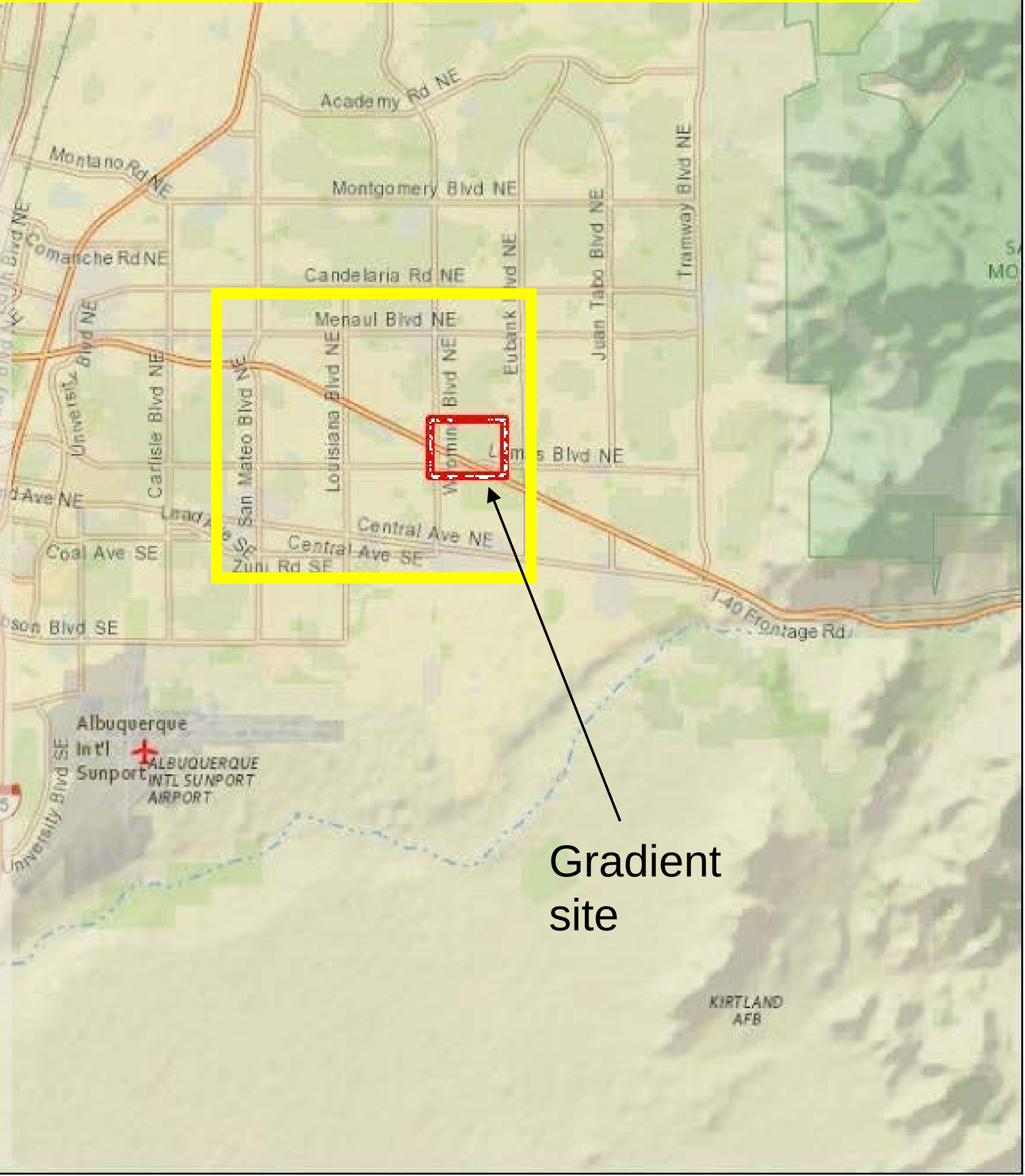




Streaming data and video

Roadway Gradient Analysis

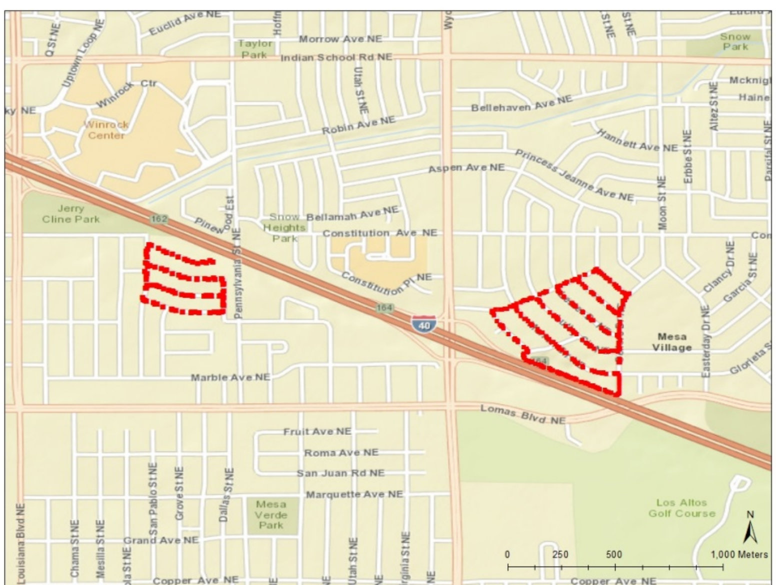
- SAC recommended a “detailed spatial/road and traffic source characteristics information”
- Developed alternative mobile sampling scheme to assess near-roadway pollutants
- Implemented this approach in all cities
- Analysis of initial test data from Albuquerque

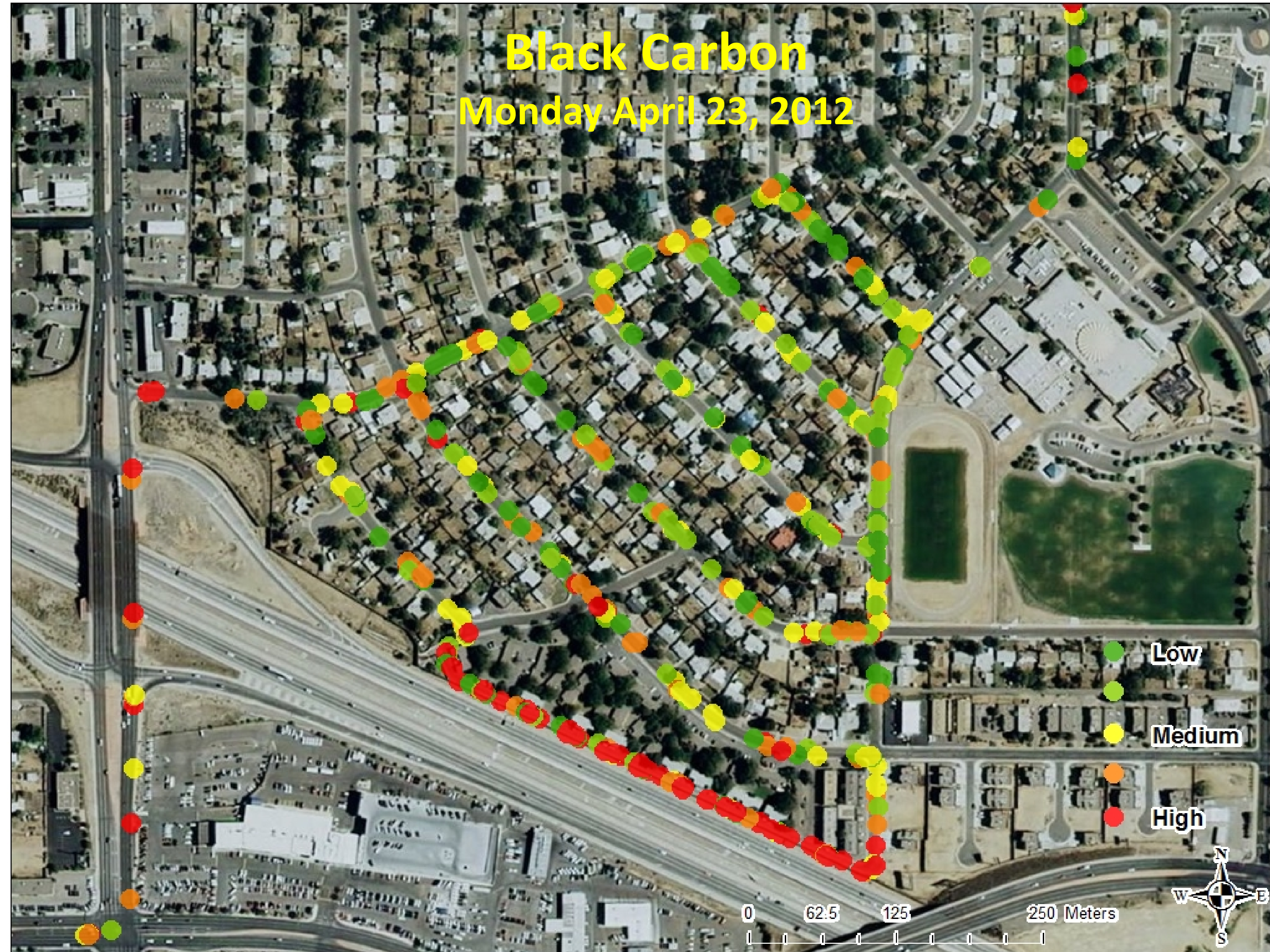


Aim 1

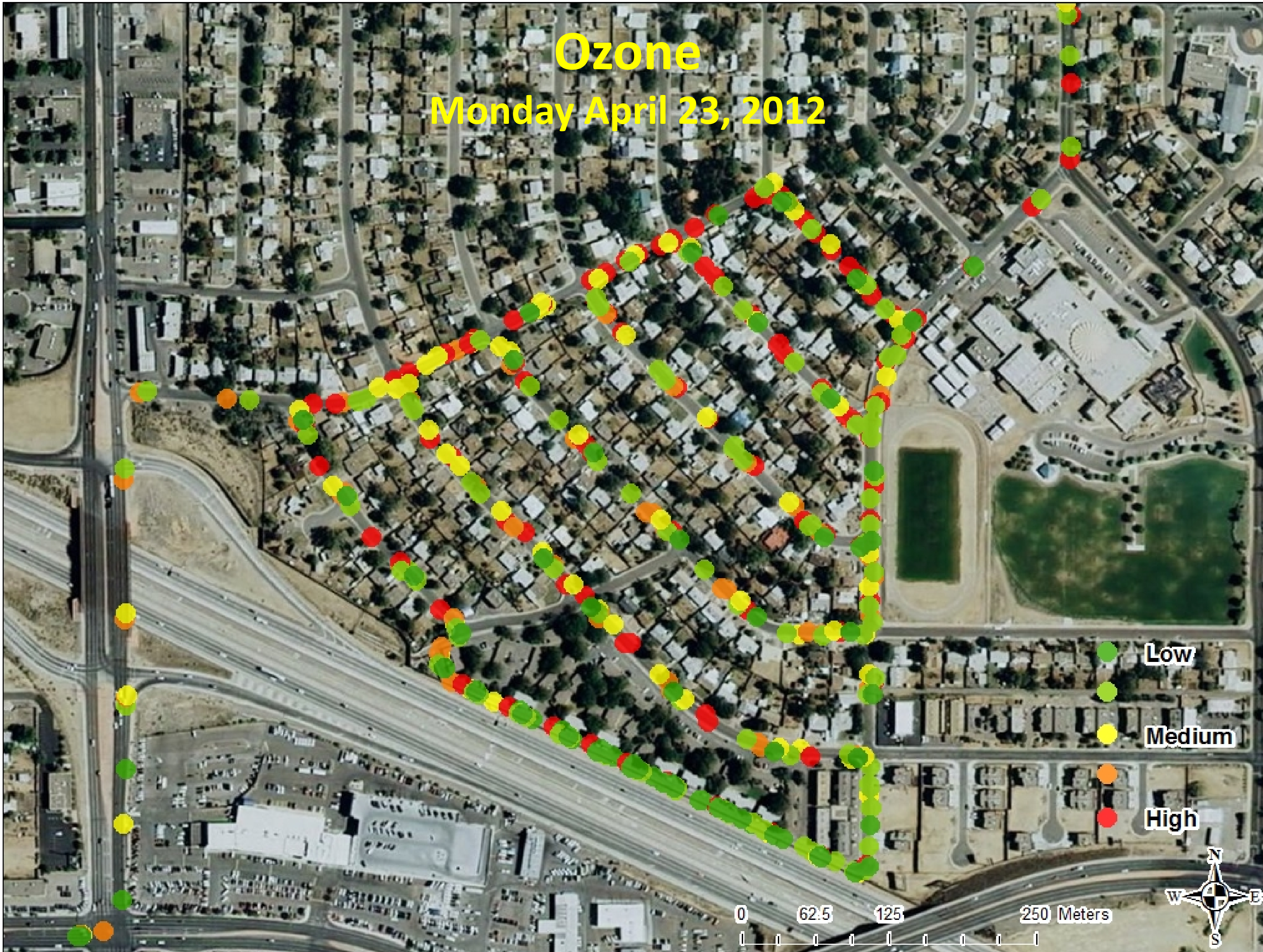
Gradient Sampling Data

- 7 sequential days (April 18 – 24, 2012)
- Evening commute timeframe (3 – 7pm)
- Roads traversed at least two times per sampling day
- North and South gradient sites spanning 30 to 500 m from I-40
- 10 second measurements

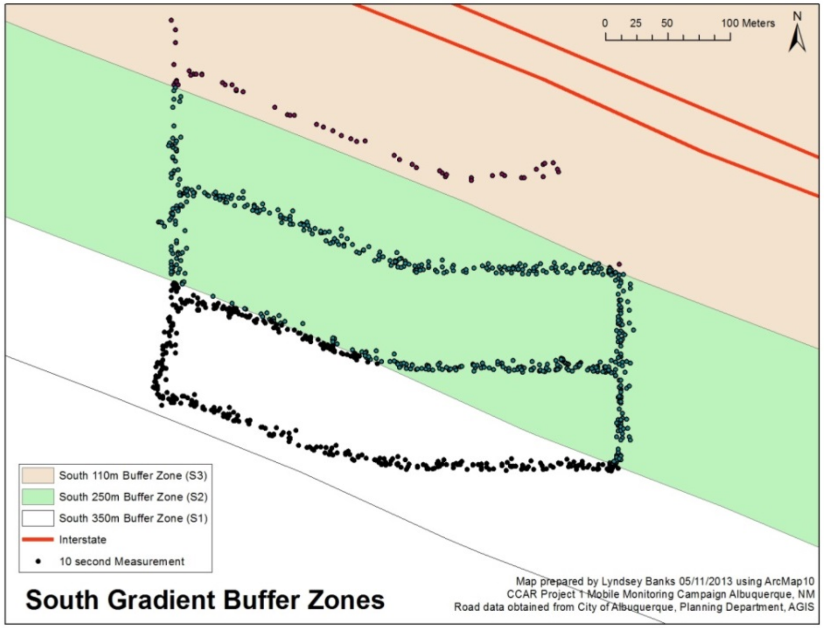


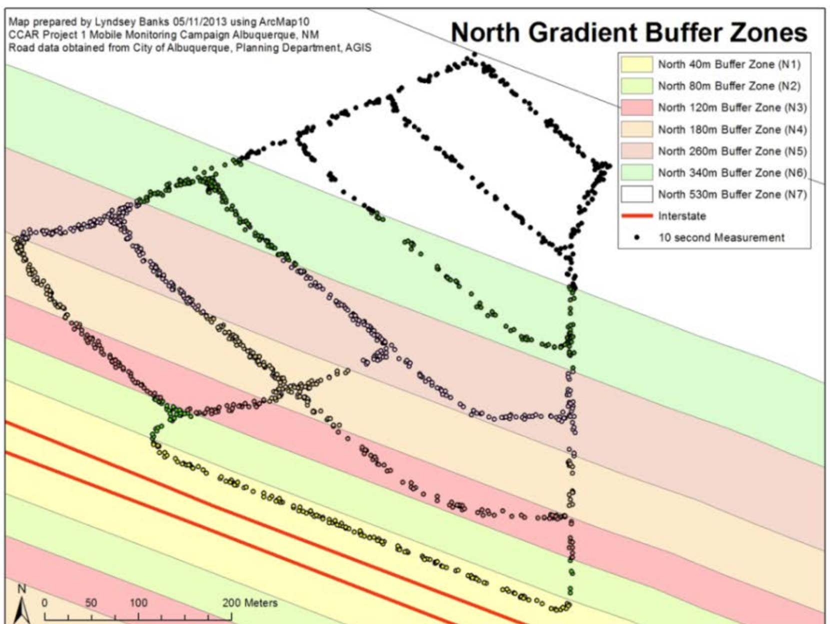


Black Carbon
Monday April 23, 2012
Low
Medium
High
0
62.5
125
250 Meters
N
S
E
W


Ozone
Monday April 23, 2012
Low
Medium
High
N
W
E
S
0
62.5
125
250 Meters

Gradient Site Analysis

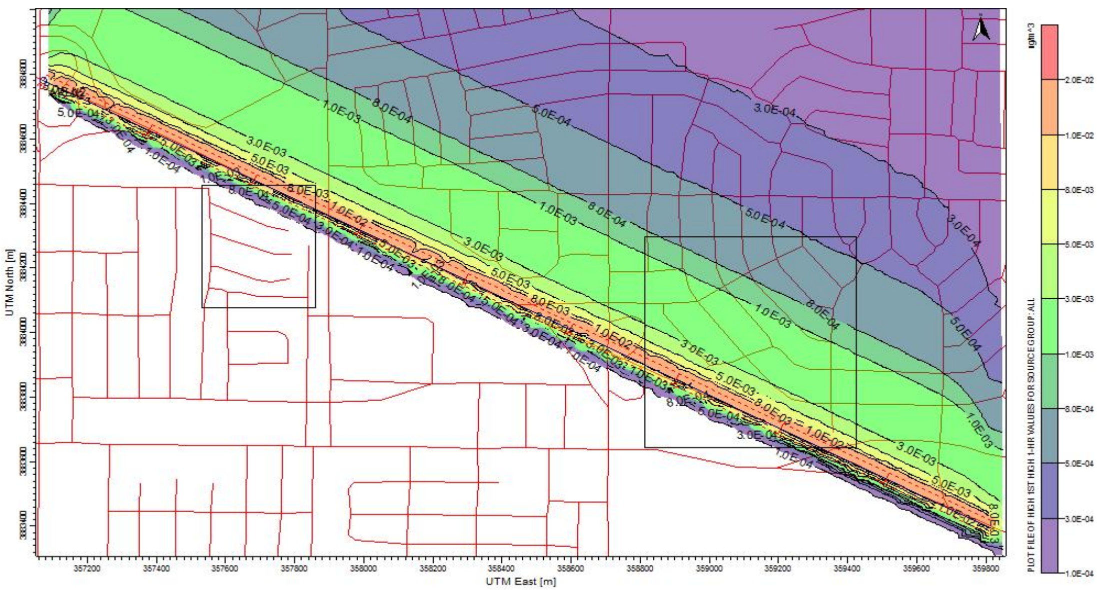




Buffer Zone	Distance From Interstate (m)	Total # of 10 s Measurements
S3	< -250 (-350)	398
S2	-100 – -250	587
S1	< -110	58
N1	< 40	148
N2	40 – 80	49
N3	80 – 120	171
N4	120 – 180	264
N5	180 – 260	332
N6	260 – 340	229
N7	> 340 (530)	265

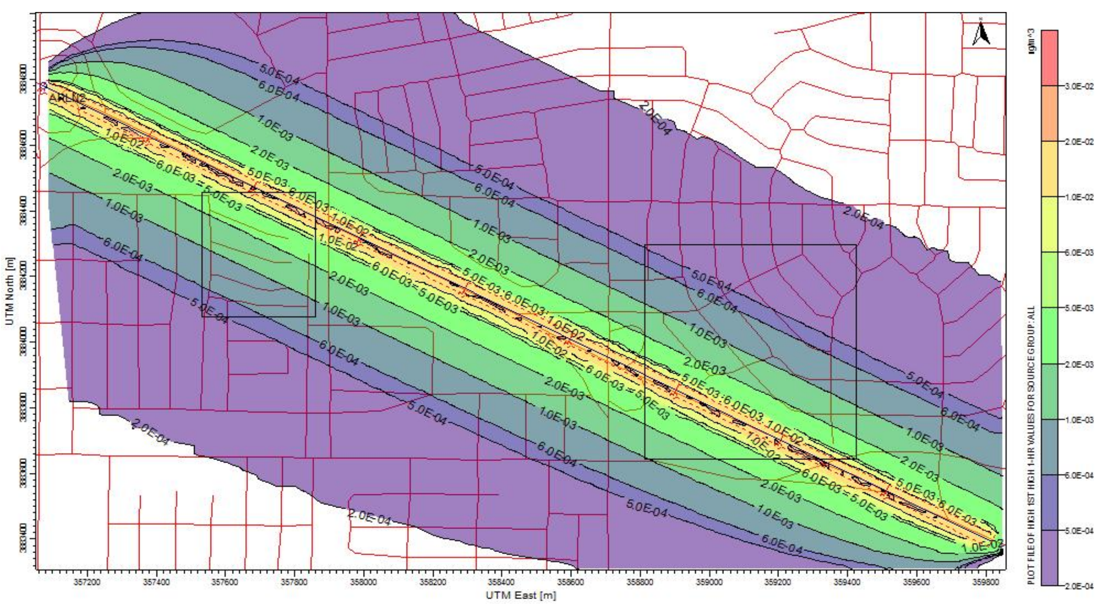
Dichotomized Dispersion Analysis

- Generated NOx dispersion maps w AERMOD View 8.2;
- EPA MOVES 2010b Emissions & NMDOT hourly traffic counts



Dominant North Dispersion

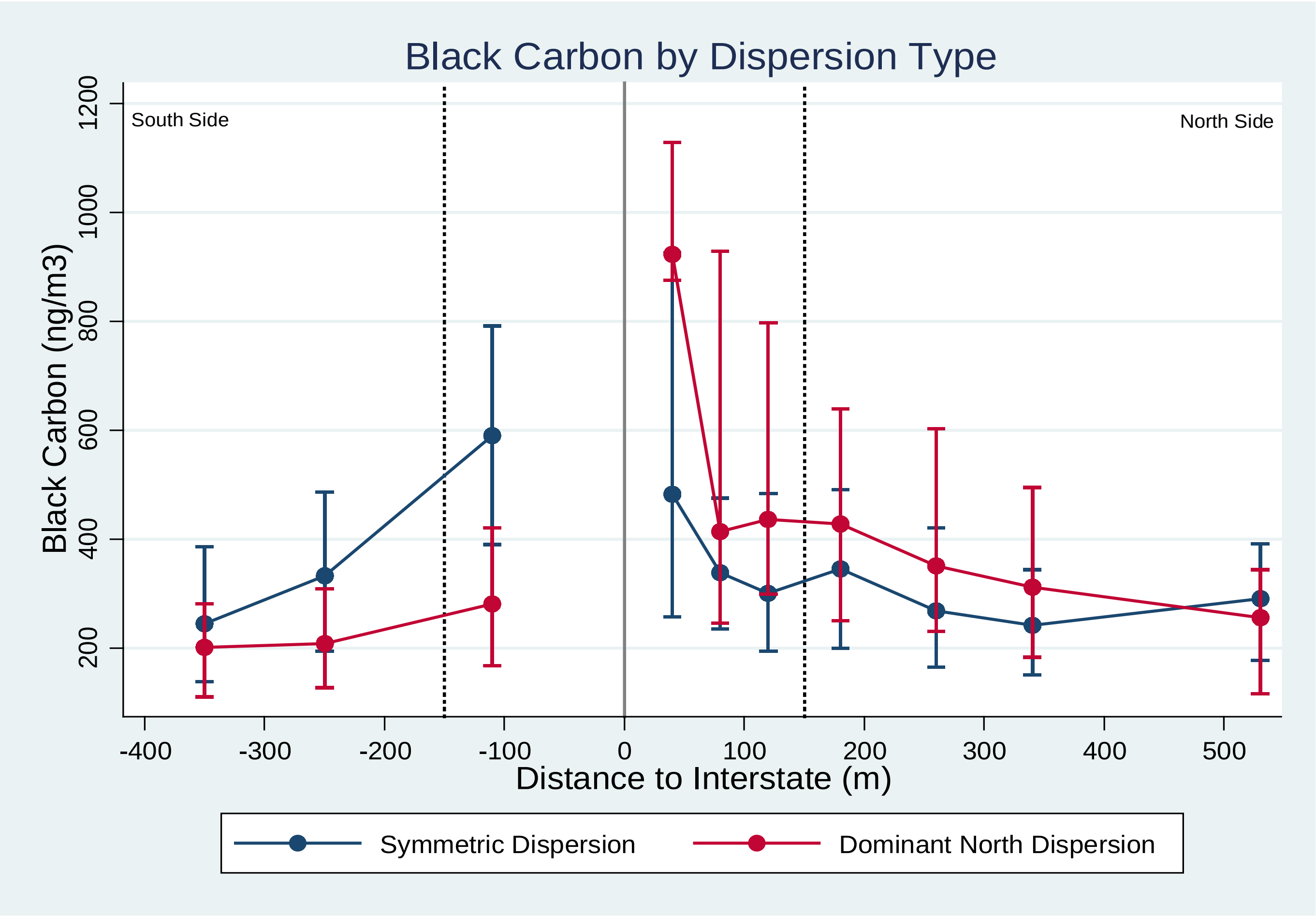
(April 18, 20 and 22)



Symmetric Dispersion

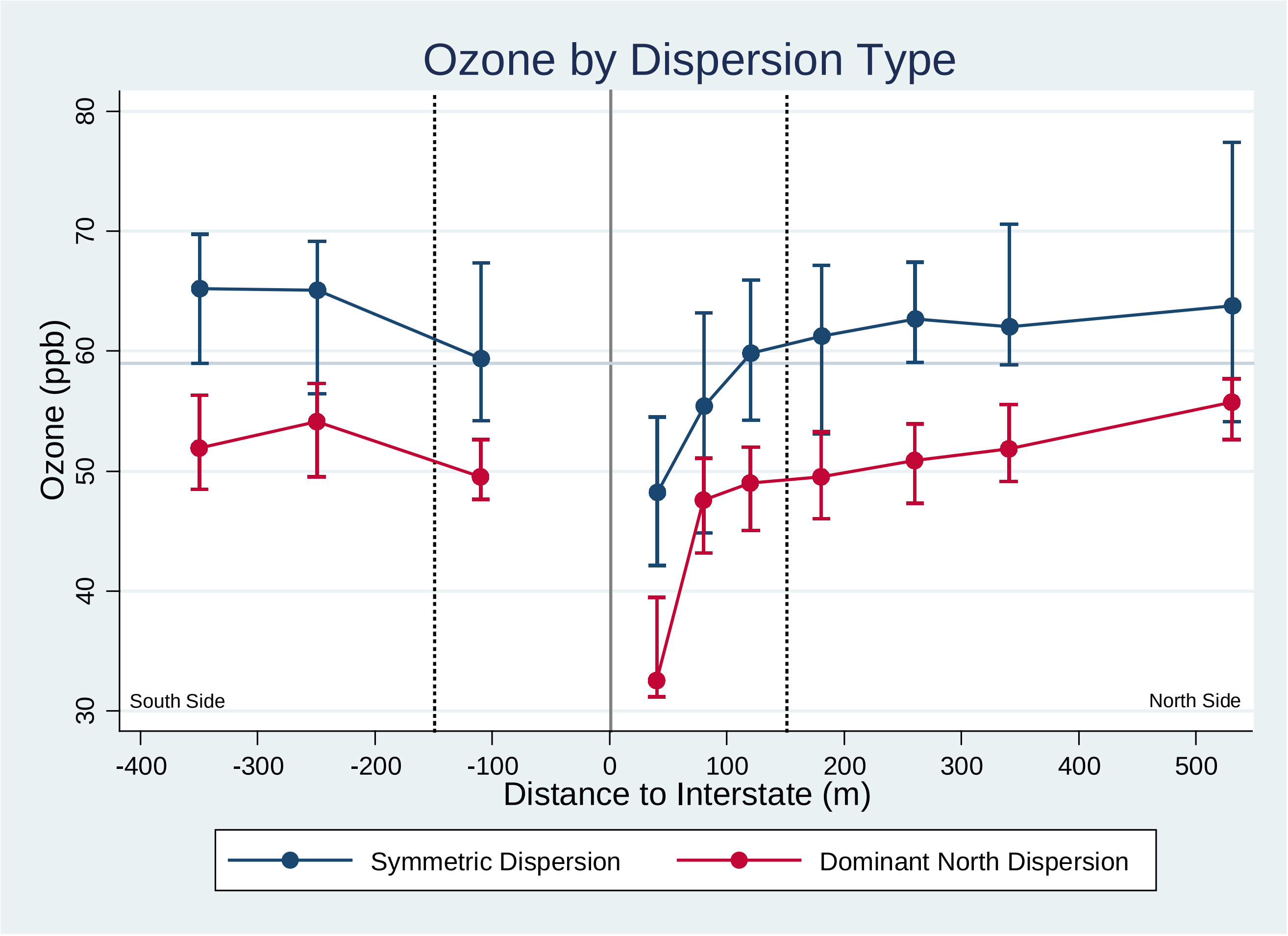
(April 19, 21, 23 and 24)

Black Carbon



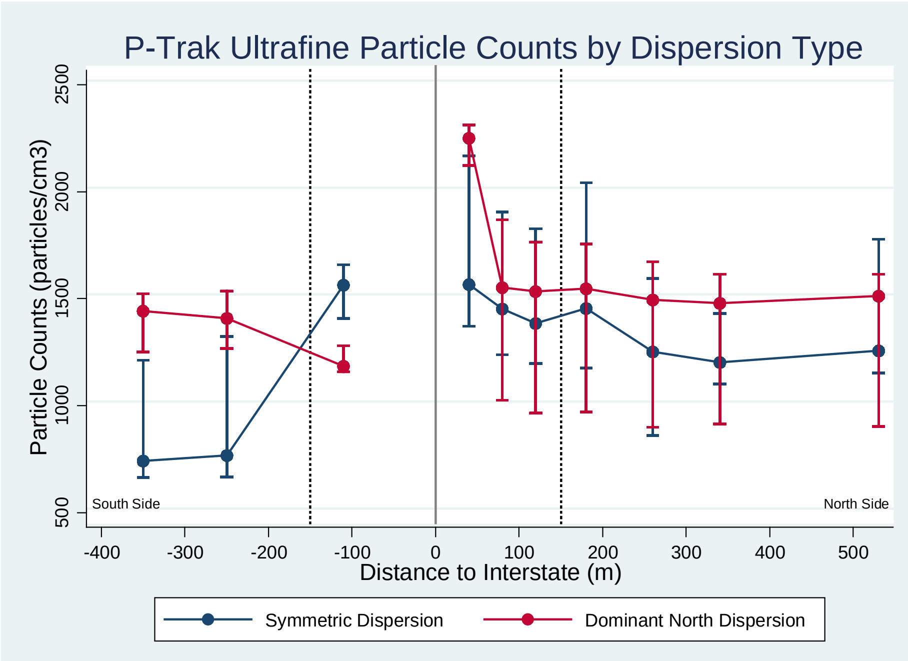
- Product of incomplete combustion of fuels
- Elevated black carbon concentrations observed downwind of interstate
- Baldauf et al. (2008) observed a similar trend: 20 meter location was 1014 ng/m^3
  200 meter location was 824 ng/m^3

Ozone

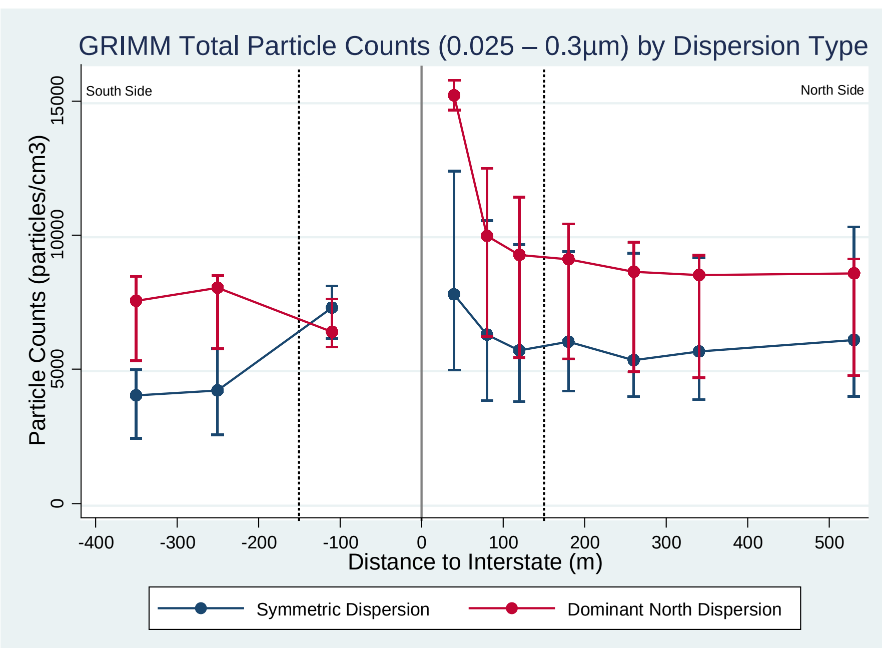


- Ozone undergoes consumption from NO on interstate ($NO + O_3 \rightarrow NO_2 + O_2$)
- Elevated ozone for symmetric dispersion consistent with stagnant meteorology
- EPA AQS Site average concentration of 59 ppb

