NIST
National Institute of Standards and Technology
U.S. Department of Commerce

FORENSICS @ NIST | #NISTForensics

A Framework for Optimizing GC-MS Methods for Analysis of Ignitable Liquid Residues

Edward Sisco[1], Marcela Najarro[1],
Dennis Leber[2], Charles Hagwood[2]
Anuj Srivastava[3]

[1]Surface & Trace Chemical Analysis Group, MML, NIST
[2]Statistical Engineering Division, ITL, NIST
[3]Florida State University

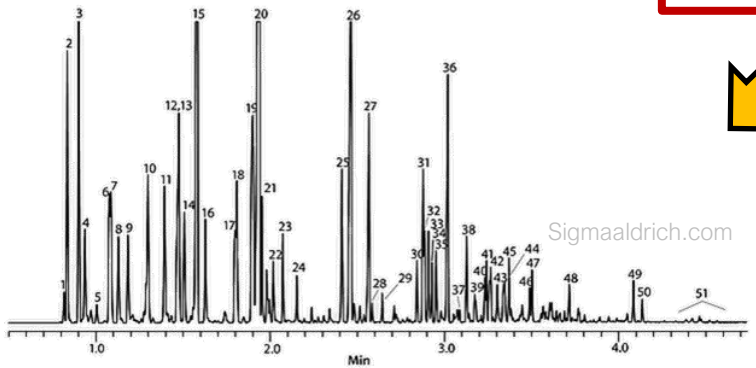
Introduction to the Project

Fire Debris Analysis

Collection

Sampling

Analysis

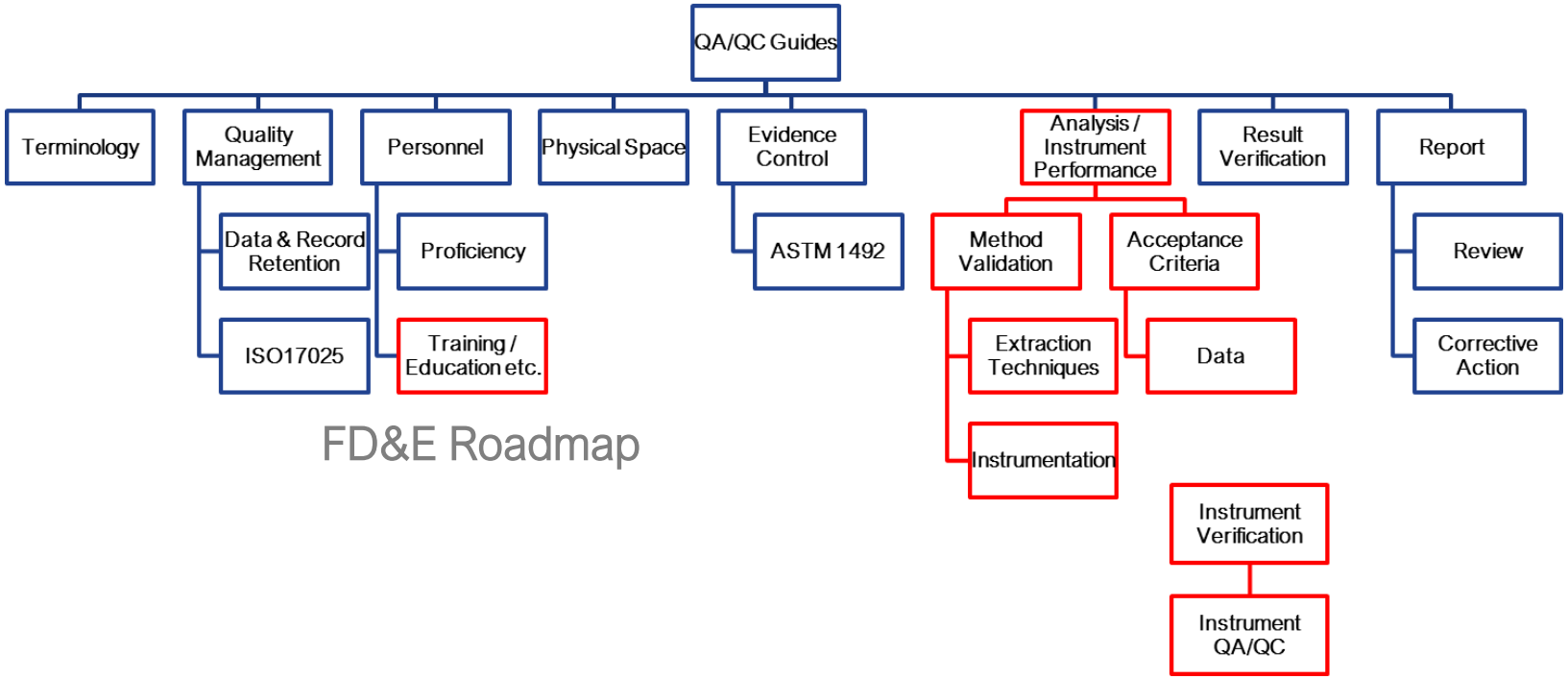


Interpretation

The quality of the analytical result (chromatographic fingerprint) can impact the interpretation of the data.

How We Got Here

- Prior to this project our group had no real experience with fire debris analysis
- Joined OSAC FD&E in 2015 / 2017 for explosives expertise
- The roadmap development suggested a need for stronger instrument performance & validation guidance



FD&E Roadmap

Benefits of Optimization

Quantitative method optimization helps address common questions.

1. How do I know my method is fit for purpose?
    - Labs have developed acceptance criteria
    - Method development varies amongst labs and is typically qualitative
2. How does this GC-MS compare to another?
    - Scientists inherently know that some instruments are more sensitive
    - Process for handling multiple instrument purchases
3. If a method is updated, do I need to re-validate?
    - Unclear whether changing setting “xx” will significant alter my results
4. Is my system still operating optimally?
    - Qualitative verification test mixtures make small changes in performance difficult to detect

Challenges in Optimization

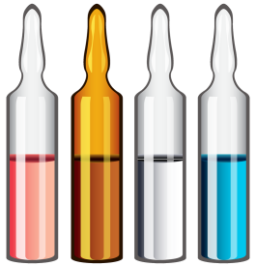
Method optimization is a difficult measurement challenge.

- GC-MS is a multi-component instrument
  - There are many factors to optimize
  - The interplay between factors means that they cannot be analyzed individually
- This process can be time consuming, resource heavy, and expensive
  - Laboratories have time and resource constraints that minimize how in-depth they can go
- There is no definition of what “optimized” means
  - Need to attach data & meaning

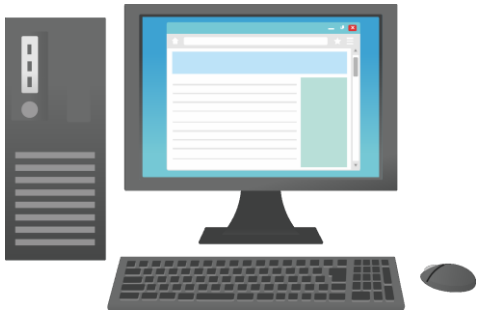
The Project

- Define “optimized” in quantitative terms
- Develop a framework to allow labs to optimize a GC-MS method
  - Investigate all parameters simultaneously instead of individually
  - Minimize impact of lab operations
  - Do not prescribe a method but measure performance of developed methods
    - Feeding into a statistical process control
- Create a performance driven platform without prescribing methods
- Include the community’s input throughout the entire process
  - Adoption of the framework is key
    - OSAC helps to facilitate this

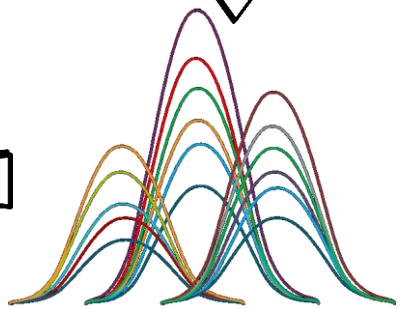
New instrument delivered & installed



Order test mix



Go to web platform & get a list of optimization runs



Complete analyses of test mix

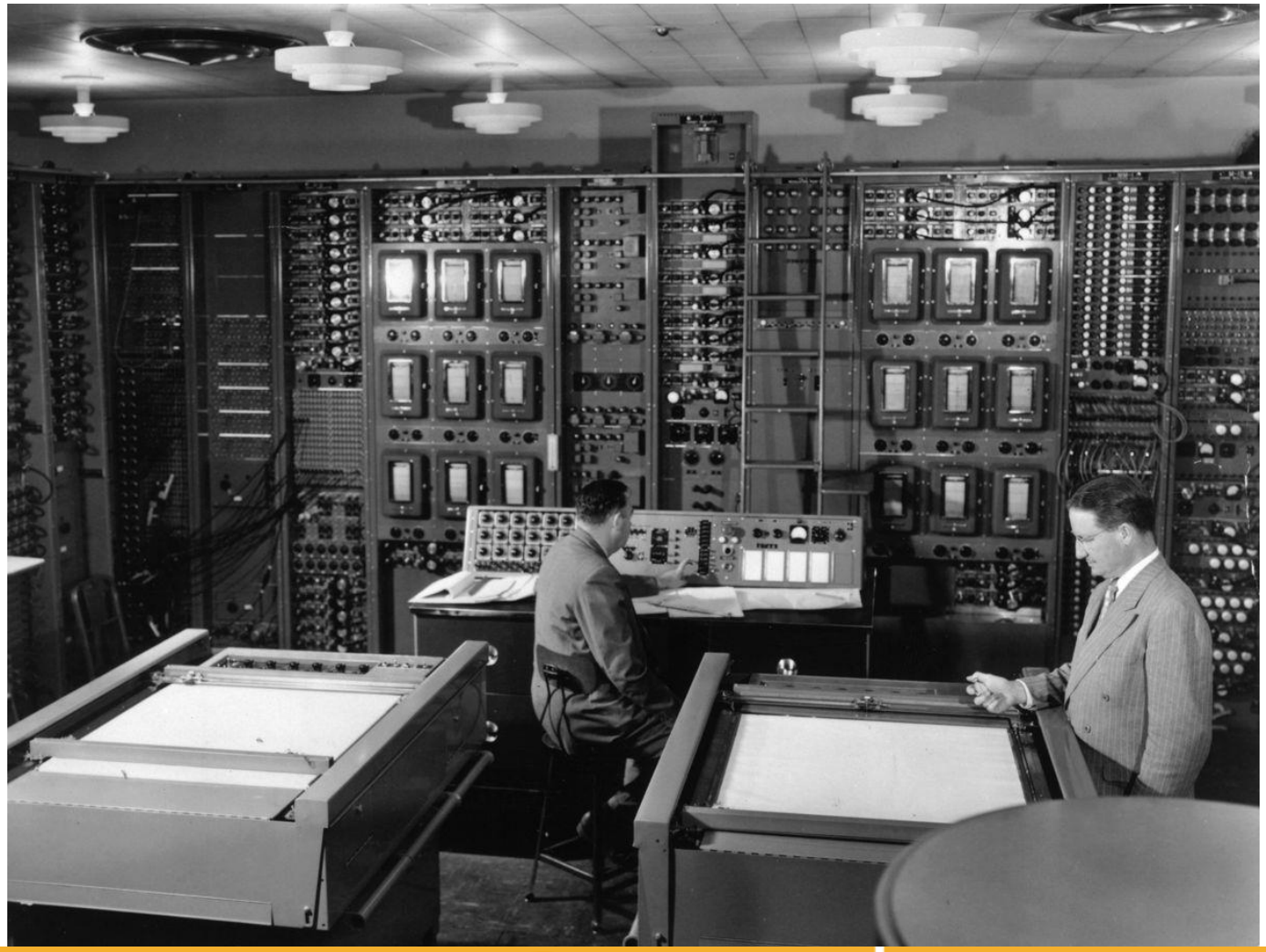
Upload datafiles to get optimal method

Repeat analysis of test mix over time to provide diagnostics & QA/QC

- *A lot of foundational work is required to get to this point*
- *The work we are presenting today if an effort to complete this foundational work*

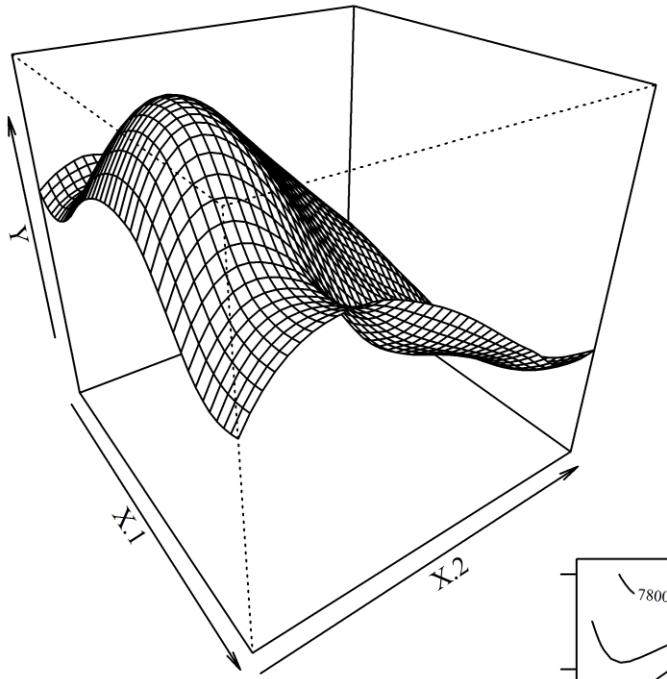
Translation to Statistical Problem

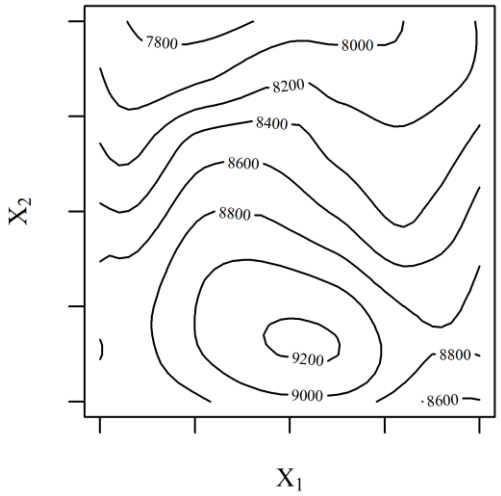
Best Settings



Response Surface Methodology

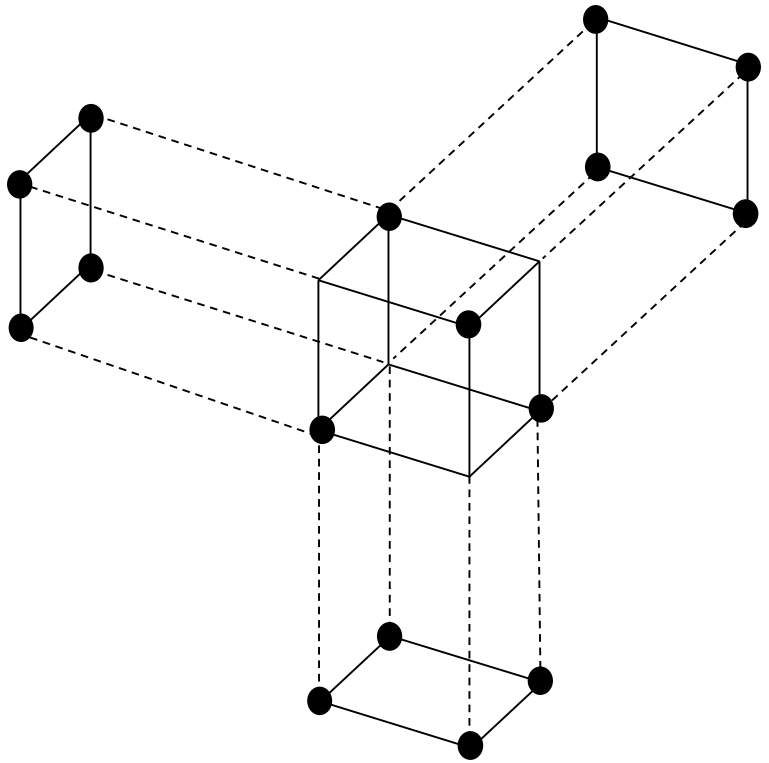
- Optimize the response
- Sequential experimentation
  - Step 1: identify important factors
  - Step 2: approximate surface
  - Step 3: locate global extreme





