Test Group Summary Report

C40B Petroleum Hydrocarbons in Water

January 2026 PT Round

Issued: March 2, 2026

PTC | Proficiency Testing Canada

Table of Contents

1.0 The Proficiency Testing Report 1

2.0 Definitions 1

3.0 Scoring System 1

3.1 Homogeneity and Stability Assessment 2

3.2 The z score 2

3.2 Composite (PT) Score 2

3.3 Identifying Bias 2

3.4 Deviations from Evaluation Procedure 3

4.0 PT Round Specific Data Summary 3

4.1 Summary statistics 3

4.2 *z* - Score Plots 3

4.3 kernel density plots 3

4.4 stability and homogeneity Plots 3

4.5 Box-and-Whisker Plots 3

4.6 Historic Comparison Plot 3

Annex A Summary by Analyte 4

1.0 The Proficiency Testing Report

The Proficiency Testing Report consists of two parts.

- *PTC Proficiency Testing Report*: This report contains participant-specific data and other confidential information. This report is emailed to participants at the end of the PT round.
- *Test Group Summary Report*: A Test Group Summary Report is created for each quantified test group at the end of the PT round. These reports contain more detailed information on the round than is found in the participant-specific PTC Proficiency Testing Report. These reports do not contain any confidential information and are made available on the PTC web site.

2.0 Definitions

The participant-specific PTC Proficiency Testing Report contains some terms that new participants may not be familiar with.

Code:	The registration code that is unique to each analyte that a participant is registered for.
App:	If a participant is accredited by CALA, this three-digit number is the appendix number that the accredited method is assigned to.
N:	The number of participants results that were used to calculate the summary statistics. This excludes qualified data (e.g., <) and any results that were flagged as outliers.
Assigned:	The Assigned Value is the robust mean of the reported results, outliers excluded. This is often referred to as the "target" value.
± u:	The uncertainty of the assigned value.
Reported:	The result reported by the participant.
s:	The Standard Deviation of Proficiency Assessment (SDPA). This value is used to determine the acceptance limits for the PT evaluation.
z-Score:	A value assigned to each reported result that is a measure of the degree to which it deviates from the Assigned Value.
Score:	The composite score of the four results reported for each analyte. It is normalized to a score out of 100.
Bias:	A flag assigned if bias is detected using the re-scaled *z*-score procedure.

3.0 Scoring System

Participant performance is evaluated for each proficiency testing sample by a quantitative method that is consistent with ISO/IEC 17043 - *Conformity assessment- General requirements for the competence of proficiency testing providers*, the *International Harmonized Protocol for Proficiency Testing of (Chemical) Analytical Laboratories* (2006), and ISO 13528:2015 *Statistical methods for use in proficiency testing by interlaboratory comparisons*.

The following is a brief description of the evaluation procedure used by PTC. The detailed evaluation procedure is described in PROC09 – PT Evaluation *Procedure*, which is available on the PTC website www.PTCanada.org).

3.1 HOMOGENEITY AND STABILITY ASSESSMENT

Homogeneity and stability are assessed using participant data. Regression analysis is performed on reported result against order of sample production (Homogeneity) and reported result against date of analysis (Stability). If the slope is significantly different than zero for either then the Standard Deviation of Proficiency Assessment (s) is increased to minimize the impact.

3.2 THE Z SCORE

A "z-score" is calculated for each reported result as follows:

$$z - Score = \frac{(x - \bar{X})}{SDPA}$$

where: x = participant result;
$\overline{X}$ = the Assigned Value;
SDPA = the Standard Deviation for Proficiency Assessment.

The assigned value $\overline{X}$ is generally estimated from the inter-laboratory Robust mean after outliers due to obvious gross errors (e.g., reported in wrong units) have been removed.

The Standard Deviation for Proficiency Assessment, s, is determined as follows:

- The inter-laboratory Robust standard deviation ($Stdev_{rob}$) is calculated using reported results, obvious outliers removed;
- The regression equation standard deviation ($Stdev_{reg}$) is estimated from regression equations derived from previous studies (see PROC11- *PT Regression Equations* for details);
- The SDPA is the higher of $Stdev_{rob}$ and $Stdev_{reg}$;
- When a laboratory reports its detection limit, s will be estimated using a pooled variance procedure that uses both the inter-laboratory data and the reported detection limit.

3.2 COMPOSITE (PT) SCORE

Since each PT round involves four or two separate samples of distinct concentration for each test, it is necessary to calculate a composite PT score for each test to determine overall performance. The composite score is calculated by first averaging the absolute *z*-scores for the four results and then calculating a final score as 100 + (-15 x avg |z|).

Acceptable PT Scores equal or exceed 70.

3.3 IDENTIFYING BIAS

The proficiency testing report provides flags for bias. These are determined using the re-scaled *z*-score procedure.

$$RSZ = \frac{\sum z}{\sqrt{N}}$$

where *z* = the *z*- score
N = the number of samples

Flags are assigned for each test group/parameter combination as follows:

RSZ $\geq$ -2 and $\leq$ 2 no flag assigned

RSZ > 2	H (High)
RSZ > 3	VH (Very High)
RSZ < -2	L (LOW)
RSZ < -3	VL (Very Low)

3.4 DEVIATIONS FROM EVALUATION PROCEDURE

Other than changes to the Standard Deviation of Proficiency Assessment due to homogeneity or stability flags, any deviation from the published evaluation procedure is described on the cover page(s) of the final *PTC Proficiency Testing Report*.

4.0 PT Round Specific Data Summary

The following pages provide more detailed information about the PT round indicated in the cover page of this report than is found in the participant-specific PTC Proficiency Testing Report. The graphical representations and the statistical summaries are based upon the data after outliers have been removed.

4.1 SUMMARY STATISTICS

In addition to some of the statistics found in the customer reports, this table includes additional summary statistics such as Median, different measures of dispersion, the number of outliers removed, the number of results in the Questionable range (|z| between 2 and 3) and the Unacceptable range (z > 3), and whether a data set was flagged for Homogeneity or Stability. This section also includes sorted scatter plots of the data for each sample.

4.2 *z* - SCORE PLOTS

The z -scores for each sample are ranked in increasing order and plotted. When the data is normally distributed, the plot should show a slight sigmoidal curve, with an equal number of points above zero as below. Each bar in these plots is colour-coded to indicate the analytical method used by the participant.

4.3 KERNEL DENSITY PLOTS

Kernel density plots are generated for each data set. These plots are a graphical way to represent the overall data distribution and are used to visualize possible deviations from normality and unimodality.

4.4 STABILITY AND HOMOGENEITY PLOTS

Plots of reported result against analysis date, and reported result against order of bottling are displayed, along with the regression line. These regression analyses are used to determine if the SDPA should be adjusted due to homogeneity or stability.

