RUTGERS
THE STATE UNIVERSITY OF NEW JERSEY

Green-Duwamish River Watershed PCB Congener Study Phase 2 Summary

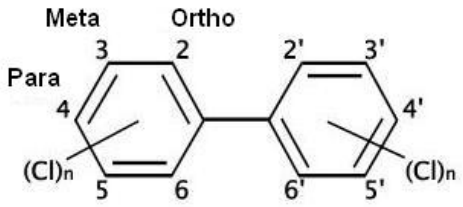
Factor Analysis Results

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

leidos

rodenburg@envsci.rutgers.edu

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

rodenburg@envsci.rutgers.edu

Introduction

- Green-Duwamish River Watershed PCB Congener Study: Phase 1
  - Described sources, fate & transport, and toxicity of PCBs
  - Discussed analytical methods and data comparability
  - Compiled congener data collected in the Green-Duwamish River watershed
  - Included 645 environmental samples collected and analyzed for 209 PCB congeners by Method 1668

rodenburg@envsci.rutgers.edu

Phase 2 - Initial Data Evaluation

Can the data be used for factor analysis?

Criteria for data to be used in factor analysis:

- Enough data: at least as many samples as congeners/peaks
- Enough data above detection
- Surrogate recoveries or duplicates needed to estimate uncertainty
  - duplicates should be assimilated with regular samples so statistical independence between samples is preserved for the analysis
- Detection limits are needed—must be estimated if not available

rodenburg@envsci.rutgers.edu

Phase 2: Source Evaluation

- 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

rodenburg@envsci.rutgers.edu

Additional considerations for fingerprinting analysis

- GC column and co-elution pattern
  - SPB-octyl is the main column used for 1668
  - SGE-HT8 and DB-5 or equivalent are alternatives
  - Very different co-elution patterns
- Blanks
- Obvious trends
  - Non-Aroclor PCBs
  - Aroclors

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.

rodenburg@envsci.rutgers.edu

Advantages of Positive Matrix Factorization

over other models, for example Principal Components Analysis

- Positive correlations only – mass balance model
  - Great for concentrations of contaminants
- Assign a point-by-point uncertainty estimate
- Missing and below detection limit values can be included by assigning them a high uncertainty
- “Robust” mode can be used so that outlier values will not skew the factor profiles
  - Data can span many orders of magnitude
- PMF provides the quantitative contribution estimate from each factor for each sample.

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

rodenburg@envsci.rutgers.edu

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)

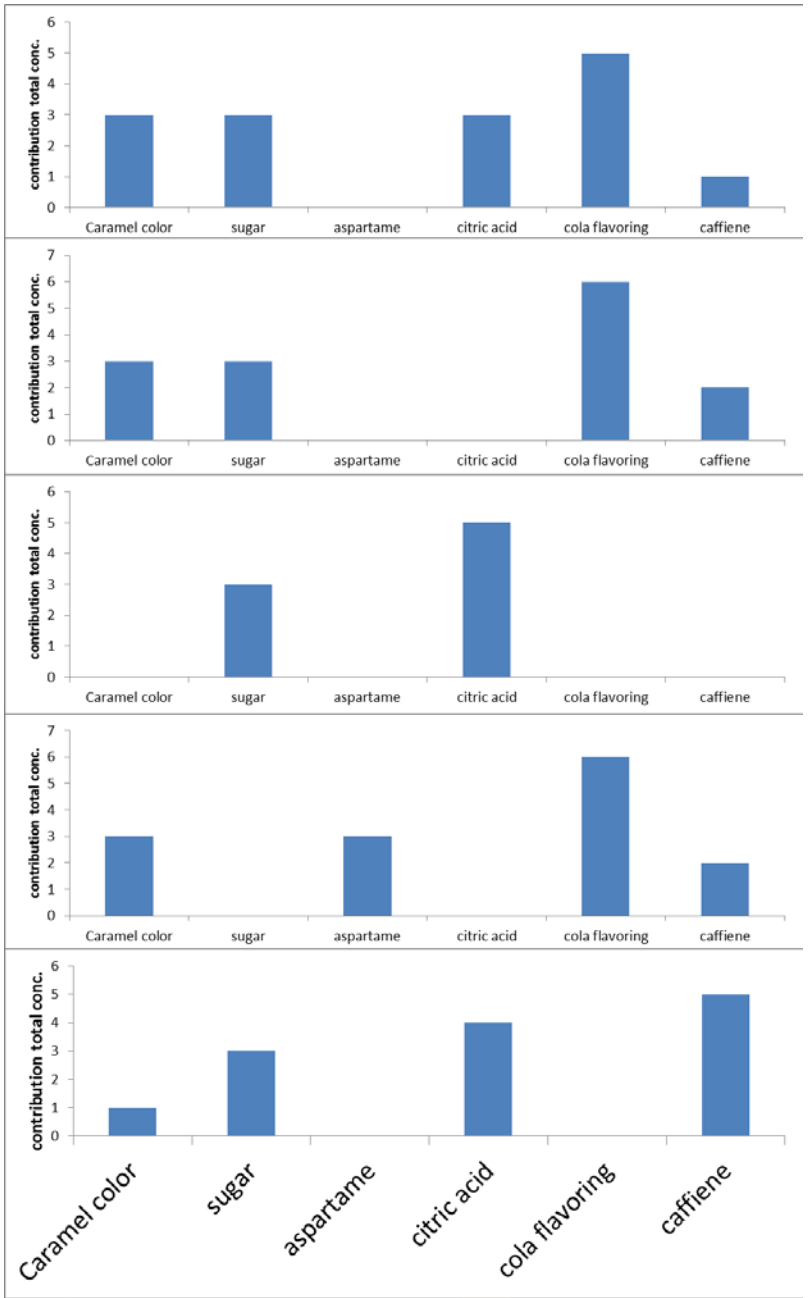
rodenburg@envsci.rutgers.edu

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"

