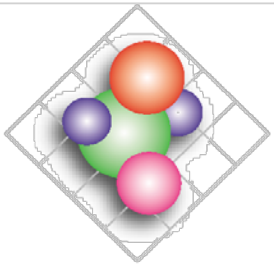
PQRI workshop on
"Sample Sizes for Decision Making in New Manufacturing Paradigms

Focus Area: Blend Uniformity

Fernando Muzzio, Professor II
Director, ERC-SOPS
Rutgers University
Department of Chemical and Biochemical Engineering,
98 Brett Road, Piscataway, NJ 08854

RUTGERS THE STATE UNIVERSITY OF NEW JERSEY PURDUE UNIVERSITY NJIT New Jersey's Science & Technology University uprm.edu universidad de puerto rico Recinto Universitario de Mayagüez



ENGINEERING RESEARCH CENTER FOR
STRUCTURED ORGANIC PARTICULATE SYSTEMS

RUTGERS UNIVERSITY
PURDUE UNIVERSITY
NEW JERSEY INSTITUTE OF TECHNOLOGY
UNIVERSITY OF PUERTO RICO AT MAYAGÜEZ

NSF

Outline

- Rules of Engagement
- Blend Uniformity or Content Uniformity?
- Blend uniformity issues
  - *What are we trying to see*
  - *How much data do we need?*
  - *What type of data do we need?*
- Thief sampling
  - *Advantages (?) and limitations (@#$%!!!!)*
  - *Case studies*
- Stratified sampling of tablets and capsules
  - *Advantages (!!!) and limitations (?)*
- PAT approaches
  - *What are we trying to see?*

Rules of Engagement

- Goal: promote a dialogue, leading to a consensus, regarding criteria for making the correct choice of blend characterization methods
- Let us take a science/engineering approach
  - *Meaningful measurements*
  - *Variability sources understood*
  - *Freedom to use the best performing/most appropriate technology*
- Let us ignore unscientific agendas
- Let us ignore regulatory legacy
  - *BU and CU, in this talk, mean the actual extent of phenomenon, not the USP/FDA legacy methods for in process and release testing*

BU or CU?

- We measure BU because it affects CU. Job 1 is to assure CU
- If we measure BU properly, we can
  - *Diagnose and mitigate causes of variability*
  - *Optimize blending, granulation, milling…*
- If we go beyond BU (and measure blend microstructure), we can
  - *Understand effect of microstructure on cohesion, hydrophobicity, compressibility, hardness, dissolution…*
  - *Support RTR*
- We can get both BU and CU from the same measurements – but it is usually too late to control process outcome

Ideal situation

- BU method is a choice driven by blend and process characteristics, not a prescription driven by legacy
- BU measurement is representative and extensive enough to support statistically significant conclusions about mixing performance and variability sources and mechanisms (measurement systems problem)
- BU measurement is instantaneous (control problem)
- Pharmaceutical scientists and regulators understand what the measurements mean
- Choice of method and mode of implementation are driven by science

Some real world problems

- Thief sampling has many sources of error (including bias), produces very little data, and is not instantaneous.
- Stratified sampling is not instantaneous and it is the most labor intensive methods
- Many PAT approaches are inaccurate, instrument dependent, not representative, and they are difficult to validate.
- Many of the people using these approaches, or regulating the use of these approaches, have limited understanding of the scientific issues
- This is great for consultants!

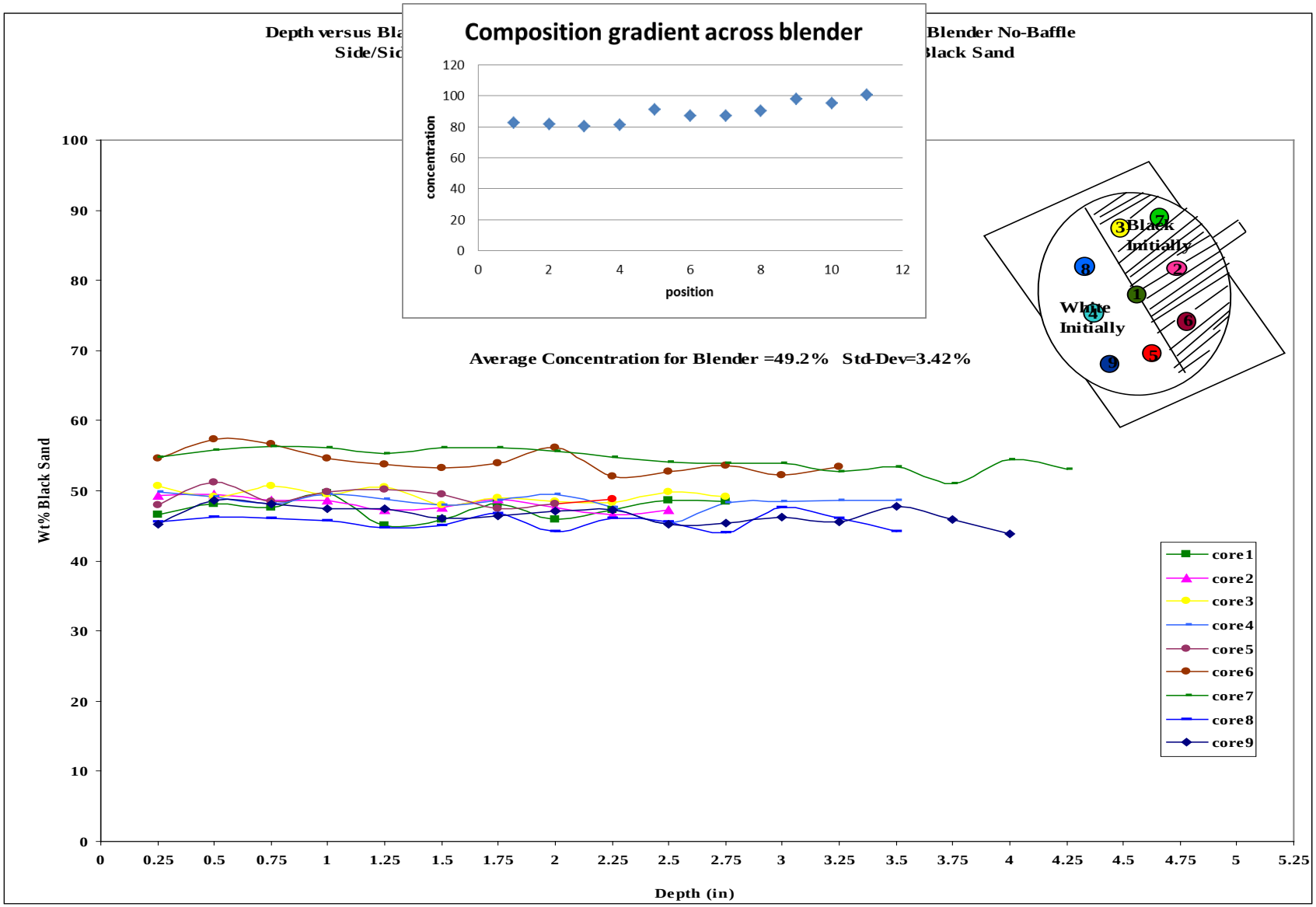
Blend uniformity: real issues

- Insufficient blending
- Segregation
- Agglomeration

Insufficient mixing:

- Scenario 1: Macroscopic heterogeneity
- Characteristic Symptoms:
  - *Composition gradients across blender*
  - *Time-dependent potency of tablets and capsules*
  - *RSD decreases with time but is independent of sample size*
- Causes:
  - *Inadequate loading*
  - *Excessive blender filling*
  - *Mixing time too short (often due to incorrect scale-up)*
- How to diagnose: need to resolve spatial gradient inside blender, or temporal gradient at the blender discharge
- Measurement requirements: System must be able to resolve spatial and/or temporal gradient in statistically significant manner