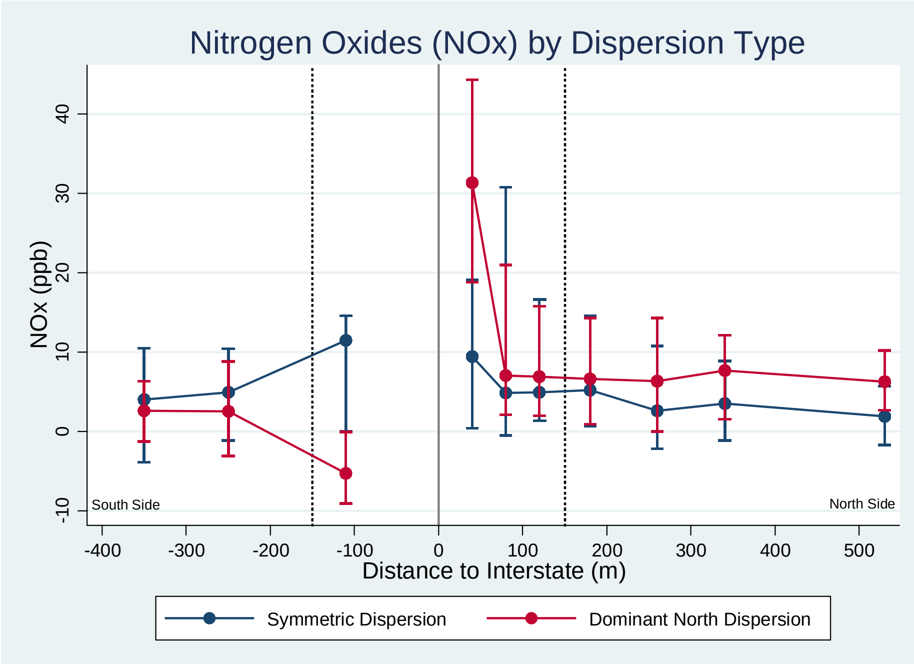
Ultrafine Particles (P-Trak)

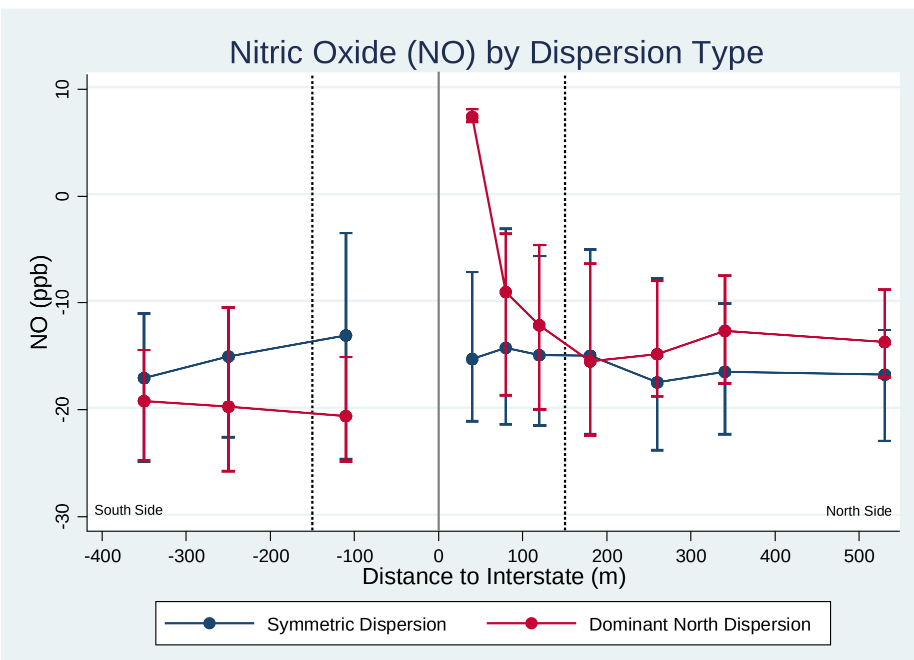


- Size selective diffusion screen used (0.05 – 1 µm)
- Approx. 1.7 times higher downwind than upwind
- Hagler et al. (2010) observed a similar trend - 1.8 times higher within the 20 to 150 m downwind region than with background ultrafine particle counts



Ultrafine Particles (GRIMM Nanocheck)





Nitrogen Oxides ($NO_x = NO + NO_2$)

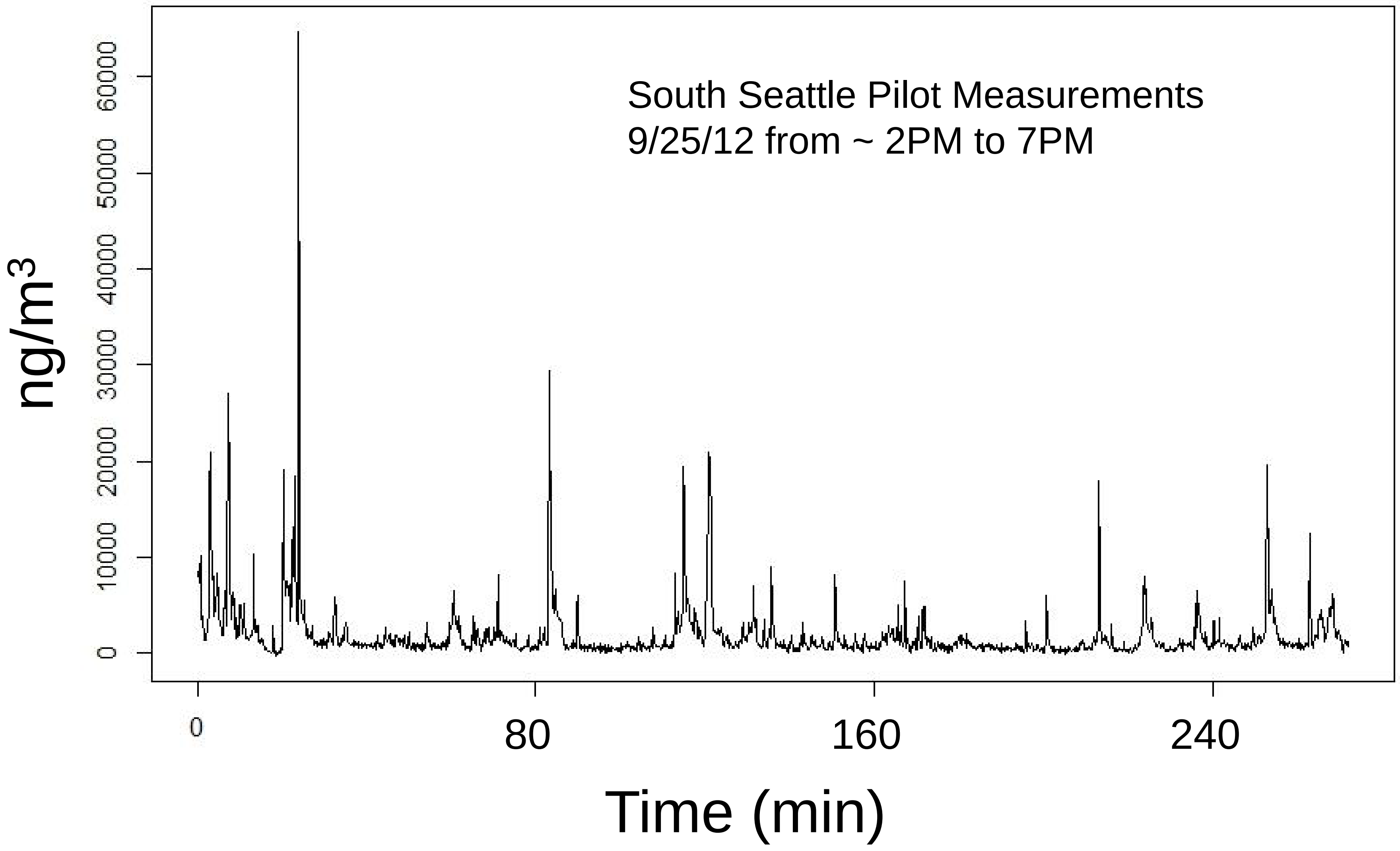
- Instrument baseline offset/interference from hydrocarbons
- Relative change indicative of the dispersion process
- EPA AQS Site average concentration of 6.4 ppb

Nitric Oxide (NO)

Analysis of mobile 10 sec. data

- Can we extend the single component analysis to the multivariate data?
- How to interpret the multivariate features?

Black Carbon time-series graph



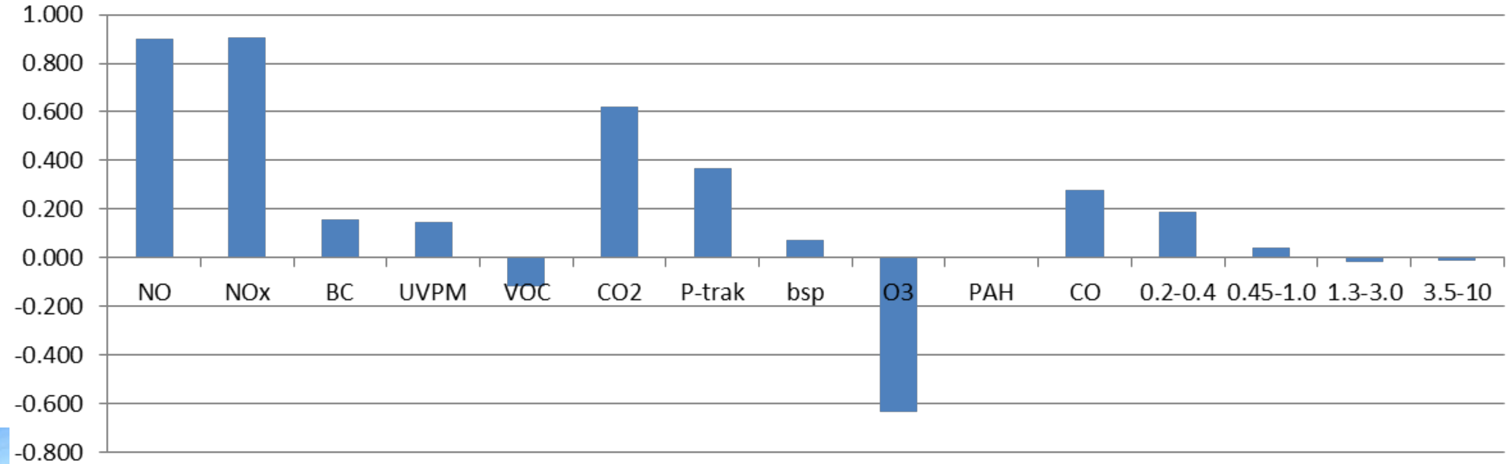
Multivariate Time-Series Analysis

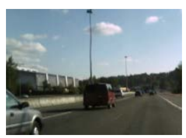
- South Seattle Pilot Study Data
- Measurements taken on 9/25/12 from ~2PM to 7PM
- Tested multivariate correlations between simultaneously measured species concentrations
- Performed simple variable reduction method on simultaneous 10-second data
- Used Principal Component Analysis with Varimax rotation
- 75% of variance represented by 5 orthogonal factors

RC1: High Nitrogen Oxides and Low Ozone Concentrations (Diesel @ load)

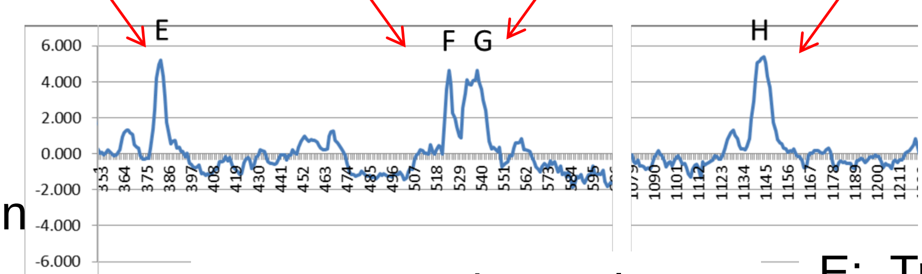
21% of total observed variance

Factor correlation with given species





Relative Factor Contribution



10 second sample #

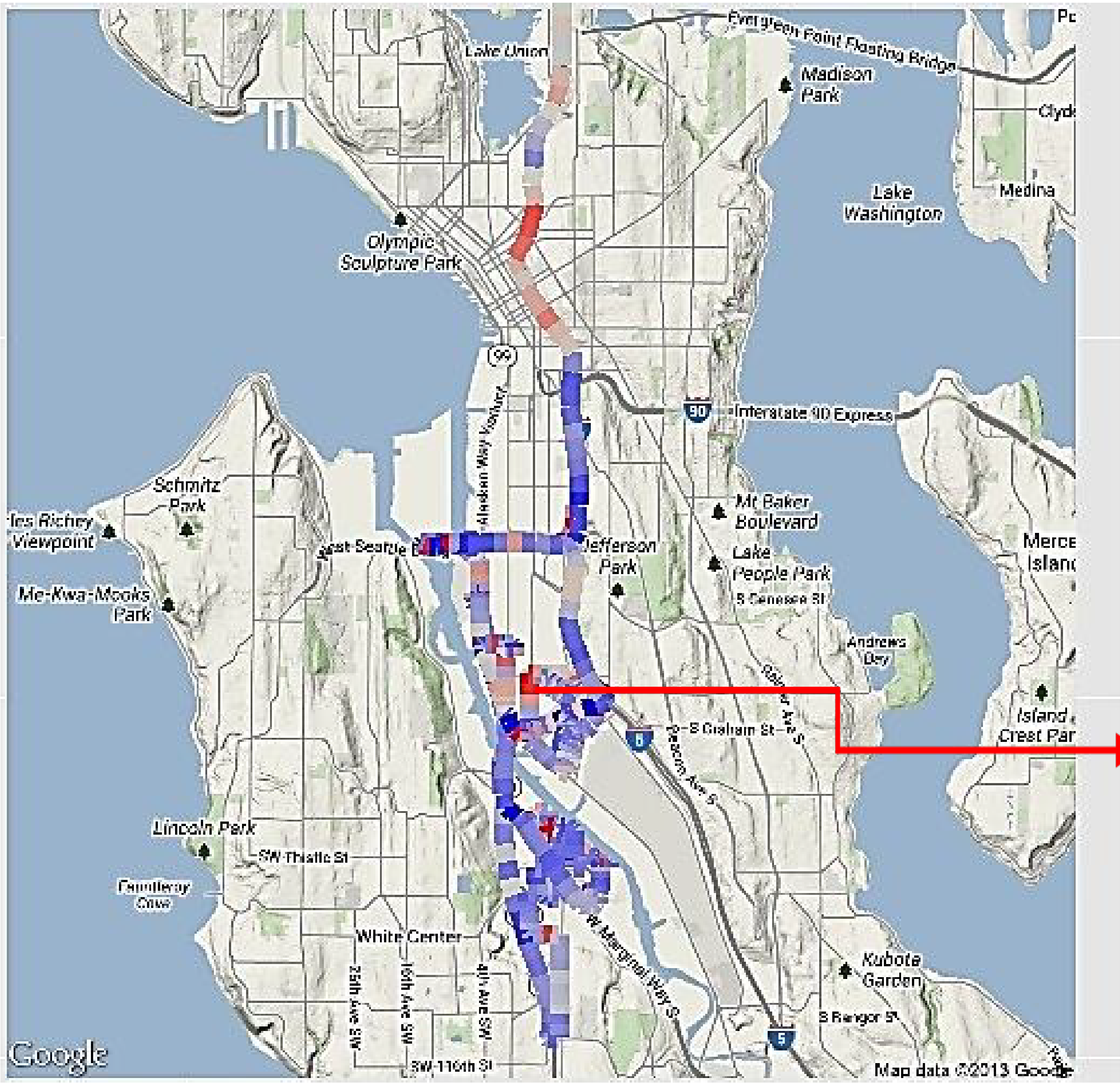
E: Truck passing uphill

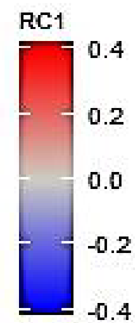
F: In uphill traffic

G&H: Roadside next to uphill traffic

Spatial Map of RC1 Score

South Seattle Pilot
9/25/12 ~ 2PM-7PM.
Darker colors indicate
higher score

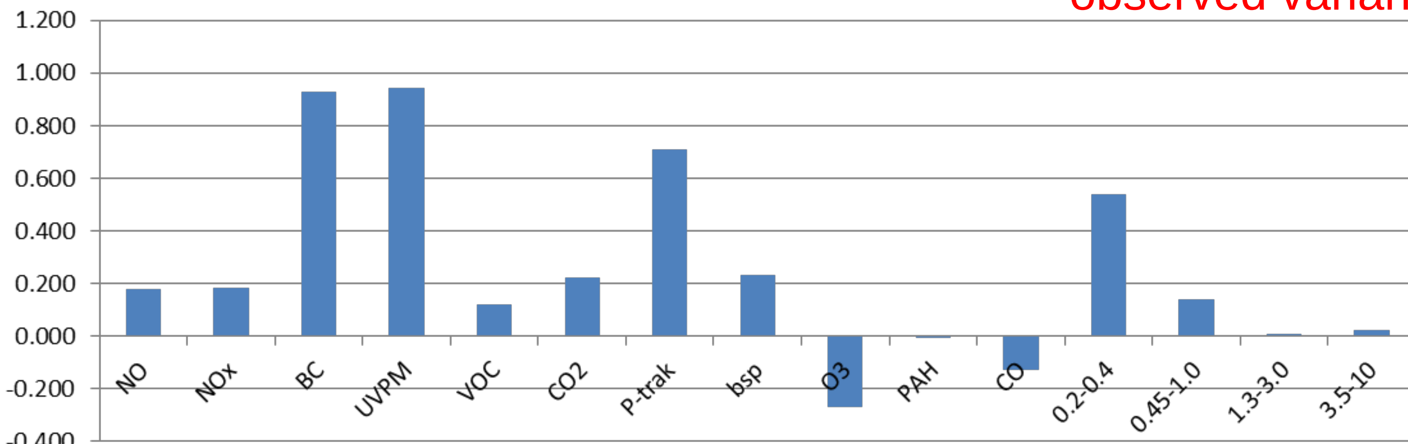




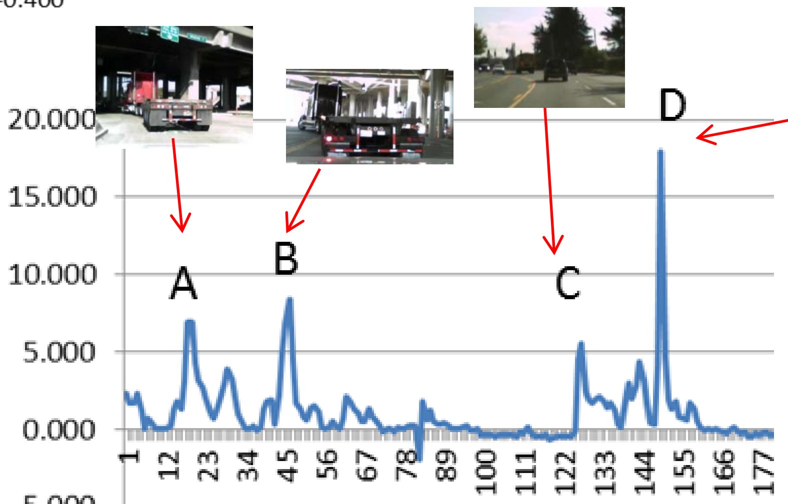
RC2: High Black Carbon and Particle Number Concentrations (Diesel Soot)

19% of total observed variance

Factor correlation with given species



Relative Factor Contribution



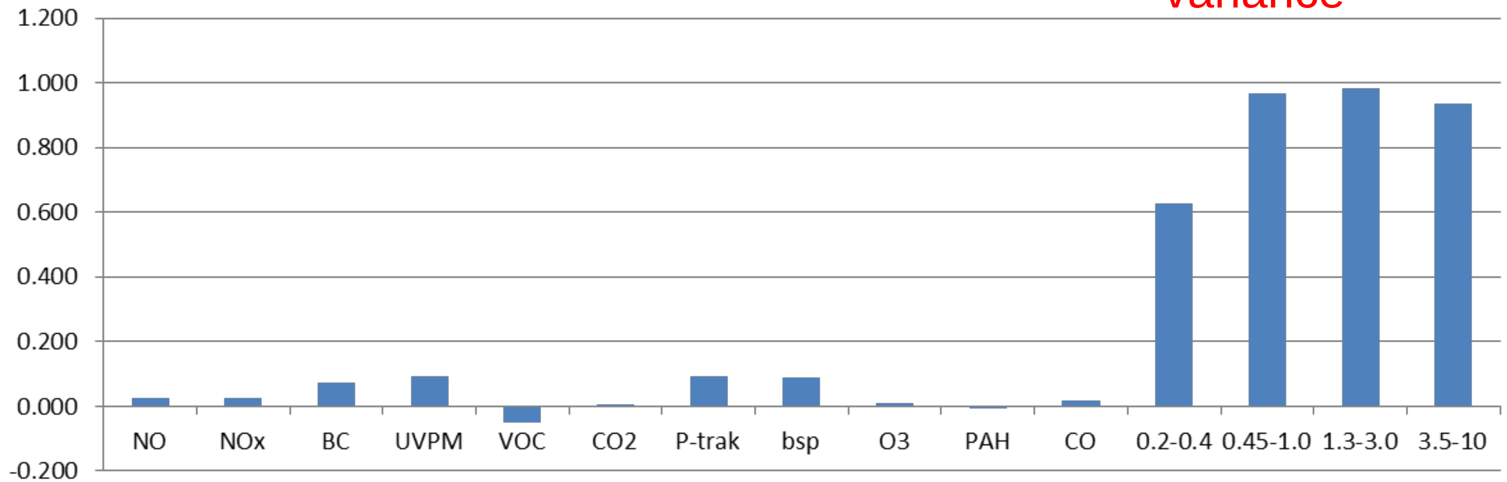
10 second sample #

A: At red light behind truck
B: Behind truck under freeway
C: Following school bus
D: At red light behind school bus

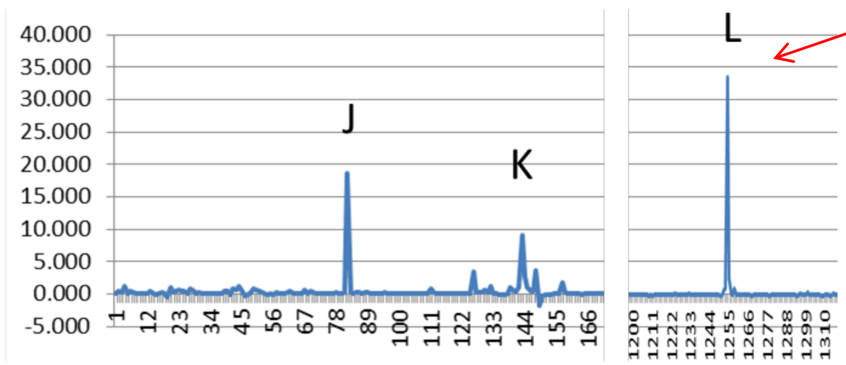
RC3: High Larger Particle Concentrations (PM Rich)

18% of total observed variance

Factor correlation with given species



Relative Factor Contribution



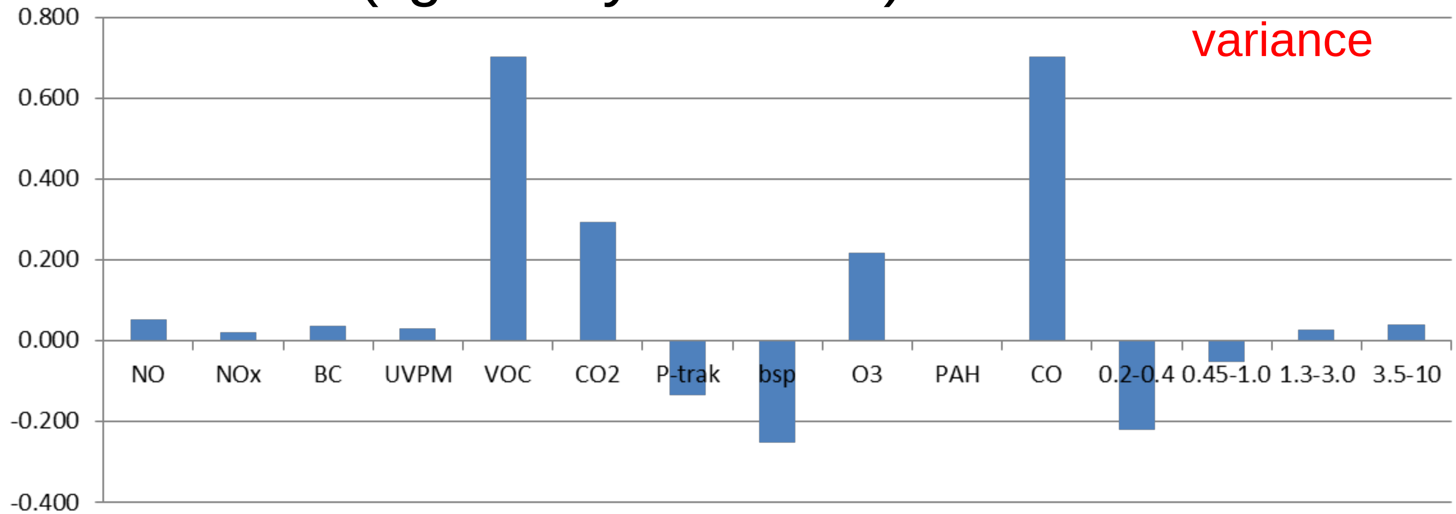
10 second sample #

J: sample inlet adjusted in field
K: behind school bus
L: dust plume from off road truck

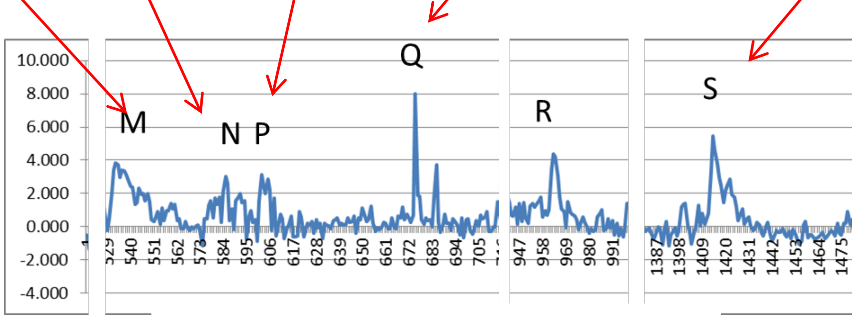
RC4: High VOC and CO correlation (light duty vehicles)

8% of total observed variance

Factor correlation with given species



Relative Factor Contribution



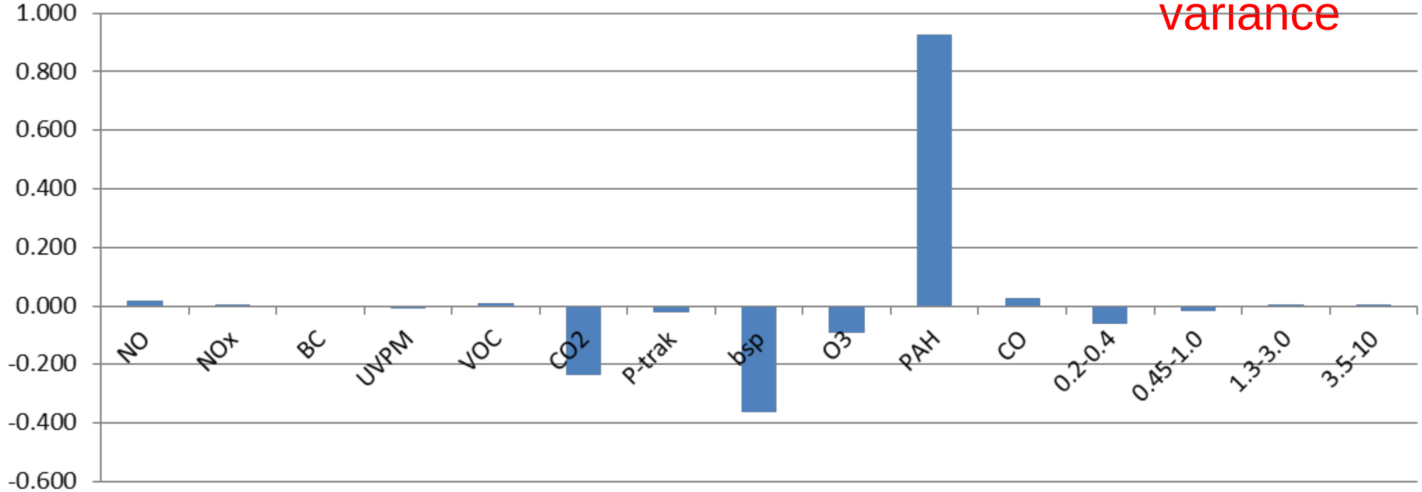
10 second sample #

M: next to minivan uphill
N: cold start vehicle
P: pickup truck
Q: industrial site (no CO)
R: residential street
S: traffic pulse in opposite direction after stop light

