Are we there yet?

How many particles do you need for Image Analysis

Jeffrey Bodycomb, Ph.D.
HORIBA Scientific
www.horiba.com/us/particle

Outline

- Intro to Image Analysis
- Effect of Number of Particles
- Experiments to Determine the Number of Particles

Why Image Analysis?

- Need shape information, for example due to importance of powder flow
- Verify/Supplement diffraction results
- Replace sieves for size distribution analysis

These may have the same size (cross section), but behave very differently.

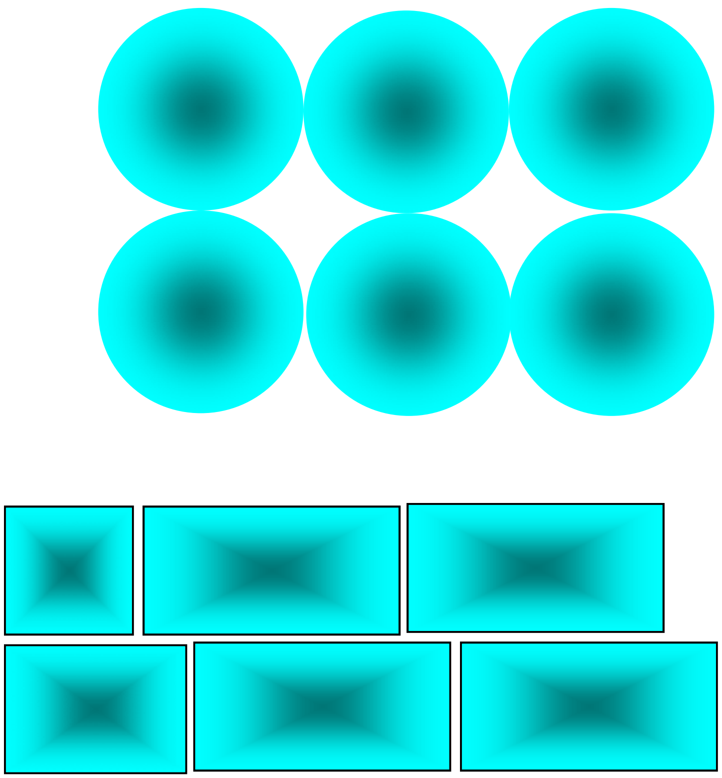
What information can we get?

- Particle size distribution
- Shape characteristics
- Differences
  between product samples
  (uncoated vs coated)
- Agreement with quality
  specifications
  (good vs bad)

- Manufacturing performance (effect of grinding on size distribution and shape)

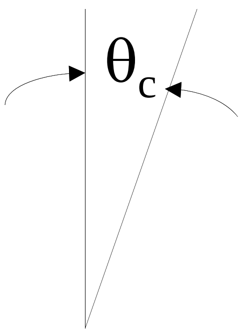
Why Image Analysis

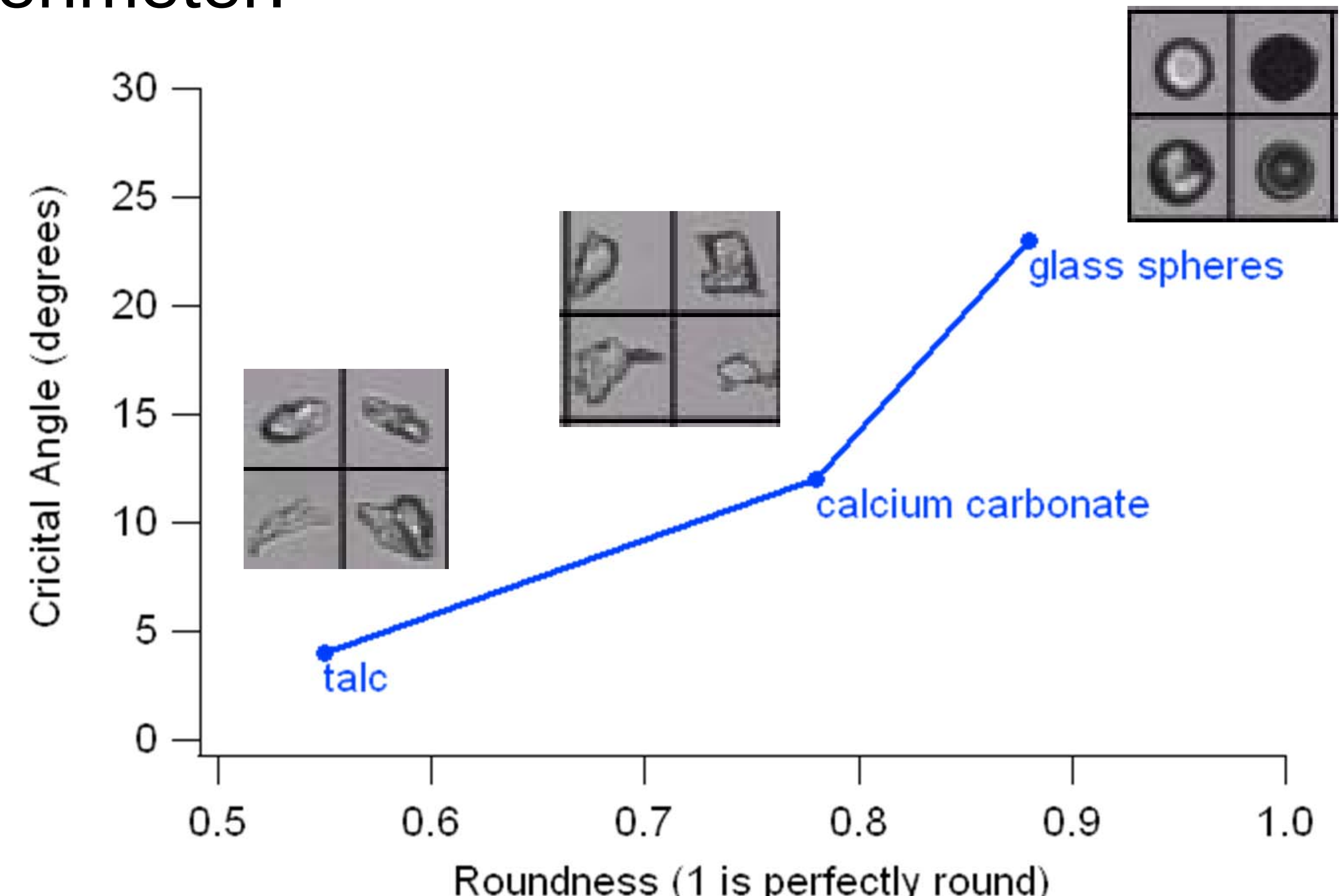
Need shape information for evaluating packing and flow.



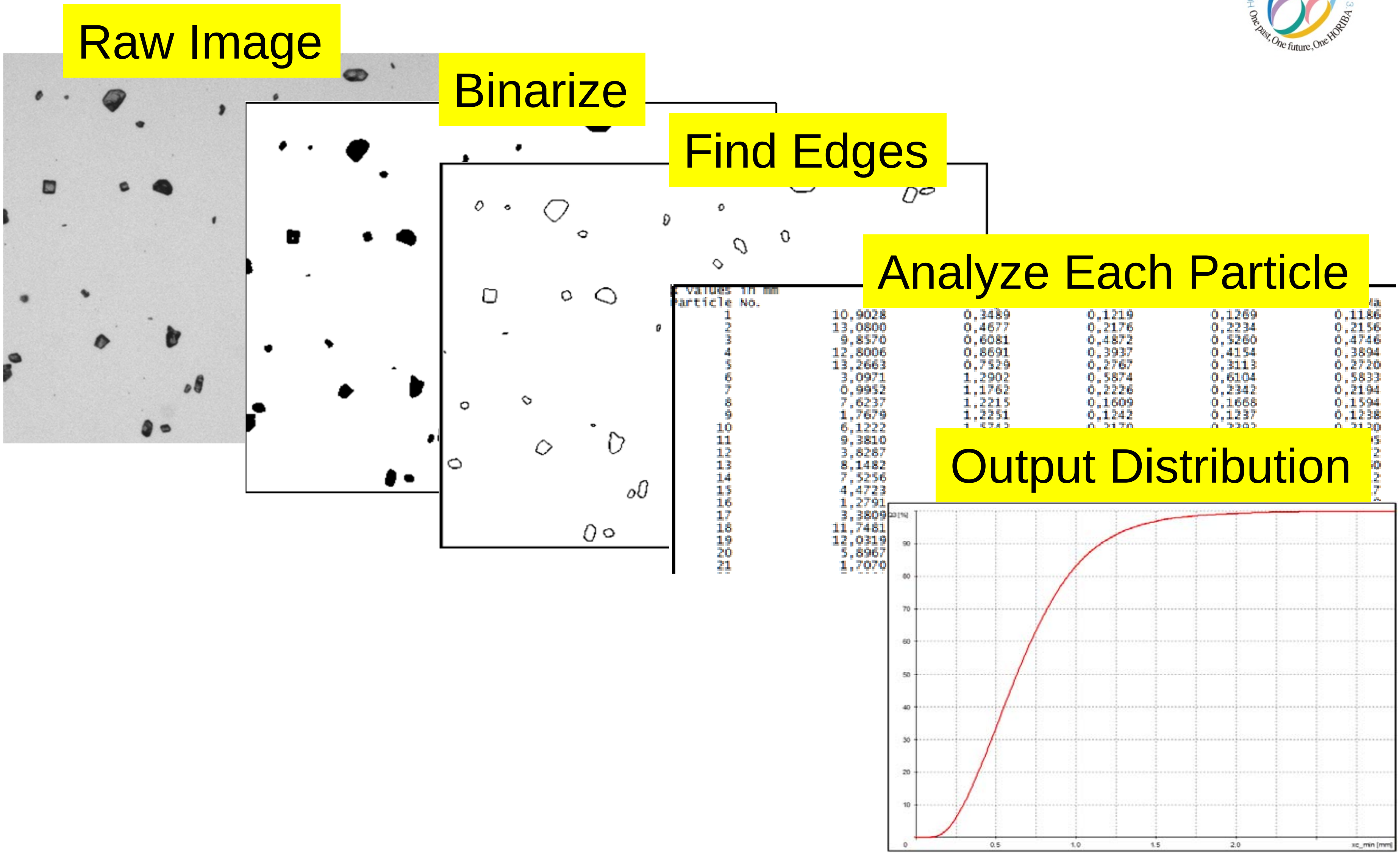
Effect of Shape on Flow

- Yes, I assumed density doesn't matter.
- Roundness is a measure based on particle perimeter.





