

Annex D: Statistical Analysis Summary

This section describes the Round Robin Experimental Design Test Matrix data for use in estimating the Repeatability (r) and Reproducibility (R) of Volume (mL) at 120 seconds in the Engine Oil Gelation Test (EOGT), as well as providing the r and R estimates themselves. This section also summarizes the statistical analysis of other estimates provided in this Research Report.

A. Data

A plan for data to be generated via a Round Robin Experimental Design Test Matrix in which eight (8) test laboratories were to run two (2) tests each, on seven (7) different matrix oils (Oil F, Oil U, Oil E, Oil R, Oil K, Oil M, Oil P). A test is defined as the average of duplicate results (determinations) on an oil. Time order effects were designed into the Test Matrix.

Several laboratories re-started their portion of the Test Matrix due to issues with the developing test procedure and equipment. Test laboratory 5 went through several test iterations with their final data captured as Lab 5.2. Test laboratory 6 also went through several test iterations with their final data captured as Lab 6.2. The final data was generated over the period from March 17, 2025 to October 31, 2025.

Test laboratories 2 and 4 ran the test with a different Funnel (larger funnel size) than the other test laboratories. A statistically significant bias was detected with the larger funnel size and the data from test laboratories 2 and 4 was deemed invalid. Both laboratories 2 and 4 were given an opportunity to re-run their tests using the correct funnel size but declined. Therefore, there were only six (6) test laboratories with valid test data (6 Valid Labs).

All of the test data generated was captured and stored in an Excel Workbook created and maintained by the ASTM Test Monitoring Center (TMC) named ALL_EOGT_ILS_ReportForm_20251215. The data were transferred to another Excel Workbook named ResearchReportData_Feb_16_2026 which was analyzed by a team of ASTM Subcommittee B Data Analysts. Workbook data referenced in the section is available from the TMC.

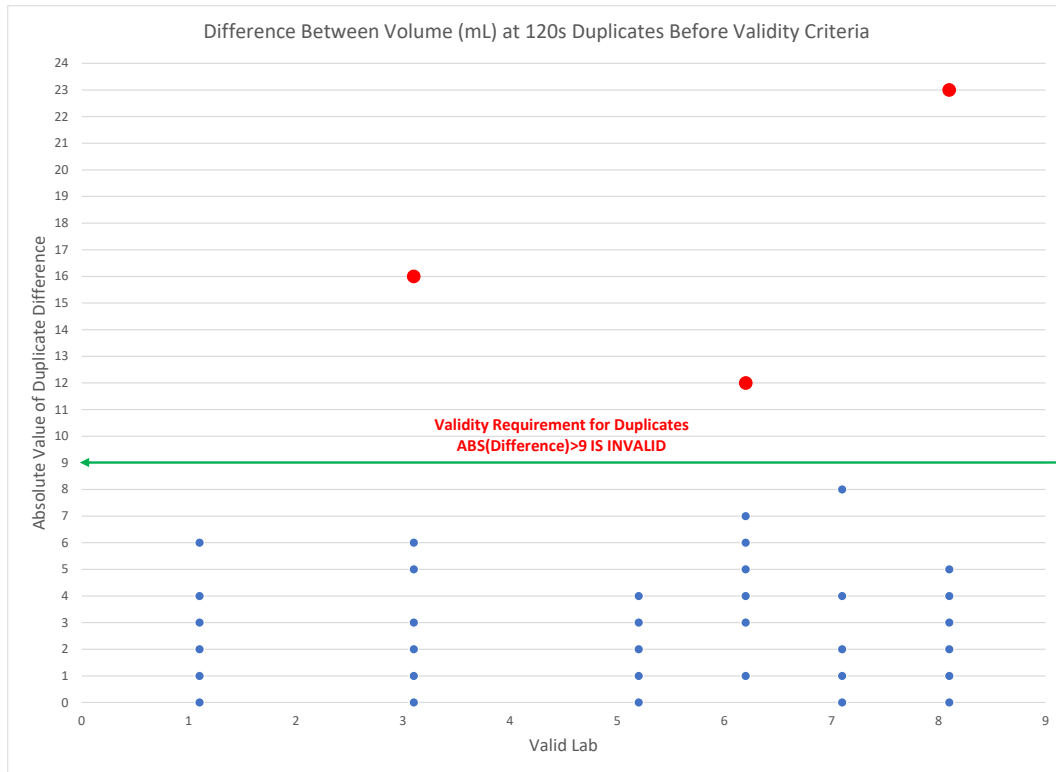
The evolution of the datasets from the ResearchReportData_Feb_16_2026 Workbook used in the data analysis is described below.