Example of insufficient mixing due to bad loading

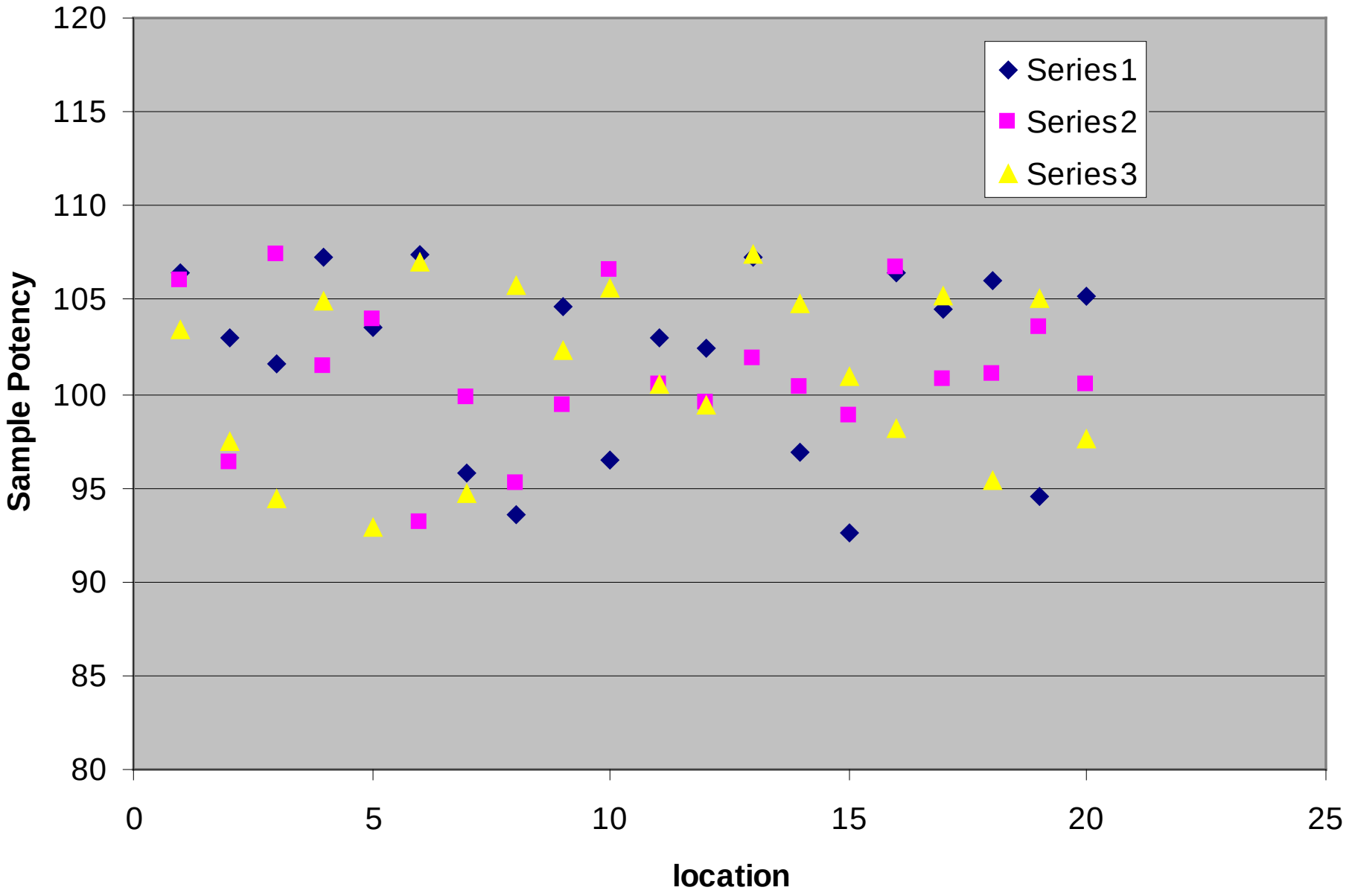


Insufficient mixing:

- Scenario 2: Coarse cohesive blend
- Characteristic Symptoms:
  - *Blend has high RSD but does not display gradients or outliers*
  - *RSD decreases with time and may be sample size dependent*
  - *Tablet or capsule potency does not display time-dependent pattern*
- Causes:
  - *Cohesion of the blend or the API*
  - *Mixing time too short (Not enough shear)*
  - *Characteristic of small scale mixers or low shear mixers*
- How to diagnose: need to demonstrate lack of spatial gradient inside blender, and/or lack of temporal gradient at the blender discharge
- Measurement requirements: System must be able to resolve spatial and/or temporal gradients in statistically significant manner

Examples – Coarse Mixture

Set 1 - POWDER SAMPLES WITH NO MAJOR DEVIATIONS - SHOWS THE EFFECT OF AVERAGING									
Case1: Random Mixture with moderate variability									
RSD from individual values is higher than RSD of averages									
Average of location RSD's lower than RSD of individual values but higher than RSD of averages									
Location	Sample1	sample2	sample3	Location Avg					
1	106.4708	105.9772	103.4237	105.2906					
2	103.0394	96.30543	97.44525	98.93001					
3	101.6582	107.3447	94.42169	101.1416					
4	107.2126	101.4732	104.9737	104.5532					
5	103.5762	103.9814	92.80578	100.1211					
6	107.3996	93.14939	106.963	102.504					
7	95.82761	99.77068	94.67383	96.75737					
8	93.54683	95.202	105.7024	98.15039					
9	104.6373	99.3774	102.2482	102.0876					
10	96.52324	106.5639	105.6518	102.913					
11	102.9206	100.4219	100.5412	101.2946					
12	102.3844	99.46263	99.31681	100.388					
13	107.2248	101.8736	107.4488	105.5158					
14	96.8296	100.291	104.7572	100.6259					
15	92.57886	98.82437	100.9265	97.44326					
16	106.4739	106.6654	98.16339	103.7676					
17	104.5106	100.7172	105.1664	103.4647					
18	106.0105	101.0751	95.34597	100.8105					
19	94.5878	103.5672	105.0579	101.0709					
20	105.2407	100.5285	97.56177	101.1103					
Mean Ind		101.397		101.397	Mean Avg				
RSD Ind		0.043922		0.024117	RSD Avg				


120
115
110
105
100
95
90
85
80
Sample Potency
0
5
10
15
20
25
location
Series1
Series2
Series3

locations										
1	2	3	4	5	6	7	8	9	10	11
106.4708	103.0393519	101.6582	107.2126	103.5762	107.3996	95.82761	93.54683	104.6373	96.52324	102.9206
105.9772	96.30542869	107.3447	101.4732	103.9814	93.14939	99.77068	95.202	99.3774	106.5639	100.4219
103.4237	97.44525284	94.42169	104.9737	92.80578	106.963	94.67383	105.7024	102.2482	105.6518	100.5412
105.2906	98.93001115	101.1416	104.5532	100.1211	102.504	96.75737	98.15039	102.0876	102.913	101.2946
0.015533	0.03643119	**0.064039**	0.027667	**0.063308**	**0.079063**	0.027622	**0.067166**	0.025798	**0.053953**	0.013914

12	13	14	15	16	17	18	19	20		
102.3844	107.2248	96.8296	92.57886	106.4739	104.5106	106.0105	94.5878	105.2407		
99.46263	101.8736	100.291	98.82437	106.6654	100.7172	101.0751	103.5672	100.5285		
99.31681	107.4488	104.7572	100.9265	98.16339	105.1664	95.34597	105.0579	97.56177		
100.388	105.5158	100.6259	97.44326	103.7676	103.4647	100.8105	101.0709	101.1103		avg of location RSDs
0.017239	0.029912	0.039497	0.044558	0.046781	0.023215	**0.052943**	**0.056038**	0.038298		0.041149

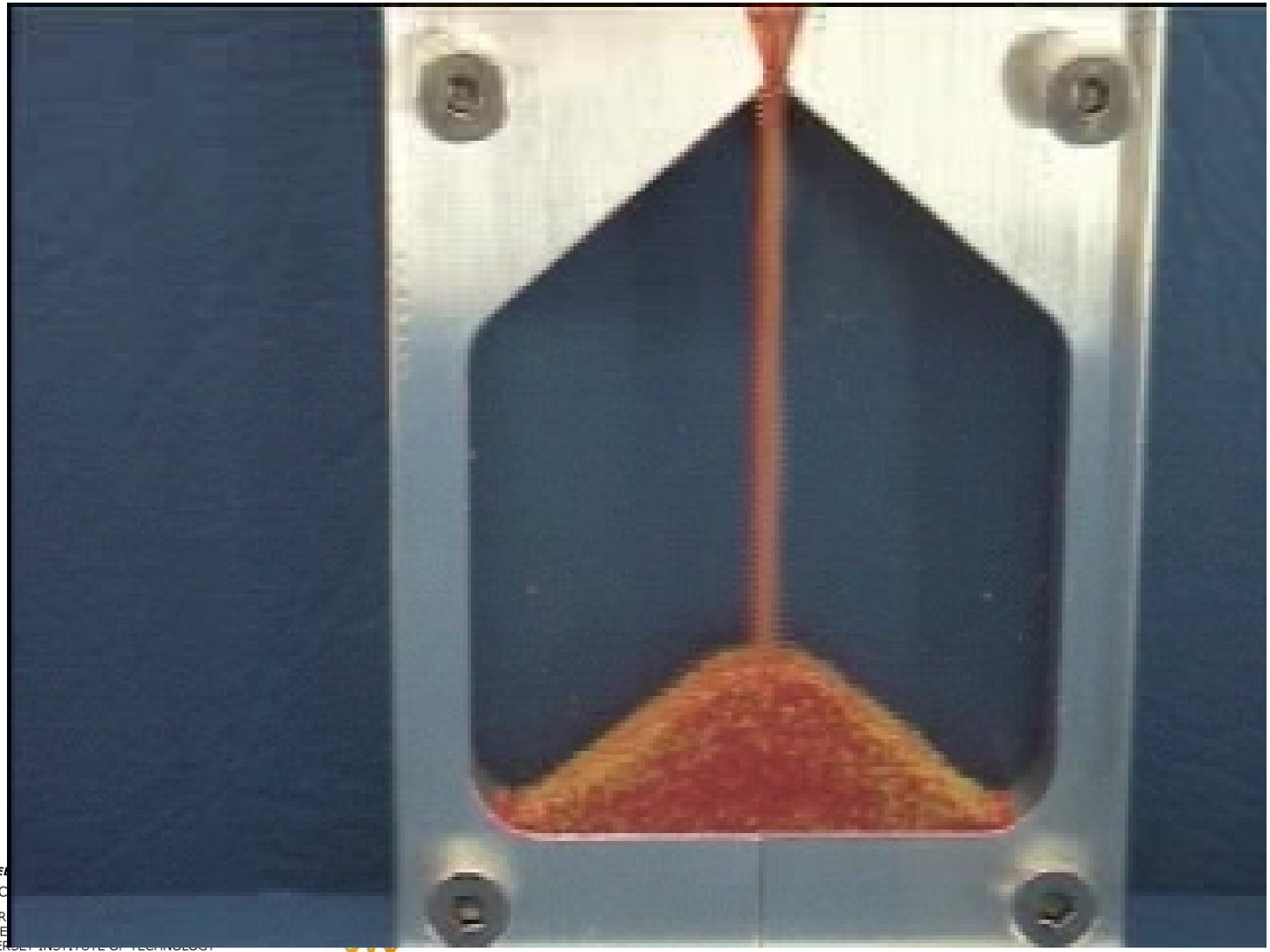
ANOVA						
ource of Variatio	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	340.8429	19	17.9391	0.86518	0.623243	1.852893
Within Groups	829.3814	40	20.73453			
Total	1170.224	59				