Data Evaluation

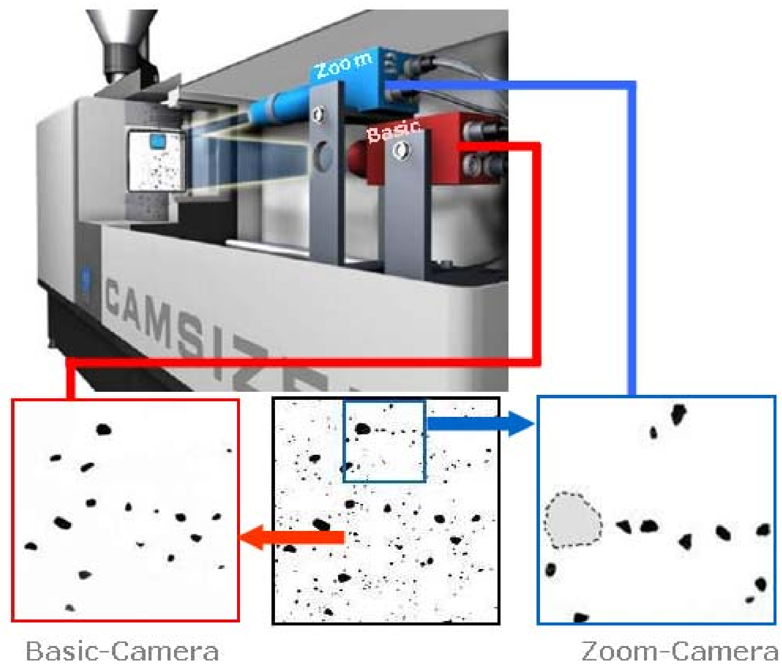


Two Approaches

Static:
particles fixed on slide,
stage moves slide

0.5 – 1000 um
2000 um w/1.25 objective

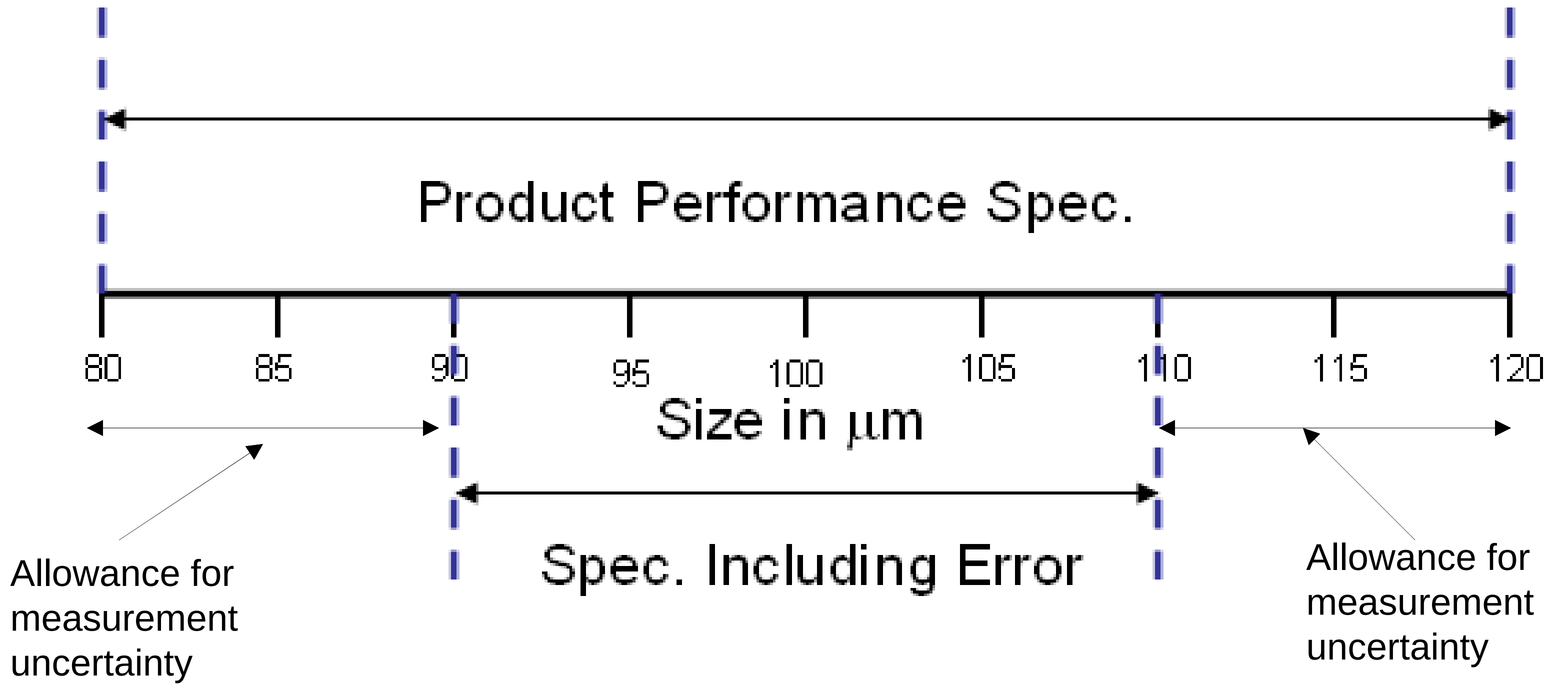
Dynamic:
particles flow past camera



1 – 3000 um

Specification with Measurement Error

Must tighten internal spec by lab error %
Then product always within performance specification

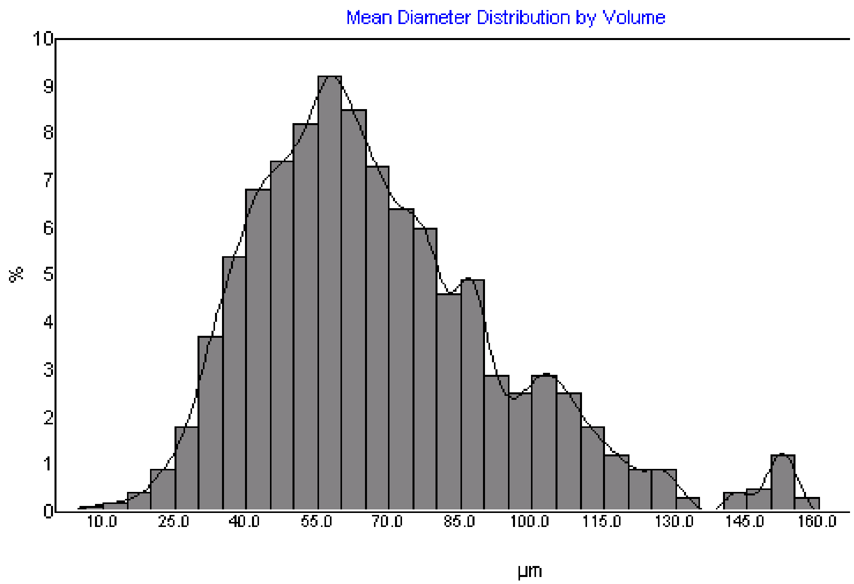


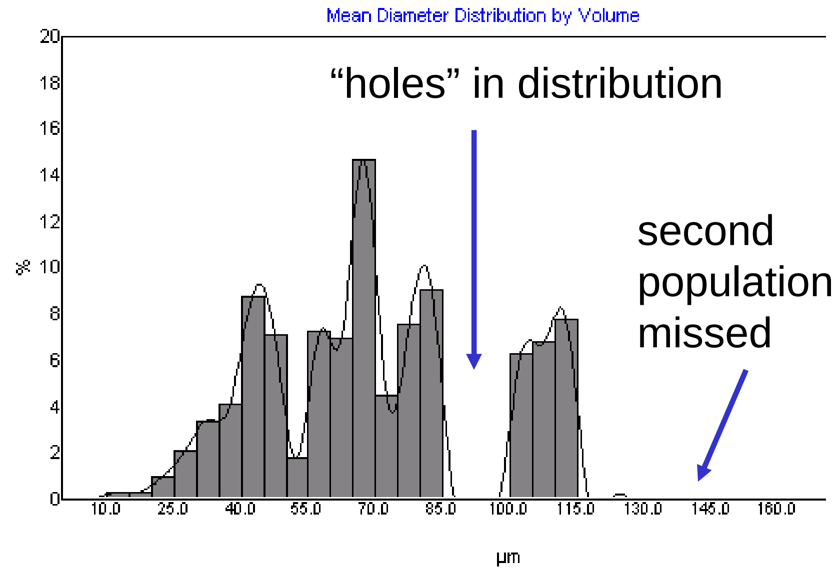
http://www.spcpress.com/pdf/Manufacturing_Specification.pdf, By David Wheeler

Effect of Number of Particles Counted

20,000 particles

200 particles





Statistics - Silica Catalyst

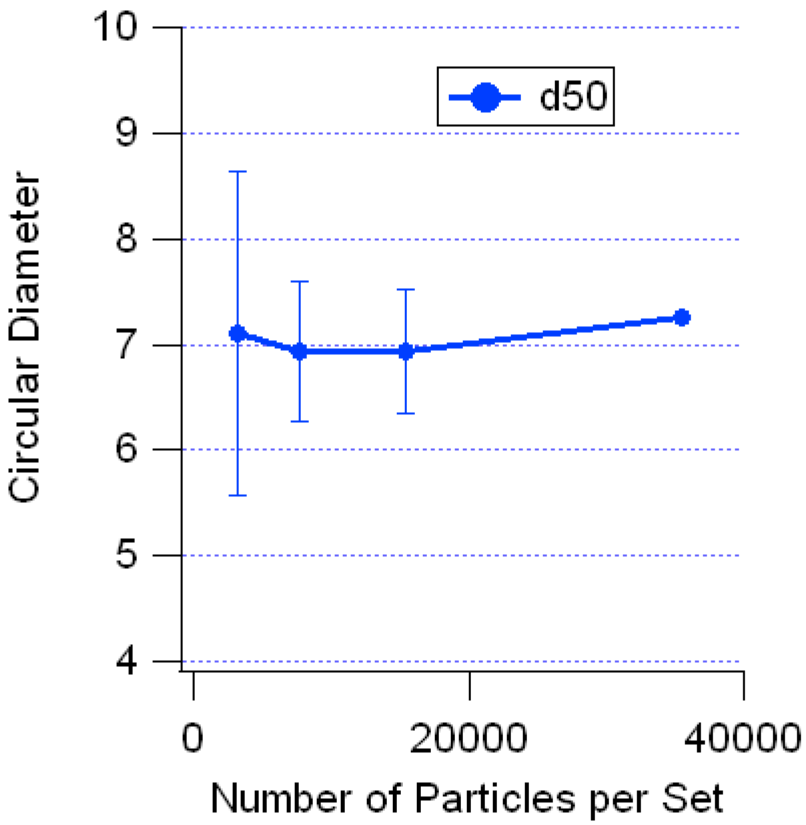
Mean:	27.477	D[4,3]:	68.088	Min, Ma
STD:	22.129	D[3,2]:	57.598	Confider
RSD:	80.539%	D[v,0.1]:	37.870	Confider
D[n,0.1]:	4.697	D[v,0.5]:	63.467	Optics:
Median:	23.824	D[v,0.9]:	104.351	Scanne
D[n,0.9]:	57.814			

Statistics - Silica Catalyst 200

Mean:	31.805	D[4,3]:	69.060	Min, M
STD:	21.927	D[3,2]:	58.947	Confid
RSD:	68.942%	D[v,0.1]:	39.418	Confid
D[n,0.1]:	7.021	D[v,0.5]:	68.941	Optics
Median:	29.488	D[v,0.9]:	105.901	Scanr
D[n,0.9]:	63.734			

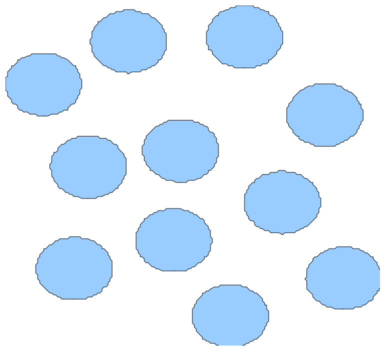
But d10, d50 &d90 may appear similar

Divide Large Data Set into Smaller Sets

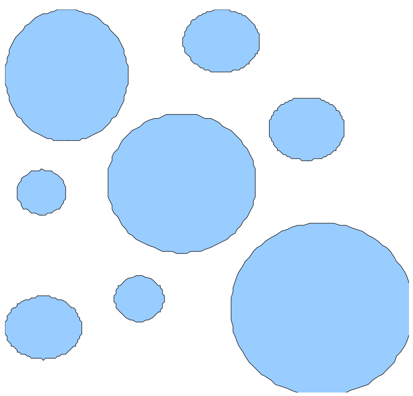


- Error bars are one standard deviation from repeated measurements of the same number of particles from different parts of the sample.
- The error bars get smaller as you evaluate more particles.

How Many Particles?



All of the particles are the same size. I only need to measure one to find the average size for distribution.



Here, particles are different sizes and none are the average size. I need to measure many particles to find the average size.