Screening Experiment

- Factor screening to identify most important factors
- Orthogonal first-order experimental designs
- Two-level designs
  - 2^k factorial
  - 2^{k-p} fractional factorial



Pilot Study – Experimental Settings

Identifying Factors

- The first step for this project was identifying the relevant factors and determining the appropriate settings (low vs. high)
- Developed list of potential factors in-house
- During in-person OSAC meetings, discussed which factors should or should not be studied
  - Wanted to ensure we investigated factors that were commonly considered and not commonly considered
- Cast a wide net in this phase that can hopefully be reduced in the second phase
  - Not be afraid to go outside of instrument manufacturers recommended settings
    - May not be most appropriate for fire debris analysis

Initial Factor List

- Settled on a list of 19 factors
  - Split between injector, GC, and MS
  - Some factors are easier to change than others

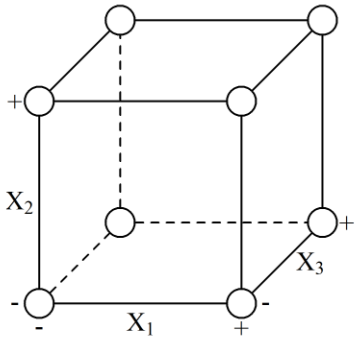
Injector			GC			MS		
Factor	Low	High	Factor	Low	High	Factor	Low	High
Injector Type	SSL	PTV	Column Length	15 m	30 m	Scan Rate	1 scan/s	4 scan/s
Injector Temp	200 °C	300 °C	Column Type	DB-1	DB-5	Source Temp	200 °C	275 °C
Split Ratio	1:1	25:1	Flow Rate	0.5 mL/min	3.0 mL/min			
Volume	0.5 µL	2.0 µL	Initial Temp	35 °C	50 °C			
Liner Material	Deact.	"SKY"	Initial Hold	1.5 min	3.0 min			
Liner Shape	Open	Tapered	Ramp Rate	10 °C/min	30 °C/min			
Liner Wool	Yes	No	Final Temp	275 °C	300 °C			
Septa Type	Green	Red	Transfer Temp	200 °C	300 °C			
Syringe Type	Fixed	Remov.						

Pilot Study – Experimental Design

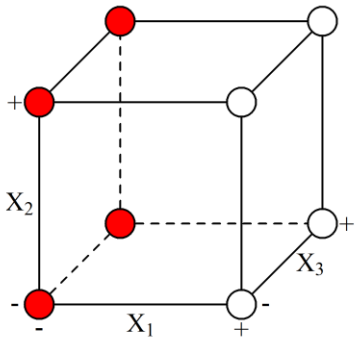
Two Level Screening Experiment

- Identify factors with largest impact on forensic fire debris measurement and analysis
- Conclusions limited to NIST Thermo system
- Two-level designs are extremely efficient and effective at identifying factors that have the largest impact
- Two-level full factorial design
  - 2^{19} = 524,288 experimental conditions
  - Three replicates = 1,572,864 observations

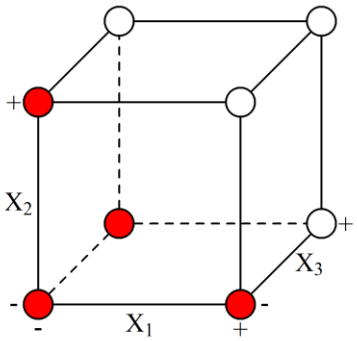
Two-level Fractional Factorial Designs



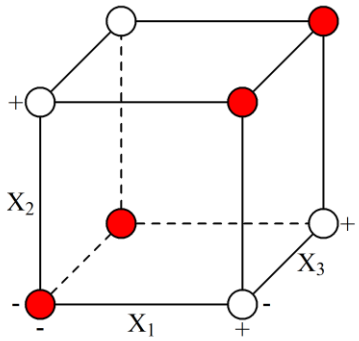
Coverage



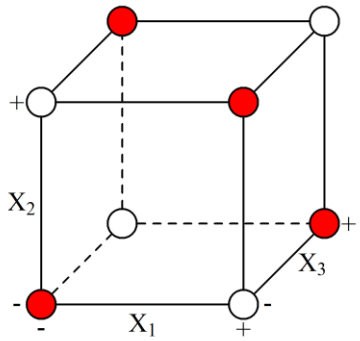
Balance



Confounding



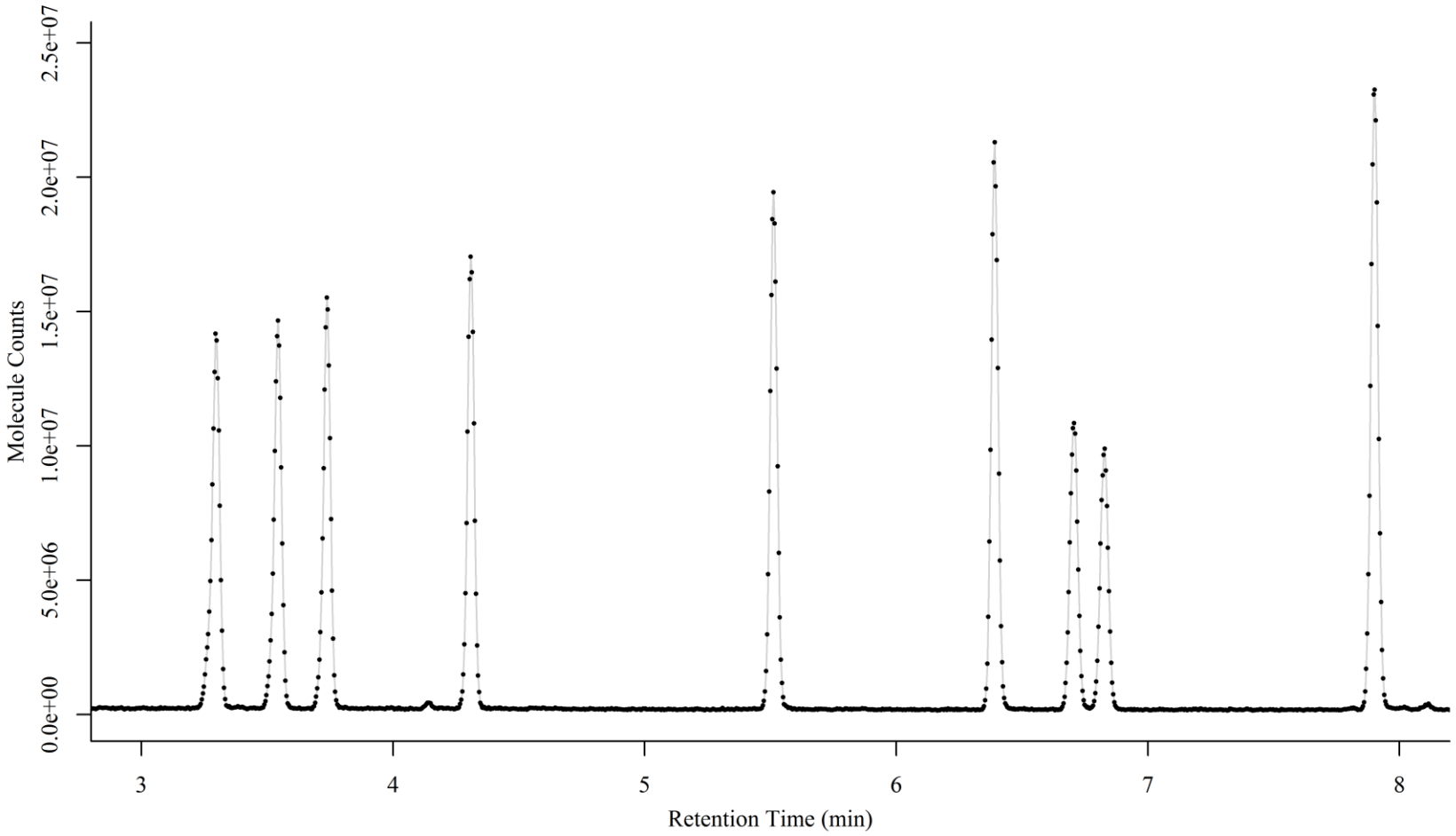
Orthogonal



Pilot Study Design

- Two-level orthogonal fractional factorial design
- Split-plot design
  - 8 hard-to-change factors
  - 11 easy-to-change factors
  - Split plot confounding (vs. Cartesian product)
- Fraction factorial split-plot design
  - 64 experimental conditions; 192 observations
  - Resolution IV design
- Logistics
  - 16 days
  - Morning, afternoon, and evening blocks

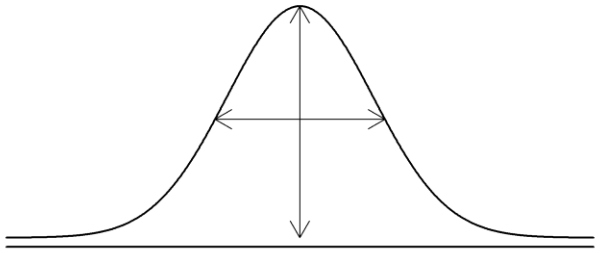
Observations: Chromatographs



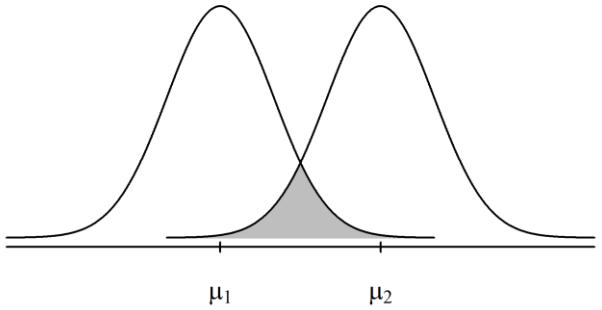
Response Variables

- Model each peak with Gaussian kernel

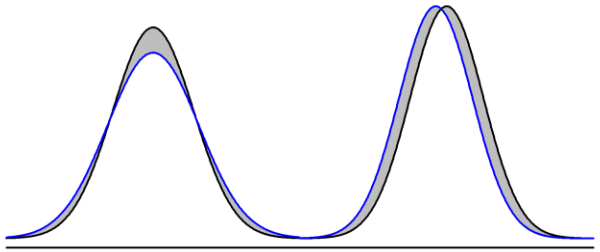
Signal-to-noise



Resolution



Repeatability

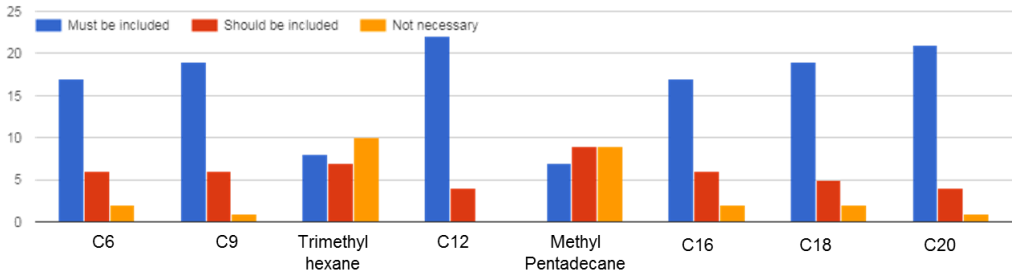


Pilot Study – Execution

Creating a Test Mixture

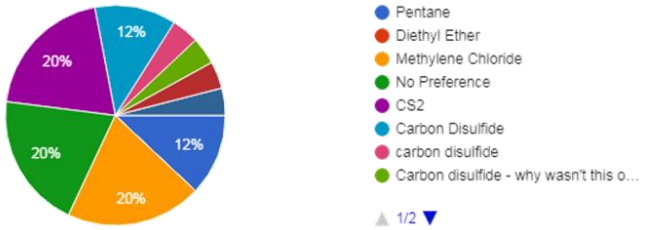
- In order to complete study we needed to develop a test mixture
- FD&E subcommittee suggested E1618 test mixture was a good starting point but was not extensive enough
- Polled FD&E, and others on the components of the test mixture

Proposed Carbon Chain:

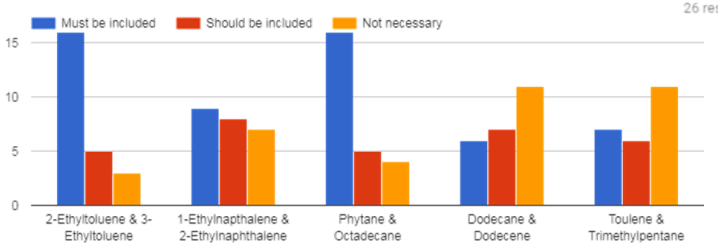


What solvent is most appropriate?

25 responses

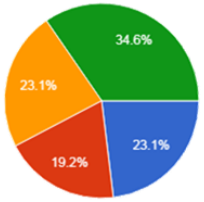


Proposed Resolution Challenge Pairs:



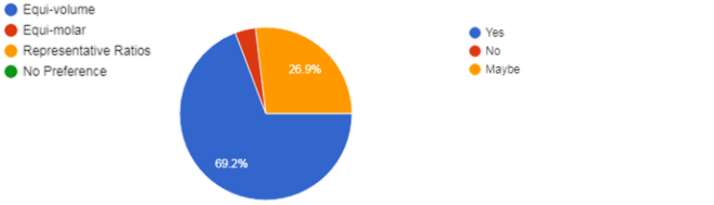
What should the composition of the test mix be?

26 responses



Is there interest in your lab for a test mixture such as this to be available for method validation, performance checks, competence, method validation, etc?