- 'All_Lab_Data' Tab
 - All Round Robin Data provided by ASTM TMC
- 'CorrectFunnelLab_Data' Tab
 - Data from Valid Labs with correct funnel size (mm) {<4.3}
- 'Lab_Data' Tab
 - Data from Valid Labs with correct funnel with Invalid Tests identified
- 'Lab_ValidData_Duplicates' Tab
 - Valid Data from Valid Labs
- 'RR_ValidData' Tab
 - Valid Test Results (Average of Duplicates) from Round Robin
- 'RR_ValidData_FM' Tab
 - Valid Test Results for Oils F and M

B. Validity Requirements for Round Robin Data

- The test must be run according to procedure
 - The **test result** (ASTM E2282) is the average test oil flow of the prepared test sample (prepared according to the test procedure) run in *duplicate*
 - A *duplicate* is equivalent to a **test determination** (ASTM E2282), which is the value of a characteristic or dimension of a single test specimen derived from one or more observed values. Each **observed value** (ASTM E2282) of the *duplicates* is from the same prepared single test specimen.
- Each of the two (2) *duplicates* used in computing the average test oil flow MUST reach 120 seconds before 70mL Volume is reached
- *Duplicate Criteria*: The absolute value of the difference between *duplicates* (**test determinations**) must be <=9

- The *Duplicate* Criteria is based upon a standard **statistical hypothesis test** (ASTM E2586) using a one-sided **type I error** rate of 1% based upon data in the 'Lab_Data' tab, that the difference between the duplicates is = 0. It is **NOT** an outlier criterion.
 - Null Hypothesis: (Duplicate 1 – Duplicate 2) = 0
 - The Root Squared Error of *duplicates* = 4.1243 based upon 98 samples
 - The **critical value** (ASTM E2586) for the **statistical hypothesis test** with 98 **degrees of freedom, df** (ASTM E2586), using the 1% upper tail of the Student's t-Distribution (value of 2.4) equals:
 - $0 + 2.4 * 4.1243 = 9.9$
 - Any *duplicate* difference beyond the critical value is number (truncated to 9) is concluded to violate the assumption that the difference in *duplicates* is randomly distributed around zero (0). The chance of being incorrect in such a conclusion is <2%, which is less than one case in twenty, and the test is declared invalid. Based on the criteria, the tests below are INVALID, not outliers, INVALID.
 - CMIR 195958, Lab 3.1, Oil F, INVALID for Failing Determinability Validity Criteria (2, 18)
 - CMIR 198097, Lab 8.1, Oil F, INVALID for Failing Determinability Validity Criteria (32, 55)
 - CMIR 201792 in Lab 6.2, Oil K, INVALID for Failing Determinability Validity Criteria (54, 66)
 - Based upon data in the 'RR_ValidData' tab
 - The Root Squared Error of *duplicates* = 2.8
 - This estimate on the valid data only is for informational purposes and does not impact the *Duplicate* Criteria



