State of California—Health and Human Services Agency
California Department of Public Health

KAREN L. SMITH, MD, MPH
Director & State Health Officer

EDMUND G. BROWN JR.
Governor

September 20, 2017

TO: Participants in the July 2017 Voluntary Proficiency Test in Forensic Alcohol Analysis

SUBJECT: Assigned Values and Expected Ranges of Results for the July 2017 Proficiency Test in Forensic Alcohol Analysis

Attached is a summary of the descriptive statistics for the July 2017 proficiency test in forensic alcohol analysis. The Department prepared four test pools (06267A, 06267B, 07067A, and 07067B) for this proficiency test. Included in the summary are the target formulation values for the pools, the test pools' true values as determined by the Department's analyses, the peer-group or consensus values and the standard deviations, and graphical summaries of the distribution of participant results.

With the recent revisions[1] to the Title 17 regulations, the Department is no longer authorized to evaluate participants' performances on proficiency tests. Instead, staff of each individual laboratory must evaluate the laboratory's results to determine whether they are consistent with expected test results [cf. 17 CCR §1220.1 (b)]. The comments below describing the procedures historically used by the Department when evaluating results are advisory in nature and intended to assist the laboratories in evaluating their own results.

Historically, the Department has determined the acceptable limits of performance based on reported results that are within the range representing ±5% of the 99% confidence interval of the peer group mean, where the range has been truncated to two significant figures (Table 1). This range is described as the "Tier #2 interval." The Department also calculated a narrower "Tier #1 interval," which represents the range of reported results that are within ±5% of the 95% confidence interval of the peer group mean where the range is based on the results reported to three significant figures. Tier #1 was expected to include those laboratories demonstrating a high degree of accuracy. The second, wider tier was intended to include those laboratories not as close to the central tendency as the first tier, but still accurate and therefore adequately competent.

One of the recent revisions to the Title 17 regulations was to permit the expression of results to either two or three decimal places. When reporting to the second decimal place, the digit in the third decimal place must be deleted [cf. 17 CCR §1220.4 (b)]. The regulations are silent with respect to the procedures for determining the third decimal place.

[1] Revised Title 17 regulations filed with the Secretary of State on 1/26/17, with an effective date of 4/1/2017.

The majority of the participants [20 out of 26] reported results to three decimal places. Under these circumstances, the wider second tier based on two decimal place results, which again historically was used by the Department to evaluate the laboratories' results, is no longer appropriate.

The IUPAC International Harmonized Protocol for the Proficiency Testing of Analytical Chemistry Laboratories (Harmonized Protocol) recommends the use of z-scores for evaluating proficiency test data. However, the Harmonized Protocol notes that that the interpretation of the z-scores is based on the normal distribution of reported results, in which case the z-scores can be expected to follow the standard normal distribution. As indicated in Table 2, none of the results in this proficiency test were found to be normally distributed. Accordingly, the use of z-scores may not be completely appropriate, but they still may be useful to identify outlier and/or warning level results. The expression for calculating a z-score is included in Table 2. Generally a score between -2 and +2 ($|z| \leq 2$) is considered satisfactory or acceptable. A score outside the range -3 to +3, inclusive ($|z| \geq 3$) is considered unsatisfactory or unacceptable and the laboratory must take corrective actions. Z-scores between -3 and -2 or +2 and +3 ($2 < |z| < 3$) are considered questionable and these two ranges should be used as warning limits. Scores within the warning limit ranges in two or more consecutive test events could be considered unacceptable.

The proficiency test results expressed as *z*-scores for the participants whose results were used to determine the peer group mean and statistics in the July 2017 test are summarized in Figure 7[2]. Participants are identified by codes. An enclosure with the current correspondence provides codes for the results submitted by your laboratory.

Another approach for evaluating proficiency test data, which is non-parametric and does not require the data to be converted to a standard normal form, divides the test data at regular intervals or quantiles[3]. The quartile is a type of quantile: the first quartile (Q_1) is defined as the middle value between the lowest value and the median of the data set. The second quartile (Q_2) is the median of the data set. The third quartile (Q_3) is the middle value between the median and the highest value of the data set. The interquartile range (IQR), a measure of the dispersion of the data, is the difference between the upper and lower quartiles ($IQR = Q_3 - Q_1$). Boundaries (called fences) are set at $Q_1 - 1.5$ IQR (lower fence) and $Q_3 + 1.5$ IQR (upper fence) to identify potential outliers in the tails of the distribution. In Figure 5, the data from pools 06267A and 06267B are presented as box and whisker or Tukey plots with the quartiles and fences shown. The median of the data is shown by a black line and the mean of the data is shown by a red line inside the box. Figure 6, presents the same data for pools 07067A and 07067B. These figures can be used by the participants to evaluate their data.