26 responses



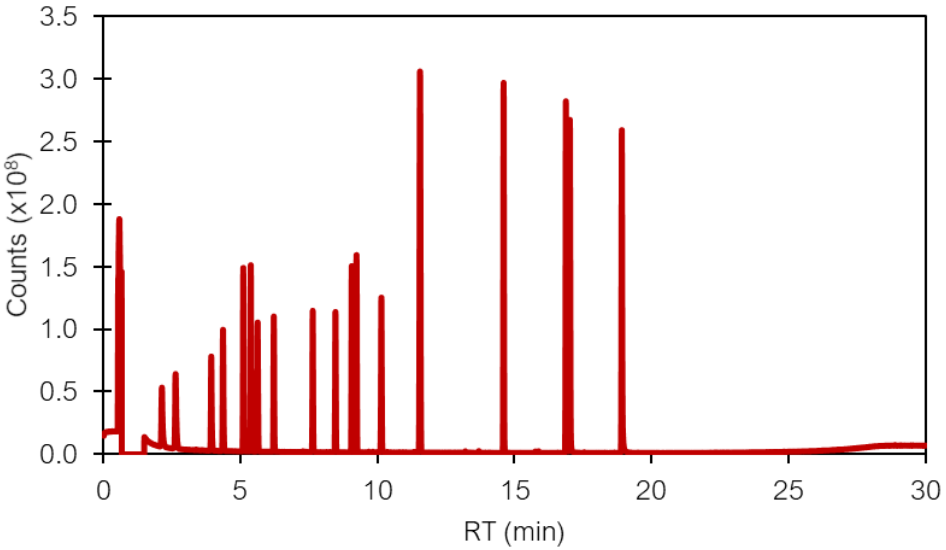
Creating a Test Mixture

- Used results of survey to determine components of text mixture
- Decided to create in-house
  - Unable to ampoule carbon disulfide and dichloromethane
    - Storing in capped vials was insufficient
  - Decided we needed to make mixture in isopropanol

n-Hexane	1,2,4-Trimethylbenzene	2-Ethylnapthalene
Toluene	Indane	1-Ethylnaphthalene
o-Xylene	1,2,4,5-Tetramethylbenzene	n-Hexadecane
2,2,4-Trimethylhexane	n-Dodecane	n-Octadecane
n-Nonane	n-Dodecene	Phytane
m-Ethyltoluene	Naphthalene	n-Eicosane
o-Ethyltoluene	2-Methylnaphthalene	

The Pilot Study

- Underestimated the amount of work involved in the study
  - 192 runs took over 6 months to complete
  - Required us to understand the nuances of software
- All data analysis completed manually
  - Due to changing settings, automated data analysis on software platform was not practical
- Disruptive to other ongoing work
  - Instrument was constantly in a different state
- Process of maximizing randomization is taxing on instrument

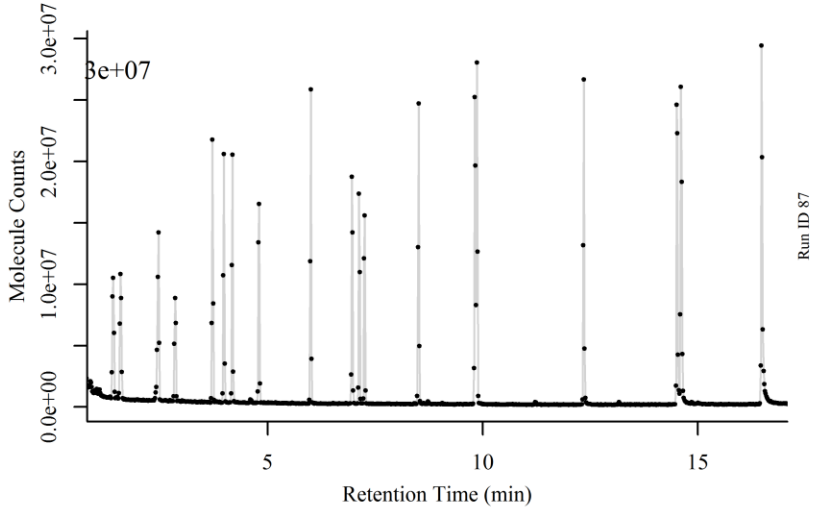


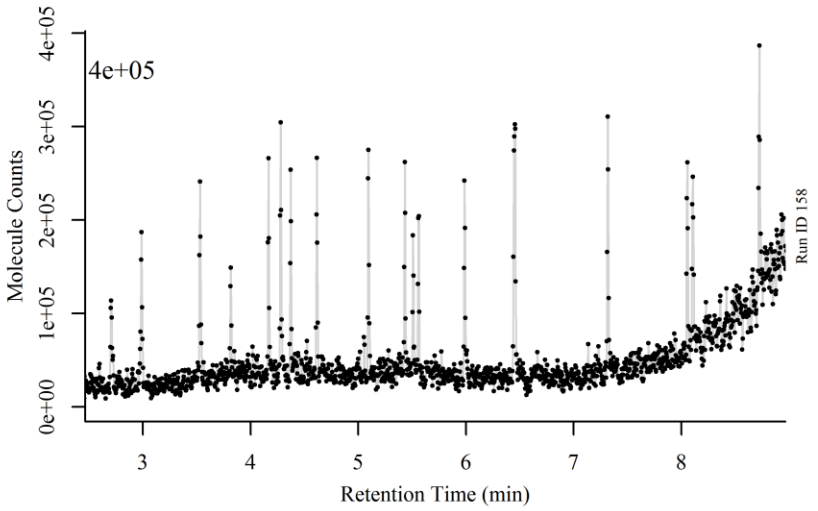
Lessons Learned

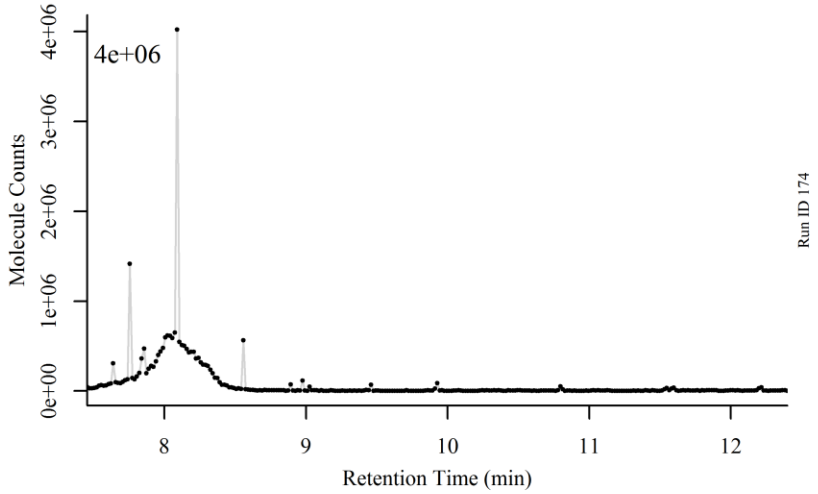
1) We don't have the infrastructure to create the test mixture
   - SRM Division was willing to help but needed a long lead time
2) We need a way to automate data analysis component
3) We need to really consider the balance between randomization, instrument health, and total time of study
4) We thought we knew what "optimal" meant, but it is more nuanced than originally thought
   - We need to bring in the experts