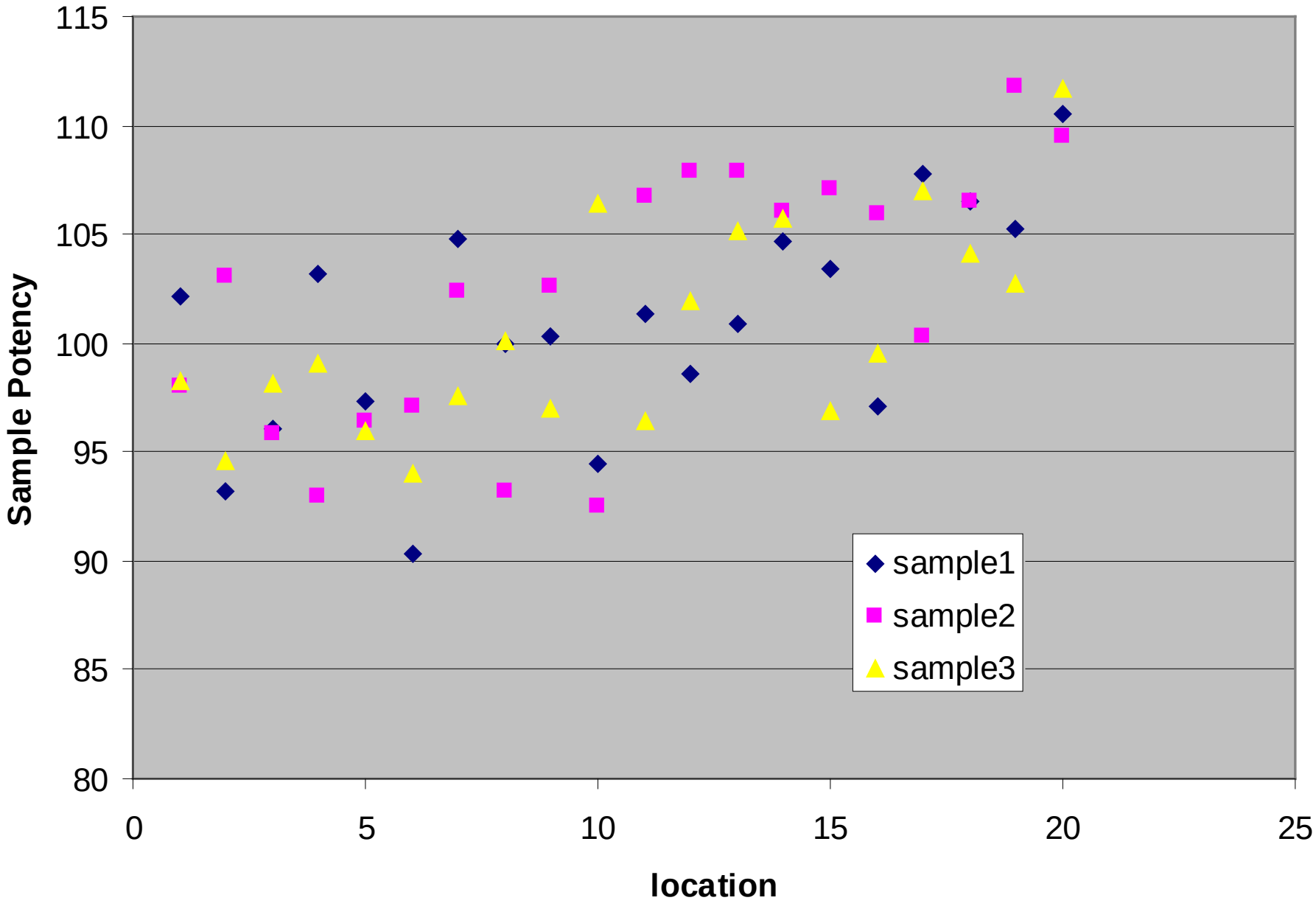
Locations are statistically similar

Segregation

- Characteristic Symptoms:
  - *Composition gradients across blender*
  - *Time-dependent potency of tablets and capsules*
  - *RSD does not decrease with time and is independent of sample size*
  - *If process is robust, effect is significant across batches*
- Causes:
  - *Formulation (particle size, density, shape)*
  - *Uncontrolled blender discharge*
- How to diagnose: need to resolve spatial gradient inside blender, or temporal gradient at the blender discharge
- Measurement requirements: System must be able to resolve spatial and/or temporal gradient in statistically significant manner

Video de segregación (Schulze)




115
110
105
100
95
90
85
80
Sample Potency
0
5
10
15
20
25
location
sample1
sample2
sample3

Location Analysis											
1	2	3	4	5	6	7	8	9	10	11	12
102.1613	93.14619	96.11046	103.1942	97.38091	90.28382	104.7371	99.95092	100.267	94.42307	101.3268716	98.53565
98.0293	103.0861	95.82992	92.97771	96.46062	97.1001	102.4056	93.14839	102.625	92.54652	106.691839	107.8806
98.24226	94.52864	98.11544	99.10535	95.9349	93.94717	97.51275	100.0957	96.99664	106.3679	96.40154083	101.8626
	6										
99.47761	96.92031	96.68528	98.42575	96.59214	93.77703	101.5518	97.73166	99.96288	97.77916	101.4734171	102.7596
0.023388	0.055554	0.012892	0.052243	0.007577	0.036377	0.036307	0.04062	0.028275	0.076673	0.050719826	0.046094

13	14	15	16	17	18	19	20				
100.9177	104.6478	103.383	97.13934	107.7914	106.4951	105.2498	110.48				
107.8351	106.0472	107.1061	105.9849	100.2845	106.4927	111.7424	109.4915				
105.1271	105.7346	96.89972	99.4883	106.9419	104.1458	102.7558	111.6179				
104.6266	105.4765	102.4629	100.8709	105.0059	105.7112	106.5826	110.5298		average of location RSDs		
0.033316	0.006964	0.050409	0.045424	0.039149	0.012824	0.043527	0.009627		0.035398		

ANOVA is used to analyze variability between locations

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Column 1	3	298.4328	99.47761	5.412861
Column 2	3	290.7609	96.92031	28.99044
Column 3	3	290.0558	96.68528	1.553711
Column 4	3	295.2772	98.42575	26.44044
Column 5	3	289.7764	96.59214	0.535715
Column 6	3	281.3311	93.77703	11.63712
Column 7	3	304.6555	101.5518	13.59439
Column 8	3	293.195	97.73166	15.76004
Column 9	3	299.8886	99.96288	7.988986
Column 10	3	293.3375	97.77916	56.20509
Column 11	3	304.4203	101.4734	26.48867
Column 12	3	308.2789	102.7596	22.43557
Column 13	3	313.8799	104.6266	12.15069
Column 14	3	316.4296	105.4765	0.539543
Column 15	3	307.3888	102.4629	26.67729
Column 16	3	302.6126	100.8709	20.99469
Column 17	3	315.0178	105.0059	16.89935
Column 18	3	317.1335	105.7112	1.837867
Column 19	3	319.7479	106.5826	21.52236
Column 20	3	331.5894	110.5298	1.132315

ANOVA

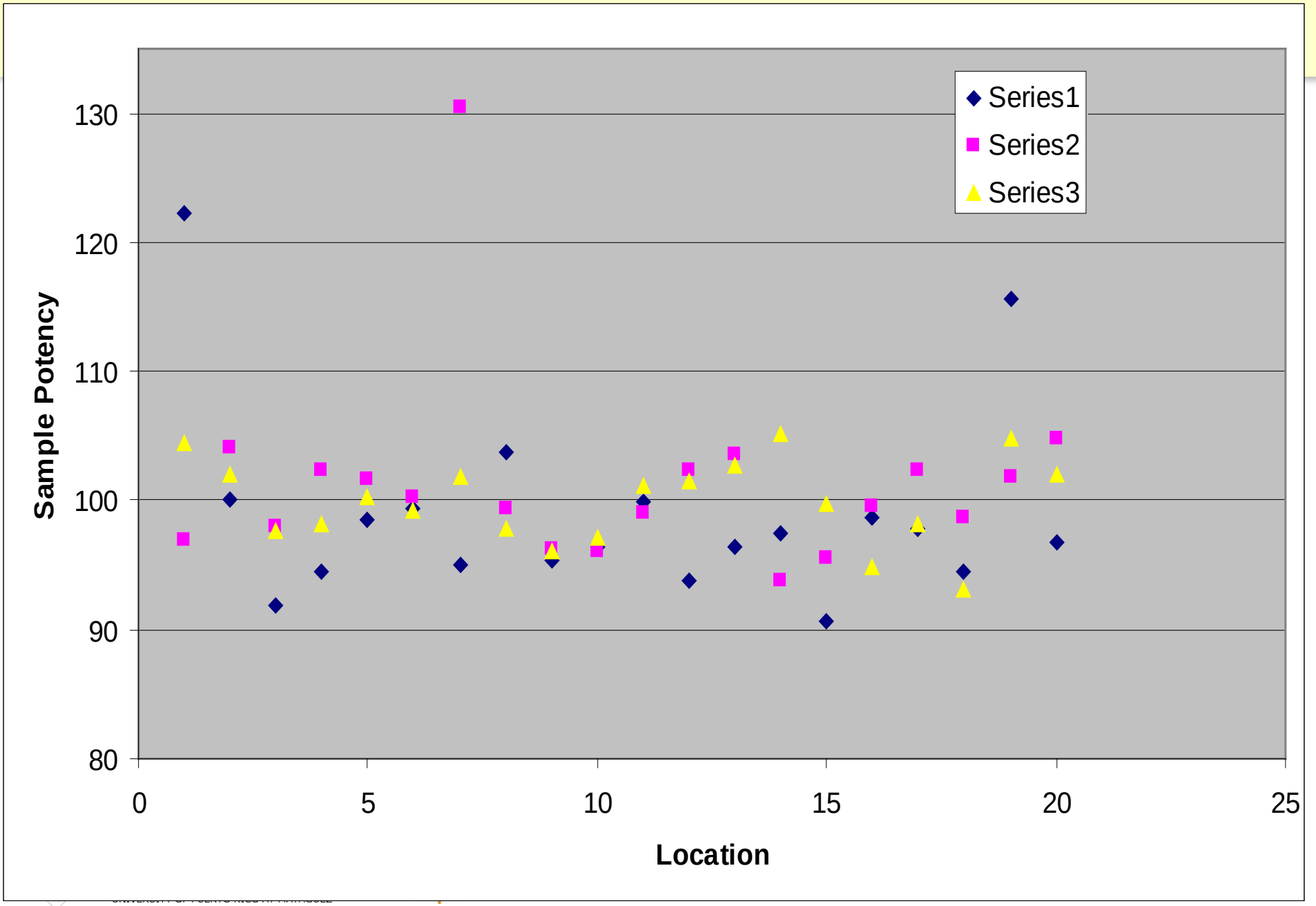
rce of Varia	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between G	1008.501	19	53.07902	3.329956	0.000669	1.852893
Within Gro	637.5943	40	15.93986			
Total	1646.096	59				