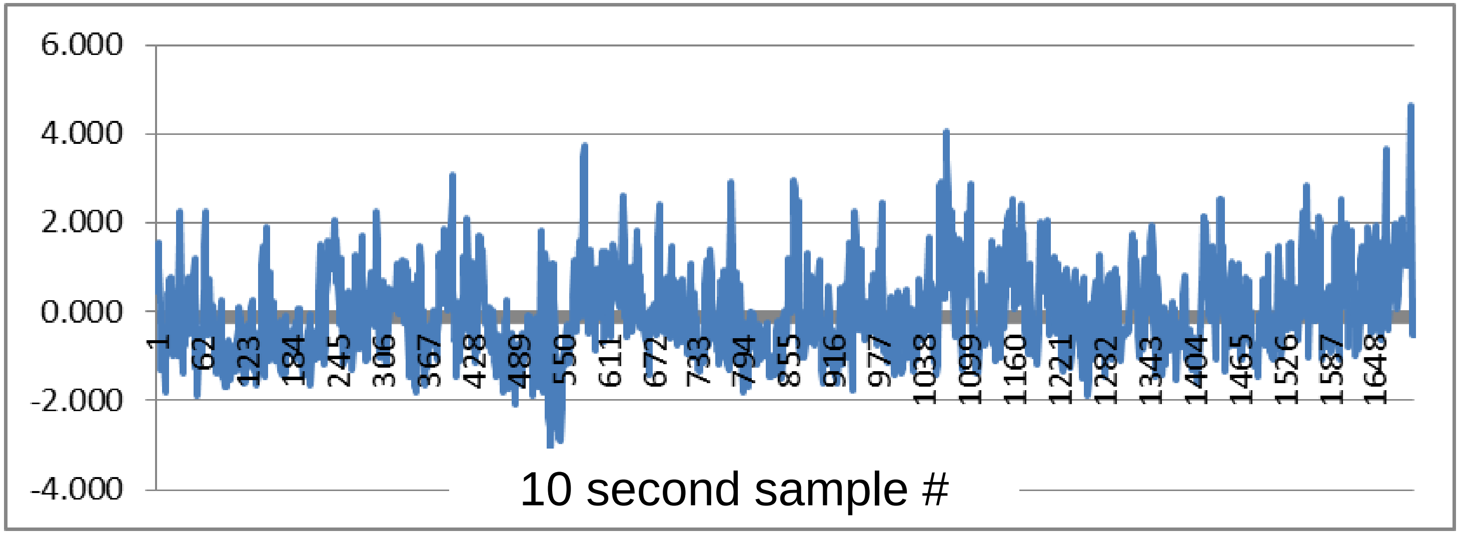
Particle-bound PAHs

7% of total observed variance

Factor correlation with given species



Relative Factor Contribution

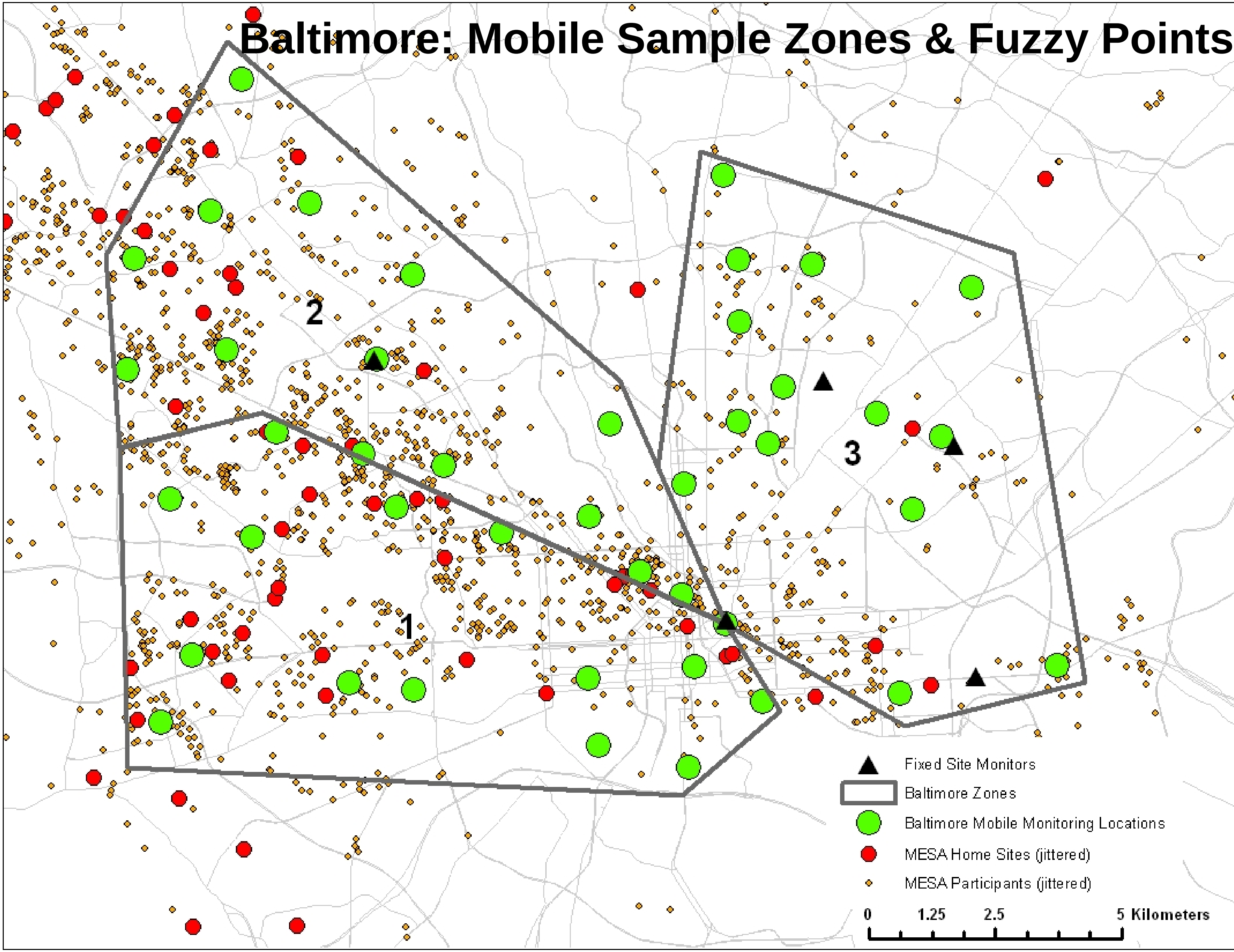


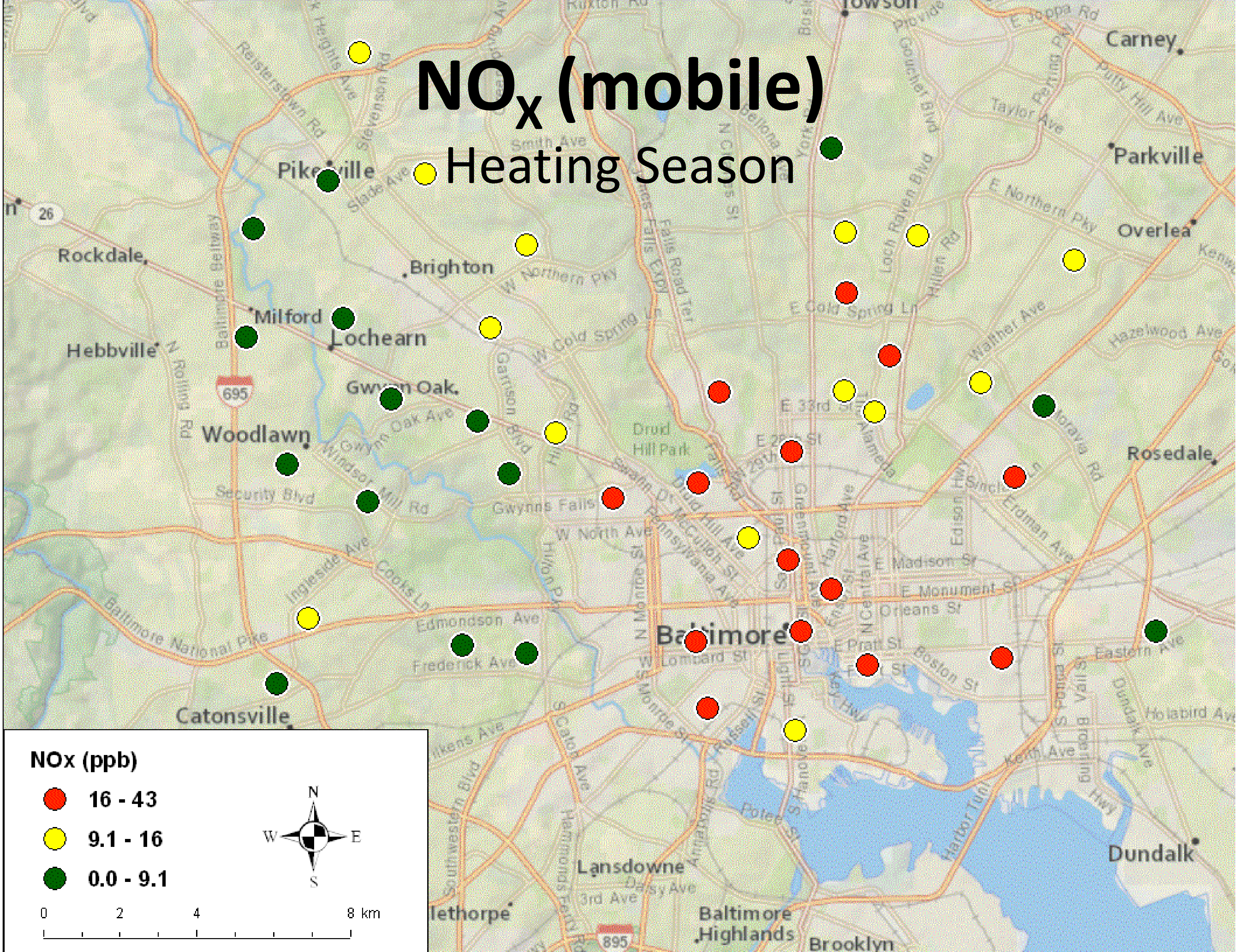
Analysis of 2-week Fuzzy Point Data

- How do we use all the mobile data over a 2-week measuring campaign in a city?
- Can we combine the passive badge data with the mobile data?
- Are there multivariate features?

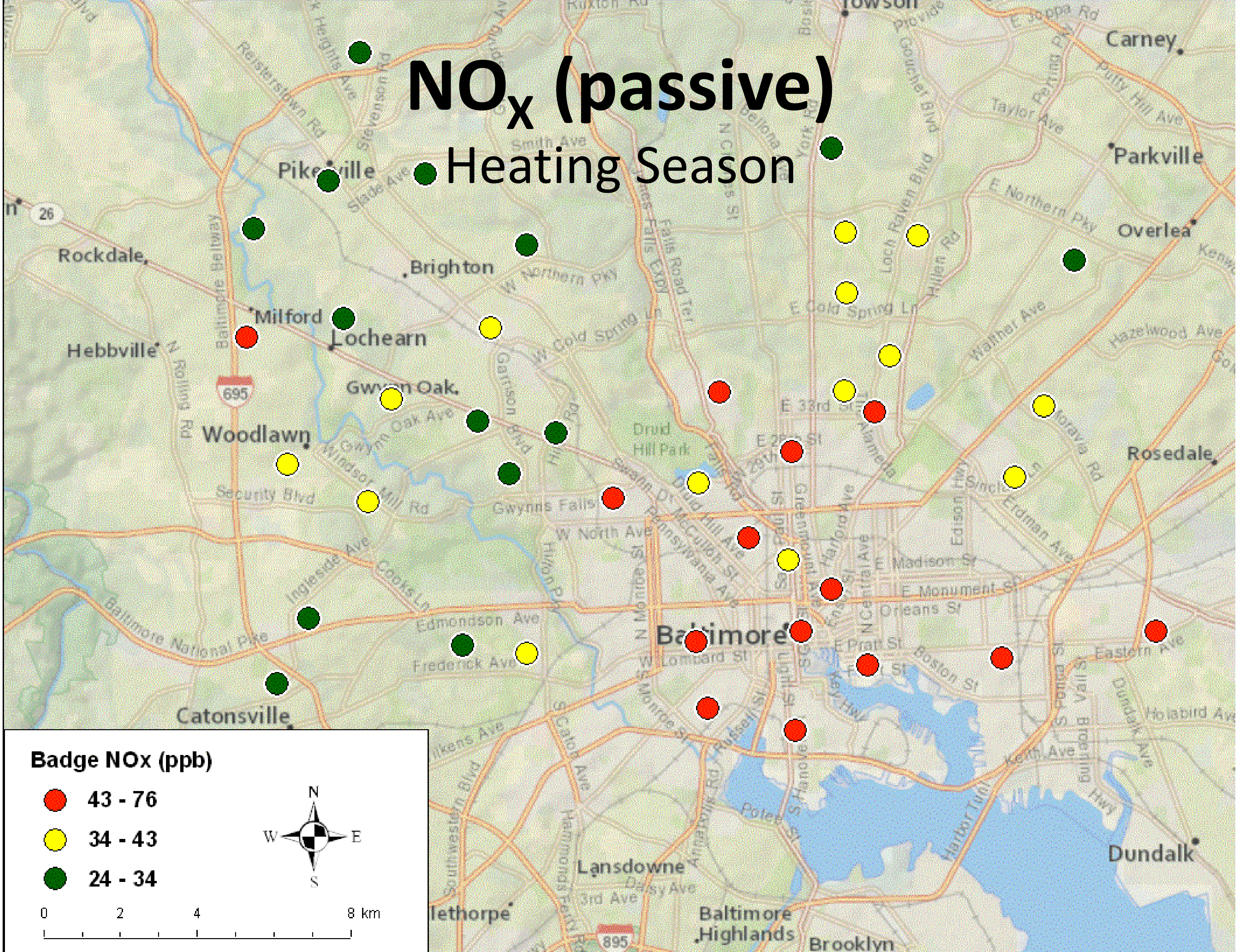
Baltimore Heating Season

- Previous approach: univariate analysis
  - Mobile data time-series with spatial reference site
  - Univariate 10-sec. values time-adjusted relative to fixed site over all routes for the 2-week sample period
- Segmented to data into fuzzy points
- Extracted 2-week medians for each pollutant
  - Plotted medians as Tertiles (High →Low)


Baltimore: Mobile Sample Zones & Fuzzy Points
2
3
1
Fixed Site Monitors
Baltimore Zones
Baltimore Mobile Monitoring Locations
MESA Home Sites (jittered)
MESA Participants (jittered)
0
1.25
2.5
5 Kilometers



NOx (mobile)
Heating Season
NOx (ppb)
16 - 43
9.1 - 16
0.0 - 9.1
N
W
E
S
0
2
4
8 km
Carney
Parkville
Overlea
Rosedale
Dundalk
Pikesville
Rockdale
Brighton
Milford
Lochearn
Hebbville
Gwynn Oak
Woodlawn
Catonsville
Baltimore
Lansdowne
Baltimore Highlands
Brooklyn
Towson
Druid Hill Park



NOx (passive)
Heating Season
Badge NOx (ppb)
43 - 76
34 - 43
24 - 34
N
W
E
S
0
2
4
8 km
Carney
Towson
Parkville
Overlea
Pikesville
Rockdale
Brighton
Milford
Lochearn
Hebbville
Gwynn Oak
Woodlawn
Rosedale
Baltimore
Catonsville
Lansdowne
Dundalk
Baltimore Highlands
Brooklyn

Preliminary multivariate analysis of Baltimore fuzzy point data

- Build on the previous univariate analysis
  - Use time-adj. mobile data at fuzzy points;
- Extracted mobile campaign medians for each pollutant at each fuzzy point;
- **Combined** mobile medians with passive badge data from the same points;
- Performed PCA analysis, plot loadings and scores

Aim 2

Data inputs: passive + mobile data (2-week)

- O_3 — by-product polutant
- vehical emissions
  - NO_2
  - NO_x
  - SO_2*

* not detected

Mobile Platform Addition

Time-adjusted medians

- NO_x
- NO_2
- NO
- O_3
- BC
- 25 - 100 nm
- 100 - 250 nm
- 250 nm - 1 um

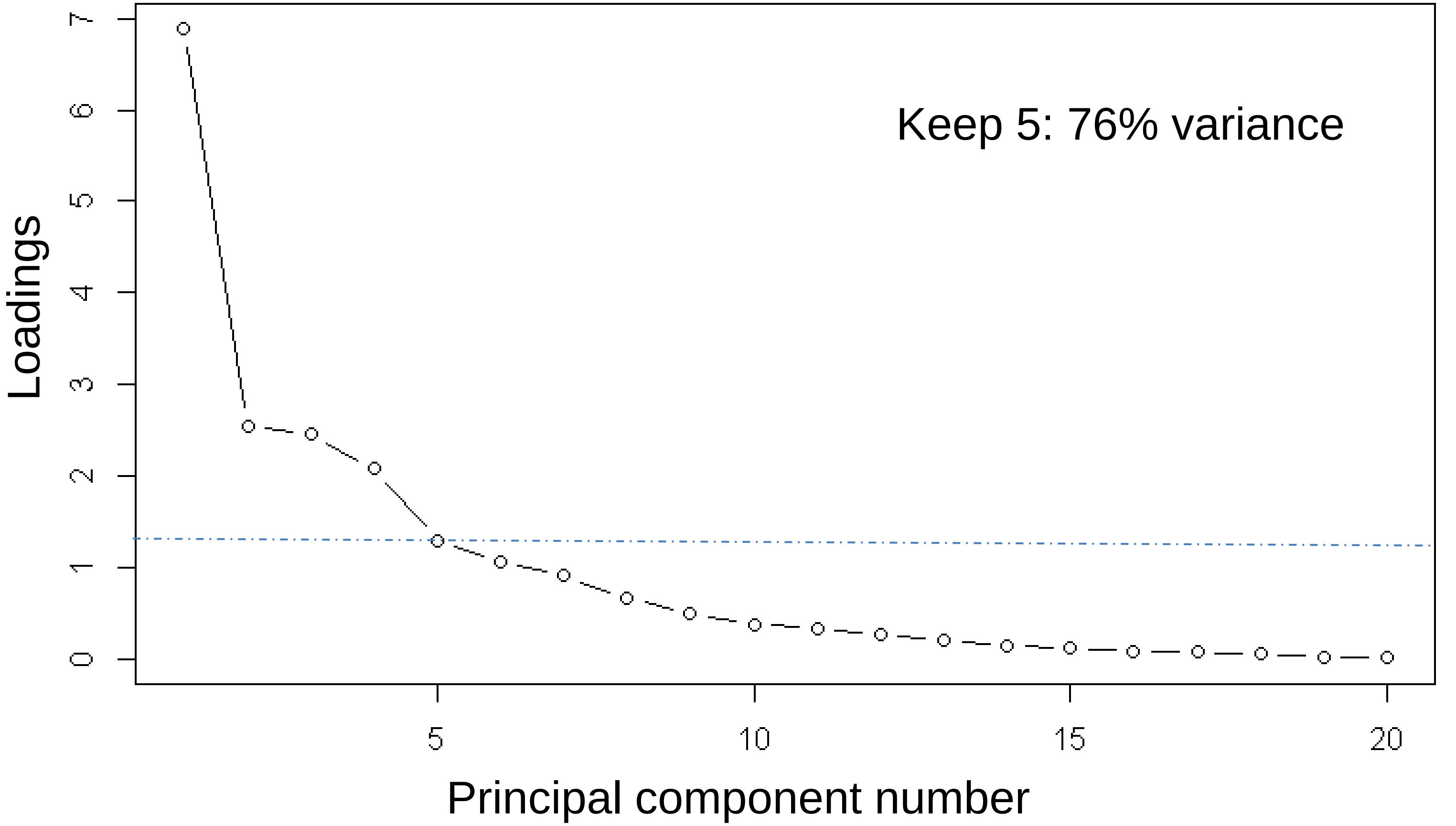
Enriched In Diesel Exhaust (DE) vs Gasoline Exhaust (GE)

- Decane
- Nonane
- m-Xylene
- Dodecane
- Undecane
- o-Xylene

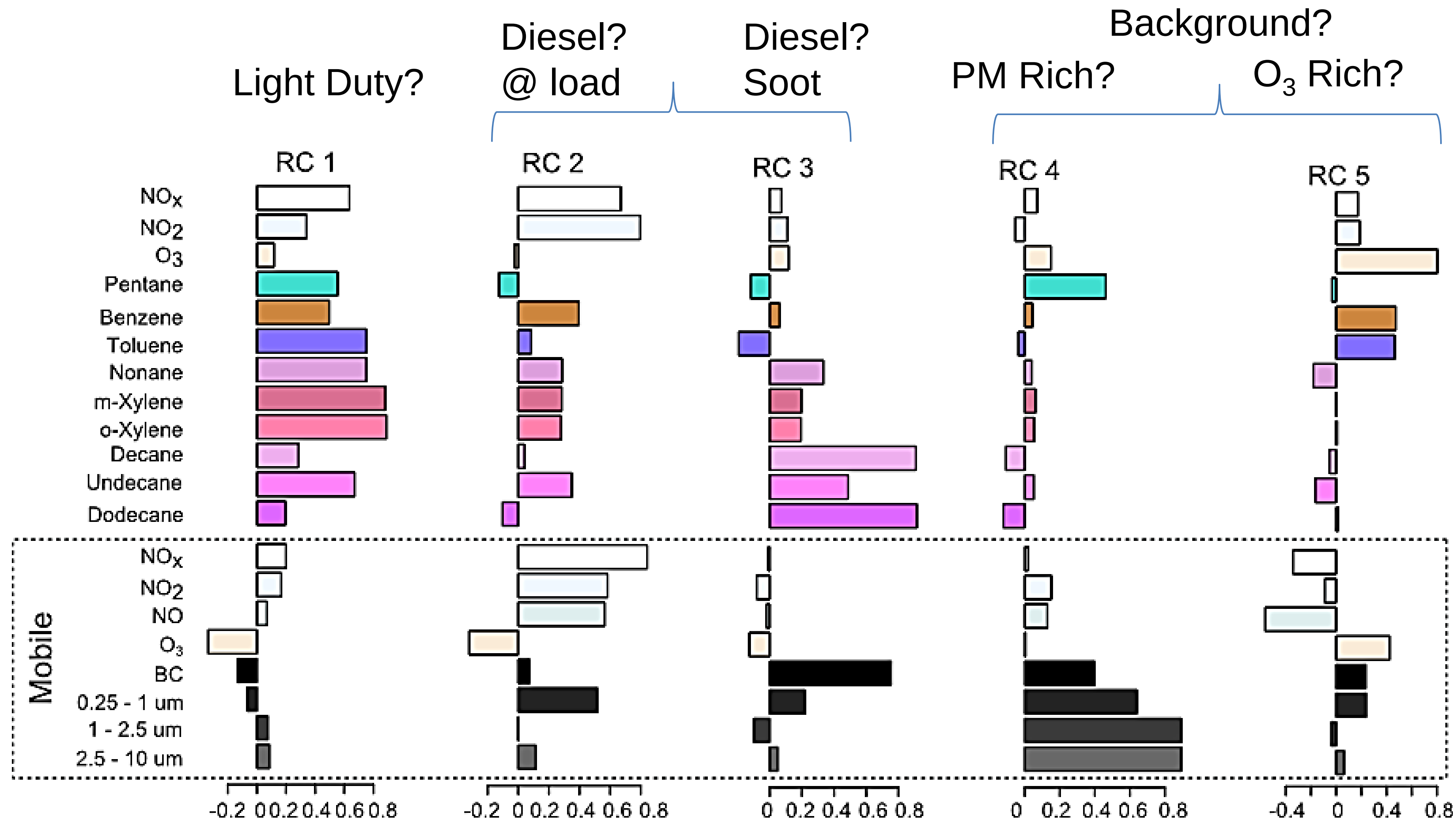
- Toluene — DE & GE
- Pentane — GE / biogenic / evaporative
- Isoprene — biogenic/ aged aerosol
- Benzene — exhaust / biomass smoke

Principal Components: Scree plot

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
SS loadings	6.88	2.55	2.46	2.07	1.30	1.06	0.90	0.65
Proportion Var	0.34	0.13	0.12	0.10	0.06	0.05	0.05	0.03
Cumulative Var	0.34	0.47	0.59	0.70	0.76	0.82	0.86	0.89

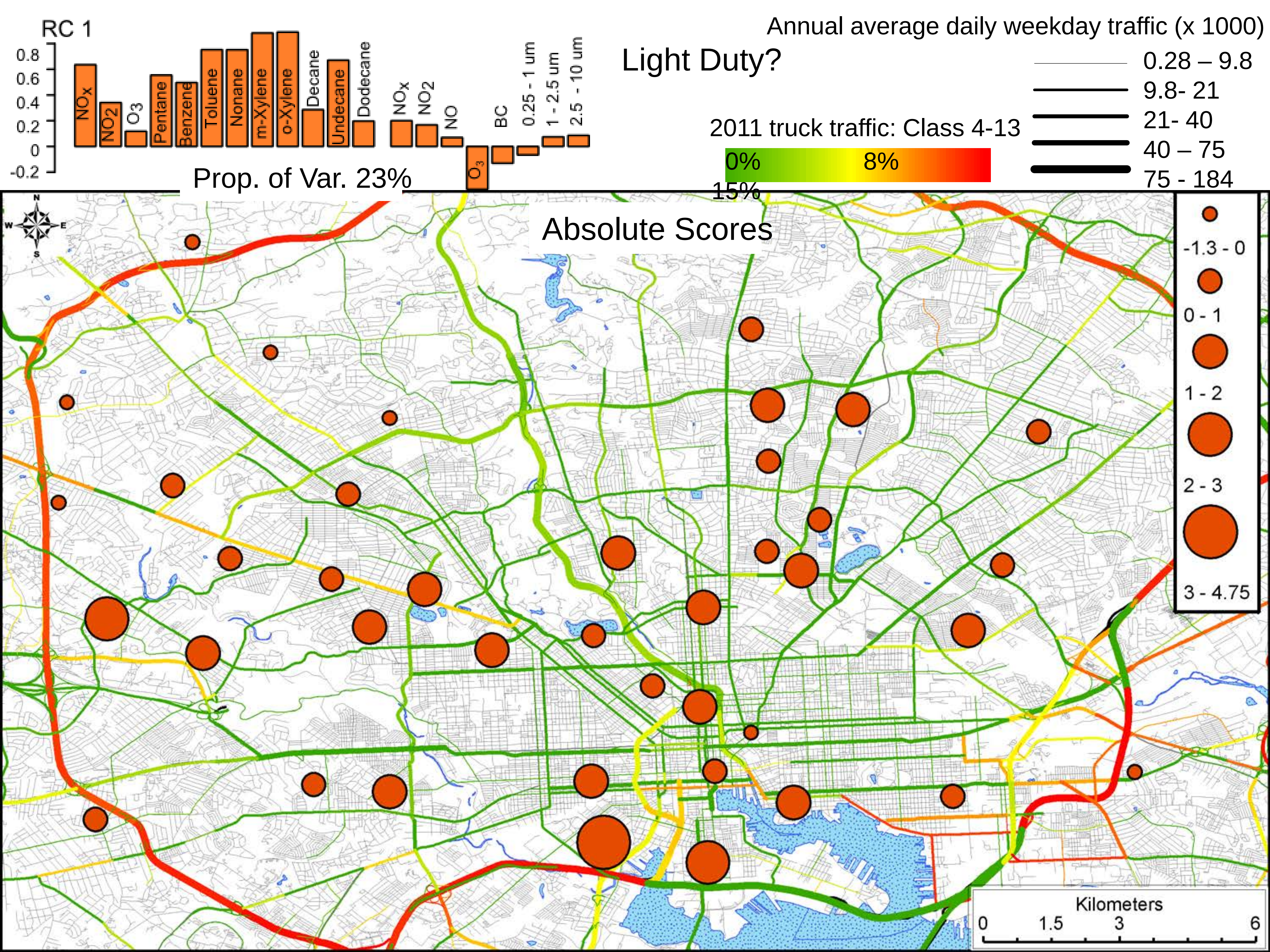


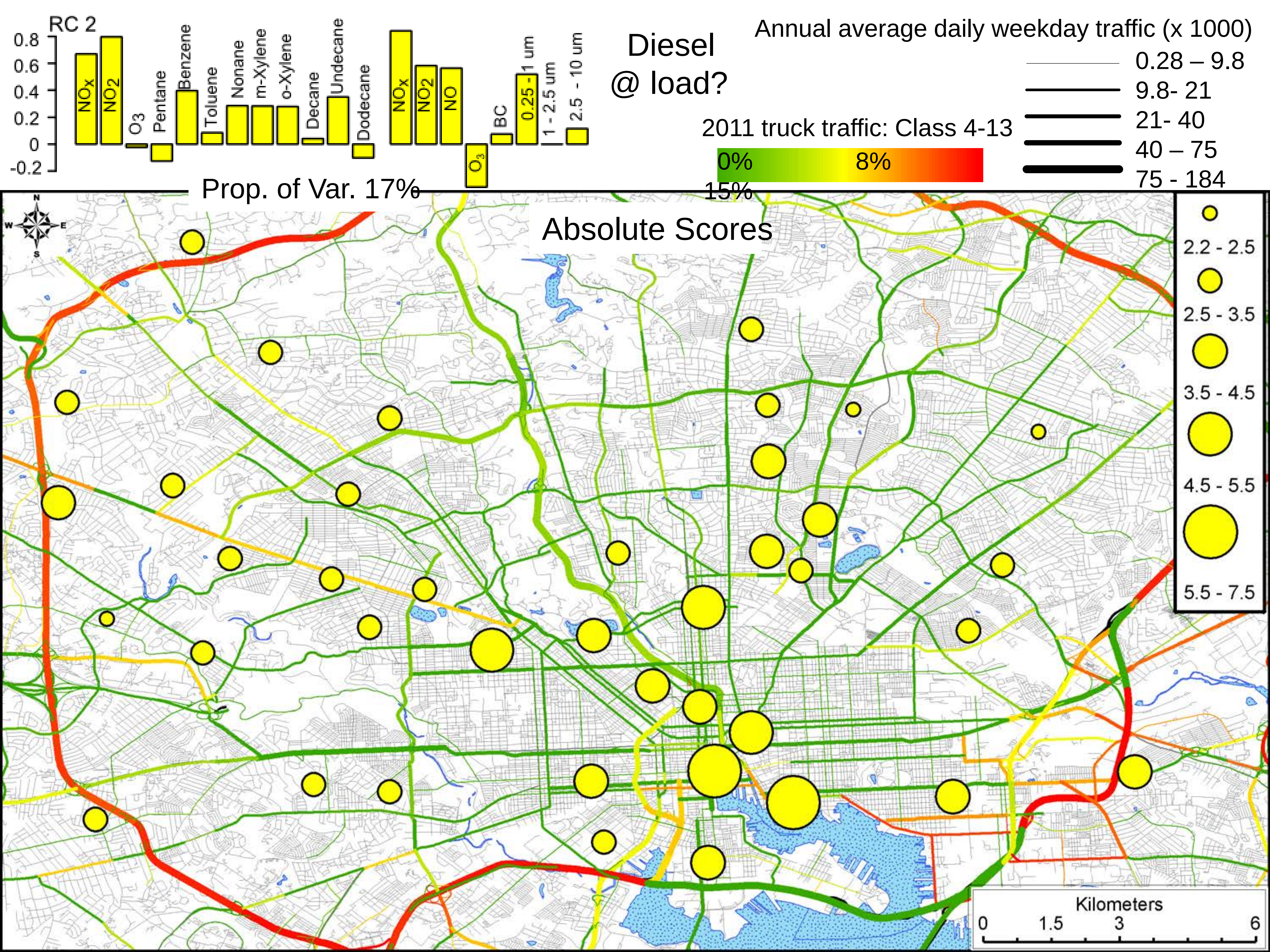
Varimax Rotated Components (RC)



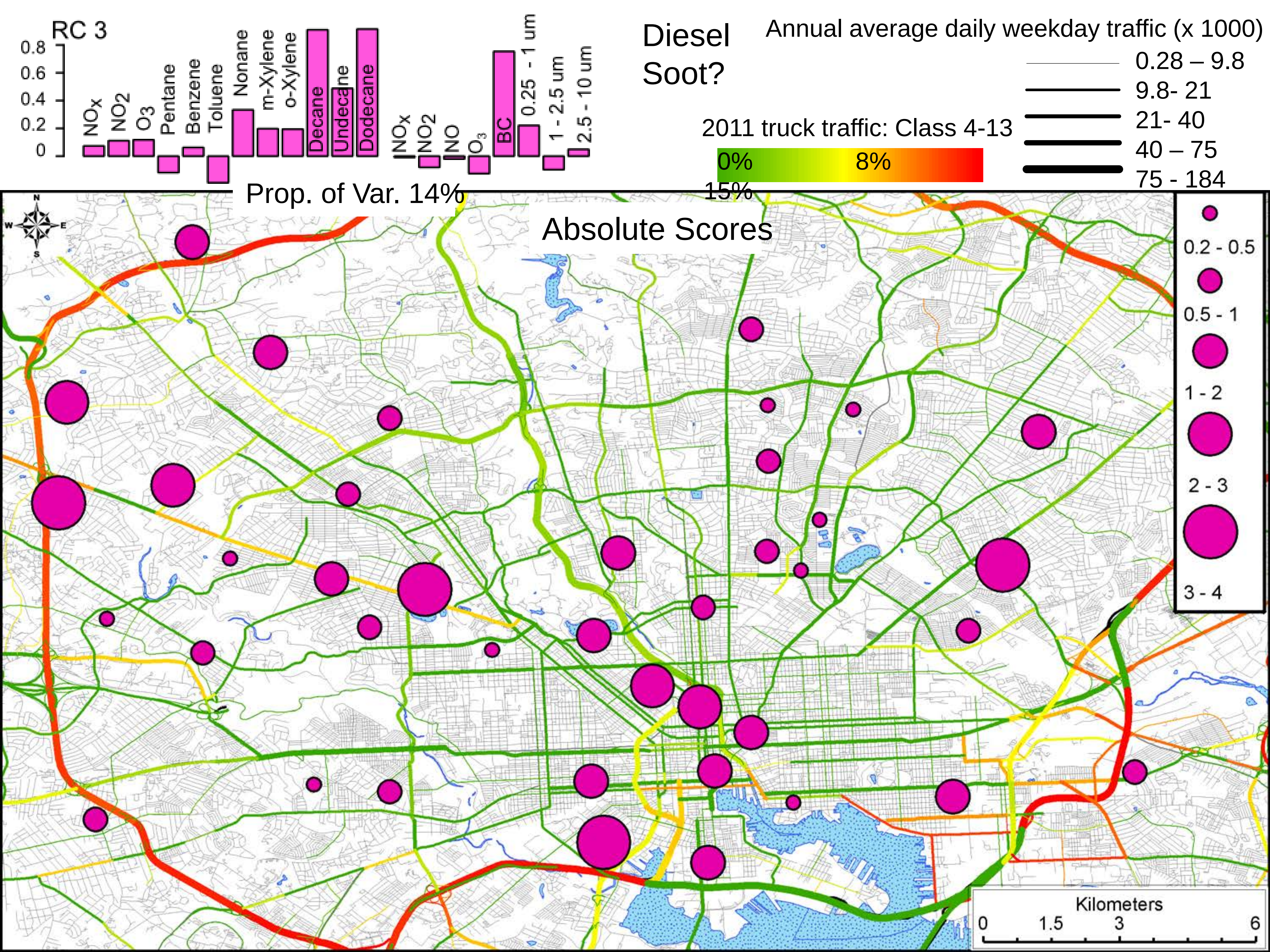
Prop. of Var.	23 %	17 %	14 %	12 %	10 %

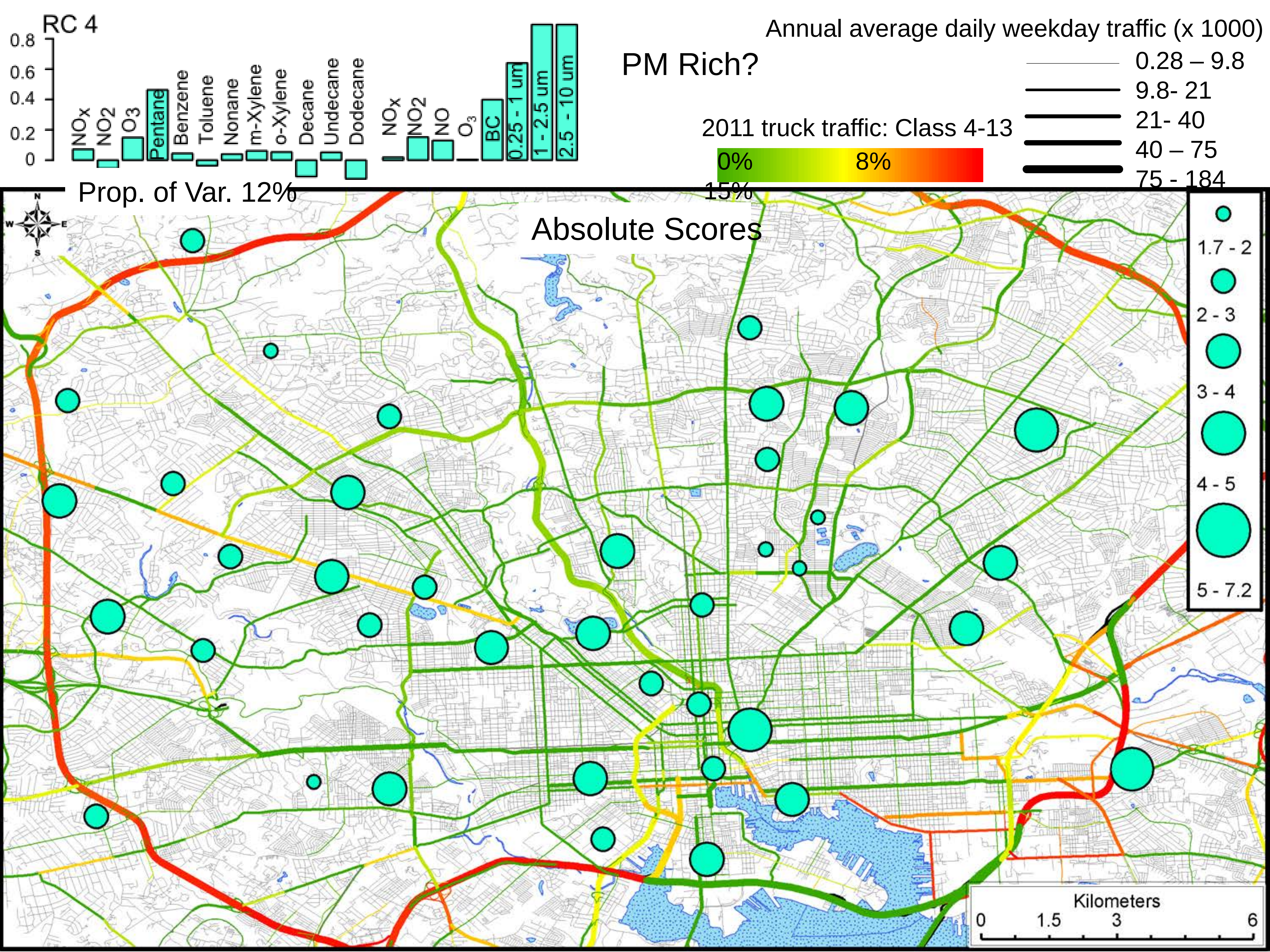
Total 76%


RC 1
0.8
0.6
0.4
0.2
0
-0.2
NOx
NO2
O3
Pentane
Benzene
Toluene
Nonane
m-Xylene
o-Xylene
Decane
Undecane
Dodecane
NOx
NO2
NO
O3
BC
0.25 - 1 um
1 - 2.5 um
2.5 - 10 um
Prop. of Var. 23%
Light Duty?
Annual average daily weekday traffic (x 1000)
0.28 – 9.8
9.8- 21
21- 40
40 – 75
75 - 184
2011 truck traffic: Class 4-13
0%
8%
15%
Absolute Scores
-1.3 - 0
0 - 1
1 - 2
2 - 3
3 - 4.75
Kilometers
0
1.5
3
6