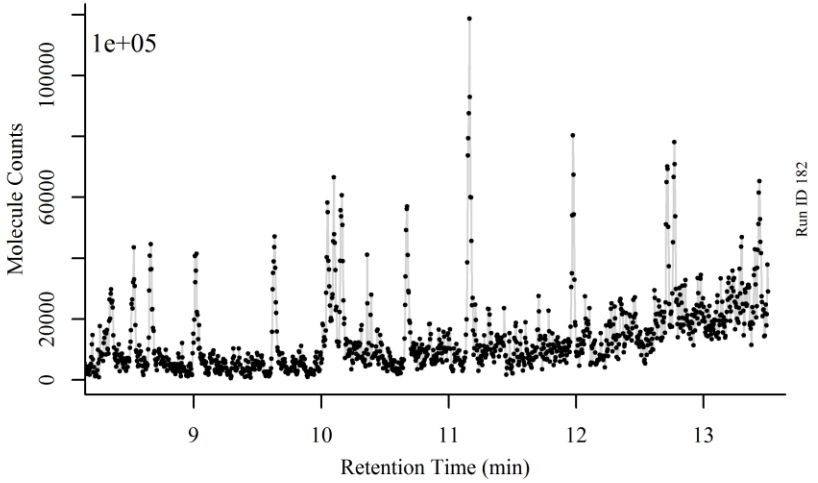
Pilot Study – Analysis

Settings Matter

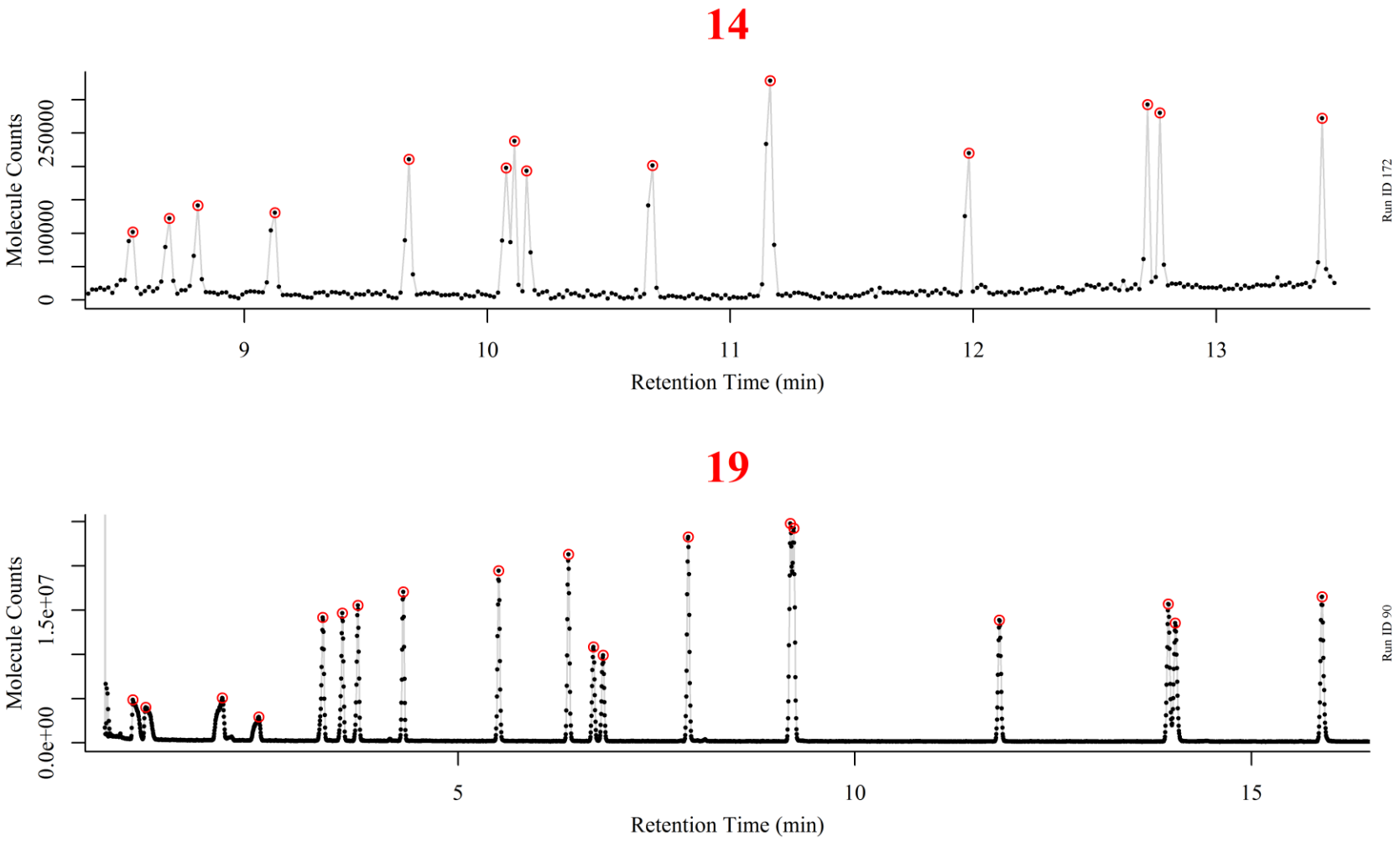




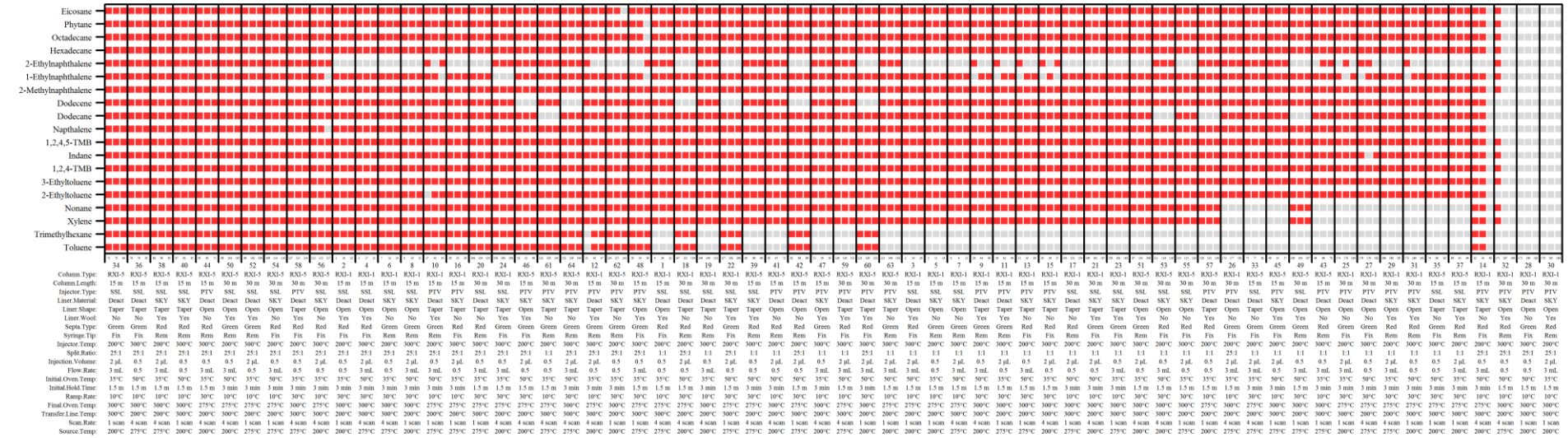




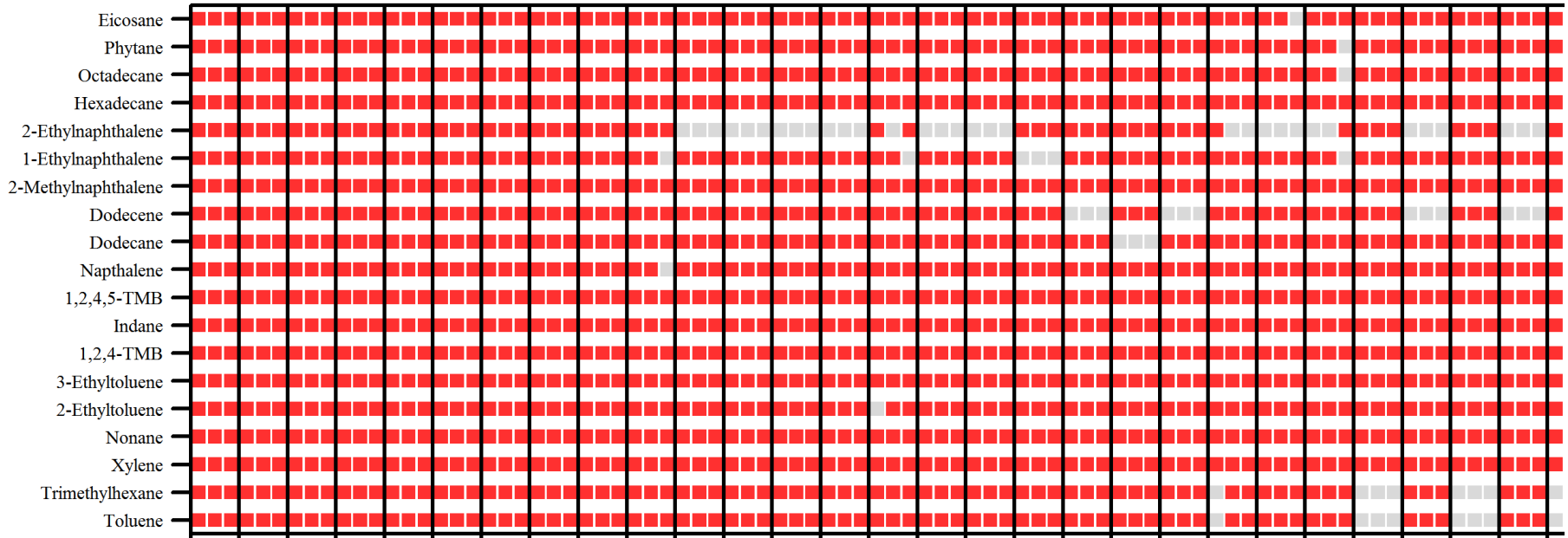
Number of Peaks



Number of Peaks

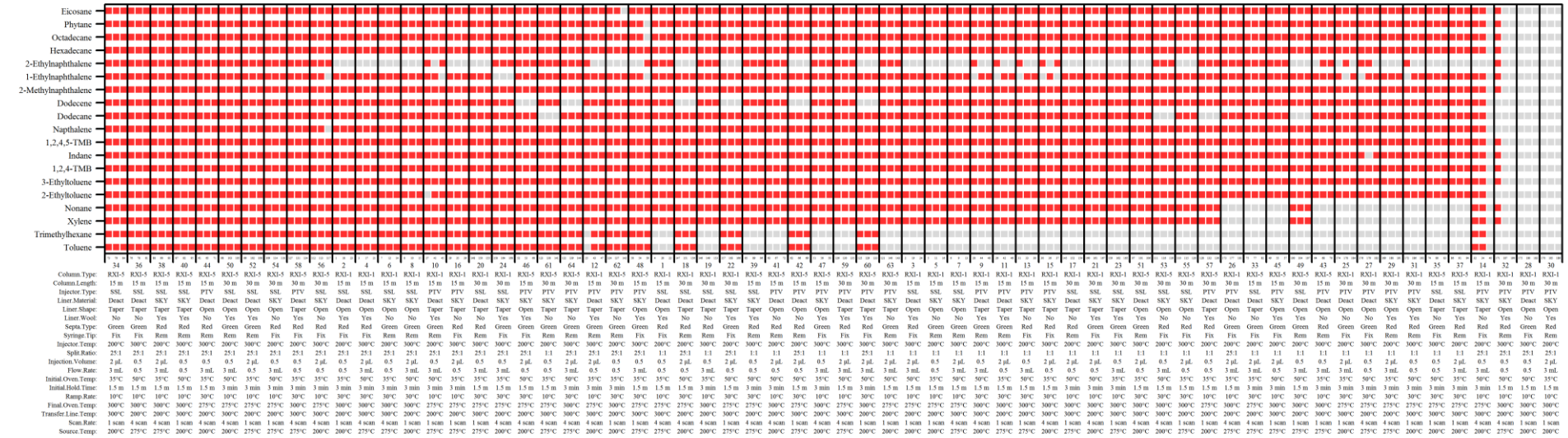


Number of Peaks

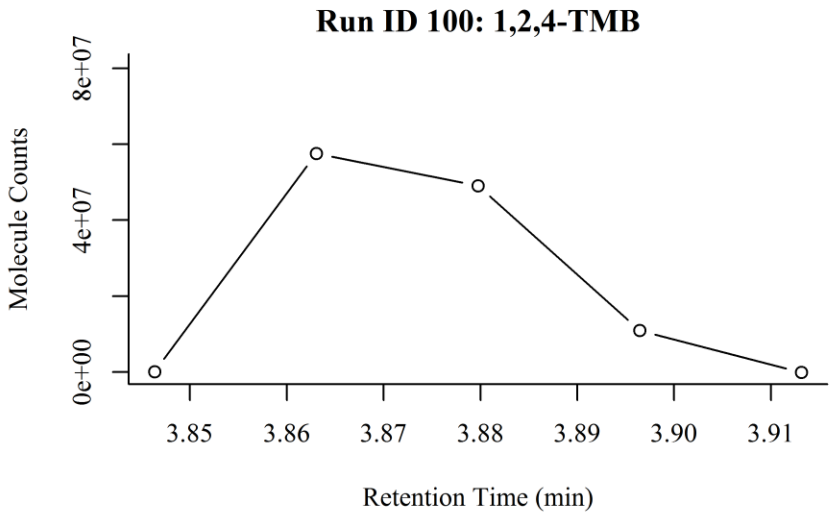


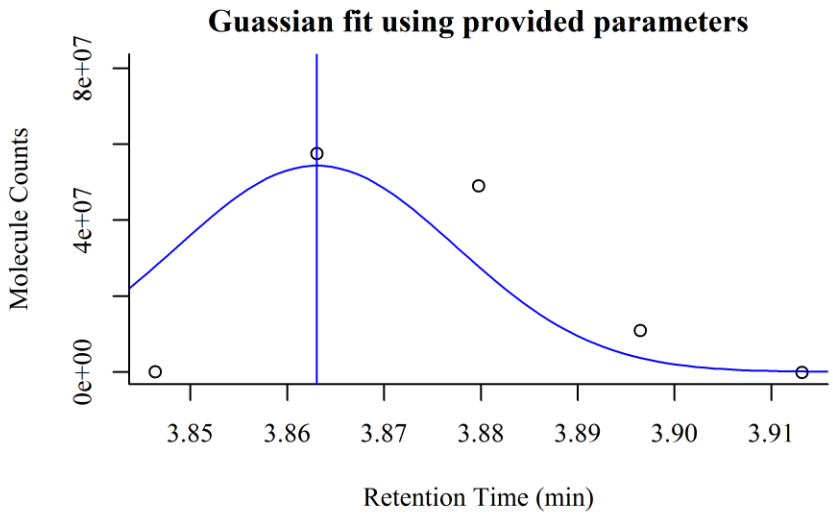
	34	36	38	40	44	50	52	54	58	56	2	4	6	8	10	16	20	24	46	61	64	12	62	48	1	18	19	22	
Column.Type:	RXI-5	RXI-5	RXI-5	RXI-5	RXI-5	RXI-5	RXI-5	RXI-5	RXI-5	RXI-5	RXI-1	RXI-1	RXI-1	RXI-1	RXI-1	RXI-1	RXI-1	RXI-1	RXI-5	RXI-5	RXI-5	RXI-1	RXI-5	RXI-5	RXI-1	RXI-1	RXI-1	RXI-1	R
Column.Length:	15 m	15 m	15 m	15 m	15 m	30 m	30 m	30 m	30 m	30 m	15 m	15 m	15 m	15 m	15 m	15 m	30 m	30 m	15 m	30 m	30 m	15 m	30 m	15 m	15 m	30 m	30 m	30 m	1
Injector.Type:	SSL	SSL	SSL	SSL	PTV	SSL	SSL	SSL	PTV	SSL	SSL	SSL	SSL	SSL	PTV	PTV	SSL	SSL	PTV	PTV	PTV	PTV	PTV	PTV	SSL	SSL	SSL	SSL	S
Liner.Material:	Deact	Deact	SKY	SKY	Deact	Deact	Deact	SKY	Deact	SKY	Deact	Deact	SKY	SKY	Deact	SKY	Deact	SKY	SKY	SKY	SKY	Deact	SKY	SKY	Deact	Deact	Deact	SKY	S
Liner.Shape:	Taper	Taper	Taper	Taper	Open	Open	Open	Open	Taper	Open	Open	Open	Open	Open	Taper	Taper	Taper	Taper	Open	Taper	Taper	Taper	Taper	Open	Open	Taper	Taper	Taper	T
Liner.Wool:	No	No	Yes	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes	No	No	Yes	No	No	Yes	Yes	No	No	Yes	No	Yes	Yes	No	No	Yes	Y
Septa.Type:	Green	Green	Red	Red	Red	Green	Green	Red	Red	Red	Red	Red	Green	Green	Green	Red	Red	Green	Green	Green	Green	Green	Green	Green	Red	Red	Red	Green	F
Syringe.Tip:	Fix	Fix	Rem	Rem	Rem	Rem	Rem	Fix	Fix	Fix	Fix	Fix	Rem	Rem	Rem	Fix	Rem	Fix	Fix	Rem	Rem	Rem	Rem	Fix	Fix	Rem	Rem	Fix	R
Injector.Temp:	200°C	300°C	200°C	300°C	300°C	200°C	300°C	200°C	200°C	300°C	200°C	300°C	200°C	300°C	200°C	300°C	300°C	300°C	200°C	200°C	300°C	300°C	200°C	300°C	200°C	200°C	300°C	200°C	3(
Split.Ratio:	25:1	25:1	25:1	25:1	25:1	25:1	25:1	25:1	25:1	25:1	25:1	25:1	25:1	25:1	25:1	25:1	25:1	25:1	25:1	1:1	25:1	25:1	25:1	25:1	1:1	25:1	1:1	25:1	
Injection.Volume:	2 µL	0.5	2 µL	0.5	0.5	0.5	2 µL	0.5	0.5	2 µL	0.5	2 µL	0.5	2 µL	0.5	2 µL	0.5	0.5	2 µL	0.5	2 µL	2 µL	0.5	0.5	0.5	2 µL	0.5	2 µL	(
Flow.Rate:	3 mL	0.5	3 mL	0.5	3 mL	3 mL	0.5	3 mL	0.5	0.5	0.5	3 mL	0.5	3 mL	3 mL	0.5	3 mL	3 mL	0.5	0.5	3 mL	0.5	0.5	3 mL	0.5	0.5	3 mL	0.5	(
Initial.Oven.Temp:	35°C	50°C	35°C	50°C	35°C	50°C	35°C	50°C	35°C	35°C	35°C	50°C	35°C	50°C	50°C	35°C	35°C	35°C	50°C	35°C	50°C	35°C	35°C	35°C	35°C	50°C	35°C	50°C	5
Initial.Hold.Time:	1.5 m	1.5 m	1.5 m	1.5 m	1.5 m	3 min	3 min	3 min	3 min	3 min	3 min	3 min	3 min	3 min	3 min	3 min	1.5 m	1.5 m	1.5 m	1.5 m	3 min	3 min	3 min	1.5 m	1.5 m	1.5 m	3 min	1.5 m	3
Ramp.Rate:	10°C	10°C	10°C	10°C	30°C	10°C	10°C	10°C	30°C	10°C	30°C	30°C	30°C	30°C	10°C	10°C	30°C	30°C	30°C	10°C	30°C	10°C	30°C	30°C	10°C	30°C	10°C	30°C	3
Final.Oven.Temp:	300°C	300°C	300°C	300°C	275°C	275°C	275°C	275°C	300°C	275°C	300°C	300°C	300°C	300°C	275°C	275°C	275°C	275°C	275°C	275°C	300°C	275°C	300°C	275°C	275°C	275°C	300°C	275°C	2
Transfer.Line.Temp:	300°C	200°C	200°C	300°C	200°C	300°C	200°C	200°C	300°C	300°C	200°C	300°C	300°C	300°C	200°C	200°C	300°C	200°C	200°C	300°C	300°C	300°C	200°C	300°C	300°C	200°C	200°C	300°C	2(
Scan.Rate:	1 scan	4 scan	4 scan	1 scan	4 scan	4 scan	1 scan	1 scan	4 scan	4 scan	1 scan	4 scan	4 scan	1 scan	1 scan	1 scan	1 scan	4 scan	4 scan	4 scan	4 scan	4 scan	1 scan	1 scan	4 scan	4 scan	4 scan	1 scan	4
Source.Temp:	200°C	275°C	275°C	200°C	200°C	200°C	275°C	275°C	275°C	200°C	200°C	275°C	275°C	200°C	275°C	275°C	275°C	200°C	200°C	275°C	275°C	200°C	200°C	275°C	275°C	200°C	200°C	275°C	2

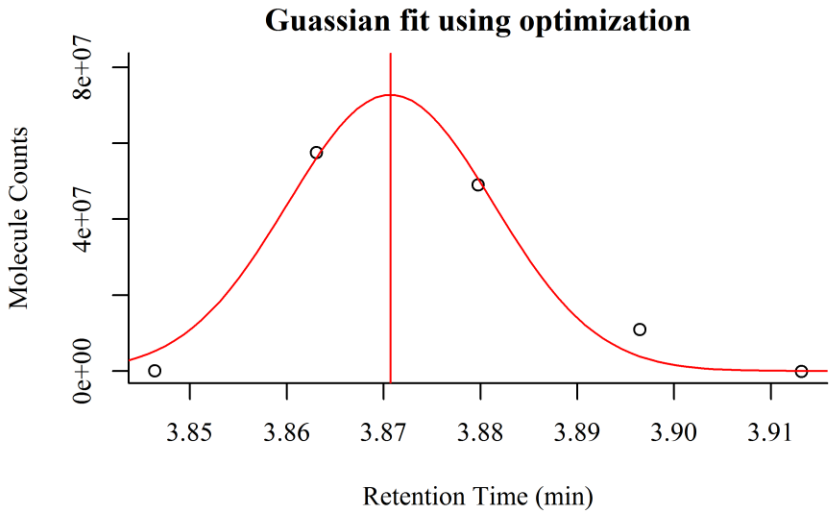
Number of Peaks

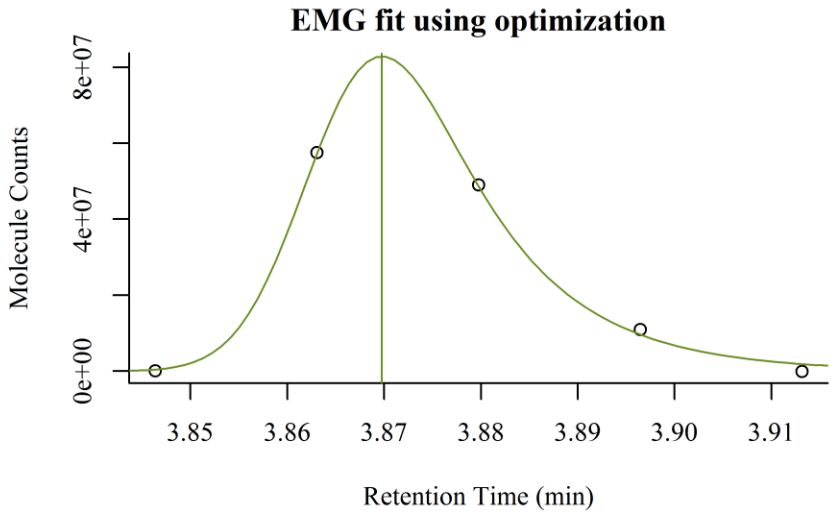


Peak Models


Run ID 100: 1,2,4-TMB
Molecule Counts
0e+00
4e+07
8e+07
3.85
3.86
3.87
3.88
3.89
3.90
3.91
Retention Time (min)


Guassian fit using provided parameters
Molecule Counts
0e+00
4e+07
8e+07
3.85
3.86
3.87
3.88
3.89
3.90
3.91
Retention Time (min)


Guassian fit using optimization
Molecule Counts
0e+00
4e+07
8e+07
3.85
3.86
3.87
3.88
3.89
3.90
3.91
Retention Time (min)


EMG fit using optimization
Molecule Counts
0e+00
4e+07
8e+07
3.85
3.86
3.87
3.88
3.89
3.90
3.91
Retention Time (min)

The 2nd Study – Experimental Design and Execution

The Second Time Around

The insight gained from the pilot study allowed to better understand the questions we should be asking and the people we should be asking.