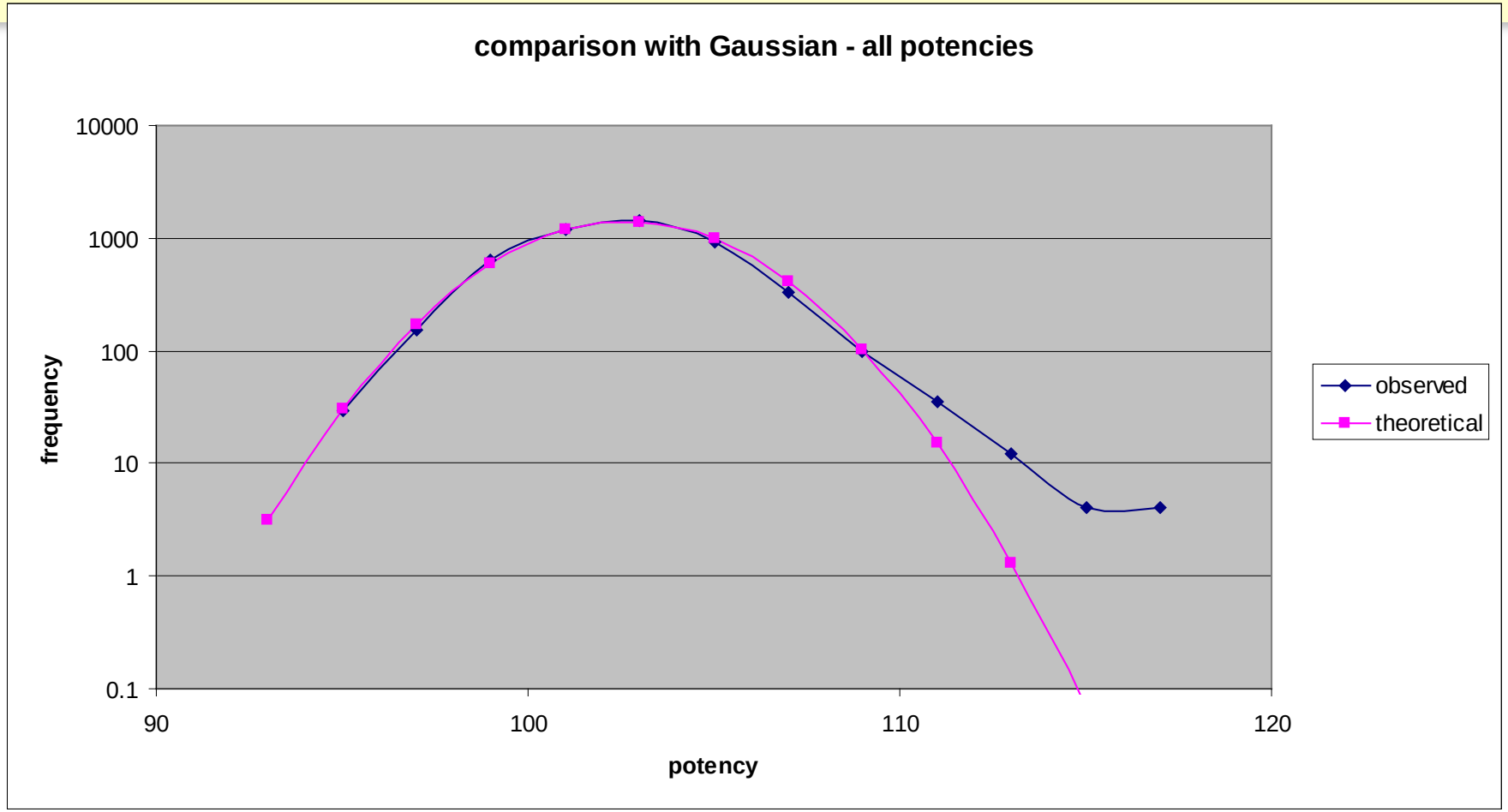
Anova between locations: Locations ARE significantly different

Agglomeration

- Characteristic Symptoms:
  - *Isolated "hot spots" where samples (or individual tablets) show significant superpotency*
  - *No "low fliers"*
  - *High values are outside the normal population (they show up as outliers)*
  - *RSD decreases with sample size*
- Causes:
  - *API not properly de-agglomerated*
  - *API agglomerates in blender due to electrostatics, moisture, MgSt softening, etc*
- How to diagnose: perform a large number of measurements. Determine both the underlying normal distribution and detect the outliers.
- Measurement requirements: System must be able to measure a large number of values (several hundred)

Set 3 - POWDER SAMPLES WITH AGGLOMERATES - SHOWS					
Case 3: Mixture with Agglomerates					
Location	Sample1	sample2	sample3	Location Avg	
1	**122.3123**	96.95018	104.5146	107.9257	
2	100.1209	104.0918	101.9234	102.0454	
3	91.79322	98.0125	97.63952	95.81508	
4	94.52912	102.2845	98.09353	98.30239	
5	98.43445	101.6035	100.2304	100.0895	
6	99.40265	100.2777	99.20356	99.62797	
7	94.96024	**130.4**	101.8727	109.0777	
8	103.6684	99.34741	97.77706	100.2643	
9	95.33494	96.24522	96.09863	95.89293	
10	96.44071	96.07732	97.18798	96.56867	
11	99.93404	99.0759	101.1966	100.0689	
12	93.86015	102.3704	101.4504	99.22699	
13	96.34526	103.5954	102.6602	100.867	
14	97.45158	93.79322	105.088	98.77759	
15	90.73452	95.51705	99.80418	95.35192	
16	98.72903	99.59913	94.88759	97.73858	
17	97.80984	102.2983	98.15117	99.41976	
18	94.50889	98.68102	93.11534	95.43508	
19	**115.622**	101.8041	104.7755	107.4005	
20	96.82955	104.7974	102.069	101.232	
Mean Ind		100.0564		100.0564	Mean Avg
				4.010655	
RSD Ind		**0.063937**		0.040084	RSD Avg


Series1
Series2
Series3
Sample Potency
130
120
110
100
90
80
0
5
10
15
20
25
Location


comparison with Gaussian - all potencies
10000
1000
100
10
1
0.1
frequency
90
100
110
120
potency
observed
theoretical

A useful tool: the q plot (normal probability plot)

- Standard score:

$$z = \frac{y - \bar{y}}{s}$$

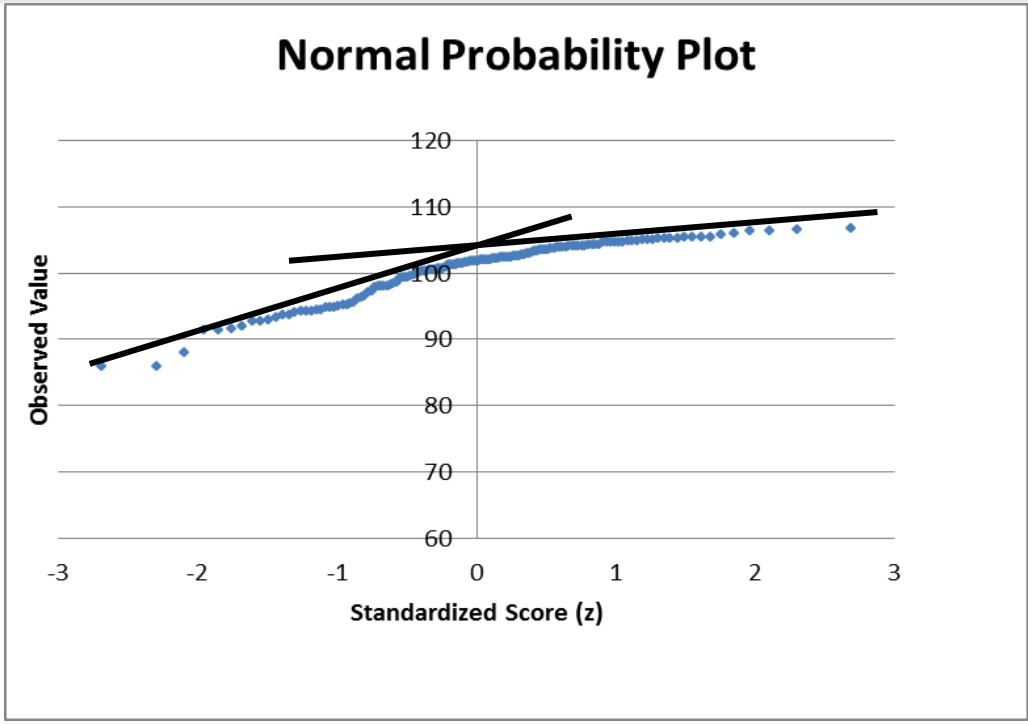
z is the normal score corresponding to the percentile of y