4.5 BOX-AND-WHISKER PLOTS

Box-and-Whisker plots are another way to display the distribution of the data. The box denotes the first and third quartile and the whiskers are the 5th and 95th percentile.

4.6 HISTORIC COMPARISON PLOT

The Historic Comparison Plot is a plot of robust mean against robust standard deviation for the previous ten PT rounds as well as the current PT round. This plot can be used to identify possible changes in the sample formulation.

F2: C10-C16

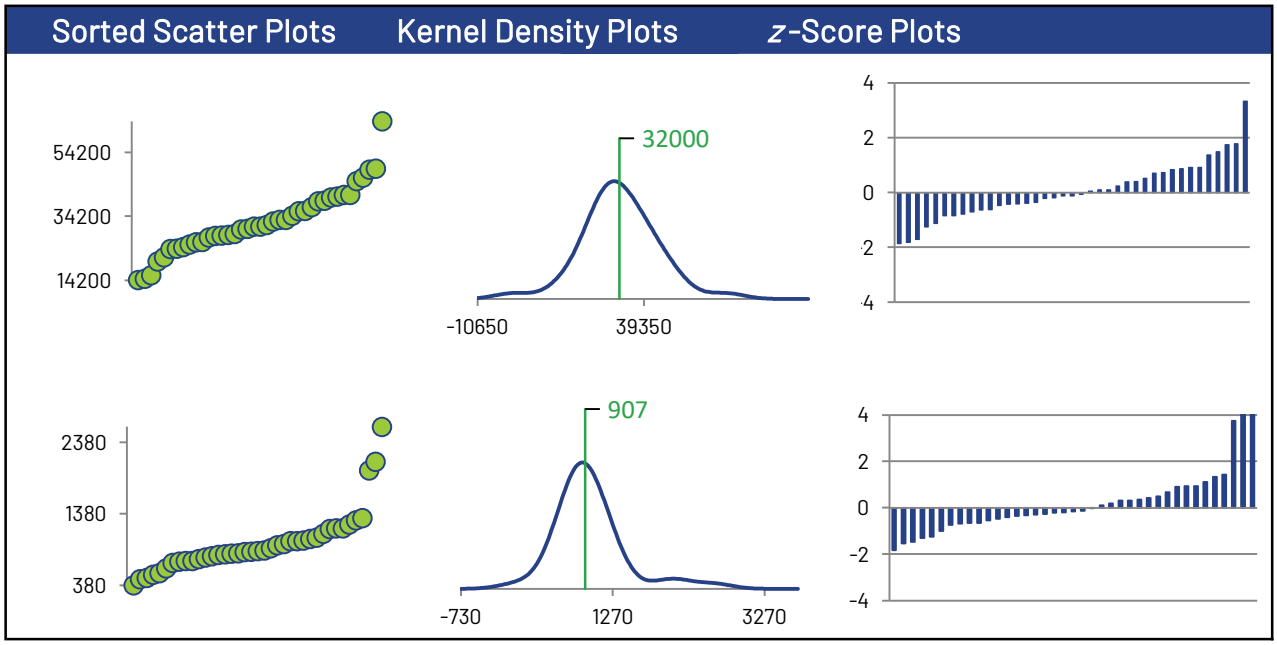
Summary Statistics

Statistic	C40B-1	C40B-2	C40B-3	C40B-4
No. of results used for summary statistics	39	39	39	38
Median µg/L	30900	860	16000	17500
Robust Mean µg/L	32000	907	15600	18100
Uncdrtainty of the Assigned Value µg/L	1910	57.8	837	1030
Robust Standard Deviation µg/L	9550	289	4180	5100
Regression Standard Deviation µg/L	4810	136	2340	2720
Stability Flag				
Homogeneity Flag				
Standard Deviation Used (SDPA) µg/L	9550	289	4180	5100
Flagged Outliers	0	0	0	1
|z|>3.0	1	3	0	0
2<|z|<3	0	0	2	0
Failure Rate %*	10.3			

Methods Used

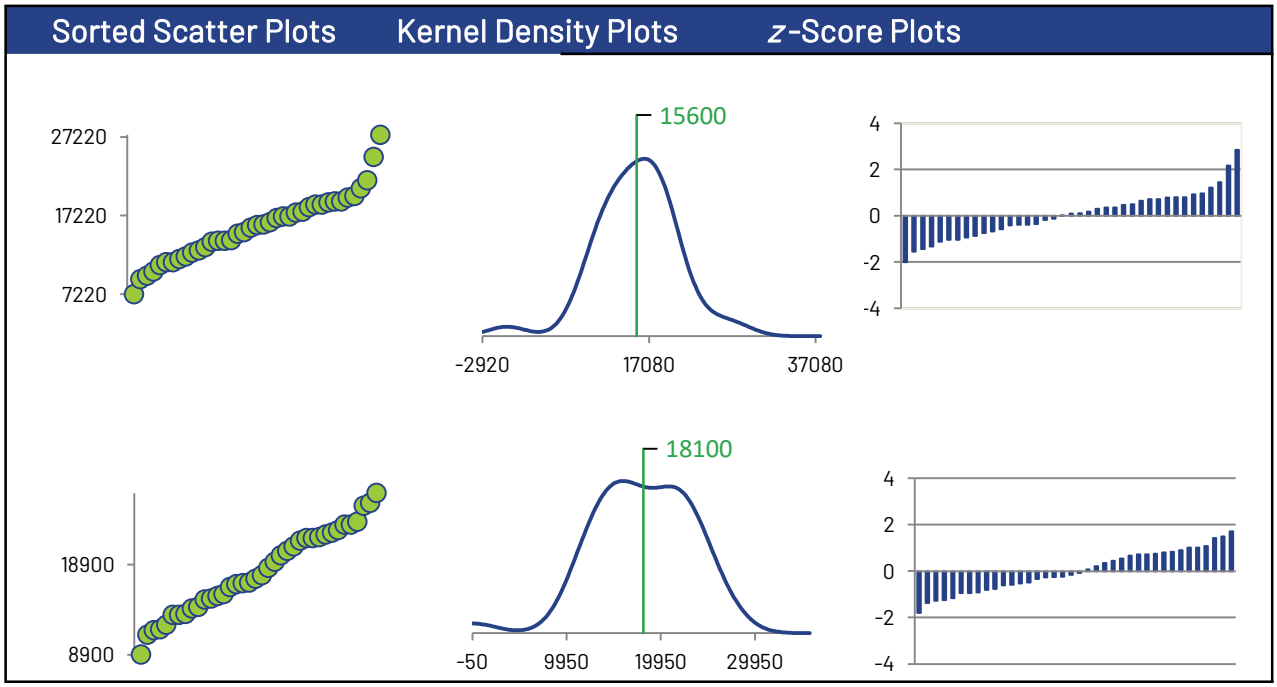
Method	C40B-1	C40B-2	C40B-3	C40B-4
GC/FID (Blue)	39	39	39	38

