Methodology to Determine the Particle Size Distribution, Mean and Standard Deviation from Sieve Data

Kelley St. Charles, Consultant to R.J. Reynolds Tobacco Company, Winston-Salem, NC
Walter T. Morgan, RAI Services, Winston-Salem, NC

Tobacco Science Research Conference, Leesburg, VA, Sept. 15-18, 2019

Introduction

- One of FDA's Frequent Deficiencies found in SE reports for smokeless tobacco is particle size information*
- Six of the 8 OPTIONS list particle size distribution
  - Portioned moist snuff, loose snus, portioned snus, loose dry snuff, dissolvables, portioned chewing tobacco
  - D10, D50, D90 (10%, 50% and 90% of particles smaller than particle size indicated)
  - Median, mean and mode values
- Two of the 8 OPTIONS (loose chewing tobacco, loose moist snuff) list particle size range
- Examples will focus on smokeless tobacco but can be applied to any measured distribution or truncated data
  - A standard methodology for soil science characterization and aggregate usage
  - Truncated data mean that the complete range is not characterized

* Common issues found in Substantial Equivalence Reports | FDA
https://www.fda.gov/tobacco-products/substantial-equivalence/common-issues-found-substantial-equivalence-reports-0

Introduction

- Standard sieves are commonly used to determine particle size distribution of solids down to ~ 20 µm
- One of the oldest physical analytical methods
  - Used in Ancient Egypt to grade grain harvests
  - Tyler Standard Sieve series introduced in 1910 followed by Ro-Tap Sieve Shaker in 1914.
- Choice of sieves is important
  - Ideally spans the majority of particle size range without having an excessive number
  - Spacing too small wastes analyst time and can promote weighing errors

Sieve Choice - 3 Examples

Excessive number of Sieves

Sieve Screen	% of Total Mass Mean	%RSD
10	0.68	97.1
12	0.50	70.0
14	1.22	59.0
16	2.26	36.3
18	4.23	28.6
20	11.08	14.9
30	30.12	4.3
40	33.95	6.7
50	13.18	19.7
60	2.05	25.9
70	0.28	50.0
80	0.22	59.1
Pan	0.22	54.5

Ideal
(From Excessive Data)

Sieve Screen	% of Total Mass
16	3.95
20	15.31
30	30.12
40	33.95
50	13.18
Pan	2.77

Missed range
~ ½ on top sieve
(Different sample)

Sieve Screen	% of Total mass
18	46.2
20	11.7
30	22.5
35	7.6
40	5.3
45	3.4
Pan	3.2

What to do with sieve data?

- Convert data to information
  - Presentation covers methodology
  - Presentation does not cover sample comparison methods
- Sieve screens descriptors translated to size units
- Data then converted to a distribution
  - Type of distribution (e.g. normal, log-normal)
  - Determine statistics (mean, standard deviation, percentiles)
- Can use even missed range data to optimize sieve sizes

Translation to size units

ASTM E-11 U.S. Standard	Basic opening	Units	ASTM E-11 U.S. Standard	Basic opening	Units
No. 4	4.75	mm	No. 50	300	µm
No. 5	4.00	mm	No. 60	250	µm
No. 6	3.35	mm	No. 70	212	µm
No. 7	2.80	mm	No. 80	180	µm
No. 8	2.36	mm	No. 100	150	µm
No. 10	2.00	mm	No. 120	125	µm
No. 12	1.70	mm	No. 140	106	µm
No. 14	1.40	mm	No. 170	90	µm
No. 16	1.18	mm	No. 200	75	µm
No. 18	1.00	mm	No. 230	63	µm
No. 20	850	µm	No. 270	53	µm
No. 25	710	µm	No. 325	45	µm
No. 30	600	µm	No. 400	38	µm
No. 35	500	µm	No. 450	32	µm
No. 40	425	µm	No. 500	25	µm
No. 45	355	µm	No. 635	20	µm

