SAND2008-1801C

Analysis of Flow-Cytometer Scattering and Fluorescence Data to Identify Particle Mixtures

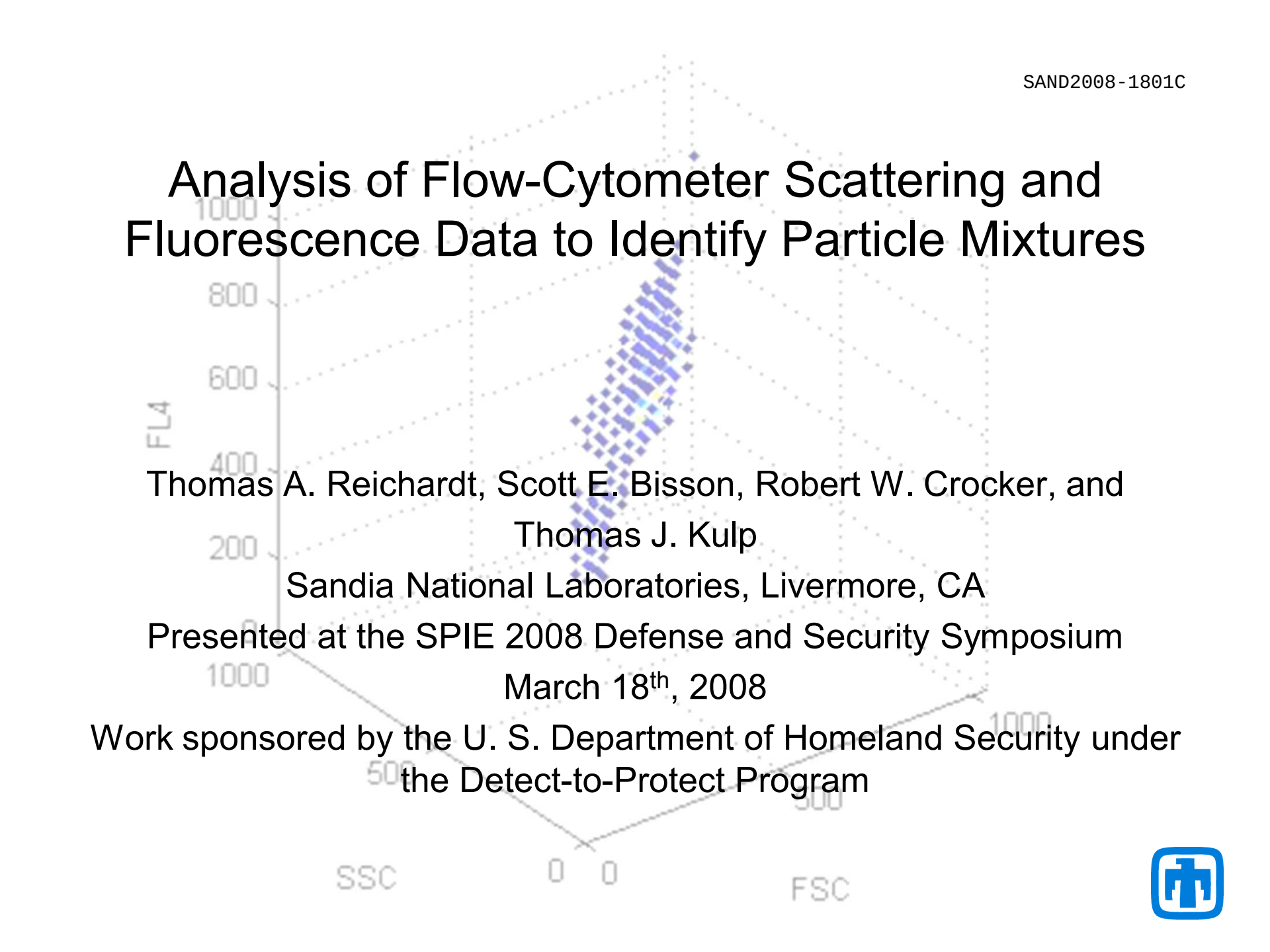
Thomas A. Reichardt, Scott E. Bisson, Robert W. Crocker, and Thomas J. Kulp