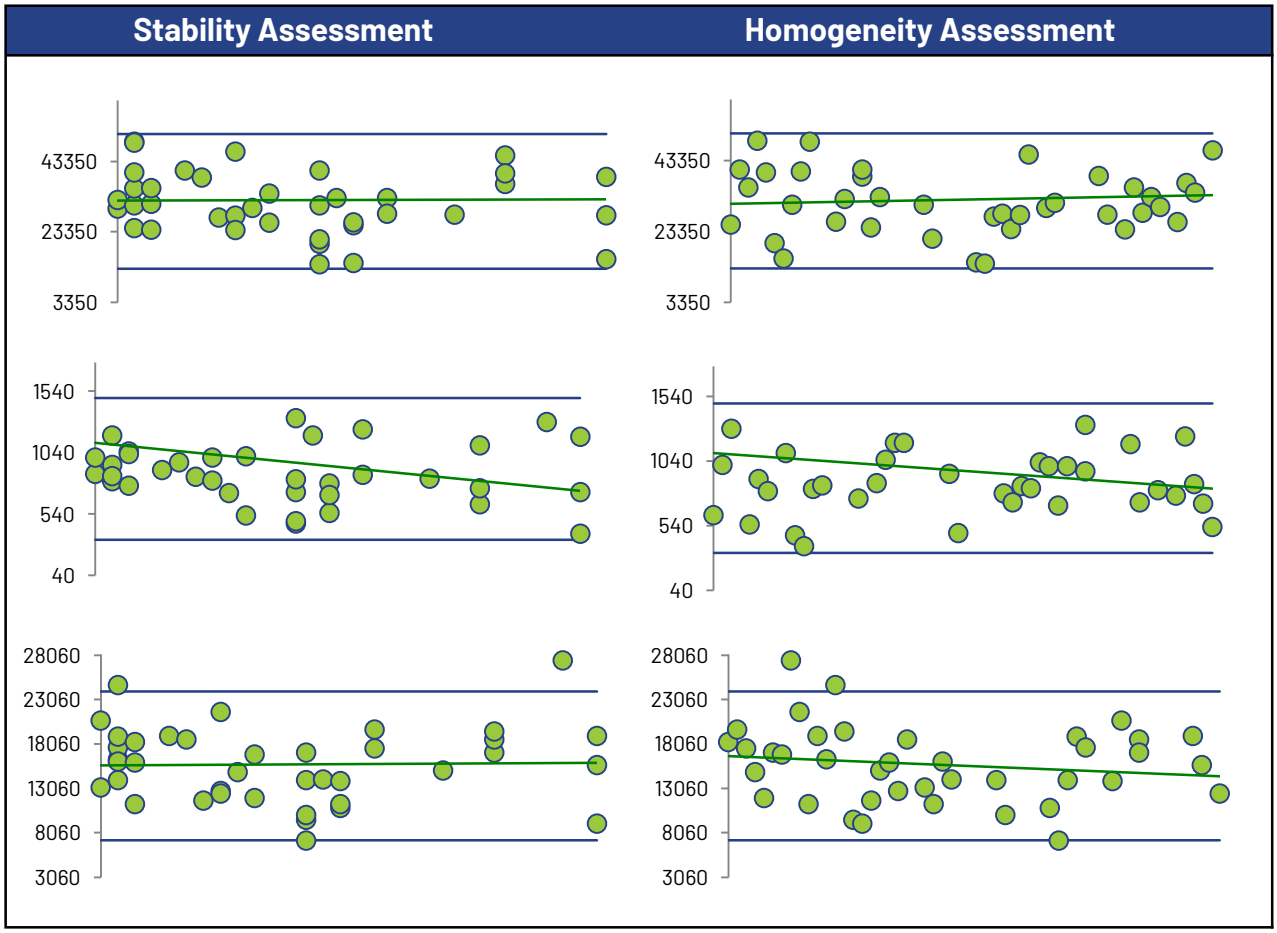
All summary stats, except the failure rate, and the plots below are based on the data excluding any flagged outliers



* Displayed failure rate does not account for the RDL modified z-Scores or inaccurately reported non-detects.

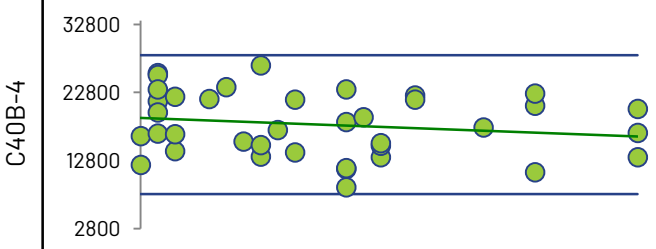
F2: C10-C16


Sorted Scatter Plots
Kernel Density Plots
z-Score Plots
C40B-3
27220
17220
7220
15600
-2920
17080
37080
4
2
0
-2
-4
C40B-4
18900
8900
18100
-50
9950
19950
29950

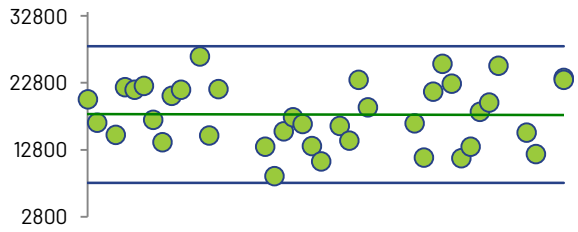

Stability Assessment
Homogeneity Assessment
C40B-1
43350
23350
3350
C40B-2
1540
1040
540
40
C40B-3
28060
23060
18060
13060
8060
3060