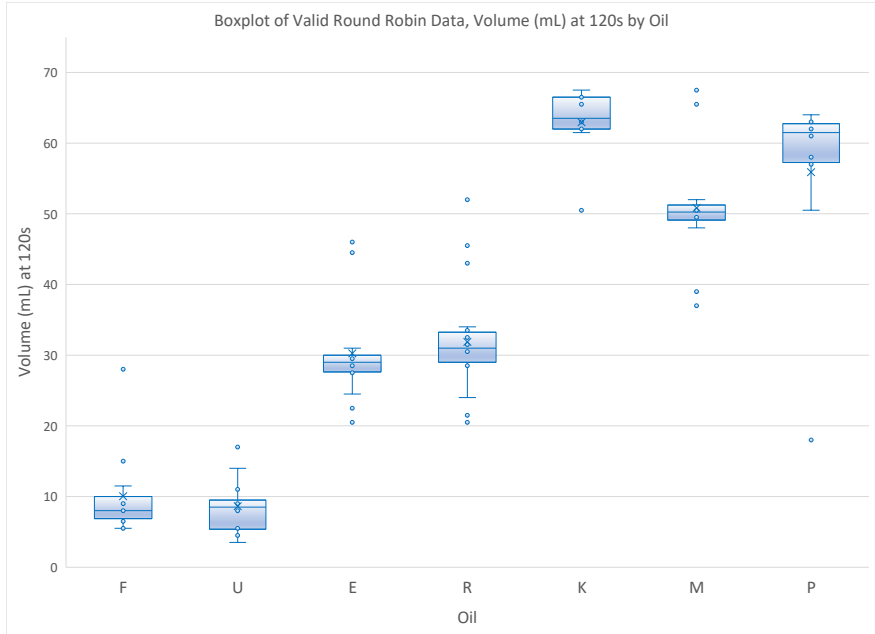
C. General Data Analysis of Valid Data

The test result (average of *duplicates*) of interest is the Oil Volume (mL) at 120 seconds (V120) in the test.

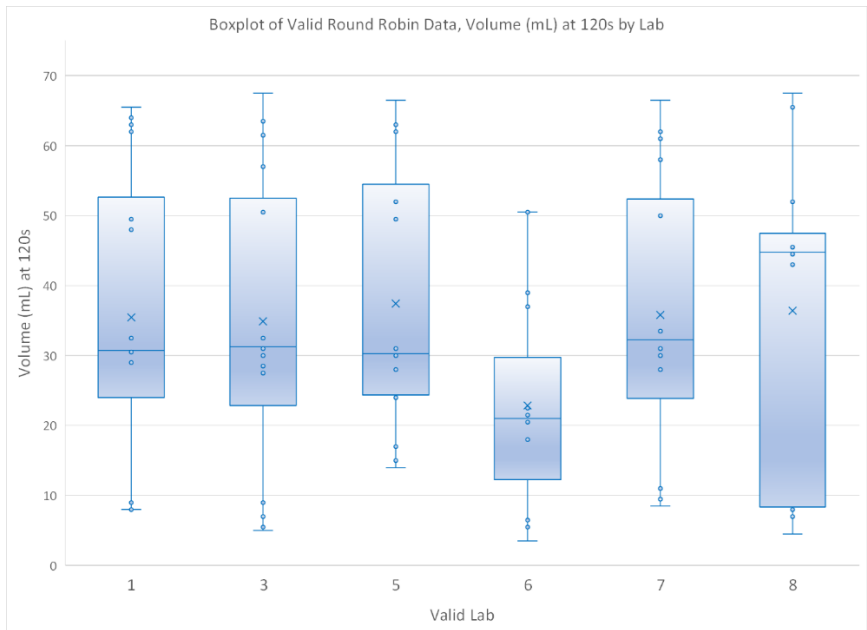
In general, Round Robin Test Oils differ in performance while Valid Labs have similar performance. Based upon a standard **statistical hypothesis test** (ASTM E2586) using a **type I error** rate of 1%, there is statistical evidence proving a **bias** (ASTM E177) among the Valid Labs depending upon the Round Robin Test Oil.

