Clash of cultures: uncertainty vs. accuracy

Chris J. Evans
Zygo Corporation, Laurel Brook Road, Middlefield, CT 06457
cevans@zygo.com

Abstract: Accuracy of measurement is desirable but, by definition, cannot be quantified. ISO gives a formalism for evaluating uncertainty, combining information from many sources, and providing a guard-band for deciding if an optic meets specification.

OCIS codes: 120.4630 Optical inspection; 220.4840 Testing

1. Introduction

"Accuracy" is a positive, desirable attribute. We all want accurate tools and accurate measurements – and we strive for high accuracy, for the pinnacle. "Uncertainty" appears to have only negative connotations; do we really not know what we are doing? Low uncertainty also lacks the cachet, the motivational hook for sales and marketing. Is "uncertainty just what NIST calls accuracy"?

2. Accuracy

Accuracy is well understood conceptually. An accurate part has dimensions very close to those called out on the drawing, but what is an accurate measurement? The ISO vocabulary of terms in metrology [1] provides some useful definitions. First, it defines a "measurand" as "the particular quantity subject to measurement"; in the context of optical testing this might be the matrix of heights used as the hit map in a deterministic finishing process or a parametric description of it, such as the rms, the PSD, or the coefficients of a Zernike fit to some specified order. The second useful definition is of "accuracy of measurement", given as the "closeness of agreement between the result of a measurement and a true value of the measurand". The true value is unknowable making it logically impossible to quantify the accuracy of a specific measurement. As Ref. 1 helpfully remarks in a footnote to the definition, "accuracy is a qualitative concept".

The definition of accuracy, therefore, provides a logical or semantic problem; if that were the only difficulty we could simply consider "accuracy" and "uncertainty" as synonyms – at least for the purposes of optical testing – and focus our attention on other issues. However, the changing nature of optics manufacturing and the rapid increase in outsourcing of individual optical elements leads to the need for conformance testing. Incoming Quality Assurance departments use a variety of instruments to check that the part supplied meets the specification. Since the instrument is imperfect, how do you determine if the part should be accepted? The so called "Gage maker's rule" suggests that the test instrument be 10x more accurate than the tolerance. Establishing the accuracy of an instrument, however, is difficult – particularly if we are considering a complex instrument. Even the simple cases are not so simple. It is easy to say that a 25 mm micrometer is "more accurate" that a 150 mm vernier caliper for measuring the diameter of a 20 mm shaft. But how accurate is it? How do you combine quantitatively the effect of roughness and roundness of the workpiece, operator skill, materials properties, etc. The complexity increases when you ask "How accurate is brand X interferometer" without defining, for example, that the task is to measure a mild asphere under a set of specified conditions including calibration, environment and so on. What do you do if a part "passes" when measured on a profilometer and "fails" when measured on an interferometer when both are claimed to be capable of $\lambda/10$ metrology?

3. Uncertainty

The conceptual difficulties summarized in Section 2, and many others, are avoided by following the ISO Guide to the Expression of Uncertainty in Measurements (GUM)[2]. This internationally agreed methodology applies as well to complex optical measurements as to the more simple mechanical measurements usually invoked to explain it.

A key philosophy [1,2] is that "…the result of a measurement is only an … estimate of the value of the measurand and thus is complete only when accompanied by a statement of the uncertainty of that estimate". The uncertainty of a measurement is defined [1] as a "parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand". Immediately these definitions dispose of one area of confusion. Uncertainty relates to a specific measurement of a specific component; the instrument contributes to the uncertainty of the measurement, but the uncertainty is not a characteristic of the instrument used. The uncertainty for the height map in a given measurement [3] is not the same as the uncertainty in the PV[4], or any other parametric description. Uncertainty is not just the standard deviation (multiplied by a coverage factor) of the measurand, but includes evaluations of all contributions to the dispersion of values that might reasonably be attributed to the measurand

The formalism of the GUM [2] is completely general and allows the use of information from any source. It is standardized, imposes no arbitrary choices over classification of error types, and provides a consistent means of combining such information. Rather than considering "random" and "systematic" errors, the GUM defines Type A uncertainties as those which the metrologist chooses to treat by statistical methods and Type B as those treated by other means, including but not limited to handbook values, analysis, and the judgment of the metrologist. One metrologist may treat a particular uncertainty source statistically (Type A) while another might apply prior knowledge (Type B), presumably resulting in similar uncertainty evaluations (for equally competent metrologists). For example, suppose that 100 mm diameter flats are being measured in a stable environment using a mechanical phase shifting interferometer. One metrologist may decide to treat the noise (electronics noise, turbulence, pzt non-linearity, etc) statistically by taking a number of independent measurements and computing the variance pixel by pixel across the aperture (Type A). Another metrologist might model each contribution from first principles (Type B). A third might have made the same measurement many times under the same conditions and know that the noise will be less than 2 nm per pixel (Type B). None of these approaches is wrong; none superior to the others.