F2: C10-C16

Stability Assessment



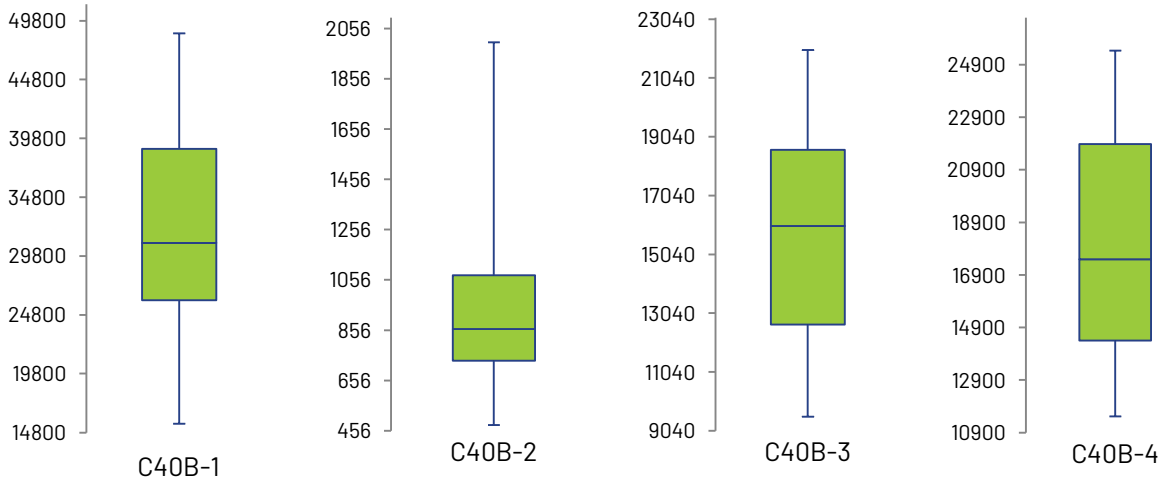
Homogeneity Assessment



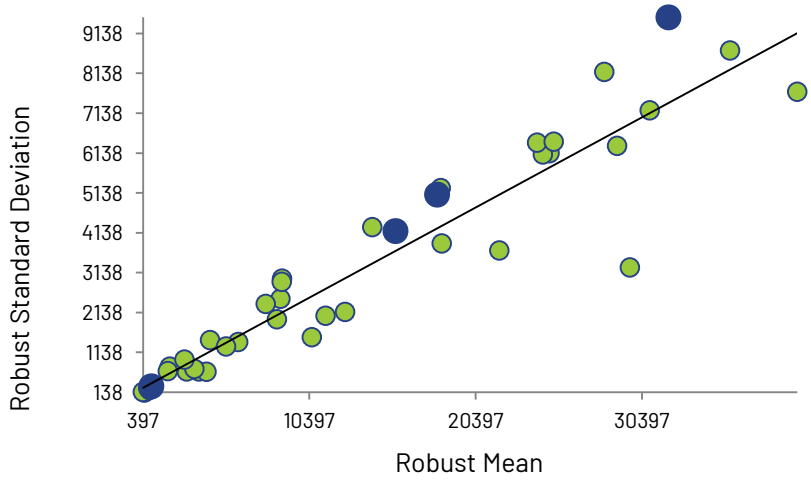
Stability assessments are regression analysis of reported result against date of analysis.

Homogeneity assessments are regression analysis of reported result against bottling order.

Box and Whisker Plots



Current Round (blue) Compared to Historic Rounds (green)



F3: C16-C34

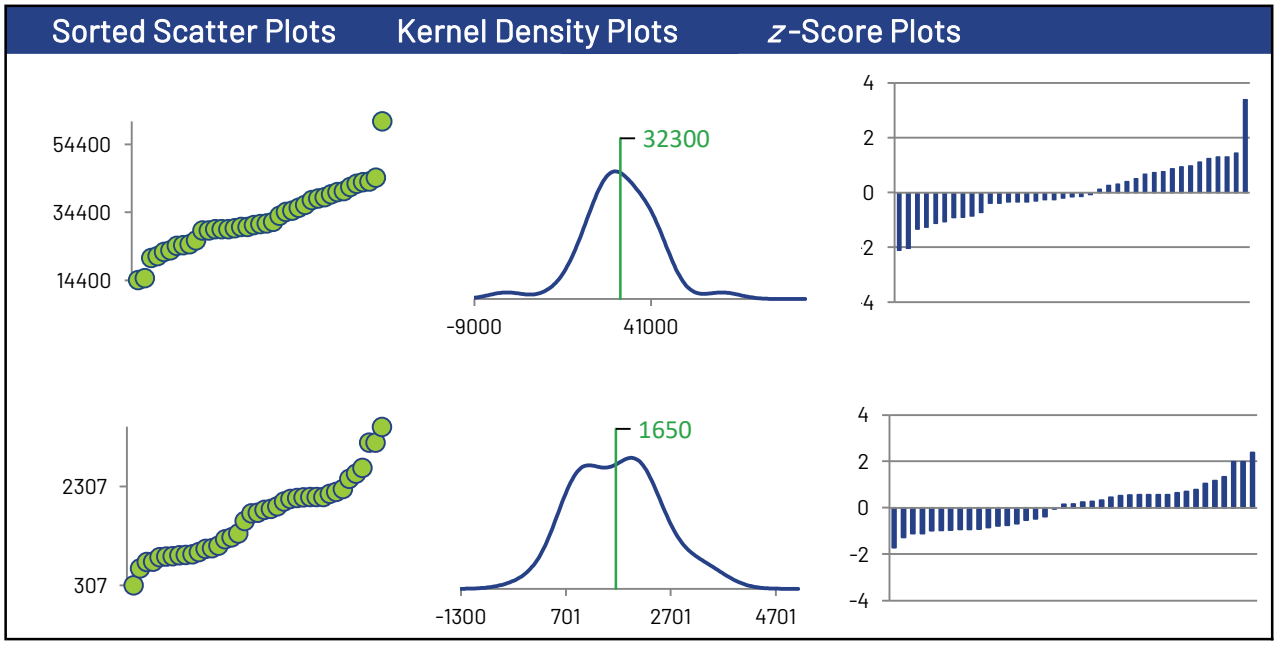
Summary Statistics

Statistic	C40B-1	C40B-2	C40B-3	C40B-4
No. of results used for summary statistics	39	39	39	38
Median µg/L	30900	1780	13500	22600
Robust Mean µg/L	32300	1650	14100	23000
Uncdrtainty of the Assigned Value µg/L	1700	157	739	1230
Robust Standard Deviation µg/L	8490	784	3690	6070
Regression Standard Deviation µg/L	4850	248	2120	3450
Stability Flag				
Homogeneity Flag				
Standard Deviation Used (SDPA) µg/L	8490	784	3690	6070
Flagged Outliers	0	0	0	1
|z|>3.0	1	0	1	0
2<|z|<3	2	1	1	1
Failure Rate %*	2.6			

