RUTGERS
THE STATE UNIVERSITY OF NEW JERSEY

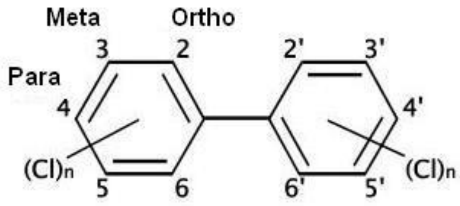
Green-Duwamish River Watershed PCB Congener Study Phase 2 Summary

Factor Analysis Results

Lisa A. Rodenburg
Department of Environmental Sciences
Rutgers, the State University of New Jersey

leidos

My research specializes in analyzing large data sets on PCBs and other pollutants



- New York/New Jersey Harbor
  - Water column, dischargers, sediment
- Delaware River PCB TMDL data
  - Water column, sediment, dischargers, air
- Integrated Atmospheric Deposition Network - Chicago
- San Francisco Bay
  - BDEs in sediment, PCBs in water
- Portland Harbor Superfund Site
  - Water column and sediment, biota
- Green River/Duwamish, Washington
  - Water, sediment, biota, air

Introduction

- Green-Duwamish River Watershed PCB Congener Study: Phase 1
- Phase 2
  - Initial Data Assessment
  - Source Evaluation

Phase 2: Objectives

- Identify PCB chemical signatures
- Determine the relative contribution of these source signatures
- Identify potentially known/unknown sources of and/or pathways for PCBs in the Green/Duwamish
- Recommend a set of PCBs (individual congeners and/or suites of congeners) to be included in modeling for the Green/Duwamish watershed PLA
- Provide recommendations for data collection and/or analysis approaches for future PCB congener data collection

Factor Analysis Equation

Applies to Principal Components Analysis, PMF etc.

View the PCB signal as a mixture of mixtures

Some of those mixtures are **Aroclors** …some are not.

Use this equation to predict concentration of each congener, based on number, fingerprint and concentration of sources.

You do NOT need any information about the sources, such as their fingerprints, or even how many there are!

$$X = G\,F + E$$

$$(m \times n) \quad (m \times p) \quad (p \times n)$$

X = input data matrix
G = matrix of conc of each factor in each sample generated by model
F = matrix of fingerprint of each factor (p) generated by model
E = leftover or residual
n = number of analytes
m = number of samples
p = number of factors (sources)

Note: in all forms of factor analysis, the user has to decide what is the 'correct' number of sources based on model output.

The Soda Analogy

- Several different soft drinks to choose from
- Sometimes kids like to mix these…

- Say we have 100 kids who made mixed drinks from the same soda fountain…

rodenburg@envsci.rutgers.edu

Analytes

- Sugar = most non-diet sodas
- Aspartame = some diet sodas
- Carmel coloring = most colas, root beer, etc.
- Citric acid = Sprite, 7-Up, some fruity drinks such as Cherry Coke, etc.
- Cola flavoring = most colas
- Caffeine = most colas

http://dchealthybites.com

Data matrix

	Caramel color	sugar	aspartame	citric acid	cola flavor-ing	caffeine
Anna	0.50	0.62	0.41	0.58	0.99	0.87
Bruce	0.58	0.86	0.25	0.78	0.35	0.14
Carlos	0.65	0.06	0.68	0.75	0.50	0.06
Donna	0.33	1.00	0.98	0.39	0.63	0.92
Emily	0.38	0.10	0.40	0.14	0.11	0.06
Francis	0.67	0.60	0.44	0.60	0.50	0.10
George	0.07	0.23	0.65	0.37	0.82	0.54
Harriet	0.95	0.53	0.02	0.25	0.51	0.86
Inga	0.46	0.67	0.19	0.92	0.23	0.45
John	0.32	0.97	0.79	0.19	0.88	0.21
Karl	0.81	0.42	0.68	0.70	0.15	0.08
Lisa	0.22	0.62	0.47	0.94	0.52	0.75
Michael	0.00	0.95	0.98	0.19	0.45	0.88
Nick	0.49	0.46	0.25	0.02	0.97	0.02
Olga	0.36	0.49	0.55	0.62	0.94	0.07

Concentrations (mg/L)