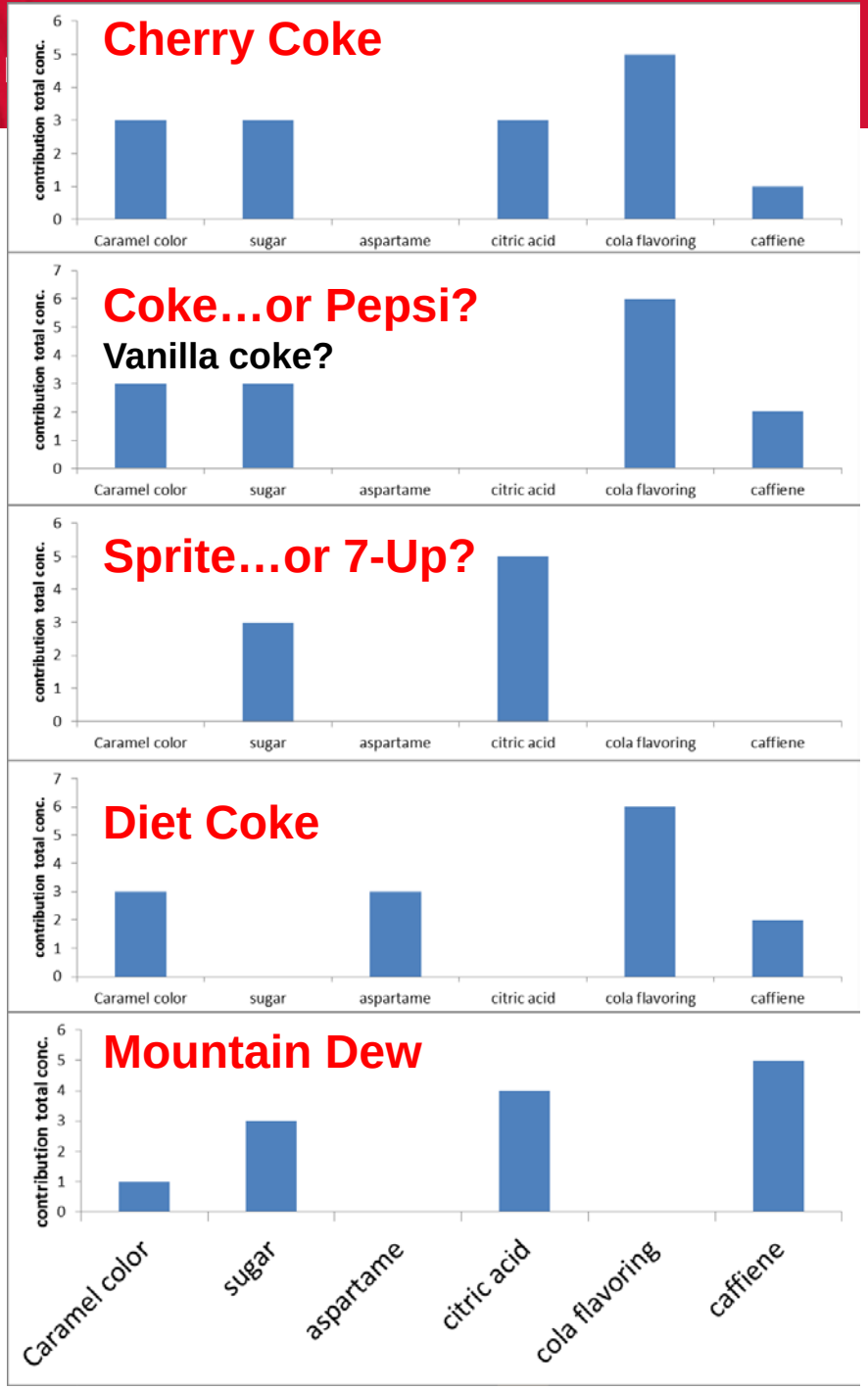


PMF results - F matrix Fingerprints

PMF **can't** tell you:

- What it all means

- **YOU** have to interpret this information



rodenburg@envsci.rutgers.edu

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%→

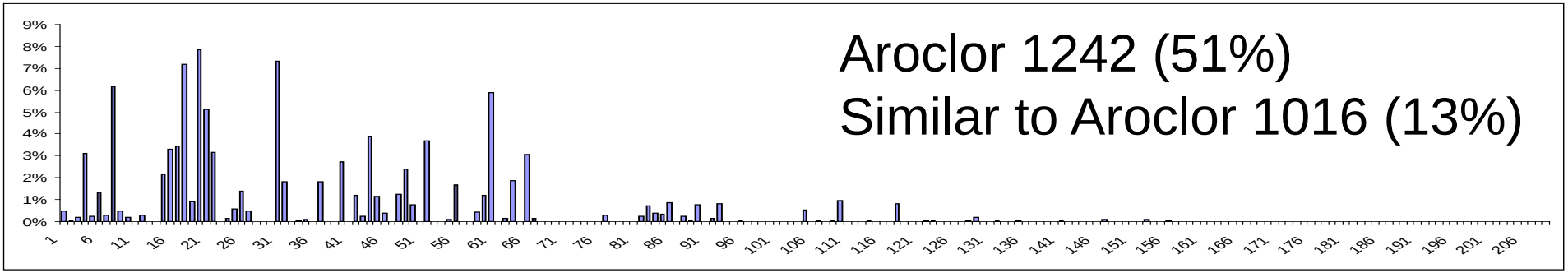
Need ancillary info, such as age, gender, time of day etc.

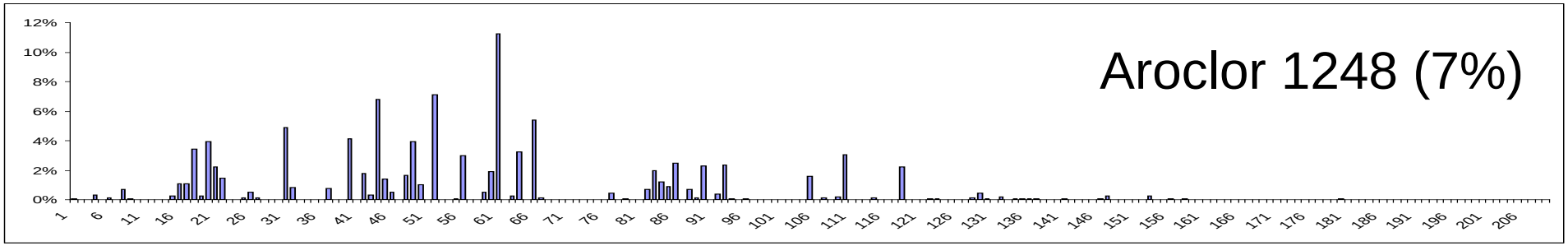
rodenburg@envsci.rutgers.edu

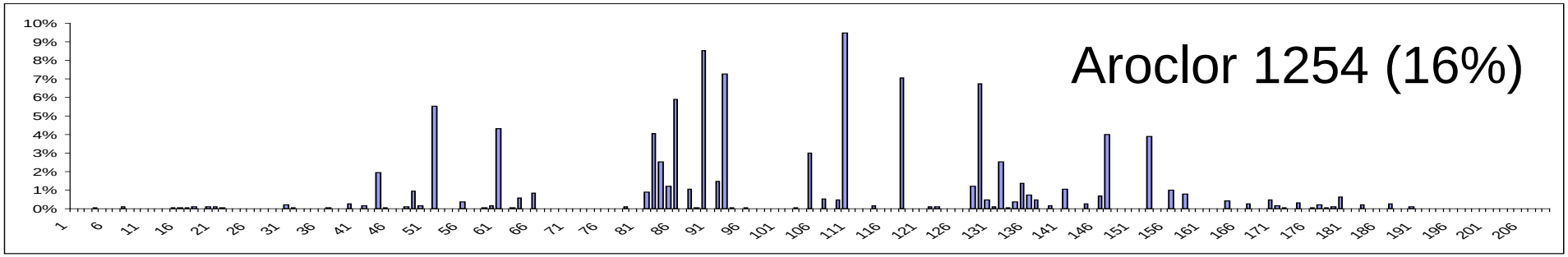
Main PCB sources in most watersheds

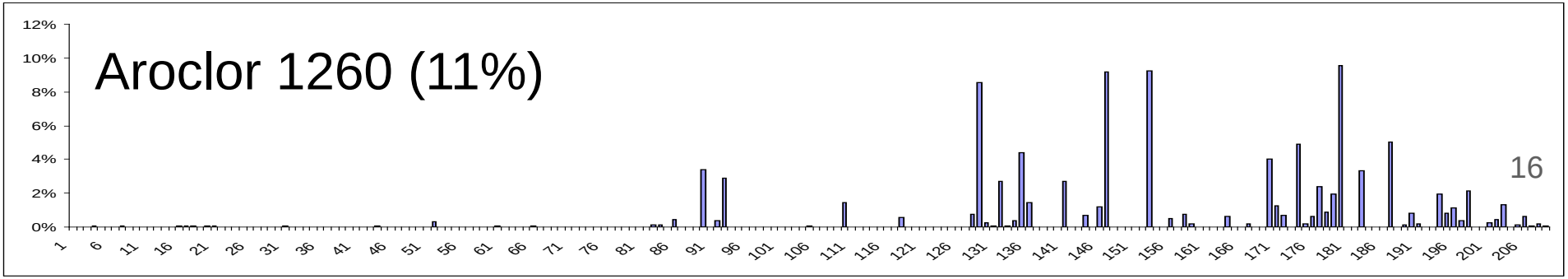
- AROCLORS!
- Non-Aroclor congeners from pigments
- Reductive dechlorination of Aroclors by bacteria

Aroclor fingerprints have been measured









Known inadvertent non-Aroclor PCB sources

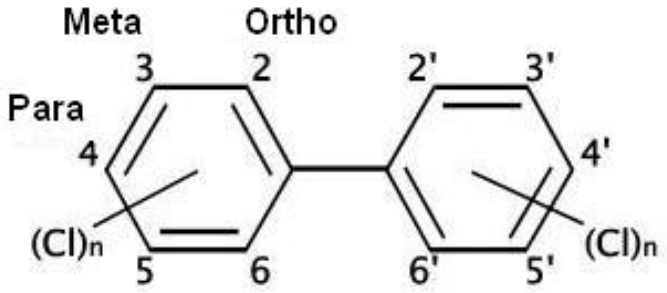
- Organic pigments, especially diarylide yellow, contains primarily PCB 11, among others
  (like 12?, 13?, 35, 77, 52 etc)

- Titanium dioxide (white pigment) may contain PCBs 206, 208, and 209

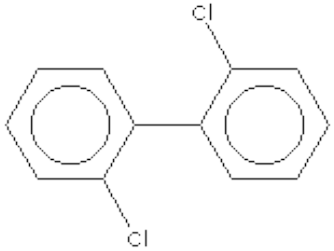
- Silicone rubber tubing produces PCBs 68, 44 and 45, etc.
  (Perdih and Jan Chemosphere 1994)
  - Don't sample using silicone rubber tubing!