Methods Used

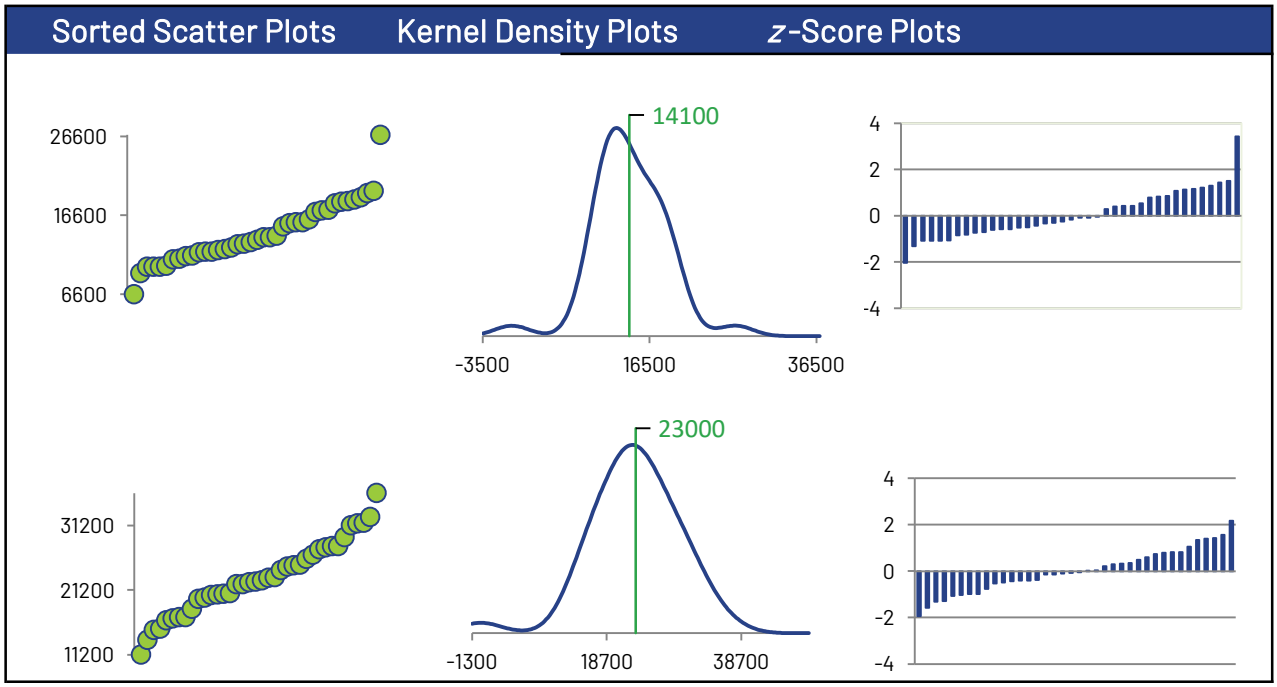
Method	C40B-1	C40B-2	C40B-3	C40B-4
GC/FID (Blue)	39	39	39	38

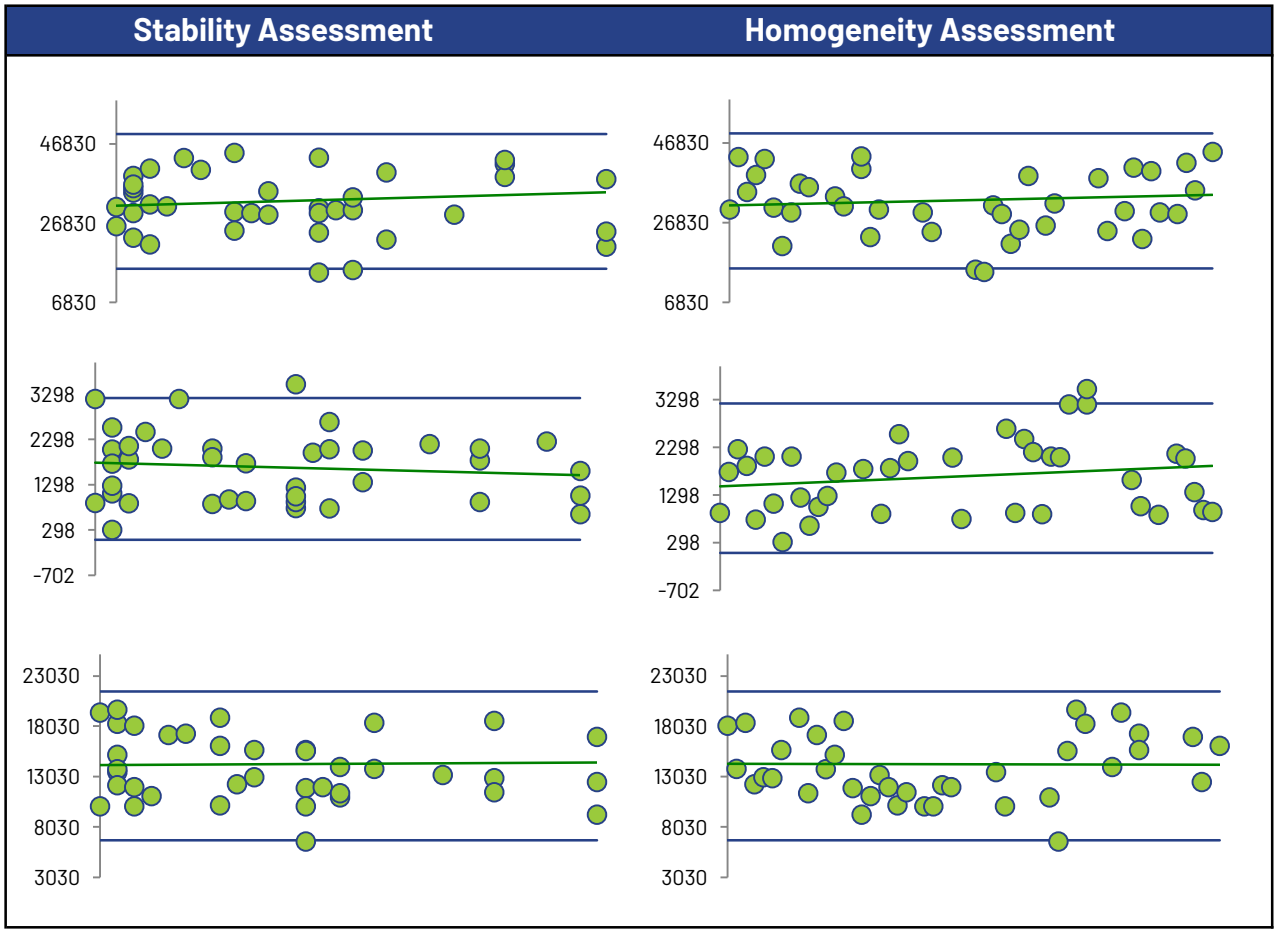
All summary stats, except the failure rate, and the plots below are based on the data excluding any flagged outliers



* Displayed failure rate does not account for the RDL modified z-Scores or inaccurately reported non-detects.