How many particles do I need to measure?

Two Kinds of Standard Deviation!

- Sample standard deviation is a property of the sample. It is the width of the size distribution.
- Measurement standard deviation is a result of the measurement and is affected by the sample standard deviation.

<USP> 776: Mean

$$u = z\left(\frac{s}{\sqrt{n}}\right) \qquad n = \left(\frac{sz}{u}\right)^2$$

u=uncertainty
z=confidence coefficient (often ~2, see a statistics book)
s=standard deviation of distribution (width)
n=number of particles measured

Example: for uncertainty of ± 5 µm with st dev= 20 µm
Must measure 61 particles

$$n = \left(\frac{sz}{u}\right)^2 = \left(\frac{(20)(1.96)}{5}\right)^2 \approx 62$$

Implies normal particle size distribution, greater than 30 particles, and known standard deviation.

<USP> 776: Standard Deviation

$$\sqrt{s^2\left(\frac{n}{\chi_a^2}\right)} < \sigma < \sqrt{s^2\left(\frac{n}{\chi_b^2}\right)}$$

χ=Moment of chi squared distribution (see a statistics book)
s=estimated standard deviation of distribution (width)
n=number of particles measured

These limits are asymmetric around the standard deviation.

Implies normal particle size distribution, greater than 30 particles, and known standard deviation.

How to think about this

- The particle measurement problem is just like the uncertainty estimates from a series of experiments (say repeated measures of salinity).

- The difficulty is that particle distributions are not always narrow or normal.

ISO 13322-1-2004 Static Image Analysis

- Tables do not match experience

Number of particles required n*, δ = 0.05, p = 0.95				
δ	GSD	n*(MMD)	n*(Sauter)	n*(MVD)
0.05	1.1	585	389	131
	1.15	1460	934	294
	1.2	2939	1808	528
	1.25	5223	3103	843
	1.3	8526	4920	1247
	1.35	13059	7355	1750
	1.4	19026	10504	2363
	1.45	26617	14457	3093
	1.5	36007	19295	3956
	1.55	47358	25093	4952
	1.6	60811	31919	6092

Number of particles required n*, δ = 0.1, p = 0.95				
δ	GSD	n*(MMD)	n*(Sauter)	n*(MVD)
0.1	1.1	146	97	33
	1.15	365	233	73
	1.2	735	452	132
	1.25	1306	776	211
	1.3	2131	1230	312
	1.35	3265	1839	438
	1.4	4756	2626	591
	1.45	6654	3614	774
	1.5	9002	4824	989
	1.55	11839	6273	1238
	1.6	15203	7980	1523

What About Real Life?

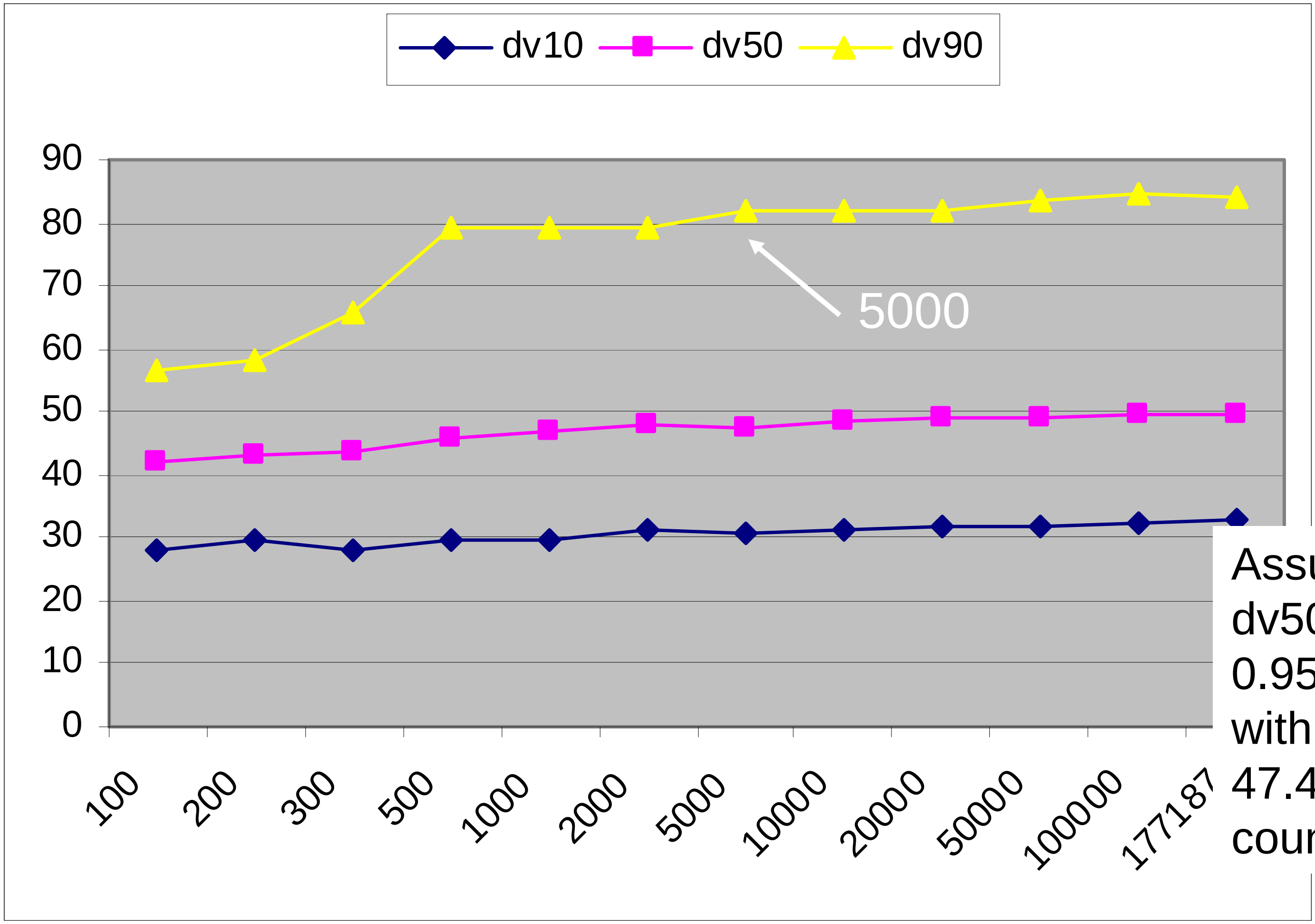
- Observations from Customer Samples

USP is too optimistic

ISO is too conservative

- Build real-life data set for checking models

More particles for more accuracy.



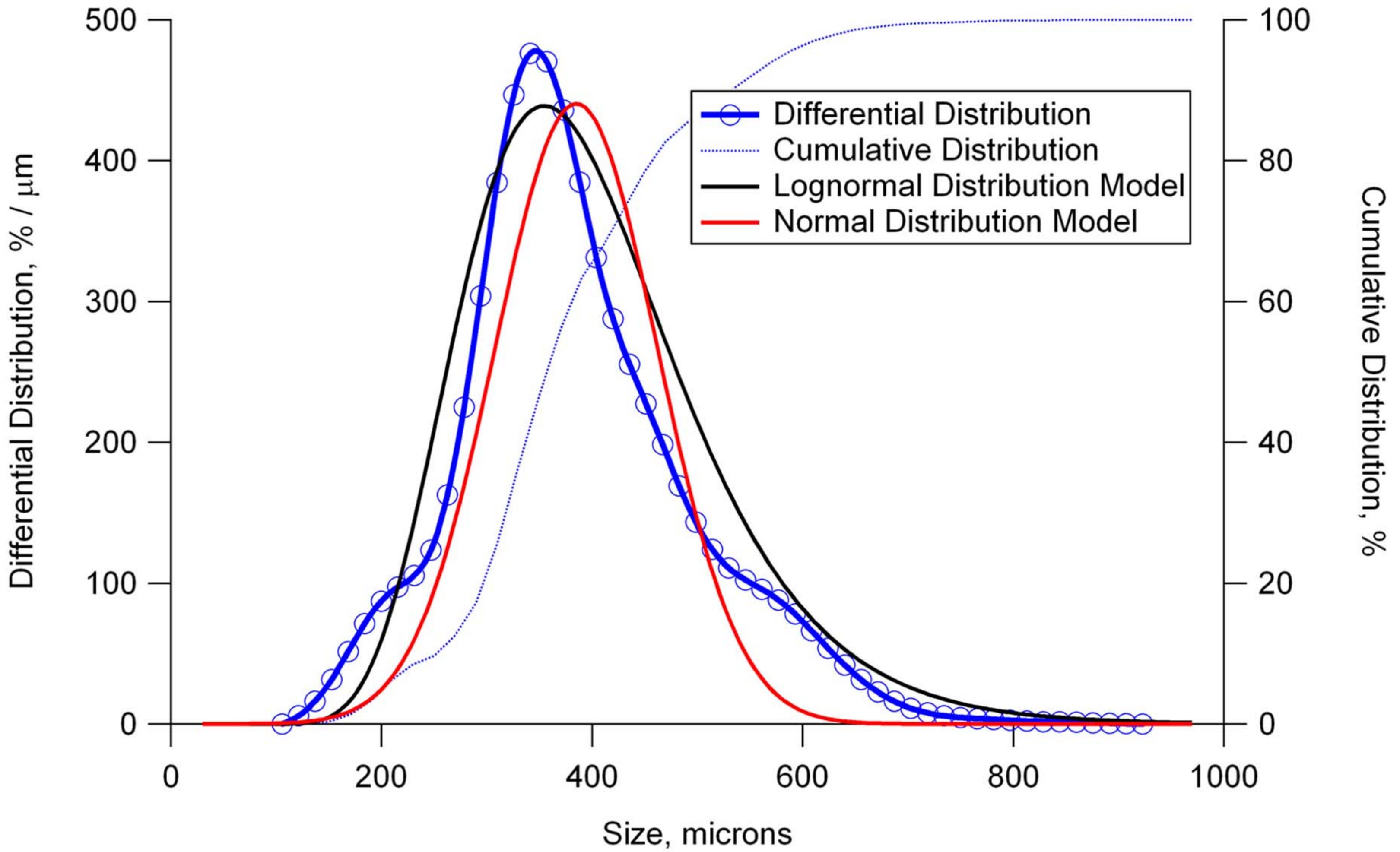
Assume 49.833 is “correct” dv50,
0.95 x 49.833 = 47.34 is within 95%,
47.463 achieved by 5000 counts!

Use this to control precision of your data (and not spend extra time on precision you don’t need.