1. Plots Demonstrating Oil Discrimination and Possible Lab Bias

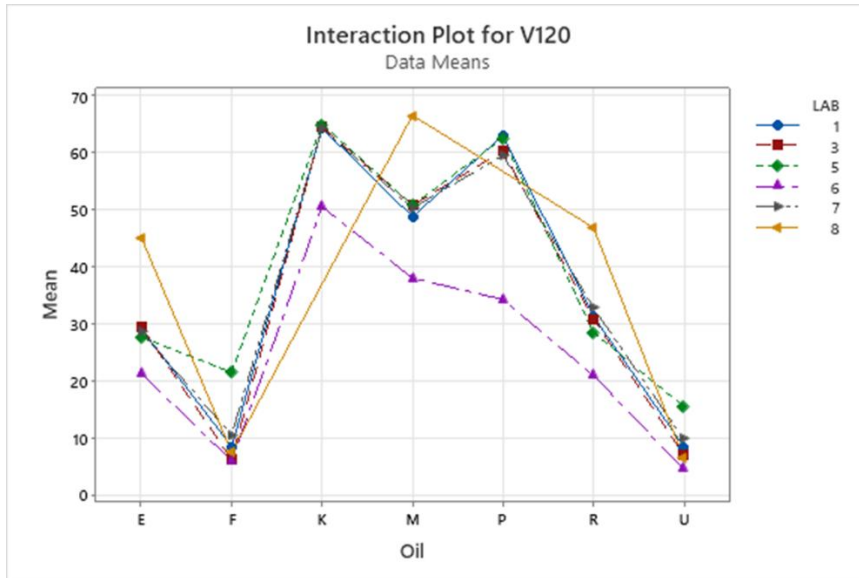
The plot below demonstrates that the Oils may be separated into three performance tiers with respect to Volume (mL) at 120 seconds. Low-tier (Oils F and U), mid-tier (Oils E and R), and high-tier (Oils K, M, and P). Note that all tests on Oils K and P in Lab 8 are invalid due to reaching 70 mL before 120 seconds and are not shown in this plot of valid data only. Also note that the risk of an invalid test for Oils K and P is high regardless of Valid Lab due to the high flow rate of the oil on average.



The plot below demonstrates that the Valid Labs differ but are similar in performance. Note that differences slightly depend upon the Oil tested which is illustrated in the interaction plot.



The interaction plot shows that while not egregious, the oil mean for each Oil does slightly depend upon the Valid Lab in which it is run.



2. General Linear Model Using All Oils

An additive General Linear Model (GLM) is fit using Minitab statistical software. Based upon a standard **statistical hypothesis test** (ASTM E2586) using a **type I error** rate of 1%, there is statistical evidence proving a **bias** (ASTM E177) among the Valid Labs depending upon the Round Robin Test Oil confirming the observations in the plots in the previous section.

GLM: $V120 = (\text{LAB}) + (\text{Oil within LAB}) + \text{error}$

The output from the analysis is provided below.

Factor Information

Factor	Type	Levels	Values
LAB	Fixed	6	1, 3, 5, 6, 7, 8
Oil(LAB)	Fixed	40	E(1), F(1), K(1), M(1), P(1), R(1), U(1), E(3), F(3), K(3), M(3), P(3), R(3), U(3), E(5), F(5), K(5), M(5), P(5), R(5), U(5), E(6), F(6), K(6), M(6), P(6), R(6), U(6), E(7), F(7), K(7), M(7), P(7), R(7), U(7), E(8), F(8), M(8), R(8), U(8)

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
LAB	5	1592.8	318.55	19.46	0.000
Oil(LAB)	34	33153.1	975.09	59.57	0.000
Error	51	834.8	16.37		
Total	90	36203.7			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
4.04569	97.69%	95.93%	*