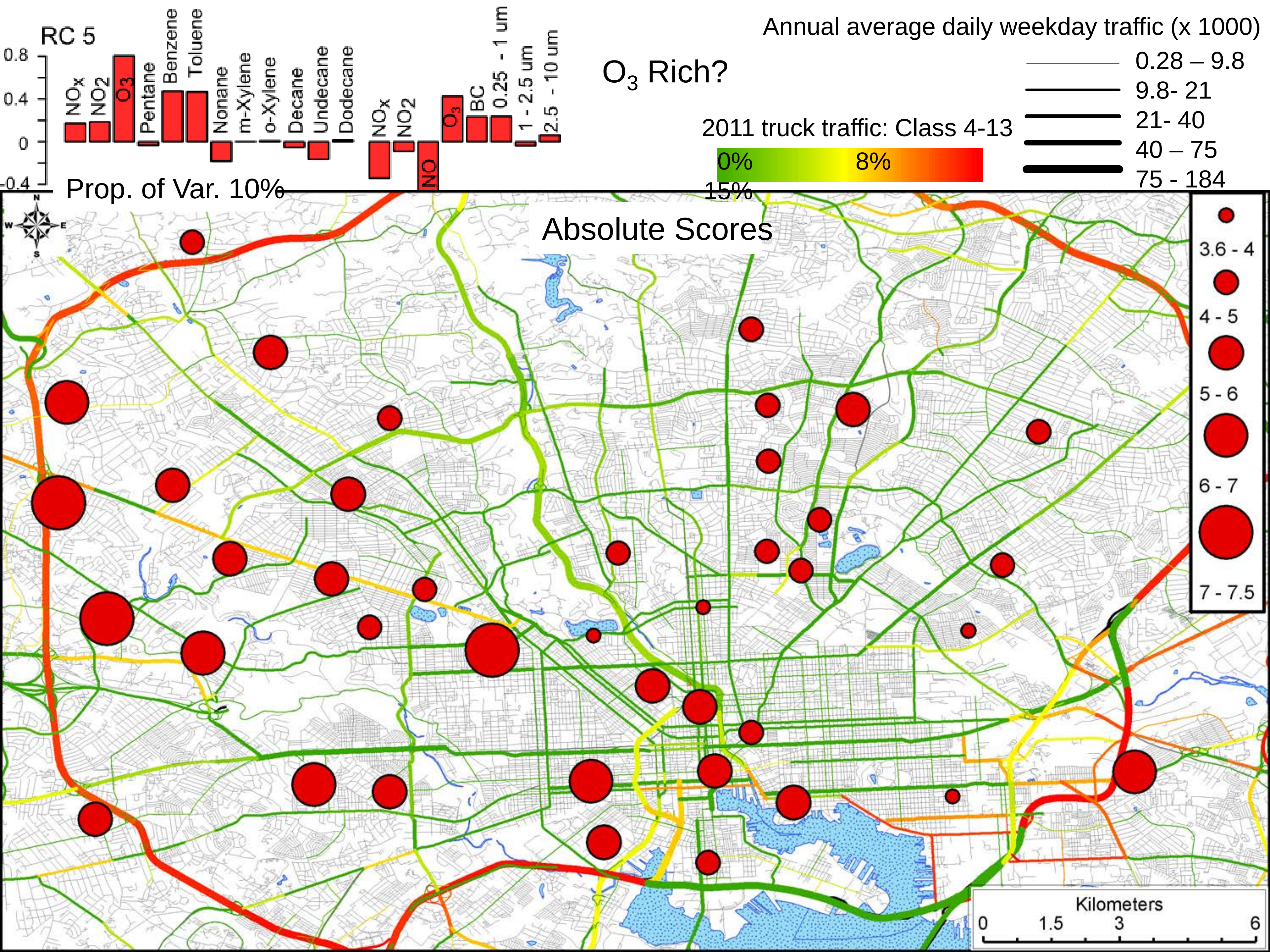


RC 2
0.8
0.6
0.4
0.2
0
-0.2
NOx
NO2
O3
Pentane
Benzene
Toluene
Nonane
m-Xylene
o-Xylene
Decane
Undecane
Dodecane
NOx
NO2
NO
O3
BC
0.25 - 1 um
1 - 2.5 um
2.5 - 10 um
Prop. of Var. 17%
Diesel
@ load?
Annual average daily weekday traffic (x 1000)
0.28 – 9.8
9.8- 21
21- 40
40 – 75
75 - 184
2011 truck traffic: Class 4-13
0%
8%
15%
Absolute Scores
2.2 - 2.5
2.5 - 3.5
3.5 - 4.5
4.5 - 5.5
5.5 - 7.5
Kilometers
0
1.5
3
6
N
W
E
S


RC 3
0.8
0.6
0.4
0.2
0
NOx
NO2
O3
Pentane
Benzene
Toluene
Nonane
m-Xylene
o-Xylene
Decane
Undecane
Dodecane
NOx
NO2
NO
O3
BC
0.25 - 1 um
1 - 2.5 um
2.5 - 10 um
Prop. of Var. 14%
Diesel Soot?
Annual average daily weekday traffic (x 1000)
0.28 – 9.8
9.8- 21
21- 40
40 – 75
75 - 184
2011 truck traffic: Class 4-13
0%
8%
15%
Absolute Scores
0.2 - 0.5
0.5 - 1
1 - 2
2 - 3
3 - 4
Kilometers
0
1.5
3
6



RC 4
0.8
0.6
0.4
0.2
0
NOx
NO2
O3
Pentane
Benzene
Toluene
Nonane
m-Xylene
o-Xylene
Decane
Undecane
Dodecane
NOx
NO2
NO
O3
BC
0.25 - 1 um
1 - 2.5 um
2.5 - 10 um
Prop. of Var. 12%
PM Rich?
2011 truck traffic: Class 4-13
0%
8%
15%
Annual average daily weekday traffic (x 1000)
0.28 – 9.8
9.8- 21
21- 40
40 – 75
75 - 184
Absolute Scores
1.7 - 2
2 - 3
3 - 4
4 - 5
5 - 7.2
Kilometers
0
1.5
3
6
N
S
E
W



RC 5
0.8
0.4
0
-0.4
NOx
NO2
O3
Pentane
Benzene
Toluene
Nonane
m-Xylene
o-Xylene
Decane
Undecane
Dodecane
NOx
NO2
NO
O3
BC
0.25 - 1 um
1 - 2.5 um
2.5 - 10 um
Prop. of Var. 10%
O3 Rich?
Annual average daily weekday traffic (x 1000)
0.28 – 9.8
9.8- 21
21- 40
40 – 75
75 - 184
2011 truck traffic: Class 4-13
0%
8%
15%
Absolute Scores
3.6 - 4
4 - 5
5 - 6
6 - 7
7 - 7.5
Kilometers
0
1.5
3
6

Next steps: PCA analysis - moving forward

- Examine univariate time adjustments to the 10 sec data
  - (What time adjustment is needed? Alternative – smoothing?)
- Segmenting data into fixed spatial buffers around the fuzzy points (FP-subset);
- Performing PCA analysis of FP-subset to extract factors;
  - Robust/ Sparse PCA; joint PCA w spatial splines etc.
  - Identify sources from PCA factor loadings (video info);
- Combine PCA scores w LUR to assess spatial structure;
- Work with Biostatistics Core to perform additional analysis of the time-series data set (see posters)

Conclusions

- Observed pollutant gradients within the first 100 meters downwind of major roadway
- Achieved temporally and spatially resolved measurements of multi-pollutant mixtures
- Observed multivariate features of traffic-related and secondary pollutants at fuzzy points

Provides us with a better understanding of how to characterize multi-pollutants for future studies

Thank You!

CENTER FOR CLEAN AIR RESEARCH

UNIVERSITY *of* WASHINGTON

CCAR

Center for Clean Air Research

Project 2: Simulated Roadway Exposure Atmospheres for Laboratory Animal and Human Studies

www.LRRI.org

SCHOOL OF PUBLIC HEALTH
UNIVERSITY of WASHINGTON

THE UNIVERSITY of NEW MEXICO

WASHINGTON STATE UNIVERSITY

Lovelace™

Objectives

- **Simulate ambient exposures in the laboratory**
  - Bridge these exposures to ambient measurements/modeling (Project 1)
- **Compare toxicity of exposures**
  - Use these results to determine mechanisms (Project 3) and to define priorities and atmospheres for human exposures (Project 4)

Principle Activities Since Last ESAC

- Refinement of approaches to Irradiation Chamber Studies with Motor Vehicle Exhaust
  - Key challenge: modulation of NOX:VOC:PM ratios with minimal particle loss
  - Key challenge : Conduct of long-term studies with constant atmospheres
- Assay development/analysis for short term determination of cardiovascular response and chemical drivers (Project 2-3 collaboration)
- Review of Project 1 Data to help inform atmosphere decisions
- STATUS: This past year we have made significant advances in the bioassays to enable reproducible, mechanistic based acute assays for screening of test atmospheres. This has been completed, and next steps are to screen new atmospheres. Focus this past year has been MVe and ozone studies

Key *Initial* Research Questions

- Does agglomeration and physical transformation of particulate motor vehicle emissions alter their toxicity (does size matter)?
  - COMMENT: tabled due to feasibility and other priorities
- Does chemical transformation, and formation of secondary organic aerosol from motor vehicle emission precursors, enhance or diminish the toxicity of roadway atmospheres?
  - COMMENT: delayed due to other priorities. Initial challenges observed. This has been overcome by technical approach to get irradiation chamber to work and to develop acute sensitive assays for comparisons.

Key *Initial* Research Questions

- Do ozone and other background co-pollutants alter or exacerbate the toxicity of motor vehicle emissions?
  - COMMENT: Campen has been characterizing response to ozone and this year also to MVe mixtures. We have not conducted a 'background' atmosphere as originally described. This is on-going
- Does road dust, a significant non-tailpipe roadway emission, confer any cardiovascular toxicity that may confound associations with tailpipe emissions?
  - COMMENT: we are including a road dust atmosphere alone and in combination for new trials with acute assays

Additional Goal: Bridge Atmospheres to Project 1

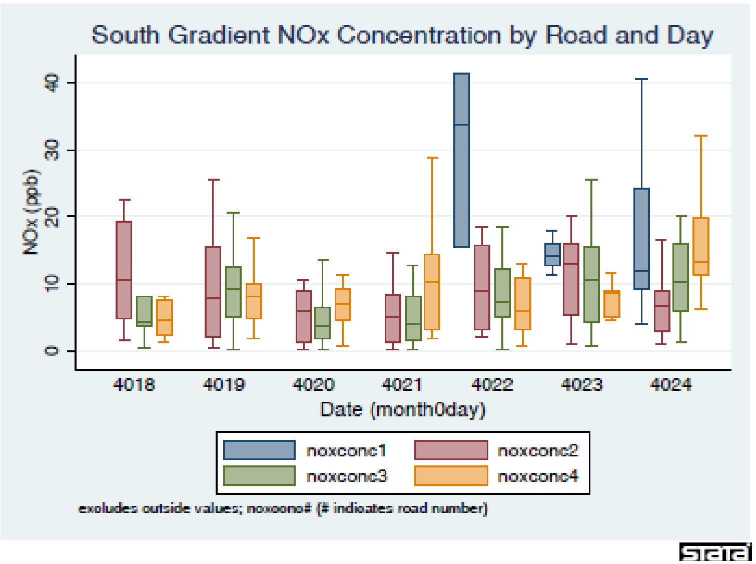
- During the past year, the analysis from the Albuquerque campaign from Project 1 was completed.
- Those results are defined here for review
- These are being considered to utilize for the following atmospheres:

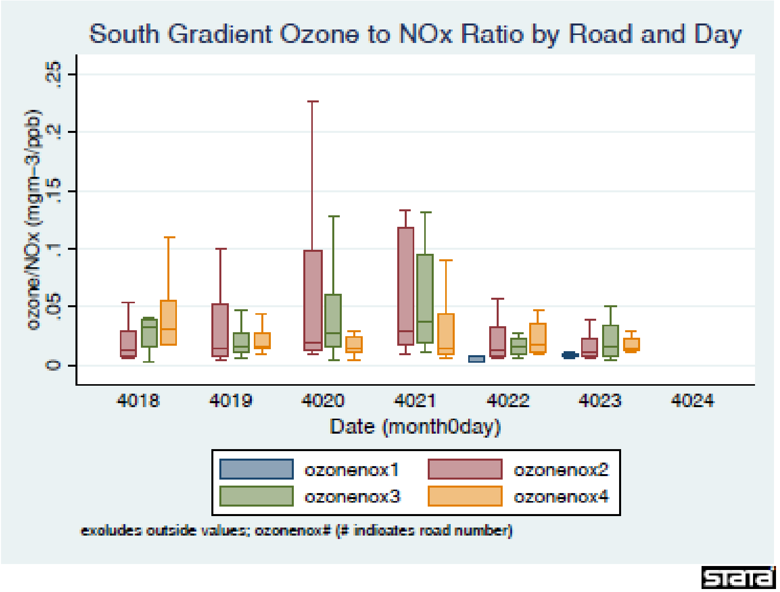
Traffic Transect Study from Project 1

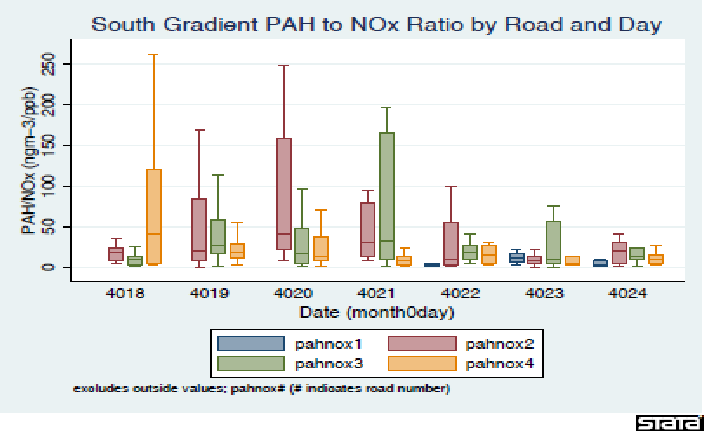
TABLE 1: Albuquerque Data Summary - All Gradient Roadways

	Mean	SD	Median	Percentiles			
				5%	25%	75%	95%
NO_x (ppb)	10.68	10.46	7.750	1.000	4.000	14.10	32.10
O_3/NO_x (mgm^{-3}/ppb)	0.03931	0.1133	0.01554	0.002768	0.008389	0.07412	0.1242
PAH/NO_x (ngm^{-3}/ppb)	46.30	190.0	13.41	1.682	5.584	30.39	148.9
CO_2 (ppm)	372.6	5.83	371.8	365.8	369.0	374.8	383.2

Results by Date (Example): NO







Outcome of analysis

- The gradients are not as large as one would like to tease out potential differences in biological response (the delta is actually pretty small)
  - Considering this:
    - How do we use monitoring data to create exposure atmospheres?
    - Is utilization of these data better then a default approach?

<u>Background</u> + <u>Traffic Emissions</u>

O_3, $(NH_4)_2SO_4$, NH_4NO_3, VOC, NI, V

Tailpipe, Evaporative, Tire & brake, Resuspended Dust

100 m 500 m 1 km ?? → <u>Exposures</u>

Distance From Roadway

Chemical Transformation

$OH^{\cdot}$, Sunlight

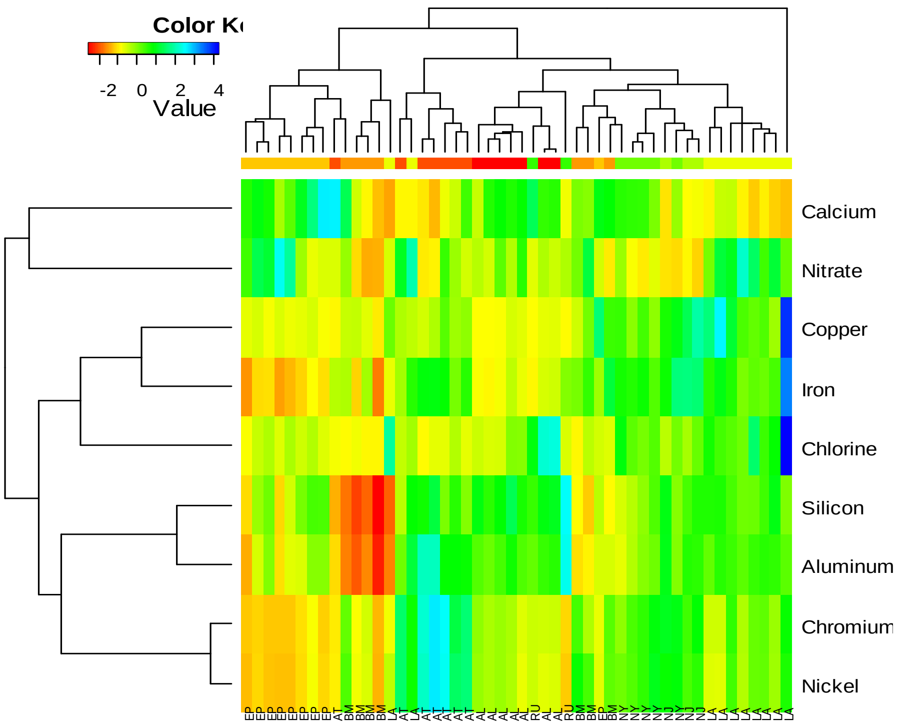
Aging

Nucleation, Agglomeration

Next Steps, Considering Prior Information

- Feedback from SAC appreciated
  - Default: MVe combinations with/without particles and gases and NOx permutations has been focus to date
  - New and on-going studies to continue to investigate MVe carbon/gas ratios and also to include:
    - Road dust (Urban composite)
    - MVe –Nox + Ozone (two different NOx/Ozone Gradations and two different Mve degradations
    - MVe + SOA
    - Urban background at same PM: Ammonium nitrate/ammonium sulfate, road dust; with and without ozone at same level of MVe

Other Updates



Part of center investigation was looking back further into data analysis of road dust. We presented that last year. These data are now in press at AQAH

1 additional manuscript submitted and 2 to be submitted next month on initial Assays with MVe and ancillary mixtures

Project 3: Cardiovascular Consequences of Immune Modification by Traffic-Related Emissions

Campen, Rosenfeld, Lund, McDonald

Project 3 Aims

- **Aim 1**, we will ascertain the potentiating effects of physical and photochemical aging on fresh emissions, in terms of driving this vascular oxidative stress.

- **In Aim 2, we will examine effects of the emissions-induced oxidative modifications to endogenous phospholipids, in terms of activating immune-modulating receptors such as LOX-1, CD-36, TLR-2, and TLR-4.**

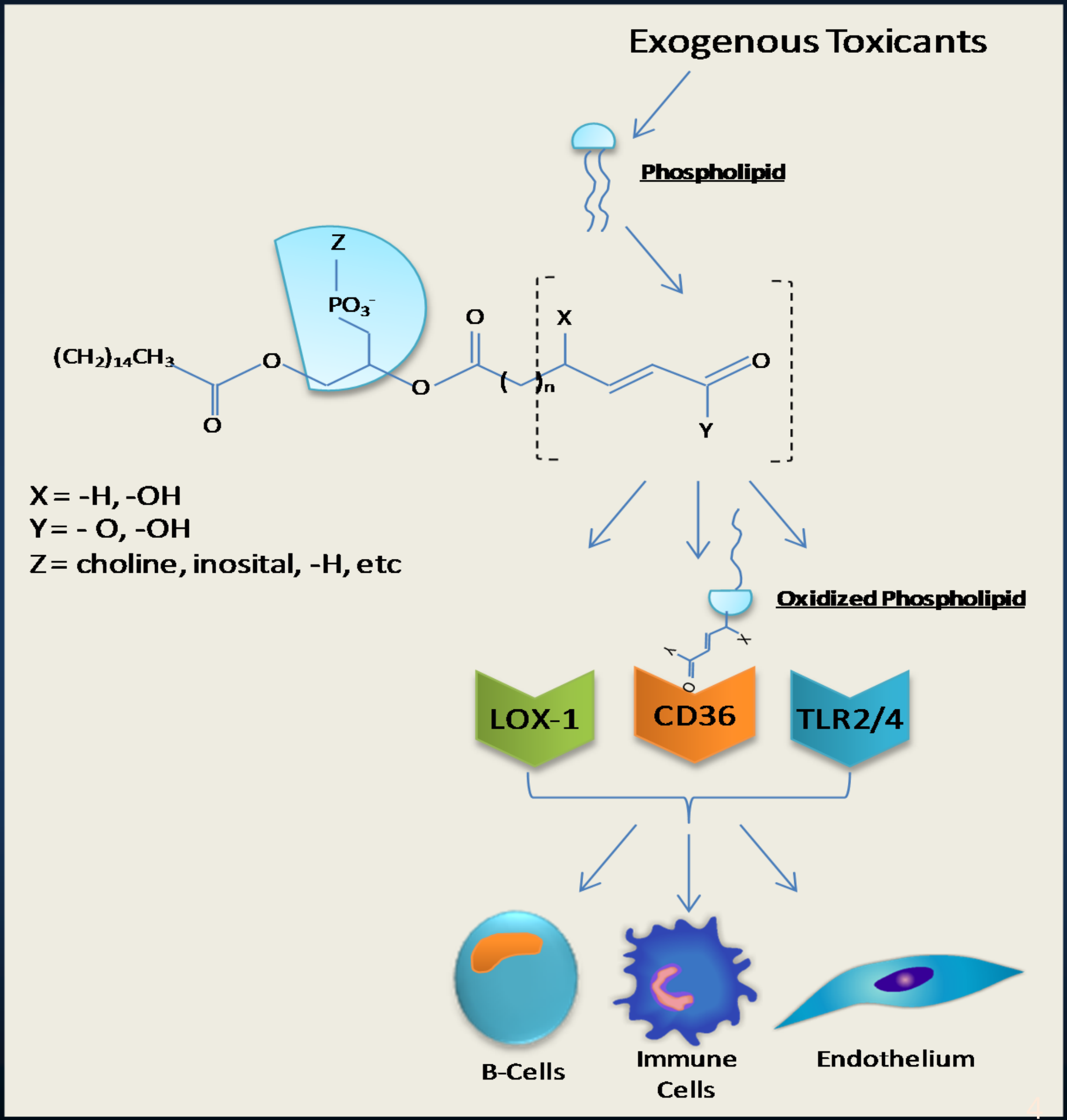
- **Aim 3**, we will further explore the role of specific immune cell populations as participants in the innate and adaptive responses to emissions-induced phospholipid modifications.

Comments on Progress for the Past Year

- Because the photochemical transformations led to a dynamic atmosphere, chronic exposures would be challenging to conduct/analyze
  - Necessitated a move to more acute outcomes
    - But still consistent with chronic vascular disease pathogenesis
- Emphasis from SAC to explore the bloodborne vasoactive factors

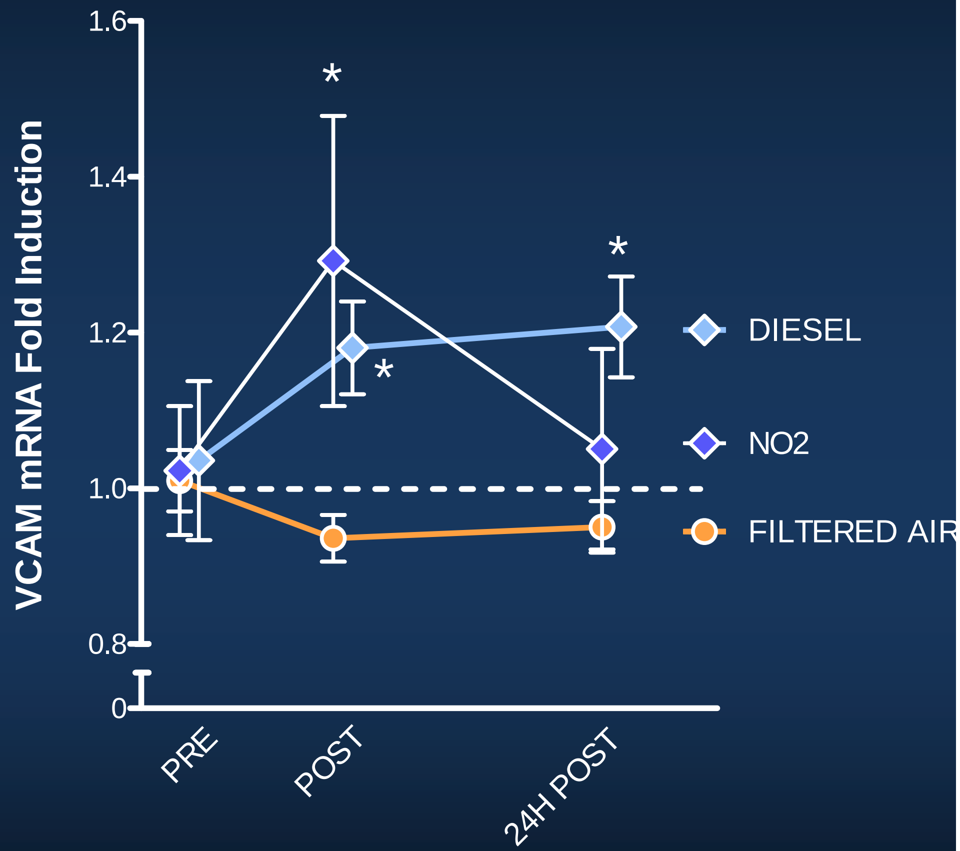
Hypothesis: Chemical modification of phospholipids ^*and/or proteins* leads to altered biological signaling

Specifically, we suspect that reaction by-products from the lung enter the circulation to drive <u>endothelial cell activation</u> via cell surface receptors, especially pattern recognition receptors such as CD36, TLR4, and LOX-1



Evidence that the signal is blood-borne:

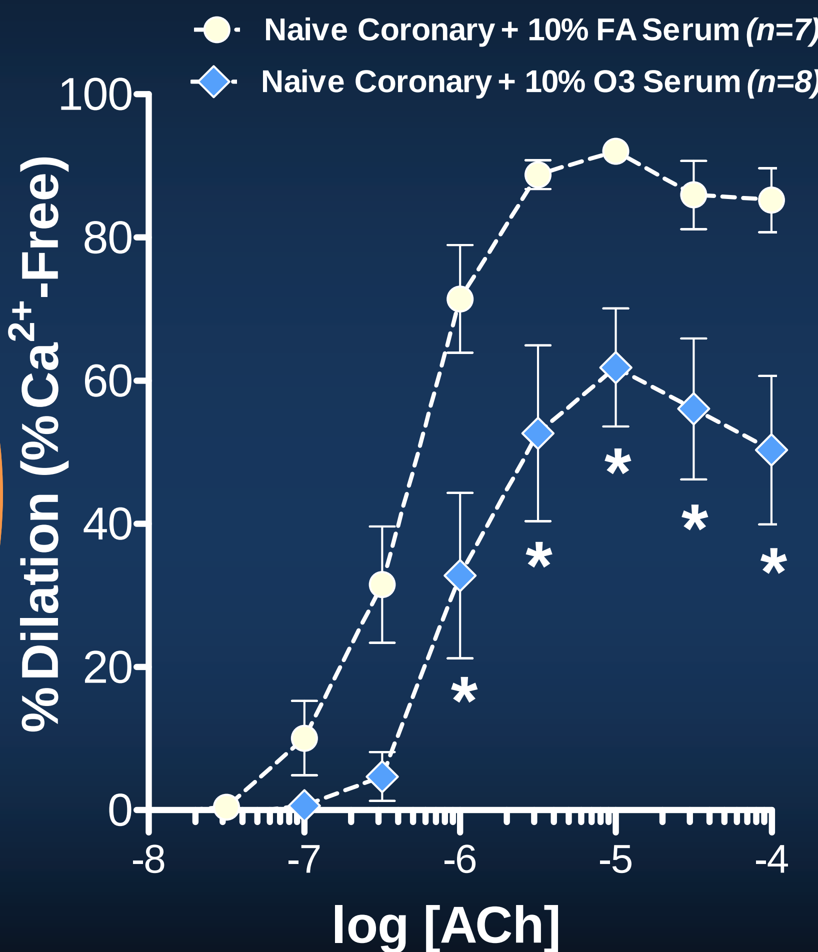
- Plasma from humans exposed to NO_2, Diesel, or filtered air (control) for 2 h
- Incubated with primary human coronary artery endothelial cells at 10% in media
- Plasma after exposures induced ICAM, VCAM, P-selectin and IL-8



Channell et al., Tox Sci, 2012

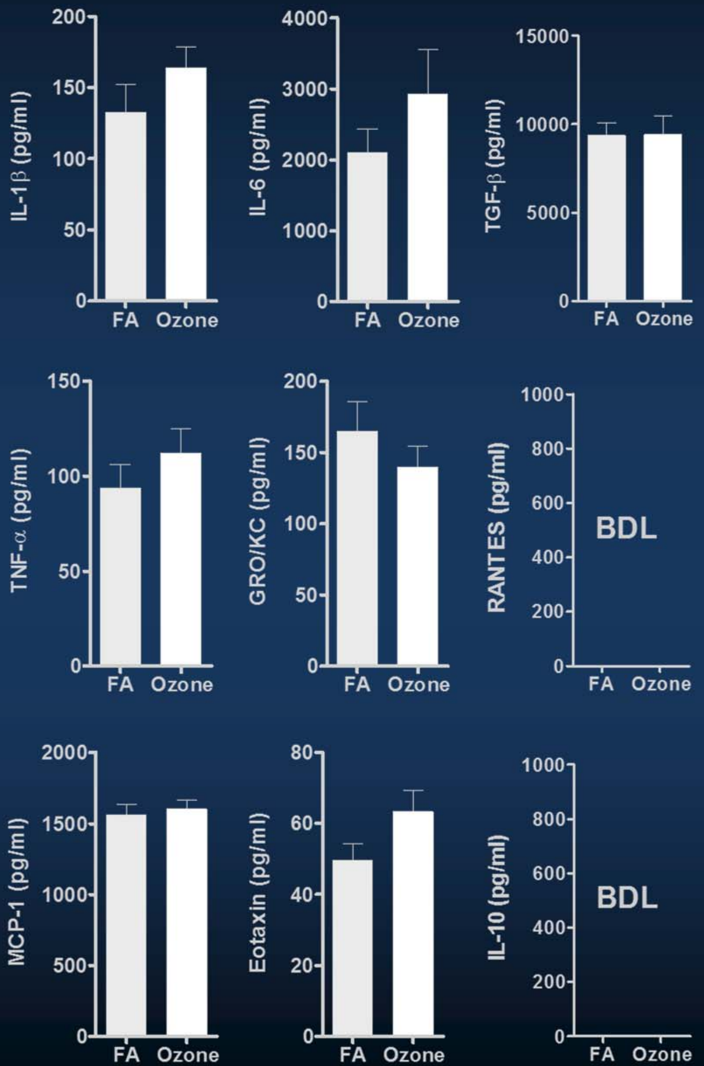
Serum from Ozone-Exposed Rats Impairs Vasodilation *Ex Vivo*

- 1ppm O_3 x 4h in rats, serum harvested 24h-post
- Infusion of a dilute (10%) serum in the lumen of isolated coronary arteries leads to impaired vasodilation
- No effect of serum from air-exposed rats (control)



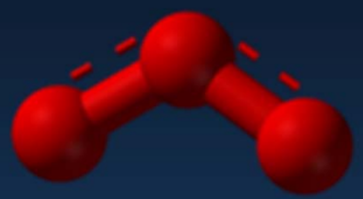
None of the Changes Appear to be "Usual Suspect" Cytokines

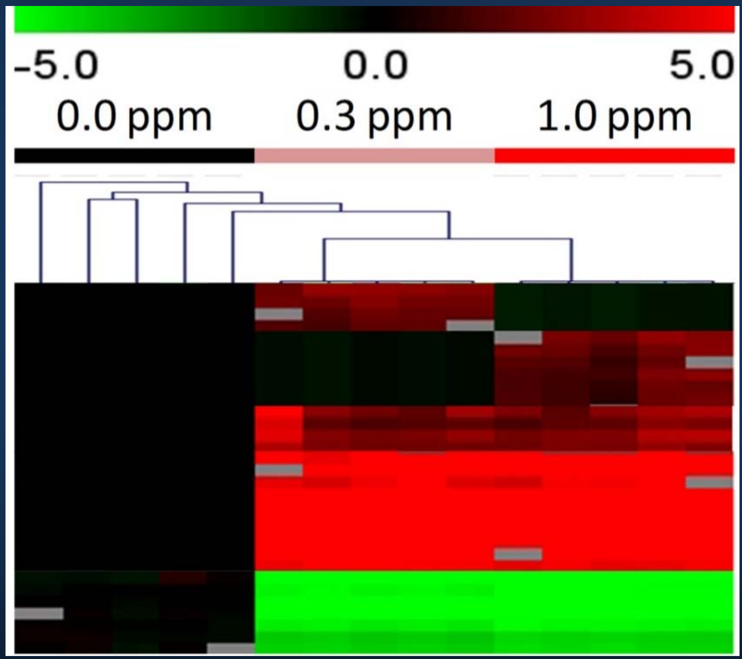
- Cytokines do not seem to be relevant
  - GM-CSF, IL-10, RANTES all BDL
- Studies that show increased circulating cytokines are either
  - Very high concentrations
  - Instillation
  - Increases are modest, not physiologically significant



What might be in the blood?