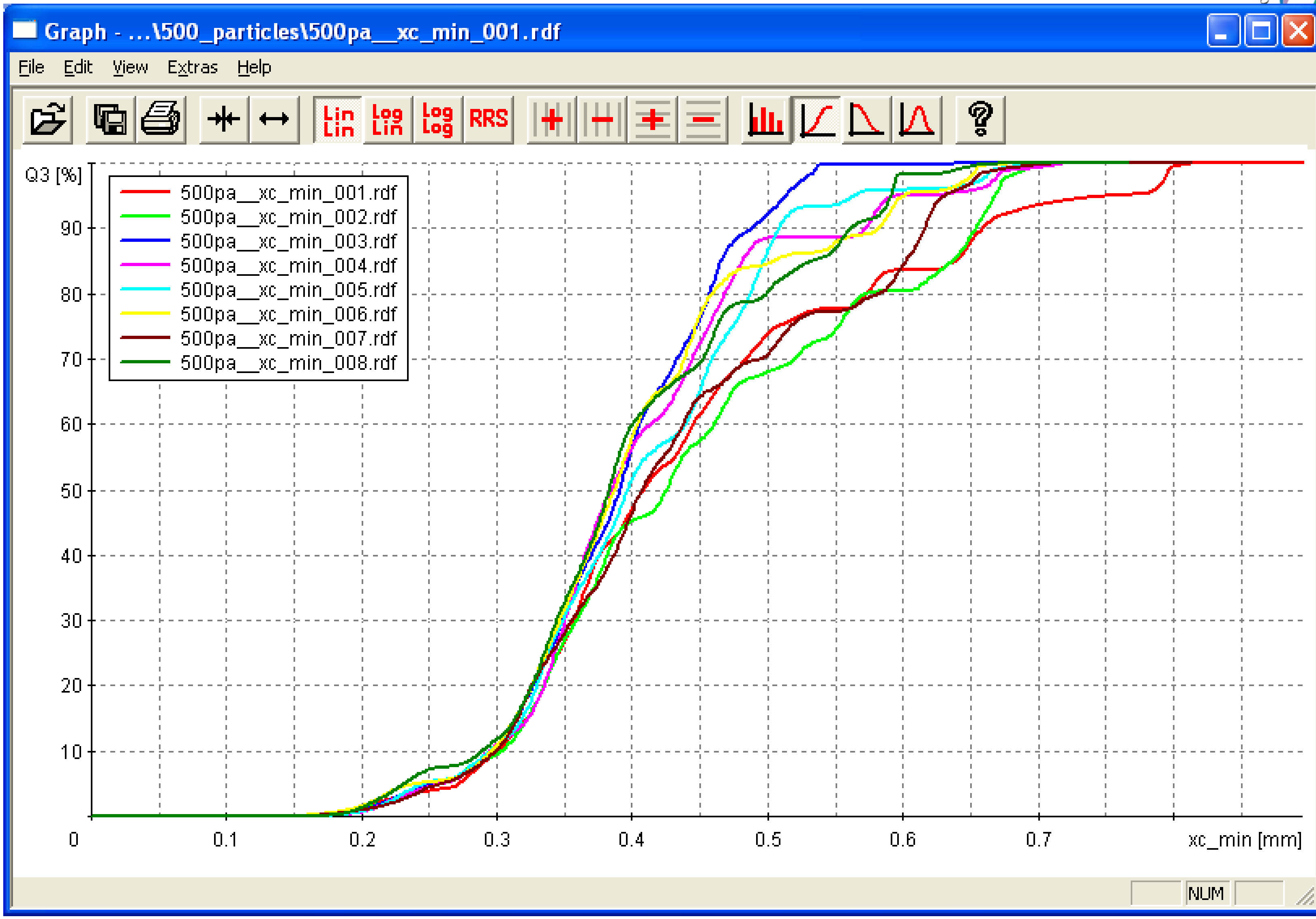
Experimental Measurement of Effect of Counting More Particles

- Use glass bead sample with somewhat broad size distribution.
- Split sample into smaller samples with varying numbers of particles. Recover and reuse samples for the next measurement.
- Run 8 trials with different numbers of particles and looked at variation in results from each trial.
- Measure all particles for real answer.

Real Data vs. Models



Data from 600 particles

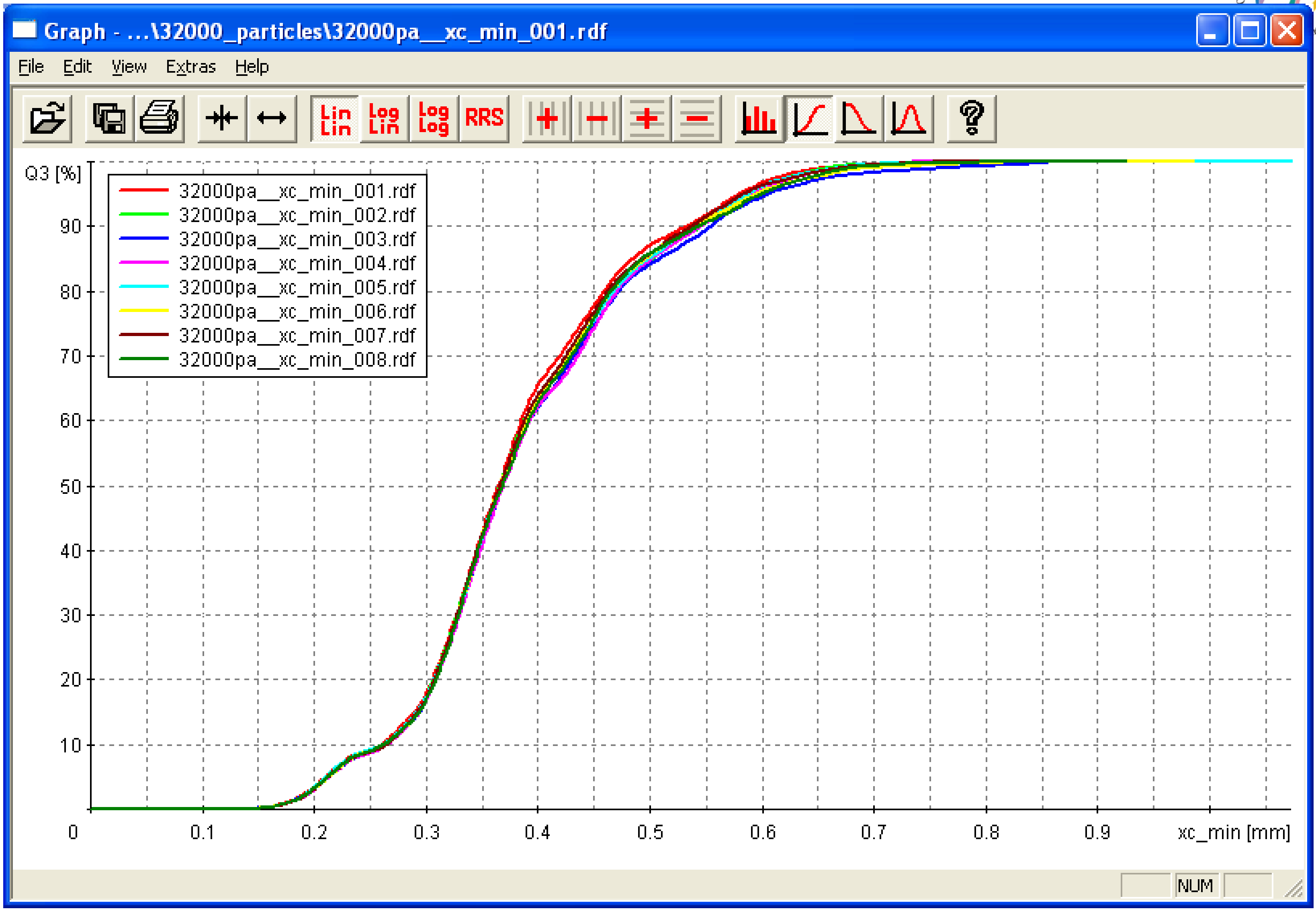


Example Results for ~500 Particles

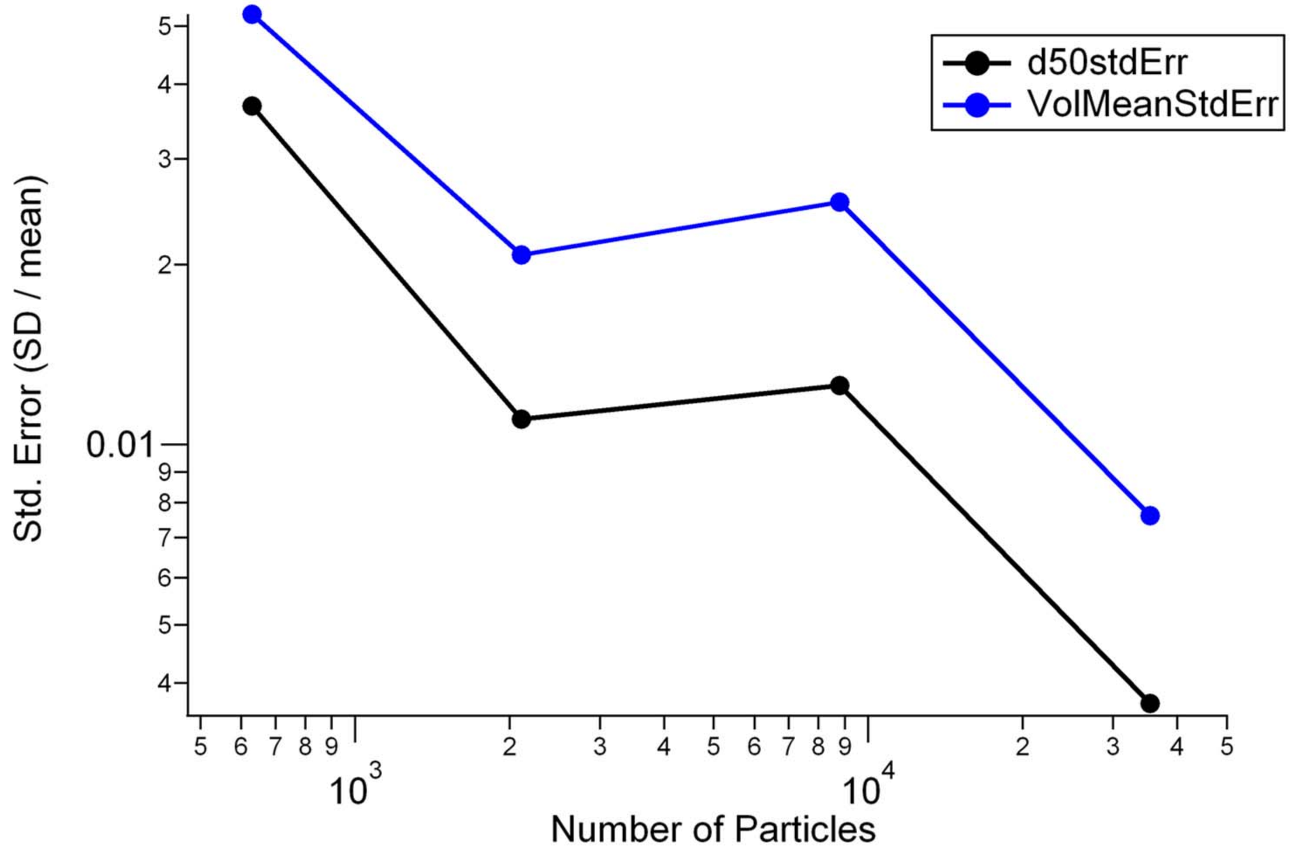
	Number of Particles	Volume Mean Diam. (µm)
500p Trial #1	516	445
500p Trial #2	612	446
500p Trial #3	669	389
500p Trial #4	603	405
500p Trial #5	660	407
500p Trial #6	629	402
500p Trial #7	699	432
500p Trial #8	652	402
Average	630.00	416.00
Std. Dev.	55.70	21.78

Repeat for increasing numbers of particles.

