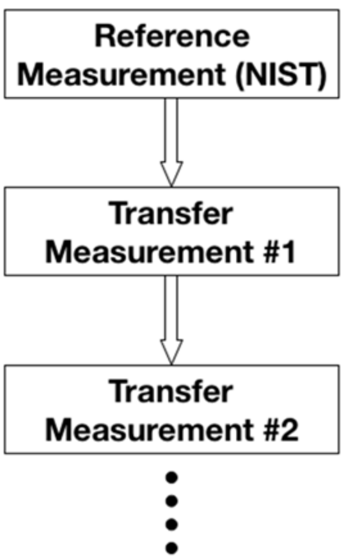
Traceability:
What is it? Why should you care?

In the United States,[1] the ultimate authority on measurement practice lies within the Department of Commerce, at the National Institute of Standards and Technology (NIST) [1]. So to start any discussion about traceability, we should start there. NIST states that the "traceability of measurement requires the establishment of an unbroken chain of comparisons to stated references each with a stated uncertainty."



Thus, traceability quite literally means the ability to be traced. Just like breadcrumbs leading you back to the place you started, the *unbroken chain* leads you back to the original reference measurement, and in the US, those measurements are carried out at NIST. So each traceable standard has a chain of measurements connecting it to measurements performed at NIST. Further, to be traceable, the documentation accompanying the traceable standard (its *certificate*) must *state* how the standard is connected back to NIST. Often certificates will only list one "link" in the chain. Assuming all of the previous links are trustworthy, one link is all that is really required to continue the chain. End users almost never continue this chain, but others will. For example, Avian Technologies, and other measurement services organizations by definition are providing physical standards and traceable measurements, each with a connection to the chain back to NIST. Often larger companies will also perform additional transfer measurements.

In summary, traceability requires a defined connection between any traceable standard, upstream through a series of measurements, ultimately ending at a reference measurement performed at an NMI. The position along the chain may carry special meaning, and therefore deserves special terminology. Standards that have been directly measured by an NMI are called *primary standards*. Avian Technology maintains many primary standards, as should any

---

[1] Since Avian Technologies resides in the US, we will refer to NIST as the ultimate reference for measurements. But note that all countries engaged in measurement programs have an analogous organization. These are collectively referred to as "National Metrology Institutes" (NMIs).

measurement services organization. After primary standards, the specific definitions get a little fuzzy; often the next level below primary is *secondary standard*, and after that, *transfer standard*. The end user has a *working standard* daily as a part of a measurement program. A standard that is supplied with a commercial instrument is an example of a working standard.

Now that we have a general idea of what traceability is and before going further, we need to understand what traceability is Not?

Laboratories who are contracting for measurement services want assurances of traceability, but they generally believe that traceability implies accuracy or implies that the measurement laboratory maintains ISO (9001 or 17025) certification.

Unfortunately, neither is necessarily true. There is no mention of accuracy in the NIST definition of traceability. While we at Avian Technologies strive for the most accurate measurements possible, this alone does not ensure, or even imply, traceability. ISO certification is a process that laboratories may choose to undergo. The certification states that they have and maintain certain procedural criteria in their workplace. Again, this does not necessarily equate to either accuracy or traceability.

What then is the process of producing a traceable standard?

A client ordering a traceable standard will receive a physical standard and a certificate. The certificate will contain the results of measurements performed on the standard, for example values of reflectance measured over a range of wavelengths. The certificate will also state uncertainties for the measurements (original statement from NIST, “each with a stated uncertainty”). The quantity that is stated for each wavelength is called the measurand and is the quantity that we seek, for example spectral reflectance or color coordinates. The uncertainty is “that parameter, associated with the result of measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand” (from the ISO guidelines to the expression of uncertainly in measurement, hereafter referred to as GUM). The uncertainty is what is transferred between instruments and each instrument has a stated uncertainty.

As commonly defined and used, “uncertainty” means the lack of a precise understanding of something. The relevant definition of *uncertain* from Merriam-Webster lists three related phrases:

- not known beyond doubt
- not having certain knowledge
- not clearly identified or defined

The *dispersion of values* refers to the fact that all instruments vary in their measurements, even when presented with an identical specimen. At its essence, uncertainty is result of statistically quantifying this dispersion.