ASTM E-11 compatible with ISO

Sieve sizes increase by the fourth (1/4) root of 2 ≈ 1.19

Every 4th sieve is 2x (or ½) the size

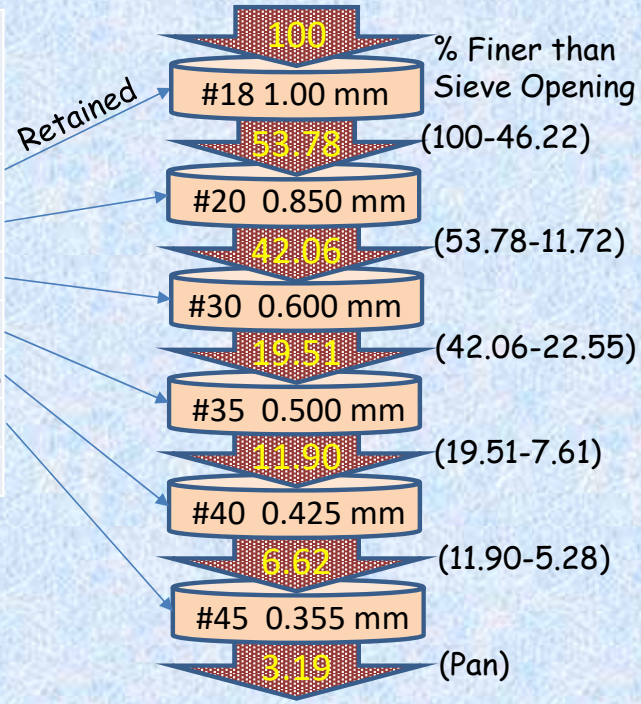
Bar/line charts of % mass vs sieve size are non-linear on X axis

Cumulative Distribution

Add size

Missed Range Data Set

Opening (mm)	Sieve screen	% of total mass retained
1.00	18	46.22
0.850	20	11.72
0.600	30	22.55
0.500	35	7.61
0.425	40	5.28
0.355	45	3.43
	pan	3.19



Cumulative Distribution

Zero to	Cumulative % of total mass
0.355	3.19
0.425	6.62
0.500	11.90
0.600	19.51
0.850	42.06
1.00	53.78

Methodology using simulated spreadsheet

Setup to calculate Mean and Standard Deviation

Data and Cumulative **Milled Snuff** Data

	A	B	C
1	Data		Milled
2	Sieve	mm	Snuff
3	18	1.00	16.57%
4	20	0.850	18.10%
5	30	0.600	33.98%
6	35	0.500	13.63%
7	40	0.425	7.59%
8	45	0.355	4.82%
9	Pan		5.32%
10			
11	**Cumulative**	Zero to	
12	pan	0.355	5.32%
13	45	0.425	10.14%
14	40	0.500	17.73%
15	35	0.600	31.36%
16	30	0.850	65.34%
17	20	1.00	83.44%
18	18	>1	100.01%