Sandia National Laboratories, Livermore, CA

Presented at the SPIE 2008 Defense and Security Symposium

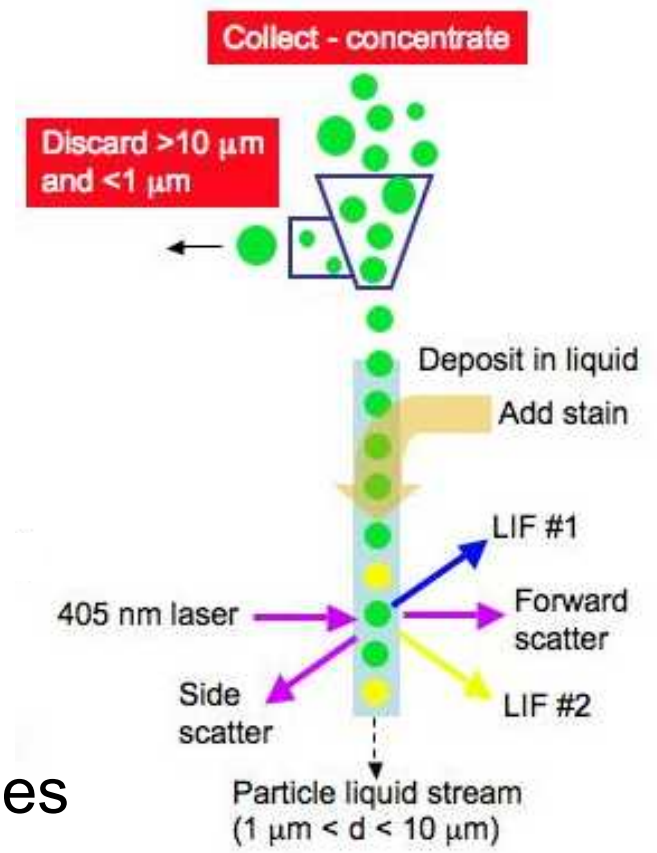
March 18th, 2008

Work sponsored by the U. S. Department of Homeland Security under the Detect-to-Protect Program

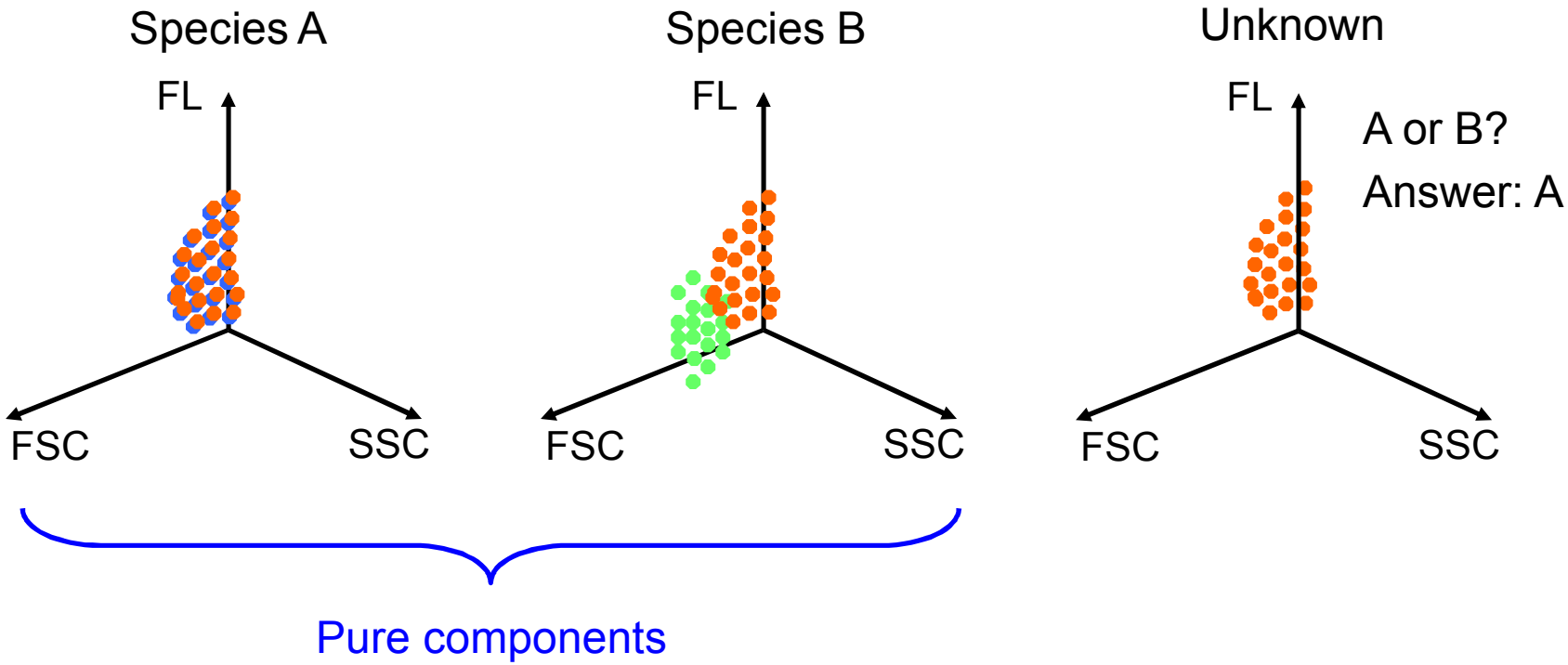


Review of flow-cytometer staining approach

- Evaluate cytometer staining to reduce false-alarm rates for early-warning sensor
- For each particle:
  - Forward scatter (FSC)
  - Side scatter (SSC)
  - 2 fluorescence (FL) channels
- Only 4 pieces of information
- 1 min, 300 L/min: 1000's of particles