- After O_3 exposure, **<u>1,480 unique changes</u>** were identified in rat serum in the <u><5kD fraction</u>
- We propose that oxidative modifications of endogenous molecules leads to reactive epitopes for Pattern Recognition Receptors

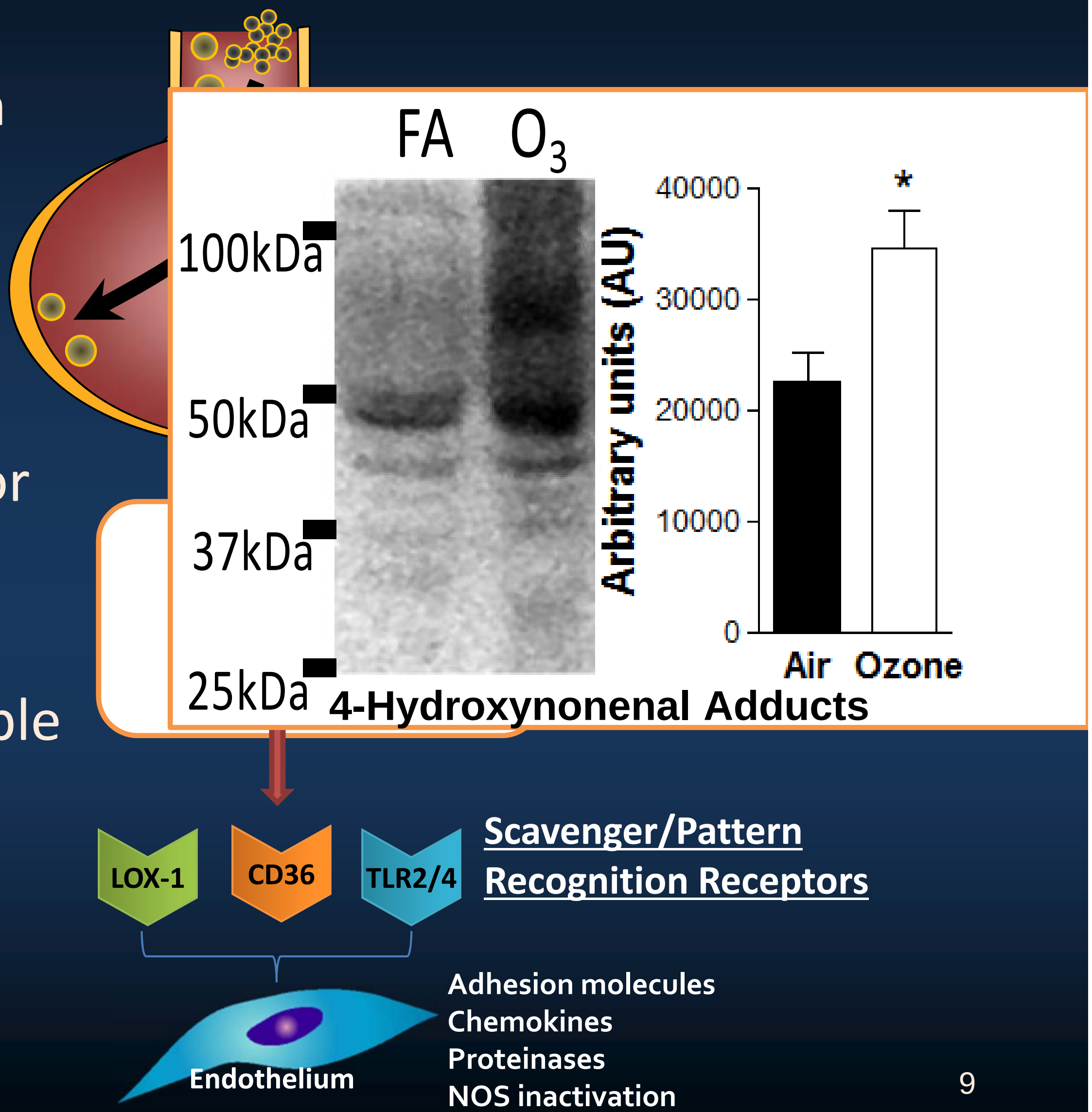




Courtesy Andrew Ottens, VCU

Hypothesis: Pollution Induces Molecular Shrapnel

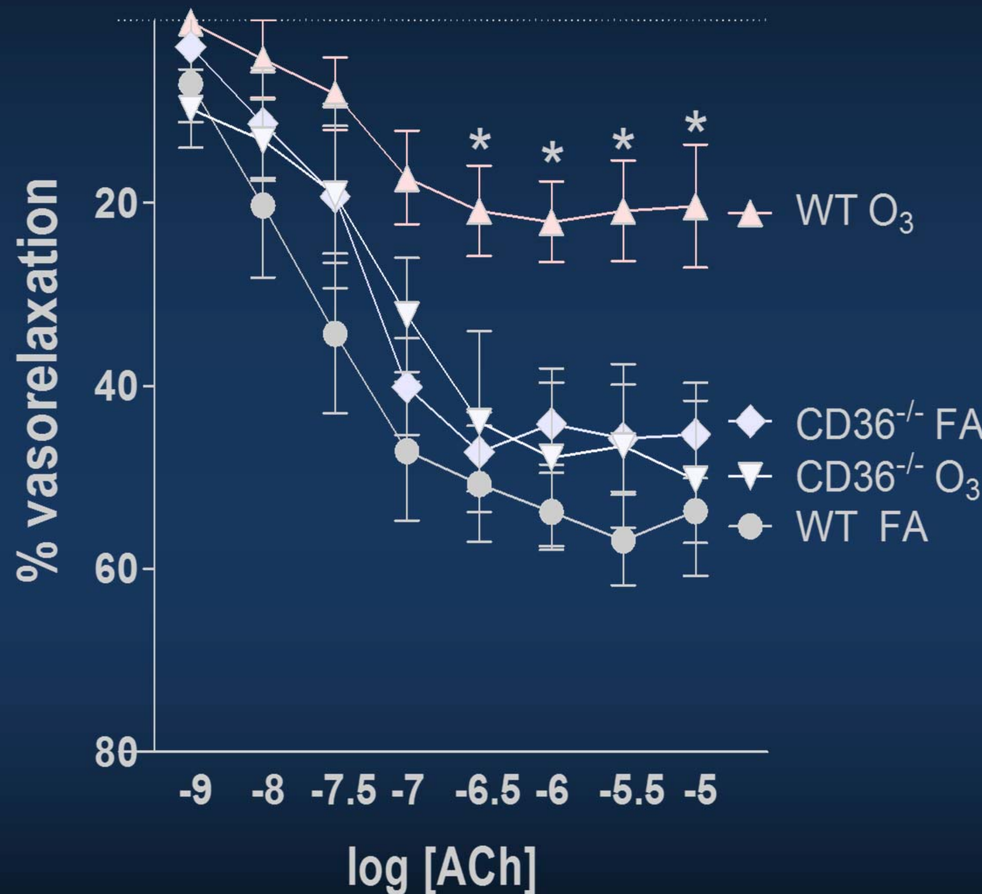
- Reactions of air pollutants in the lungs create reactive intermediates
  - DAMPs, AGEs, Hyaluronan Fragments, 4-HNE
- Intermediates adduct onto or modify existing endogenous proteins
- Modifications are recognizable epitopes for Pattern Recognition Receptors



Ozone induces impairments of vasorelaxation in aortas from WT, but not CD36-null mice

- 1 ppm O3 x 4 h, harvested aortas 24h post-exposure
- Female C57 mice, ~8-16 wks old paired with CD36$^{-/-}$ mice
- Force-tension myography on aortic rings
- Preconstricted (U46619), then relaxation to acetylcholine assessed

Robertson et al, Tox Sci, 2013

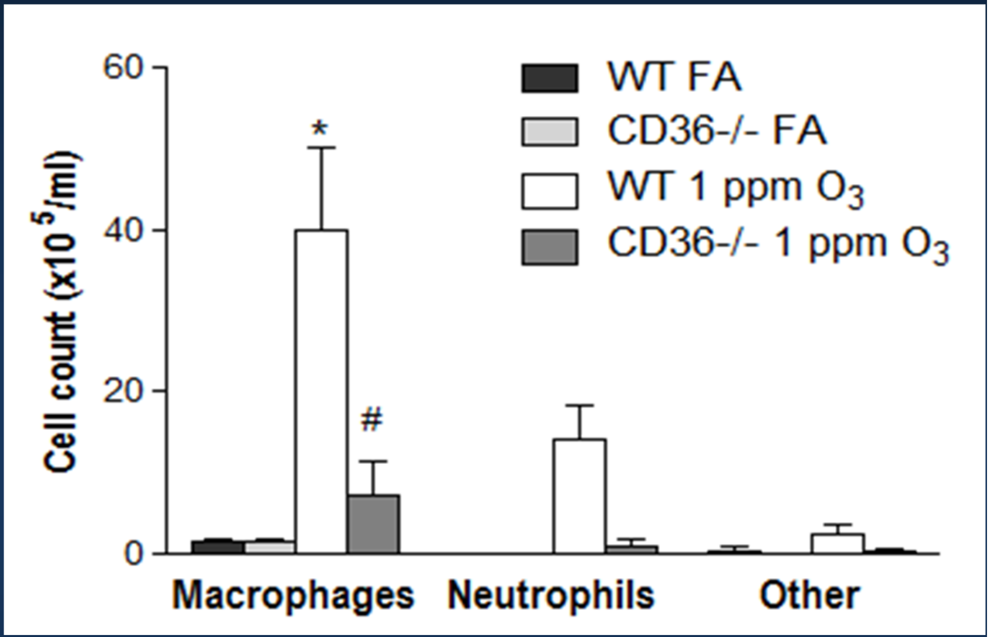


CD36-Deficiency Also Protected Against Pulmonary Inflammation

- Significant reduction in O3-induced macrophages and neutrophils in CD36-null compared to WT

- Not a caveat –

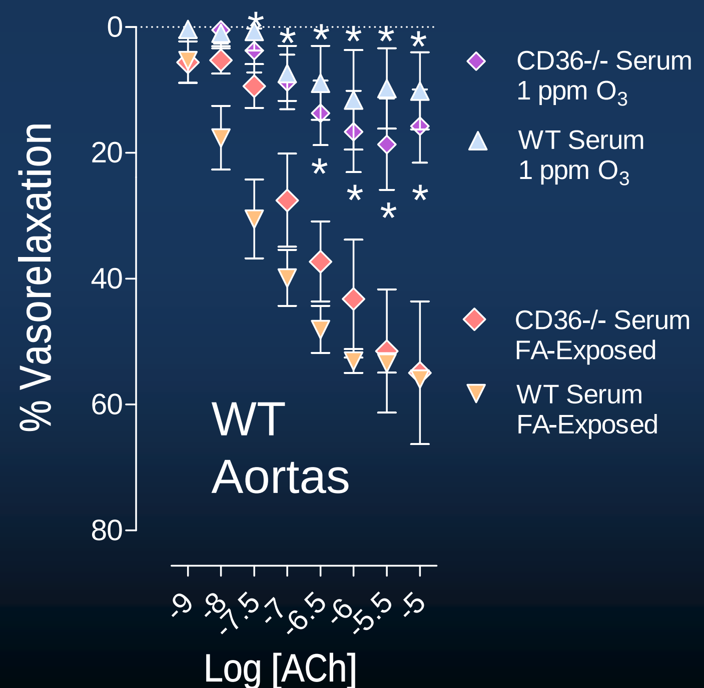
an OPPORTUNITY!!

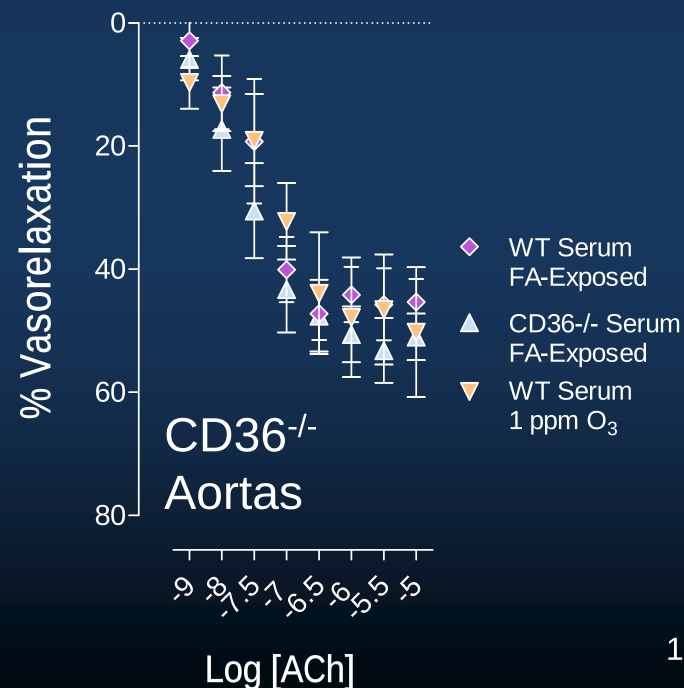


Robertson et al. in prep

Using Homologous Serum Assays with Ex Vivo Aortas…

- Naïve WT aortas treated with 2.5% serum from CD36-null mice recapitulated O3-induced vascular impairments
- Serum modifications are independent of CD36 AND pulmonary inflammation!
- CD36 on aortas mediates response to serum factors!



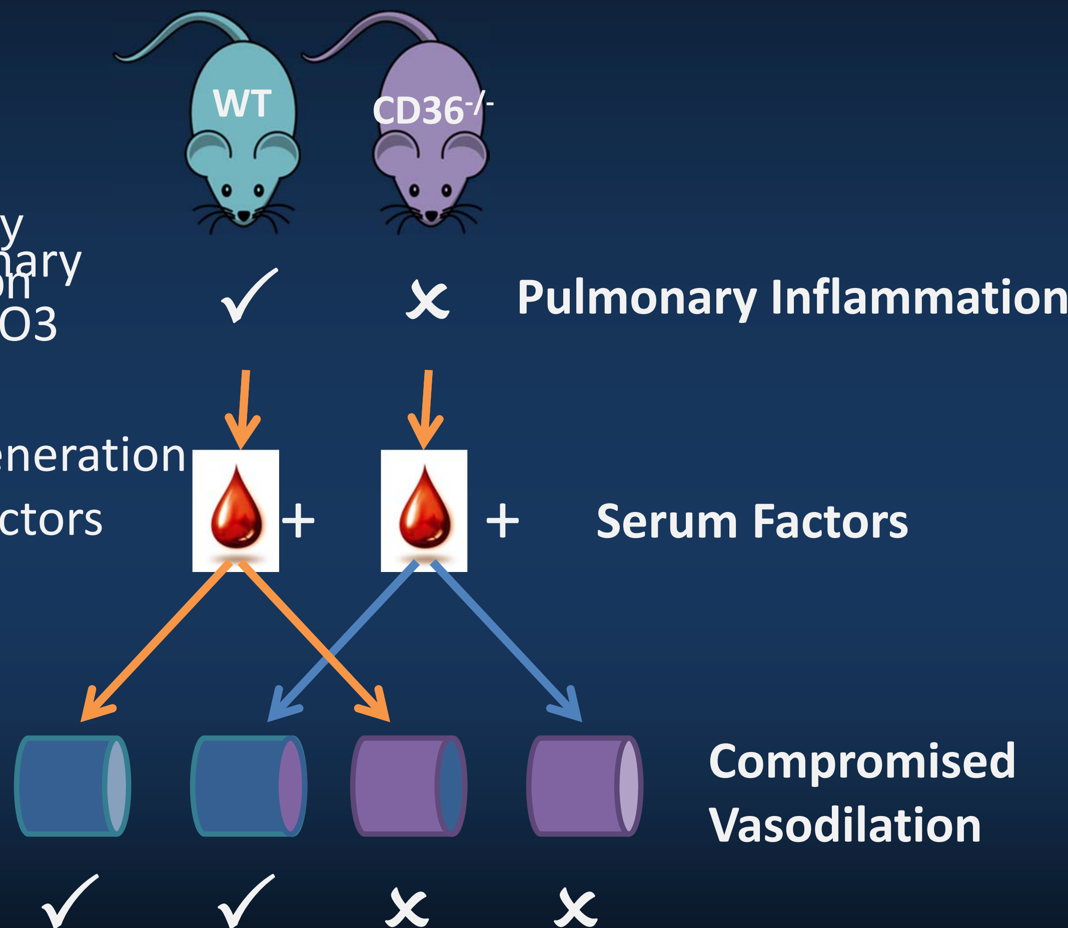


Pattern Recognition Receptor CD36 Role in Mediating Ozone-Induced Vascular Outcomes

Pulmonary inflammation may help REDUCE the generation and release of circulating factors!

CD36 *is* needed for pulmonary inflammatory response to O3

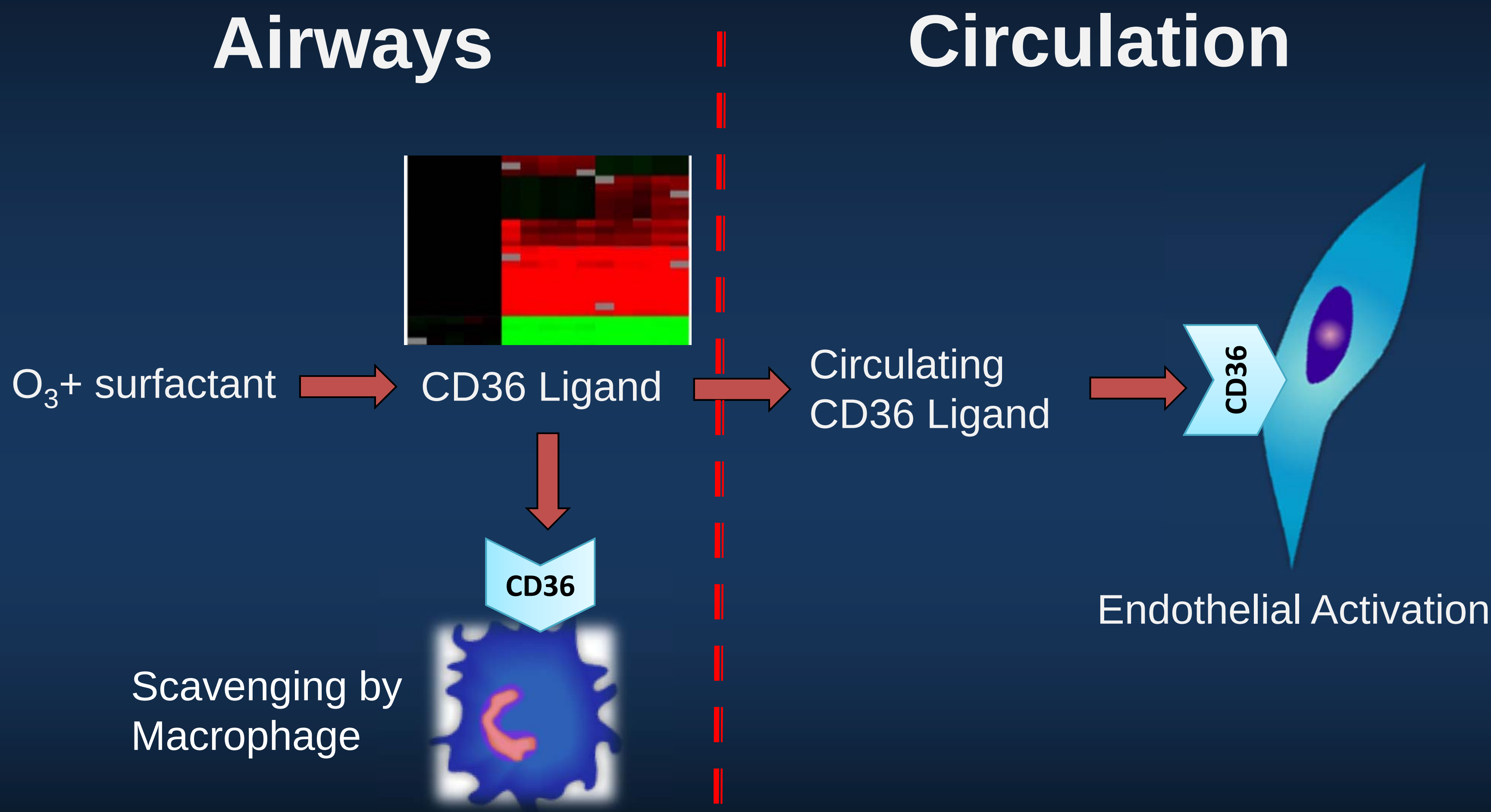
CD36 *is NOT* needed for generation of circulating vasoactive factors



Vascular CD36 *is* needed for the response to O3-induced vasoactive factors

Robertson et al, in prep

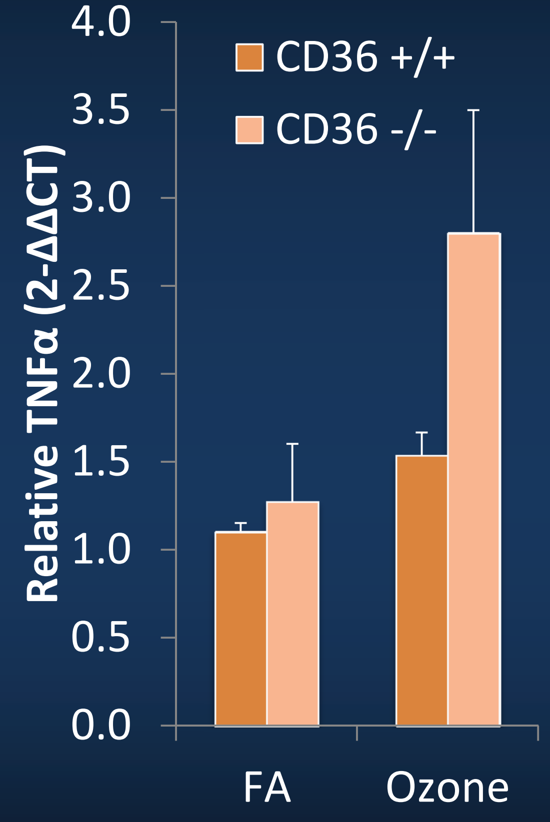
Do Alveolar Macrophages Have a Role in Scavenging By-Products?



Activation of Microglia is most pronounced with serum from CD36-null mice

- Preliminary findings of O_3 induced neuroinflammation suggest that expression of TNFα in the brain of $CD36^{-/-}$ mice is enhanced compared to WT

Michelle Block, VCU

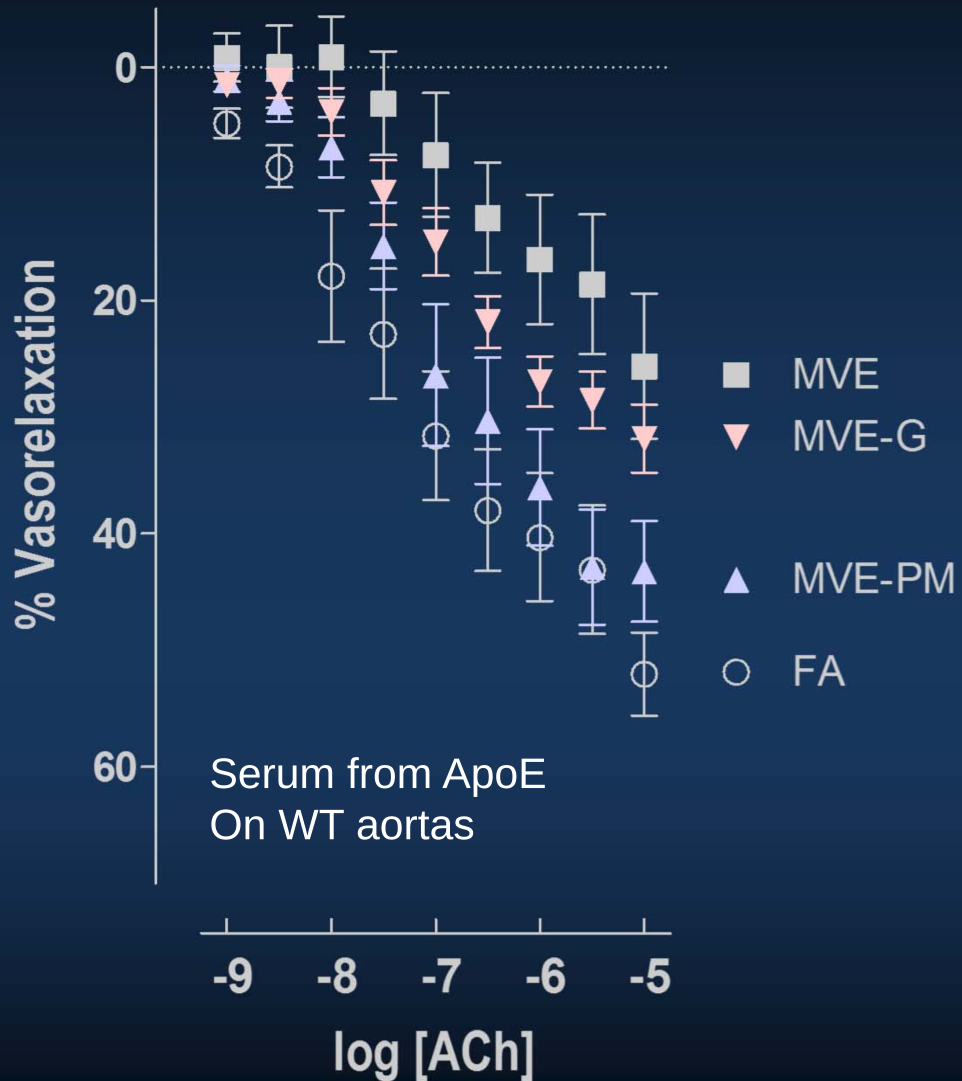


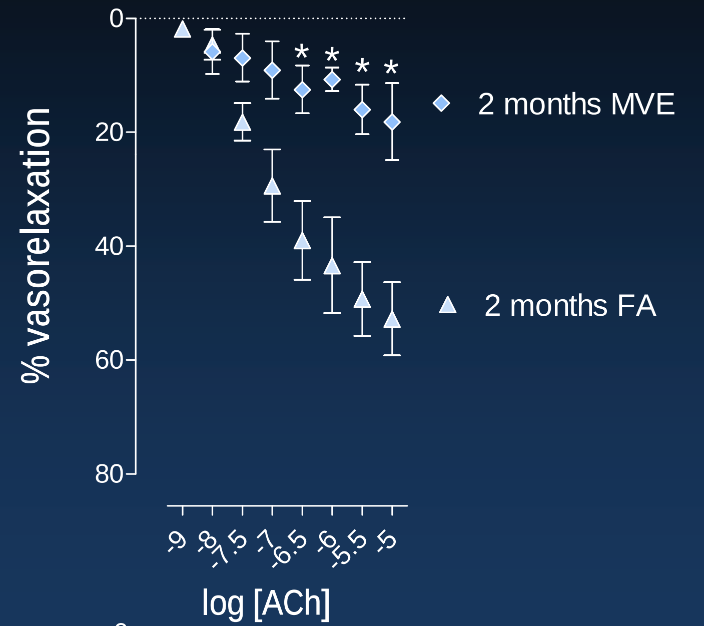
Courtesy Michelle Block, VCU

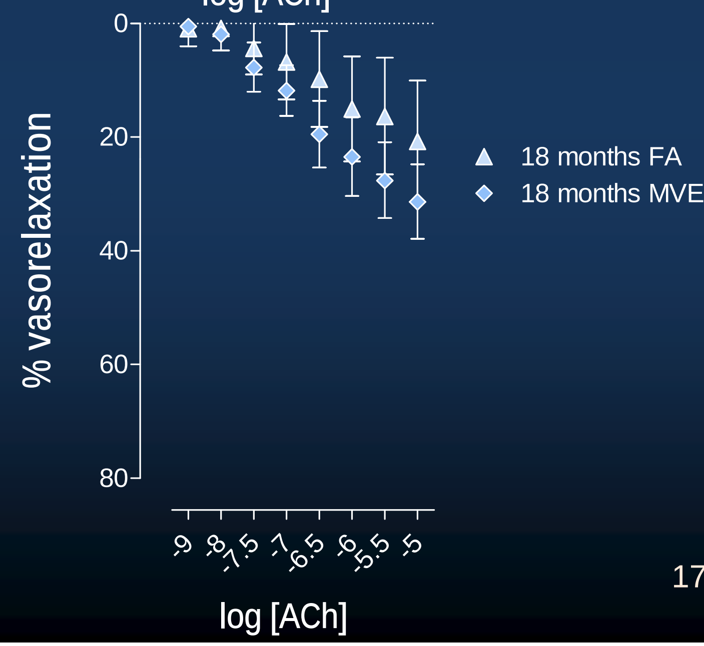
Are serum factors playing a role with complex mixtures?