3 Rep Std Dev
2.96%
0.92%
0.49%
0.72%
1.28%
0.78%
1.17%

Pooled Std Dev 1.41%

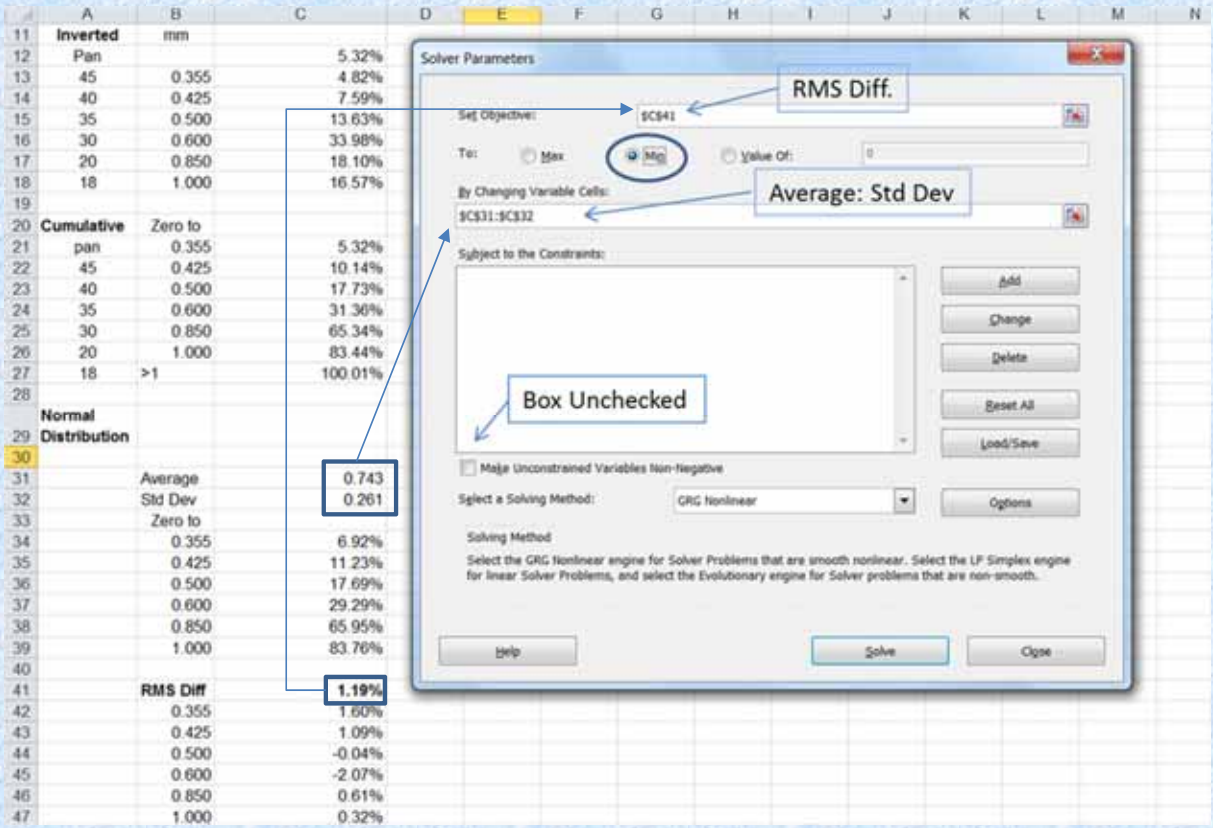
	B	C	C Formula
29			**Normal Distribution**
30			
31	**Mean**	**0.743**	Initial estimate from data ~0.75
32	**Std Dev**	**0.261**	Initial estimate of ~ 1/2-1/3 Mean
33	Zero to		**Cumulative calculation**
34	0.355	6.92%	=NORM.DIST($B34,C$31,C$32,TRUE)
35	0.425	11.23%	=NORM.DIST(Size, Mean, Std Dev, TRUE)
36	0.500	17.69%	(Copy to other cells)
37	0.600	29.29%	
38	0.850	65.95%	
39	1.00	83.76%	
40			**Root Mean Square Difference**
41	**RMS Diff**	**1.19%**	=SQRT(SUMSQ(C42:C47)/COUNT(C42:C47))
42	0.355	1.60%	=C34-C12 **(Calculated – Measured)**
43	0.425	1.09%	(Copy to other cells)
44	0.500	-0.04%	
45	0.600	-2.07%	
46	0.850	0.61%	
47	1.00	0.32%	

Procedure for Mean and Std. Dev.

- Vary Mean and Standard deviation values to minimize RMS difference
- Can be done manually
  - Vary Mean up and down and notice whether RMS difference goes up or down
  - Switch to Standard Deviation cell and do the same until RMS goes up regardless
  - Repeat with Mean, then Standard Deviation until changes are less than desired precision
- Can use Excel™ Solver function
  - Part of the spreadsheet Add-ins
  - Located in the Data tab after adding in