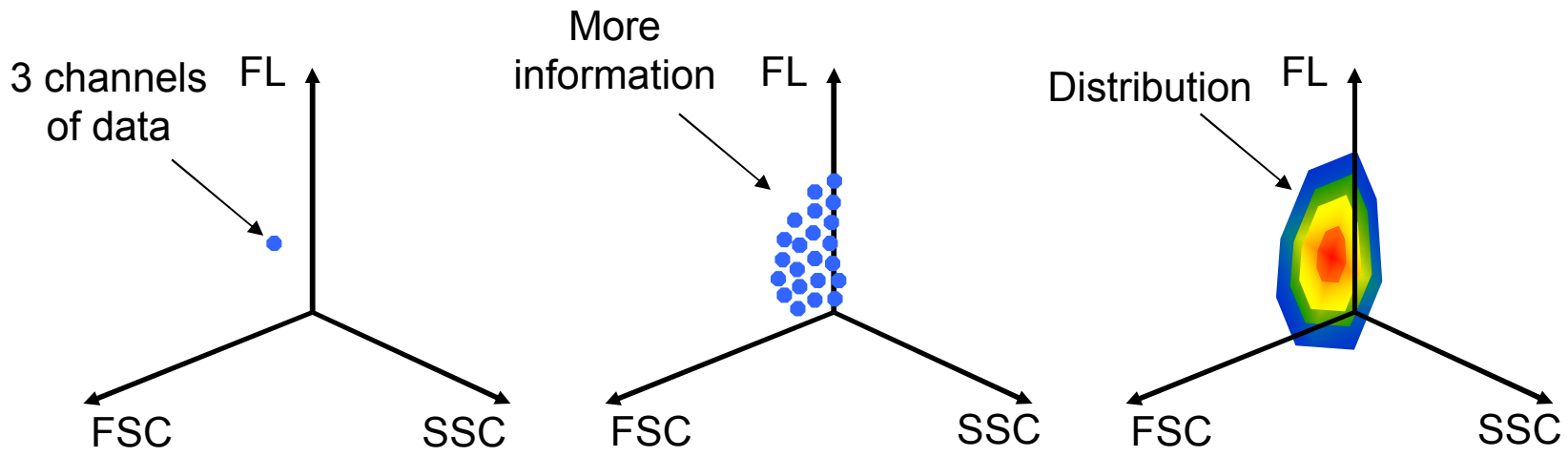
Bioaerosol characterization with 3 channels



- 300 L/min × 1 min × 100 particles/L = 30,000 particles
- No need to classify individual particles upon acquisition