1) Went back to OSAC subcommittee with pilot results for more discussions
2) Brought others from NIST with expertise in different areas into the project
   - GC experts, MS experts, data analytics experts, LC-MS experts
3) Brought practitioners and NIST scientists together to better understand what it means to have an “optimized method”
4) Brought it vendors to aid in the creation of the test mixture

Key Takeaways

Discussions with all who participated led to number of a key takeaways:

- Need to redefine the factors and the test mixture
  - Need to be cognizant of the levels we are choosing
- Need to understand differences between instrument manufacturers
  - Are the factors that most affect response universal?
- Need to better define what the “optimized” response is
- Need to consider the mass spectral data, not just the chromatographic data
- Laboratories were willing to aid in running a round-robin version of the study, but with minimal disruption to casework
- There are ways to automate data extraction and analysis
  - Leverage AMDIS

Results of Community Discussions

Discussions with the community at the OSAC in-person meeting led to increased discussions on the chosen factors.

- Removed Injector Type
  - Not a feasible factor to change for most labs
- Increased Split Ratios
- Updated septa choice to include Merlin Microseal
  - This forced a fixed needle choice
- Added column film thickness
- Fixed final column temperature, transfer line temperature, mass scan ranges, purge, and wash parameters

Revised Factor List

- Reduced to 18 factors
  - Maintained that some factors are easier to change than others
  - Had to consider ways to minimize operational downtime

Injector			GC			MS		
Factor	Low	High	Factor	Low	High	Factor	Low	High
Injector Temp	200 °C	300 °C	Column Length	15 m	30 m	Scan Rate	1 scan/s	4 scan/s
Split Ratio	10:1	50:1	Column Type	DB-1	DB-5	Source Temp	200 °C	276 °C
Volume	0.5 µL	2.0 µL	Column Thickness	0.25 µm	1.0 µm	Tune Type	Auto	Standard
Liner Material	Deact.	"SKY"	Flow Rate	0.5 mL/min	3.0 mL/min	Signal Threshold	0 cts	200 cts
Liner Shape	Open	Tapered	Initial Temp	35 °C	50 °C			
Liner Wool	Yes	No	Initial Hold	1.5 min	3.0 min			
Septa Type	Green	Merlin	Ramp Rate	10 °C/min	30 °C/min			

Results of Community Discussions

Discussions with the community at the OSAC in-person meeting also led to changes in the test mixture.

- Initial mixture did not consider high volatility polar compounds (i.e. ethanol and acetone)
  - Necessary for screening method to detect
- Use of isopropanol as the matrix was not practical
  - Ideally it would be a component in the mixture
  - Preferred dichloromethane as matrix
- Decided it would be best to have others create the test mixture
- Decided keeping the original compounds in the test mixture was crucial
- Mixture needed to be ampouled to keep stability

Revised Test Mixture

- Test mixture created and ampouled by Restek and shipped directly to participating laboratories
  - Create the E1618 mix most laboratories use
- Concentrations ranged from 10 µg/mL to 60 µg/mL for individual components

Ethanol	m-Ethyltoluene	2-Methylnaphthalene
2-Propanol	o-Ethyltoluene	2-Ethylnapthalene
Acetone	1,2,4-Trimethylbenzene	1-Ethylnaphthalene
n-Hexane	Indane	n-Hexadecane
Toluene	1,2,4,5-Tetramethylbenzene	n-Octadecane
o-Xylene	Naphthalene	Phytane
2,2,4-Trimethylhexane	n-Dodecane	n-Eicosane
n-Nonane	n-Dodecene	

Current Plan

We are currently running the 2nd study.

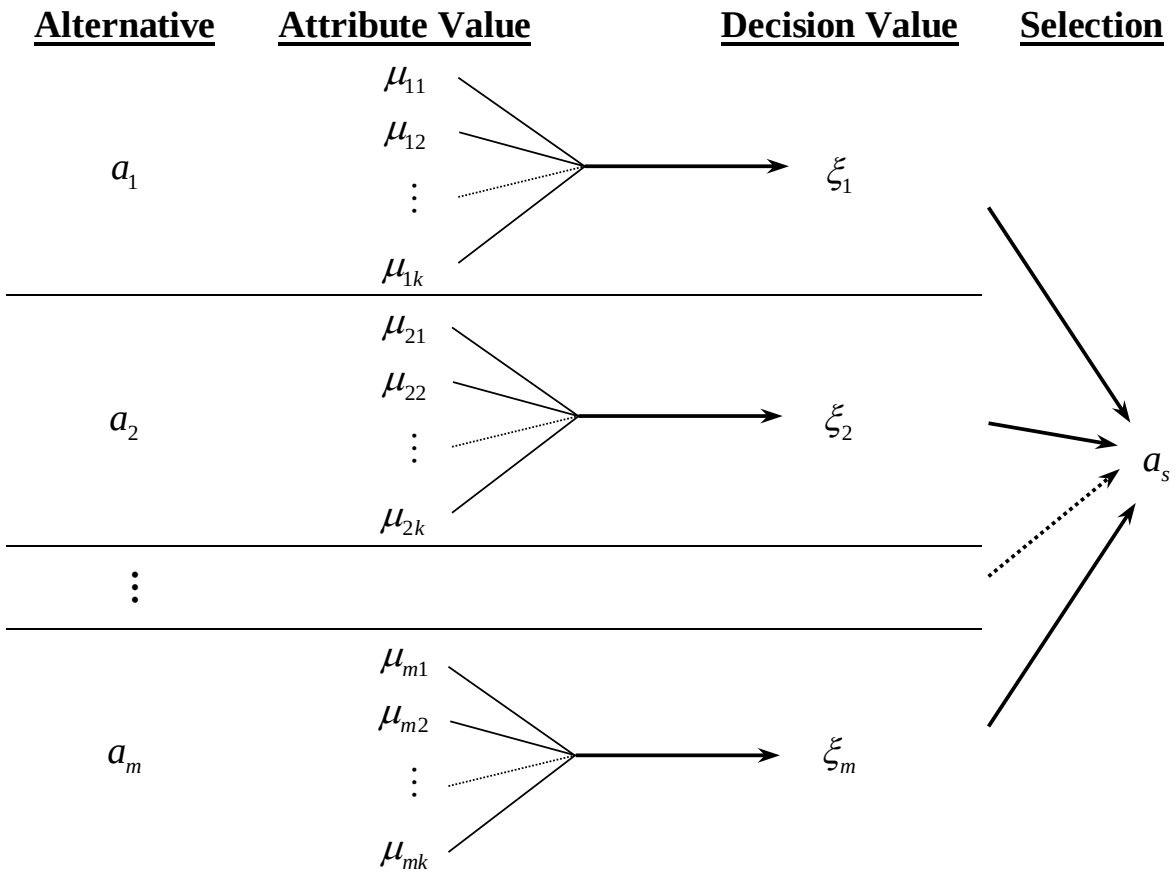
- Two forensic laboratories agreed to participate in the study
- Found an Agilent system in house to run the study on
  - Analysis being completed on both an Agilent and a Thermo
- All laboratories running the identical run matrix
  - Worked to minimize the time for the analysis by grouping hard-to-change factors even further
- Test mixture was shipped directly to participating laboratories
- Consumables kit provided to participating laboratories:
  - Columns, liners, nuts, septa, microseals, gold seals, needles, etc.

Vent Instrument, Cut 30 m Column in Half								
Change Liner, Septa, & Syringe								
Determine NEW solvent window for methods 3, 4, 11, 12, 21, 22, 29, 30								
G	**49**	12	15 m	RXI-1	1.00 µm	SKY	Tapered	No	Green
G	**50**	29							
G	**51**	3							
G	**52**	22							
G	**53**	3							
G	**54**	22							
G	**55**	12							
G	**56**	29							
Change Liner, Septa, & Syringe								
H	**57**	30	15 m	RXI-1	1.00 µm	Deactivated	Tapered	Yes	Merlin
H	**58**	4							
H	**59**	21							
H	**60**	11							
H	**61**	4							
H	**62**	30							
H	**63**	21							
H	**64**	11							
H	**65**	21							
H	**66**	11							
H	**67**	30							
H	**68**	4							
Change Liner, Septa, & Syringe								
I	**69**	12	15 m	RXI-1	1.00 µm	SKY	Tapered	No	Green
I	**70**	3							
I	**71**	22							
I	**72**	29							
Vent Instrument & Change Column								
Change Liner, Septa, & Syringe								
Determine solvent window for methods 7, 8, 15, 16, 17, 18, 25, 26								
J	**73**	18	30 m	RXI-1	0.25 µm	Deactivated	Tapered	No	Merlin
J	**74**	7							
J	**75**	16							
J	**76**	25							
J	**77**	7							
J	**78**	18							
J	**79**	25							
J	**80**	16							
Change Liner, Septa, & Syringe								
K	**81**	8	30 m	RXI-1	0.25 µm	SKY	Tapered	Yes	Green
K	**82**	17							
K	**83**	15							
K	**84**	26							
K	**85**	26							
K	**86**	17							
K	**87**	8							
K	**88**	15							
K	**89**	17							
K	**90**	8							
K	**91**	26							
K	**92**	15							
Change Liner, Septa, & Syringe								
L	**93**	25	30 m	RXI-1	0.25 µm	Deactivated	Tapered	No	Merlin
L	**94**	7							
L	**95**	18							
L	**96**	16							

The 2nd Study – Analysis

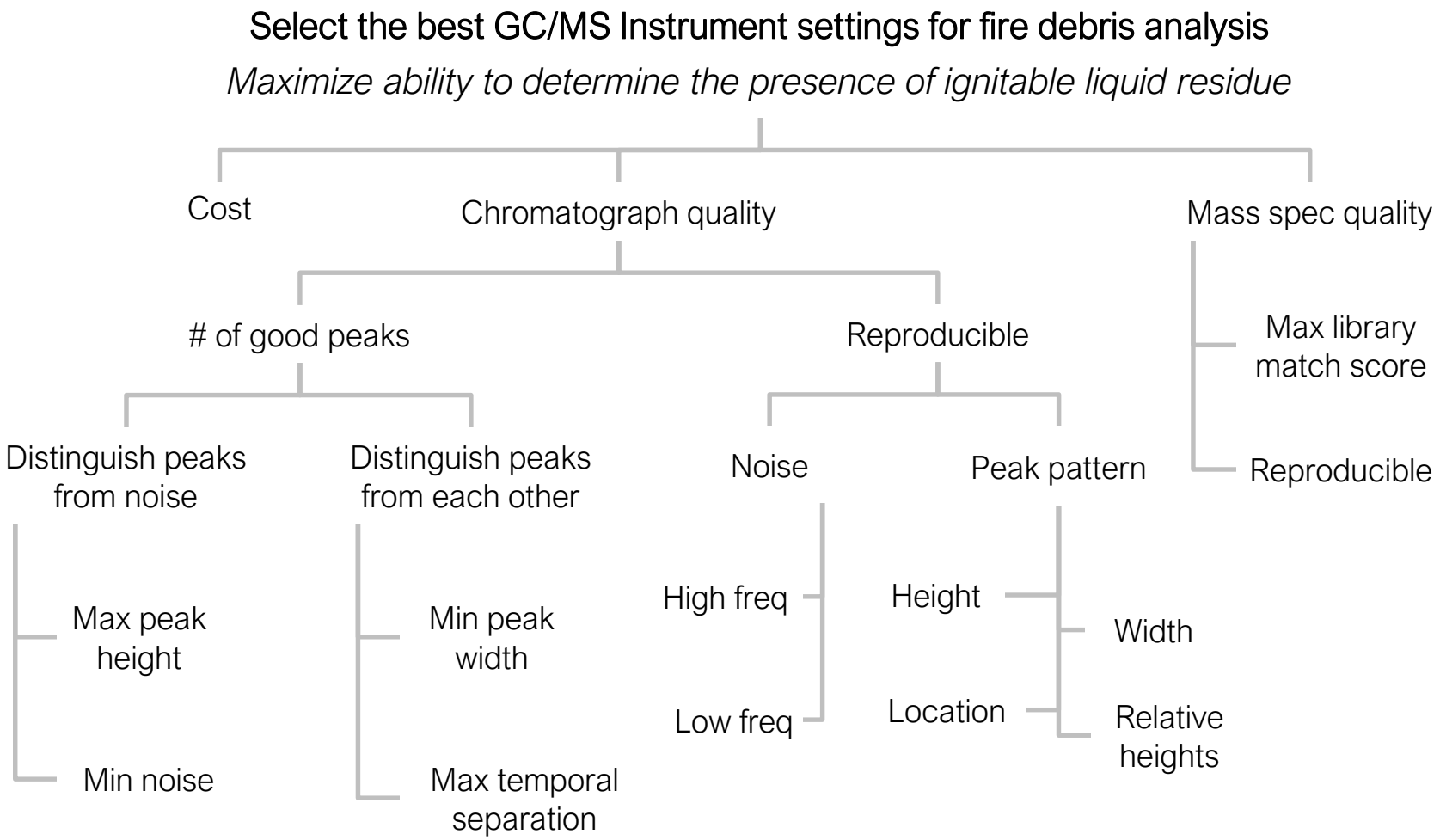
Decision Analytics

- Multi-attribute selection decision
- Incorporate forensic analysts' preferences
- Generalizable approach