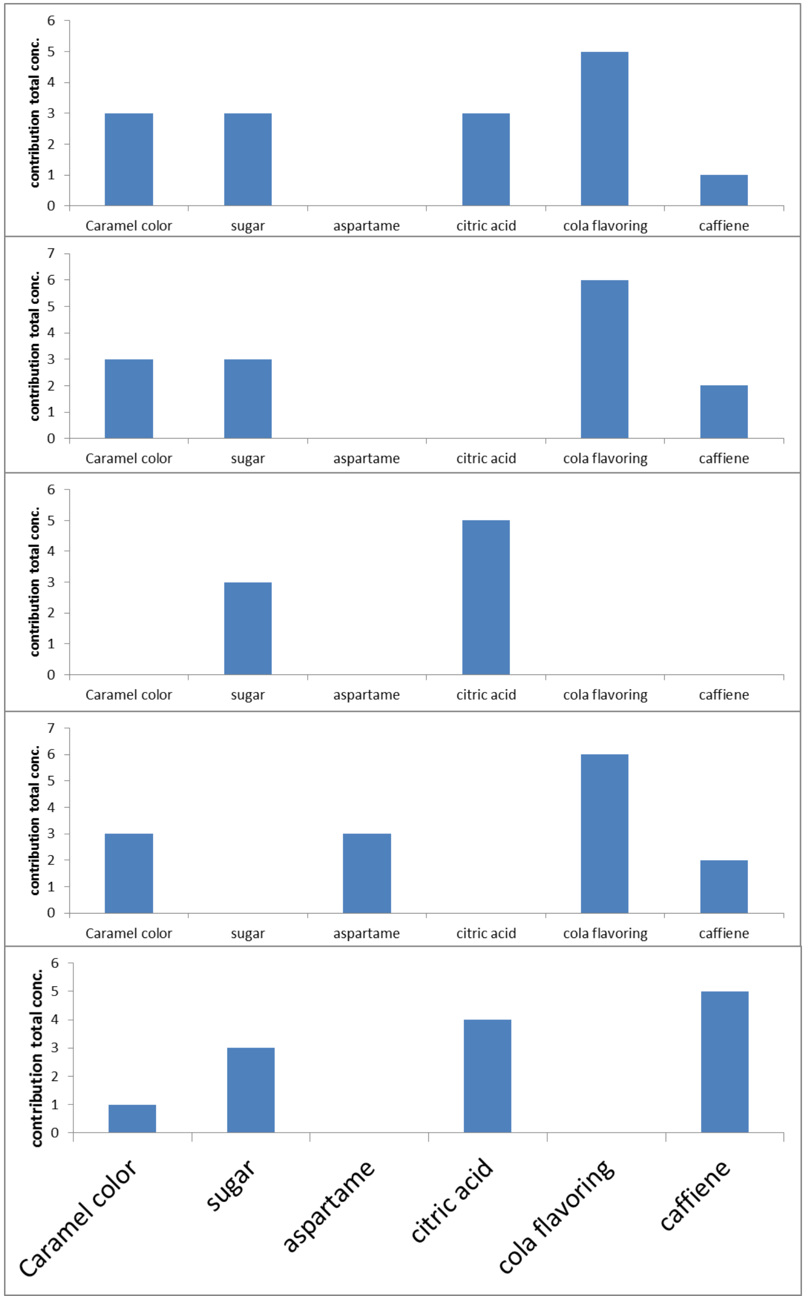
rodenburg@envsci.rutgers.edu

PMF results

PMF can tell you:

- How many sources (fingerprints, factors)
- Their fingerprints (F matrix)
- How abundant each fingerprint is in each sample (G matrix)

"F matrix"

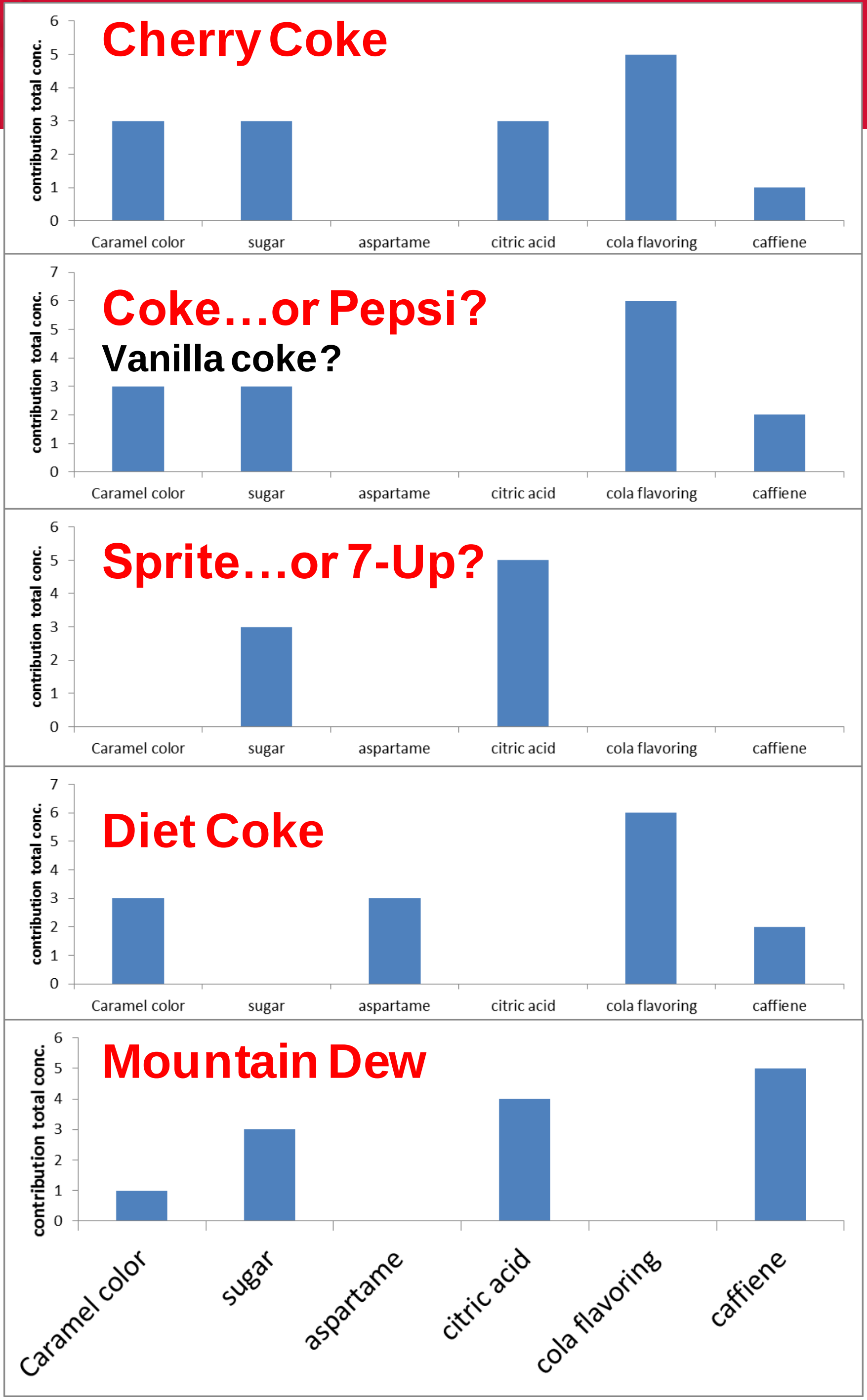


PMF results - F matrix Fingerprints

PMF **can't** tell you:

- What it all means

- **YOU** have to interpret this information



PMF Results – G matrix

- G matrix: abundance of each factor in each sample

- Helps with questions like:
  - Older people prefer diet soda?
  - Women prefer non-caffeinated drinks?
  - More caffeine consumed later at night?