[2] When calculating z-scores, the Department used the round even mean of the three decimal place duplicate results reported by the participants since this represents the best estimate of the sample concentration.
[3] See Statistics and Chemometrics for Analytical Chemistry Sixth Edition, Miller and Miller (p. 158)

A copy of this report is available on Food and Drug Laboratory webpage:

Sincerely,

Clay Larson, Chief
Abused Substances Analysis Section
Food and Drug Laboratory Branch

For questions or additional information, contact the Food and Drug Laboratory Branch:

Phone – (510) 412-6220

Web – https://www.cdph.ca.gov/Programs/CEH/DFDCS/Pages/FALP.aspx

Email - fdlb.info@cdph.ca.gov

Statistical Data for July 2017 Proficiency Test in Forensic Alcohol Analysis

Table 1 CDPH Tier #1 and Tier #2 Acceptable Ranges (grams%)

Pool #	Pool Date Code	Peer Group Mean	Tier #1	Tier #2
#1A	06267A	0.059	0.054 – 0.064	0.05 – 0.06
#1B	06267B	0.104	0.097 – 0.111	0.09 – 0.11
#2A	07067A	0.161	0.150 – 0.172	0.14 – 0.17
#2B	07067B	0.255	0.239 – 0.271	0.23 – 0.27

Table 2 Summary of Test Pool Data

Parameter		Pool1A (06267A)	Pool 1B (06267B)	Pool 2A (07067A)	Pool 2A (07067B)
Pre-distribution Data	Target Value	0.060	0.110	0.160	0.260
	True Value[4]	0.057	0.103	0.160	0.255
	Standard Deviation	0.0002	0.0006	0.0007	0.0013
Descriptive statistics	Mean	0.059	0.104	0.161	0.255
	Adjusted Mean[5]	0.059	0.104	0.161	0.256
	Standard Error[6]	0.0002	0.0004	0.0005	0.0009
	Median	0.059	0.104	0.161	0.256
	Standard Deviation	0.0016	0.0025	0.0037	0.0061
	Minimum	0.055	0.098	0.149	0.234
	Maximum	0.063	0.110	0.172	0.266
	Count	47[7]	47[7]	47[7]	47[7]
Descriptive statistics (box plot)	Q1 (25%)	0.058	0.103	0.160	0.252
	Q3 (75%)	0.060	0.106	0.162	0.258
	IQR	0.002	0.003	0.002	0.006
	Lower Fence	0.055	0.098	0.157	0.243
	Upper Fence	0.063	0.110	0.165	0.267
Histogram		Figure 1	Figure 2	Figure 3	Figure 4
Normal distribution?[8]		No (p=0.017)	No (p<0.001)	No (p<0.001)	No (p<0.001)
Box Plot (SigmaPlot™)		Figure 5	Figure 5	Figure 6	Figure 6
Robust mean, X*[9]		0.0590	0.1042	0.1612	0.2558
Robust standard deviation, σ_{rob}		0.0014	0.0016	0.0027	0.0035
Fitness-for-purpose standard deviation, σ_p[10]		0.0018	0.0029	0.0042	0.0063
Consensus value (X_a) determined as Mode ($\mu_{1/2}$) of Gaussian Kernel distribution		0.0590	0.1042	0.1612	0.2558
Uncertainty of the consensus value, X_a , S.E.[11]		0.00024	0.00023	0.00033	0.00062
$X_a \pm$ S.E.		0.0590 ± 0.0002	0.1041 ± 0.0002	0.1611 ± 0.0003	0.2558 ± 0.0006
z-score		$z = \frac{X - X_a}{\sigma_p}$	$z = \frac{X - X_a}{\sigma_p}$	$z = \frac{X - X_a}{\sigma_p}$	$z = \frac{X - X_a}{\sigma_p}$

[4] Based on CDPH's Headspace Gas Chromatographic Method
[5] Mean determined from participant data after the removal of outlier(s)
[6] Standard Error of the Mean
[7] A total of 26 laboratories participated and analyzed a total of 47 sample sets.
[8] Shapiro-Wilk test used at 0.05 significance level.
[9] Robust mean of the results reported by the participants was calculated using Algorithm A in Annex C of ISO 13528:2005.
[10] The Department has determined a value for σ_p as 2.5% of robust mean for roughly symmetrical distributions based on the uncertainties associated with the reported results on recent tests together with the 5% accuracy and precision standard of performance requirements set forth in the regulations. In case of skewed, non-normal distributions, the revised, derived Horwitz equation (σ_p') is used **: $\sigma_p' = 0.02 * \mu_{1/2}^{0.8495}$**
[11] Determined as the Standard Error of Mode using bootstrap simulation technique with bandwidth of $0.75*\sigma_p$