Linear unmixing

- High throughput enables fitting of distributions

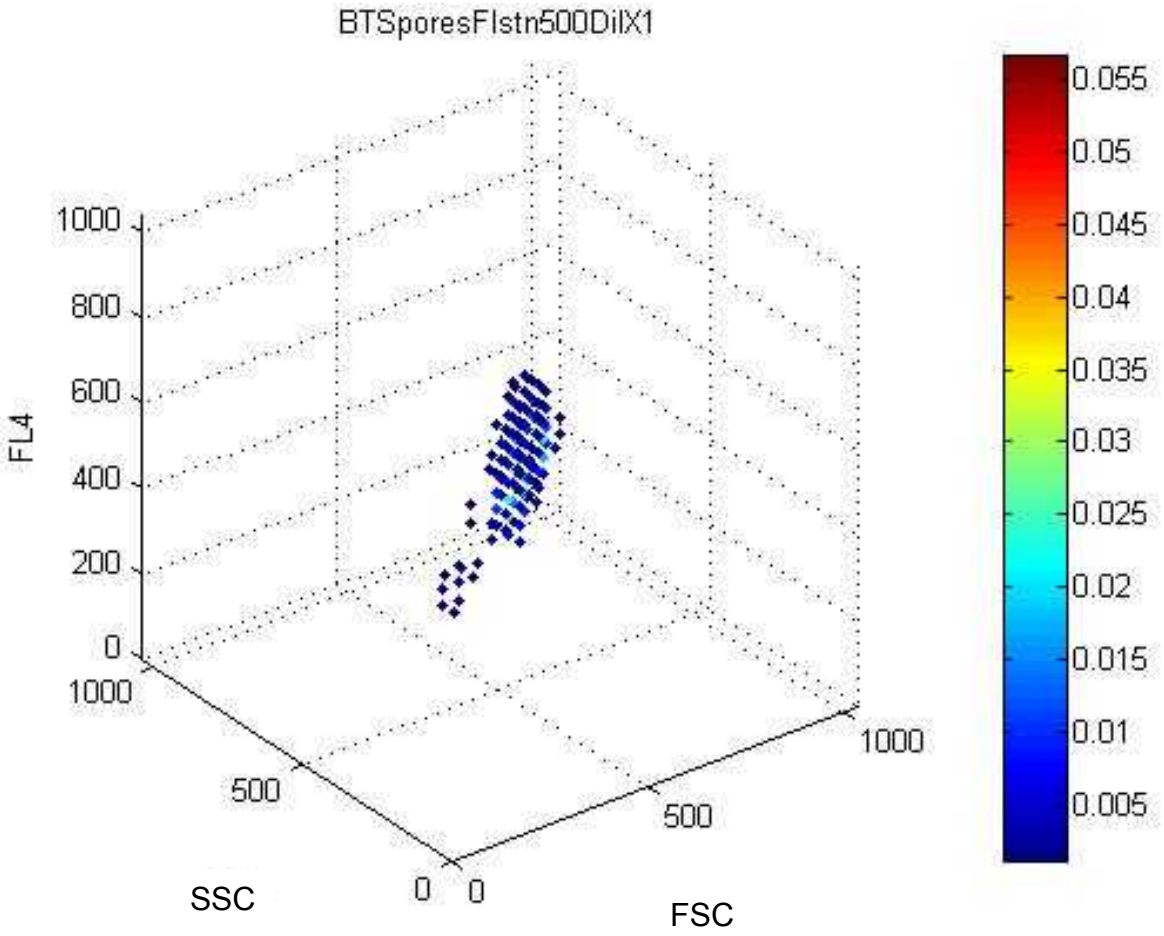


- Linear Unmixing:

  Assemble *library* of known sample distributions

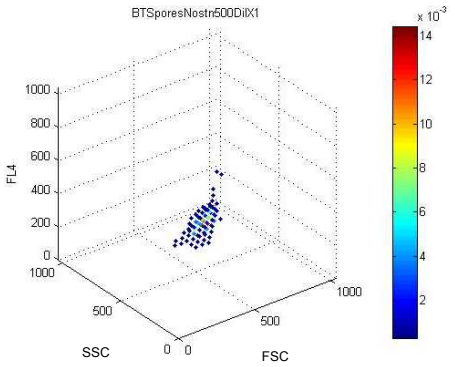
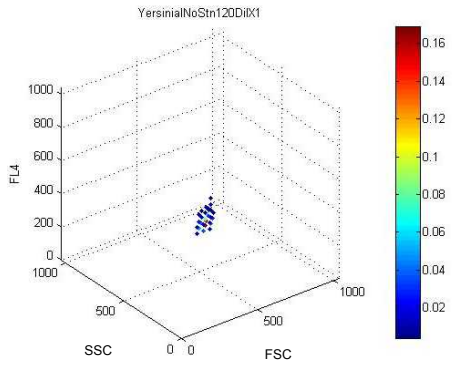
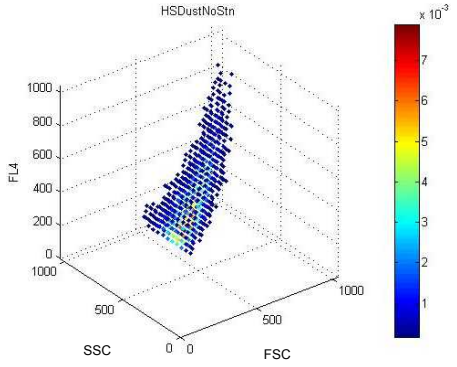
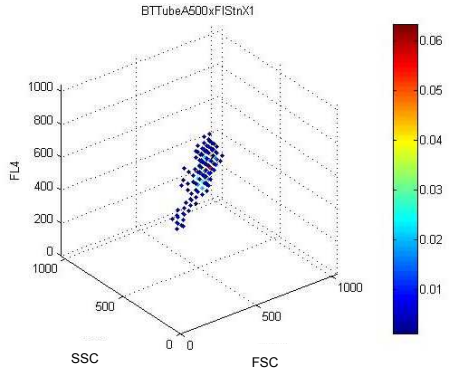
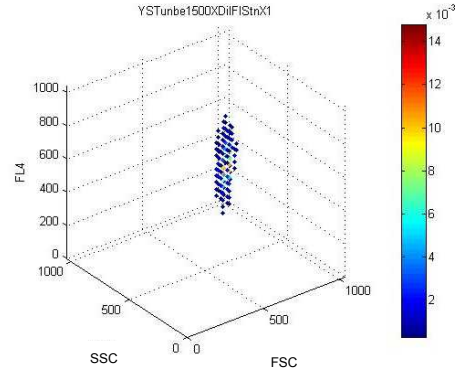
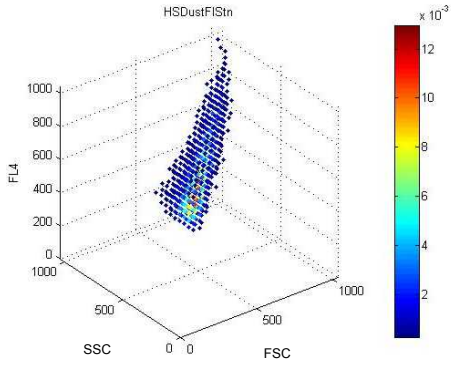
  Unknown mixture = *linear combination* of known samples