	Cherry Coke	Coke	Sprite	Diet Coke	Mt Dew
Anna	16%	20%	13%	19%	32%
Bruce	20%	30%	9%	28%	13%
Carlos	25%	2%	26%	29%	19%
Donna	10%	30%	29%	12%	19%
Emily	34%	9%	35%	12%	10%
Francis	24%	21%	16%	21%	18%
George	3%	11%	30%	17%	38%
Harriet	42%	23%	1%	11%	23%
Inga	19%	27%	8%	37%	9%
John	10%	31%	25%	6%	28%
Karl	29%	15%	25%	25%	5%
Lisa	8%	22%	17%	34%	19%
Michael	0%	37%	38%	7%	18%
Nick	22%	21%	11%	1%	44%
Olga	12%	16%	19%	21%	32%

Rows sum to 100%→

Need ancillary info, such as age, gender, time of day etc.

rodenburg@envsci.rutgers.edu

Main PCB sources in most watersheds

- AROCLORS!
- Non-Aroclor congeners from pigments
- Reductive dechlorination of Aroclors by bacteria

rodenburg@envsci.rutgers.edu

Green-Duwamish Data sets analyzed by PMF

- % of mass = % of the total mass contained in all the data that was included in the PMF analysis
- Air and storm drain congener lists limited by number of samples
- Water congener list limited by large numbers of Below Detection Limit (BDL) values
- **Model solution must be consistent with everything you know of the system.**

compartment ->	air	sediment	surface water	tissue	storm drain
columns	SPB-octyl	SPB-octyl & SGE-HT8	SPB-octyl	SPB-octyl	SPB-octyl & DB-5
samples	64	146	209	128	74
peaks	64	80	42	90	73
congeners	100	154	69	135	142
% of mass	88%	94%	60%	96%	92%
% data points Below Detection Limit	18%	9%	30%	1.4%	15%

Includes duplicates

RUTGERS

Sediment

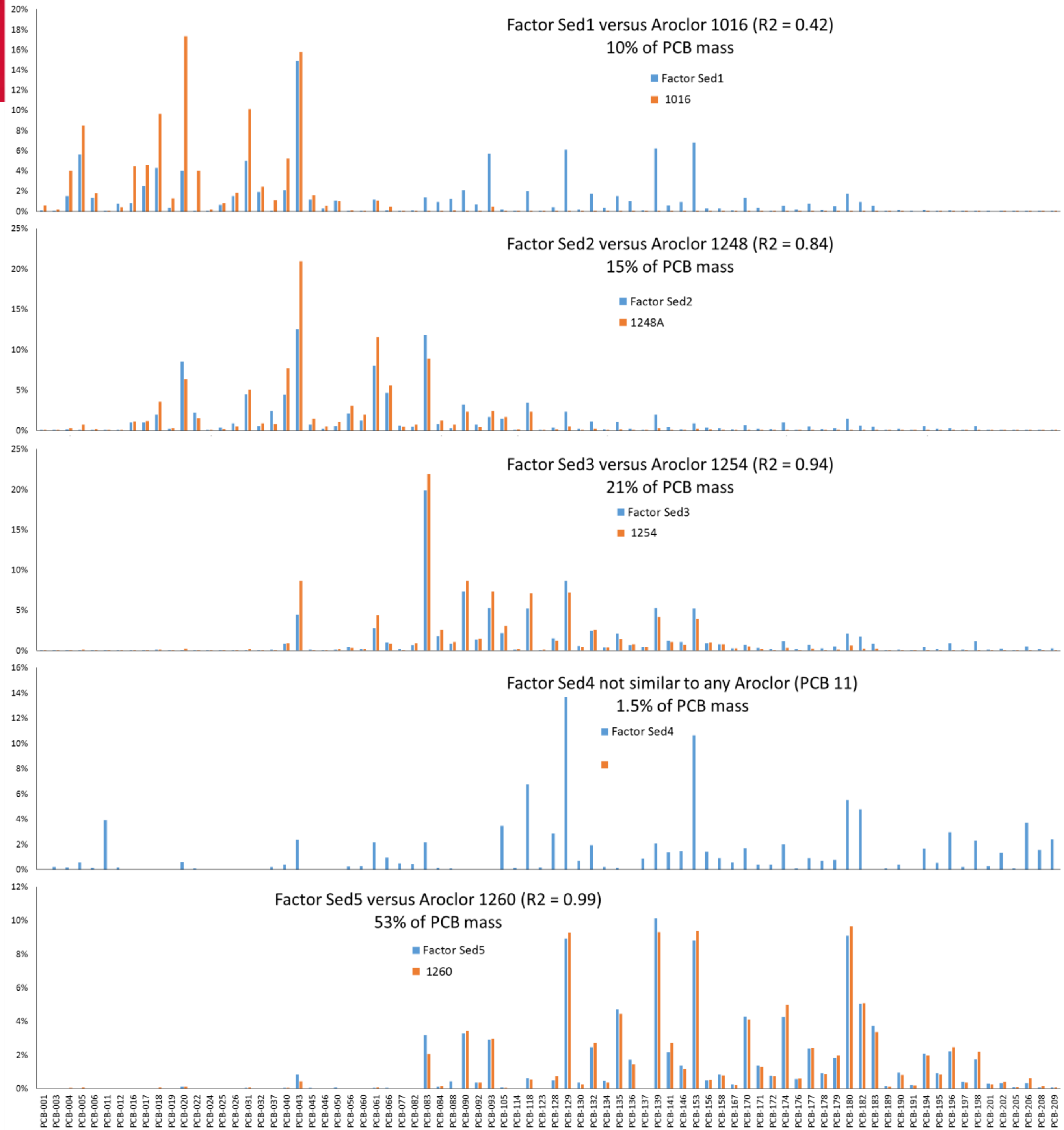
5 factors found:

4 similar to Aroclors

1260>>1254>1248>1016

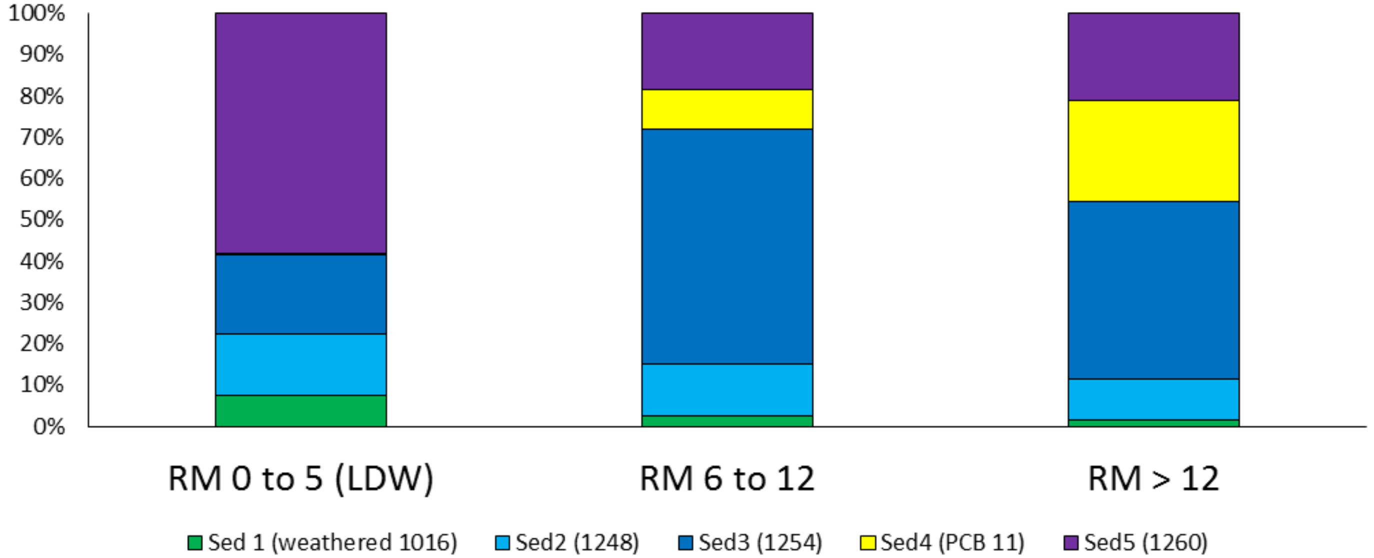
Sed4 not similar to Aroclors, contains a lot of PCB 11

- Wastewater/stormwater/CSO?
- Or atmospheric deposition?



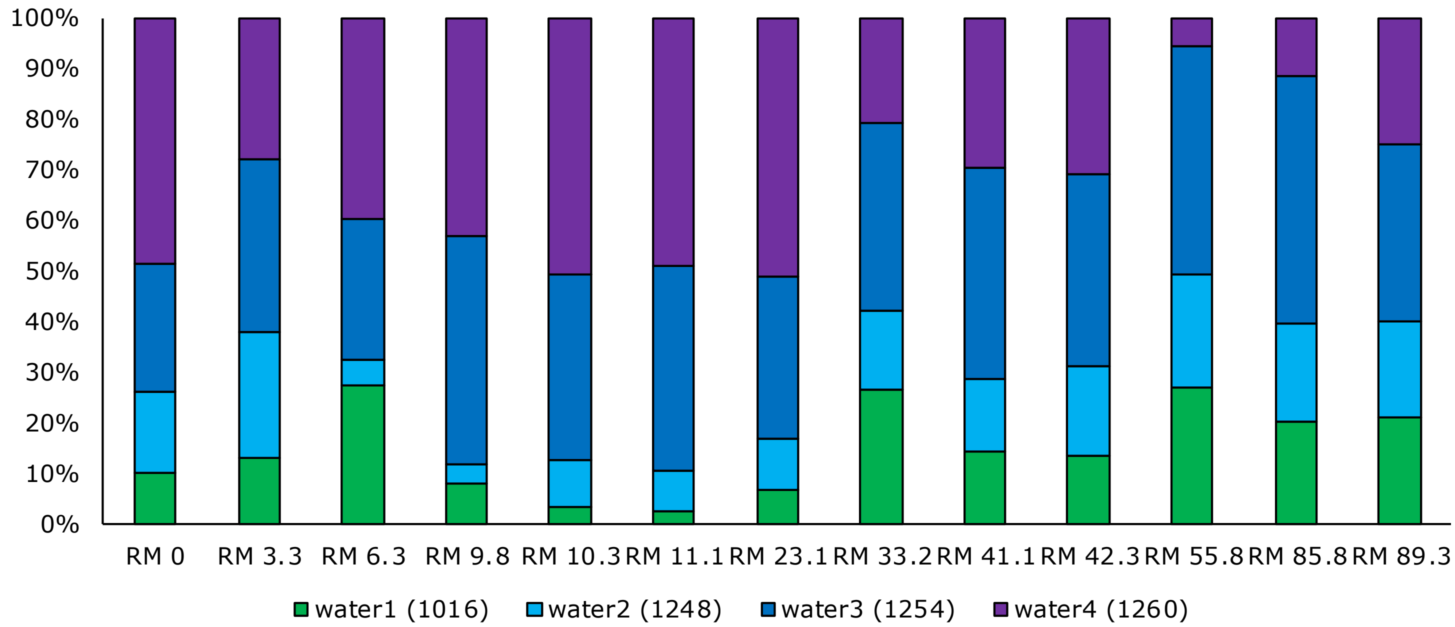
rodenburg@envsci.rutgers.edu

Sediment – spatial trends



- Sed5 (Aroclor 1260) dominates near river mouth
- Sed4 (PCB 11) more important upstream