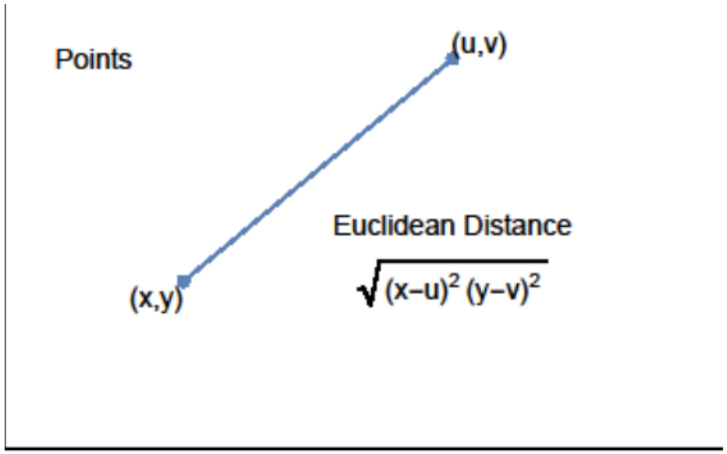
Decision model: $\xi_i = f(\mu_{i1},\ldots,\mu_{ik}) = \sum_{j=1}^{k} \lambda_j v_j(\mu_{ij}), \quad \sum_{j=1}^{k} \lambda_j = 1$

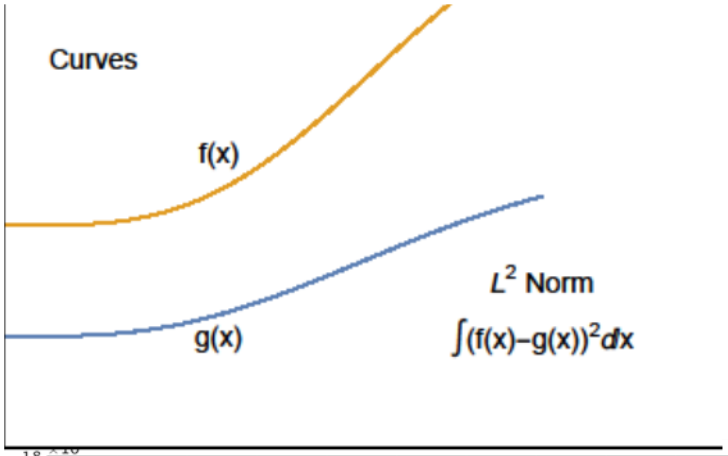
Objective Hierarchy

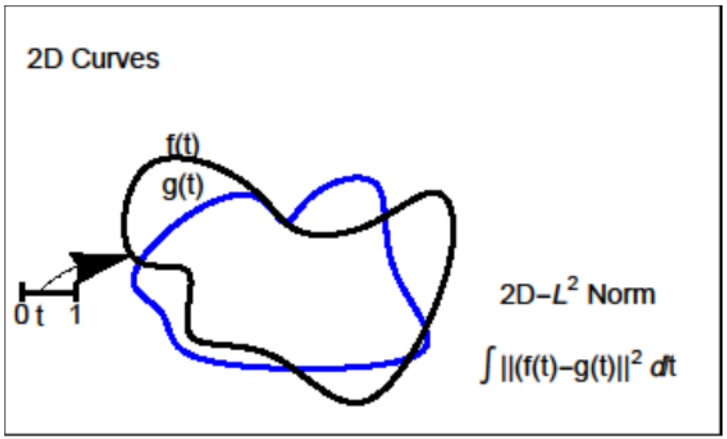


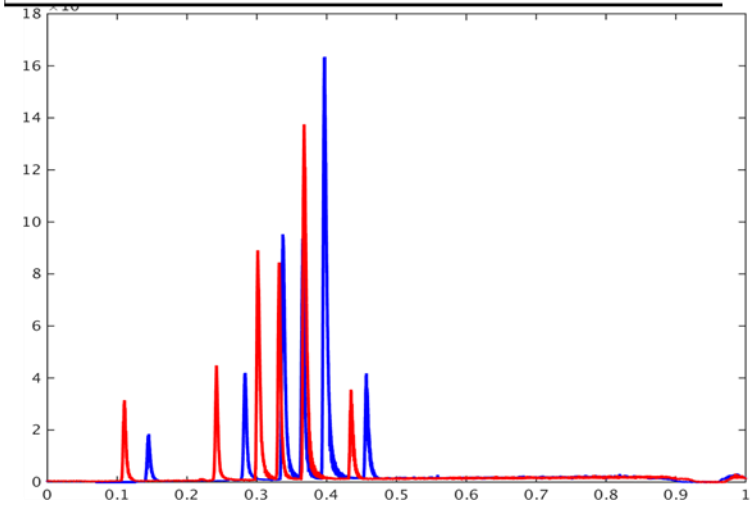
Novel Data Treatment – Shape Analysis

Distance Metrics





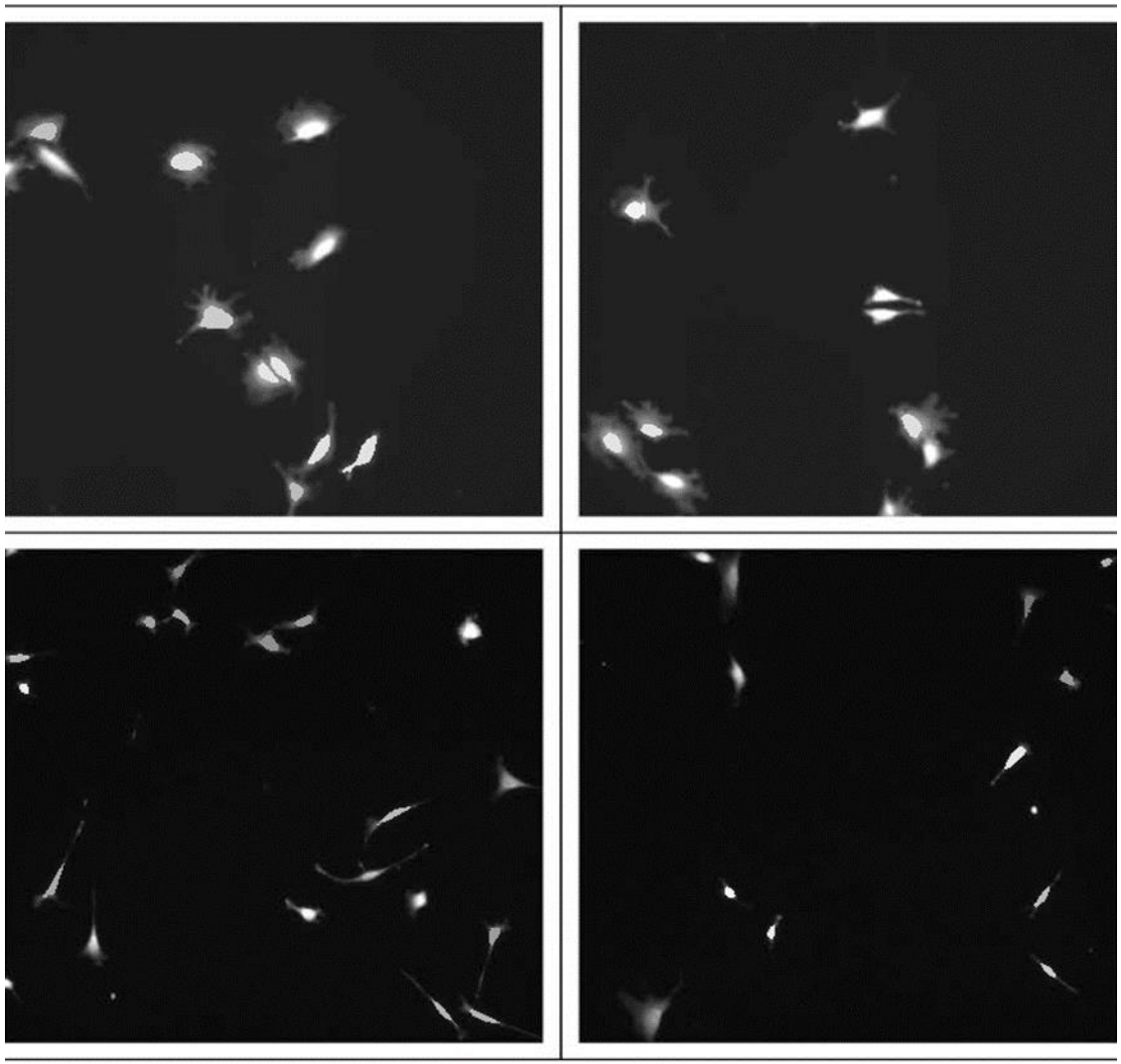




The L^2 Metric

- L^2 is a fine mathematical metric
- L^2 says the two objects differ, but it does not capture the differences our mind sees
- We see shape differences

Two Muscle Cell Populations



For Chromatographs

1. We see differences in time locations of matched peaks across measurements
2. Height differences among matched peaks across measurements
3. Relative heights of selected peaks

Shape as a Metric

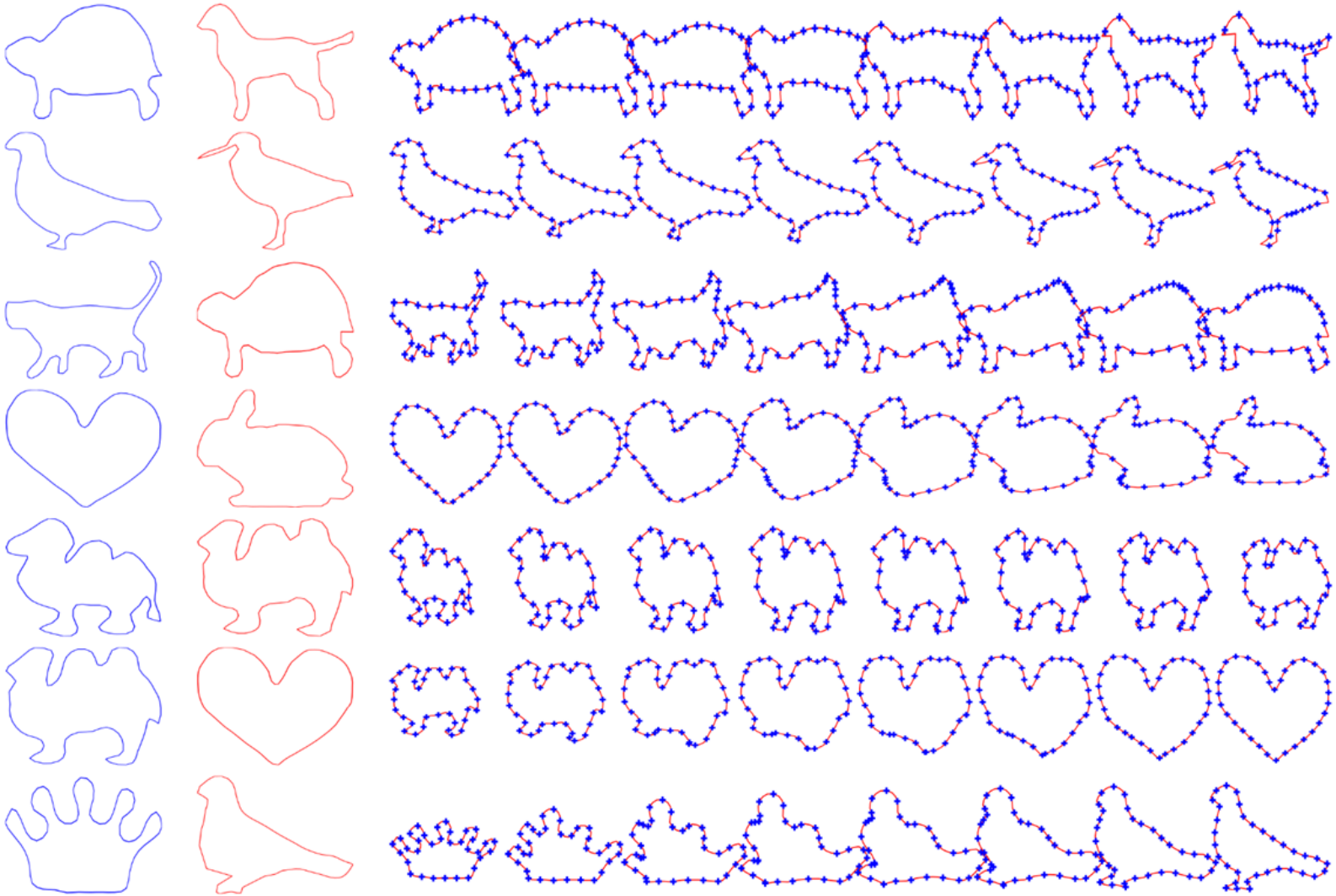
What if we could:

1. Mathematically quantify the shape of an object!
2. Formally, define a distance metric between two shapes

Shape as a Metric

- We could then define a space of shapes
- We could define a path in this space between two shapes
- We could define the shortest path between two shapes and call it shape distance

Path of shapes between two curves



What Happened in Plots

- The speed as one goes around one curve is changed in order to better align its points with points on the other curve

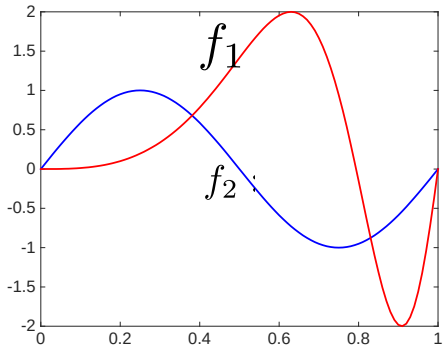
  A speed change means replacing f(t) by f(**ϒ(t)**)

- In addition, it is rotated to get a better alignment

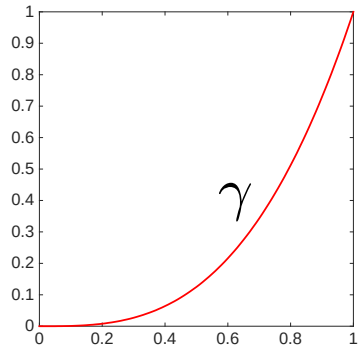
  Rotation means replacing f(t) by O(f(t))

  O a rotation matrix

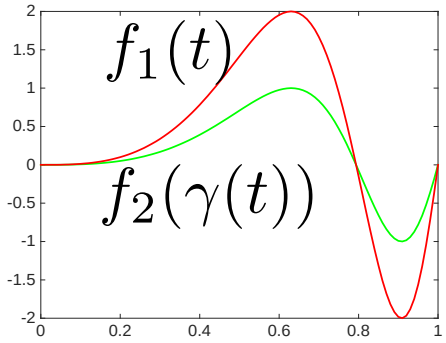
Diffeomorphism (Time Warping)



Given functions



Time-warping function



Aligned functions

Peaks and Valley have been aligned

The Shape Metric

Given two curves f and g with shapes q_f ,q_g

$$d_s(q_f,q_g) = \min_{\gamma,o} \int_0^1 \left\| q_f(t) - \sqrt{\dot{\gamma}}\, O(\,(q_g\gamma(t))) \right\|^2 dt$$