Figure 1

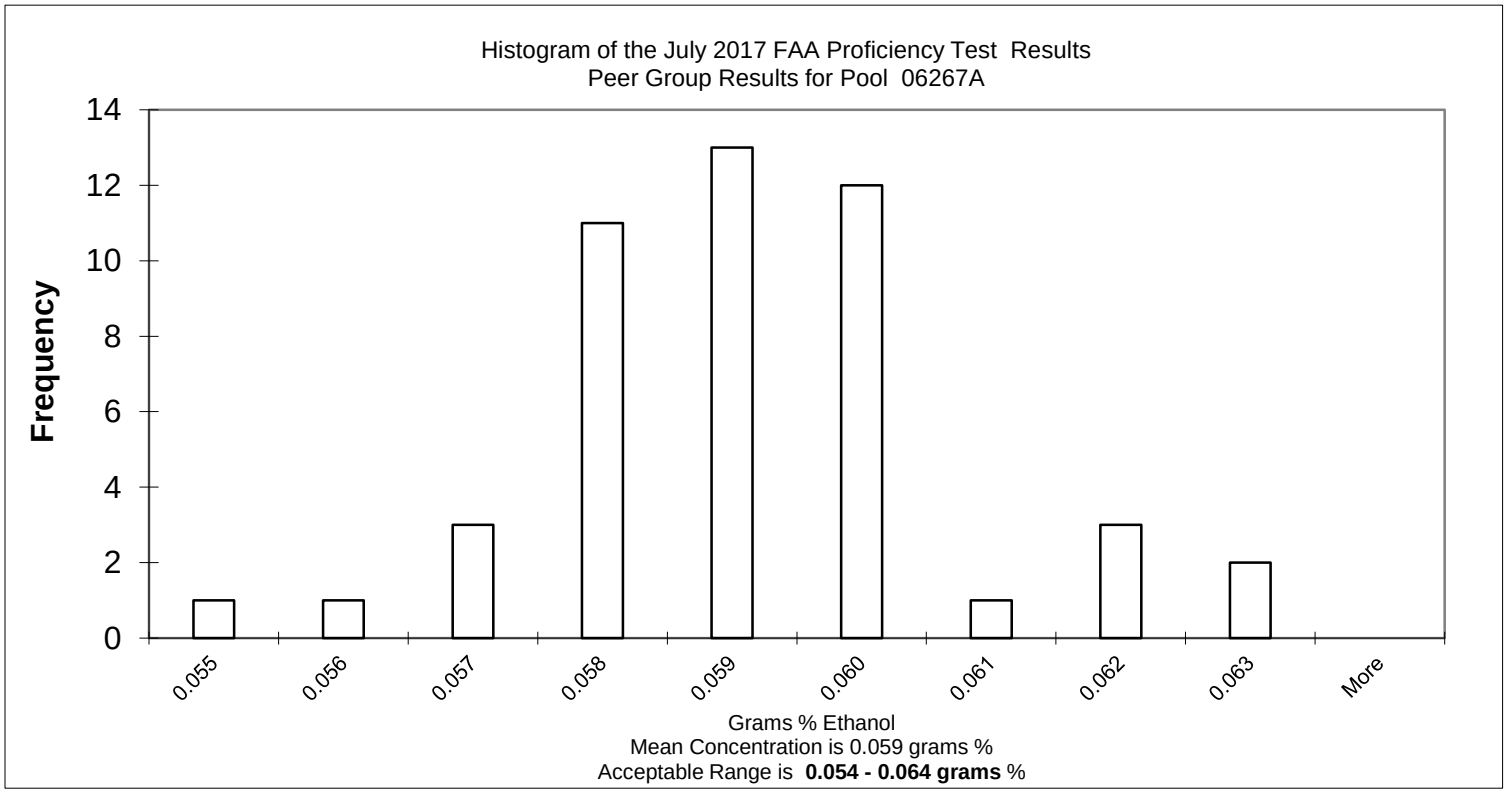


Figure 2