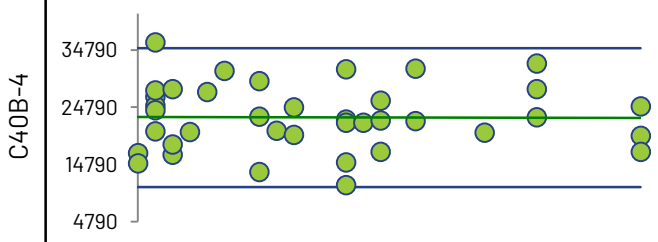
F3: C16-C34


Sorted Scatter Plots
Kernel Density Plots
z-Score Plots
C40B-3
26600
16600
6600
14100
-3500
16500
36500
4
2
0
-2
-4
C40B-4
31200
21200
11200
23000
-1300
18700
38700
4
2
0
-2
-4

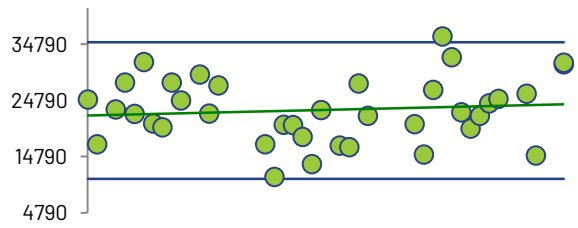

Stability Assessment
Homogeneity Assessment
C40B-1
46830
26830
6830
46830
26830
6830
C40B-2
3298
2298
1298
298
-702
3298
2298
1298
298
-702
C40B-3
23030
18030
13030
8030
3030
23030
18030
13030
8030
3030

F3: C16-C34

Stability Assessment



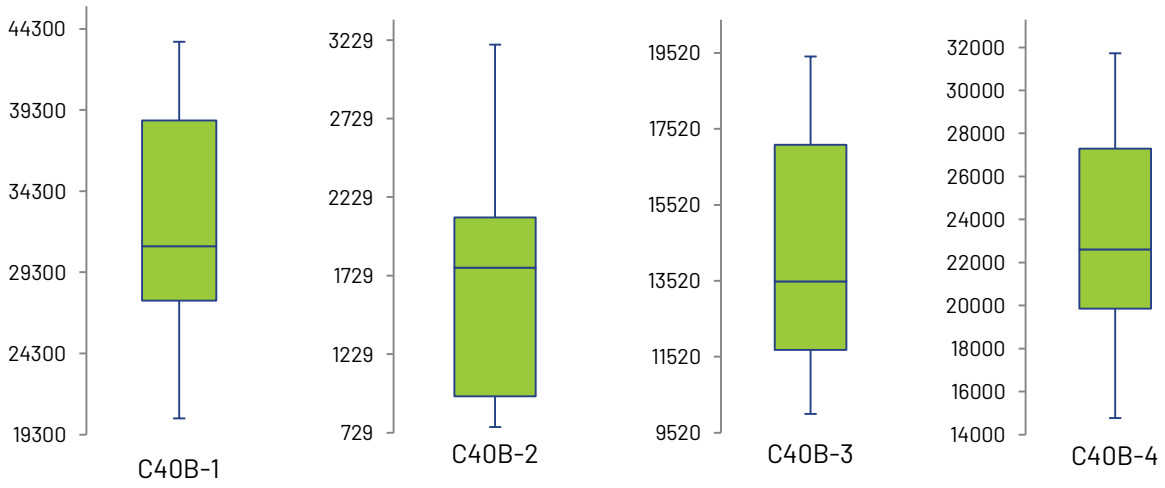
Homogeneity Assessment



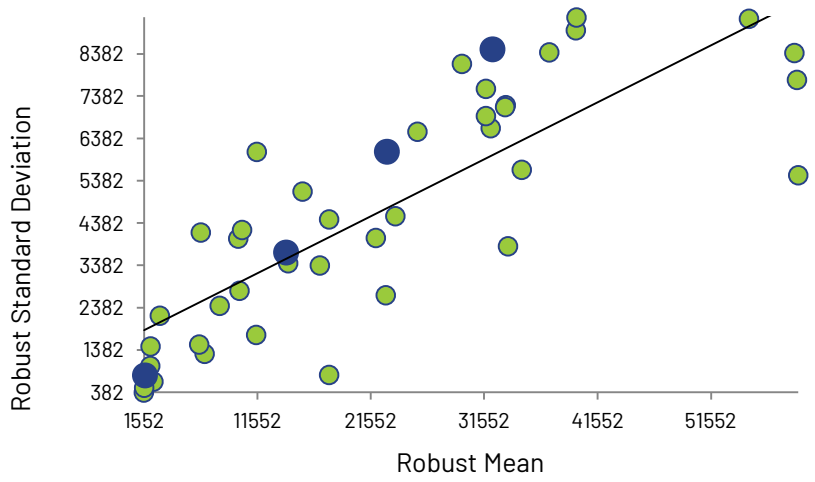
Stability assessments are regression analysis of reported result against date of analysis.

Homogeneity assessments are regression analysis of reported result against bottling order.

Box and Whisker Plots



Current Round (blue) Compared to Historic Rounds (green)



