

Test Method (Im)precision

How to Account for It

BY ALEX T. LAU

Q: How can you account for test method (im)precision in the adjudication of a petroleum product specification conformance dispute?

A: In the petroleum industry, commercial petroleum product specifications are generally articulated in terms of maximum or minimum limits as measured by specifically referenced standardized test methods (STM) developed by ASTM International or other standards development organizations. The current de facto industry practice is that an aliquot taken from a homogeneous batch of finished product will be tested once using the specification-referenced method by the supplier, and the product is deemed to be conforming and fit for release if this single test result meets the specification limit.

Since we do not live in a perfect world, replicate execution of the same STM on the same material by different operators in different laboratories will not always yield numerically identical results.¹ Hence, there will be occasions where a supplier will ship a product based on a single conforming test result, but a retest by the customer at the receiving facilities will yield a nonconforming result.

This article introduces an objective protocol in ASTM D3244 on how to address the product conformance issue when faced with two conflicting results as described in the above scenario.

D3244, Practice for Utilization of Test Data to Determine Conformance with Specifications, is under the jurisdiction of ASTM D02.94, Coordinating Subcommittee on Quality Assurance and Statistics, a part of ASTM International Committee D02 on Petroleum Products and Lubricants. The standard covers guidelines and statistical methodologies with which two parties, usually a supplier and a receiver, can compare and combine independently obtained test results to obtain an assigned test value for the purpose of resolving a product quality

dispute. The technique for determining the acceptance limit, against which the *ATV* is compared, is an integral part of this protocol. The protocol applies only to STMs with published repeatability and reproducibility limits that conform to requirements of *Form and Style for ASTM Standards*.

A brief overview of the D3244 protocol is described below to handle the case when there are two results: one is conforming (usually from the supplier) and the other is nonconforming (usually from the receiver or an independent third party retest). Each result is obtained from a single application of the specification-referenced STM by the supplier's and receiver's (or independent) laboratory, respectively.

The D3244 protocol calls for the following steps:

1. The supplier and receiver agree a priori on the probability of acceptance, P , if the true value of the property is exactly at the specification limit. This P value is required for the calculation of the AL in step 3.
2. Compare the difference between the two results to the published reproducibility, R , of the STM. If this difference is less than or equal to R , then calculate the ATV by taking the average of both results. If not, reject both results and obtain a new set of results using a mutually agreed upon aliquot. Repeat step 2.
3. Calculate the AL based on the ATV in step 2, the P value agreed upon in step 1, and R for the STM using Equation 2 in the D3244 practice. If the ATV meets the AL , the product is accepted. Otherwise the product is rejected.

Determination of the P value in step 1 is guided by the mutually agreed upon criticality of the specification. In the absence of an agreement to the contrary, practice D3244 recommends the following:

- For noncritical specifications, set the AL such that there is 95 percent probability of product acceptance if the true value of the property is exactly at the specification limit value.

- ▶ For critical specifications, set the *AL* such that there is 5 percent probability of product acceptance if the true value of the property is exactly at the specification limit value.

ASTM D3244 defines critical specifications as those specifications that, due to the product characteristic or the end use of the product, or both, require that the receiver have a high degree of assurance that the true value of the product property actually meets or exceeds the quality level indicated by the specification limit value. Noncritical specifications are defined as those that only require reasonable assurance that the product property is not substantially poorer than indicated by the specification limits.

It should be noted that for $P = 0.05$ (critical specification), the *AL* will actually be numerically inside the specification limit values, which will result in a lower consumer's risk of unknowingly accepting nonconforming product. For $P = 0.95$, the *AL* will be outside the specification limit values, which will result in a lower supplier's risk of falsely rejecting conforming product.

When $P = 0.5$, the *AL* coincides exactly with the specification limit. This means that there is a 50 percent probability that the product will be accepted if the true value of the property is exactly at the specification limit. The practical implication with $P = 0.5$ is that the receiver and supplier equally share the risk associated with test method (im)precision.

This is also the delineation point between critical and noncritical specification as chosen by the practice.

As a prerequisite for acceptance for lab test results to be used in the calculation of *ATV*, practice D3244 also requires the following:

- ▶ Long-term standard deviation for the STM, as practiced by each lab, for material typical of the product in dispute, is statistically indistinguishable or better than the published

method standard deviation under reproducibility conditions.

- ▶ Each lab must be able to demonstrate, by way of results from proficiency testing programs, a lack of a systemic bias in the execution of the STM in question.

Both requirements can be substantiated by in-house quality control programs involving appropriate statistical control charts that meet the requirements of ASTM D6299, Practice for Applying Statistical Quality Assurance and Control Charting Techniques to Evaluate Analytical Measurement System Performance. Practice D6299 is also under the jurisdiction of D02.94.

Interested readers are encouraged to study the examples in the annex of practice D3244.

REFERENCE

1. Lau, Alex, "What Are Repeatability and Reproducibility? Part 1: A D02 Viewpoint for Laboratories," *ASTM Standardization News*, March/April 2009, pp. 15-18, www.astm.org/SNEWS/MA_2009/datapoints_ma09.html.

ALEX T. LAU, *TCL-Consulting, Whitby, Ontario, Canada, is chairman of D02.94, the Coordinating Subcommittee on Quality Assurance and Statistics, and of D02.01B on Precision, which are part of ASTM Committee D02 on Petroleum Products and Lubricants. An ASTM International fellow, Lau is also a member of Committees E11 on Quality and Statistics, E36 on Accreditation and Certification and F08 on Sports Equipment and Facilities.*

DEAN V. NEUBAUER, *Corning Inc., Corning, NY, coordinates the DataPoints column; an ASTM International fellow, he is chairman of Committee E11 on Quality and Statistics and chairman of E11.90.03 on Publications.*

sonline

Get more tips for ASTM standards development at www.astm.org/sn-tips.

Find other DataPoints articles at www.astm.org/standardization-news/datapoints.