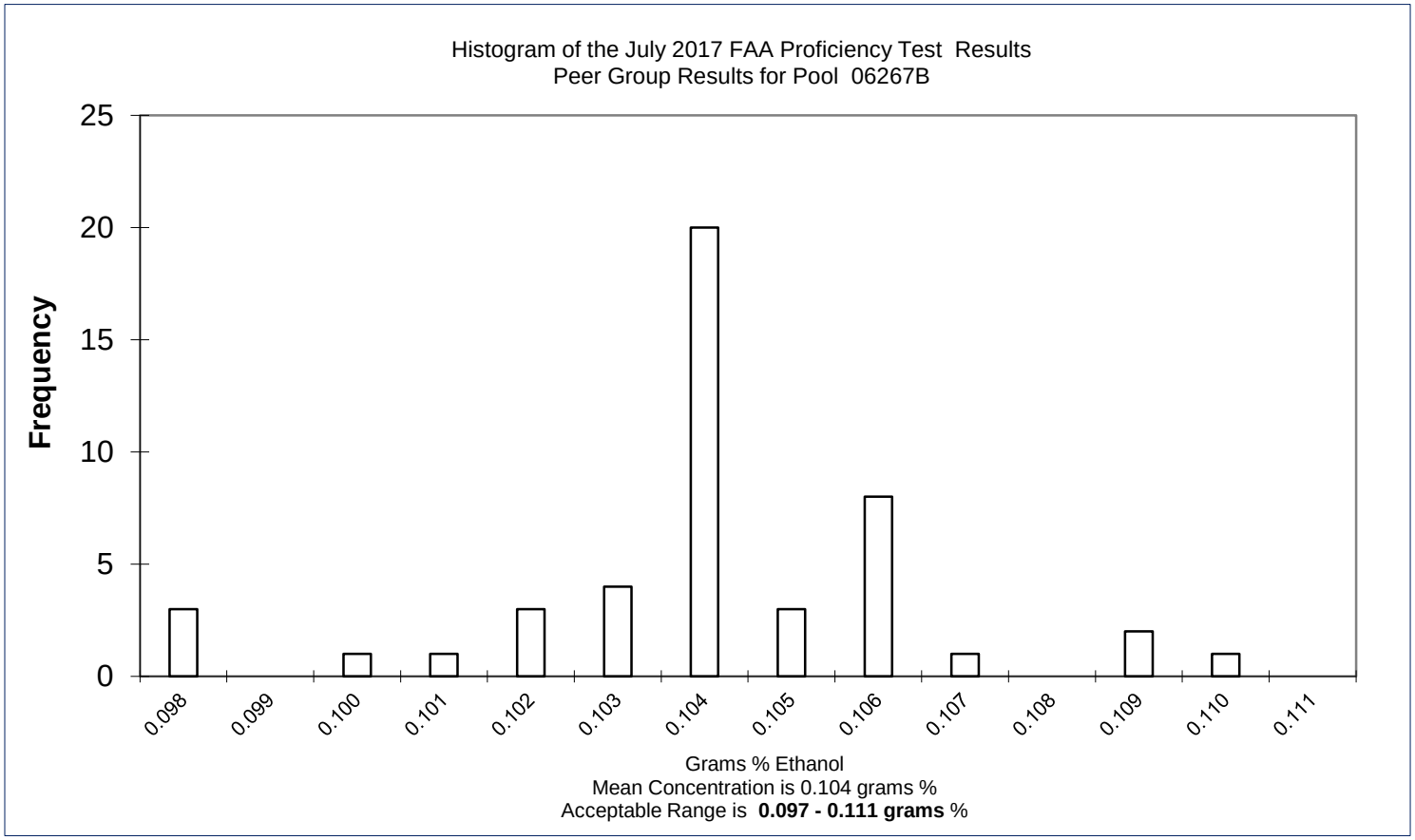


Figure 3