- $ApoE^{-/-}$ mice on high fat diet (8 weeks old, male)
- Exposed 6h/d for 50d to:
  - MVE (300ug/m3)
  - MVE minus PM
  - MVE minus gases
  - Filtered Air Control
- Serum used for ex vivo vascular myography
- Aortas sectioned for immunohistochemistry

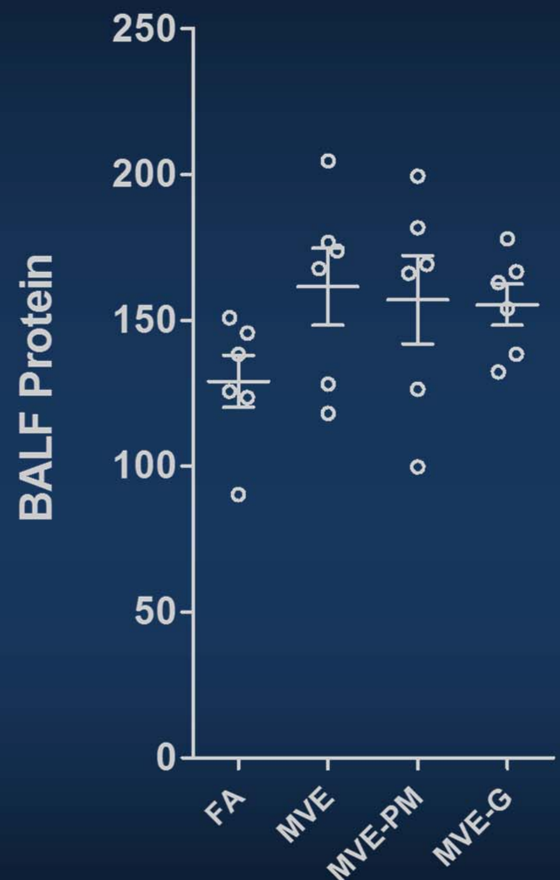
MVE Induces Vasoactive Factors

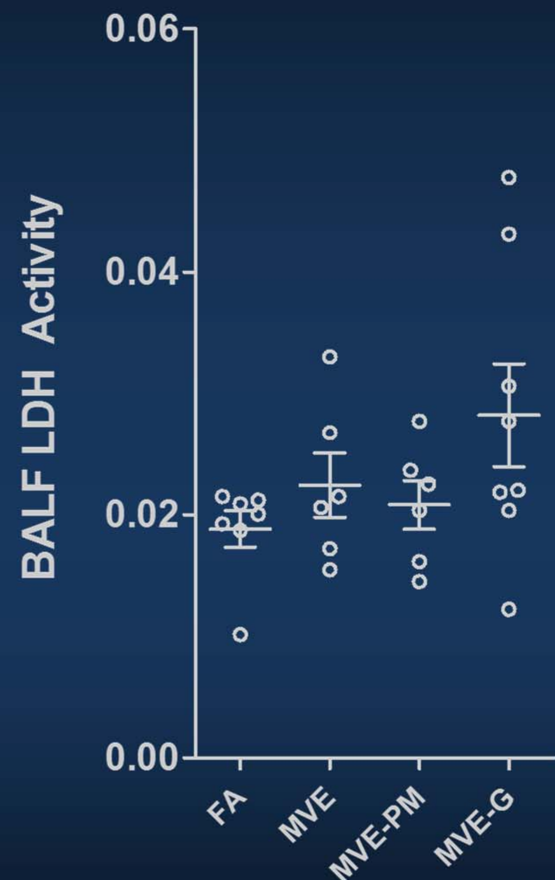




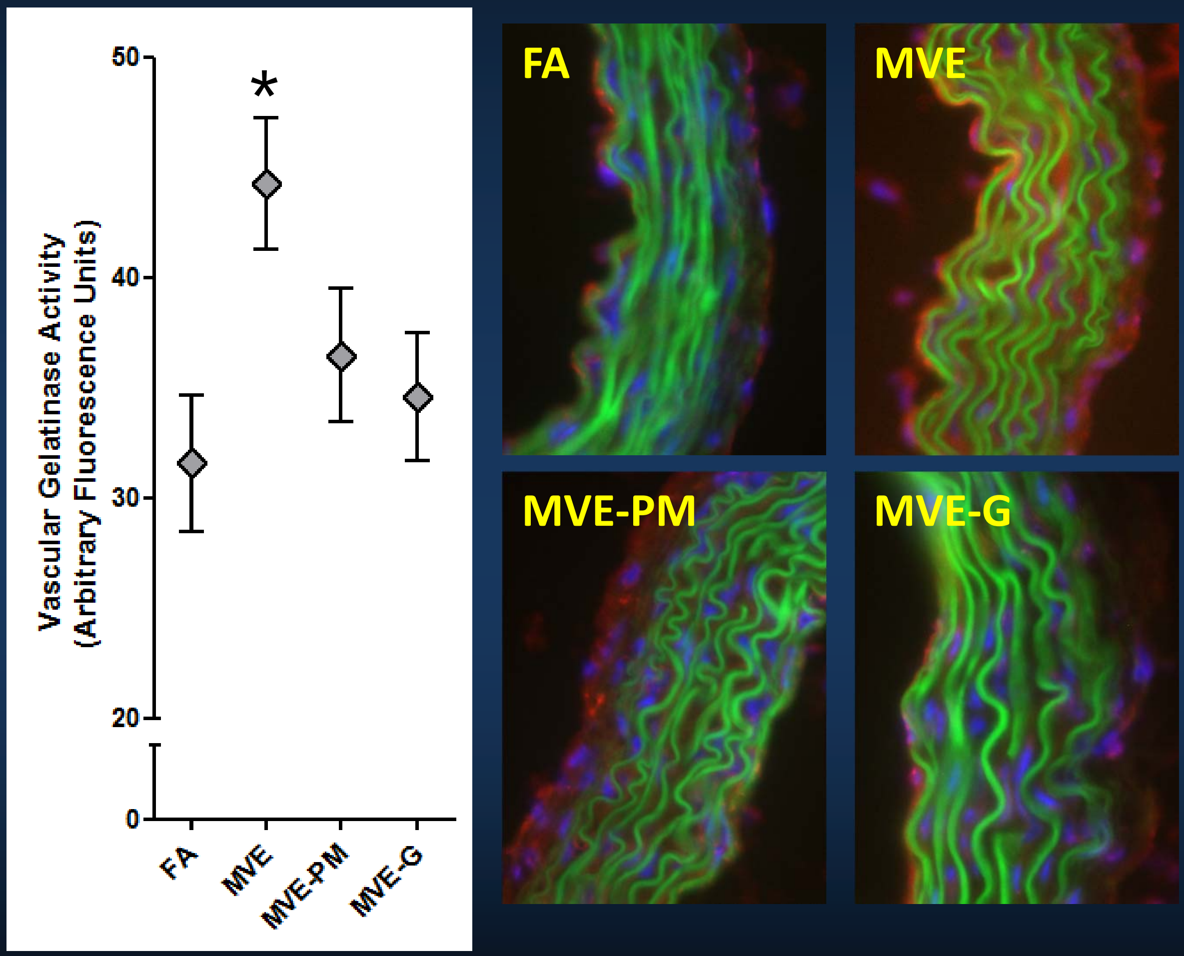


MVE Induces No Pulmonary Inflammation, Assessed by Bronchoalveolar Lavage

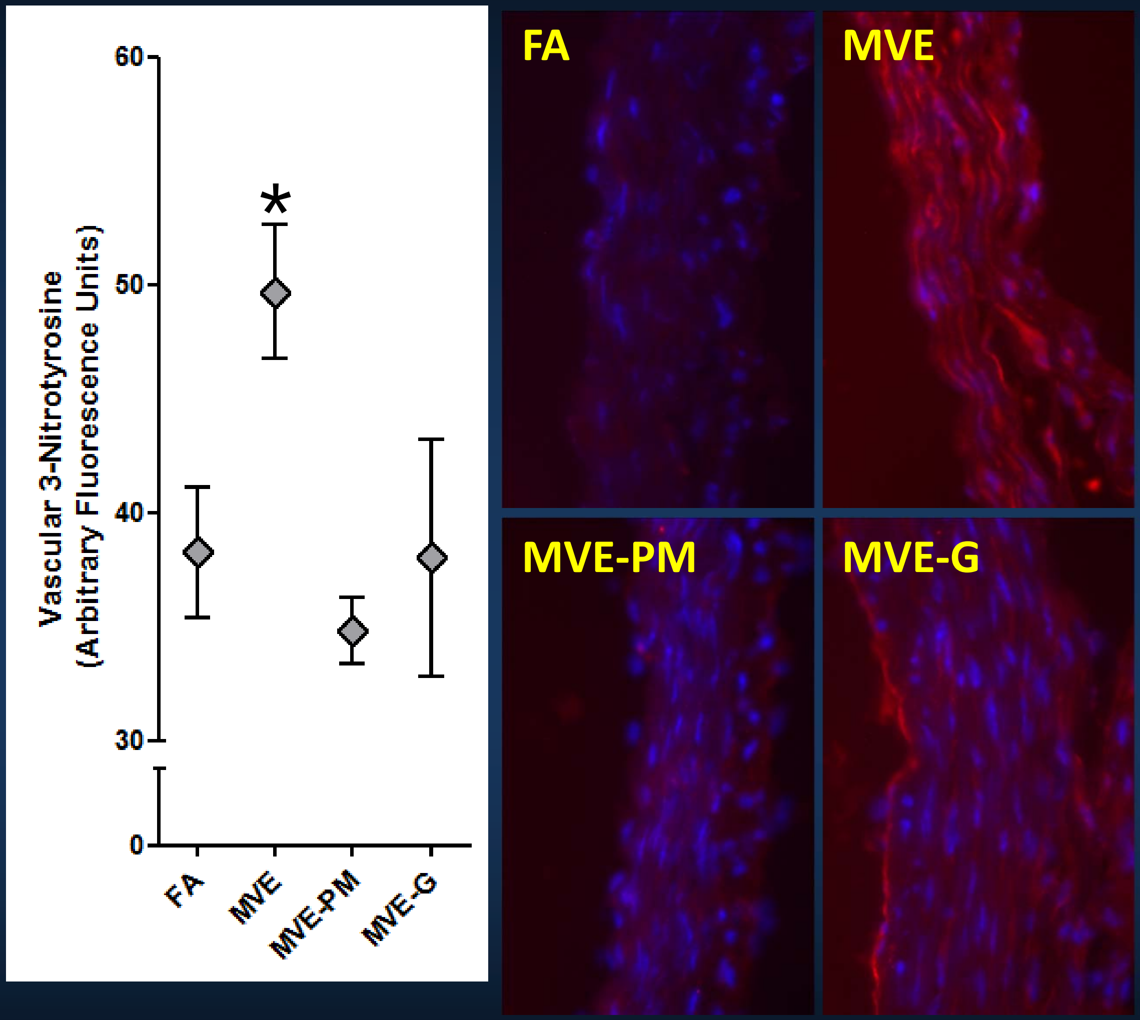




MVE Induces MMP9 Protein Levels, Reduced by Filtration/Denudation

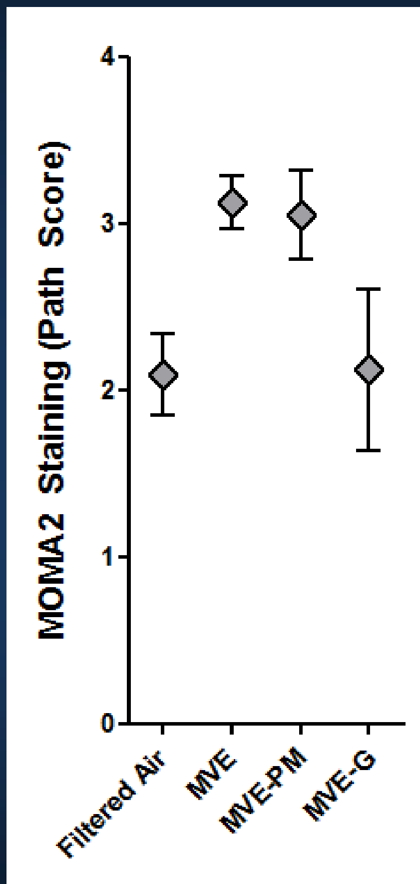


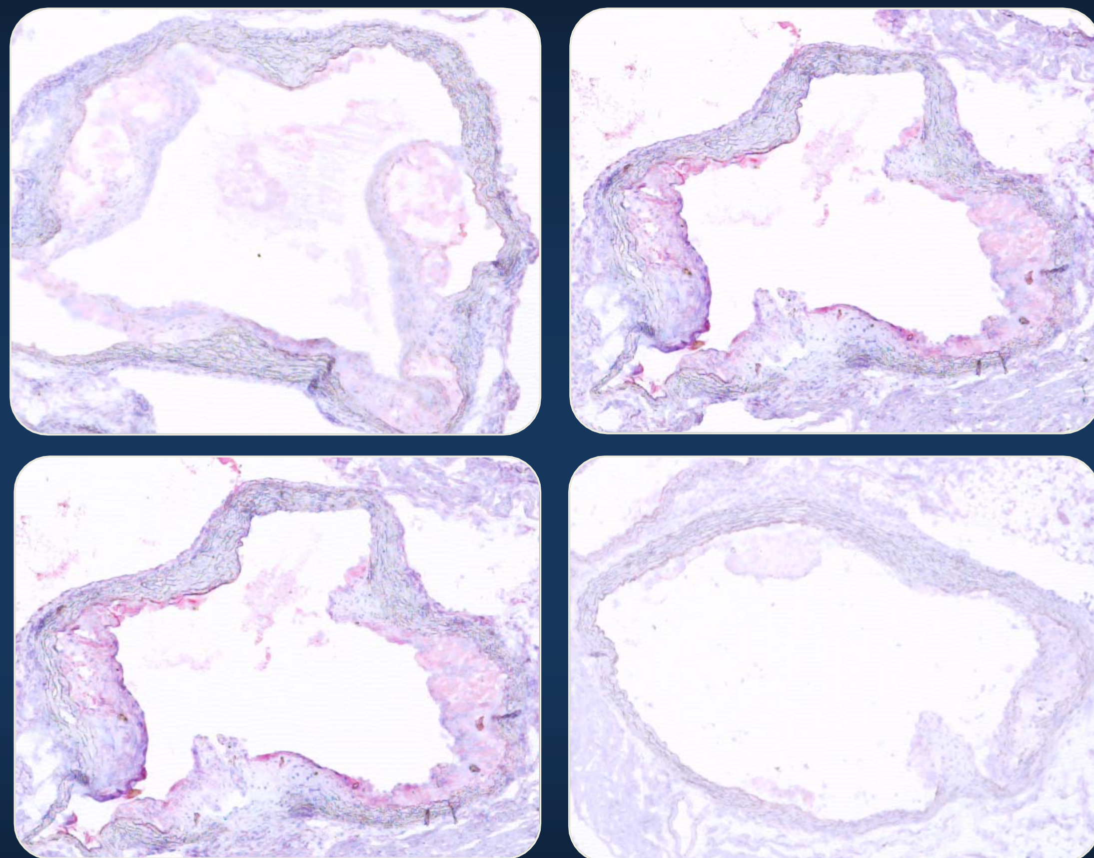
3-NT



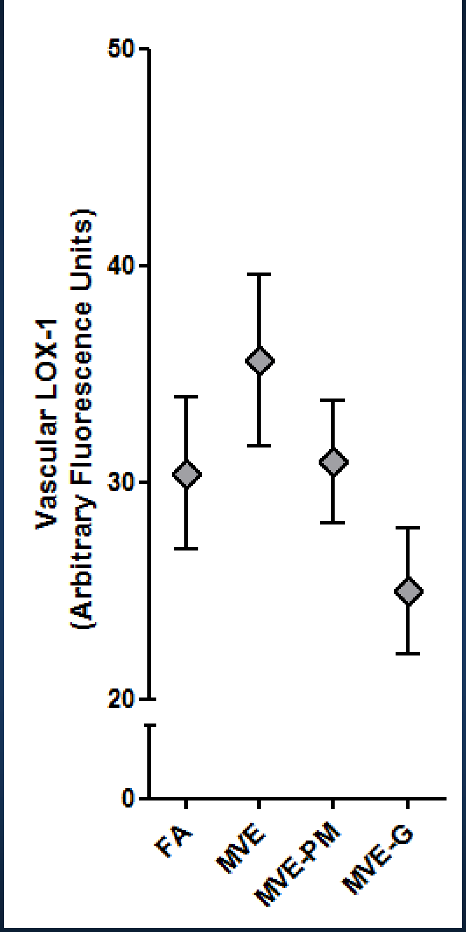
Vascular 3-Nitrotyrosine
(Arbitrary Fluorescence Units)
60
50
40
30
0
*
FA
MVE
MVE-PM
MVE-G
FA
MVE
MVE-PM
MVE-G

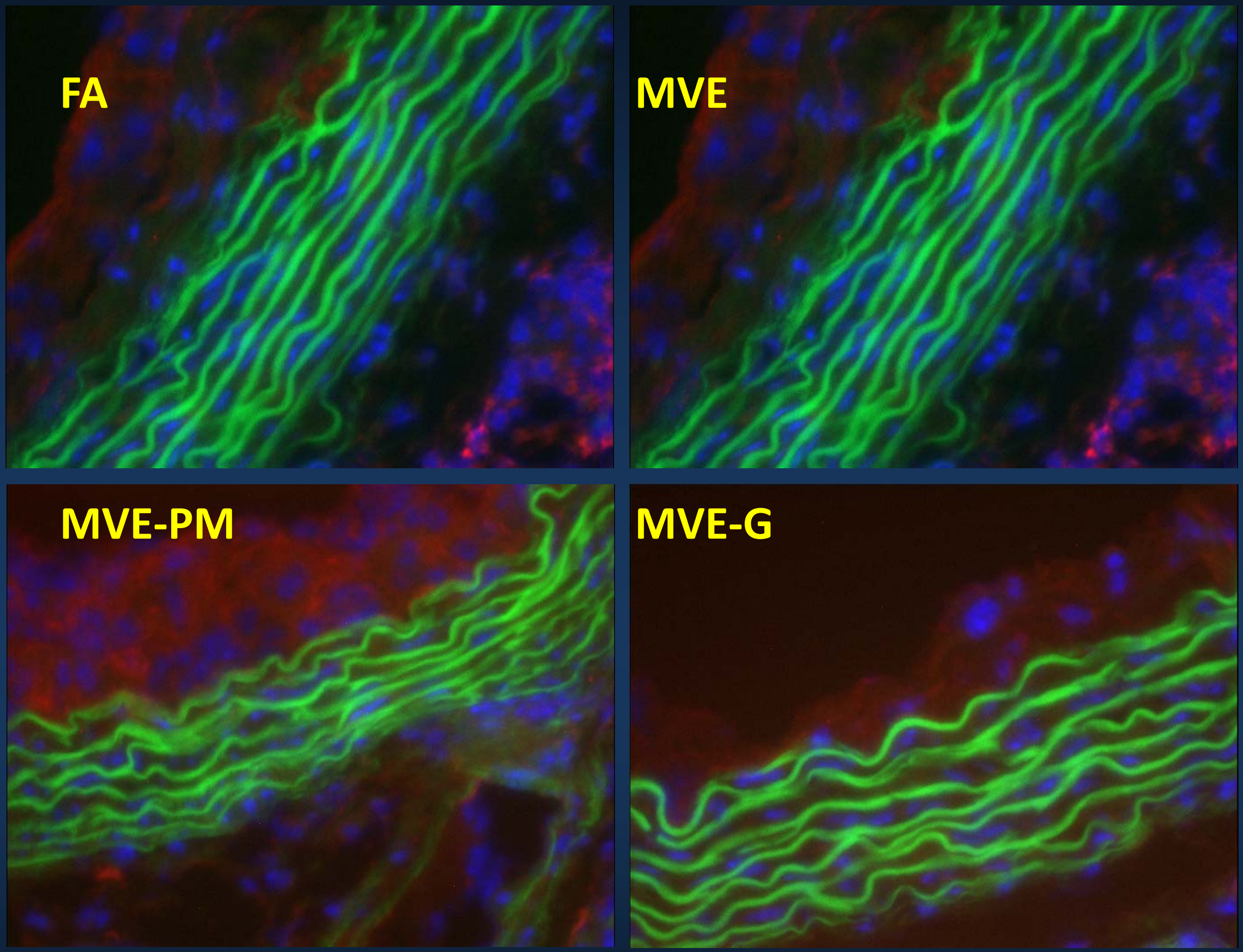
MOMA2


MOMA2 Staining (Path Score)
0
1
2
3
4
Filtered Air
MVE
MVE-PM
MVE-G

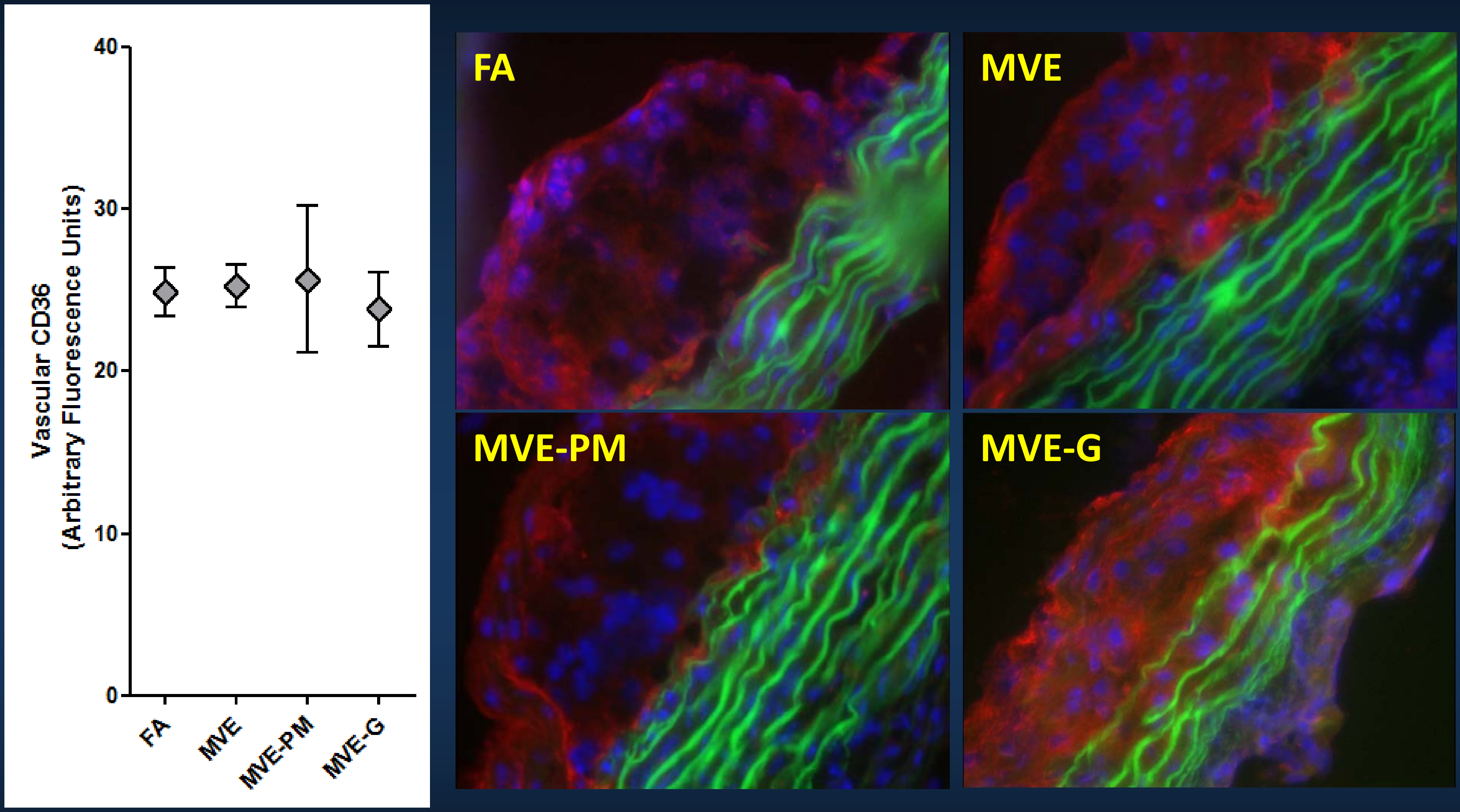


LOX-1


Vascular LOX-1
(Arbitrary Fluorescence Units)
50
40
30
20
0
FA
MVE
MVE-PM
MVE-G

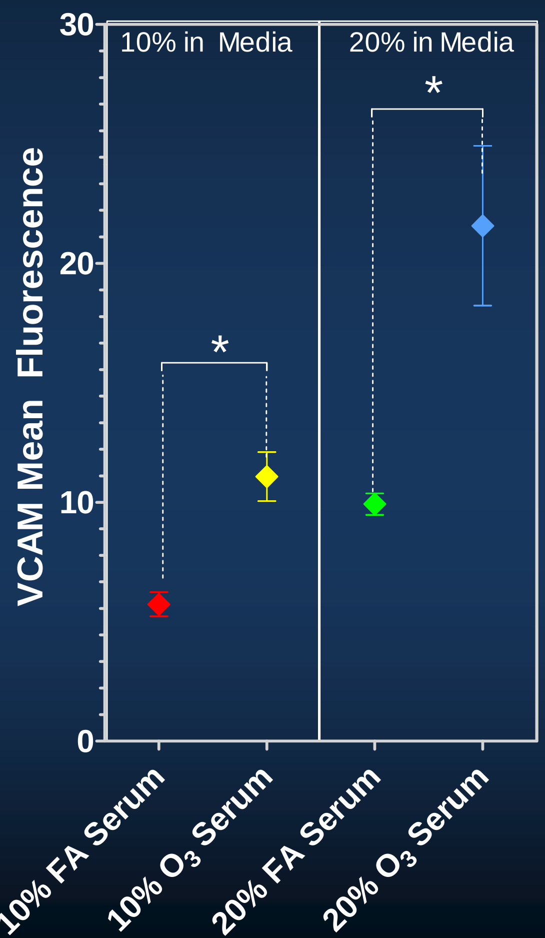

FA
MVE
MVE-PM
MVE-G

CD36 Unchanged

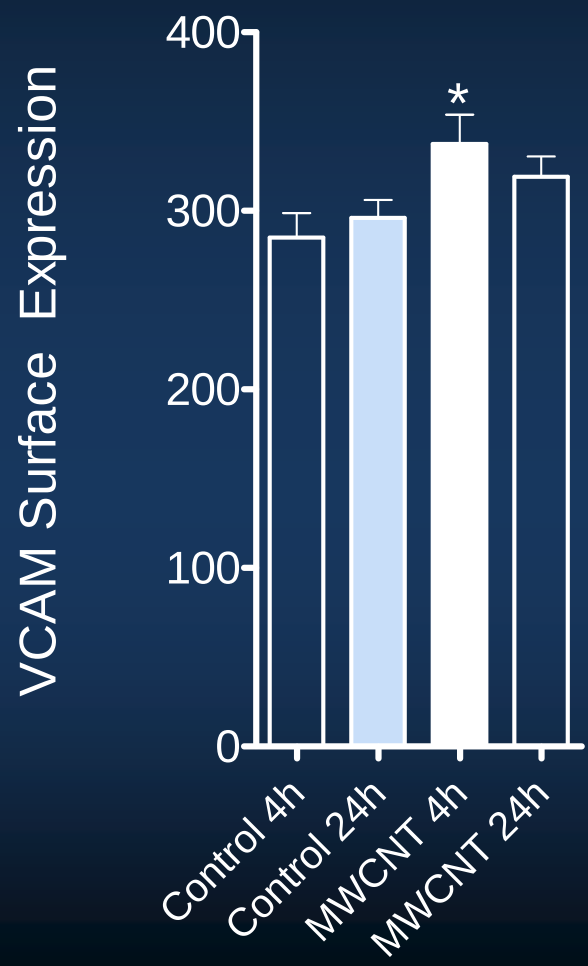


Further Exploration on Bioactivity in Serum: Inflammatory Markers

Rats, O3, 1ppm x 4h

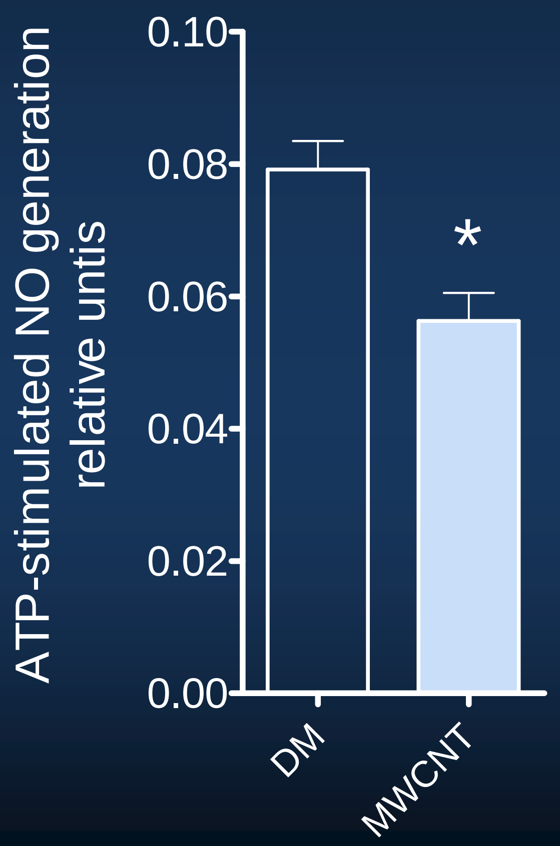


Mice, MWCNT, 40 ug IT

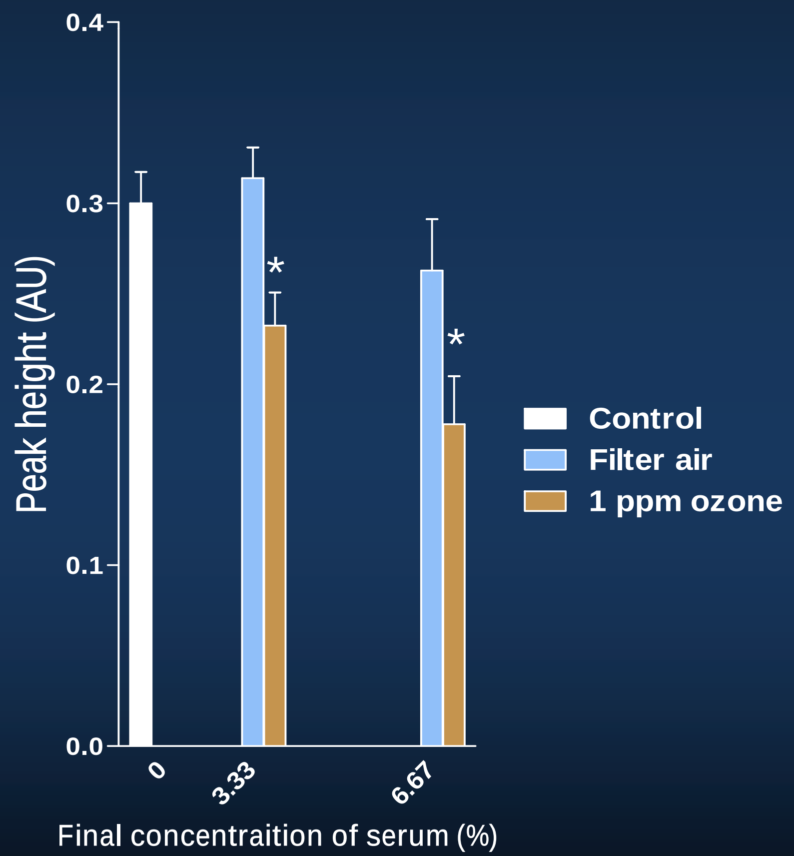


Nitric Oxide Generation and Scavenging by Electron Paramagnetic Resonance

NO Generation in Murine Endothelial Cells



Scavenging of NO (Spermine NONOate) in Acellular Assay



Getting back to actual Aims

- Now have approval at LRRI to conduct a series of 1-day exposures to MVE and related mixtures in rats and mice, to assess relative serum inflammatory potential changes.
- Can also consider treating circulating monocytes with serum and looking for priming and adhesive effects
  - Readily move to TLR/CD36/LOX-1 knockout mice for cells

Serum Compositional Changes: Summary

- Bioactivity
  - Impairs ACh-mediated vasodilation
    - Reduces NO generation, scavenges free NO
  - Induces battery of inflammatory responses
    - Increased surface expression and mRNA for VCAM, ICAM, mRNA and release of IL-8
  - Macrophage activation Aim 3
  - TBD: Endothelial Barrier Integrity
- Common responses induced by
  - Ozone, MVE, MWCNT, graphene, nitrogen dioxide, diesel Aim 1
- Compositional analysis indicates
  - No cytokines altered
  - 1500 other things – fragments and adducts

CENTER FOR CLEAN AIR RESEARCH

UNIVERSITY *of* WASHINGTON

Project 5: Effects of Long-Term Exposure to Traffic-Derived Particles and Gases on Subclinical Measures of Cardiovascular Disease in a Multi-Ethnic Cohort

PI: Joel Kaufman

Aims

- Aim 1: To build a multi-pollutant exposure model for traffic-derived air pollutants for use in epidemiological analysis
- Aim 2: To determine the effect of time-in-transit on personal exposure in this cohort
- Aim 3: To estimate the effect of individual-level exposure to traffic-derived air pollution on subclinical cardiovascular disease in MESA Air

Current Focus is on Field Work

- <u>Aim 1:</u> To build a multi-pollutant exposure model for traffic-derived air pollutants for use in epidemiological analysis
- **<u>Aim 2:</u> To determine the effect of time-in-transit on personal exposure in this cohort**
- <u>Aim 3:</u> To estimate the effect of individual-level exposure to traffic-derived air pollution on subclinical cardiovascular disease in MESA Air

Focusing on In-Vehicle Exposures

- Studying actual travel patterns
  - GPS trackers and proximity sensors provide gold standard
  - Can be combined with more specific self-reported time-location diary data
  - Can then be compared with data acquired for summer and winter from the MESA Air Questionnaire
- Measuring concentrations of TRAP in vehicles
  - Goal to determine importance of the in-vehicle “compartment”
  - Determine whether we need to add an in-vehicle component to MESA Air individual exposure model

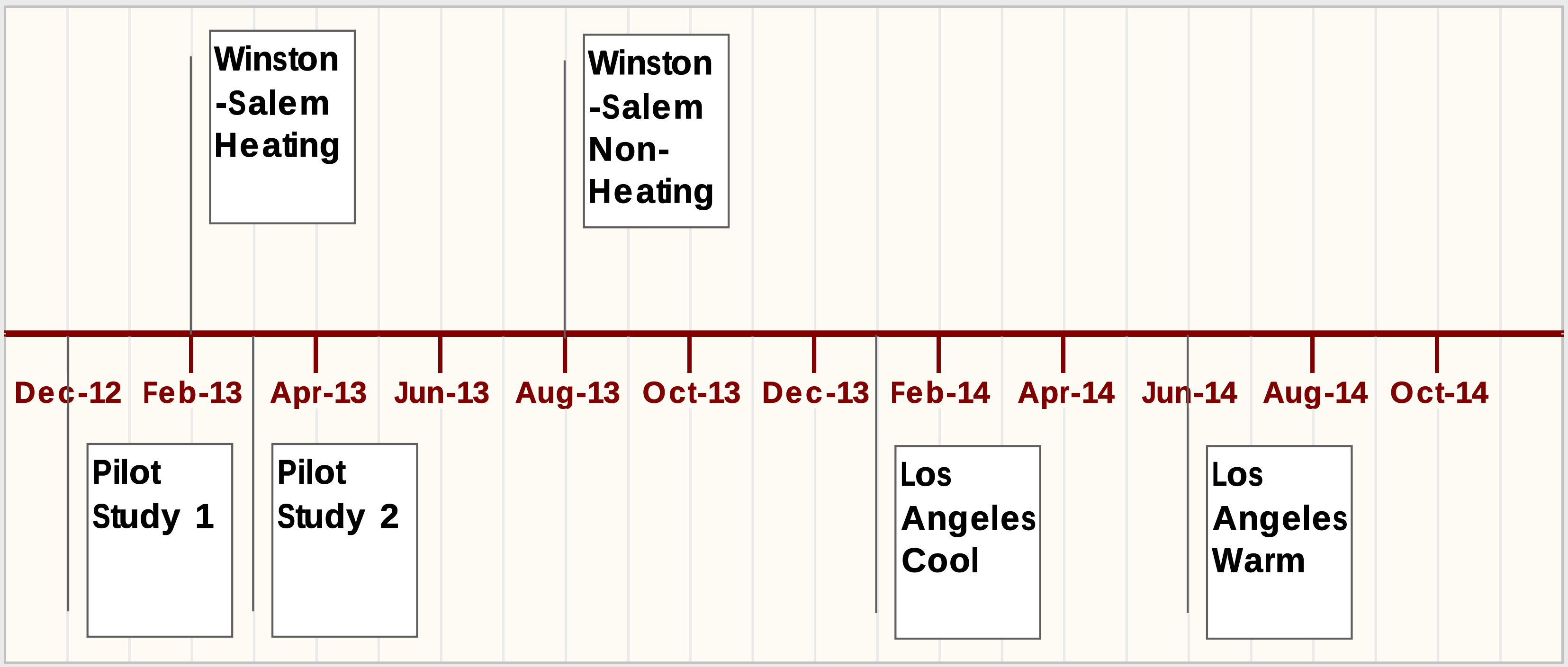
Monitoring Campaign

- Two-week duration
- Sample ~50 participants in each of two cities (Winston-Salem and Los Angeles) in each of two seasons
  - February 2013 (Winston-Salem) and January 2014 (LA)
  - August 2013 (Winston-Salem) and June 2014 (LA)
- Location logging
  - GPS tracking unit
  - Proximity monitor
  - Self-reported time-location diary
- Passive monitoring
  - Ogawas
  - Organic Vapor Monitor

Passive monitoring

- Personal monitoring plus three “compartments”
  - Indoor
  - Outdoor
  - In-vehicle
- In-vehicle monitoring set up is portable
  - Participants asked to take it with them if they travel in other personal vehicles
  - Participants asked to open and close lid at beginning and end of their trips
  - Also includes a proximity sensor, a timer, and a temperature and humidity logger

Project 5 Schedule



Primary Questions for Pilot Studies

- How much time do participants need to drive?
- Which analytes are most likely to provide useful information in the full study?
- How reproducible are the samples?