Microbial dechlorination of PCBs

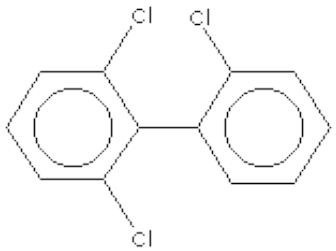
- Previously, seen ***only*** in aquatic sediments, but we found it in:
  - **Sewers** (esp. combined)
  - **Landfills**
  - **Groundwater** at contaminated sites
- Mediated by *chloroflexi*
  - Use organochlorine compounds as electron acceptors
- Usually removes chlorines at *meta* and *para* but not *ortho* positions
  - Several pathways identified
  - Main products are PCB 4 and 19



PCB 4 (2,2')



PCB 19 (2,2',6)



Green-Duwamish Data sets analyzed by PMF

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

- % 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

rodenburg@envsci.rutgers.edu

Uncertainty

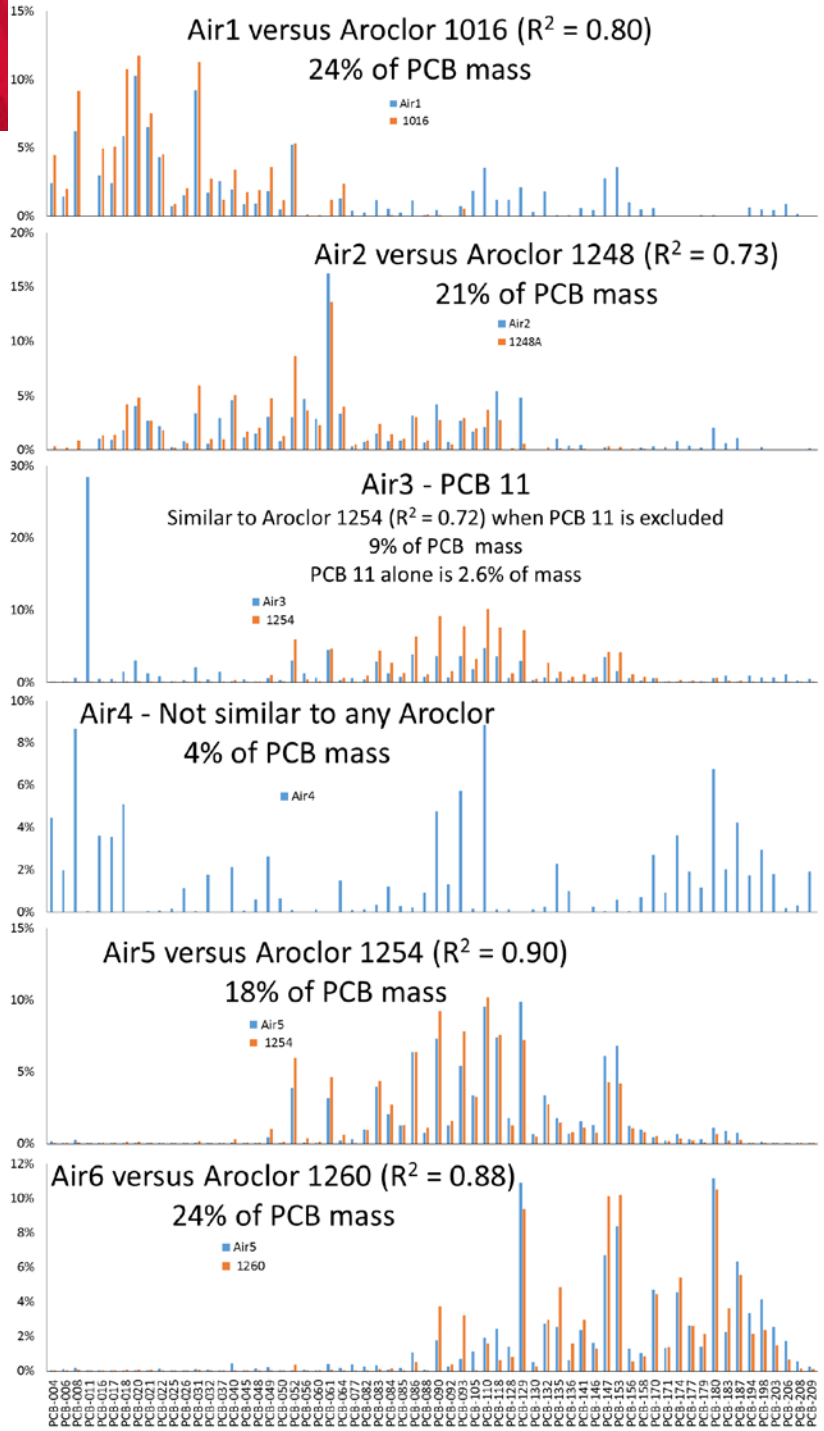
The PMF model and results are highly reproducible. The results might still be uncertain.

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)

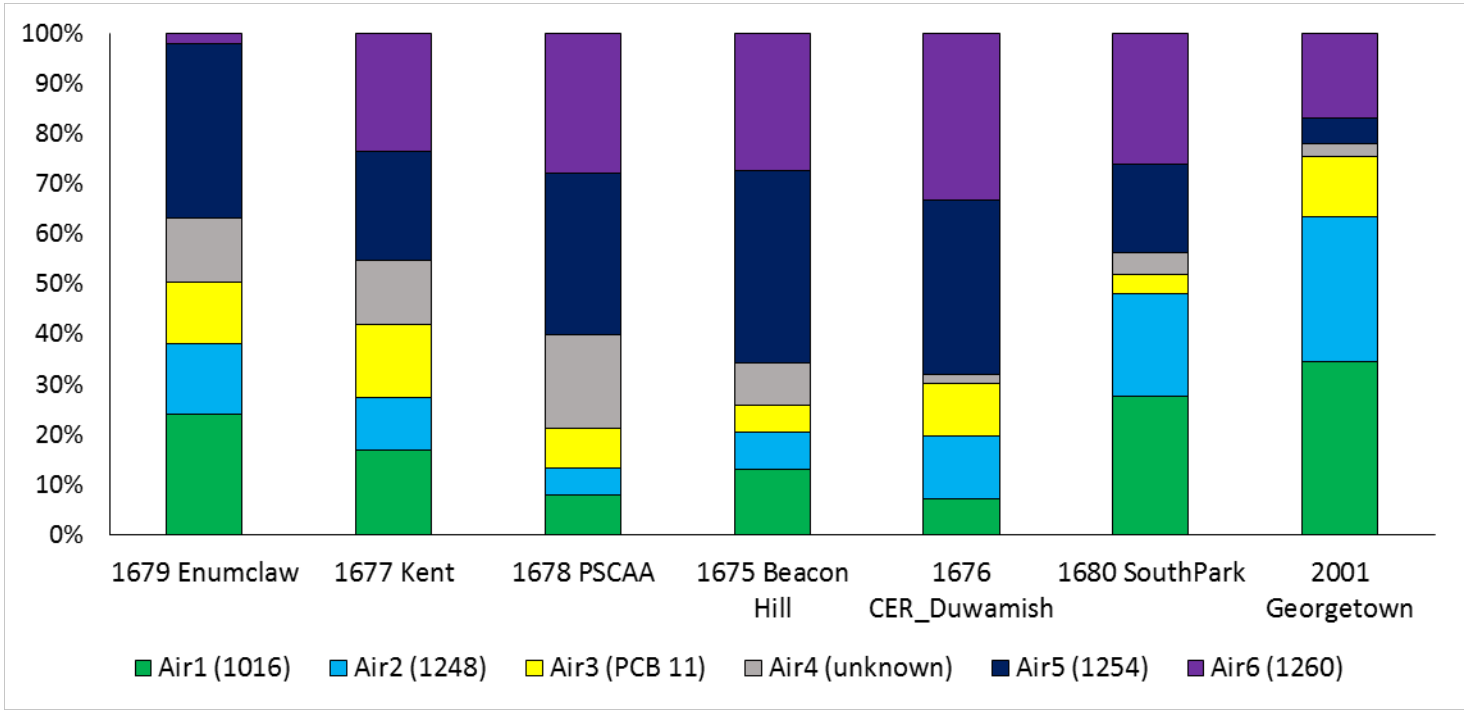
- 6 factors found:
- 4 surprisingly similar to Aroclors
- 1016 > 1260 > 1248 > 1254
- Lower MW formulations more abundant in the atm dep
- Air4 (5% of mass) does not resemble any Aroclor – composition is variable



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?

