doi:10.1017/S1431927616002841
Microsc. Microanal. 22 (Suppl 3), 2016

Uncertainty Is Our Friend - Rethinking Microanalysis Around Uncertainty Metrics

Nicholas W. M. Ritchie[1] and Dale E. Newbury[1]

[1.] NIST, Materials Measurement Science Division, Gaithersburg, MD, USA

A good measurement [electron probe microanalysis or other] is one that provides sufficient information to assist with the answer to a specific question. A good measurement isn't necessarily an accurate measurement and it isn't necessarily a precise measurement. It is however a measurement that is accurate enough and precise enough and reliable enough to answer a question that is important to you or your customer. Yet, how often do we ask ourselves when is good enough, good enough? How do we know whether the measurement we just made is really good enough? How would we improve our measurement protocol to make the current inadequate measurement good enough? Uncertainty metrics provide a rigorous means to answer each of these questions.

By necessity, a useful measurement must be associated with a realistic uncertainty metric. This may seem like an ivory tower ideal to many and inconsistent with real world constraints. We constantly make useful measurements of mass, length, time and any number of other quantities without explicitly associating uncertainties. It is true but there is a difference. We don't explicitly state uncertainties with many everyday measurements and yet we trust them because we have a realistic intuitive sense of the uncertainties. Microanalytical measurements are different. Our intuition fails us when it comes to microanalytical measurements. We don't naturally have a good intuitive feel for the uncertainties (though experts do develop one with time and experience.) Our tools fail us too. In the rare instances in which they do provide uncertainty metrics, the uncertainty metrics are over-optimistic and unrealistic. Most quantitative analysis packages limit themselves to reporting metrics of the measurement precision and totally avoid the question of measurement accuracy.

As a result, we are consistently over-optimistic in our expectations. The software reports contains the same three digits or more of precision regardless of whether the measurement was performed carefully or performed carelessly. There is nothing inherent in the software to motivate us to take care.

We need to change this. As a community, we need tools that guides us towards better measurements. We need *pessimistic software* rather than *optimistic software*. Our current software assumes that we are going to do everything right and that every break is going to go our way. Instead we need software that assumes that unless we verify otherwise that we are making the easy choices rather than the optimal choices. It needs to assume that the sample geometry is sub-optimal and the measurement protocol is sloppy. It needs to report the consequences of these sub-optimal choices as an uncertainty (albeit initially a very large uncertainty). Then it needs to step us through the process of designing a better measurement protocol to reduce the uncertainty. It needs to make specific suggestions and demonstrate how these suggestions will influence the quality of the measurement. It needs to report how the result will improve with better sample preparation like embedding, coating or cross-sectioning. It needs to report how much more accurate the result will be if the sample is polished with 600 grit paper, a 250 nm grit or a 50 nm grit. The software needs to report how much of difference carbon coating or other conductive coatings could make. It should report the level of accuracy we can expect with standardless analysis and the level of accuracy we could expect with standards – simple or similar. It short, it needs to provide the analyst with the information necessary to turn a sloppy, inadequate measurement into a sufficiently good measurement. It needs to inform the analyst of the consequences of various choices and allow the analyst to decide how best to make the measurement good enough through a series of well informed choices.

Furthermore, the software needs to use all the tools available to verify everything we know about spectra. It needs to quantify using complementary algorithms like k-ratio based $\varphi(\rho z)$ algorithms and peak-to-background based algorithms. K-ratio-based algorithms tend to be more precise but peak-to-background algorithms tend to be less susceptible to sample geometry. If the results don't agree to within their associated uncertainties, we need to find out why. We need to quantify the spectra using all available lines and verify the consistency. If the results for the K and L lines don't agree to within the uncertainties, we need to know why. The software needs to simulate the resulting composition and compare the simulated spectra to the measured spectra[1]. It needs to verify every piece of information entered by the analyst. Are the standards really what the analyst believes? (Quantify them.) Are the spectra consistent with the conditions under which they were supposedly collected? (Perform sanity checks.) In short, the software needs to identify as many failure modes as possible and suggest corrective action. We need to build tools that not only make it possible to make good measurements but *tools that make it likely that we will make good measurements* - not only for the expert but also for the novice and the occasional user too. For every measurement made by an expert, there are many times more made by novices. We need to ensure that they are all sufficiently reliable.

There is a lot of work to be done. As a community, while we have considered uncertainty metrics based on precision for decades [2], we have only just started to consider measurement accuracy[3, 4] and we have only scratched the surface. In addition to uncertainty resulting from algorithmic choice, mass absorption coefficients and backscatter coefficients, we need to address uncertainties due to sample morphology, sample preparation, standardless analysis and others. We need to encapsulate the intuition of experts into the software. We need to change the culture from a culture that assumes the best to one that is realistic about human foibles and our preferences for shortcuts. Making [bad] microanalytical measurements is quick and easy but making reliable microanalytical measurement will always require care and diligence.

References:

[1] Duncumb, Peter, Ian R. Barkshire and Peter J. Statham. Microscopy and Microanalysis **7.04** (2001): 341-355.
[2] Marinenko, R. B. and S. Leigh. "Uncertainties in electron probe microanalysis." IOP Conference Series: Materials Science and Engineering. Vol. 7. No. 1. IOP Publishing, 2010.
[3] Armstrong, John T., J. Donovan and P. Carpenter. Microscopy and Microanalysis **19.S2** (2013): 812-813.
[4] Ritchie, Nicholas WM and Dale E. Newbury. Analytical chemistry **84.22** (2012): 9956-9962.

https://doi.org/10.1017/S1431927616002841 Published online by Cambridge University Press