Data from 32000 particles

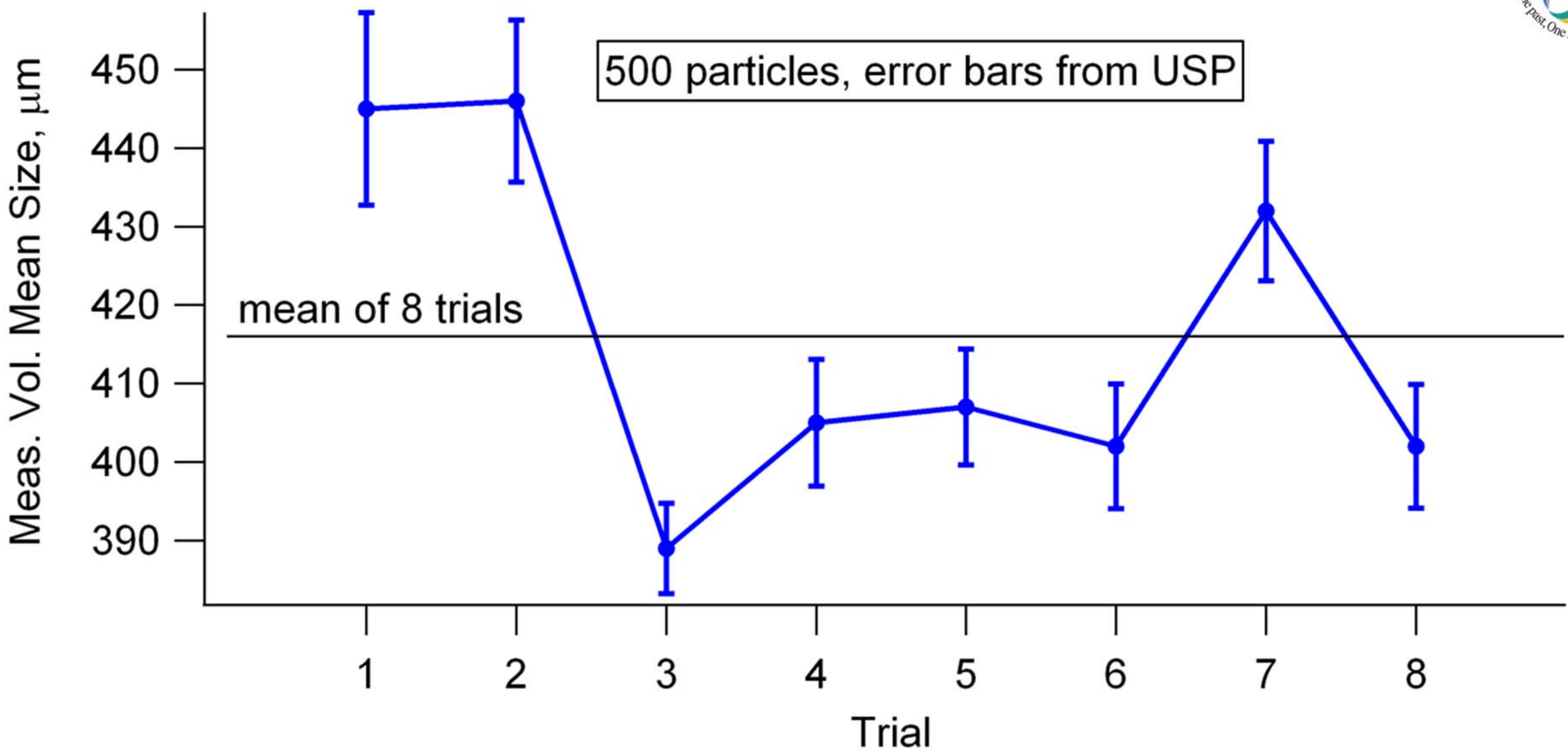


Dv50, Volume Mean



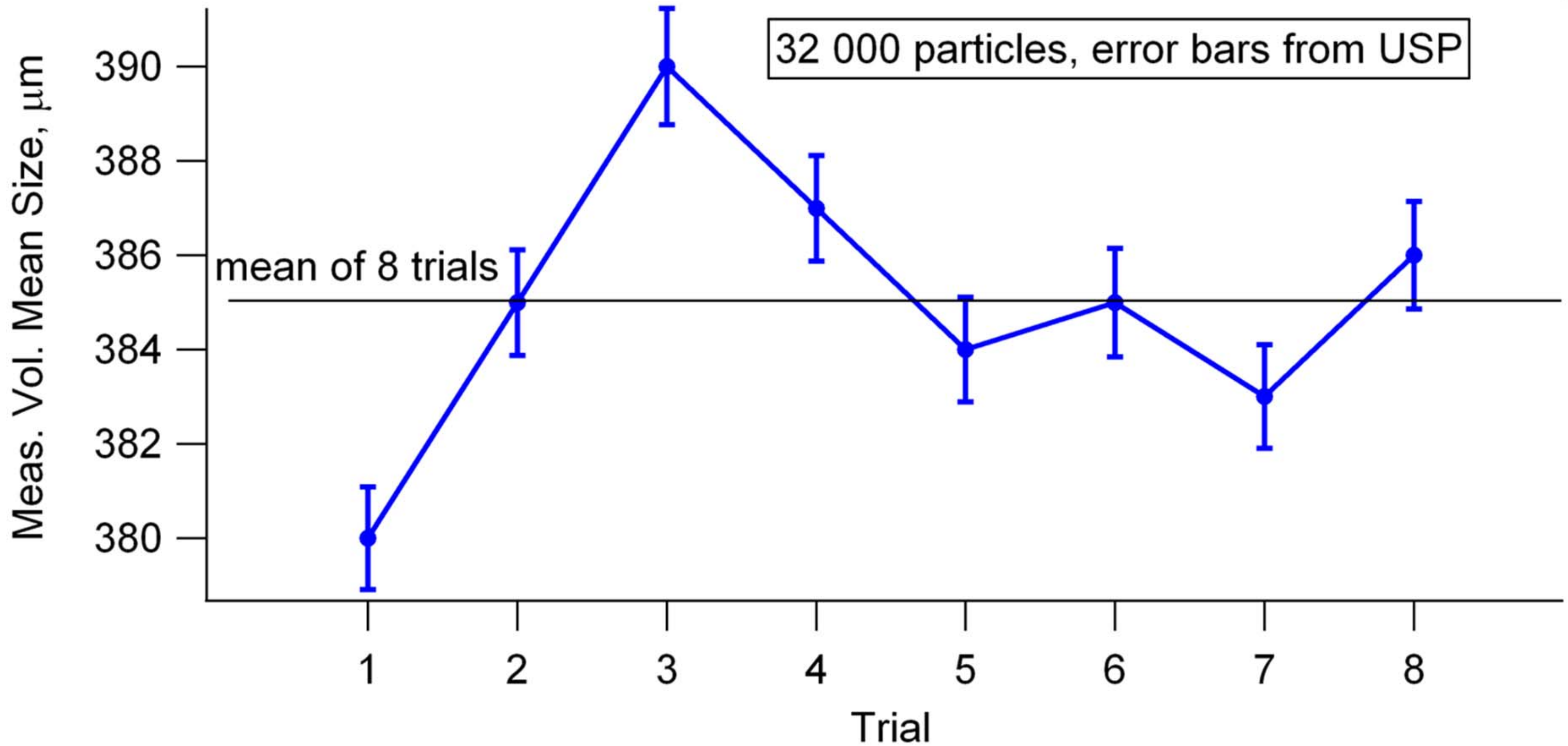
- Data here shows that reproducibility of d50 is <u>always</u> better than the reproducibility of mean.

USP confidence intervals



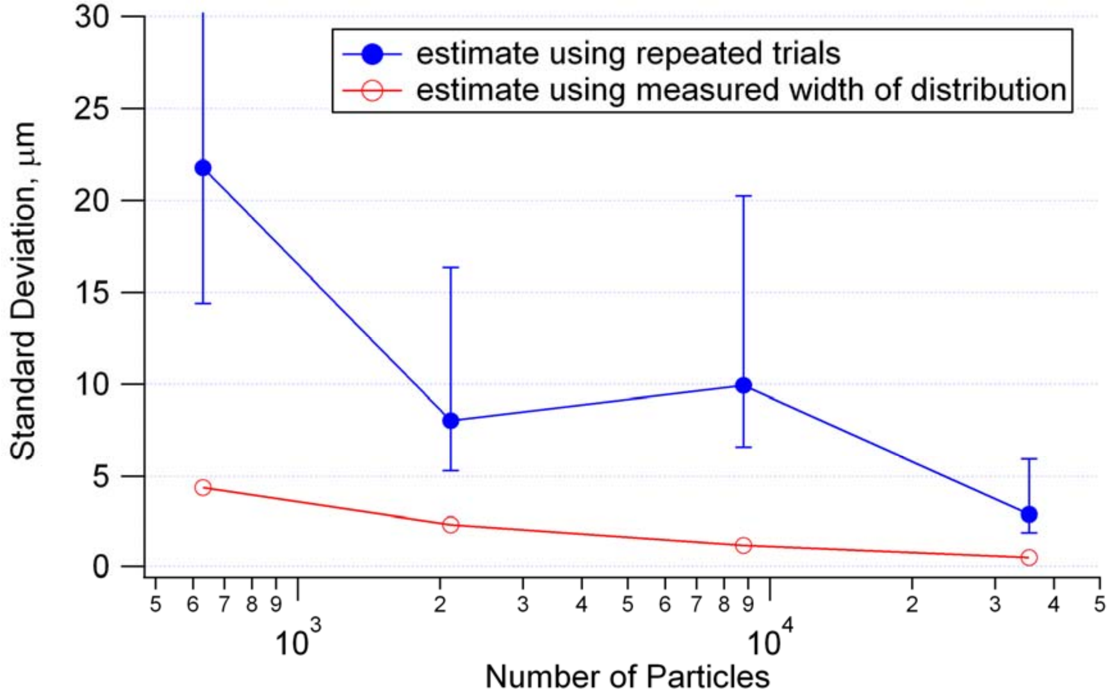
- For 500 particles, the USP confidence interval derived from measured volume mean and standard deviation never included the volume mean of the entire population.

32 000 particles



- For 32 000 particles, the USP confidence interval derived from measured volume mean and standard deviation included the volume mean of the entire population one half of the time.
- USP model improves with more particles

Look at Standard Deviations



- USP method estimates standard deviation of measurement as $s_{pop}/(n^{1/2})$
- We made repeated trials to find true standard deviation.
- Repeated trials vary more than expected.

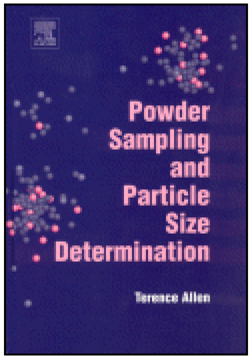
Excess Variance (Standard Deviation)

- Postulate an "excess" SD_{ex}
- $SD_{Tot}^2 = SD_{dist}^2 + SD_{ex}^2$
- Excess variance is ~ 2- 20 microns or ~1 to 5%
- Right in range for splitting/sampling effects.

- Need to account for splitting/sampling effects in error estimates.