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	34.448	0.437	78.83	0.000	
LAB					
1	1.766	0.947	1.87	0.068	1.52
3	1.135	0.947	1.20	0.236	1.52
5	4.325	0.947	4.57	0.000	1.52
6	-9.37	1.00	-9.33	0.000	1.66
7	2.087	0.947	2.20	0.032	1.52
Oil(LAB)					
E(1)	-7.05	2.23	-3.17	0.003	1.51
F(1)	-27.71	2.63	-10.55	0.000	1.69
K(1)	28.04	2.63	10.67	0.000	1.69
M(1)	12.54	2.63	4.77	0.000	1.69
P(1)	26.79	2.63	10.19	0.000	1.69
R(1)	-4.88	2.23	-2.19	0.033	1.51
E(3)	-6.08	2.23	-2.73	0.009	1.51
F(3)	-29.33	2.63	-11.16	0.000	1.69
K(3)	28.92	2.63	11.01	0.000	1.69
M(3)	15.17	2.63	5.77	0.000	1.69
P(3)	24.67	2.63	9.39	0.000	1.69
R(3)	-4.75	2.23	-2.13	0.038	1.51
E(5)	-11.11	2.23	-4.99	0.000	1.51
F(5)	-17.27	2.63	-6.57	0.000	1.69
K(5)	26.23	2.63	9.98	0.000	1.69
M(5)	11.98	2.63	4.56	0.000	1.69
P(5)	23.73	2.63	9.03	0.000	1.69
R(5)	-10.27	2.23	-4.62	0.000	1.51
E(6)	-3.75	2.26	-1.66	0.104	1.56
F(6)	-19.08	2.66	-7.18	0.000	1.73
K(6)	25.42	3.59	7.07	0.000	2.36
M(6)	12.92	2.66	4.86	0.000	1.73
P(6)	9.17	2.66	3.45	0.001	1.73
R(6)	-4.08	2.26	-1.80	0.077	1.56
E(7)	-7.87	2.23	-3.54	0.001	1.51
F(7)	-26.04	2.63	-9.91	0.000	1.69
K(7)	27.71	2.63	10.55	0.000	1.69
M(7)	13.71	2.63	5.22	0.000	1.69
P(7)	22.96	2.63	8.74	0.000	1.69
R(7)	-3.70	2.23	-1.66	0.102	1.51
E(8)	10.67	2.17	4.92	0.000	1.43
F(8)	-27.00	2.52	-10.73	0.000	1.55
M(8)	32.00	2.52	12.72	0.000	1.55
R(8)	12.33	2.17	5.69	0.000	1.43

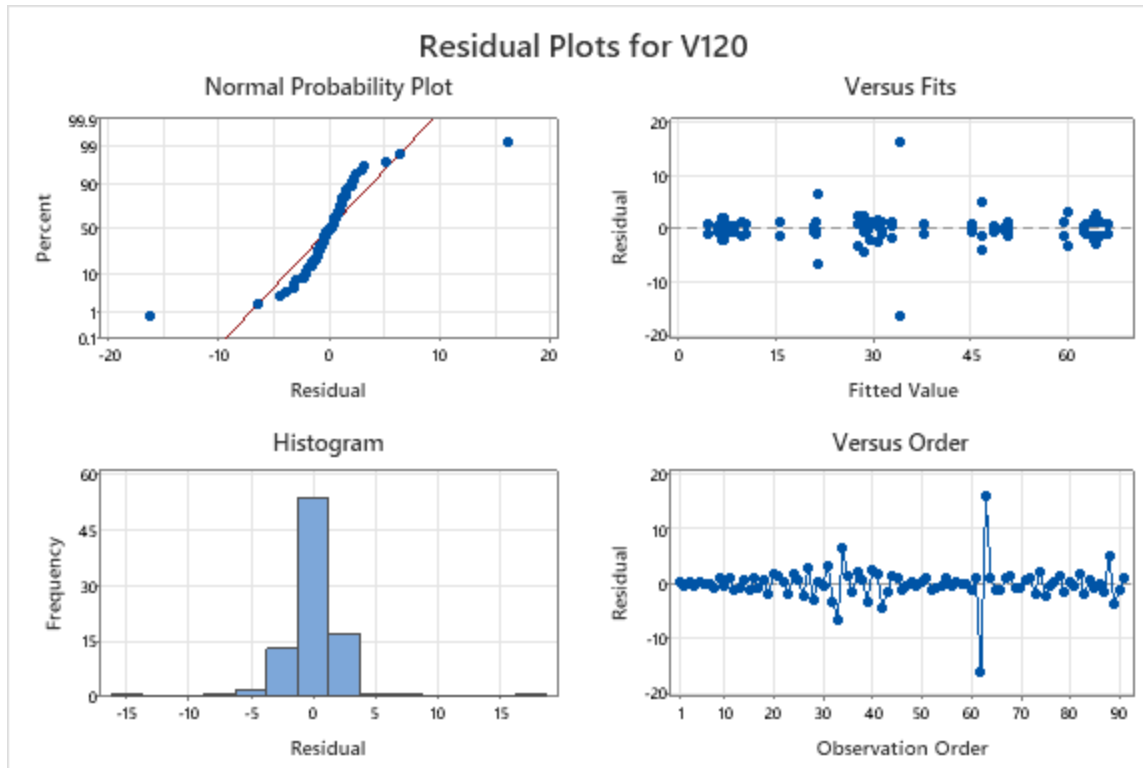
Fits and Diagnostics for Unusual Observations