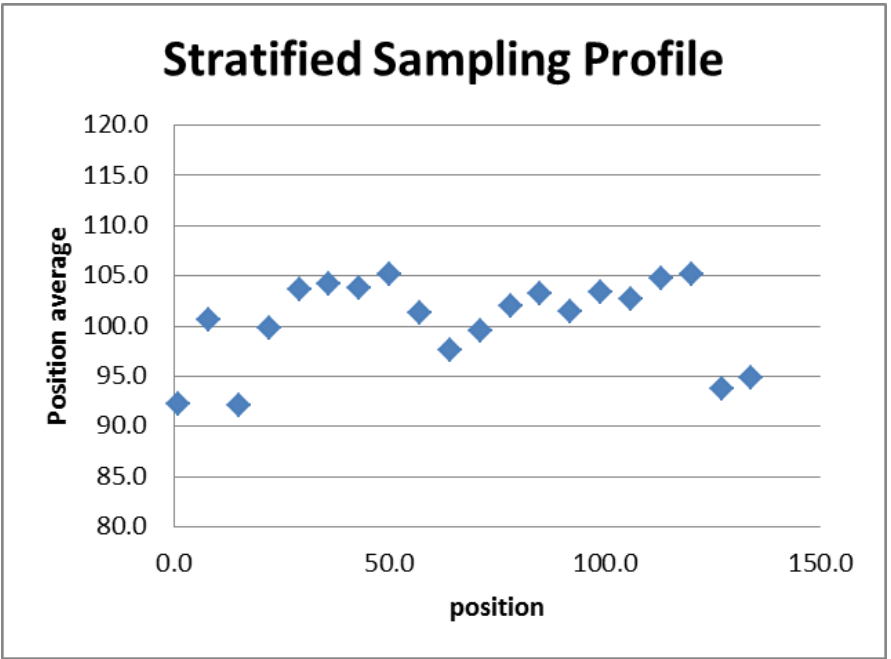
Thus, a plot of y vs. z has a slope of s and an intercept of $\bar{y}$

1	100.6	SUMMARY OUTPUT						
2	99.7							
3	100.3	*Regression Statistics*						
4	97.3	Multiple R	0.043947					
5	100.9	R Square	0.001931					
6	104.6	Adjusted R Square	-0.01803					
7	99.7	Standard Error	2.76406					
8	100.4	Observations	52					
9	100.6							
10	100.3	ANOVA						
11	103.6		*df*	*SS*	*MS*	*F*	*ignificance F*	
12	107.5	Regression	1	0.739205	0.739205	0.096754	0.757053	
13	105.8	Residual	50	382.0014	7.640027			
14	102.8	Total	51	382.7406				
15	112.4							
16	100.4		*Coefficients*	*tandard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
17	102.4	Intercept	102.1029	0.777804	131.2707	3.85E-65	100.5407	103.6652
18	104.2	X Variable 1	0.007944	0.02554	0.311053	0.757053	-0.043354	0.059242
19	101.3							
20	98.6							
21	99.9							
22	100.9	PROBABILITY OUTPUT				normsinv(L363/100)		
23	103.1							
24	102.7	*Percentile*	*Y*	*Z*				
25	109.7	0.961538462	97.3	-2.34103				
26	101.2	2.884615385	98.6	-1.89803				
27	104.4	4.807692308	99.7	-1.66379				
28	102	6.730769231	99.7	-1.49615				
29	101	8.653846154	99.8	-1.36238				
30	102.5	10.57692308	99.9	-1.24935				
31	101	12.5	100.3	-1.15035				
32	102.3	14.42307692	100.3	-1.0615				
33	106.2	16.34615385	100.3	-0.98033				
34	101.8	18.26923077	100.4	-0.90515				
35	101.6	20.19230769	100.4	-0.83477				
36	106.6	22.11538462	100.6	-0.7683				
37	101.6	24.03846154	100.6	-0.70507				
38	100.7	25.96153846	100.7	-0.64453				

Examples from the real world

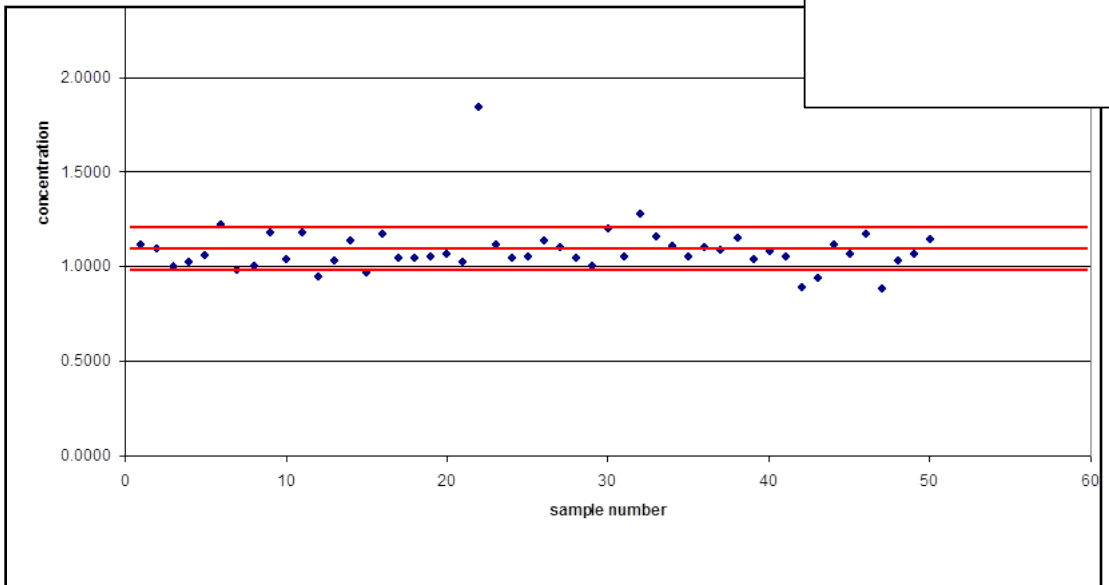
A large tablet data set from a single batch, displaying the multi-slope pattern characteristic of segregation or incomplete macromixing





Examples from the real world

Multi-batch tablet dataset containing several thousand measurements, displaying agglomeration fingerprint



Measurement methods: thief sampling

- Advantages
  - *Cheap to implement*
  - *Easily accepted by regulators*
  - *Measures variability at unit dose level*
- Disadvantages
  - *Bias (samples can be consistently subpotent or consistently superpotent)*
  - *API can stick to thief*
  - *Small number of samples (10-30), combined with large sampling error, produce unreliable results with low statistical significance*
  - *Slow, useless for control*
  - *Exposes operators to powder*
  - *Method does not detect segregation after blending*
  - *Hard (or impossible) to resolve spatial gradients or detect agglomerates*
  - *Hard (or impossible) to validate method*
  - *Difficult to assure consistent location sampling*

Measurement methods: stratified sampling of tablets and/or capsules

- Advantages
  - *Cheap, easy to implement*
  - *Highly accurate, unbiased*
  - *Highly representative (devoid of sampling problems)*
  - *Produces large number of samples, statistically significant results*
  - *Allows to discriminate sources of variability – detects gradients and agglomerates*
  - *Does not expose operators*
  - *Easy to validate*
  - *Measures variability at unit dose level*
  - *It is an effective approach for validating sampling and PAT!*
- Disadvantages
  - *Slow, useless for control*
  - *Not readily accepted by some regulators*