Sediment

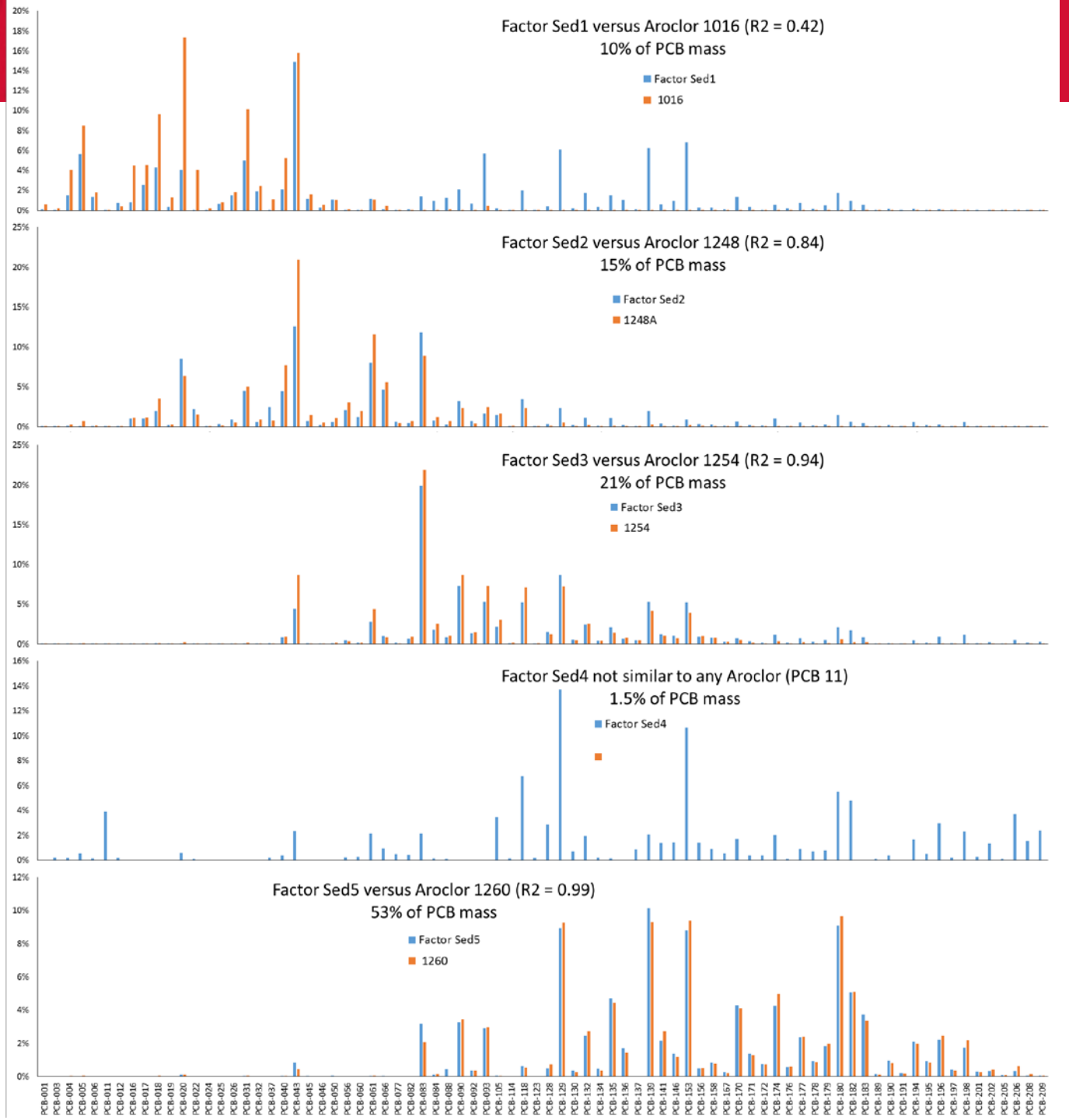
5 factors found:

4 similar to Aroclors

1260>>1254>1248>1016

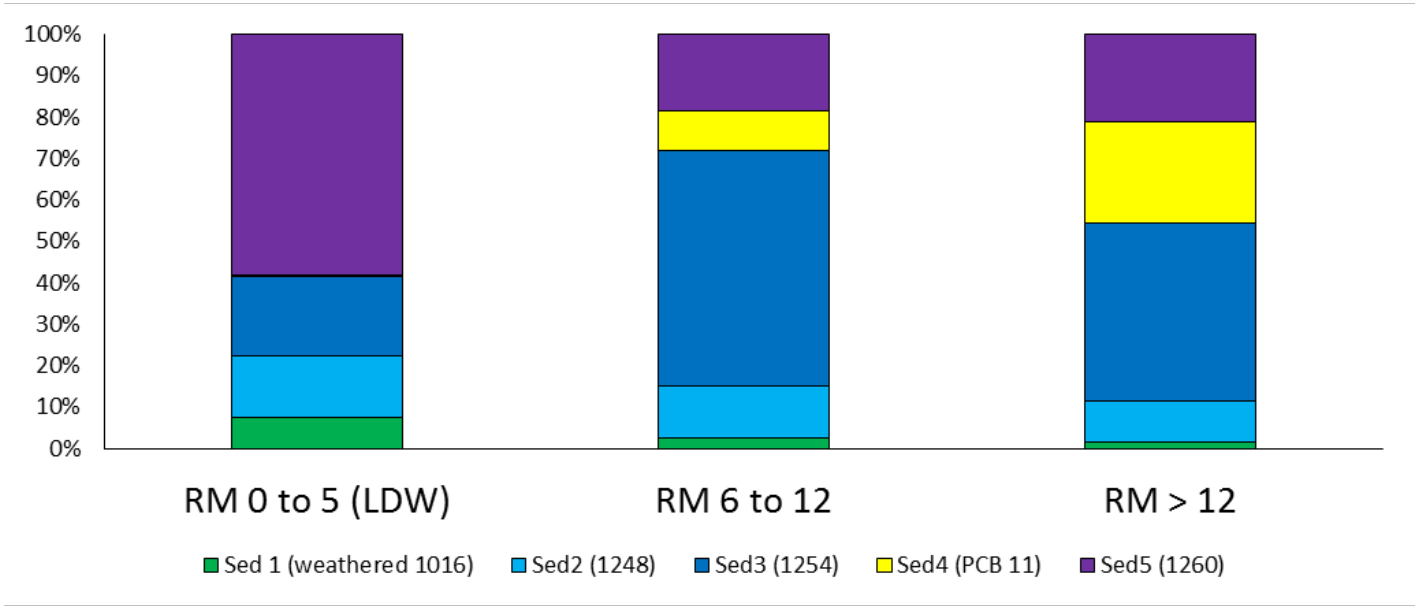
Sed4 not similar to Aroclors, contains a lot of PCB 11