Step #1: Bin channel data on known samples



26 × 26 × 26 = 17,576 bins

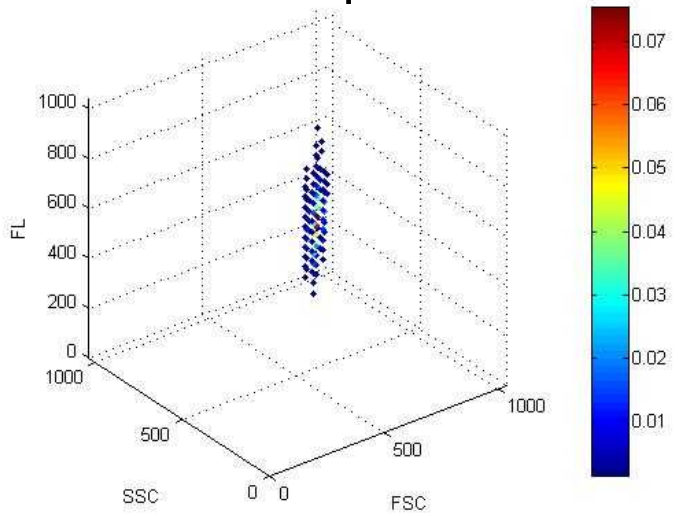
Step #2: Assemble known distributions into a library

	BT	Yersinia	SNL house dust
Native			
Stained			

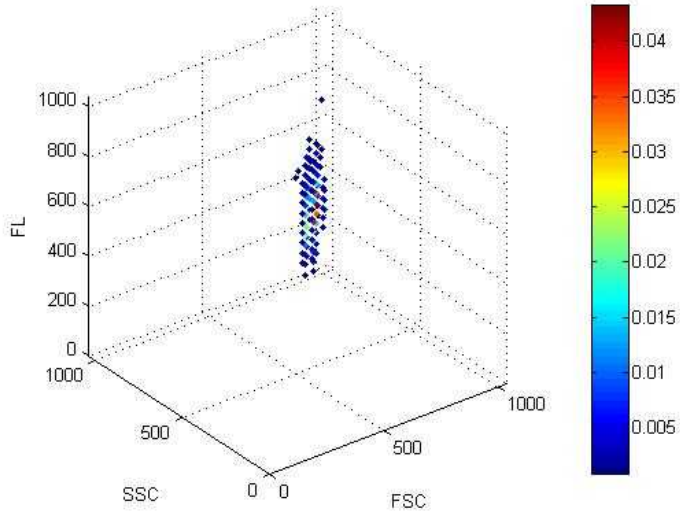
Distributions are unfolded to form 17,576-element vectors

Step #3: Acquire/bin unknown sample distribution, fitting as linear combination of library distributions

Sample



Fit



$$\mathbf{a} = \mathbf{f}\,\mathbf{P}^{T} + \mathbf{e}$$

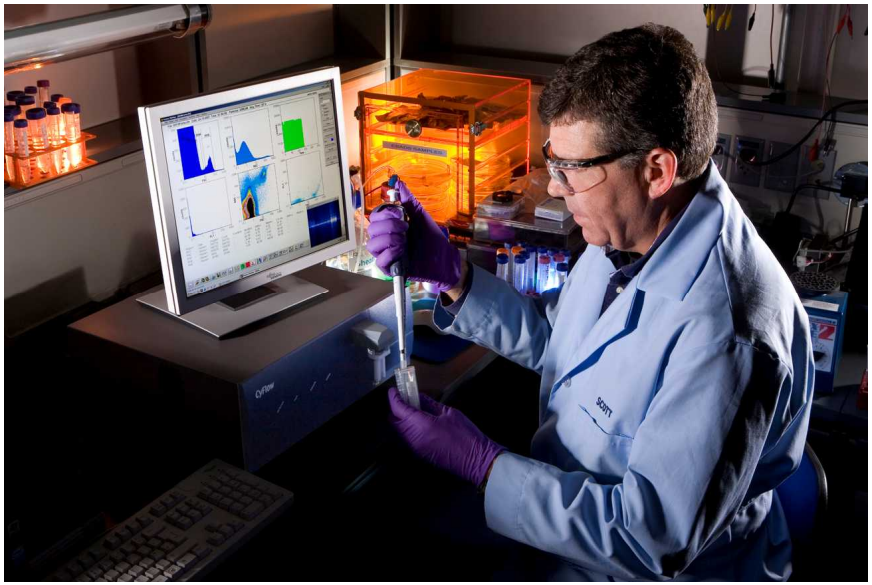
- **a**: acquired distribution
- **f**: fractional concentrations
- **P**^{T}: pure component distributions
- **e**: residual error