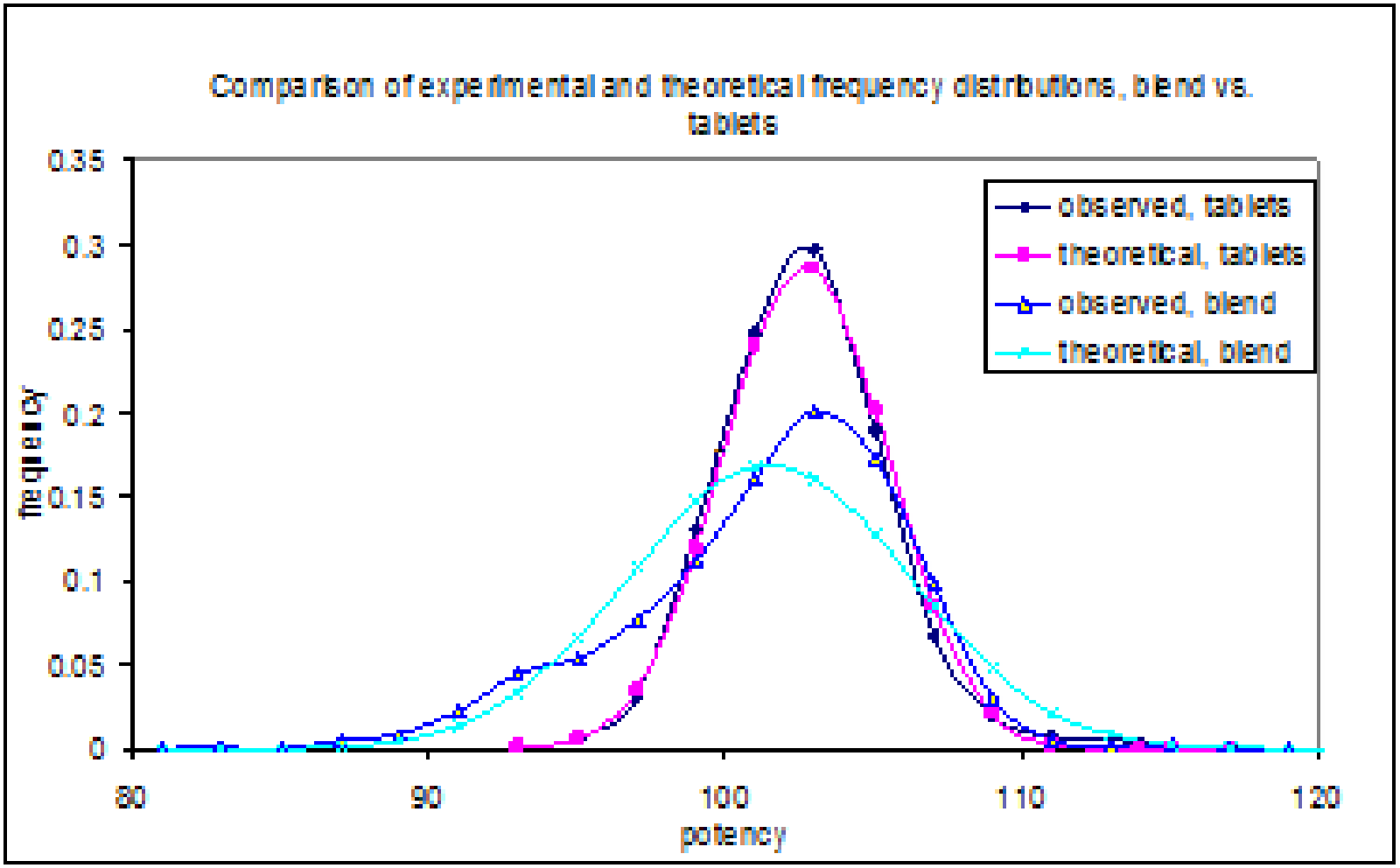
PAT methods

- Advantages
  - *Fast, useful for control*
  - *Produces large amount of statistically significant data*
  - *If multiple sensors are used, can resolve spatial gradients*
  - *If used at the right spot, can resolve temporal gradients at discharge*
  - *Ideal for continuous processing*
- Disadvantages
  - *Can be difficult to assure sample size*
  - *Can be difficult to validate*
  - *Expensive to implement, requires expertise and maintenance*
  - *Sensors can get fouled*
  - *If samples segregate, NIR method can be difficult to establish*

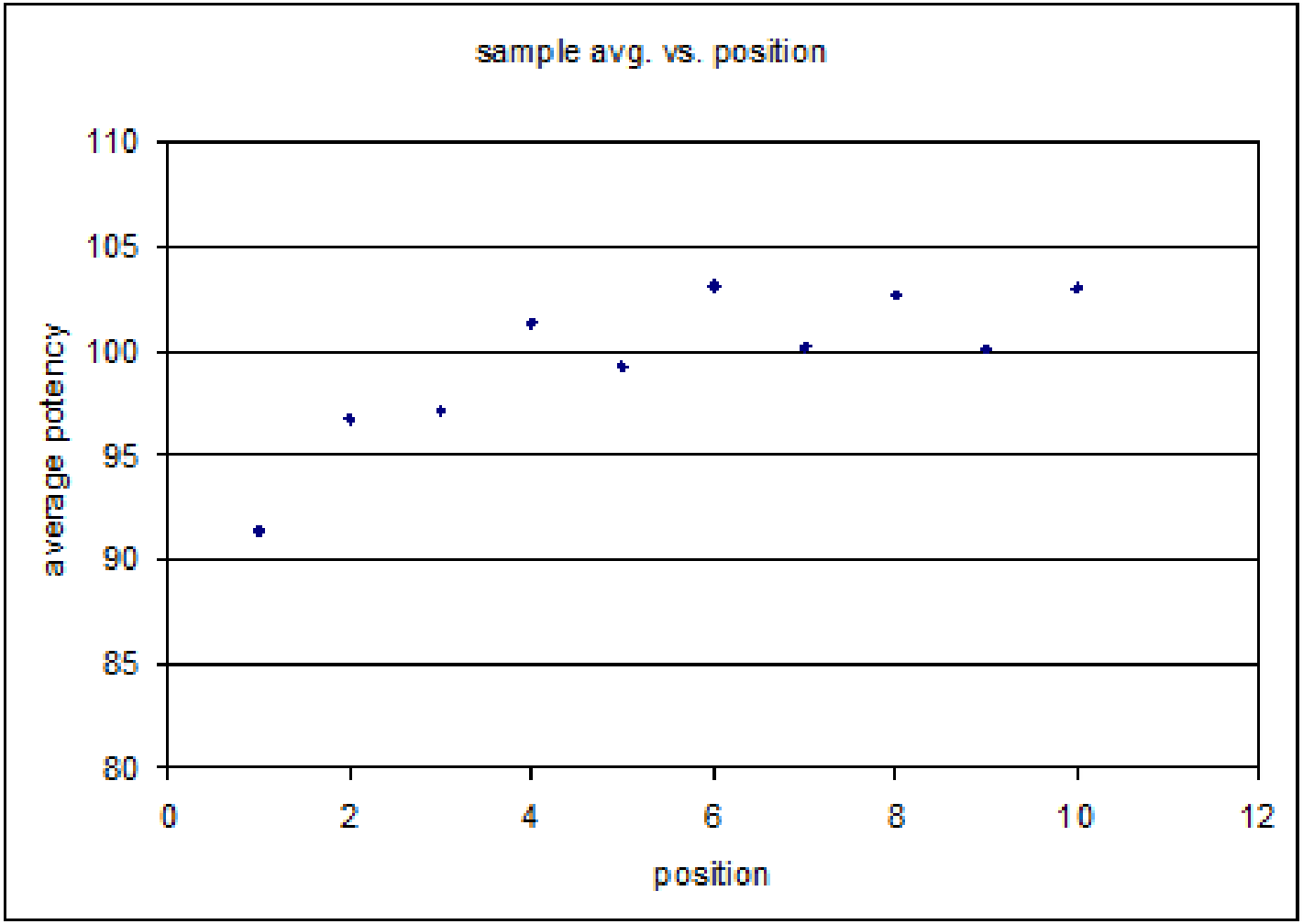
Case study 1: API sticking to thief

- Low drug content dry blend application
- Large number of batches (>150)
- Thief sampling of blender (12 positions)
- Large number of tablets analyzed

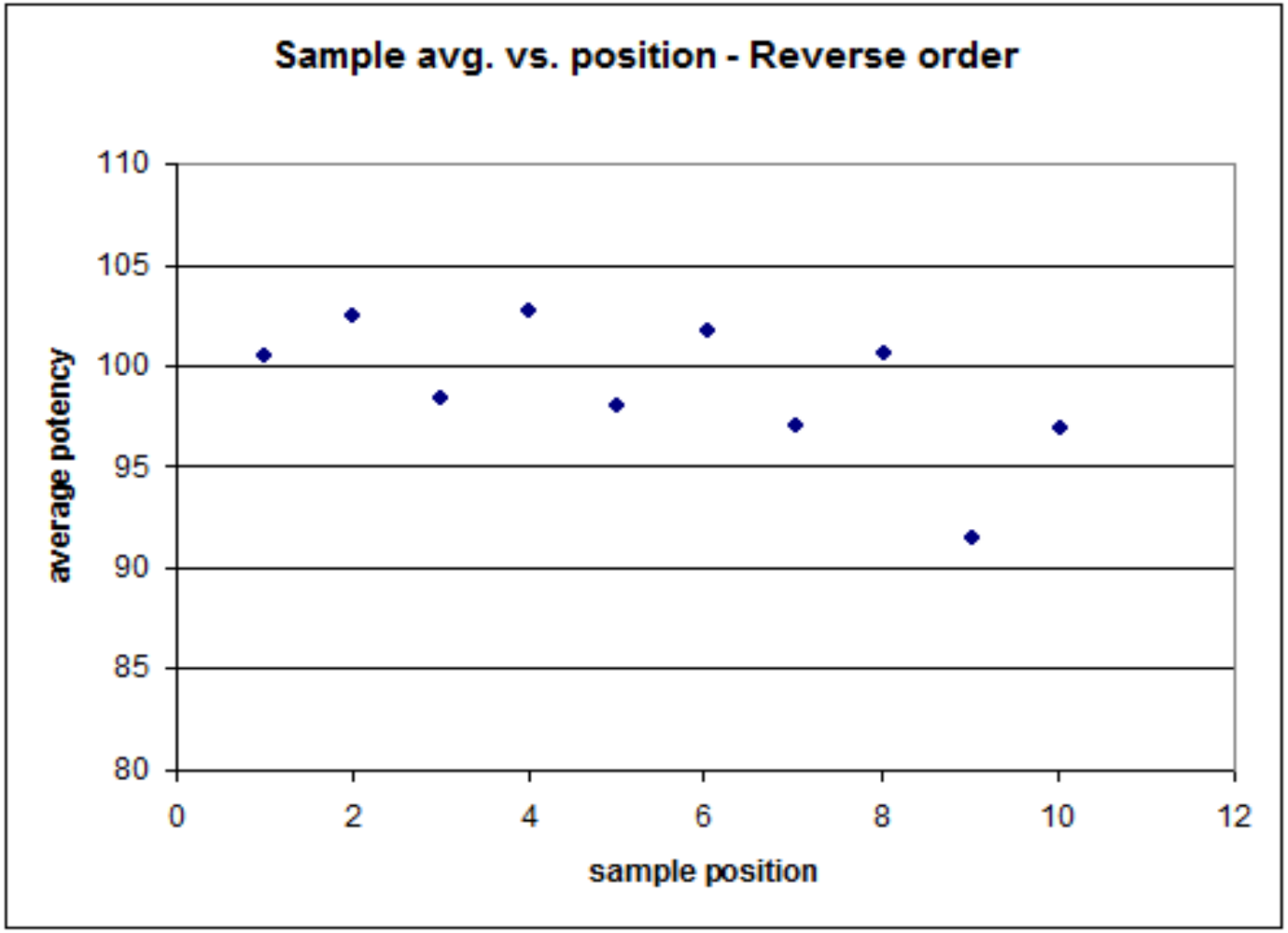
Strong deviations between samples and tablets