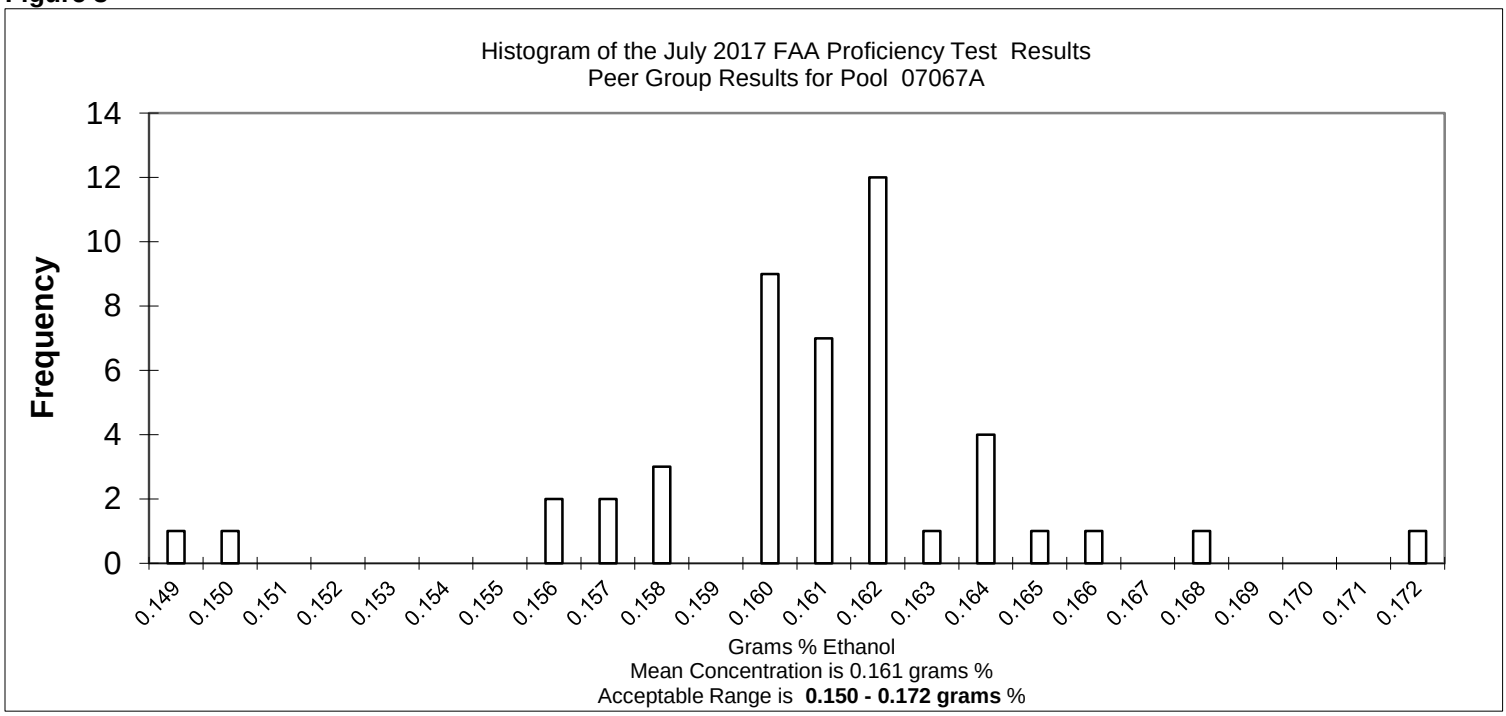


Figure 4

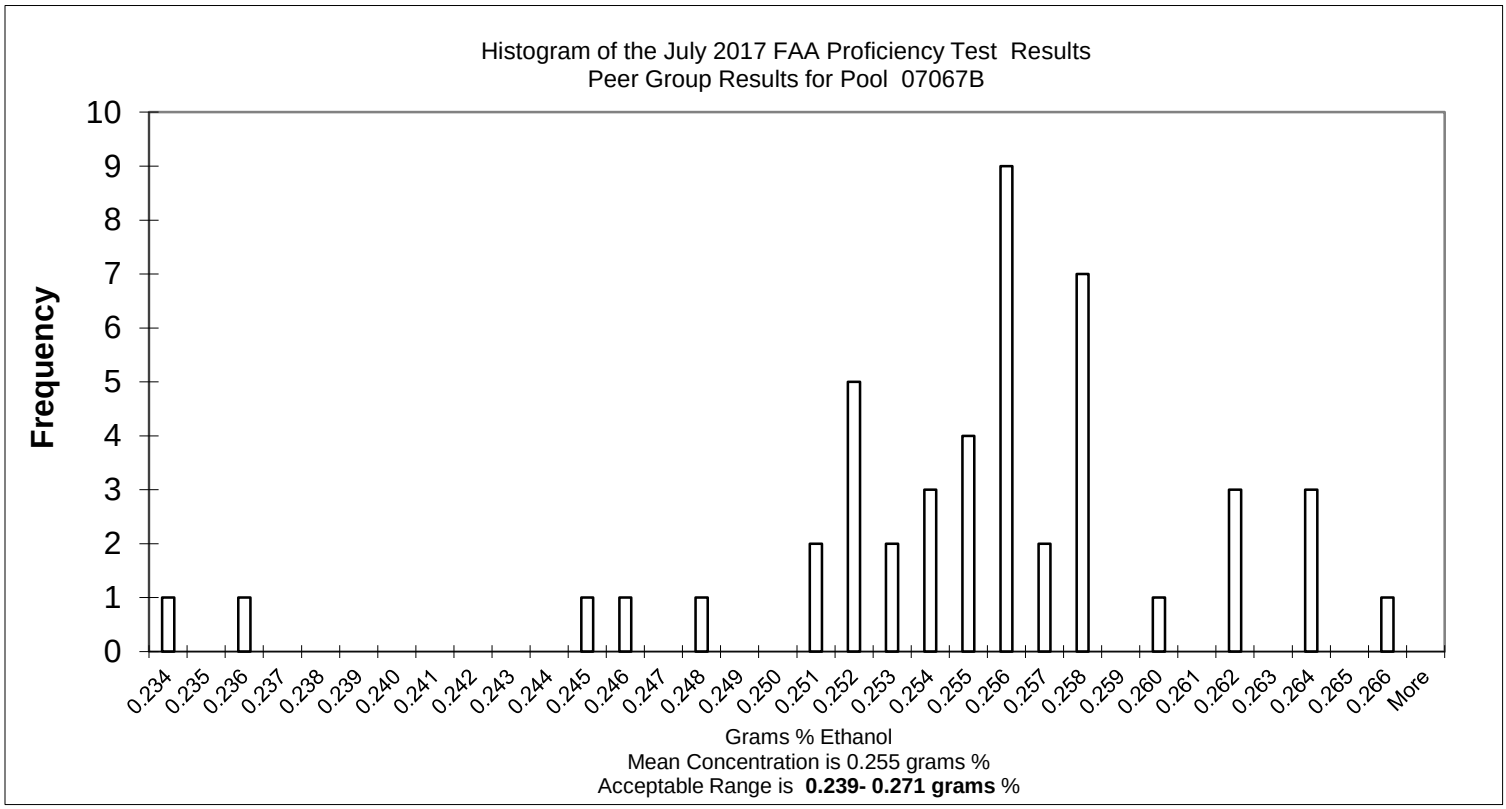