F4: C34-C50

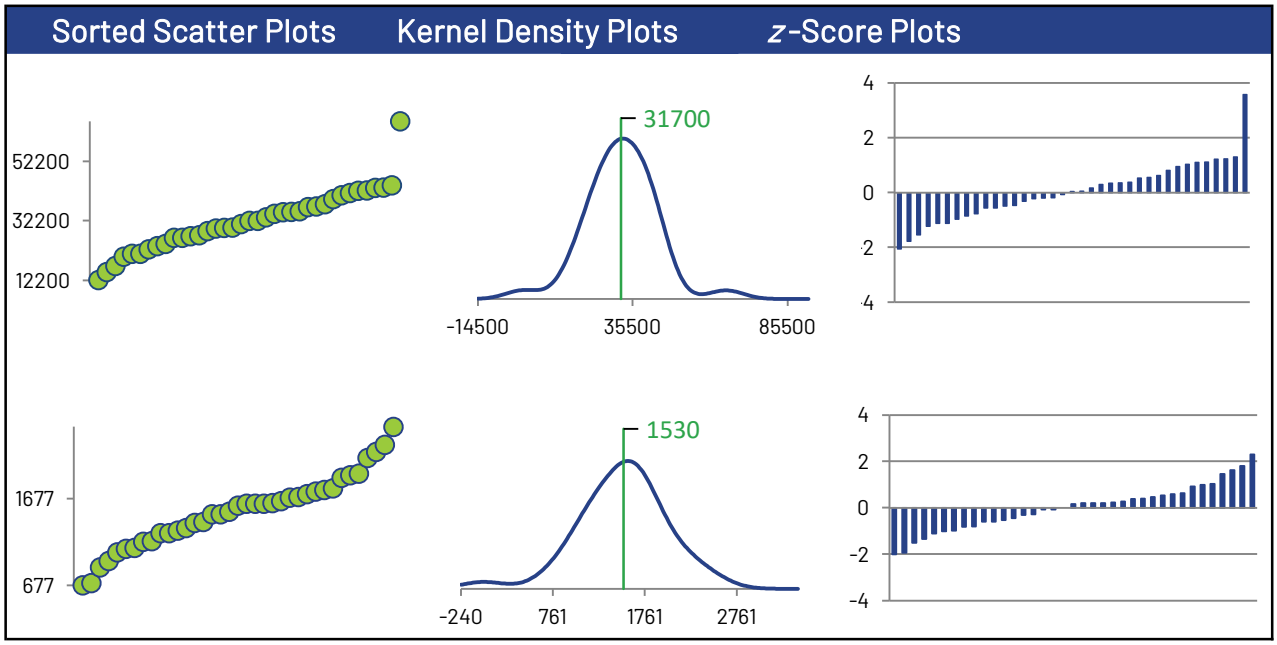
Summary Statistics

Statistic	C40B-1	C40B-2	C40B-3	C40B-4
No. of results used for summary statistics	37	37	37	36
Median µg/L	32100	1600	11400	27100
Robust Mean µg/L	31700	1530	11600	25900
Uncdrtainty of the Assigned Value µg/L	1940	87.5	682	1570
Robust Standard Deviation µg/L	9460	426	3320	7540
Regression Standard Deviation µg/L	4750	229	1730	3890
Stability Flag				
Homogeneity Flag				
Standard Deviation Used (SDPA) µg/L	9460	426	3320	7540
Flagged Outliers	0	0	0	1
|z|>3.0	1	0	1	0
2<|z|<3	1	1	0	0
Failure Rate %*	2.7			

Methods Used

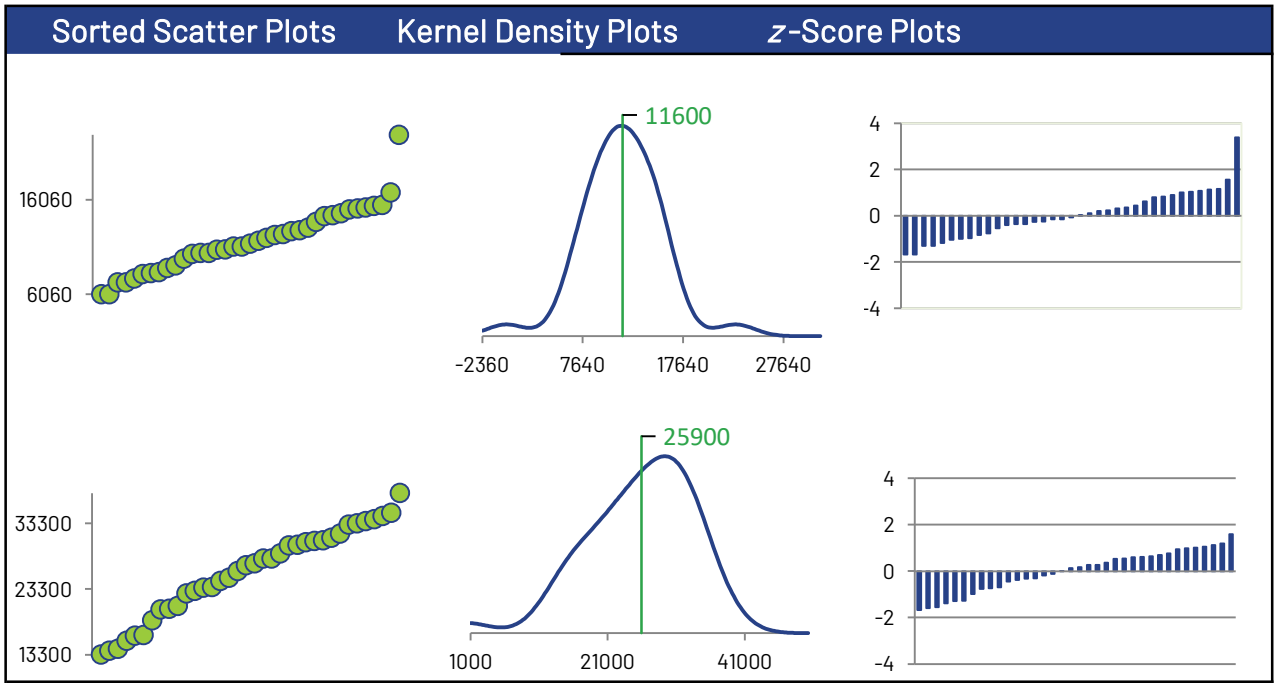
Method	C40B-1	C40B-2	C40B-3	C40B-4
GC/FID (Blue)	37	37	37	36

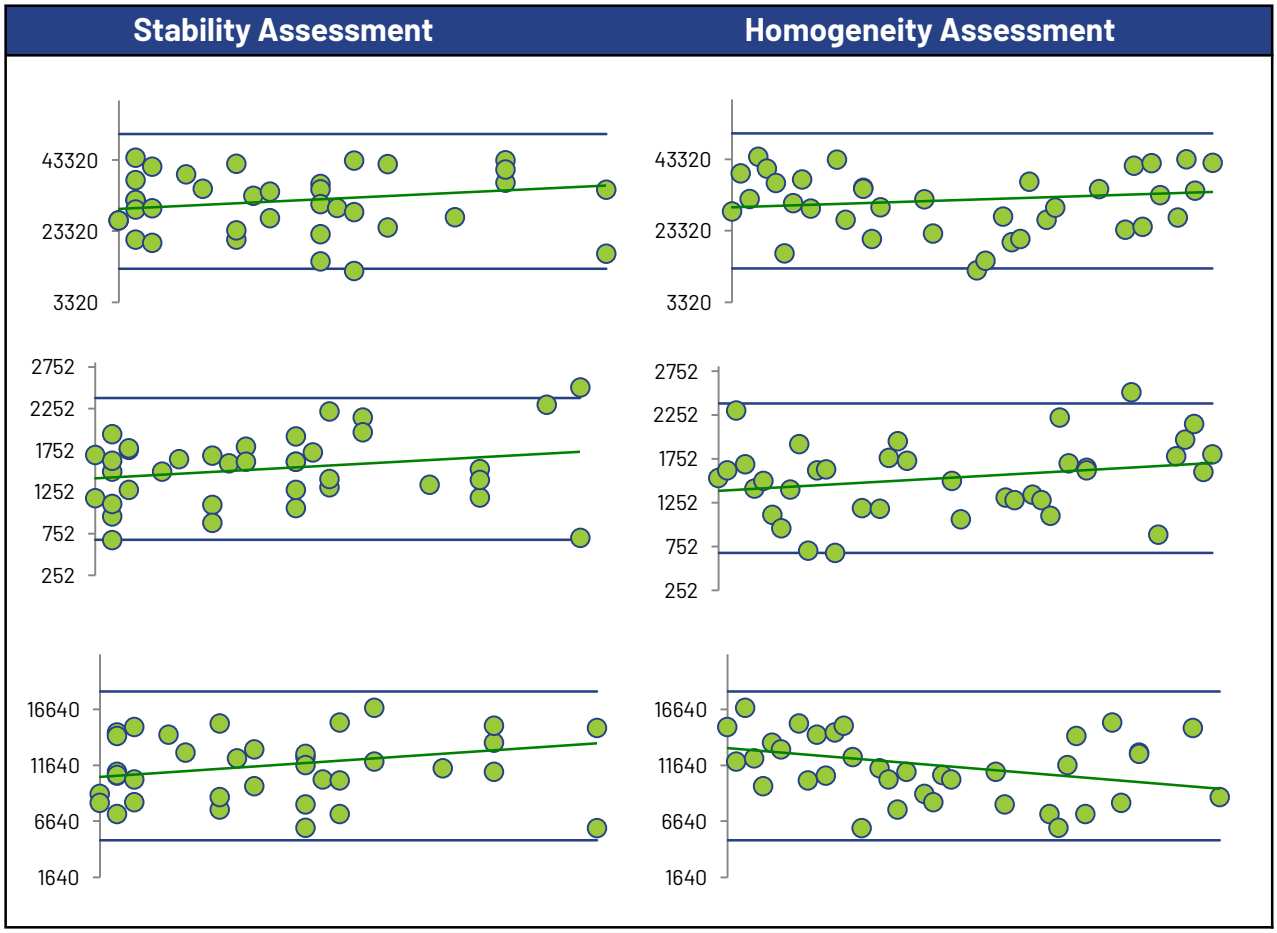
All summary stats, except the failure rate, and the plots below are based on the data excluding any flagged outliers