MATLAB built-in functions solve for **f** to minimize |**e**|.

73% Yr; <10^{-5} Bt, house dust (27% “other”)

Preliminary Design Review (PDR)

- Samples provided by ECBC
- 4 agent simulants (Bt, Yr, Ova, MS2) + house dust
  - Provided pure component distributions
  - Non-solubles: Bt, Yr, and house dust
  - Solubles: Ova, MS2 (detected by another method)



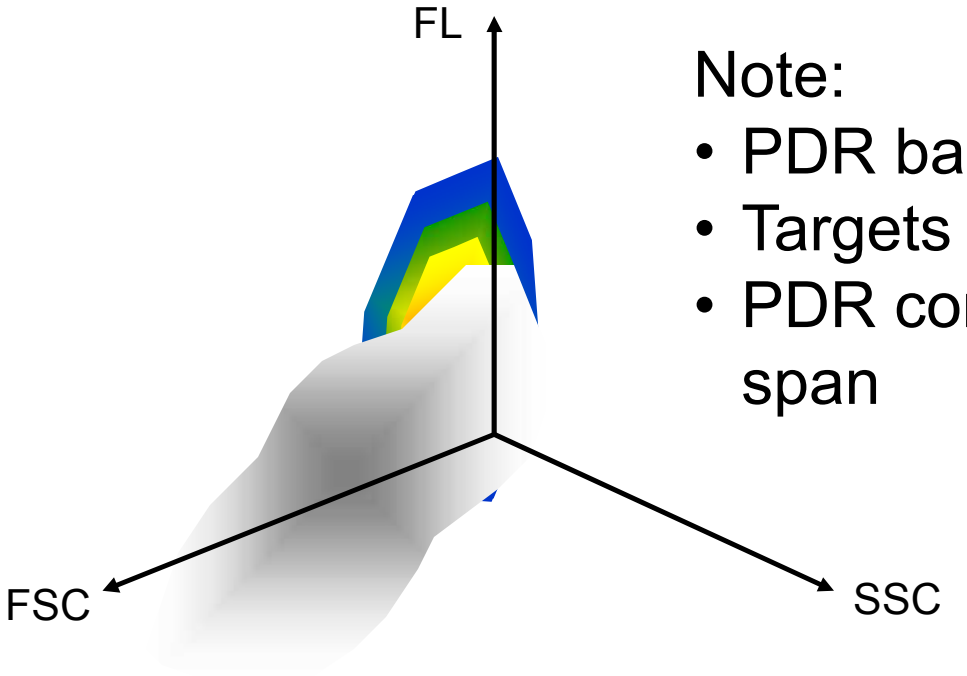
- 20 unknown mixtures of simulants and backgrounds
- Tested both native and stained

Results: 20 for 20 correct (8 for 8 non-solubles)

ECBC sample number	Cell/spore assignment	Soluble protein	Antigen	Background	
1	--	--	None	L	
2	--	Yes	Ovalbumin	H	
3	Yersinia	Yes	Yersinia	H	✓
4	--	--	None	H	
5	--	Yes	Ovalbumin	H	
6	--	--	None	H	
7	--	Yes	MS2	L	
8	--	Yes	MS2	L	
9	--	Yes	Ovalbumin	L	
10	Bt	--	Bt	L	✓
11	Yersinia	Yes	Yersinia	L	✓
12	--	Yes	Ovalbumin	L	
13	Bt	Yes	Bt	L	✓
14	--	Yes	MS2	H	
15	Bt	--	Bt	H	✓
16	Bt	--	Bt	H	✓
17	Yersinia	Yes	Yersinia	H	✓
18	Yersinia	Yes	Yersinia	L	✓
19	--	--	None	L	
20	--	Yes	MS2	H	

Success! But also potential challenges…

- *Specificity:* Unknown background distributions could overlap those of target species
- *Sensitivity:* Target distributions could vary with sample prep and shift with time