Let's first consider the behavior of an instrument. Figure 2 shows spectral reflectance data of a white ceramic tile, measured 30 times. The instrument used here is a benchtop spectrophotometer. Note that the vertical scale is greatly expanded to show the spread of data. We can see that the measured reflectance varies even when the instrument is presented with an identical sample. In this case the sample was not moved between measurements, so the exact same region of the sample was presented for all measurements. Just looking at this plot we see that the width of the band of data is about 0.0004. We already get a hint that we quantify the spread of data using the standard deviation. The use of standard deviation to quantify the statistical behavior of instruments is fundamental to the philosophy of the GUM and all derived methods. We can think of standard deviation (usually denoted with a lower case sigma "σ") as the distribution of a set of data around the mean of the data.

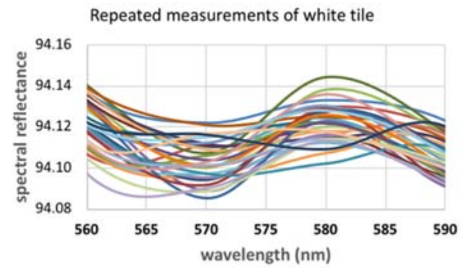


Figure 2

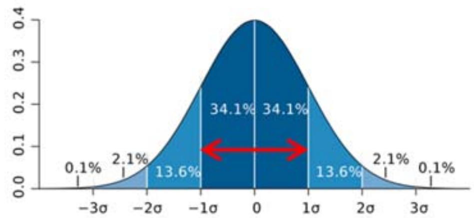


Figure 3

Figure 3 shows that for normally distributed data[2] about 68% of the data lie within one σ of the mean, about 95% within 2σ, and >99% within 3σ.[3] Figures 4 and 5 show the mean and standard deviation of the reflectance data in the upper chart.

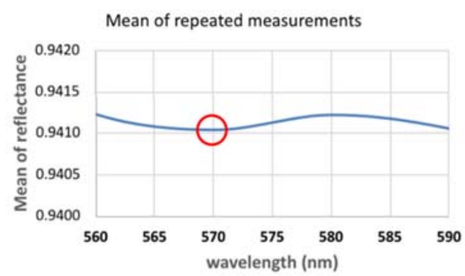


Figure 4

Consider the implications of Figures 3-5. At 570nm, the mean of the measurements is about 0.941, and standard deviation of the measured data is about 0.0010 (red circles). Therefore, according to Figure 3, there is a 68.2% probability any given measurement is in the range of 0.941±0.001.

In other words, the central dark blue region of Figure 3 (red arrow) shows that if we were to take a large number of measurements, we would expect that 68.2% (2x34.1%) of the data lie in that central region ±1σ either below or above the mean. In our case, the mean is not zero, but 0.9410, and the standard deviation is 0.0010. So our 68.4% probability range is 0.9410±0.0010.

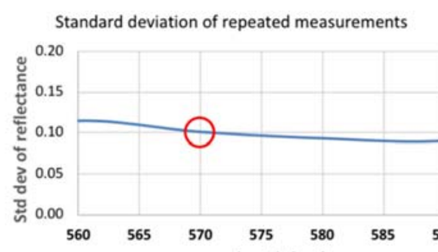


Figure 5

Figure 6 attempts so summarize this concept. Just remember that the goal here is to understand the distribution (or *dispersion*, as used in the GUM) of the replicate measurements. For normally distributed data, the mean and standard deviation do this quite nicely.

[2] "Normally distributed data" means a very specific thing to statisticians. While we make this assumption throughout the article, it is not universally agreed upon in the measurement community that this is best basis for quantifying uncertainty.

[3] The data leading to this plot were normalized to set the mean to 0.0 and the standard deviation to 1.0.