* Displayed failure rate does not account for the RDL modified z-Scores or inaccurately reported non-detects.

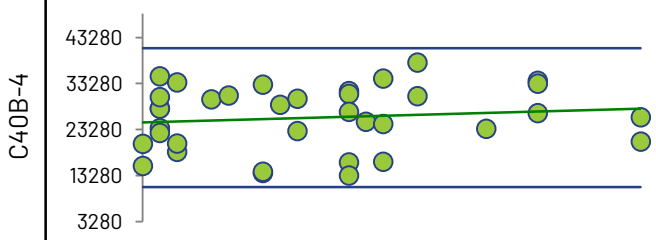
F4: C34-C50


Sorted Scatter Plots
Kernel Density Plots
z-Score Plots
C40B-3
16060
6060
11600
-2360
7640
17640
27640
4
2
0
-2
-4
C40B-4
33300
23300
13300
25900
1000
21000
41000

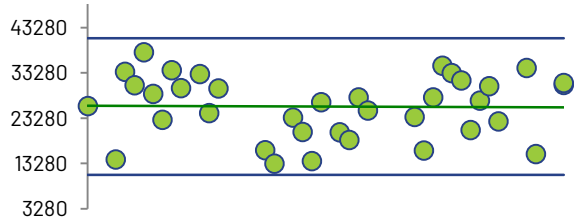

Stability Assessment
Homogeneity Assessment
C40B-1
43320
23320
3320
C40B-2
2752
2252
1752
1252
752
252
C40B-3
16640
11640
6640
1640

F4: C34-C50

Stability Assessment



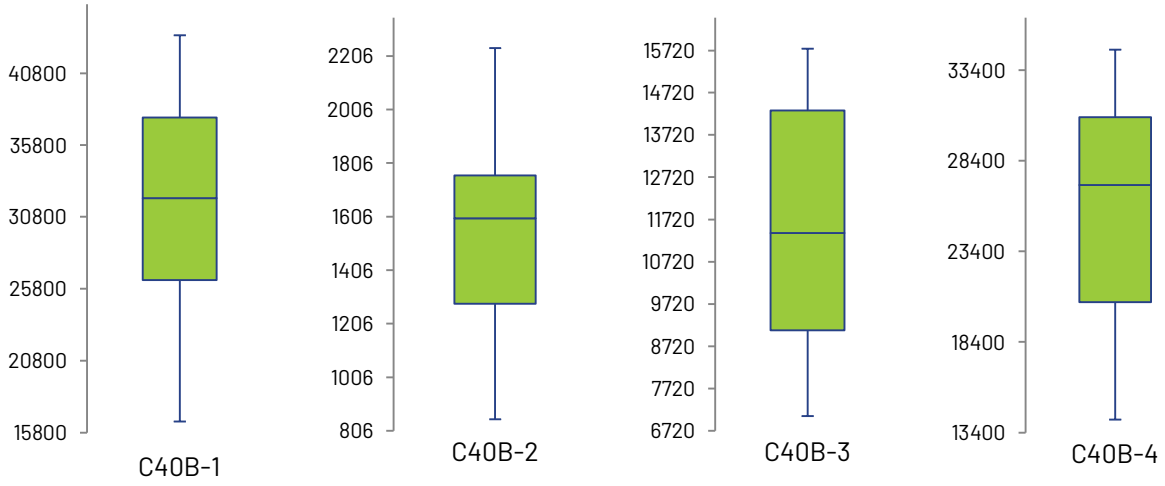
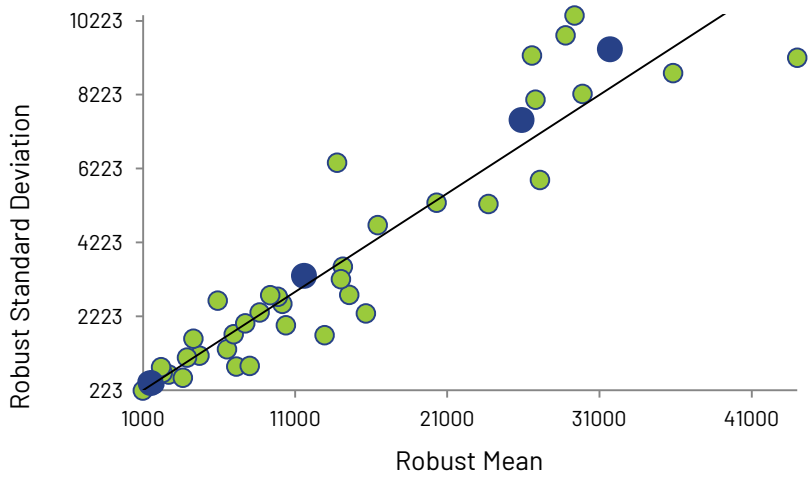
Homogeneity Assessment



Stability assessments are regression analysis of reported result against date of analysis.

Homogeneity assessments are regression analysis of reported result against bottling order.

Box and Whisker Plots



Current Round (blue) Compared to Historic Rounds (green)