Pilot Studies

- Pilot Study 1
  - December 2012 in Seattle
  - 10 samples and 10 duplicates
- Pilot Study 2
  - March 2013 in Seattle
  - 16 $NO_x/NO_2/SO_2$ samples and 22 VOC samples
- Results
  - 30 minutes of driving time per day is sufficient for achieving detection limits for most agents of interest
  - Generally high reproducibility in duplicate samples

Pilot Studies: Chemicals Detected by Driving Time

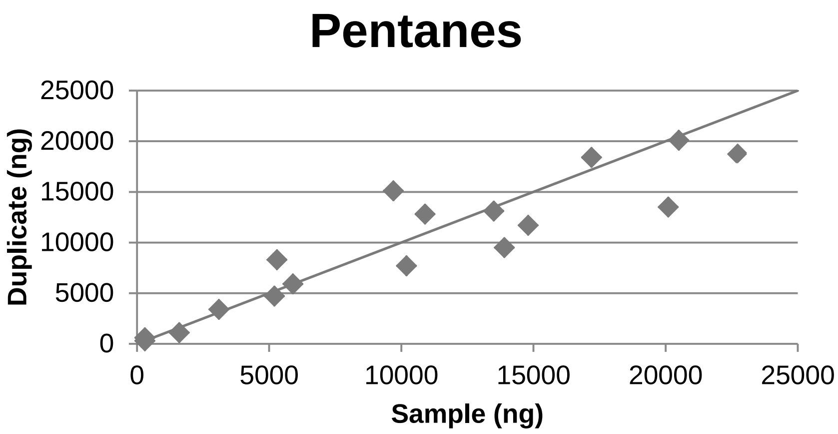
	Pentanes	n-Nonane	n-Decane	n-Undecane	n-Dodecane	Benzene	Toluene	m-Xylene	o-Xylene	NO_X	NO_2
30 min/day	●	●	●	●	●	○	●	●	●	●	○
45 min/day	●	○	●	●	○	○	●	●	●	●	○
>45 min/day	●	●	●	●	○	○	●	●	●	●	●

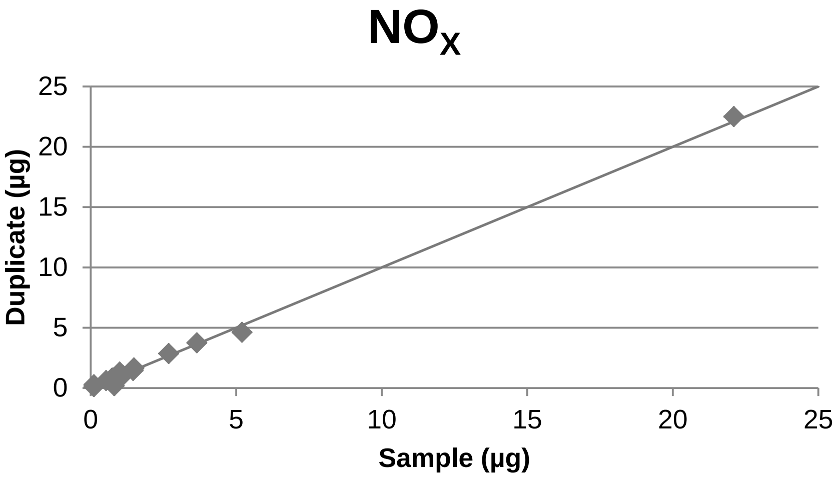
● Detected in all samples

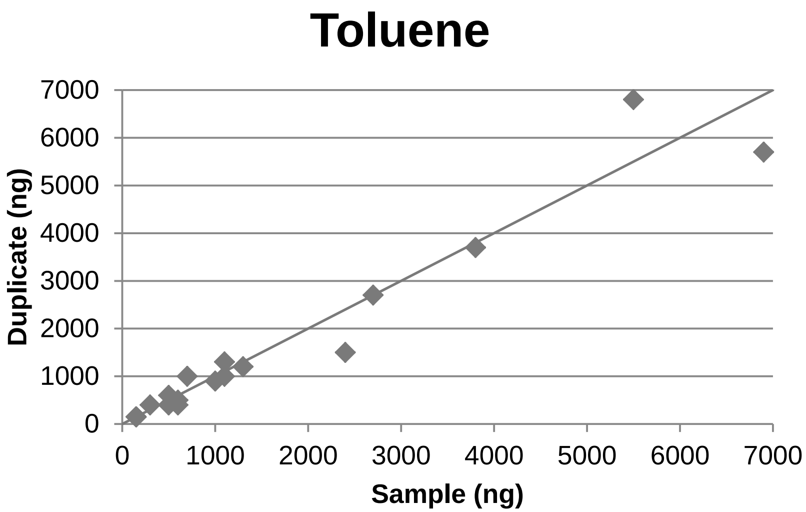
○ Detected in some but not all samples

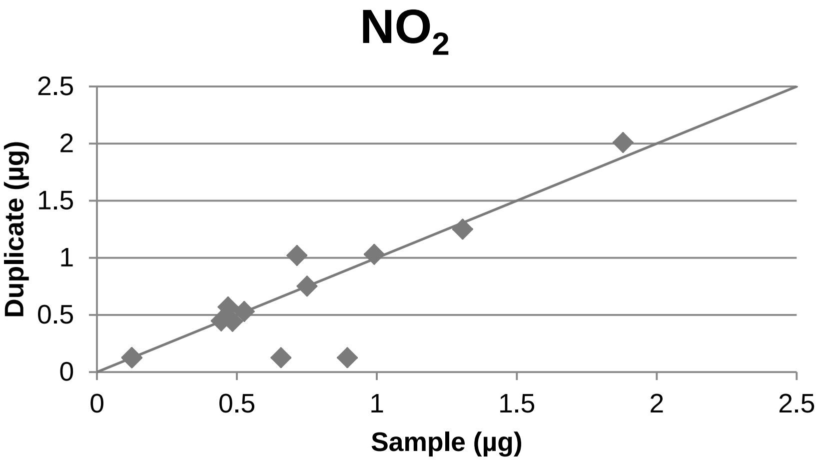
Isoprene, SO_2, and ozone were not detected in any samples

Pilot Studies: Duplicates

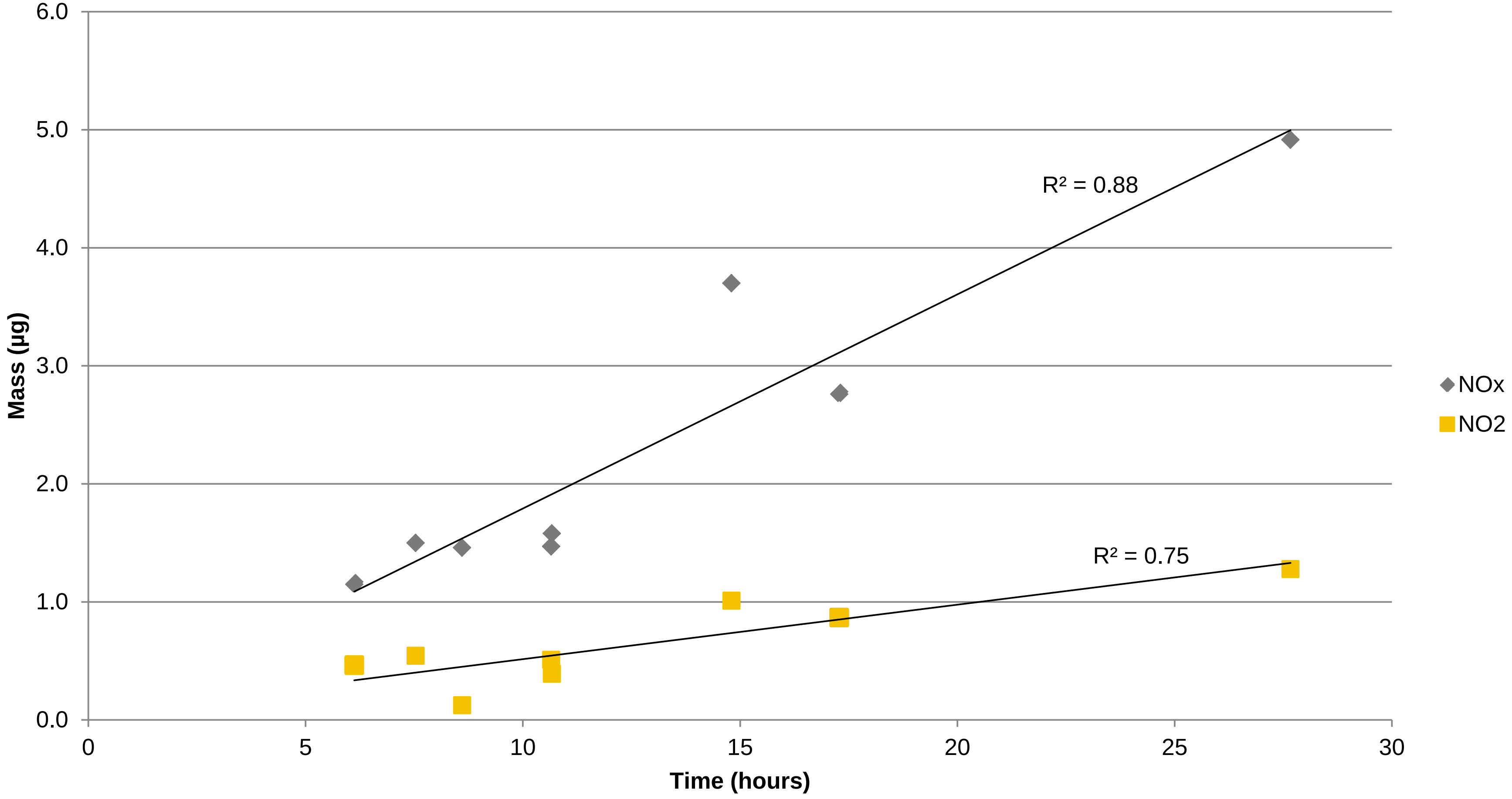




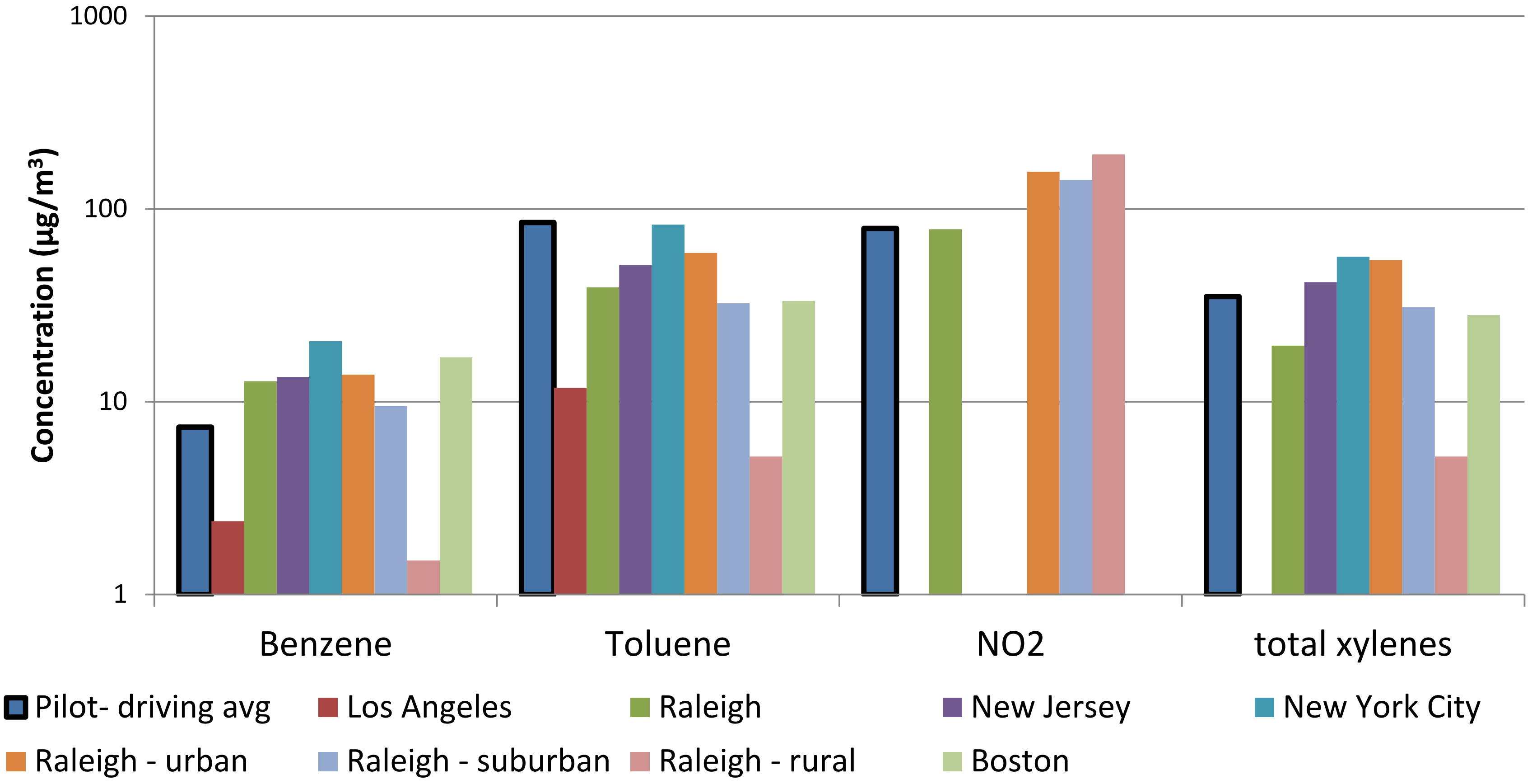




Pilot Studies: NO_x and NO_2 vs. Time Driving



Pilot Study 1: Comparison to Literature Values



Data from: Chan et al. (1991a,b), Fedoruk & Kerger (2003), Lawryk & Weisel (1996), Riediker et al. (2003)

Winston-Salem Heating

- January 27th through February 21st
- 46 participants (96% of goal)
- Mailed results letters in early July

COURTESY: VISITWINSTONSALEM/WACHOVIA

Winston-Salem Heating: Participant Demographics

Category	Project 5		W-S Cohort at Exam 5	
	Number	Percent	Number	Percent
Gender				
Male	21	46%	348	46%
Female	25	54%	415	54%
Race				
White, Caucasian	20	43%	413	54%
Black, African-American	26	57%	348	46%
Age Group*				
45-54	1	2%	8	1%
55-64	9	20%	236	31%
64-74	18	39%	256	34%
75-84	15	33%	215	28%
85+	3	7%	48	6%
Median Age*	72		70	
Age range*	54 - 89		54 - 93	

*at exam 5

Driving Time (hours per week)

	Questionnaire	Time Diary
Range	0 – 64	1.3 – 18.1
Mean ± std dev	8.7 ± 10.2	7.1 ± 4.0
Median	7.0	6.6
95th Percentile	16.8	14.1

Winston-Salem Heating: Number of Samples

	Number of samples deployed	Duplicates (% of samples deployed)
Indoor	46	11 %
Outdoor	46	9 %
Vehicle	46	9 %
Personal	46	--

	Number of samples deployed	Field Blanks (% of samples Deployed)
Total Samples	184	9 %

Winston-Salem Heating: GPS

- Over 50% of the units ran the whole study period
- Over 80% of the units ran for at least 1 week
- Only 1 unit did not run at all
- Conducted extensive testing – confident that we will have more complete data in the next field campaigns

Example Map

N
W
E
S
90
90
90
405
90
5
5
705
Your House
GPS Tracks
USA Major Roads
Road Classification
Freeways
U.S. and State Highways
Other Major Road
Secondary Road
Local Connecting Road
Important Local Road
Parks
Lakes, Rivers, Streams
0
2.5
5
10
Miles

Example Map – Close-up

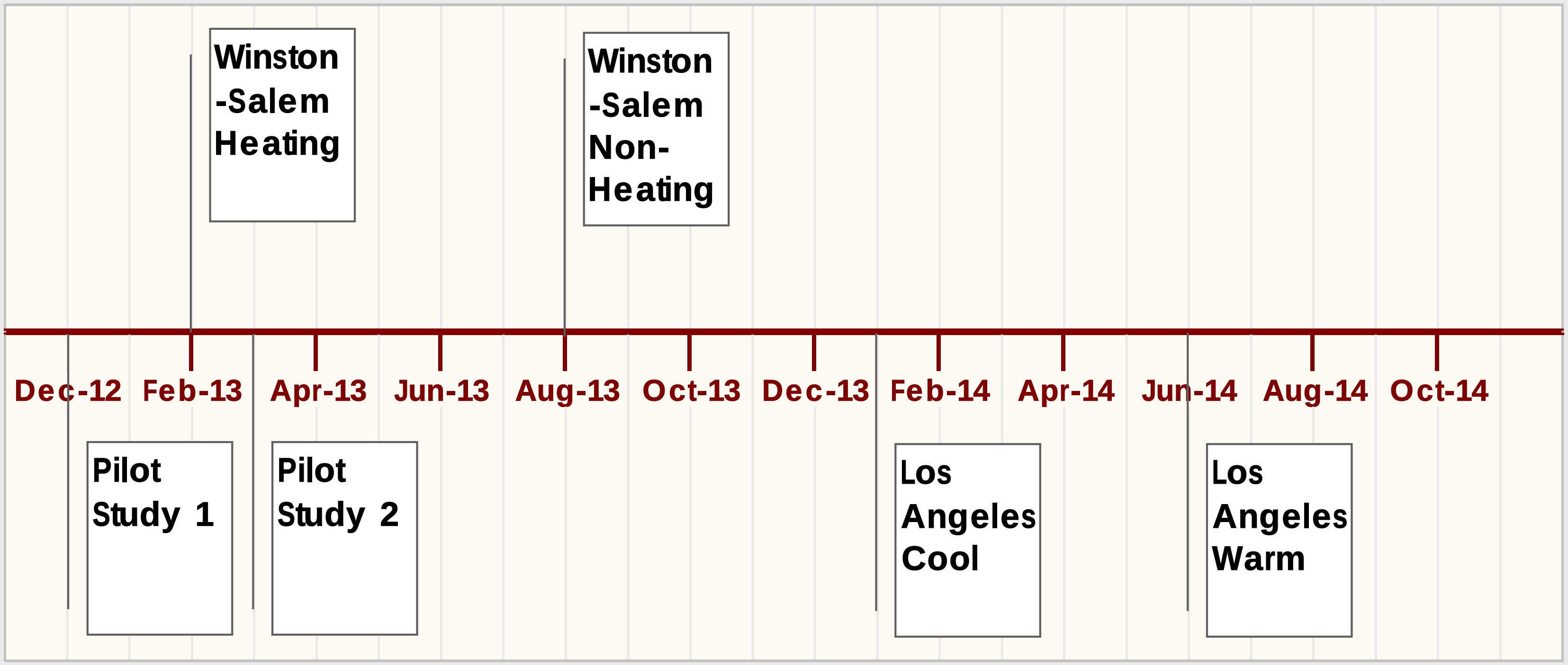
Improvements to Sampling Procedures

- Simplified Time Diary format
- Lighter personal monitor
- “Do’s and Do Not’s” Sheet for participants
- GPS units fixed

Summary

- Implemented novel sampling equipment and study design
- Successfully completed our first field campaign and start our second campaign *next week*
- Results so far suggest that the highest exposure concentrations occur in vehicles

Project 5 Schedule



Biostatistics Core Update (public release version)

UW CCAR
SAC Meeting
23 July 2013

Biostatistics core overview of activities

- Project 1 data analysis and cleaning
- Inter-center collaborations (UW, Harvard, Emory/GT)
- Multi-pollutant methodology for cohort studies with misaligned monitoring data

Multi-pollutant methodology

- **Goal**: Statistical framework for assessing health effects of long-term exposure to multi-pollutant mixtures
  - **Dimension reduction** of the multi-pollutant exposure surface based on monitoring data
  - **Spatial prediction** of the multi-pollutant exposure surface (focus on low-rank spatial models)
  - **Health effect inference** that accounts for uncertainty from prediction and dimension reduction in the first two steps
- **Application:**
  - Current: Annual averages from STN/IMPROVE + blood pressure from NIEHS Sisters Study
  - Planned: Time-adjusted concentrations from CCAR Project 1 + CVD outcomes from MESA Air

Three step sequential procedure

High dimensional monitoring data → PC profiles + scores at monitor locations → PC scores at subject locations → Health effect analysis

1. (Sparse) PCA

2. Multivariate spatial prediction modeling (**low dimensional**)

3. Linear regression (with measurement error correction)

- Current progress and plan
  - Steps 1 + 2: Predictive sparse PCA (Jandarov/Szpiro)
  - Step 2: Low-rank version of common component model for co-kriging (Bergen/Szpiro)
  - Step 3: Measurement error correction with multiple spatially misaligned pollutants (Bergen/Szpiro)

See: *Roman Jandarov slides*

- *Review of sparse PCA*
- *Predictive sparse PCA algorithm*
- *Implementation details*
- *Application to national air quality monitoring data (CSN+IMPROVE)*

Three step sequential procedure

High dimensional monitoring data → PC profiles + scores at monitor locations → PC scores at subject locations → Health effect analysis

1. (Sparse) PCA

2. Multivariate spatial prediction modeling (**low dimensional**)

3. Linear regression (with measurement error correction)

- Current progress and plan
  - Steps 1 + 2: Predictive sparse PCA (Jandarov/Szpiro)
  - Step 2: Low-rank version of common component model for co-kriging (Bergen/Szpiro)
  - Step 3: Measurement error correction with multiple spatially misaligned pollutants (Bergen/Szpiro)

Multivariate spatial prediction

- Previously discussed options
  - Unlinked models for each pollutant: kriging or low-rank splines
  - Linked models for multiple pollutants: co-kriging or splines with correlated coefficients
- None of the existing models are quite what we want
  - Co-kriging models only allow sharing information for spatial smoothing, not GIS covariates
  - Full-rank models are difficult to incorporate in predictive sPCA and measurement error correction
  - Straightforward correlation between spline/GIS coefficients does not accommodate pollutants sharing some, but not all, spatial structure
- Proposed new method (LR-CCM)
  - Low-rank version of common component co-kriging (Diggle and Ribeiro)
    - Essentially a clever low-rank spline model with correlated coefficients

Potential roles for LR-CCM

- Enhance predictive sparse PCA (steps 1 and 2) by linking spatial models for individual components
  - Requires two (non-trivial) generalizations of predictive sparse PCA: L^2 penalty on $\boldsymbol{\beta}$ and joint selection of multiple PCs
- Broaden applicability of multivariate measurement error correction (step 3) by including linked prediction models
  - We will see that bias correction fundamentally depends on covariance of coefficients in prediction models (conditional on penalty parameters)

Three step sequential procedure

High dimensional monitoring data → PC profiles + scores at monitor locations → PC scores at subject locations → Health effect analysis

1. (Sparse) PCA

2. Multivariate spatial prediction modeling (**low dimensional**)

3. Linear regression (with measurement error correction)

- Current progress and plan
  - Steps 1 + 2: Predictive sparse PCA (Jandarov/Szpiro)
  - Step 2: Low-rank version of common component model for co-kriging (Bergen/Szpiro)
  - Step 3: Measurement error correction with multiple spatially misaligned pollutants (Bergen/Szpiro)

Measurement error correction

- By the time we get to health effect estimation in Step 3, we will have already
  - Reduced the dimension of the exposure surface from p to k
  - Used spatial modeling to predict the k unmeasured exposure scores at subject locations
- Ideally account for uncertainty from both of these steps
- Dealing with dimensions reduction step conceptually awkward
  - We are not assuming a true latent reduced dimension exposure
  - Interpretation of regression coefficients changes with different reduced dimension representations
- We focus on measurement error from spatial misalignment, taking the dimension reduction as given

Relationship with previous measurement error work

- Previous presentation on measurement error (Bergen, 2012 CCAR SAC in Seattle)
- $PM_{2.5}$ components (S, Si, EC, OC) and carotid intima medial thickness (CIMT) in MESA
  - Separate analyses for each component
  - Full-rank universal kriging prediction models
  - Computationally efficient parameter bootstrap (PB) approximates parametric bootstrap; version of simulation extrapolation (PB-SIMEX)
- Crucially, **assume** kriging model corresponds to a random spatial exposure surface in the data-generating mechanism
  - Extension to multivariate analyses is “straightforward” but also “problematic” (Szpiro, 2012 CLARC meeting in Boston)

Extending to multi-pollutant framework

- To extend parametric methods, we "just" need to specify a multivariate exposure model and the corresponding health model
  - Inference by parametric bootstrap works as in single-pollutant case
- However …
  - The assumption that the single pollutant exposure surface is random is not particularly plausible
  - Extending this to multiple pollutants is even less plausible, and requires a fairly good model for how the pollutants are linked

Semi-parametric measurement error framework

- Assumptions about randomness
  - Regard the exposure as an unknown spatial surface, but one that is fixed across repeated experiments
  - For multiple pollutants, we have multiple surfaces that are likely to be correlated, **but we do not need to correctly model the correlation**
  - Subject and monitor locations are what change across experiments
- Exposure modeling by low-rank penalized spatial models
  - With sufficient degrees of freedom and well chosen penalty, works essentially as well as full-rank modeling
  - Extension to multiple pollutants is conceptually straightforward; can use linked exposure models (i.e., borrow information between multiple pollutants), but this is not essential
  - Asymptotic analysis is tractable
  - Connects well with predictive sPCA

See: *Silas Bergen slides*

- *Measurement error with spatially misaligned data and penalized regression exposure model*
- *Bias estimate that accounts for penalty parameter, use for*
  - *penalty parameter selection*
  - *post-hoc bias correction*
- *Bootstrap standard errors*
- *Simulation studies*
- *Application to Sister Study blood pressure analysis*
- *Plans for extension to multiple pollutants*

Summary of methods development

High dimensional monitoring data → PC profiles + scores at monitor locations → PC scores at subject locations → Health effect analysis

1. Sparse PCA

2. Multivariate spatial prediction modeling (**low dimensional**)

3. Linear regression (with measurement error correction)

- Current progress and plan
  - Steps 1 + 2: Predictive sparse PCA (Jandarov/Szpiro)
  - Step 2: Low-rank version of common component model for co-kriging (Bergen/Szpiro)
  - Step 3: Measurement error correction with multiple spatially misaligned pollutants (Bergen/Szpiro)

A novel principal component analysis for spatially-misaligned multivariate air pollution data

Roman Jandarov

UW CCAR

SAC Meeting

23 July 2013

Objective

- **This talk**: Develop a multi-pollutant exposure model
- Estimate health effect of multi-pollutant exposure on blood pressure (using Sister Study data)
- Generalize the methodology to data from Project 1

Description of multi-pollutant data

- Multi-pollutant data from EPA for 2009-2010
  - Monitors are dispersed throughout the lower 48 states
  - Measured pollutants $(P_1, \ldots, P_m)$:
    - Particles: PM25, PM10
    - Gases: NO2, NOx, SO2, O3, EC, OC, SO4, NO3
    - Elements: Al, As, Br, Cd, Ca, Cr, Cu, Co, Fe, K, Mn, Na, S, Si, Ni, V, Zn
- GIS covariates and coordinates at monitor locations
  - Let **Z** be a matrix of 30 PCs from geographical covariates and thin-plate splines from location coordinates
  - Use **Z** for prediction

Data for health analysis

- Y - blood pressure
- Data on Y and subject-specific covariates from NIEHS Sister Study data (cohort study on risk factors for breast cancer)
  - $> 50,000$ sisters of women with breast cancer enrolled from across the U.S.
  - Early analysis: association between PM2.5 exposure and Y (Van Hee et al, in preparation)

Challenges

- Dimensionality of multi-pollutant data
  - General health model is not practical

$$Y = \beta_0 + \sum_{l=1}^{m} \beta_l \widehat{P}_l + \text{interactions} + \text{covariates} + \ldots$$

  - Pollutant concentrations are potentiality correlated
  - Large number of main effects and interactions: hard to estimate and interpret
- Spatial misalignment of exposure and subject locations
  - Naive approach: one may need to build m exposure models to predict $(P_1, \ldots, P_m)$ for health analysis

A possible solution

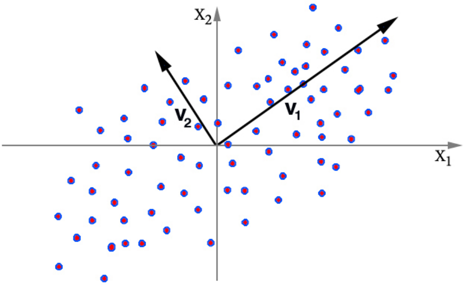
1. Dimension reduction
   - Compute first few principal components of multi-pollutant data
2. Predict scores obtained from principal components at participant locations using GIS covariates and splines
3. Fit a health model with smaller number of variables
   - Interpret coefficient of the model to identify important mixtures

This talk: Steps 1 and 2: Dimension reduction and prediction

Review of principal component analysis

- Principal component analysis (PCA) is a popular dimension reduction technique.
- A version of unsupervised learning
- Goal of PCA: Reduce the number of variables of interest into a smaller set of components
- PCA transforms the original variables into a new set of components (linear combinations of originals) equal to the number of original variables

PCA: Example

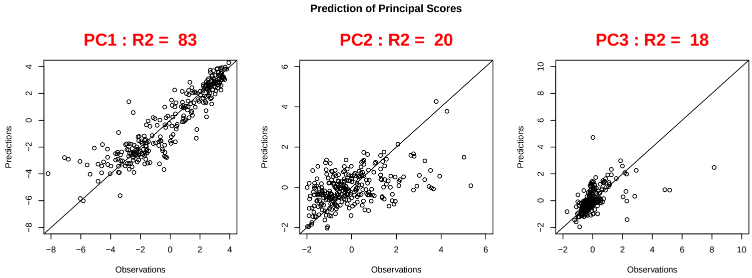


- Let $\mathbf{X}$ be a $n \times p$ matrix with standardized columns
- PCA finds direction $\mathbf{v}_1$ and $\mathbf{v}_2$ (also called loadings)
- Principal components: PC1 = $\mathbf{X}\mathbf{v}_1$, PC2 = $\mathbf{X}\mathbf{v}_2$

Sparse PCA

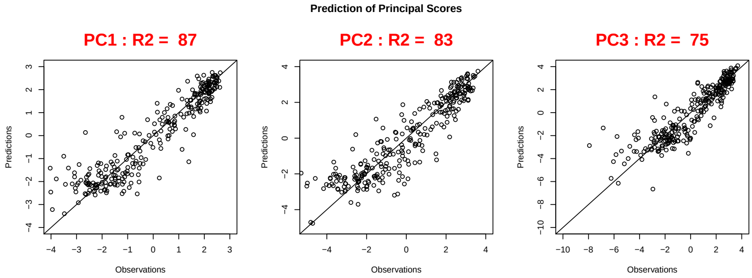
- Principal components (PCs) can sometimes be difficult to interpret
- Sparse PCA produces modified PCs with sparse loadings: loadings with only a few nonzero elements
- In sparse PCA, penalty parameter, λ, controls sparsity of loadings.
- In the context of spatial misalignment:
  - **sPCA**: λ is chosen to maximize predictability of pollutants from principal scores (obtained by projecting $\mathbf{X}$ onto $\mathbf{v}$ at monitor locations).
  - **sPCA-S**: λ is chosen to maximize spatial predictability of principal scores.