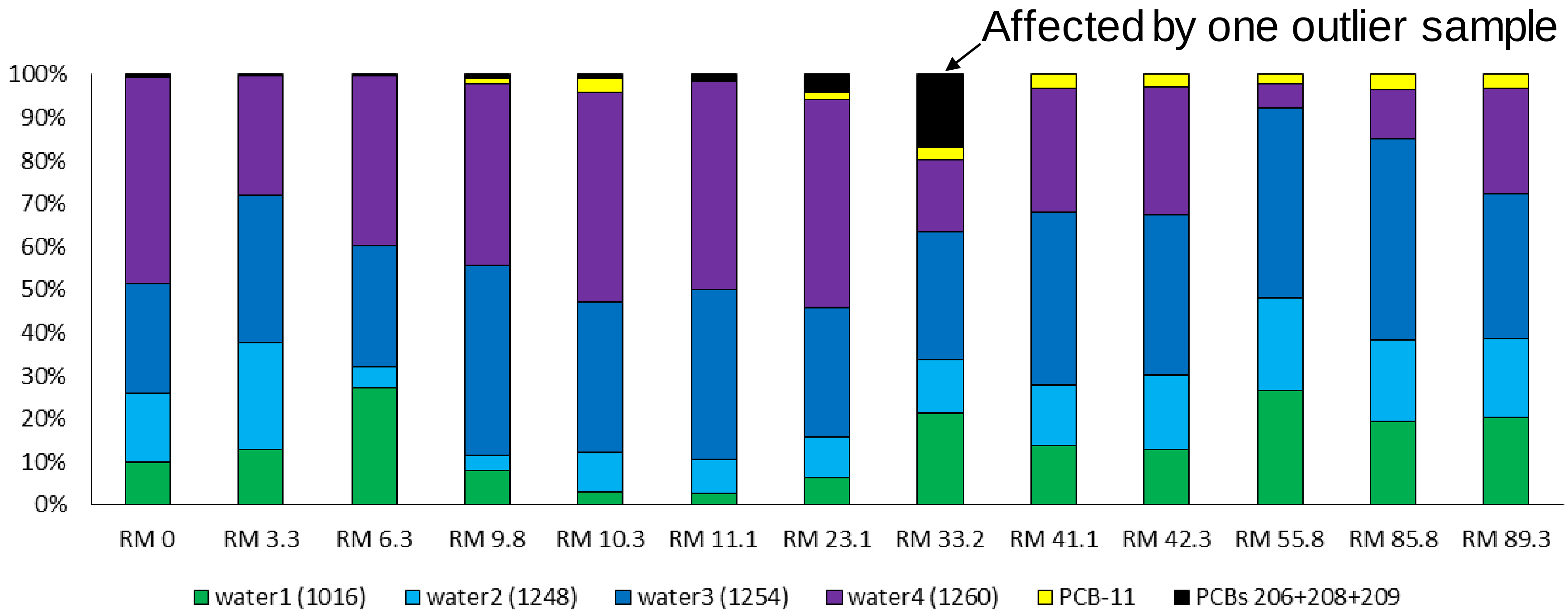
Surface water – spatial trends



- Mass-weighted average contribution to PCBs at each RM location
- Aroclor 1260 dominates nearer to river mouth
- PCB11, PCBs 206+208+209 were not included in the PMF model