Subpotent positions in blender…



… due to API sticking in the thief for the first two samples



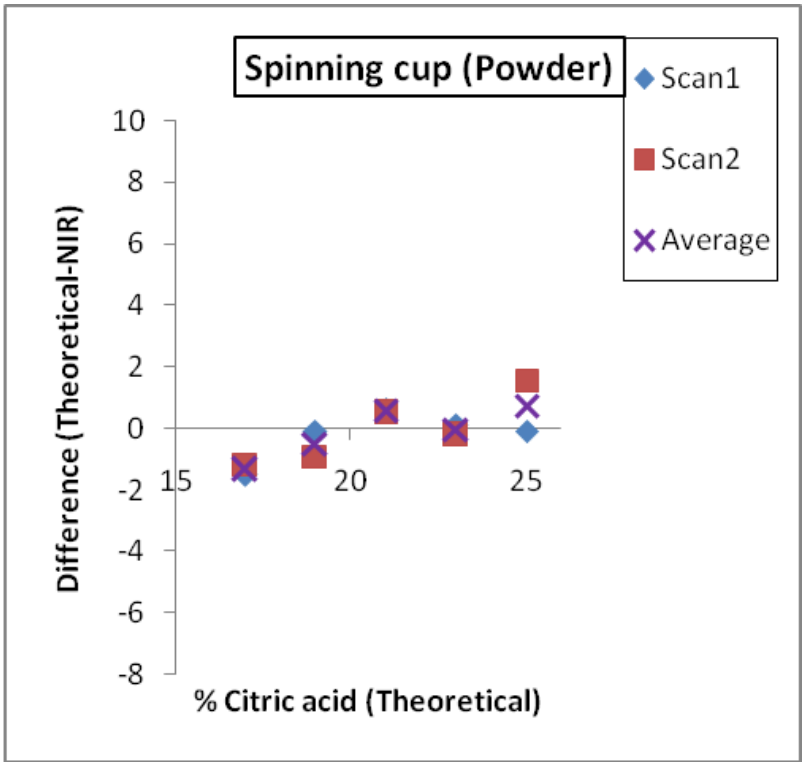
Case Study 2: comparison of NIR and thief sampling for dry beverage blend

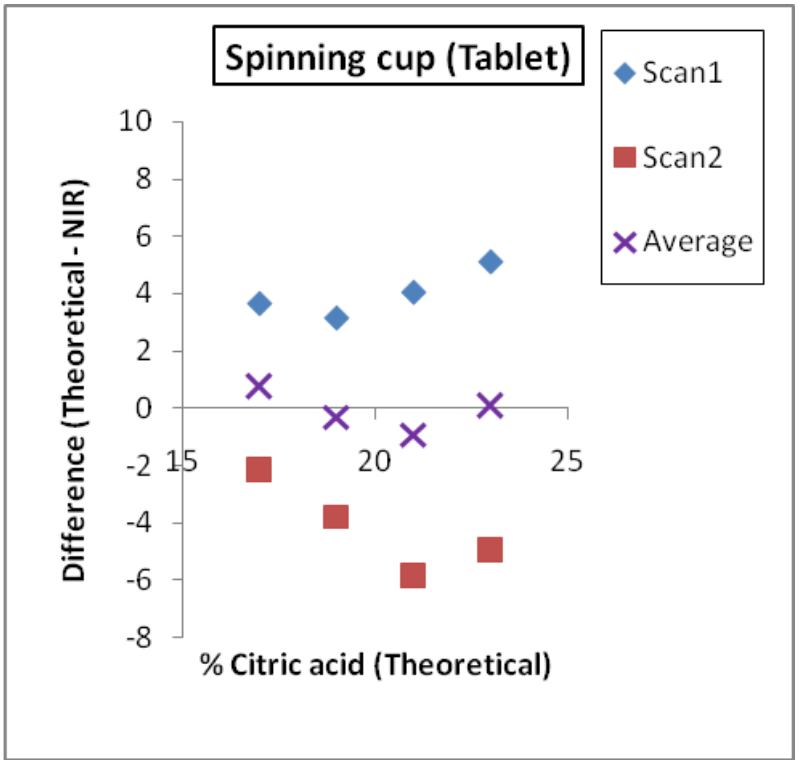
- Application was in dry beverage mix
- Two “actives” (trisodium citrate and lemon flavor) mixed in a bin with multiple “ingredients” (flavors, preservatives), then mixed with a third “active” (citric acid) and with “major ingredients” (tea, sugar) in a continuous mixer.
- NIR methods used at the bench, and on-line for both the bin and the continuous mixer
- Samples extracted using thief sampling
- Studies examined and compared PAT method to bench analysis of thief samples

NIR was difficult

- Blend segregation was a major issue
- Samples segregated intensely, affecting NIR accuracy
- Multiple bench methods attempted
  - *Repeated measures with shaking episodes in between*
  - *Rotating cup for powders*
  - *Tableting, followed by rotated tablet*

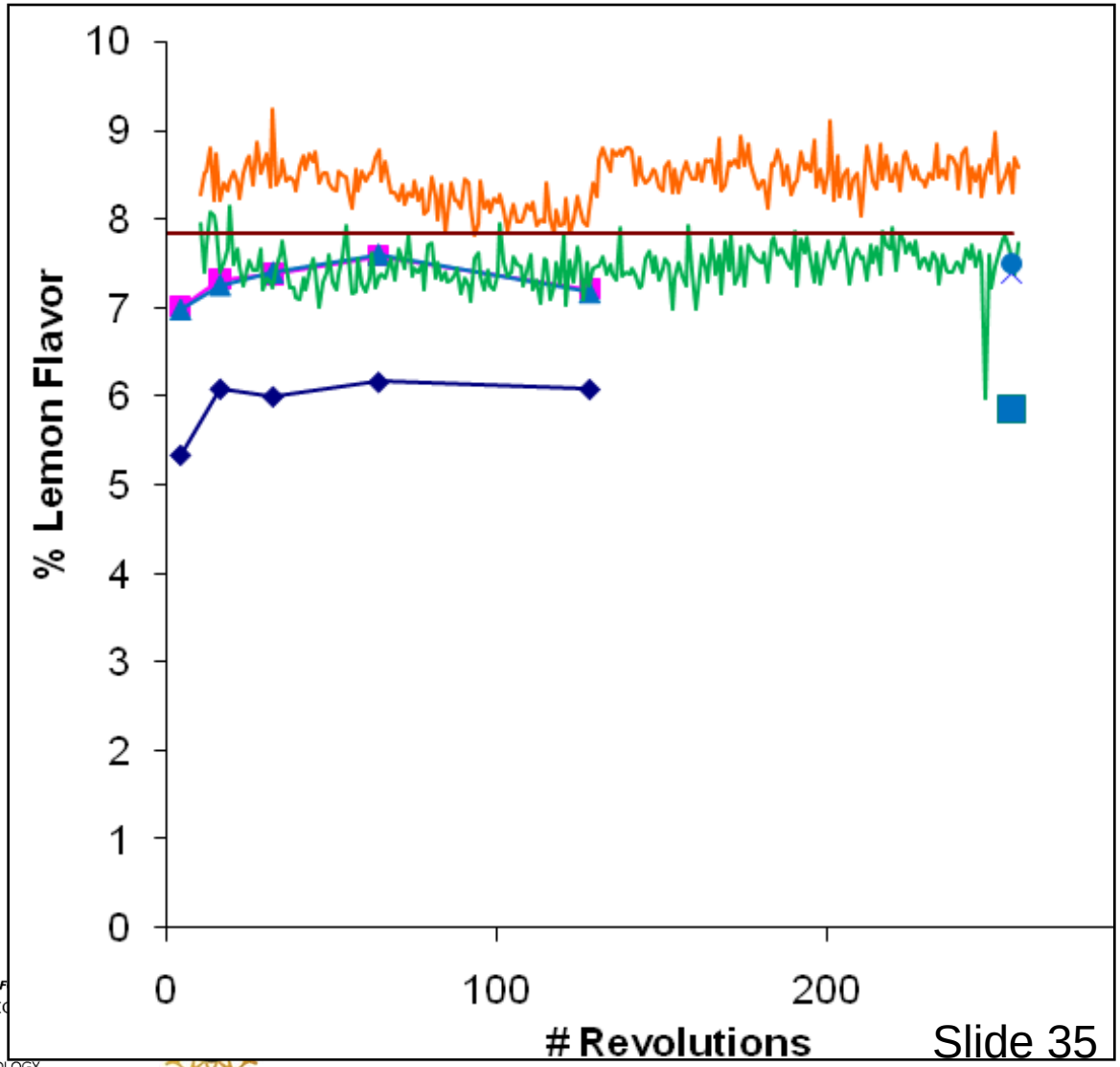
Comparison of rotating sample vs. tablet