A common problem in optical testing – hinted at in the preceding section – is described by various phrases including "measurement divergence", "uncertainty in the definition of the measurand", and "uncertainty in the realization of the measurand". The result of an optical measurement of a surface will be different from a stylus measurement of the same surface; the difference will depend on the roughness and slopes of the surface, the wavelength of light used, stylus tip radius, stylus force, Hertzian deformation, and so on. Uncertainty in the measurement result may arise from an incomplete definition of the measurand, given the characteristics of the part under test. Equally, the instrument in use may have spatial frequency cut-offs incommensurate with the definition of the measurand. For any of these examples, the resulting uncertainty can be evaluated, given some information about the instrument used and the spatial frequency content of the surface under test. Agreement or disagreement between measurements made on different instruments can be evaluated in the context of the uncertainties of the measurements.

The GUM provides the methodology for evaluating uncertainty, although as we have seen, it is not an exact science. The goal is to develop a good estimate of the uncertainty, based on knowledge of the measurement process. An excruciatingly detailed model is not required – but an understanding of the significant contributors to the uncertainty is needed.

4. Conformance testing and decision rules

Section 3 above has summarized some of the many conceptual advantages of "uncertainty" over "accuracy". The most important advantage, however, comes in the area of conformance testing. In the absence of any other agreement, it may be assumed (and is explicit is some contracts) that the default decision rule of ISO 14253 Part 1[5] applies. This standard states, simply, that in order to prove conformance to specification a vendor shall reduce the tolerance by the measurement uncertainty in the acceptance test. To prove non-conformance, the buyer must increase the tolerance by the uncertainty of his incoming inspection.

All the confusion of Section 2 about what is – and is not – conforming product has now disappeared. Optical measurements of form have single sided tolerances; nobody ever complained about receiving a $\lambda/100$ part against a $\lambda/10$ specification (unless it was a cost-plus contract or late). Suppose that this $\lambda/10$ (63 nm) part is made at a shop where the uncertainty in final inspection is $\lambda/40$ (15 nm); if the vendor measures it as better than $\lambda/13.2$ (48 nm), it should ship. Now suppose the buyer has incoming QA with an uncertainty of $\lambda/30$ (21 nm); if the part measures better than $\sim\lambda/7.5$ (84 nm), it must be accepted.

This simple example illuminates not only the clarity in commercial transactions that comes from the uncertainty formalism embedded in ISO standards, but points out the commercial advantage of improved metrology. For the vendor, improved metrology increases manufacturing headroom; $\lambda/100$ (6 nm) metrology means he only has to make a $\lambda/11$ (57 nm) part to meet a $\lambda/10$ specification. For the buyer, $\lambda/100$ metrology in incoming QA means that a $\lambda/11$ (57 nm) specification is all that is required to ensure a $\lambda/10$ performance.

5. Summary

Accuracy in measurement is desirable, but a source of confusion. Uncertainty calculations are internationally standardized, allow the use of many different sources of information, and provide a simple decision rule for acceptance testing. Evolution in the market for optical components and systems ensures that uncertainty analyses will become the norm.

6. References

[1] ISO, *International Vocabulary of Basic and General Terms in Metrology, 2nd ed*, 1993

[2] ISO, *Guide to the Expression of Uncertainty in Measurement* , 1995

[3] C. Evans, S. Smith, J. Soobitsky, and M. Woynar,"Full area calibration of large optical flats" Proc. ASPE Topical Meeting, April 2009, pp7-10

[4] C. Evans, "Uncertainty evaluation for measurements of peak-to-valley surface form errors" CIRP Annals, **57**, 2008, pp509-12

[5] ISO 14253 Geometrical Product Specification (GPS) - Inspection by measurement of workpieces and measuring equipment - Part 1: Decision rules for proving conformance or non-conformance with specifications, 1st ed., 1998.