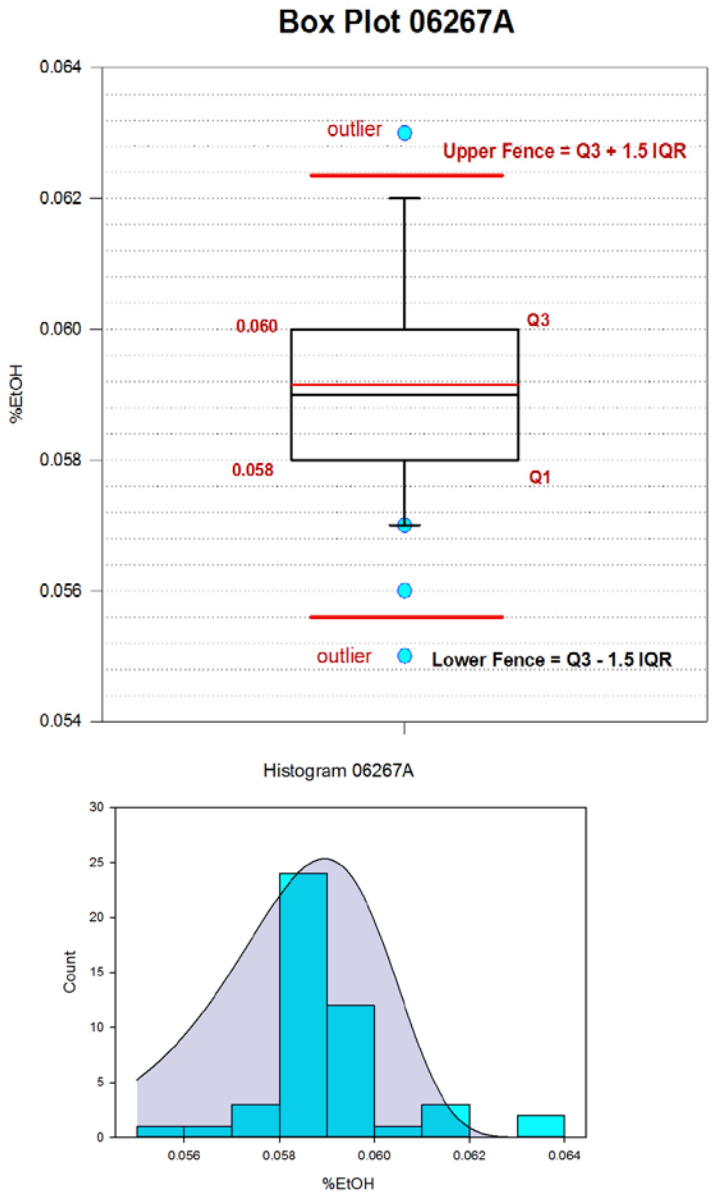