Obs	V120	Fit	Resid	Std Resid
33	15.00	21.50	-6.50	-2.27 R
34	28.00	21.50	6.50	2.27 R
59	50.50	50.50	-0.00	* X
62	18.00	34.25	-16.25	-5.68 R
63	50.50	34.25	16.25	5.68 R

R Large residual
X Unusual X

Means

Term	Fitted Mean	SE Mean
LAB		
1	36.21	1.03
3	35.58	1.03
5	38.77	1.03
6	25.08	1.11
7	36.54	1.03
8	34.50	1.19
Oil(LAB)		
E(1)	29.17	2.34
F(1)	8.50	2.86
K(1)	64.25	2.86
M(1)	48.75	2.86
P(1)	63.00	2.86
R(1)	31.33	2.34
U(1)	8.50	2.86
E(3)	29.50	2.34
F(3)	6.25	2.86
K(3)	64.50	2.86
M(3)	50.75	2.86
P(3)	60.25	2.86
R(3)	30.83	2.34
U(3)	7.00	2.86
E(5)	27.67	2.34
F(5)	21.50	2.86
K(5)	65.00	2.86
M(5)	50.75	2.86
P(5)	62.50	2.86
R(5)	28.50	2.34
U(5)	15.50	2.86
E(6)	21.33	2.34
F(6)	6.00	2.86
K(6)	50.50	4.05
M(6)	38.00	2.86
P(6)	34.25	2.86
R(6)	21.00	2.34
U(6)	4.50	2.86
E(7)	28.67	2.34
F(7)	10.50	2.86
K(7)	64.25	2.86
M(7)	50.25	2.86
P(7)	59.50	2.86
R(7)	32.83	2.34
U(7)	9.75	2.86
E(8)	45.17	2.34
F(8)	7.50	2.86
M(8)	66.50	2.86
R(8)	46.83	2.34
U(8)	6.50	2.86



3. General Linear Model Using Reference Oils F and M

An additive General Linear Model (GLM) for the Reference Oils (Oil F and Oil M) is fit using Minitab statistical software. Based upon a standard **statistical hypothesis test** (ASTM E2586) using a **type I error** rate of 1%, there is statistical evidence proving a **bias** (ASTM E177) among the Valid Labs. In general, from the Tukey Pairwise Comparisons, Labs 5 and 8 tend to get higher results while Lab 6 tends to get lower results. In general, because the **bias** does depend in part on the Oil.

GLM: V120 = (LAB) + (Oil) + (LAB*Oil) + error

Factor Information

Factor	Type	Levels	Values
LAB	Fixed	6	1, 3, 5, 6, 7, 8
Oil	Fixed	2	F, M

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
LAB	5	614.6	122.92	15.11	0.000
Oil	1	9983.8	9983.76	1227.20	0.000
LAB*Oil	5	557.2	111.44	13.70	0.000
Error	12	97.6	8.14		
Total	23	11253.2			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
2.85227	99.13%	98.34%	96.53%

Grouping Information Using the Tukey Method and 95% Confidence

LAB	N	Mean	Grouping
8	4	37.000	A
5	4	36.125	A
7	4	30.375	A B
1	4	28.625	B C
3	4	28.500	B C
6	4	22.000	C

Means that do not share a letter are significantly different.