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



rodenburg@envsci.rutgers.edu

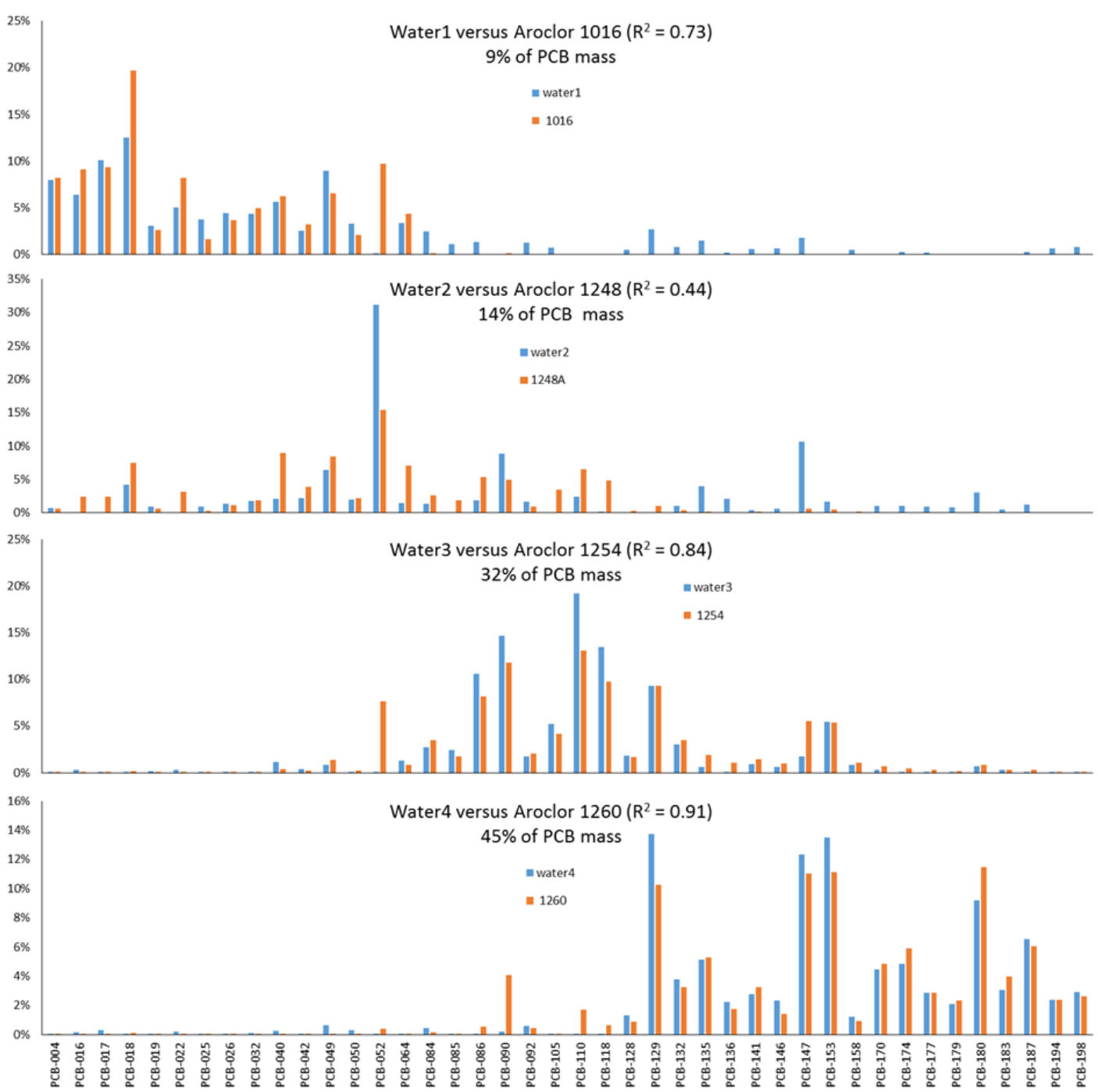
Sediment – spatial trends



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

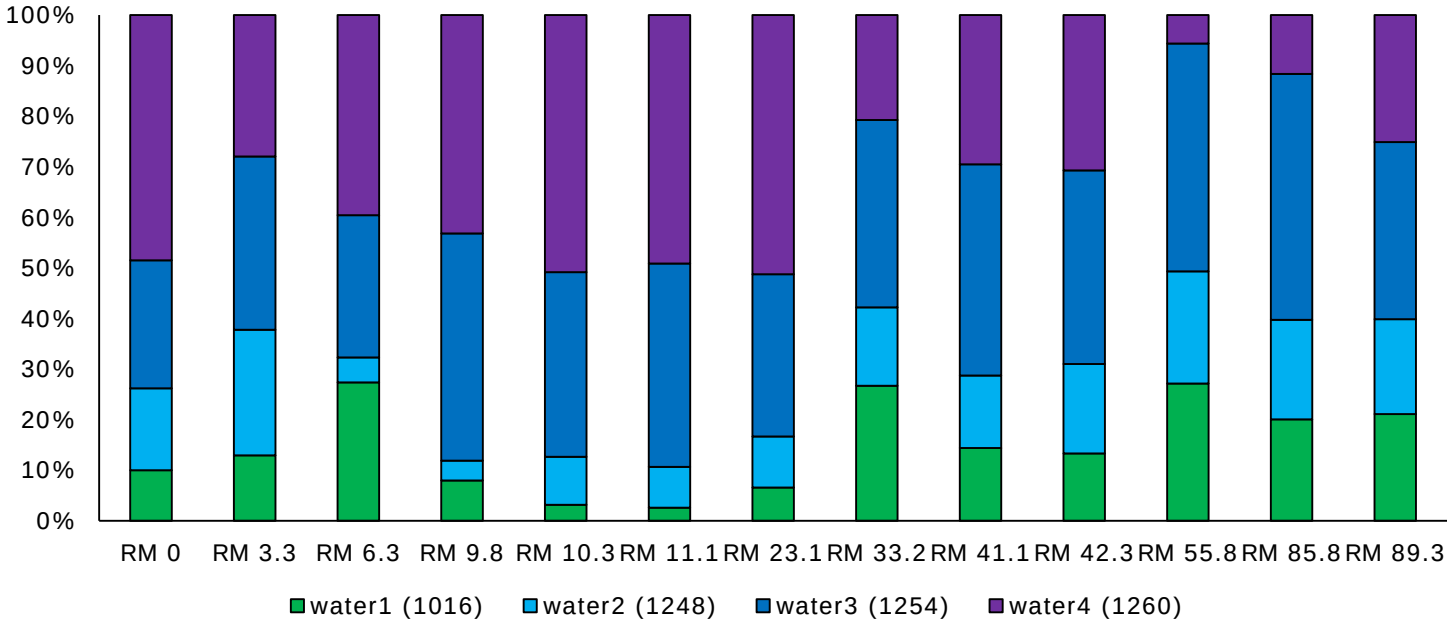
Surface Water

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



rodenburg@envsci.rutgers.edu