	B	C
29		
30		
31	Mean	0.743
32	Std Dev	0.261
33	Zero to	
34	0.355	6.92%
35	0.425	11.23%
36	0.500	17.69%
37	0.600	29.29%
38	0.850	65.95%
39	1.00	83.76%
40		
41	RMS Diff	1.19%
42	0.355	1.60%
43	0.425	1.09%
44	0.500	-0.04%
45	0.600	-2.07%
46	0.850	0.61%
47	1.00	0.32%

Solver Routine Screenshot



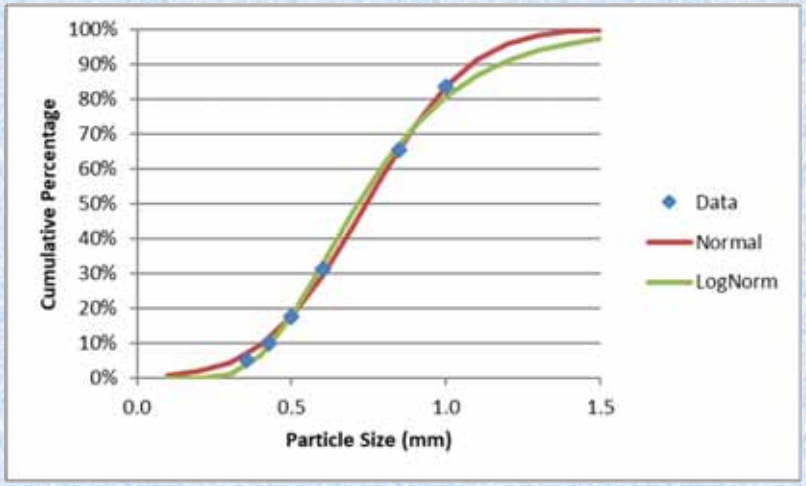
	A	B	C
11	**Inverted**	mm	
12	Pan		5.32%
13	45	0.355	4.82%
14	40	0.425	7.59%
15	35	0.500	13.63%
16	30	0.600	33.98%
17	20	0.850	18.10%
18	18	1.000	16.57%
19			
20	**Cumulative**	Zero to	
21	pan	0.355	5.32%
22	45	0.425	10.14%
23	40	0.500	17.73%
24	35	0.600	31.36%
25	30	0.850	65.34%
26	20	1.000	83.44%
27	18	>1	100.01%
28			
29	**Normal Distribution**		
30			
31		Average	0.743
32		Std Dev	0.261
33		Zero to	
34		0.355	6.92%
35		0.425	11.23%
36		0.500	17.69%
37		0.600	29.29%
38		0.850	65.95%
39		1.000	83.76%
40			
41		**RMS Diff**	**1.19%**
42		0.355	1.60%
43		0.425	1.09%
44		0.500	-0.04%
45		0.600	-2.07%
46		0.850	0.61%
47		1.000	0.32%

Procedure

- Check for lognormal distribution
  - Can use ln(sieve opening) and recalculate
  - Alternatively replace NORM.DIST function with LOGNORM.DIST function
- The mean value will be ln(mean particle size) so transform by EXP[ln(mean particle size)] to get mean particle size*
- Using same data set as before, Normal Distribution provides best fit

	Normal	Lognormal
Mean	0.743 mm	-0.335*
Std. Dev.	0.261 mm	0.387
RMS Diff.	1.19%	1.69%

* Lognormal Transform
Mean size = EXP(-0.355) = 0.715 mm
-1 SD = EXP (-0.355-0.387) = 0.486 mm
+1 SD = EXP (-0.355+0.387) = 1.05 mm



Other Smokeless Types

	Fine Cut Snuff	Long Cut Snuff	Milled Snuff	Dry Snuff
Normal				
Mean Particle Size (mm)	0.72	0.92	**0.74**	**0.11**
Std. Dev.	0.26	0.40	0.26	0.026
RMS Diff	2.93%	5.45%	1.19%	1.05%
Lognormal				
Mean	-0.364	-0.148	-0.335	-2.218
Std. Dev.	0.392	0.527	0.387	0.249
RMS Diff	1.00%	1.60%	1.69%	2.17%
Log Transform				
Mean Particle Size (mm)	**0.70**	**0.86**	0.72	0.11
-1 SD	0.47	0.51	0.49	0.085
+1 SD	1.03	1.46	1.05	0.140