Sampling Technique Error Levels

Standard Deviation (σ) in % Sugar-Sand Mixture



SCOOP SAMPLING	**6.31**
TABLE SAMPLING	**2.11**
CHUTE RIFFLER	**1.10**
SPINNING RIFFLER	**0.27**

Density of sand and sugar respectively 2.65 and 1.64 g/ml

Reference: Allen, T. and Khan, A.A. (1934), Chem Eng, 238, CE 108-112

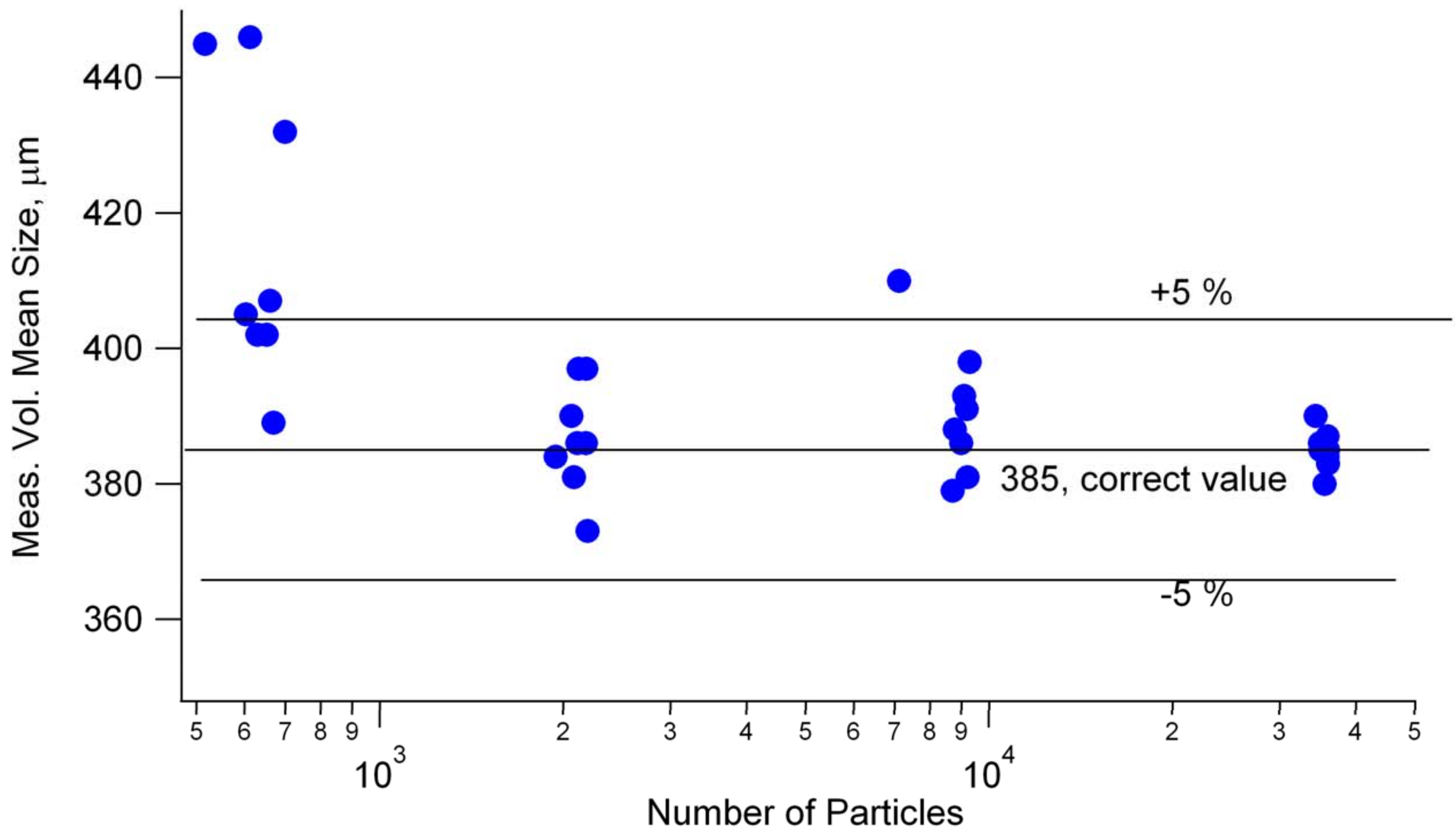
Method	Relative Standard Deviation (%)
Cone & Quartering	6.81
Scoop Sampling	5.14
Table Sampling	2.09
Chute Riffling	1.01
Spin Riffling	0.125

Sampling and Splitting

- Hugely important for this and all particle analysis.
- Learn to love your riffler
- Webinar at http://www.horiba.com/us/en/scientific/products/particle-characterization/download-center/webinars/:

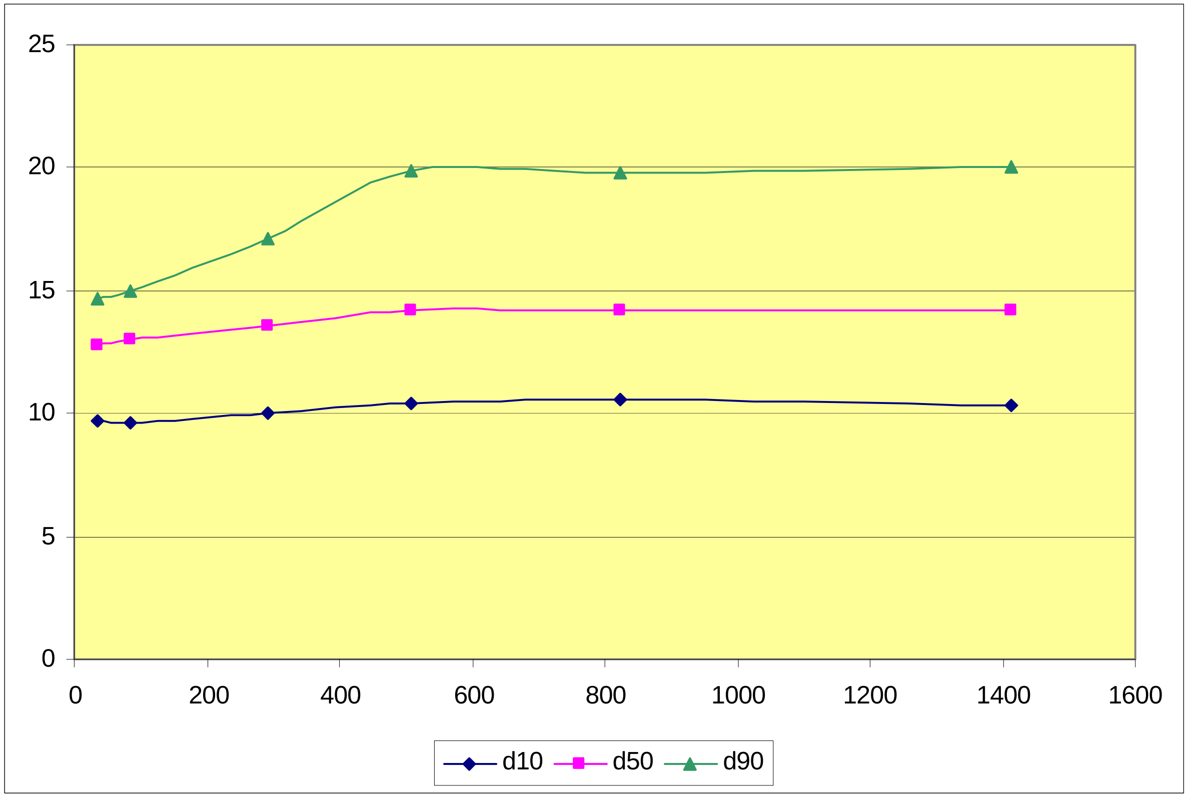
Sampling: A Critical Factor in Particle Size Analysis, TR011

Count until values stabilize



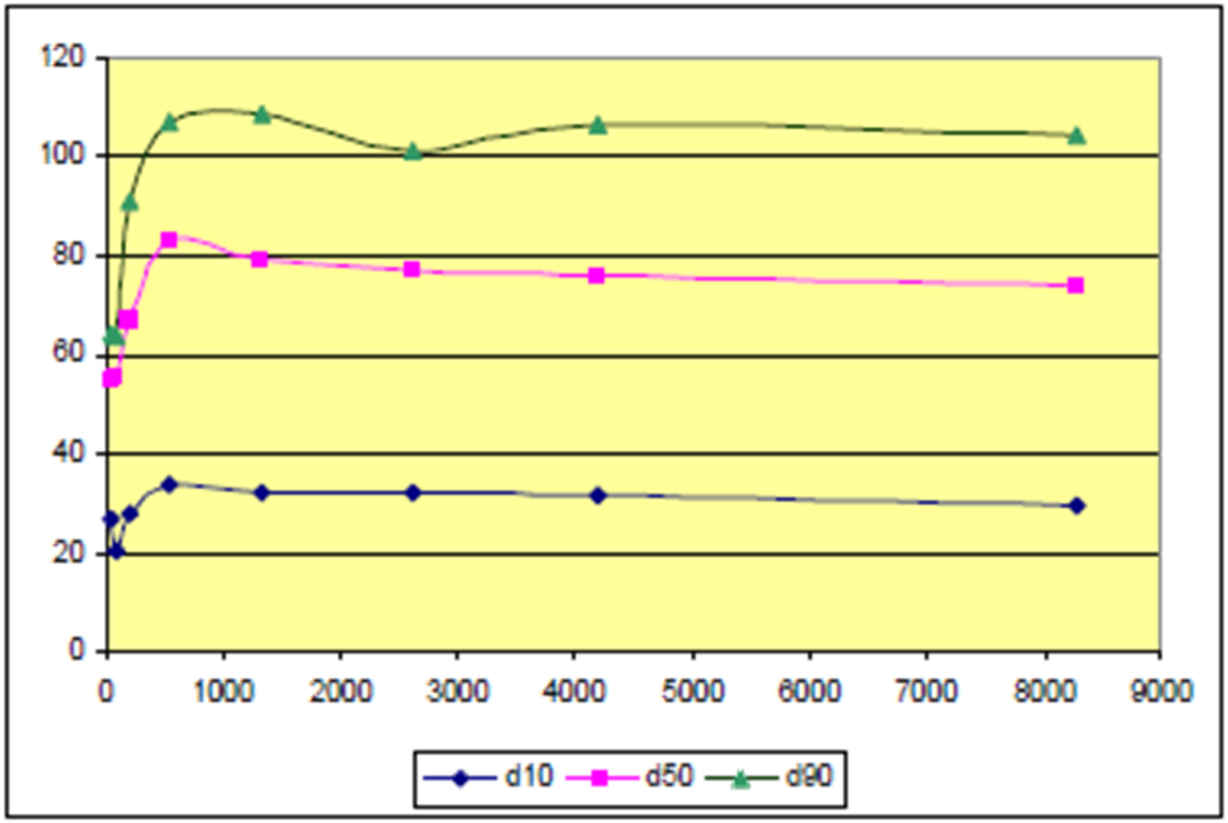
Number of Particles Measured

Count	d10	d50	d90
36	9.7	12.8	14.7
84	9.6	13	15
292	10	13.6	17.1
509	10.4	14.2	19.9
824	10.6	14.2	19.8
1412	10.3	14.2	20



Number of Particles Measured

Count	d10	d50	d90
47	26.7	54.8	64.5
82	20.3	55.8	64
207	27.8	66.8	91.3
534	33.5	82.8	107.3
1336	32.1	79.2	108.6
2627	32.3	77	101.4
4203	31.5	76.2	106.7
8270	29.7	73.8	104.2



The perfect splitter

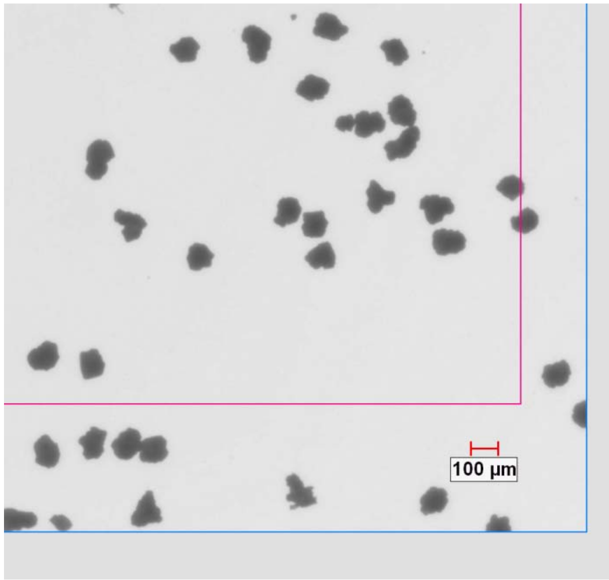
Take alumina sample and sieve it to obtain narrow size fraction.

Analyze 3865 particles in PSA300.