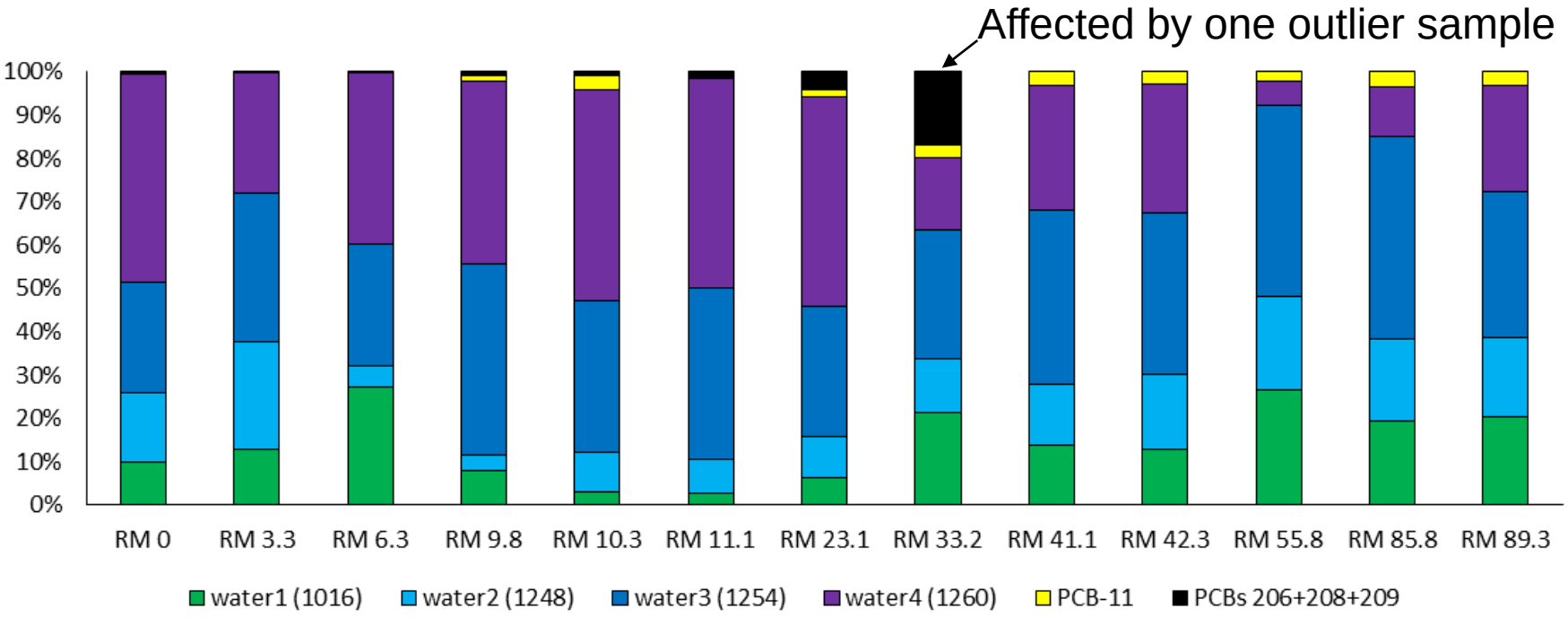
Surface water – spatial trends



- 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

RUTGERS

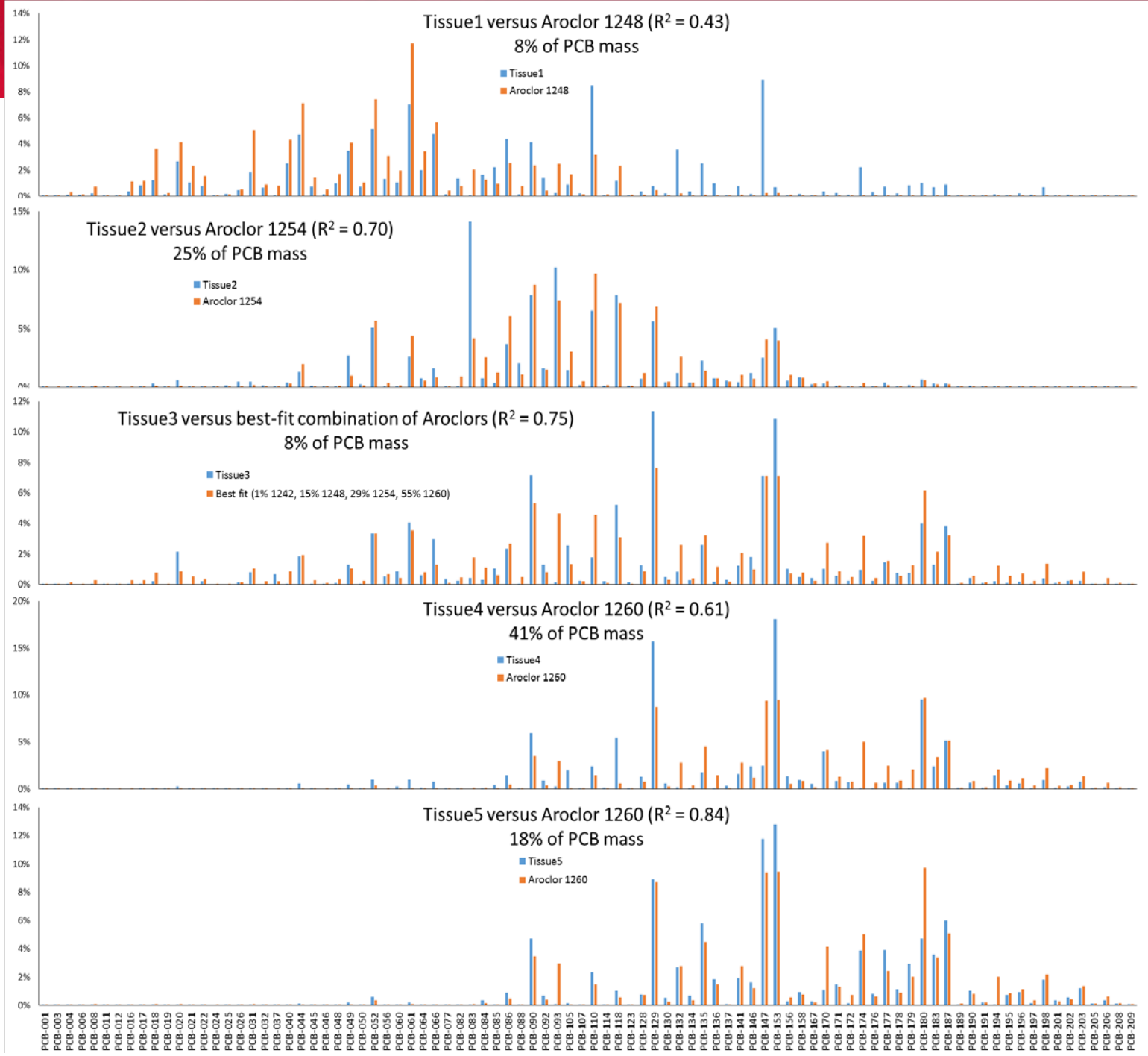
Tissue

Five factors:

All resemble single Aroclor or mix

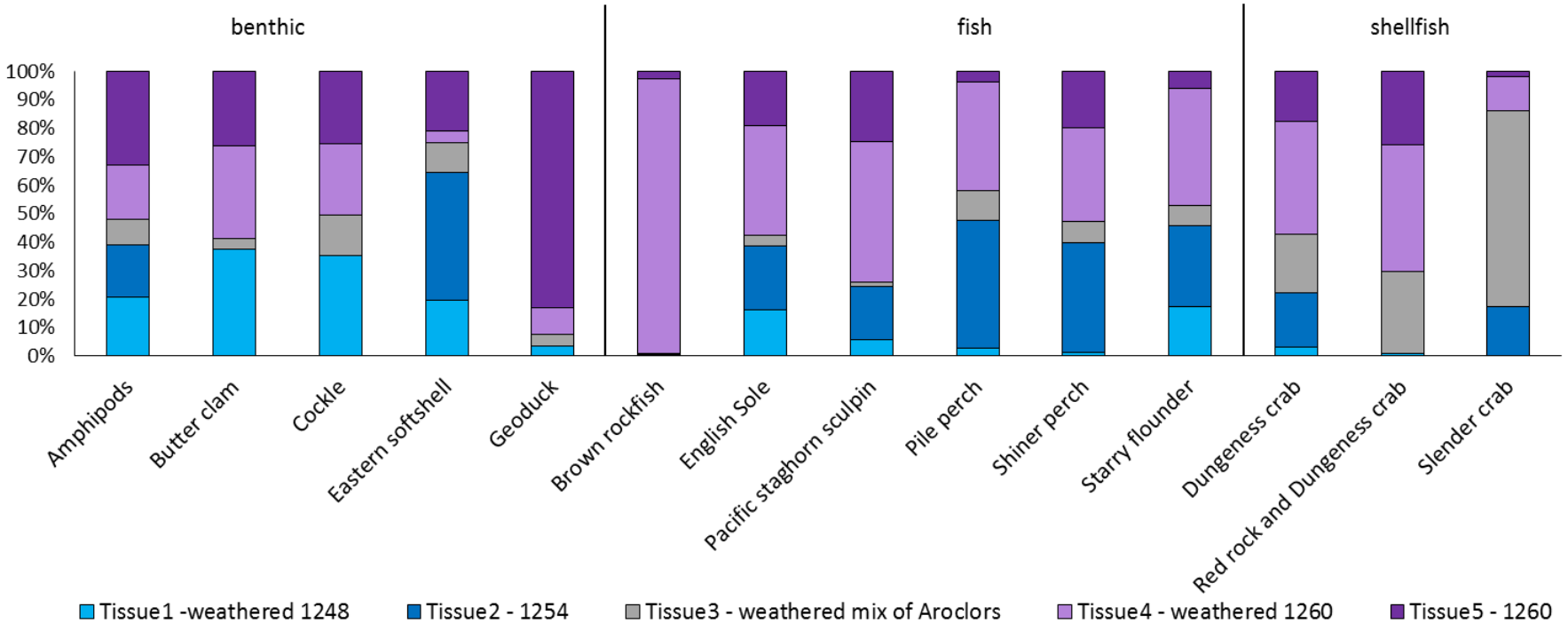
1260>1254>>mix=1248

Tissue4 and Tissue5 both resemble 1260, one more weathered (ADME processes)