rodenburg@envsci.rutgers.edu

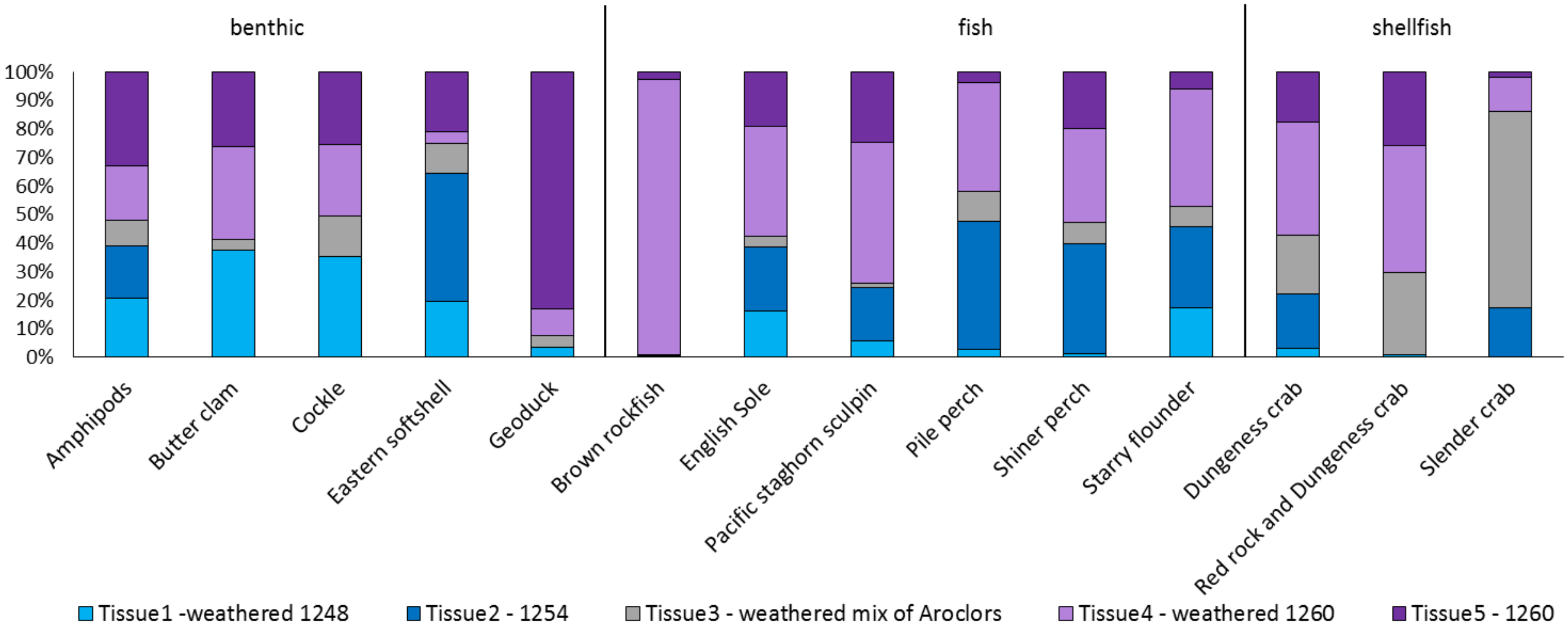
Surface Water – non-Aroclor congeners



- PCB11, PCBs 206+208+209 not very abundant in the water column

rodenburg@envsci.rutgers.edu

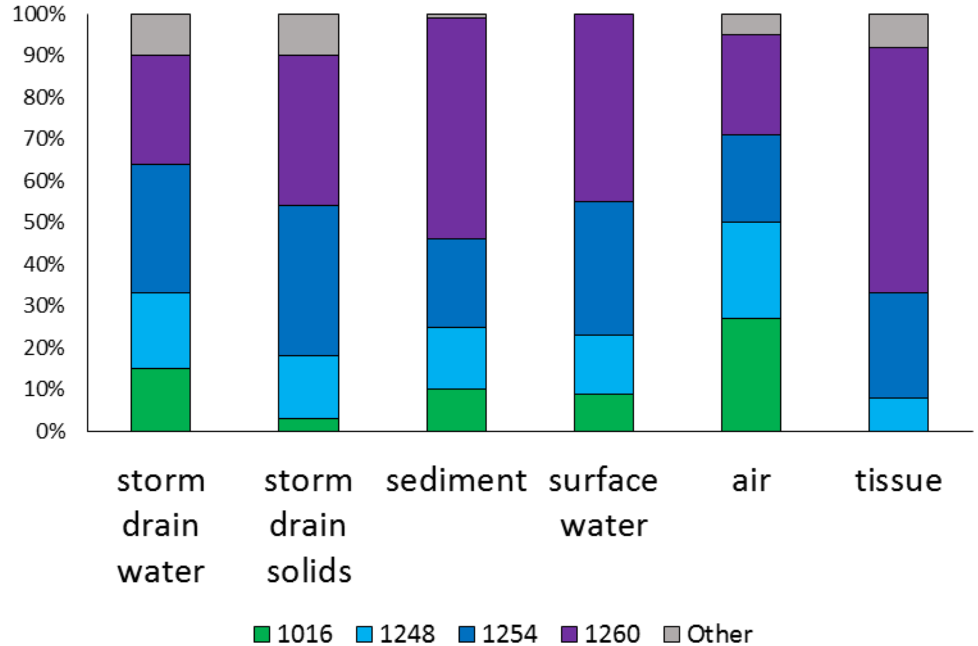
Tissue – by species



- Species vary in their ability to metabolize PCBs

rodenburg@envsci.rutgers.edu

Summary



Match (R^2) between Aroclors and factors for each compartment:

	compartment	1016	1248	1254	1260
closer	**storm drain**	0.96	0.86	0.86	0.98
↑	**sediment**	0.42	0.84	0.94	0.99
sources	**surface water**	0.73	0.44	0.84	0.91
↓	**air**	0.81	0.57	0.85	0.88
further	**tissue**	NA	0.43	0.7	0.84

Green-Duwamish Results

Types of sources:

- Across all five compartments, Aroclors are the dominant PCB sources
  - 1260 > 1254 >> 1248 > 1016/1242
- Non-Aroclor PCB sources are minor
- No dechlorination – probably due to salinity

Spatial trends in sources:

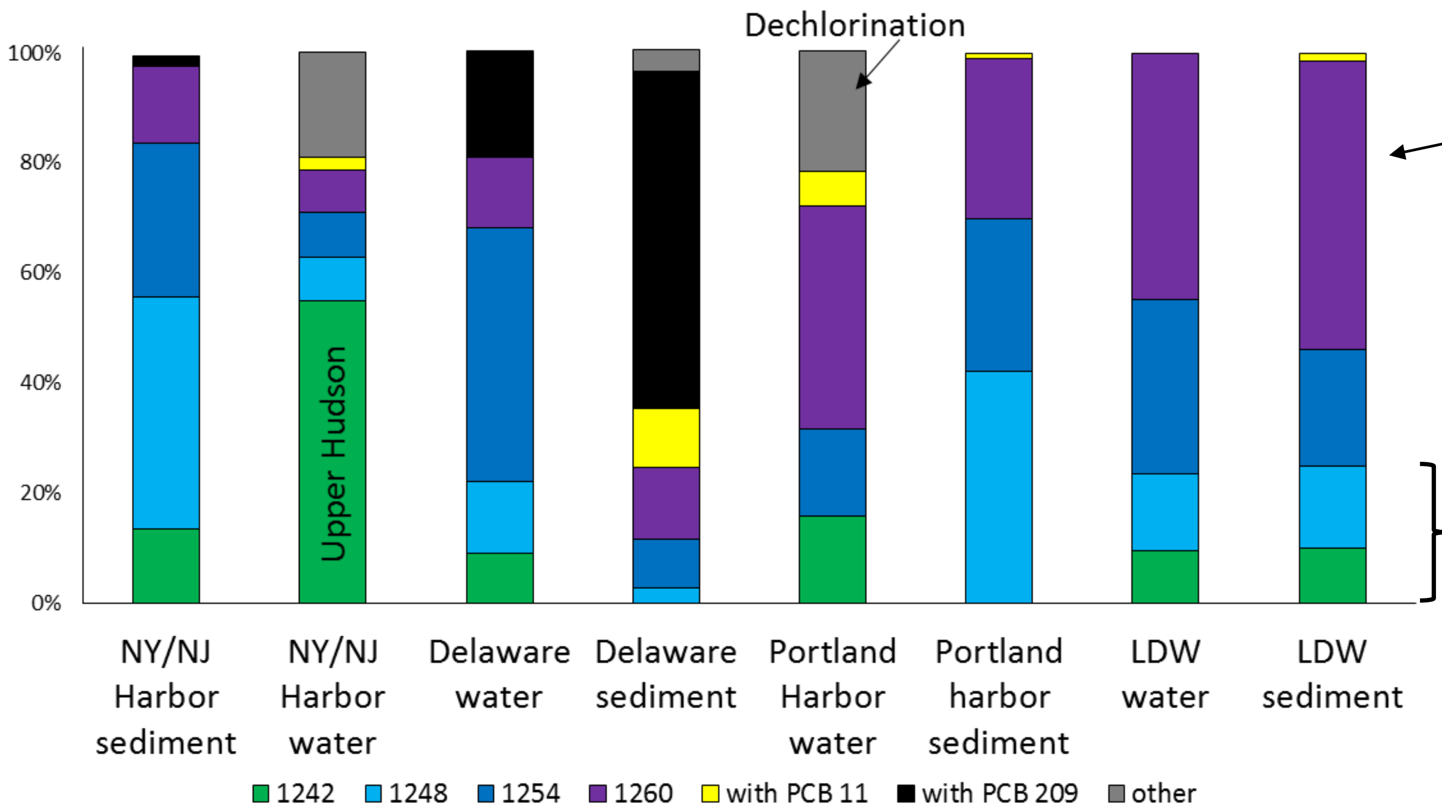
- Spatial trends are consistent across water, sediment, biota

Recommended options for modeling:

- Homologs 3 through 8
- Total PCBs

rodenburg@envsci.rutgers.edu

Comparisons to other watersheds



In Delaware and Portland Harbor, Aroclor 1260 was associated with shipyards

Often difficult for the model to discern between Aroclors 1242 and 1248

- Other systems have more ‘other’, more non-Aroclor, and often more dechlorination

Acknowledgments

- Work performed under sub-contract to Leidos and for WA Department of Ecology.

leidos

DEPARTMENT OF ECOLOGY State of Washington

- EPA – esp. Gary Norris

- Brittney Blackburne, undergraduate intern, wrote code to automate much of the PMF2 analysis

Further information

Lisa Rodenburg - Rutgers University
Lisa.Rodenburg@Rutgers.edu 848-932-5774

Iris Winstanley - Leidos
Iris.Winstanley@leidos.com 425-205-5189

Rick Thomas – Ecology Toxics Cleanup Program
rith461@ecy.wa.gov 425-649-7208

Rachel McCrea
rmcc461@ecy.wa.gov 425-649-7223

Uncertainty

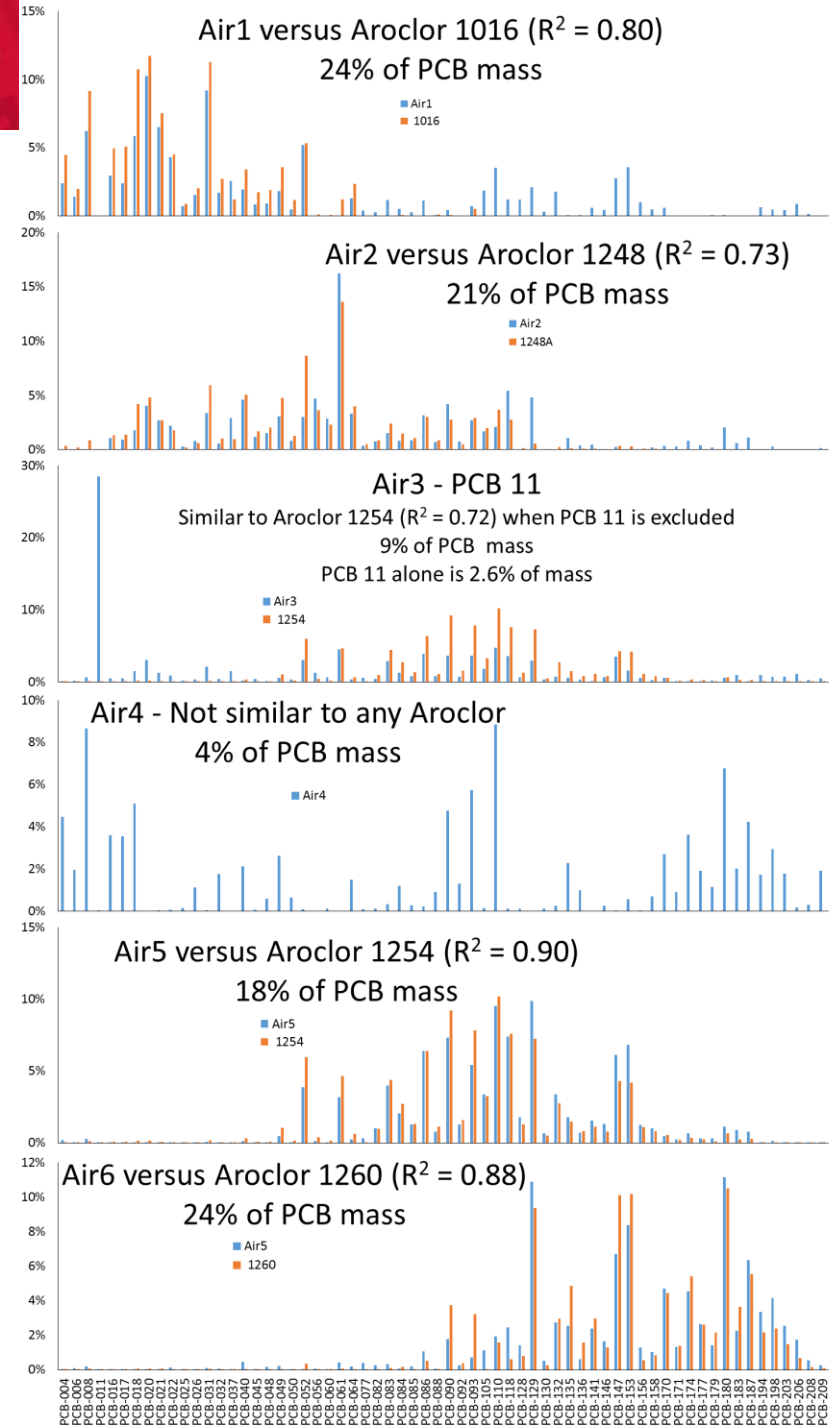
The PMF model and results are highly reproducible. This does not necessarily imply low uncertainty.