Chi Squared Goodness of Fit Test

- Is Observed value significantly different from Predicted Value?
- Lower p-values (<0.05) imply Observed values NOT consistent with Predicted
- Test Statistic = Sum [(Observed – Predicted)^2 / Predicted]
- Degrees of Freedom = # of cells - 1 - # of Parameters Estimated
- Two parameters estimated (Mean and Std. Dev.)
- Degrees of Freedom = # of cells – 3

		Milled Snuff	Fine Cut Snuff	Dry Snuff
Normal	Chi Sq	1.19	6.59	2.76
	DF	4	4	2
	H0: Data are from the specified distribution			
	p-value	**0.880**	0.159	**0.252**
LogNormal	Chi Sq	3.27	1.09	13.91
	DF	4	4	2
	H0: Data are from the specified distribution			
	p-value	0.514	**0.895**	0.001

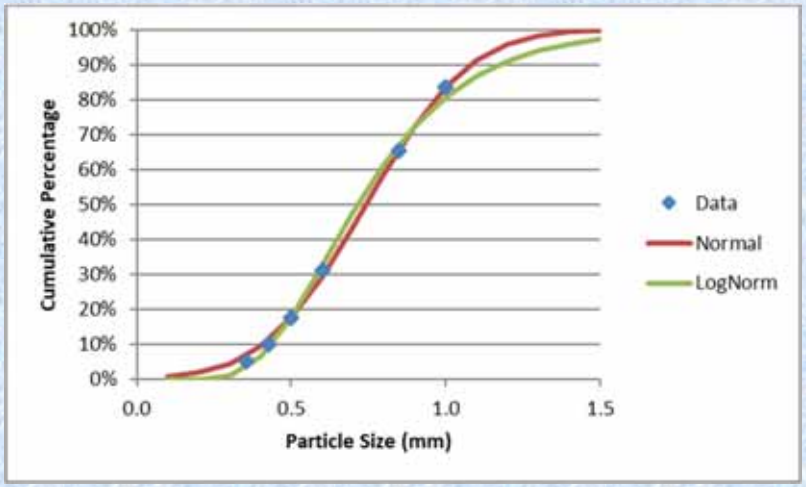
Providing information

- Distribution type, mean and standard deviation used to provide more information
- Calculation of smooth curves for graphical representation
- Percentiles especially useful
  - Providing D10, D50 and D90 values
  - Optimizing sieve selection
  - Providing descriptive information

Smooth curves for graphics

Previous Milled Snuff Results

	Normal	Lognormal
Mean	0.743 mm	-0.335
Std. Dev.	0.261 mm	0.387
RMS Diff.	1.19%	1.69%



Cumulative Distribution

=NORM.DIST(Size, Mean, Std Dev, TRUE) or

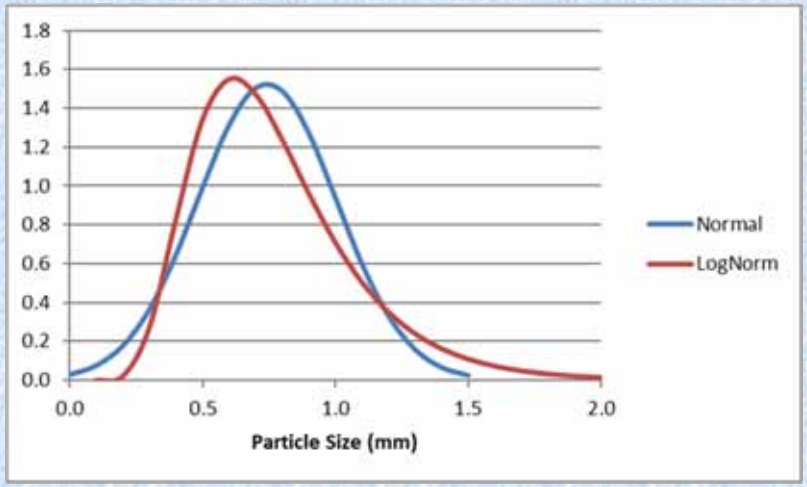
=LOGNORM.DIST(Size, Mean, Std Dev, TRUE)