rodenburg@envsci.rutgers.edu

Tissue – by species



- Species vary in their ability to metabolize PCBs

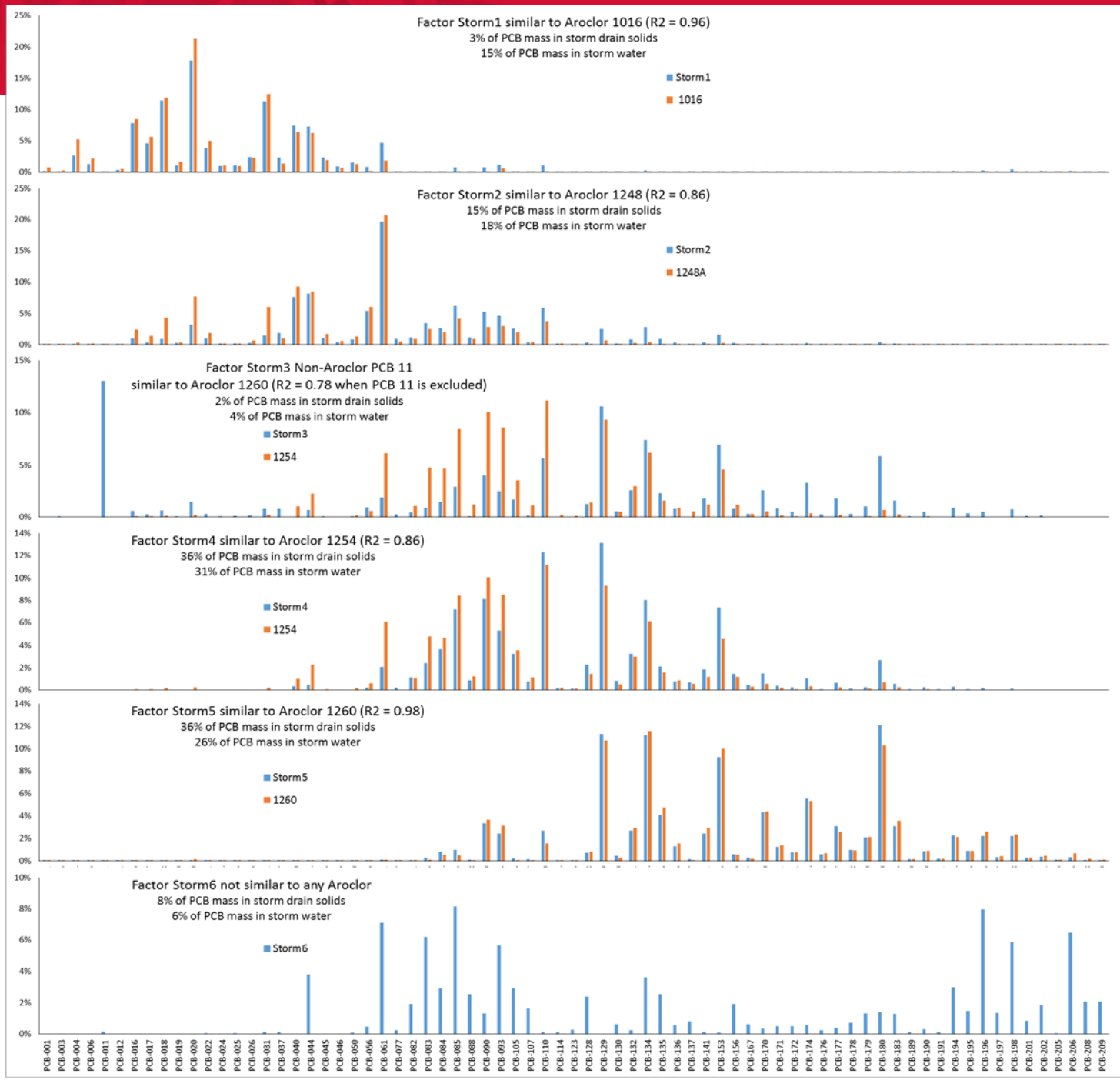
Storm drain solids and water

Six factors:

Four resemble Aroclors

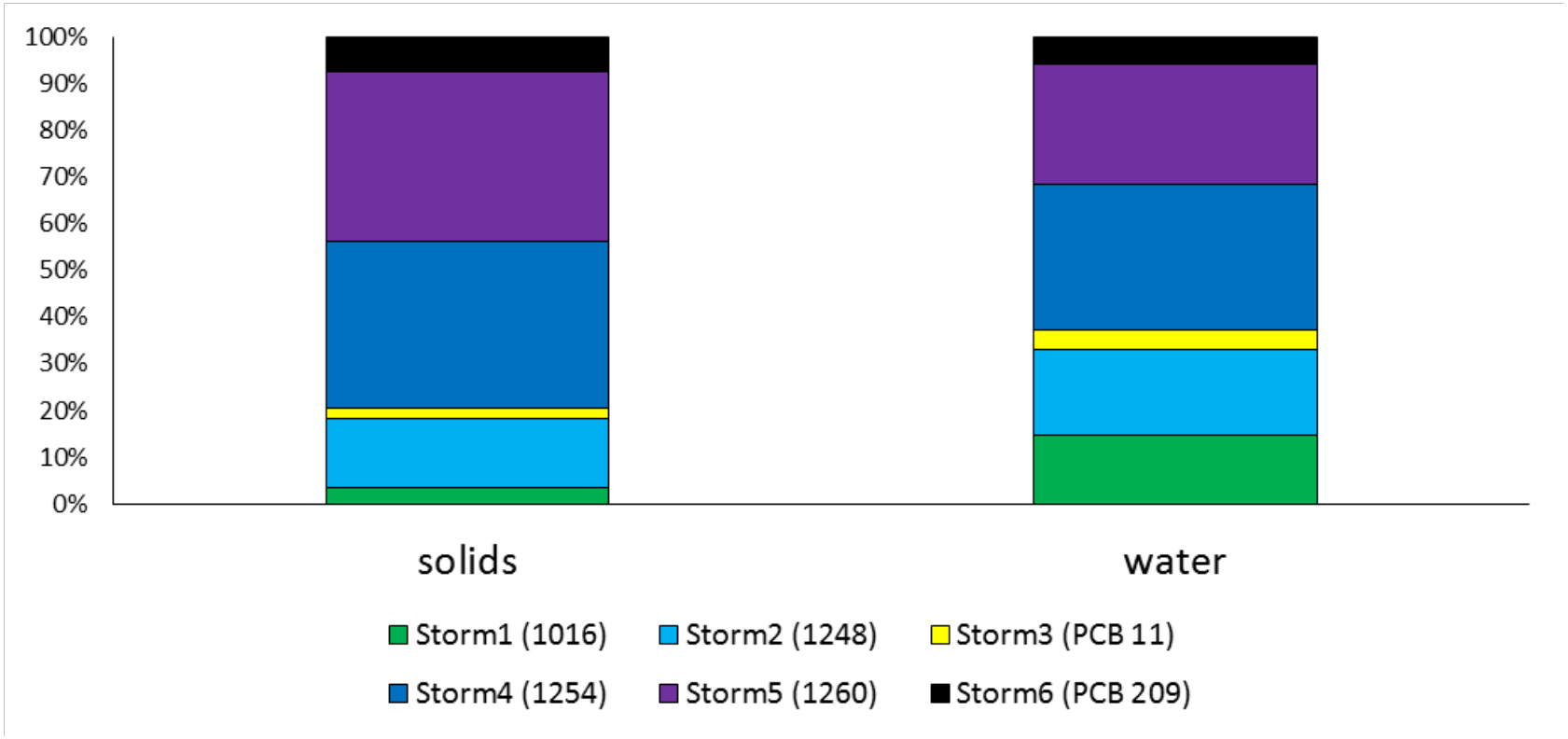
Storm3 = PCB 11

Storm6 = 206+208+209



rodenburg@envsci.rutgers.edu

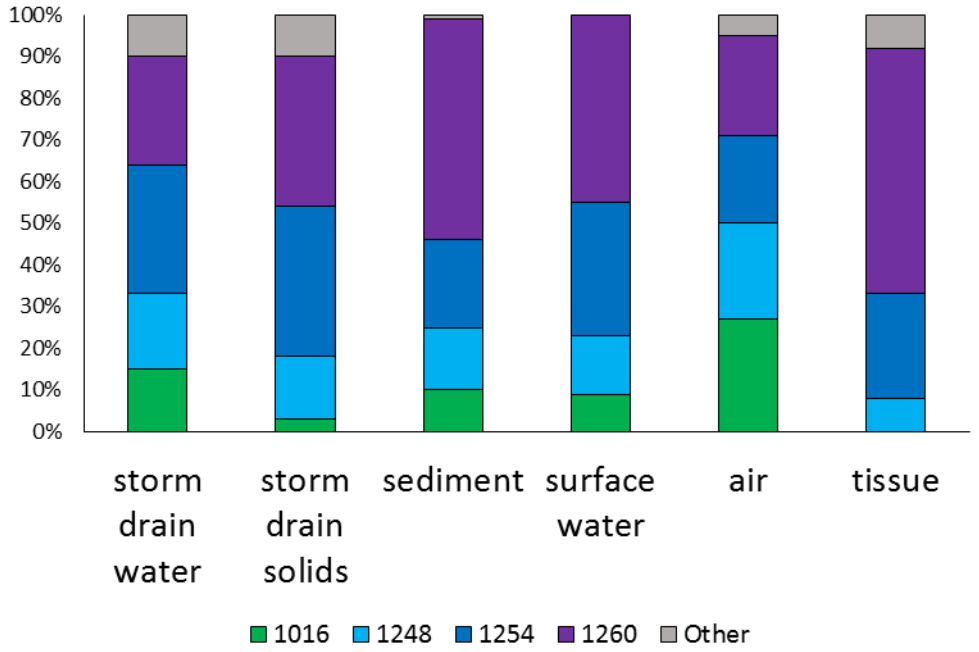
Storm drain water vs. storm drain solids



- Storm water shifted toward lower MW Aroclors

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

rodenburg@envsci.rutgers.edu

Green-Duwamish Results

Types of sources:

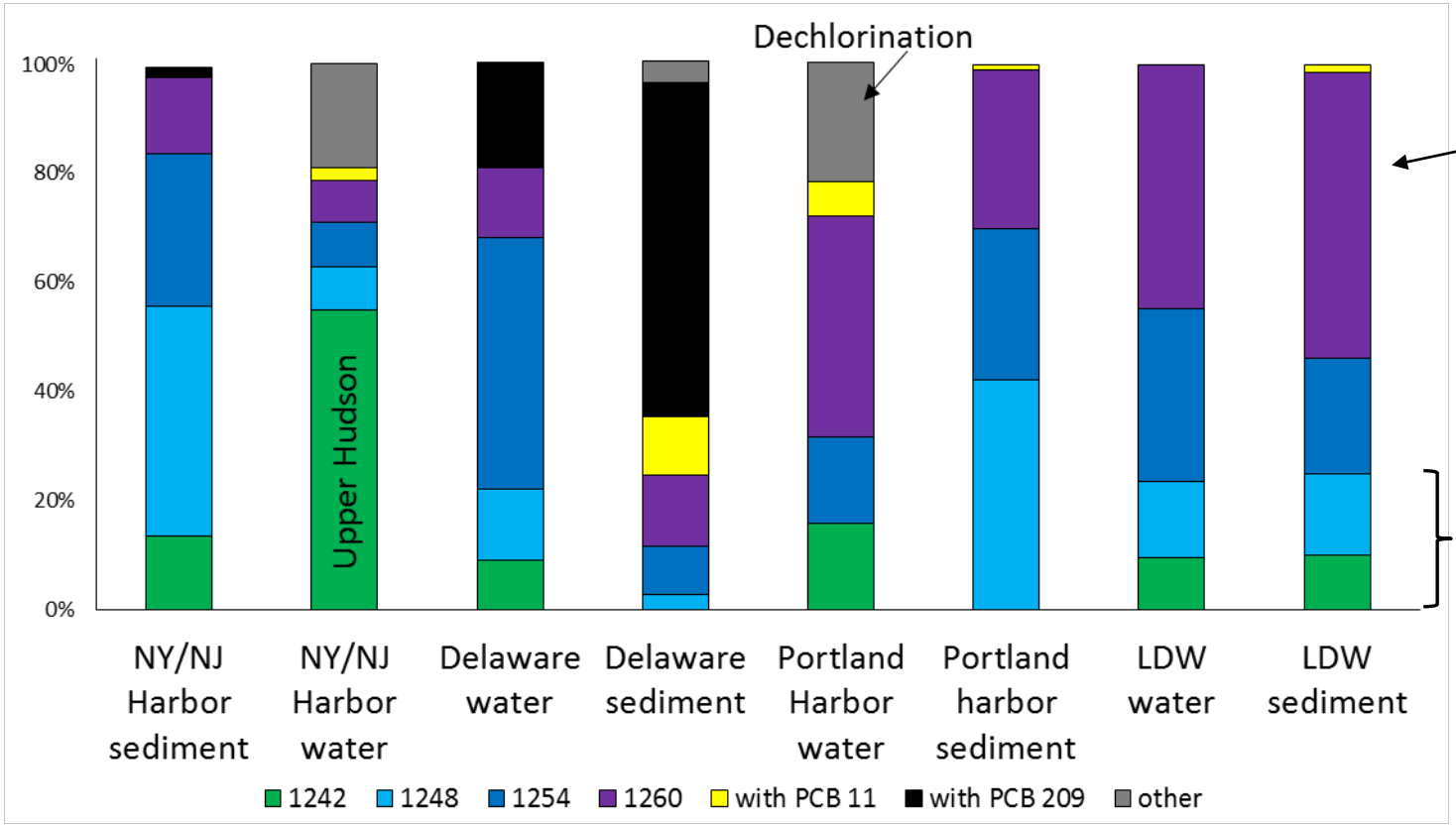
- Across all five compartments, Aroclors are the dominant PCB sources
  - 1260 > 1254 >> 1248 > 1016/1242
- Non-Aroclor PCB sources are minor
- No dechlorination – probably due to salinity

Spatial trends in sources:

- Spatial trends are consistent across water, sediment, biota

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

Implications for modeling

- **What is your endpoint?**

- **Model homologs or total PCBs?**
- When 1668 data is available, many systems have modeled PCB homologs, for example:
  - New York/New Jersey Harbor
  - Delaware River
- When 1668 data is not available, systems model either total PCBs or a subset
  - Upper Hudson River models: total and Tri+ and a few congeners
  - Green Bay model: total and 5 congeners

Recommendations for homolog modeling

- Get input from the fate modeling team before sampling, not after

- Always measure PCBs using 1668 with SPB-octyl column
  - How to incorporate Aroclor data?
- Pick a short model calibration period of about one year
  - Opposite of monitoring
- Characterize loads
  - Head of tide, ocean/sound boundary, point sources (WWTPs, CSOs, other dischargers), non-point sources
- Hundreds of samples of water, sediment, and biota needed for the calibration of the model across full range of flow conditions

If you do homologs…

- Probably don't need to model all 10
- Aroclors 1254 and 1260 are dominant sources in the Green-Duwamish
  - As a result, homologs 3 through 8 dominate
  - Homolog 3 through 8 concentrations are well described by the PMF models
- Could potentially ignore homologs 1, 2, 9, and 10
  - Not very abundant in water, sediment, biota
  - Difficult to model due to non-Aroclor sources
  - 1 & 2 are subject to aerobic degradation which is hard to model

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