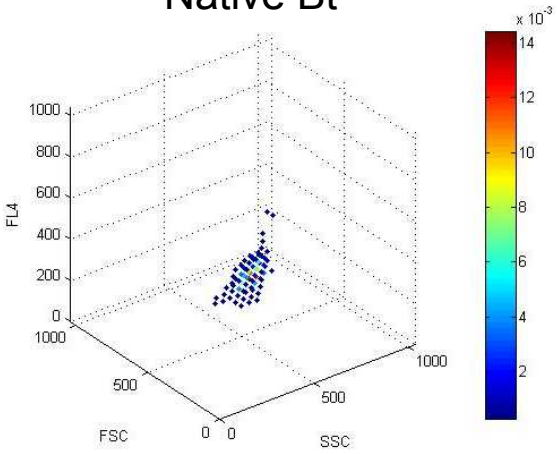


Note:
- PDR backgrounds ≠ targets
- Targets had identical prep
- PDR conducted over short time span

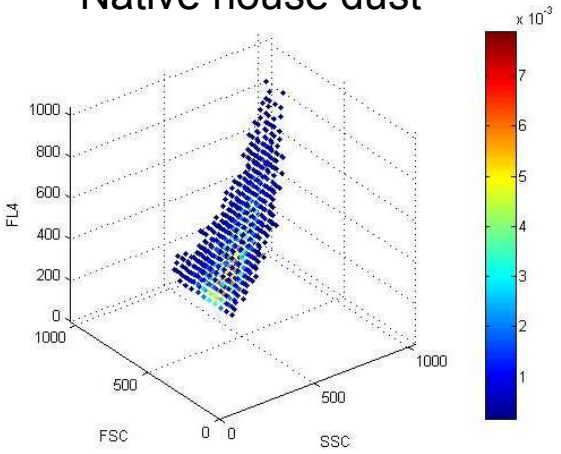
What if instrument did not “know” house dust?

- If unknown, house dust would be a problem
  - Respirable range → similar FSC, SSC
  - Contains biologicals (dust mites, dander, etc.) → similar FL

Native Bt



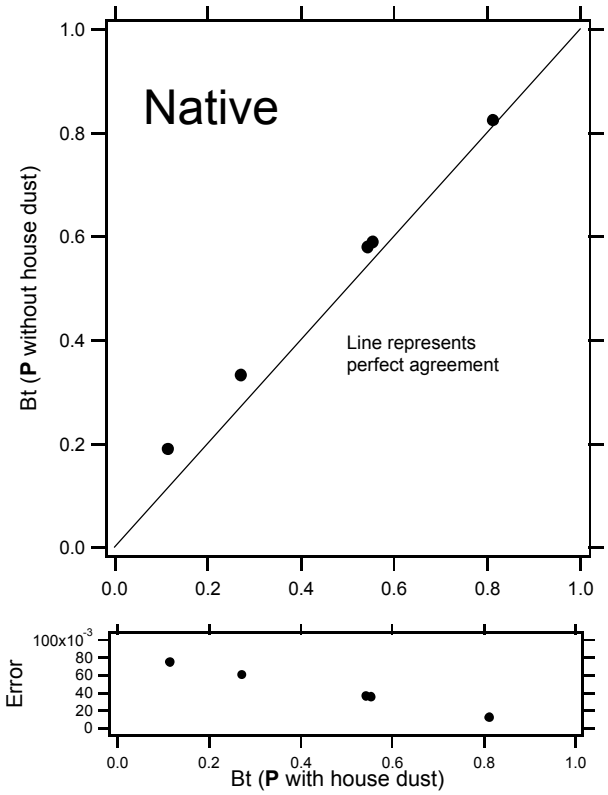
Native house dust



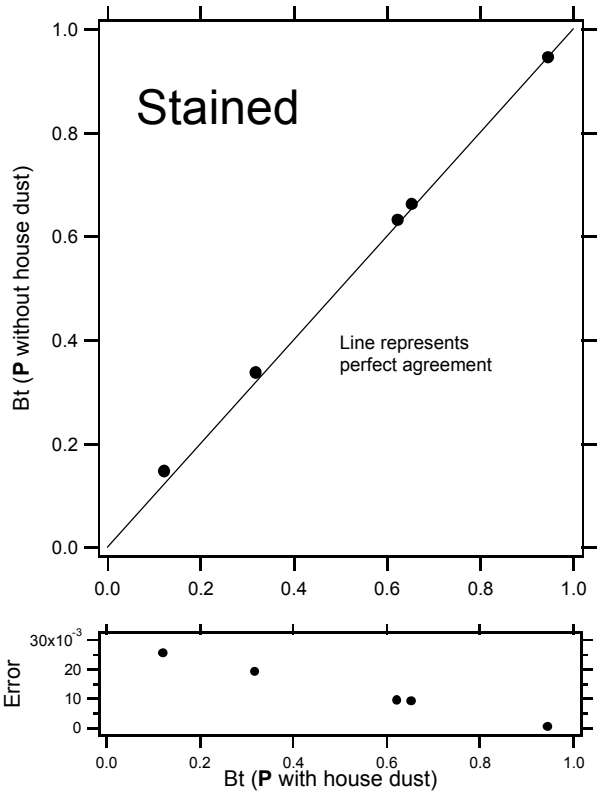
- Fit a series of Bt and house dust mixtures
  - Compare house dust known vs. unknown
  - Compare native vs. stained

Cross term of Bt vs. house dust

Fit distributions both with and without house dust in **P**



10% house dust mis-taken as Bt.