Issues with sparse PCA: Bad predictions



Penalty optimized to predict pollutants: principal components are difficult to predict (by Universal Kriging)

Issues with sparse PCA-S: Bad predictions



Penalty optimized to predict PCs: principal components are still difficult to predict (by Universal Kriging)

New approach: Idea

- We want a sparse PCA algorithm that results in principal components that can be predicted well
- Develop an algorithm that forces PCs to be close to spatial covariates

Low rank approximation to find PCs

- First, find $(\widetilde{\mathbf{u}}_1, \widetilde{\mathbf{v}}_1)$, s.t. $\|\widetilde{\mathbf{u}}_1\| = 1$ that minimizes

$$\|\mathbf{X} - \widetilde{\mathbf{u}}_1 \widetilde{\mathbf{v}}_1^T\|$$

  Define PC1 by $\mathbf{u}_1 = \widetilde{\mathbf{u}}_1 * \|\widetilde{\mathbf{v}}_1\|$. Loadings by $\mathbf{v}_1 = \widetilde{\mathbf{v}}_1 / \|\widetilde{\mathbf{v}}_1\|$
- Subsequently, find $(\widetilde{\mathbf{u}}_k, \widetilde{\mathbf{v}}_k)$ by approximating the corresponding residual matrices. Define corresponding PCs and loadings.
- Sparse PCA (Shen and Huang, 2008):
  Minimize a function of the form $\|\mathbf{X} - \widetilde{\mathbf{u}}\widetilde{\mathbf{v}}^T\| + P_\lambda(\widetilde{\mathbf{v}})$, where $P_\lambda()$ is an L_1 penalty parameterized by λ

Motivation: predictive sparse PCA (P-sPCA-S)

- Recall: $\mathbf{Z}$ - matrix of PCs from geographical covariates (and spatial splines)
  - Modify sPCA-S so that our PCs can be predicted well by $\mathbf{Z}$ (PCs $\approx \mathbf{Z}\beta$ by $\widetilde{\mathbf{u}} = \mathbf{Z}\beta$).
- At first step of the algorithm, minimize the following with respect to β and $\widetilde{\mathbf{v}}$ (using a fast alternating algorithm): $\|\mathbf{X} - \mathbf{Z}\beta\widetilde{\mathbf{v}}^T\| + P_\lambda(\widetilde{\mathbf{v}})$ with constraint $\|\mathbf{Z}\beta\|^2 = 1$
- Subsequent steps are based on using residual matrices $(\mathbf{X} - \mathbf{Z}\beta\widetilde{\mathbf{v}}^T)$
- Choose penalty parameter to maximize predictability of scores

Algorithm and definitions

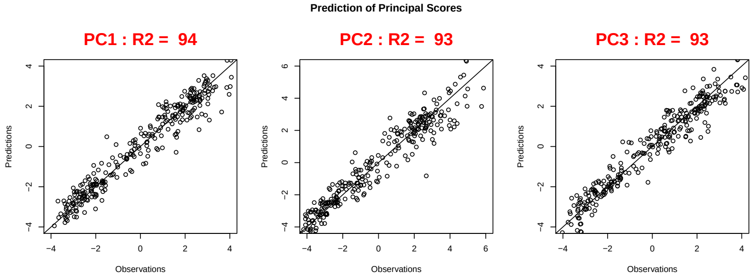
1. Obtain βs and $\widetilde{\mathbf{v}}$s
2. Loadings: $\mathbf{v} = \frac{\widetilde{\mathbf{v}}}{\|\widetilde{\mathbf{v}}\|}$
3. PC: $u_1 = X\mathbf{v_1}^T$, $u_2 = X\mathbf{v_2}^T$, ...

   (not observable at subject locations)
4. PC proxy:

   $u_{1,proxy} = Z\beta_1$, $u_{2,proxy} = Z\beta_2$, ...

   (observable at both monitor and subject locations)
5. Prediction using regression calibration for each PC:
   - Regress PC on PC proxy (all slopes < 1)
   - Use observed values of PC proxy to predict PC at subject locations

Predictive sPCA: Predictions



Obtained principal components can be predicted well

Predictive sPCA - S: Loadings for 3 PCs

	PC1	PC2	PC3
Mn	0	0	0
Ca	0	0	0
Na	0	0	0
Si	0	0	0
Al	0	0	0
Cr	0	0.25	0.06
Fe	0	0.14	0
Cu	0	0.4	0
K	0	0.4	0.11
Zn	0.11	0.34	0.24
V	0	0.37	0
Br	0.22	0.31	0.26
S	0.11	0	0.76
OC	0.45	0.2	0.26
As	0.47	0.3	0.15
EC	0.49	0.32	0.14
PM25	0.51	0.15	0.42

Prediction of pollutants from three PCs

	PCA-based			Individual	
	sPCA	sPCA-S	P- sPCA-S	Regression Splines	UK
Mn	0.88	0.47	0.44	0.45	0.29
Ca	0.45	0.11	0.14	0.53	0.49
Na	0	0.05	0.1	0.52	0.5
Si	0.74	0.09	0.12	0.59	0.59
Al	0.66	0.07	0.11	0.59	0.59
Cr	0.34	0.08	0.6	0.62	0.52
Fe	0.34	0.71	0.66	0.67	0.57
Cu	0.41	0.43	0.78	0.65	0.58
K	0.59	0.56	0.7	0.7	0.62
Zn	0.1	0.58	0.72	0.73	0.7
V	0.33	0.36	0.65	0.79	0.8
Br	0.74	0.87	0.81	0.77	0.84
S	0.54	0.73	0.9	0.9	0.97
OC	0.82	0.95	0.89	0.87	0.86
As	0.84	0.86	0.91	0.89	0.87
EC	0.86	0.9	0.85	0.91	0.88
PM25	0.87	0.95	0.9	0.91	0.93

Poor predictions from Regression and UK

sPCA is better than sPCA-S and P-sPCA-S

P-sPCA-S places zeros

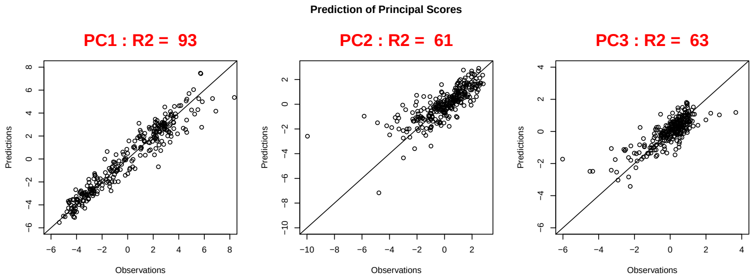
P-sPCA-S is better than sPCA and sPCA-S

All methods are close

Predictive sPCA (no penalty): Loadings

	P-sPCA with penalty			P-sPCA without penalty		
	PC1	PC2	PC3	PC1	PC2	PC3
Mn	0	0	0	0.2	-0.1	0.247
Ca	0	0	0	0.1	-0.5	0.129
Na	0	0	0	0	-0.1	-0.8
Si	0	0	0	-0.1	-0.6	-0.05
Al	0	0	0	0	-0.6	-0.05
Cr	0	0.25	0.06	0.25	0.02	0.09
Fe	0	0.14	0	0.24	-0.2	0.153
Cu	0	0.4	0	0.26	0.01	-0.01
K	0	0.4	0.11	0.26	-0.1	-0.15
Zn	0.11	0.34	0.24	0.28	0.02	0.179
V	0	0.37	0	0.25	0.05	-0.42
Br	0.22	0.31	0.26	0.29	-0.1	-0.08
S	0.11	0	0.76	0.26	0.06	0
OC	0.45	0.2	0.26	0.3	0.1	-0.04
As	0.47	0.3	0.15	0.31	0.08	0.046
EC	0.49	0.32	0.14	0.32	0.06	0
PM25	0.51	0.15	0.42	0.31	0.01	-0.02

Predictive sPCA (no penalty): Predictions



Summary of the approach

- Developed by adding a constraint to the traditional sparse PCA
- Results in improved predictability of PCs:

	PC1	PC2	PC3
Sparse PCA (penalty maximizes pollutants)	0.83	0.20	0.18
Sparse PCA –S (penalty maximizes scores)	0.87	0.83	0.75
Predictive sPCA –S (penalty maximizes scores)	**0.94**	**0.93**	**0.93**

- Can be used to predict most of the original pollutants well

Current and future work

- Current work:
  - Additional penalty parameter to penalize regression coefficients β can be added
  - Number of principal components to use in further analysis
  - A simulation study
  - Interpretation of obtained mixtures
    - Loadings are not orthogonal
    - Principal components are correlated
  - Health model analysis
- Future work:
  - Spatial all-at-once dimension reduction approach (reduced rank regression)
  - Generalization of methods to Project 1 data

Thank you!

Measurement error with penalized regression exposure modeling

UW CCAR
SAC Meeting
July 23, 2013

Background: 2-stage air pollution epidemiology studies

- Goal: Assess association between long-term air pollution exposure and continuous health outcome
- Exposures at subject locations are unobserved
- Stage 1: Exposure modeling
  - Build exposure models using monitoring data at, e.g., EPA AQS monitoring locations
  - Predict at subject locations
- Stage 2: Health modeling
  - Use predicted exposures in health model
  - Measurement error

Stage 1: Exposure modeling

- Full-rank models (e.g. universal kriging, full-rank thin-plate splines)
  - Can be thought of as penalized regression with same number of basis functions as monitoring locations
  - Can be computationally demanding
  - Standard asymptotics don't apply: number of exposure model parameters increases with sample size
- Low-rank models (e.g. low-rank kriging, thin-plate regression splines)
  - Similar to full-rank, but with fixed number of basis functions
  - Good approximation to full-rank models
  - Penalization (λ) necessary if number of basis function large relative to sample size

Stage 2: Measurement error

What happens?

- Small $\lambda \implies$ more variable exposure model coefficients
- "Estimation error"
- Large $\lambda \implies$ smoother predicted surface
- "Smoothing error"
- Both errors can bias health effect estimate, inflate its SE

Where are we headed?

1. Estimate bias as function of λ
2. Choose λ to balance bias from smoothing and estimation error
3. Correct for residual bias

Modeling assumptions

Health model:

$$y' = \beta_0 + \beta x' + \beta_z^T z' + \epsilon'$$

- Do not observe $x' = \Phi(s') + \eta'$ at subject locations s'
- Observe locations $x = \Phi(s) + \eta$ at monitoring locations s
- $s, s' \overset{iid}{\sim} G(\cdot)$ for some unknown $G(\cdot)$
- Also observe $\mathbf{r}(s)$, $\mathbf{r}(s')$ (geographic covariates, spatial splines)

Penalized exposure model

- For $\lambda \geq 0$, positive semi-definite $\mathbf{D}$:

$$\hat{\gamma}_\lambda = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \frac{1}{n} \sum_{i=1}^{n} \left(x_i - \mathbf{r}(s_i)^T \boldsymbol{\theta} \right)^2 + \lambda \boldsymbol{\theta}^T \mathbf{D} \boldsymbol{\theta}$$

- Note that $\hat{\gamma}_\lambda$ estimates $\boldsymbol{\gamma}_\lambda$, where:

$$\boldsymbol{\gamma}_\lambda \equiv \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \int (x - \mathbf{r}(s)^T \boldsymbol{\theta})^2 dG(s) + \lambda \boldsymbol{\theta}^T \mathbf{D} \boldsymbol{\theta}$$

- Smoothed surface at subject locations: $\mathbf{r}(s')^T \boldsymbol{\gamma}_\lambda$
- Observed predictions: $\mathbf{r}(s')^T \hat{\gamma}_\lambda$

Measurement error

$$\begin{aligned} x' - \mathbf{r}(s')^T\hat{\boldsymbol{\gamma}}_\lambda &= \left(x' - \mathbf{r}(s')^T\boldsymbol{\gamma}_\lambda\right) \qquad &+ \left(\mathbf{r}(s')^T\boldsymbol{\gamma}_\lambda - \mathbf{r}(s')^T\hat{\boldsymbol{\gamma}}_\lambda\right) \\ &= u_\lambda^S(s') \qquad &+ u_\lambda^E(s') \end{aligned}$$

u_λ^S: *Smoothing* error

- Difference between true surface and surface smoothed with infinite n
- Large $\lambda \implies$ larger smoothing error
- Reminiscent of Berkson error

u_λ^E: *Estimation* error

- Error from having finite n to estimate $\boldsymbol{\gamma}_\lambda$
- Small $\lambda \implies$ larger estimation error
- Reminiscent of classical error

Both types can cause bias: $E(\hat{\beta}) = \beta + \beta\psi_\lambda^S + \beta\psi_\lambda^E$

$E(\hat{\beta}) = \beta + \beta\psi_\lambda^S + \beta\psi_\lambda^E$

Bias from smoothing error:

$$\psi_\lambda^S = \frac{\int \left(x - \mathbf{r}(s)^T\boldsymbol{\gamma}_\lambda\right)\left(\mathbf{r}(s)^T\boldsymbol{\gamma}_\lambda\right) dG(s)}{\int \left(\mathbf{r}(s)^T\boldsymbol{\gamma}_\lambda\right)^2 dG(s)}$$

- Smoothing error $(x - \mathbf{r}(s)^T\boldsymbol{\gamma}_\lambda)$ becomes residual in health model
- $\lambda = 0$:
  - Smoothing error is uncorrelated with $\mathbf{r}(s)^T\boldsymbol{\gamma}_0$
  - No bias ($\psi_0^S = 0$)
- $\lambda > 0$:
  - Smoothing error is correlated with $\mathbf{r}(s)^T\boldsymbol{\gamma}_\lambda$
  - "Omitted variable" bias ($\psi_\lambda^S \neq 0$)
- ψ_λ^S estimable with monitoring data, plug-in estimates of $\boldsymbol{\gamma}_\lambda$

$E(\hat{\beta}) = \beta + \beta\psi_\lambda^S + \beta\psi_\lambda^E$

Bias from estimation error:

$$\psi_\lambda^E = v_\lambda^T E_\lambda(\hat{\boldsymbol{\gamma}}_\lambda - \boldsymbol{\gamma}_\lambda) + tr\,(A_\lambda Cov_\lambda(\hat{\boldsymbol{\gamma}}_\lambda - \boldsymbol{\gamma}_\lambda))$$

- As $n \to \infty$, $\psi_\lambda^E \to 0$
- Small $\lambda \implies$ large $Cov_\lambda(\hat{\boldsymbol{\gamma}}_\lambda - \boldsymbol{\gamma}_\lambda)$
- $Cov_\lambda(\hat{\boldsymbol{\gamma}}_\lambda - \boldsymbol{\gamma}_\lambda)$ estimable by sandwich covariance in a way that incorporates λ
- $E_\lambda(\hat{\boldsymbol{\gamma}}_\lambda - \boldsymbol{\gamma}_\lambda)$ estimable by multinomial Taylor expansion
- v_λ and A_λ: functions of $\int \mathbf{r}(s)\mathbf{r}(s)^T dG(s)$, $\int (x - \mathbf{r}(s)^T \boldsymbol{\gamma}_\lambda)\mathbf{r}(s)dG(s)$ and $\boldsymbol{\gamma}_\lambda$; estimable with monitoring or subject data and plug-in estimates of $\boldsymbol{\gamma}_\lambda$

Measurement error correction

1. Select penalty parameter (λ), using standard approach (e.g. REML) or to optimize health modeling*
2. Predict exposures and plug into health model
3. Correct for bias using estimates of ψ_λ^S and ψ_λ^E
4. Calculate SE using non-parametric bootstrap (re-sampling subject and monitoring data), carrying out Steps 1–3 in each bootstrap sample

*Choosing λ to optimize health modeling:

(a) Derive estimate of $Var(\hat{\beta})$ valid for infinite n' (accounts for variability from estimation error only)

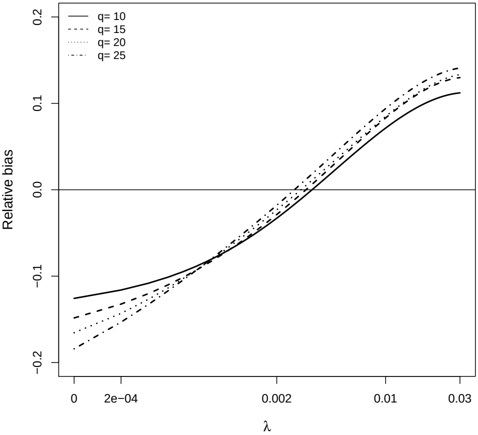
(b) Choose λ to minimize infinite-n' MSE of $\hat{\beta}$

(c) Aims to minimize bias from smoothing/estimation error while paying attention to variance from estimation error

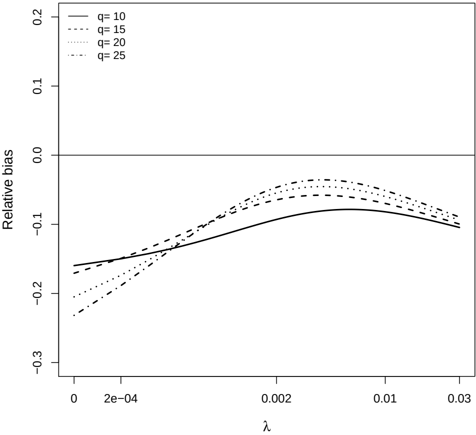
Simulations

- 4500×4500 grid
- $\Phi(s)$: fixed realization of spatially correlated stochastic process, plus 6 uncorrelated "geographic covariates"
- Exposure models: low-rank kriging (Kamman & Wand, 2003); $n = 100$; 10, 15, 20, 25 knots
- Health models:
  - Scenario 1: $y' = \beta_0 + 0.1x' + \epsilon'$
  - Scenario 2: $y' = \beta_0 + 0.1x' + \beta_Z^T z' + \epsilon'$
  - z': thin-plate spline with 8 degrees of freedom
  - $n' = 1000$

Relative bias

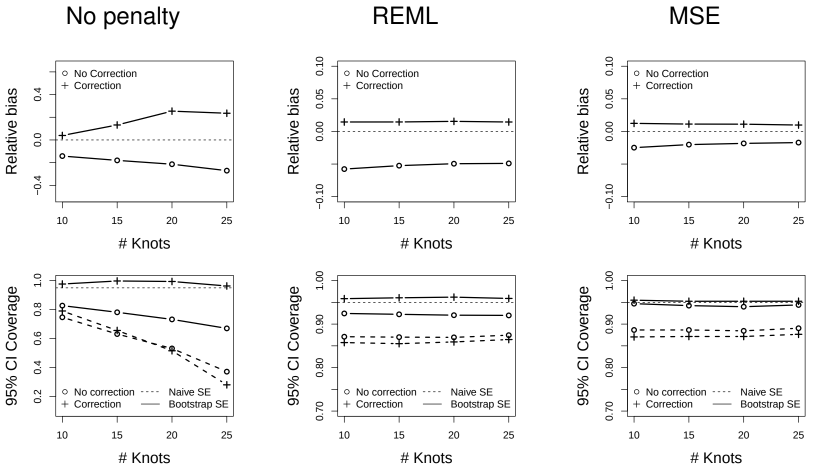


Scenario 1: No TPS in health model

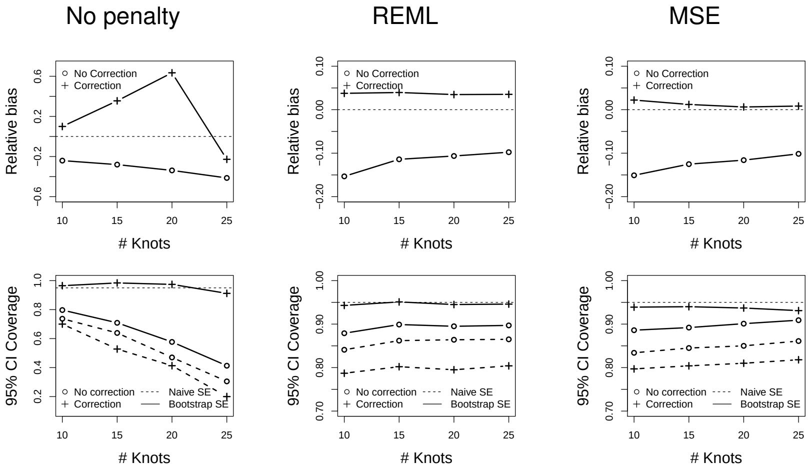


Scenario 2: With TPS in health model

Scenario 1: No TPS in health model

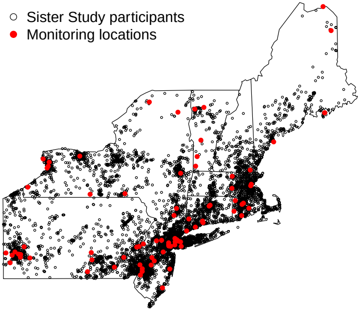


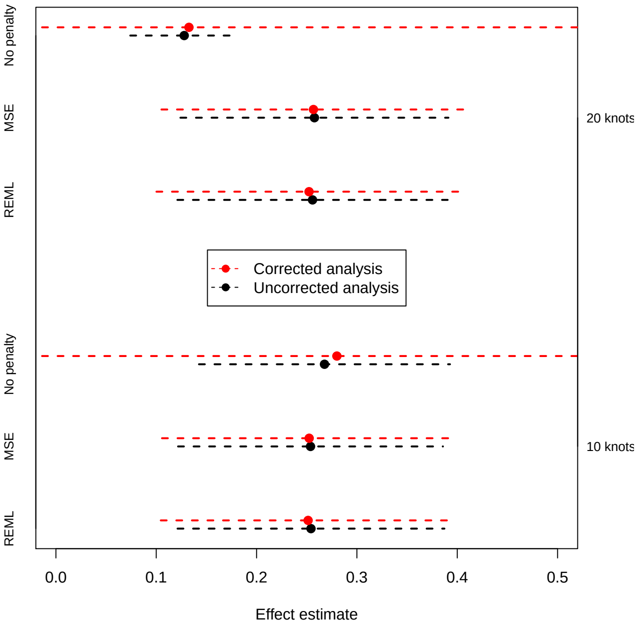
Scenario 2: With TPS in health model



Application: NIEHS Sister Study cohort

- Nationwide prospective cohort study
- Sisters of women with breast cancer
- VanHee et al: $10\mu g/m^3$ increase in annual avg $PM_{2.5}$ associated with 1.2mmHg increase in SBP (95% CI: 0.5, 1.8)
- Our application: 9 northeastern states (nonlinear health effect on national scale)





No penalty
MSE
REML
No penalty
MSE
REML
20 knots
10 knots
Corrected analysis
Uncorrected analysis
0.0
0.1
0.2
0.3
0.4
0.5
Effect estimate

Summary

- Penalization necessary to reduce drastic bias/variance introduced by estimation error
- Can estimate bias correction from available data; needed for accurate CI coverage
- Methods can be readily extended to multipollutant settings; just need moments of $(\hat{\gamma}_\lambda - \gamma_\lambda)$
- Multipollutant setting: $\hat{\gamma}_\lambda$ now vector of combined coefficients for the different pollutants
- Immediate application: nonlinear association of SBP with $PM_{2.5}$ in nationwide Sister Study; must predict $PM_{2.5}$ and $(PM_{2.5})^2$

References

Kammann EE and Wand MP. (2003). Geoadditive models. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*. 52(1):1-18.

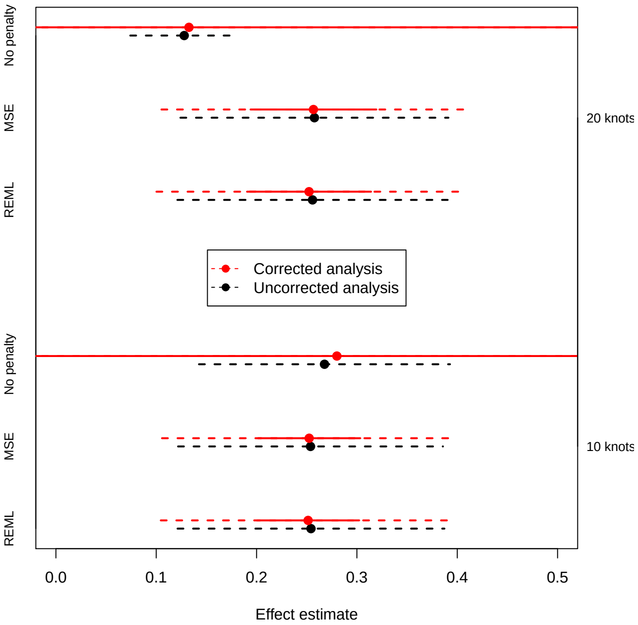
Van Hee VC, Chan SH, Szpiro AA, Oron AP, DeRoo LA, London SJ, et al. Long term air pollution exposure and blood pressure in the Sister Study. Submitted.

Yu Y and Ruppert D. (2002). Penalized spline estimation for partially linear single-index models. *Journal of the American Statistical Association*. 97(460):1042-1054.

Acknowledgements

This presentation was made possible by USEPA grant (RD-83479601-0). Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the USEPA. Further, USEPA does not endorse the purchase of any commercial products or services mentioned in the publication.

This presentation was also made possible by NIEHS Grants T32-ES015459 and R01-ES009411.


No penalty
MSE
REML
No penalty
MSE
REML
20 knots
10 knots
Corrected analysis
Uncorrected analysis
0.0
0.1
0.2
0.3
0.4
0.5
Effect estimate

Lemma 1: Suppose either Assumption 1 or 2 is met. Let $\mathbf{r}^c(\mathbf{s})$ contain elements $r_k^c(\mathbf{s}) = r_k(\mathbf{s}) - \Theta(\mathbf{s})^T\varphi_k$, where $\varphi_k = \text{argmin}_\omega \int (r_k(\mathbf{s}) - \Theta(\mathbf{s})^T\omega)^2 dG(\mathbf{s})$ for $k \in \{1, ..., p+q\}$. Let $w_\lambda^c(\mathbf{s}_i) = \mathbf{r}^c(\mathbf{s}_i)^T\boldsymbol{\gamma}_\lambda$, $\hat{w}_\lambda^c(\mathbf{s}_i) = \mathbf{r}^c(\mathbf{s}_i)^T\hat{\boldsymbol{\gamma}}_\lambda$, and $u_\lambda^S(\mathbf{s}) = \boldsymbol{\Phi}(\mathbf{s}) - \mathbf{r}(\mathbf{s})^T\boldsymbol{\gamma}_\lambda$.

With $\psi_\lambda = \dfrac{\int u_\lambda^S(\mathbf{s}) w_\lambda^c(\mathbf{s}) dG(\mathbf{s})}{\int w_\lambda^c(\mathbf{s})^2 dG(\mathbf{s})}$;

$$
\begin{aligned}
E_{[n^*]}\left(\frac{\hat{\beta}_{n^*} - \beta}{\beta} - \psi_\lambda\right) &= 2\frac{\int w_\lambda^c(\mathbf{s}_1) w_\lambda^c(\mathbf{s}_2) \mathrm{Cov}_{[n^*]}(\hat{w}_\lambda^c(\mathbf{s}_1), \hat{w}_\lambda^c(\mathbf{s}_2)) dG(\mathbf{s}_1) dG(\mathbf{s}_2)}{(\int w_\lambda^c(\mathbf{s})^2 dG(\mathbf{s}))^2}(1+2\psi_\lambda) \\
&- \frac{\int \mathrm{Var}_{[n^*]}(\hat{w}_\lambda^c(\mathbf{s})) dG(\mathbf{s})}{\int w_\lambda^c(\mathbf{s})^2 dG(\mathbf{s})}(1+\psi_\lambda) - \frac{\int E_{[n^*]}(\hat{w}_\lambda^c(\mathbf{s}) - w_\lambda^c(\mathbf{s})) w_\lambda^c(\mathbf{s}) dG(\mathbf{s})}{\int w_\lambda^c(\mathbf{s})^2 dG(\mathbf{s})}(1+2\psi_\lambda) \\
&+ \frac{\int u_\lambda(\mathbf{s}) E_{[n^*]}(\hat{w}_\lambda^c(\mathbf{s}) - w_\lambda^c(\mathbf{s})) w_\lambda^c(\mathbf{s}) dG(\mathbf{s})}{\int w_\lambda^c(\mathbf{s})^2 dG(\mathbf{s})} - 2\frac{\int u_\lambda(\mathbf{s}_1) w_\lambda^c(\mathbf{s}_2) \mathrm{Cov}_{[n^*]}(\hat{w}_\lambda^c(\mathbf{s}_1), \hat{w}_\lambda^c(\mathbf{s}_2)) dG(\mathbf{s}_1) dG(\mathbf{s}_2)}{(\int w_\lambda^c(\mathbf{s})^2 dG(\mathbf{s}))^2} \\
&\equiv n^* b_{(\hat{\beta}_{n^*} - \beta)/\beta}
\end{aligned}
$$

- Estimable from available data
- $\lambda = 0 \implies \psi_\lambda = 0$
- Small $\lambda \implies$ large $\int w_\lambda^c(\mathbf{s})^2 dG(\mathbf{s})$
- Small $\lambda \implies$ large $\mathrm{Cov}_{[n^*]}(\hat{w}_\lambda^c(\mathbf{s}_1), \hat{w}_\lambda^c(\mathbf{s}_2))$

Cross-center collaborations

Other EPA Clean Air Research Centers (CLARCs)

- Emory/Georgia Tech (“SCAPE”)
- Harvard (“Harvard”)
- Michigan State/Michigan (“GLACIER”)

Collaboration specifics

- $50,000 per center per year
- Involves 2 or more CLARCs

Planning

- Discussed at the first CLARC annual meeting in May

UW CLARC collaborations

- Mobile sampling in Atlanta (with Emory) – Tim Larson, Mike Yost
- Toxicology (Michigan State) – Matt Campen
- Exposure measurement error correction (with Harvard and Emory) – Adam Szpiro
- Satellite (remote sensing) data for $PM_{2.5}$ (with Emory and Harvard) – Paul Sampson
- [Chamber characterization] (with Harvard and Emory) – Jake McDonald

Mobile & fixed site characterization in Atlanta

- With Emory (SCAPE) – Sept 2013

- 2-week sampling at central site and 20 (mobile and passive) fuzzy points x 3 routes, incl 2 roadside gradient routes + 3 fuzzy points with full instrumentation

- compare to CMAQ predictions (4 km grid, downscaled to 250 m using LUR model)

- Aims:
  1. validate mobile fuzzy point measurements
  2. complete another near roadway campaign

Animal toxicology

1. Michigan State (GLACIER) coarse PM human experimental subject serum and BAL samples (R Brook, U Michigan)
   - Campen ex vivo endothelial cell assays

2. Campen collaboration with Jesus Araujo (UCLA & GLACIER)
   - HDL dysfunction & oxidized lipids on samples from LRRI mouse studies

Exposure measurement error correction

- With Harvard and Emory

- Georgia birth cohort (low birth weight) and EPA $PM_{2.5}$

- Common $PM_{2.5}$ exposure predictions based on UW spatio-temporal model

- 3 statistical approaches for measurement error correction (with simulations for insight into differences):
  - parameter bootstrap - UW
  - simulation extrapolation - Harvard
  - Bayesian – Emory

Satellite $PM_{2.5}$ estimation

- With Emory and Harvard

- Standard set of data for North Carolina, 2006-08, 3-km grid; MODIS AOD data downloaded

- 6 candidate models for $PM_{2.5}$ prediction
  - Harvard x 2 (mixed effects, multi-level)
  - Emory x 3 (spatial downscaler, mixed effects, CMAQ)
  - UW x 1 (spatio-temporal model)
    - assess added value of satellite data

- commons metrics for model evaluation