The integrand presents the distance between the shape of curve f and the shape of an aligned g curve, $O(\,(q_g\gamma(t)))$

$$q_f(t) = \frac{f(\dot{t})}{\sqrt{||f(\dot{t})||}}$$ (Shape Transform)

$d_s(q_f,q_g)$ is usually computed numerically

Shape Metric in 1D

In 1D the shape transform reduces

$$q_f(\mathrm{t}) = \mathrm{sign}(\dot{f}(\mathrm{t}))\sqrt{|\dot{f}(t)|}$$

Since there are no rotations

$$d_s(q_f, q_g) = \min_{\gamma} \int_0^1 (q_f(t) - \sqrt{\dot{\gamma}}\, q_g(\gamma(t)))^2 dt$$

Proposed Metrics

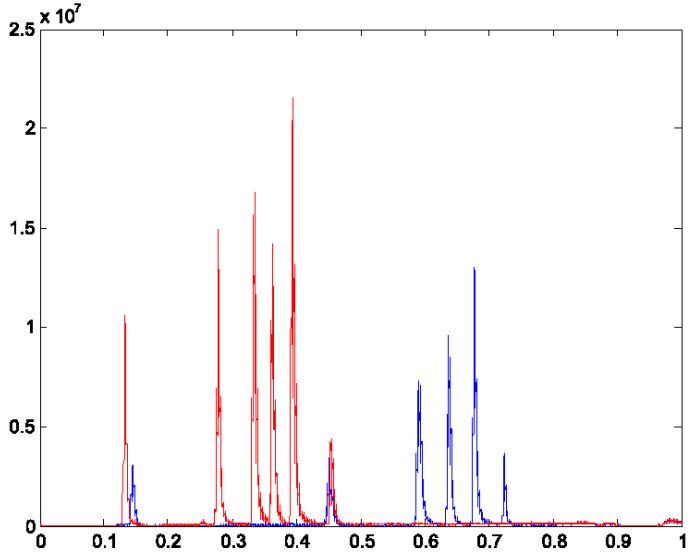
- We define two distances to compare chromatograms
- Vertical or Amplitude Distance: $d_a(f_1, f_2) \doteq \|q_1 - (q_2, \gamma^*)\|$

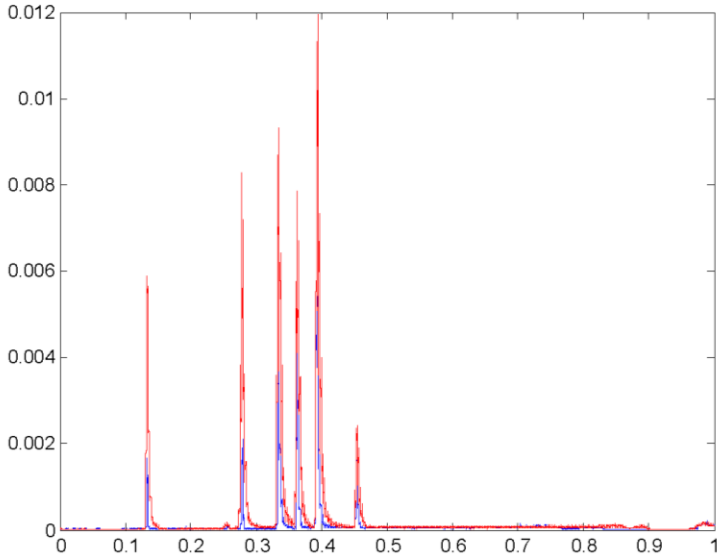
  where: $\gamma^* = \underset{\gamma}{\operatorname{arginf}} \|q_1 - (q_2, \gamma)\|$

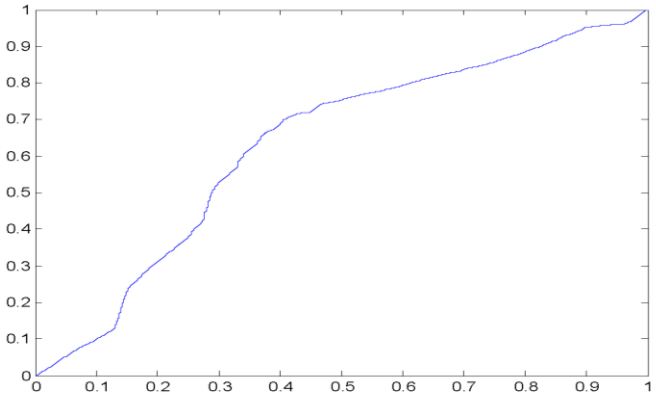
  - This is the L^2 norm between the aligned . In a sense it measures the height differences between matched peaks.
  - It can take any value between 0 and infinity.
- Horizontal or Phase Distance:

  $$d_p(f_1, f_2) = \cos^{-1}\left(\int \sqrt{\dot{\gamma}^*(t)}dt\right)$$

  - This measures the amount of warping needed to align the two chromatograms. In other words, it measures the time differences between the matched peaks.
  - It takes value between 0 and pi/2.



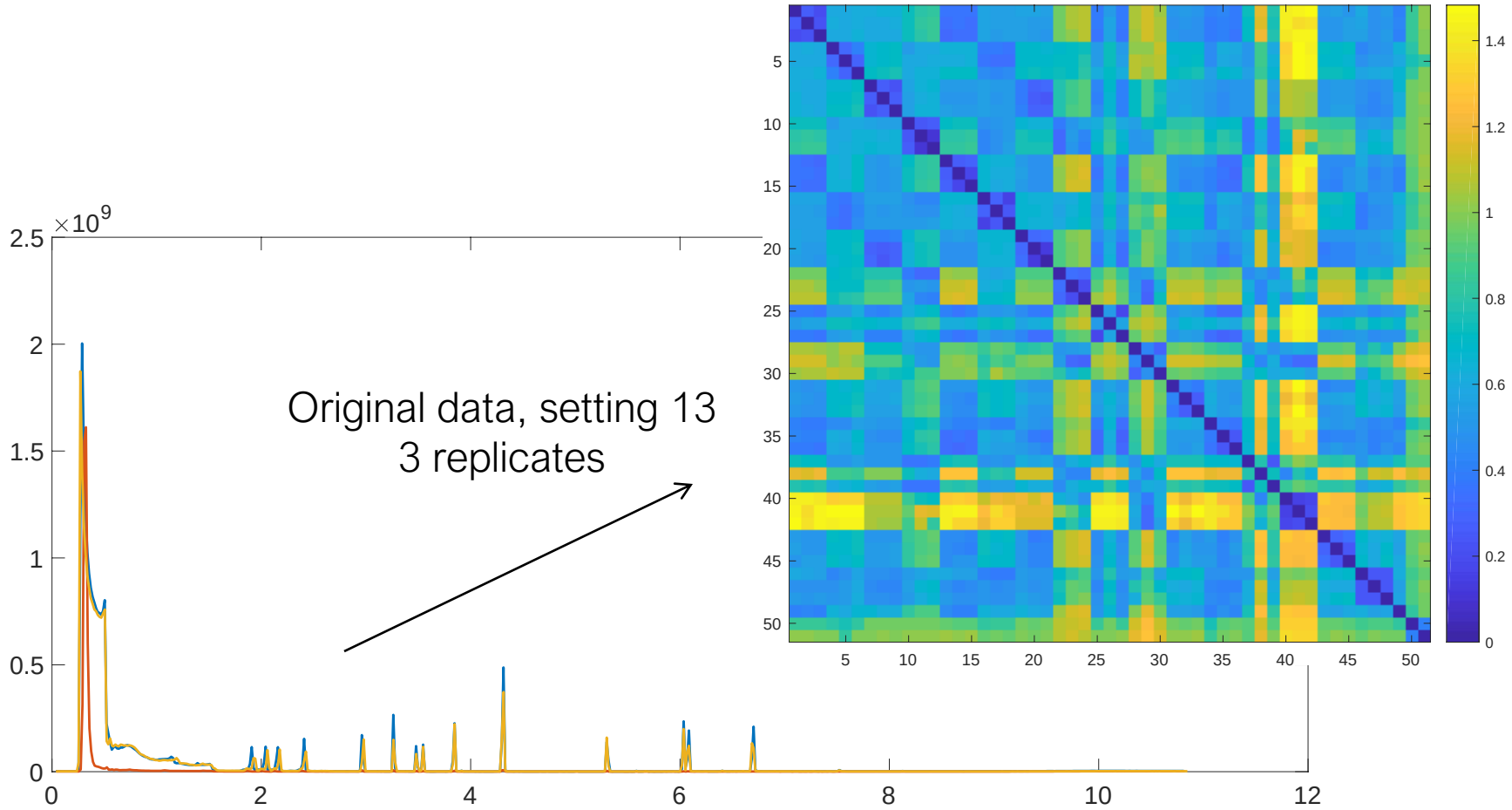




Aligned Chromatographs

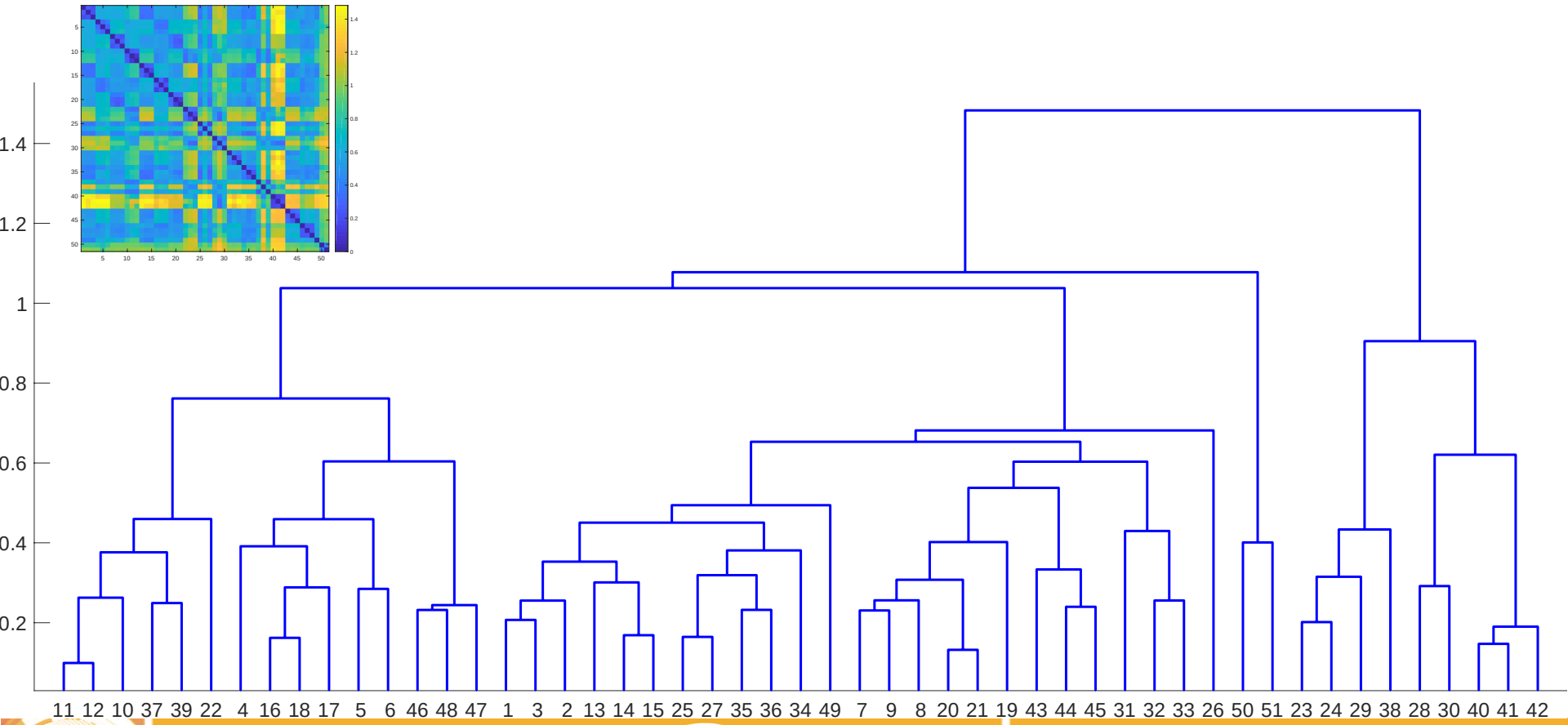
Pairwise Distance Matrices

Using Shape Distance, d_a: with 1000 sample points along each chromatogram
Using only first 51 chromatograms



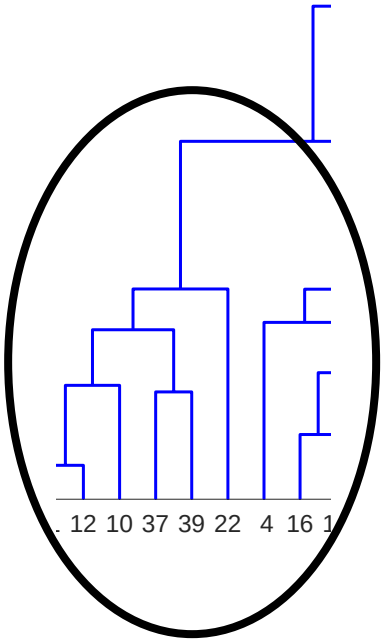
Clustering Chromatograms

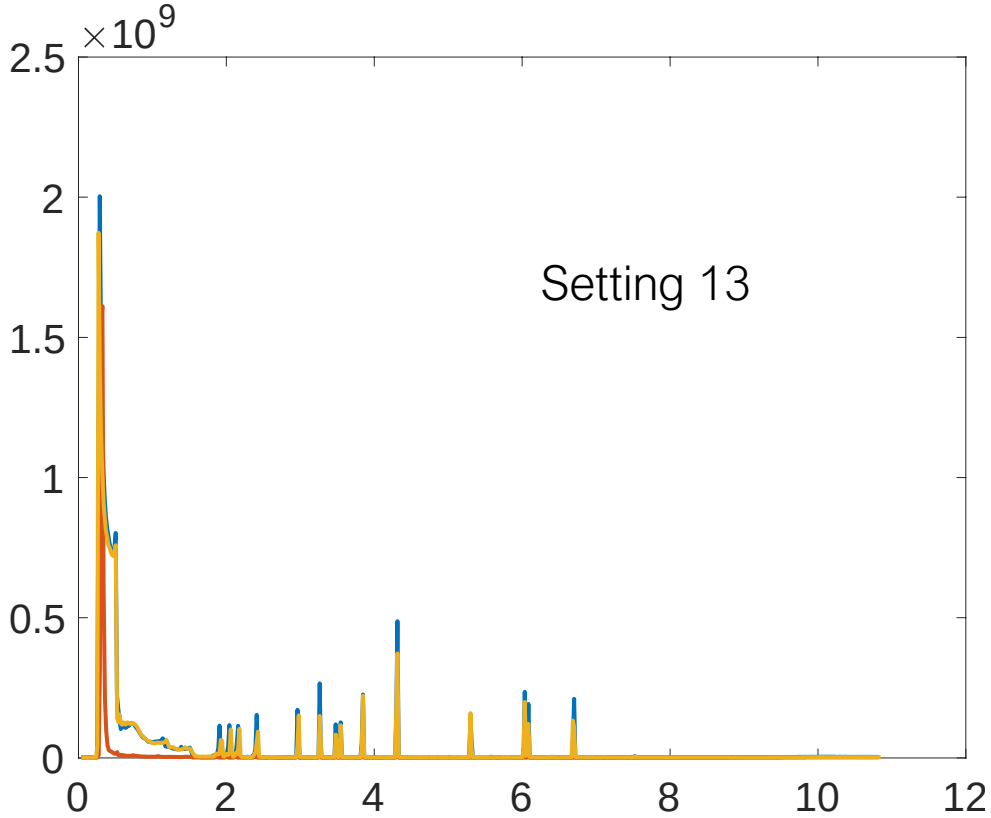
- Using Shape Distance, d_a: with 1000 sample points along each chromatogram
  - Using only first 51 chromatograms