3% house dust mis-taken as Bt.

Qualifiers

- 3% overlap when stained?
- Assumes house dust is not included in pure components
  - Local backgrounds could be sampled and included in pure component database
- Properties of house dust are most difficult to mitigate
- Rigorous test of fitting routine

Conclusions and future work

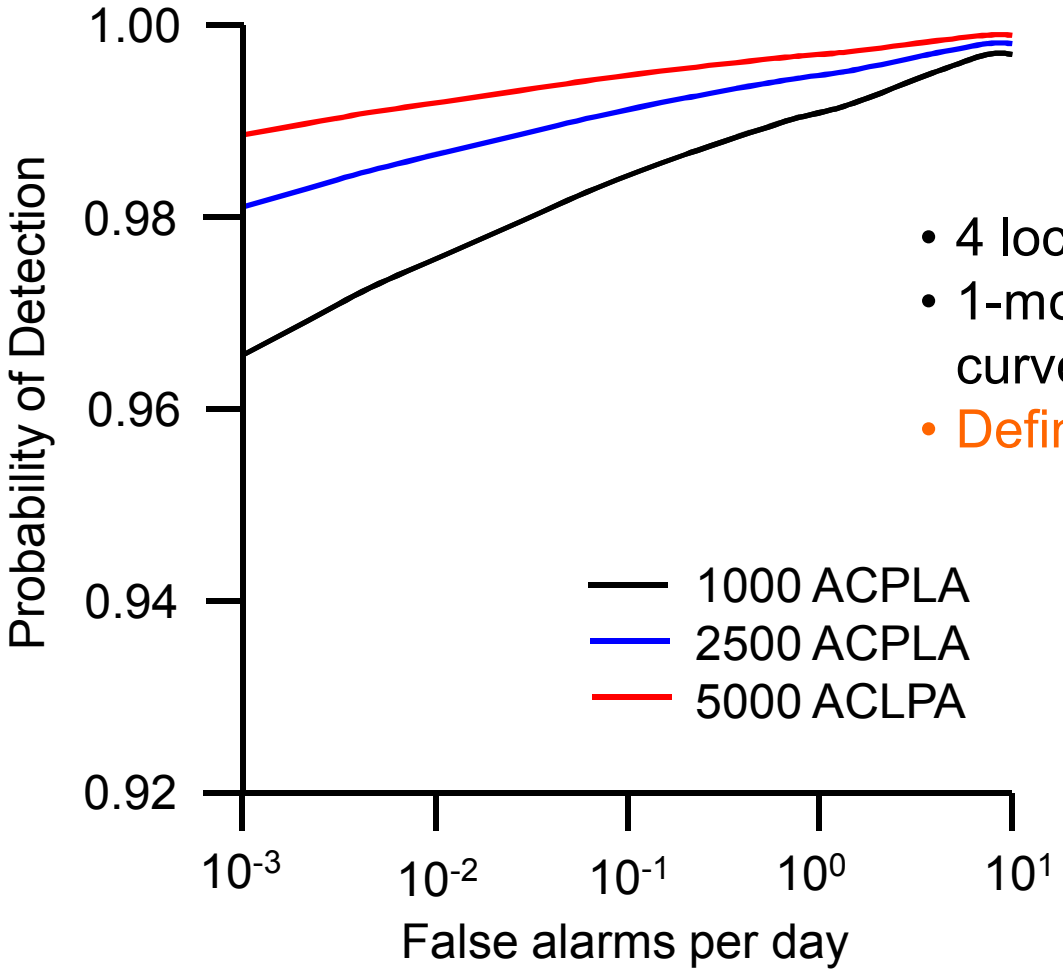
Conclusions:

- Least-squared linear unmixing used to identify components in flow cytometer data
- Staining improves discrimination of Bt vs. house dust (3% vs. 10%)

Future Work:

- Test on field-collected backgrounds → *specificity*
- Long-term evaluation of target distributions → *sensitivity*
- Develop receiver operator characteristic (ROC) curve of detection performance
  - *False alarm rate* vs. *detection limit* vs. *probability of detection*

Field sampling underway



- 4 locations, 1 month each
- 1-month site-specific ROC curves for bio-detector
- Define shift in ROC curves