Figure 5 – SigmaPlot™ analysis of pools 06267A & 06267B



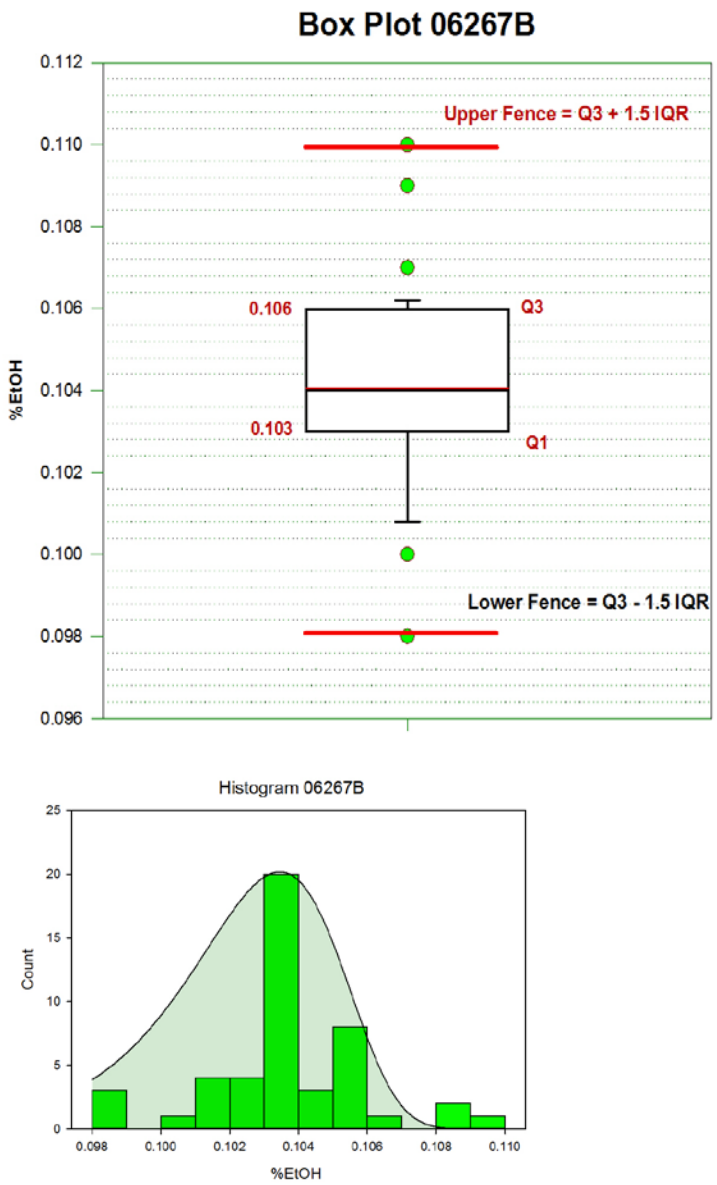
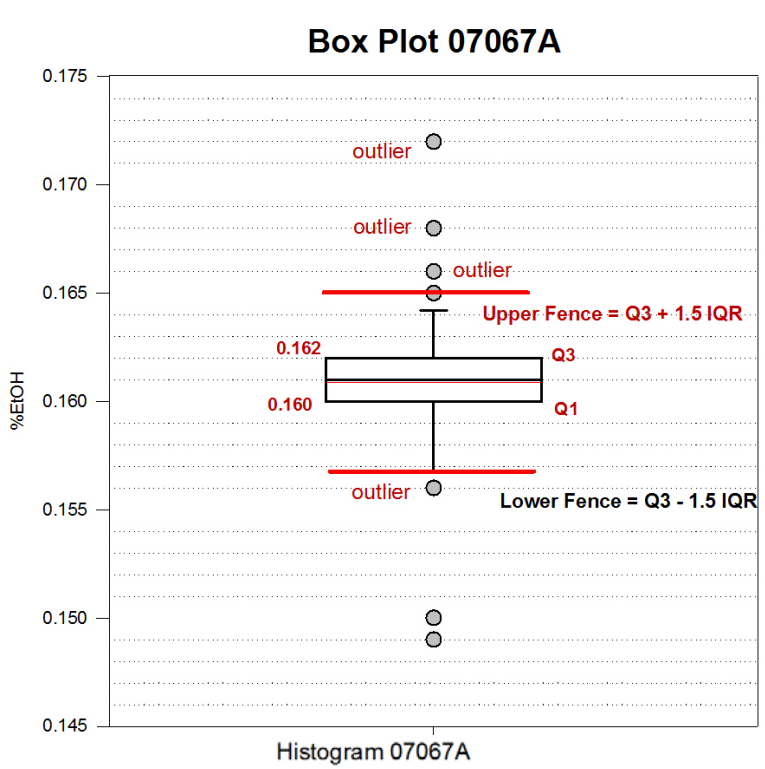
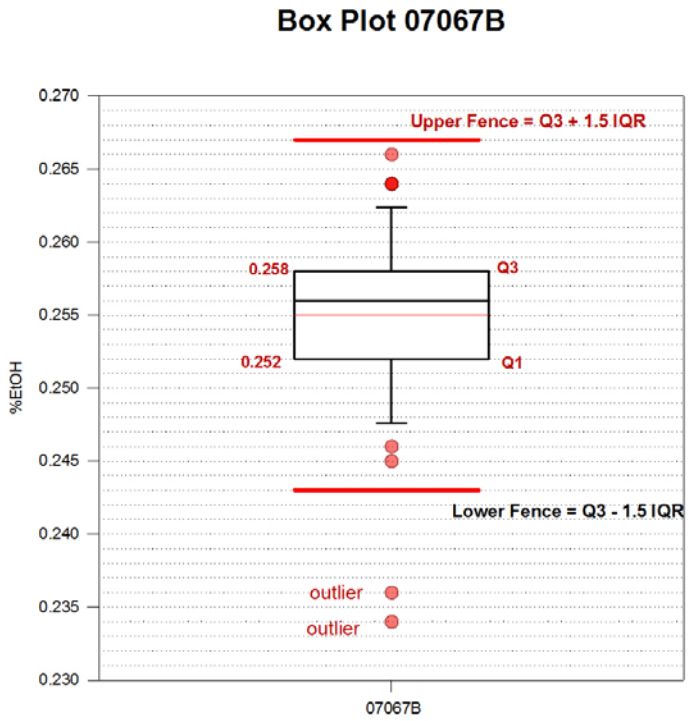
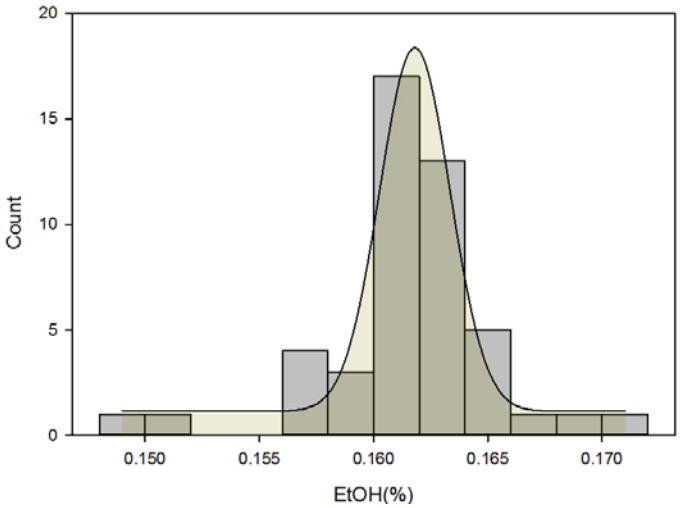