Uncertainty arises from:

- Insufficient data: not enough samples or detected analytes
  - Esp. for water compartment
- Different models may give different results for the same data
  - Tried PMF2 and PMF 5.0 – very different results
- Various permutations of the same data set may give different model results, even when the same model is used
  - We ran many permutations and got essentially the same results giving us higher confidence
- Choosing a sub-optimal number of factors
  - # of factors was relatively obvious for most compartments, less so for water
- Factors may be misinterpreted
  - Similarity between Aroclors?

Air (atmospheric deposition)

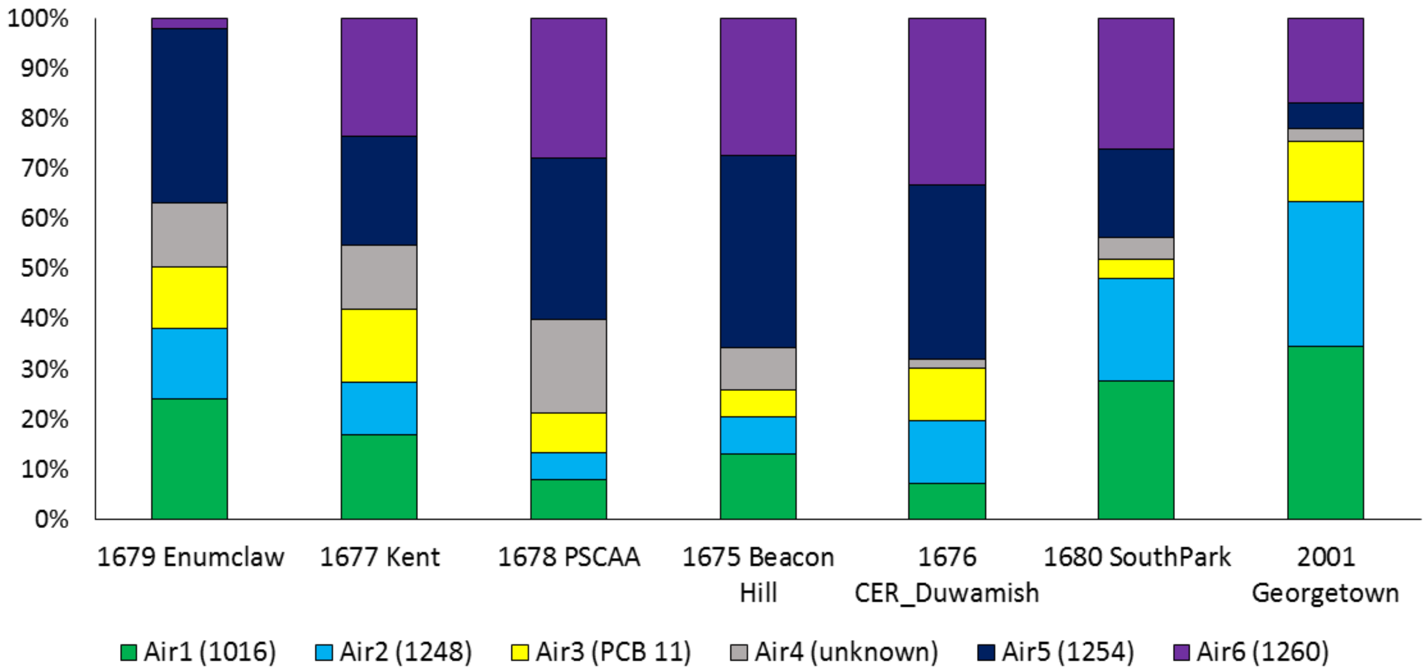
- 6 factors found:
- 4 surprisingly similar to Aroclors
- 1016 > 1260 > 1248 > 1254
- Lower MW formulations more abundant in the atm dep
- Air4 (5% of mass) does not resemble any Aroclor – composition is variable



rodenburg@envsci.rutgers.edu

Air – spatial trends

Higher PCB flux → more urban/industrial



- More 1260 in the more urban/industrial areas?
- No 'urban fractionation effect' – local sources?

Surface Water

- Four factors
- All resemble Aroclors
- Non-Aroclor congeners excluded