Clustering Chromatograms

- Using Shape Distance, d_a: with 1000 sample points along each chromatogram
  - Using only first 51 chromatograms





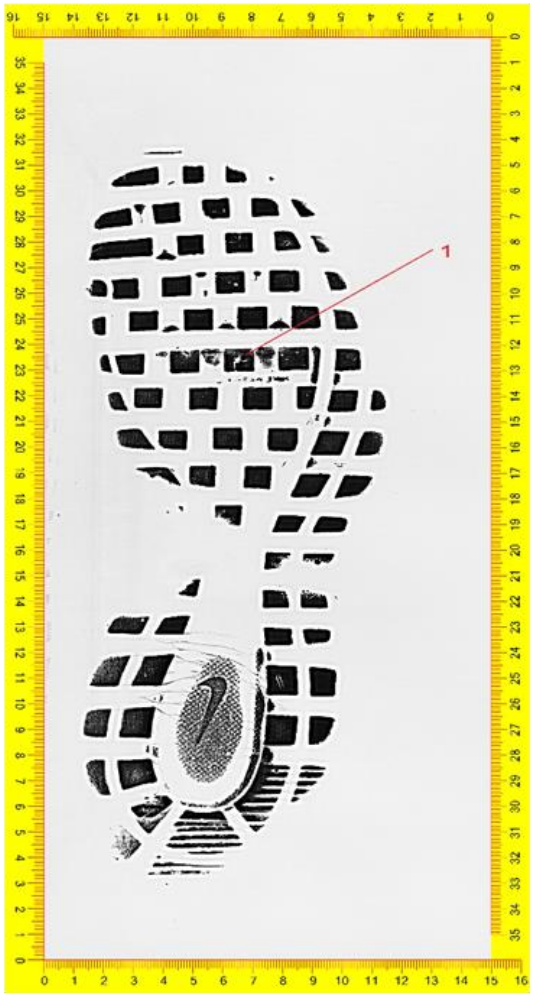
One replicate is different from others in shape.

Fire Debris Data

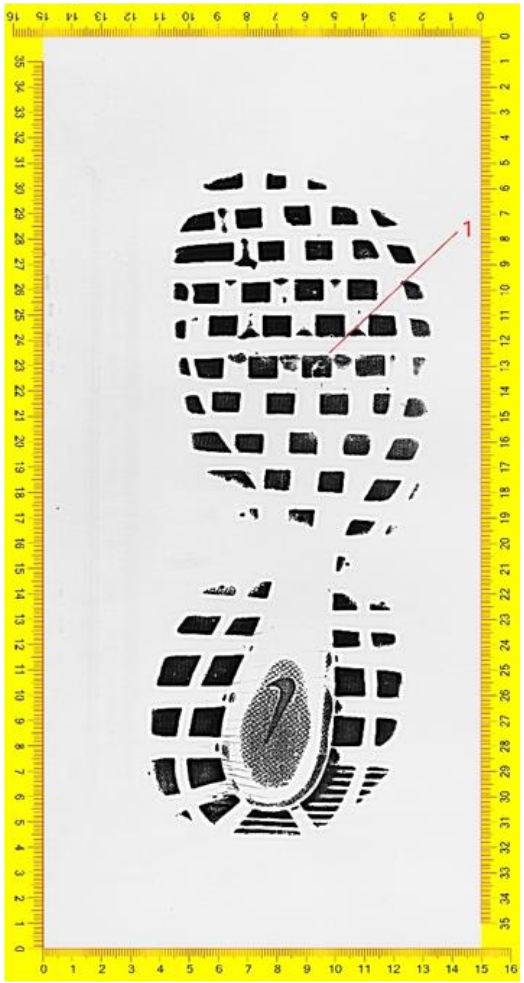
Setting	Rep 1	Rep 2	Rep 3	Dist 1	Dist 2	Dist 3	Dist Sum	Diff 1	Diff 2	Diff 3	Diff Sum
1	4	20	22	0.0054	0.0038	0.0044	0.0135	0.0166	0.0157	0.0155	0.0478
2	1	18	23	0.0188	0.0043	0.0207	0.0438	0.0208	0.0204	0.0192	0.0604
3	2	19	24	0.0022	0.0012	0.0016	0.005	0.0051	0.0046	0.004	0.0137
4	3	17	21	0.0108	0.0022	0.004	0.017	0.0157	0.009	0.0067	0.0315
5	5	11	16	0.0022	0.0014	0.0034	0.0069	0.0016	0.0012	0.0021	0.005
6	7	12	14	0.0056	0.009	0.002	0.0166	0.0147	0.0068	0.007	0.0285
7	6	9	13	0.0084	0.0033	0.0036	0.0152	0.0327	0.0166	0.0176	0.0669
8	8	10	15	0.0232	0.0022	0.0138	0.0392	0.0201	0.0142	0.014	0.0483
9	28	44	48	0.026	0.0302	0.0324	0.0886	0.077	0.1076	0.066	0.2506
10	27	41	45	0.0601	0.1552	0.0647	0.28	0.095	0.0944	0.1103	0.2997
11	26	42	46	0.1777	0.0448	0.0493	0.2718	0.1097	0.0608	0.0594	0.2299
12	25	43	47	0.0159	0.0074	0.0081	0.0315	0.0769	0.0209	0.0228	0.1206
13	31	33	38	0.0322	0.0974	0.0181	0.1477	0.048	0.0676	0.0479	0.1635
14	32	34	40	0.0109	0.0122	0.0285	0.0516	0.0082	0.0063	0.0057	0.0201
15	30	36	37	0.0247	0.0052	0.0111	0.0409	0.0068	0.0045	0.005	0.0163
16	29	35	39	0.0113	0.0126	0.0024	0.0263	0.0024	0.0024	0.0022	0.0069
17	145	152	154	0.0826	0.0424	0.0687	0.1938	0.049	0.0526	0.0526	0.1542
18	147	151	156	0.026	0.0355	0.0244	0.0859	0.0381	0.0485	0.0504	0.137
19	146	149	153	0.0196	0.0066	0.0204	0.0466	0.0113	0.0078	0.0056	0.0246
20	148	150	155	0.0369	0.0204	0.0633	0.1206	0.0075	0.006	0.0106	0.0241
21	160	161	167	0.0045	0.0046	0.0049	0.014	0.0056	0.0058	0.0063	0.0177
22	157	163	166	0.0026	0.0003	0.0036	0.0064	0.0133	0.0119	0.0112	0.0365
23	159	162	165	0.0009	0.0014	0.0006	0.0028	0.0118	0.0147	0.0196	0.0461
24	158	164	168	0.0099	0.0012	0.0048	0.016	0.0045	0.0044	0.0053	0.0142
25	174	178	184	0.0006	0.001	0.0005	0.0021	0.0275	0.0356	0.0201	0.0832
26	173	177	181	0.0024	0.0021	0.0041	0.0087	0.0038	0.0041	0.0055	0.0133

Shoe Print Application

Shoe Print 1



Shoe Print 2



Shoe Print Application

RAC 1

RAC 2

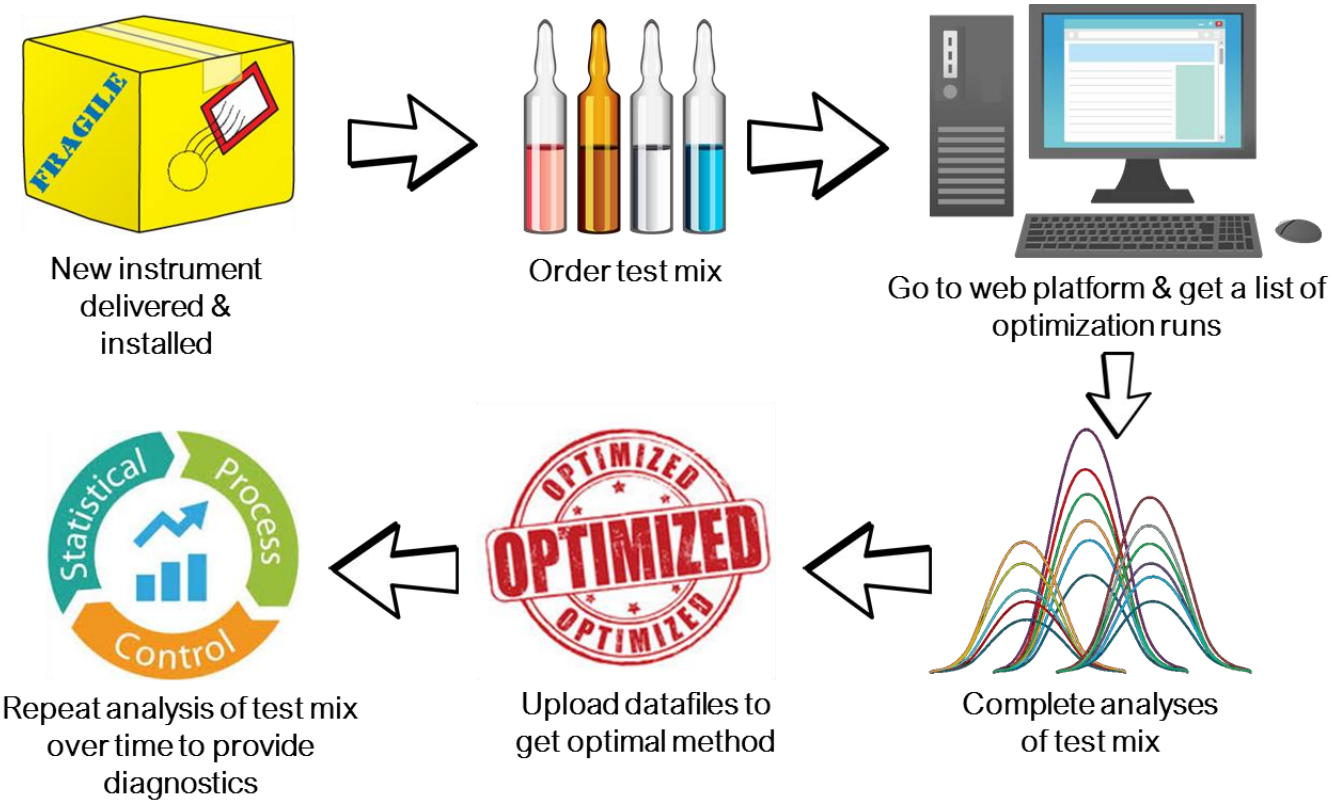
Next Steps

Where We Are

- Currently running samples using redesigned matrix and test mixture
- Working with practicing laboratories to run these samples
- Developing the objective hierarchy approach for data interpretation
- Developing the necessary methods to automate data analysis through AMDIS
- Working on the development of the Response Surface Methodology structure

- All of this work is just for to determine what settings matter
  - Not part of what is envisioned for laboratories to have to run themselves
  - This will minimize the work other laboratories will have to do

Where We Hope To Go



While this is the vision of where we want to be, we are still acquiring the necessary foundational research to make this possible.

Applying To Other Fields

Taking the lessons learned from the pilot study, we are also applying this process to optimization of a screening method for drugs of abuse.

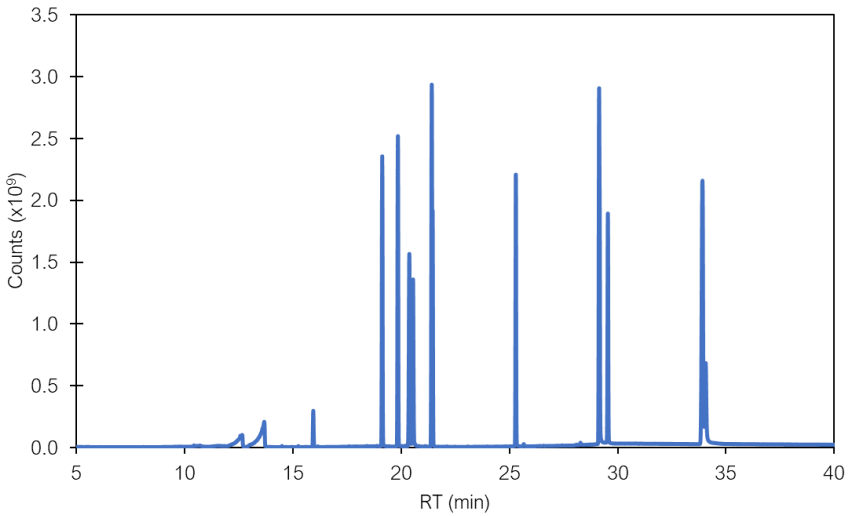
- Same process, same factors, new test mixture
  - Looking at DB-35 column instead of DB-1
- One forensic laboratory is also taking part in the study
- Drug analysis presents other challenges
  - Carryover, solvent compatibility, and high boiling point compounds
- Believe the relative weighting of response factors will be much different than fire debris analysis

Applying To Other Fields

Worked with Cayman Chemical to develop a custom test mixture.

- 15 Compounds in isopropanol matrix
- Concentration of 250 µg/mL per compound
- Tried to span compound classes and retention times

(±)-Methamphetamine	(S)-5-fluoro ADB
α-PVP (HCl)	Nandrolone
α-PBP (HCl)	Heroin
Ethylone (HCl)	Furanyl Fentanyl (HCl)
Butylone (HCl)	Furanyl Fentanyl 3-furancarboxamide Isomer (HCl)
PCP (HCl)	Alprazolam
TCP (HCl)	Stanozolol
Cocaine	



Acknowledgements

This project would not be possible without the value input, time, and effort of:

- Reta Newman, Pinellas County Forensic Laboratory
- Brenda Christy, VA Dept. of Forensic Science
- Members of the OSAC FD&E Subcommittee
- Ben Place, NIST
- Gary Mallard, NIST
- Bill Wallace, NIST
- Bill MacCreehan, NIST

Thank you.

edward.sisco@nist.gov
dennis.leber@nist.gov
charles.hagwood@nist.gov