Figure 6 - SigmaPlot™ analysis of pools 07067A & 07067B







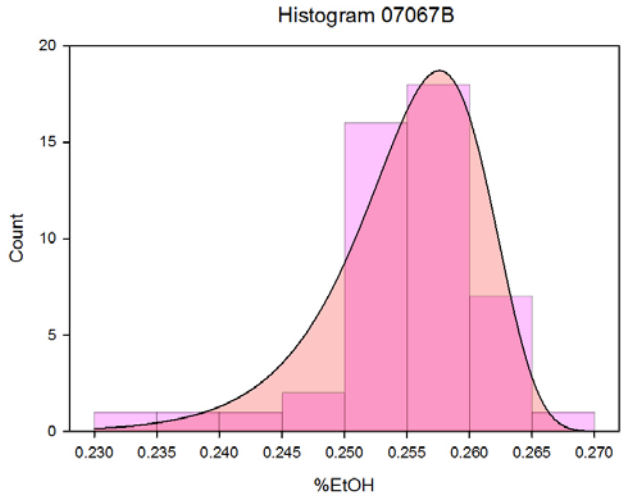
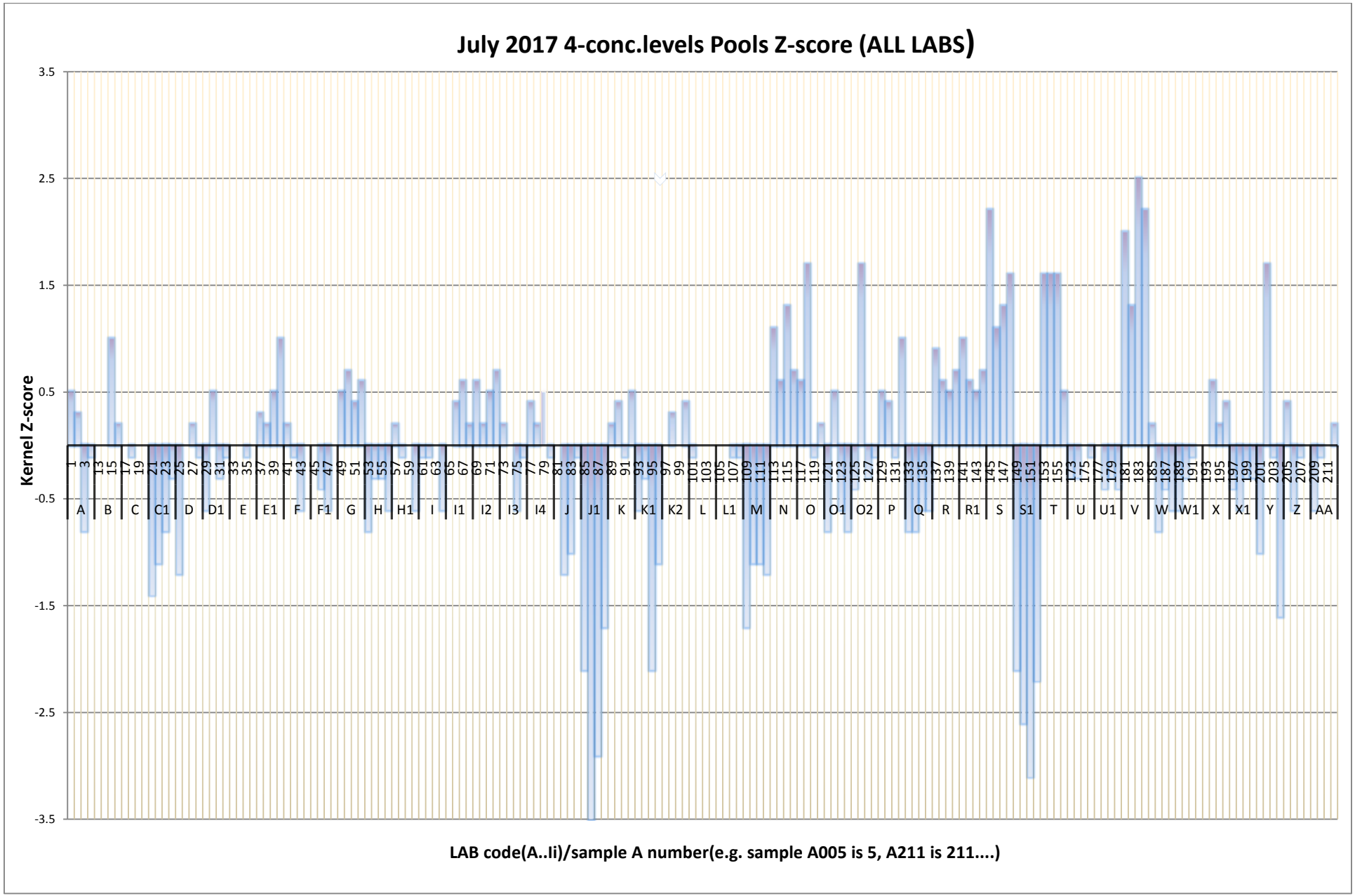


Figure 7


July 2017 4-conc.levels Pools Z-score (ALL LABS)
Kernel Z-score
3.5
2.5
1.5
0.5
-0.5
-1.5
-2.5
-3.5
A B C C1 D D1 E E1 F F1 G H H1 I I1 I2 I3 I4 J J1 K K1 K2 L L1 M N O O1 O2 P Q R R1 S S1 T U U1 V W W1 X X1 Y Z AA
LAB code(A..Ii)/sample A number(e.g. sample A005 is 5, A211 is 211....)