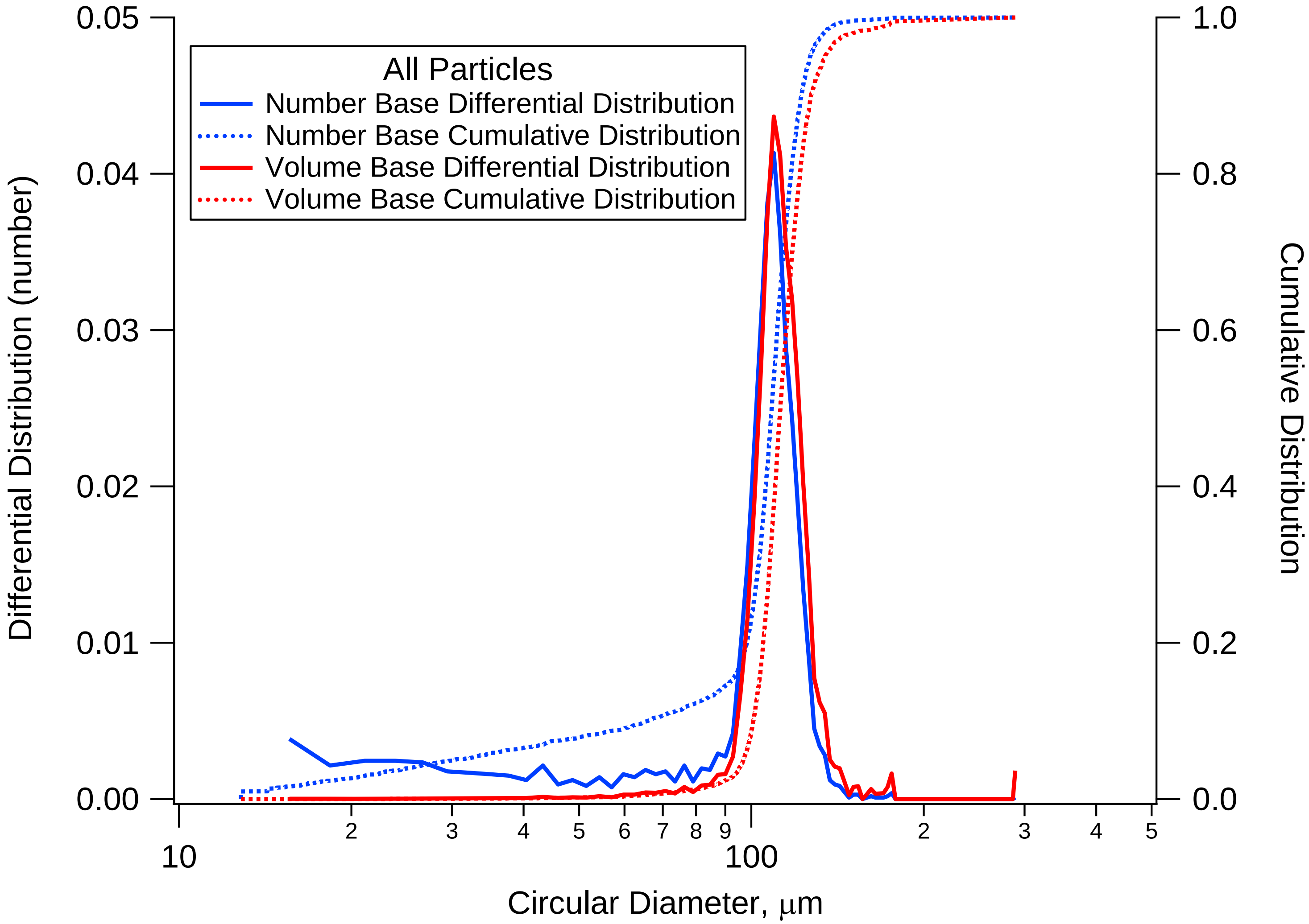
Treat 3865 particles as entire population.

“Split” by selecting particles at random in computer.

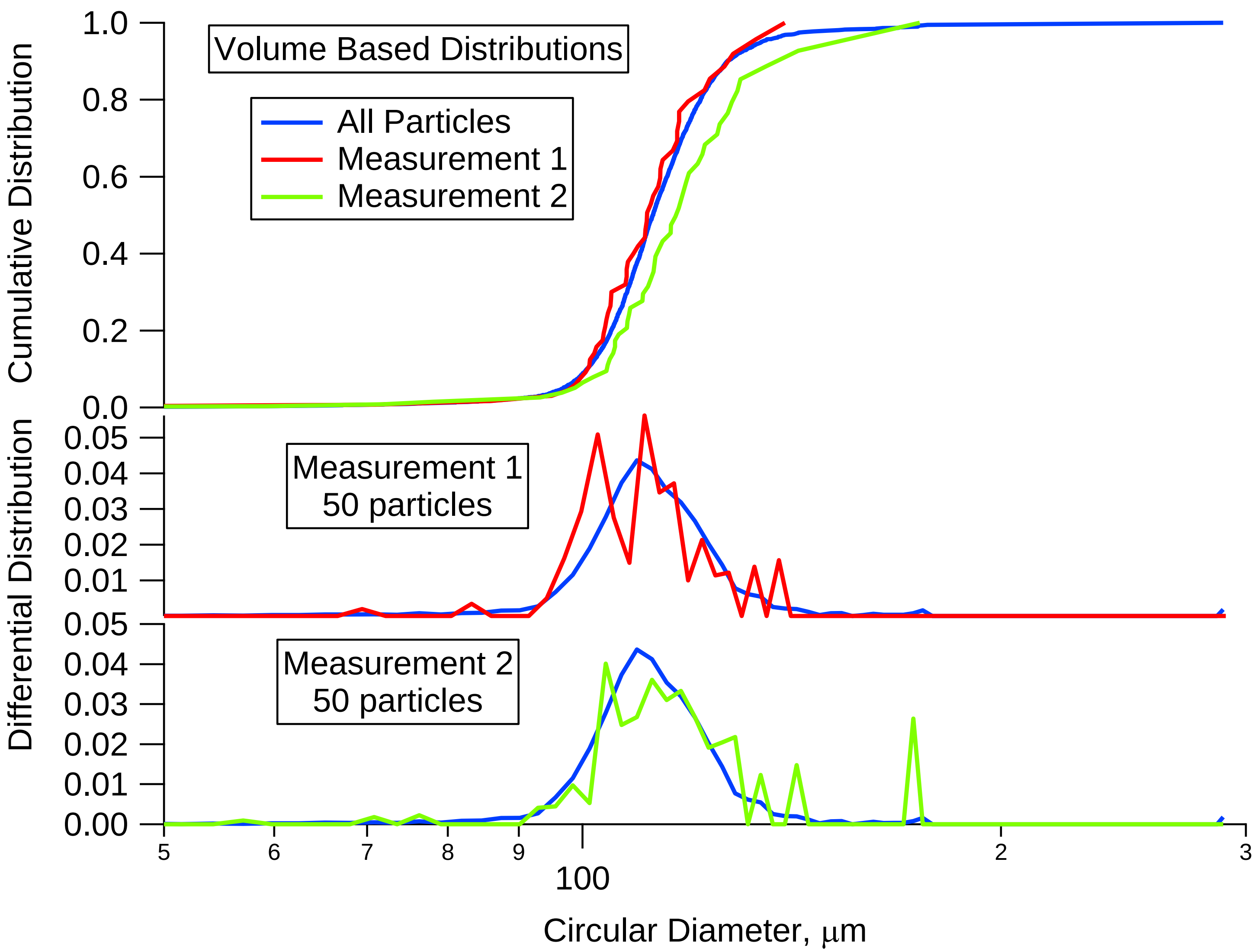
Splitting errors are eliminated, but data is less realistic.



