Application Note

Aqua Lab Instrument Performance

Accuracy, Precision, Repeatability, Reproductibility, Uncertainty: What does it all mean?

Decagon Devices

Date: 2008

Summary

Decagon reports an accuracy of ±0.003 a_w and a precision of ±0.001 for its AquaLab chilled mirror benchtop instruments. Customers who want to report accuracy, precision, and uncertainty according to ISO definitions sometimes need to know exactly what these numbers mean. This application note was written to define that as specifically as possible.

Defining Accuracy and Precision

The words "accuracy" and "precision" are often used interchangeably, but scientifically speaking, they judge the quality of a measurement in two very different ways.

Accuracy: Systematic Errors

Accuracy refers to systematic errors—errors associated with the method of measurement. Accuracy is evaluated by comparing a measurement to an accepted or defined standard. For example, a thermometer in boiling water at sea level should read 100 °C. The accuracy of any specific thermometer could be evaluated using this standard. If a measurement has no "true value," an expression of accuracy is just a qualitative number.

Calculating Accuracy

Accuracy is typically quantified as the standard error of measurement. Differences between measured values and true values across all observations are squared and summed. Accuracy is then calculated as the square root of that sum divided by the degrees of freedom. If the accuracy value from this calculation was 0.01, it would be reported as an accuracy of "±0.01." This indicates a 95% confidence interval, meaning that 95% of all measured values will be within ±0.01 of the true value.

AquaLab Accuracy

Decagon reports an accuracy of ±0.003 for its AquaLab chilled mirror benchtop instruments. In practice, the actual accuracy of each instrument as it measures each water activity standard will be slightly different. Decagon reports a general accuracy value that has been confirmed to be a conservative estimate of the actual accuracy for any instruments released by Decagon for sale. This ±0.003 accuracy value should be used when determining if an instrument is reading the standards correctly. For a more in depth discussion of AquaLab accuracy, please refer to Decagon's Application Note titled "Chilled Mirror Water Activity Instruments: What is the real story about the accuracy and speed of chilled mirror water activity instruments?"

Precision: Random Variability

Precision refers to random errors—errors that arise from the random variability of any experimental situation. Because these errors are random, they can be evaluated statistically as the standard deviation among readings on a population of equivalent samples or subsamples. (Yu et al., 2009). A precise reading would have a small standard deviation.

Accurate but Not Precise?

It's possible to be accurate without being precise, and precise without being accurate, as shown by the dart boards below. Your shots are accurate when they hit the bulls-eye, precise when they are grouped closely together. Only when they are closely grouped within the bulls-eye are they both.

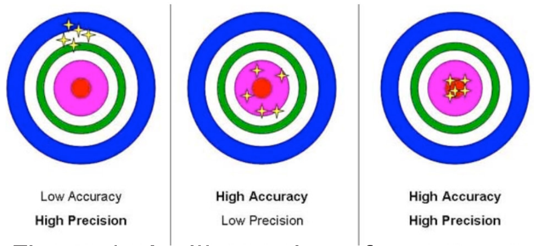

Figure 1. An illustration of the differences between accuracy and precision.

Uncertainty

ISO defines uncertainty as characterizing the dispersion of measured values around the mean (ISO 1995). In other words, it is essentially a measure of precision. The difference between uncertainty and precision is that precision typically refers to a single source of error. Uncertainty can embrace many sources of error.

Breaking Down Uncertainty

Identifying the different sources of variation can be very helpful in developing strategies to reduce uncertainty. See Yu et al. (2009) for a thorough discussion of the many sources of variation in water activity measurements.

Decagon reports the precision of its bench-top AquaLab water activity instruments as ±0.001. This is a Type A measurement of uncertainty (ISO 1995) because it is the standard deviation across multiple measurements taken on salt standards. The uncertainty of each reading of each instrument on each standard is a unique value, but as with accuracy, Decagon reports the overall uncertainty. If you need an expanded uncertainty, multiply the reported precision by 2 to represent the 95% confidence interval.

Yu et al. (2009) showed variation in an AquaLab instrument reading comes from several sources. One is from variation in the actual measurement that determines the difference between dew point temperature and sample temperature. Another comes from variation in the salt standard used as a sample.

This second uncertainty is negligible—in the range of ±0.000005. (This measurement comes from unpublished data recorded at Decagon using a NIST traceable densitometer to verify concentration of salt standards). Consequently, the measurement of the difference between the dewpoint and sample temperatures, as well as any uncertainty due to hysteresis or resolution, is the largest source of variation in AquaLab's uncertainty. A model describing this uncertainty would be

$$\mu m = \mu Td\text{-}Ts + \mu std + \mu hys + \mu res$$

where μm is the combined instrument uncertainty, $\mu Td\text{-}Ts$ is the uncertainty due to measuring the difference between the dewpoint temperature and the sample temperature, μstd is the uncertainty due to the standards, μhys is the uncertainty due to hysteresis and μres is the uncertainty due to resolution. However, it is unnecessary to measure these uncertainties individually as they are all contained within the Type A uncertainty calculated across multiple readings on salt standards.

Measuring Uncertainty

To correctly measure the Type A uncertainty at any salt standard level requires testing and calculating the standard deviation across multiple samples of each standard and not just measuring the same salt standard sample multiple times in the instrument without removing the sample as this is not a true repeated measurement but simply an indication of stability or equilibrium.

It's important to distinguish between the uncertainty of the AquaLab instrument and the uncertainty of a water activity measurement made by the AquaLab on a product. When measuring something other than homogenous salt standards, the inherent uncertainty of the product itself is a significant source of variation.

In fact, Yu et al. (2009) found that the uncertainty due to the product was always larger than the uncertainty of the instrument itself. Consequently, you should not expect AquaLab to read a product with the In fact, Yu et al. (2009) found that the uncertainty due to the product was always larger than the uncertainty of the instrument itself. Consequently, you should not expect AquaLab to read a product with the same uncertainty (±0.001) that it has when reading salt standards.

Uncertainty in Product Measurements

When determining the uncertainty or precision of water activity on a product, handling of the samples becomes very important. The standard deviation should be calculated using measurements taken on subsamples taken from an individual sample. If it is desirable to include the variation attributable to sampling, comparisons should be made of water activity means for each sample determined individually across subsamples from each sample. This can be done using a statistical tool for comparing sample means such as a Student's T test and the standard error of the mean. For batch to batch comparisons, measurements should still be made on subsamples from multiple samples in each batch, but comparisons are made using batch means as averaged across the samples in each batch and the overall standard error of the batch mean.

One further concept is important. Precision or uncertainty encompass both repeatability and reproducibility.

Repeatability

Repeatability is the goodness of agreement between successive measured values all measured under the same conditions (ISO 1995). This is calculated as the standard deviation across readings on multiple subsamples all tested under the same conditions.

Reproducibility

Reproducibility represents the goodness of agreement between measurements taken under different measurement conditions (ISO 1995). The changing conditions may include: a different time, method, lab, ambient conditions, or observer. Reproducibility is calculated as the standard deviation across readings on multiple subsamples tested under different conditions.

Reference List

International Organization for Standardization (ISO) (1995).

Decagon Devices
2365 NE Hopkins Court
Pullman Washington 99163
Printed in USA

05-25-2016