Grouping Information Using the Tukey Method and 95% Confidence

LAB*Oil N Mean Grouping

8 M	2	66.50	A
3 M	2	50.75	B
5 M	2	50.75	B
7 M	2	50.25	B
1 M	2	48.75	B C
6 M	2	38.00	C
5 F	2	21.50	D
7 F	2	10.50	D E
1 F	2	8.50	E
8 F	2	7.50	E
3 F	2	6.25	E
6 F	2	6.00	E

Means that do not share a letter are significantly different.

D. Repeatability (r) and Reproducibility (R) of Volume (mL) at 120 seconds

The estimates of **Repeatability (r)** and **Reproducibility (R)** (ASTM E177) of Volume (mL) at 120 seconds are based upon the valid data in tab 'RR_ValidData'. The estimates, along with the estimates of the mean (average) performance of each Oil are provided below and in tab 'rR'.

In the table below, s_r is the **Repeatability Standard Deviation** (ASTM E177) with r as the Repeatability, and s_R is the **Reproducibility Standard Deviation** (ASTM E177) with R as the Reproducibility. LS Mean is the Least Squares mean and is the estimate of the mean (average) performance of the Oil. Each Oil is run in each Valid Lab the same number of times except for Oil K. The Restricted Maximum Likelihood (REML) is provided for this unbalanced case.

Parameter	Target	Oil F	Oil U	Oil E	Oil R	Oil K	Oil K (REML)	Oil M	Oil P
Volume at 120s (ml)	s _r	3.8541	2.0867	1.5501	2.7412	2.9896	3.1368	1.1902	10.5475
Volume at 120s (ml)	s _R	6.4527	4.0891	8.0117	8.7264	5.3400	6.0818	9.1474	14.2923
Volume at 120s (ml)	r	10.90	5.90	4.38	7.75	8.46	8.87	3.37	29.83
Volume at 120s (ml)	R	18.25	11.57	22.66	24.68	15.10	17.20	25.87	40.42
Volume at 120s (ml)	LS Mean	10.04	8.63	30.25	31.89	61.70	62.01	50.83	55.90

The Repeatability (r) is interpreted as the maximum difference between two (2) test results on the same oil in the same test laboratory using the same operator. This maximum

difference is not expected to be exceeded in more than 5% of comparison cases between the two (2) test results.

The Reproducibility (R) is interpreted as the maximum difference between two (2) test results in two (2) different test laboratories. This maximum difference is not expected to be exceeded in more than 5% of comparison cases between the two (2) test results.

Note that each Oil differs in the estimates of r and R. Oil F is the oil chosen to represent the Repeatability (r) and Reproducibility (R) of Volume (mL) at 120 seconds for the Engine Oil Gelation Test (EOGT).

The statistical method used to estimate the **Repeatability (r)** and **Reproducibility (R)** (ASTM E177) of Volume (mL) at 120 s is based upon the Variance Components Table from the fit Mixed Effects Model where Lab is fit as a random effect. In case of balanced data, the Least Squares estimate is equivalent to the Restricted Maximum Likelihood (REML) estimate. The only unbalanced case is for Oil K, due to some tests declared as INVALID, so both LS and REML estimates are provided. The estimates are generated using Minitab statistical software. The Minitab output for Oil F and Oil M is provided below. The **Repeatability Standard Deviation** is represented by StDev for Error while the **Reproducibility Standard Deviation** is represented by StDev for Total (within plus between Lab Variance).

Oil F

Variance Components, using Adjusted SS

Source	Variance	% of Total	StDev	% of Total
LAB	26.7833	64.33%	5.17526	80.20%
Error	14.8542	35.67%	3.85411	59.73%
Total	41.6375		6.45271	

Oil M

Variance Components, using Adjusted SS

Source	Variance	% of Total	StDev	% of Total
LAB	82.2583	98.31%	9.06964	99.15%
Error	1.41667	1.69%	1.19024	13.01%
Total	83.675		9.14740	