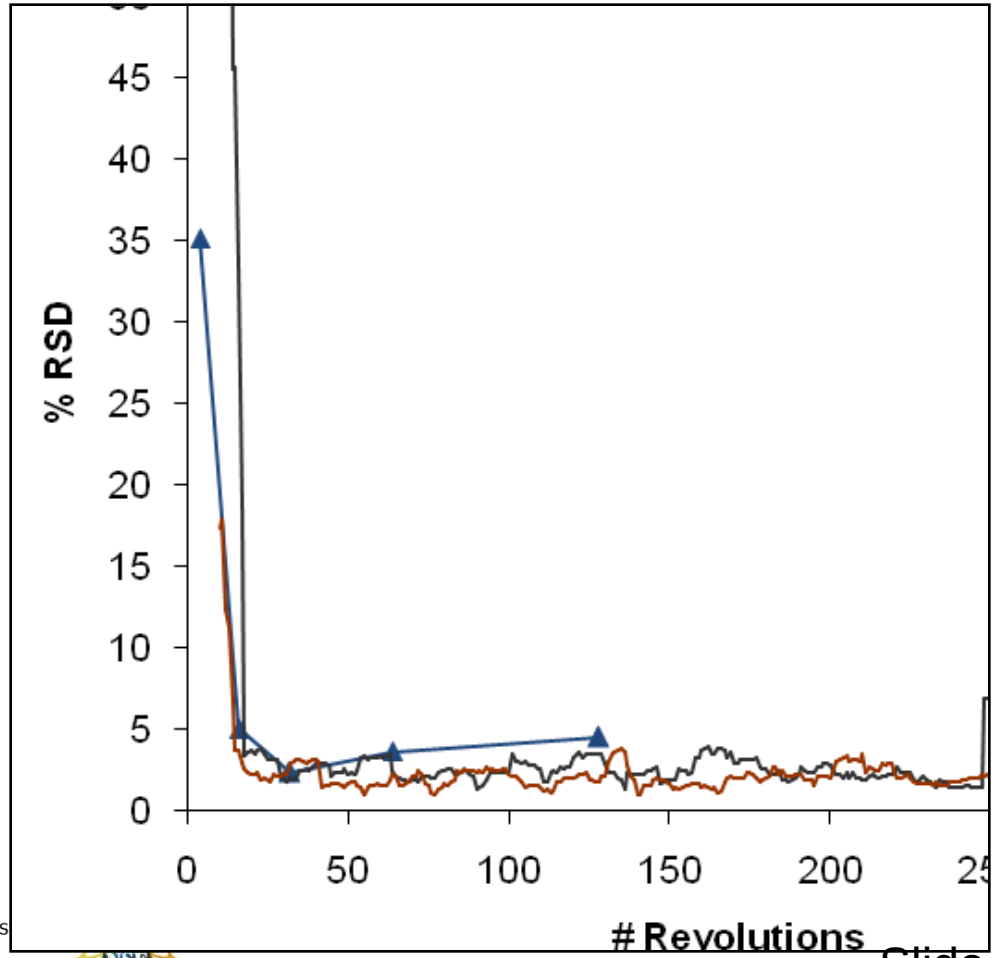
- Tablet method was promising, but was not selected by sponsor. Work continued for powder method

Comparison online vs. bench – mean – ingredient 1

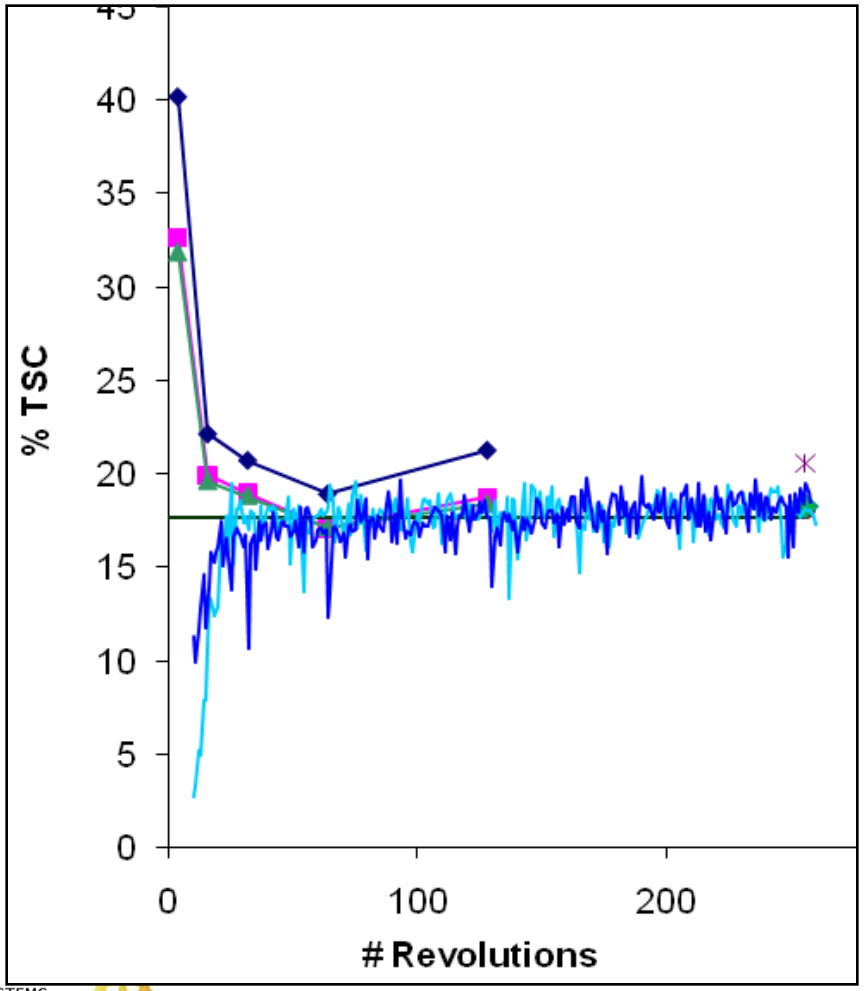


Slide 35

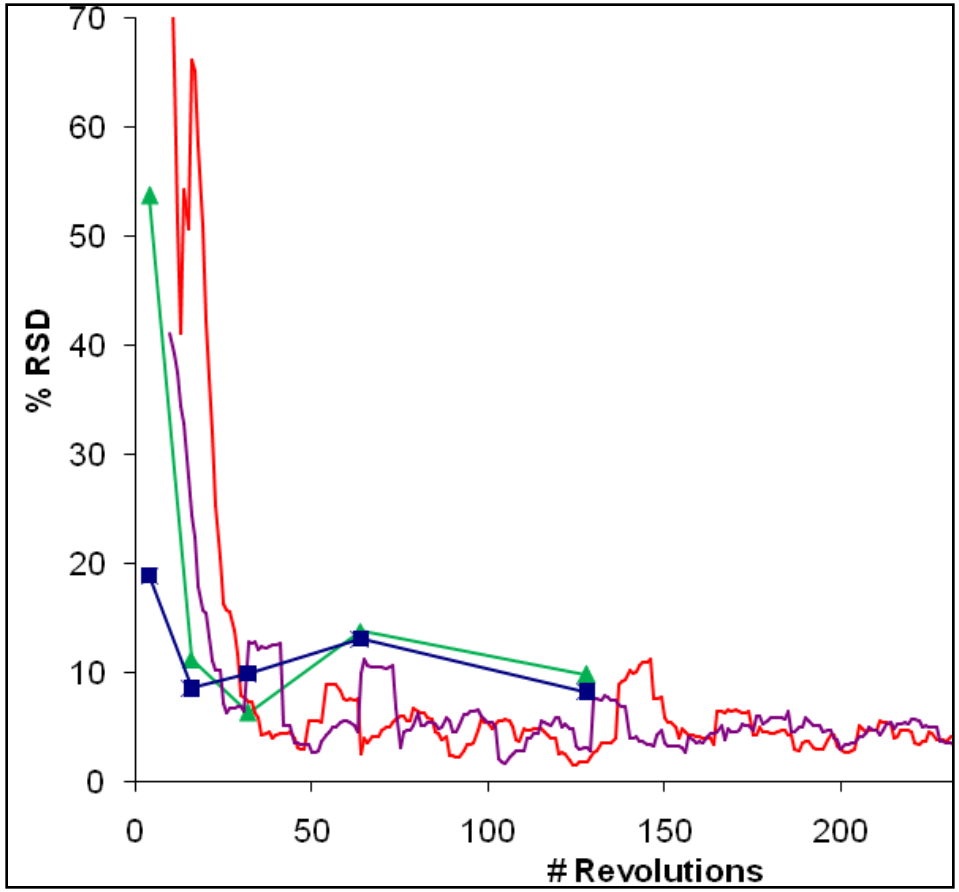
Comparison online vs. bench – RSD – Ingredient 1



Comparison online vs. bench – Mean – Ingredient 2



Comparison online vs. bench – RSD – Ingredient 2



So what is the right method?

- Proposal 1: Rational approach for developing and validating methods
  - *Use stratified sampling to determine mixing time, ensure that there is no segregation, rule out agglomeration*
  - *If company wants to use thief sampling, then use stratified sampling to validate thief sampling, ensure there is no sampling bias, etc.*
  - *If company wants to use PAT, use stratified sampling to validate PAT method, ensure there is no sensor bias, etc.*
  - *Companies should provide sound rationale for selection*
  - *Agency should allow flexibility*

So what is the right method?

- Proposal 2: Study group to determine method selection criteria and acceptance criteria
  - *Convene a group of industry, agency, and academia to review the field and harvest new knowledge*
  - *Determine criteria for selecting a method*
  - *Determine proper AC for each method*
  - *Work out the statistics so that all methods provide equivalent assurance*