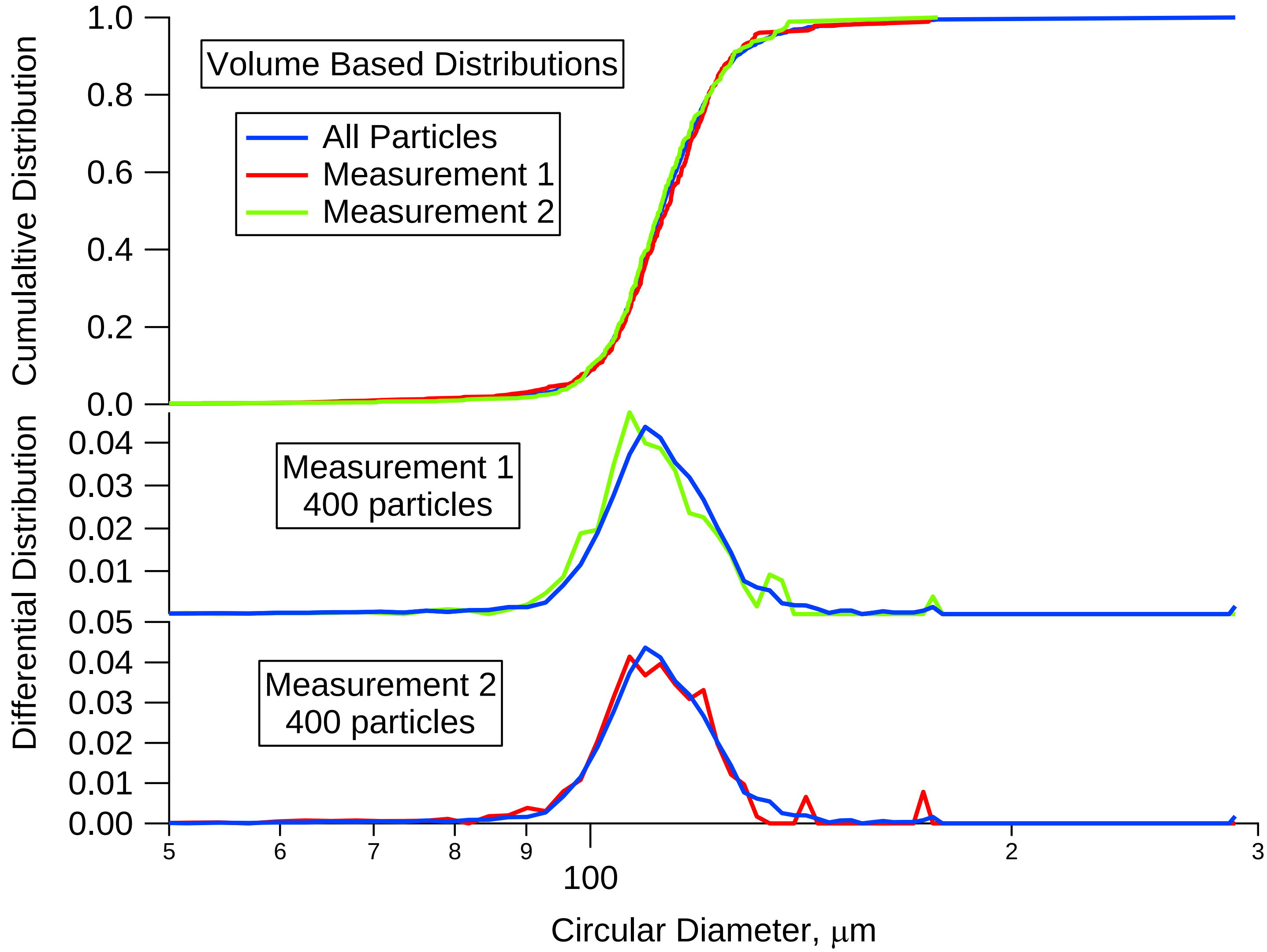
The Entire Data Set



“Split” in software: 50 particles



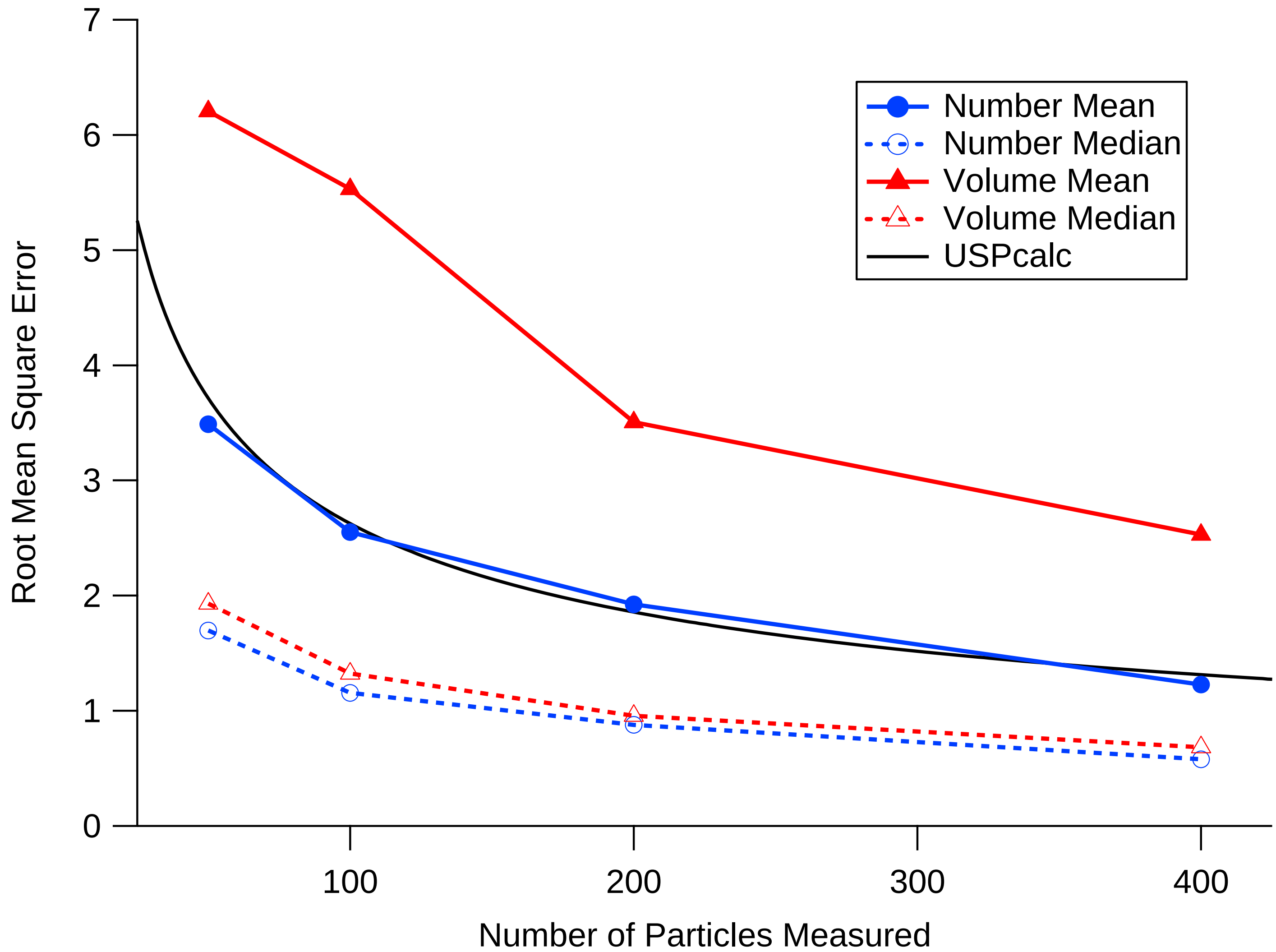
“Split” in software: 400 particles



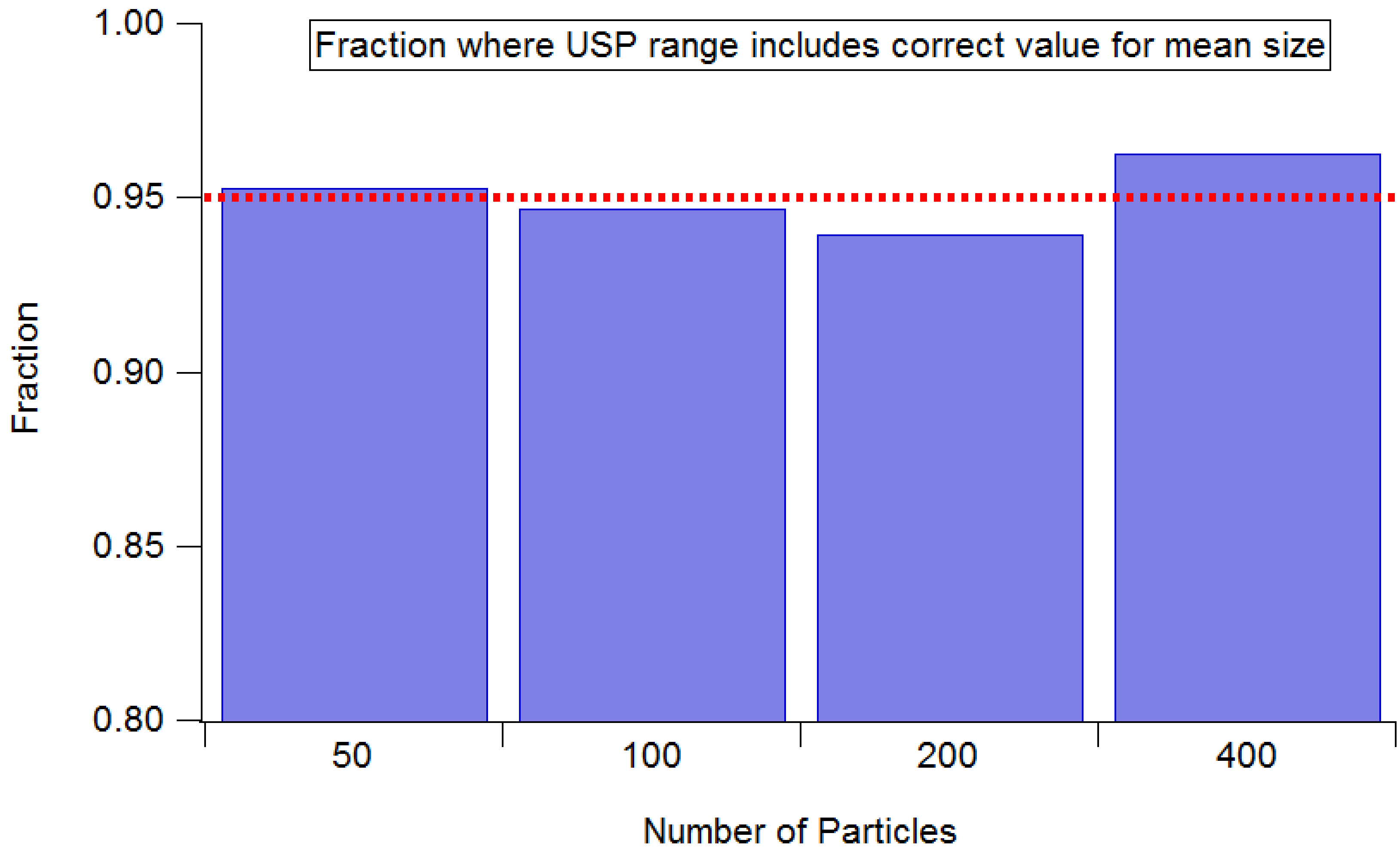
Calculations

- How to compare to USP
  - Root mean square error should match population standard deviation over the square root of the number of particles.
  - If 95% confidence chosen, 95% of measured results with added uncertainty should include correct answer.
- How to get root mean square error
  - Randomly select desired number of particles.
  - Calculate parameter (e.g., mean)
  - Compare to correct value for entire population.
  - Square the difference.
  - Average the squared differences.
  - Take the square root.

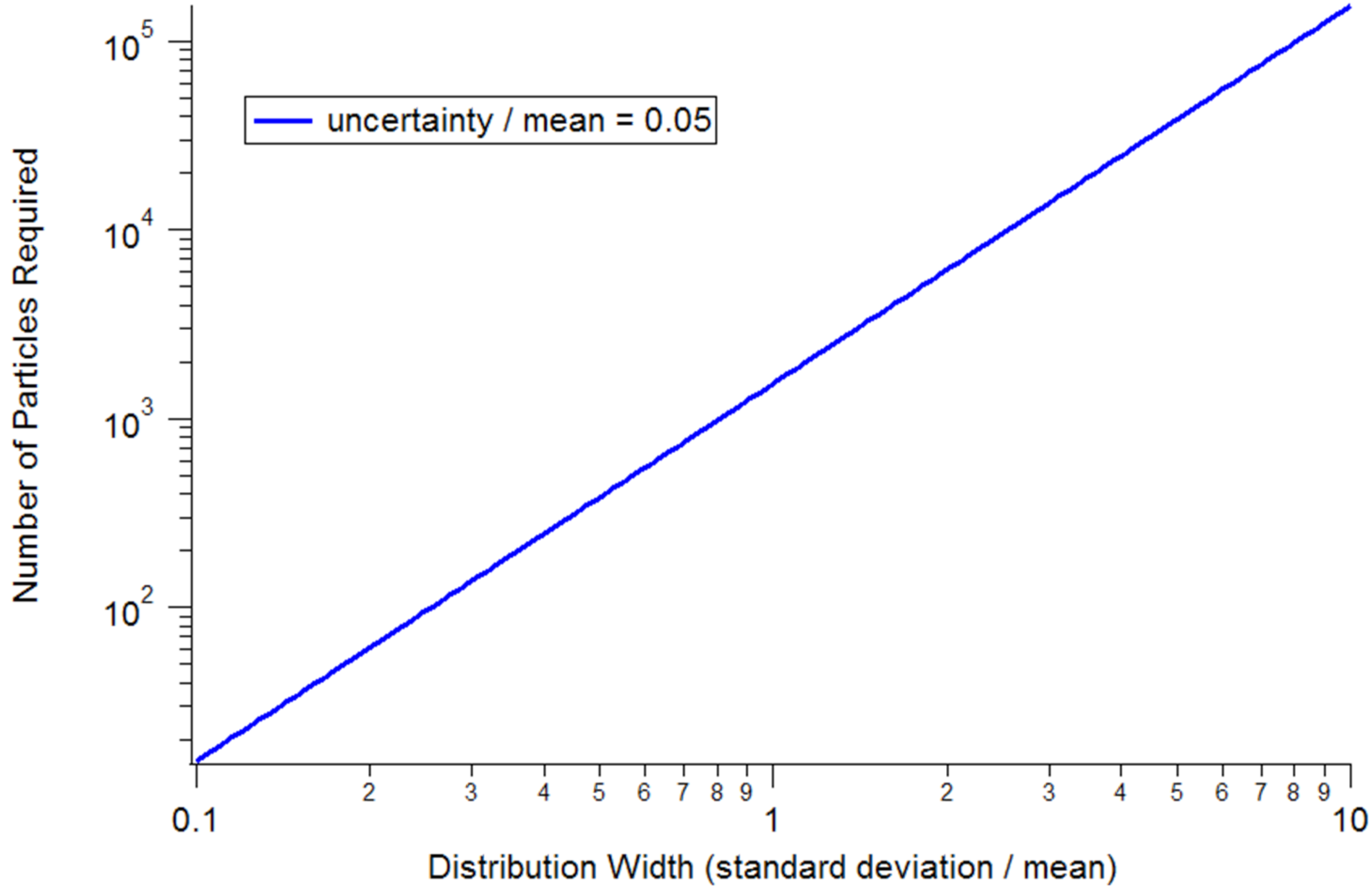
Compare With USP



Compare with USP



How Many Particles



Some materials have a distribution such that SD/Mean ~ 1.
To obtain reliable mean values, measure ~1500 particles.
To obtain more details about the distribution, more particles need to be measured.

Conclusions

- To describe a population of particles, you need to measure a number of different members of the population.
- This can be much more easily achieved using automated measurements.
- USP 776 Calculation is OK for ideal conditions, a bit optimistic for real conditions.
- Sampling (and splitting) is important.

Questions?

www.horiba.com/us/particle

Jeffrey Bodycomb, Ph.D.
Jeff.Bodycomb@Horiba.com
866-562-4698