Size (mm)	Normal	LogNormal
0.1	0.7%	0.00%
0.2	1.9%	0.00%
0.3	4.5%	1.2%
0.4	9.5%	6.6%
•	•	•
•	•	•
•	•	•
1.4	99.4%	95.9%
1.5	99.8%	97.2%

Smooth curves for graphics

Previous Milled Snuff Results

	Normal	Lognormal
Mean	0.743 mm	-0.335
Std. Dev.	0.261 mm	0.387
RMS Diff.	1.19%	1.69%



Non-cumulative Distribution

=NORM.DIST(Size, Mean, Std Dev, FALSE) or

=LOGNORM.DIST(Size, Mean, Std Dev, FALSE)

Size (mm)	Normal	LogNormal
0.1	0.075	0.000
0.2	0.177	0.023
0.3	0.364	0.275
0.4	0.645	0.833
•	•	•
•	•	•
•	•	•
1.4	0.065	0.163
1.5	0.023	0.110

Percentiles

Using Missed Range data set

	Normal	Lognormal
Mean	0.946 mm	-0.053
Std. Dev.	0.374 mm	0.537
RMS Diff.	1.85%	0.17%

Zero to	Cumulative % of total mass
0.355	3.19
0.425	6.62
0.500	11.90
0.600	19.51
0.850	42.06
1.00	53.78

	Percentile	Size (mm)
		Lognormal
	5%	0.39
D10	10%	0.48
	25%	0.66
D50	50%	0.95
	75%	1.36
D90	90%	1.89
	95%	2.29

=LOGNORM.INV(%,Mean, Std Dev)

Determine better sieve selection

U.S. Standard	Size (mm)	Cumu -lative	On Sieve
Pan	0.00		6.7%
No. 45	0.355		
No. 40	0.425	6.7%	22.7%
No. 35	0.500		
No. 30	0.600		
No. 25	0.710	29.5%	24.4%
No. 20	0.850		
No. 18	1.00	53.9%	22.7%
No. 16	1.18		
No. 14	1.40	76.6%	18.9%
No. 12	1.70		
No. 10	2.00		
No. 8	2.36	95.5%	4.5%

=LOGNORM.DIST(Size, Mean, Std Dev, TRUE)

Descriptors

- Used extensively in aggregates and soil science
  - USDA sand (0.05-2.0 mm), silt (0.002-0.05 mm), clay (<0.002 mm)
  - USDA sand: Classified into very coarse, coarse, medium, fine, very fine
- Assume the following for smokeless: Fine (<0.355 mm), Medium (0.355-0.500 mm), Coarse (0.500-1.00 mm), Large (>1.00 mm)
- Use same cumulative distribution calculations as with sieve selection but with 0.355, 0.500 and 1.00 mm size

Size (mm)	Cumu-lative	On Sieve	Size
Pan		5.32%	Fine
0.355	5.32%	12.37%	Medium
0.500	17.69%	66.07%	Coarse
1.00	83.76%	16.24%	Large

=NORM.DIST(Size, Mean, Std Dev, TRUE)

Previous Milled Snuff Results

	Normal	Lognormal
Mean	0.743 mm	-0.335*
Std. Dev.	0.261 mm	0.387
RMS Diff.	1.19%	1.69%

Conclusions

Summary

- Methodology described to calculate distribution type, mean and standard deviation from sieve data
- Both normal and lognormal distributions can be fit
- Results used to calculate and plot
  - Particle size distribution
  - Cumulative distribution
  - Percentiles
  - Descriptors
- Preliminary results can be used to optimize sieve selection
- Chi squared goodness of fit tests found observed values were consistent with predicted values

Thanks to

- R.J. Reynolds Tobacco Company Research and Development
- You for listening

Questions ?