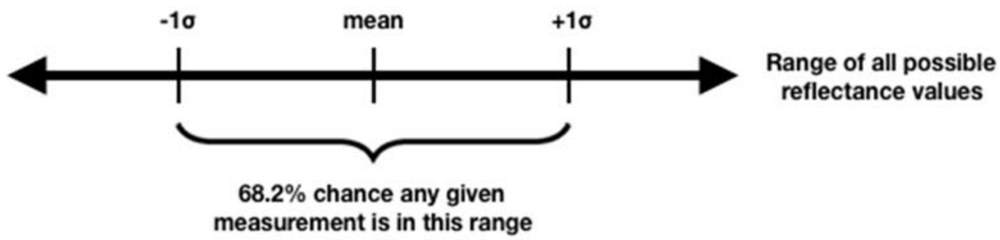


Figure 6

The reported uncertainty given by a National Metrology Institute does not simply account for the standard deviation of one measurement, but rather accounts for the uncertainty of many other parameters such as

- Variation of the light source in the instrument
- Angle of positioning of the sample
- Changes in sensitivity of the instrument due to changes in detector temperature
- … and on and on …

All of these individual uncertainties, collectively referred to as the *uncertainty budget*) are combined into one uncertainty. For the case of reflectance, we really take a bunch of individual measurements at each wavelength so there is a combined uncertainty for each wavelength, or in some cases ranges of wavelengths.

Practical Considerations

In Figure 1, we showed that for each instrument in the chain from the NMI to the laboratory environment, we transfer the uncertainty of the measurement on to the next instrument. What this means is that each person using each instrument in the chain needs to first understand the uncertainty of her/his instrument, and then combine that with the uncertainty from the previous instrument in the chain.

Practically speaking, it is very difficult to develop a complete uncertainty budget for a commercial instrument. The reference instruments at major NMI laboratories are usually not simply off-the-shelf commercial instruments, but carefully constructed research devices. Reference 4 describes the complete procedure applied at NIST to create the uncertainty budget for their reference 45:0 reflectometer and gives an example of the complexity of the process.

There are three components of uncertainty that you *can* examine with a typical commercial instrument [5]:

- Instrument uncertainty: the intrinsic variation present in the instrument when presented with a sample that is not moved between measurements. This shown in Figure 2
- Operator uncertainty: the variation imposed by the operator of the instrument. This is quantified by having a single operator repeatedly place, measure, and remove a sample.
- Uniformity uncertainty: the variation in the reflectance of the sample itself. This is quantified by repeatedly measuring the sample at various locations on its surface.

The details of this procedure are outside of the scope of this paper, but once you have the uncertainties for your instrument, you are ready to make the transfer. All you do now is measure a sample, record the reflectance, and associate the combined uncertainty with the measurement. You have now successfully added to the "*unbroken chain of comparisons*" specified in the NIST documentation.

Closing thoughts

We hope to have explained what traceability is, and what it is not. You may want and need traceable measurements, but (using our example reflectance data in Figures 2, 4, and 5) if you receive data with a reflectance uncertainty of 0.2, that traceable data is most likely not really useful.

Fortunately, all reputable high-accuracy secondary laboratories have uncertainties much smaller than that, and most should approach (but cannot be smaller) than their local National Metrological Institute. You might expect lower uncertainties to cost more money, and they should. NMIs spend countless hours characterizing and maintaining their instruments and documentation, and that does cost something. Secondary labs also have significant costs, such as purchasing and maintaining expensive primary standards. A calibration certificate will give relevant information about maintenance of all primary standards used in the measurements performed.

And finally, just keep in mind that we are all part of the same large family of metrologists, tracing our roots back to our NMIs, each of us maintaining our link in the unbroken chain of uncertainty.

Other useful places to learn about traceability

- NPL Good Practices Guides are well-written and mostly quite practical. Best of all, they are freely available for download here: http://www.npl.co.uk/publications/guides/
- NIST Technical Note 1297 1994 Edition, "Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results." Available for download: https://www.nist.gov/pml/nist-technical-note-1297

References

[1] Refer to www.nist.gov for more details than you could imagine about measurements of all kinds.

[2] Guide to the Expression of Uncertainty in Measurement (International Organization for Standardization, Geneva, Switzerland, 1993).

[3] Image used by permission from M. W. Toews [CC BY 2.5 (https://creativecommons.org/licenses/by/2.5)], from Wikimedia Commons.

[4] EA Early, ME Nadal, Uncertainty Analysis for the NIST 0:45 Reflectometer, Color Research and Application, **33**, Number 2, p100-107 (2008).

[5] HS Fairman, The Assessment of Uncertainty in Spectrophotometry, ISCC Annual Meeting, Manchester, NH (2012).