Chemometrics and Intelligent Laboratory Systems 159 (2016) 164–173

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemolab

Improved uncertainty quantification in nondestructive assay for nonproliferation

Tom Burr[a,*], Stephen Croft[b], Ken Jarman[c], Andrew Nicholson[b], Claude Norman[a], Stephen Walsh[d]

[a] Information Management, Safeguards Department, International Atomic Energy Agency, Austria
[b] Nuclear Security and Isotope Technology, Oak Ridge National Laboratory, USA
[c] Applied Statistics and Computational Modeling, Pacific Northwest National Laboratory, USA
[d] Quality Management, Department of Safeguards, International Atomic Energy Agency, Austria

ARTICLE INFO

Anniversary Issue in Chemometrics and Intelligent Laboratory Systems.

Keywords:
Approximate Bayesian computation
Calibration
Item-specific bias
Non-destructive assay
Uncertainty quantification
Random error variance
Systematic error variance

ABSTRACT

This paper illustrates methods to improve uncertainty quantification (UQ) for non-destructive assay (NDA) measurements used in nuclear nonproliferation. First, it is shown that current bottom-up UQ applied to calibration data is not always adequate, for three main reasons: (1) Because there are errors in both the predictors and the response, calibration involves a ratio of random quantities, and calibration data sets in NDA usually consist of only a modest number of samples (3–10); therefore, asymptotic approximations involving quantities needed for UQ such as means and variances are often not sufficiently accurate; (2) Common practice overlooks that calibration implies a partitioning of total error into random and systematic error, and (3) In many NDA applications, test items exhibit non-negligible departures in physical properties from calibration items, so model-based adjustments are used, but item-specific bias remains in some data. Therefore, improved bottom-up UQ using calibration data should predict the typical magnitude of item-specific bias, and the suggestion is to do so by including sources of item-specific bias in synthetic calibration data that is generated using a combination of modeling and real calibration data. Second, for measurements of the same nuclear material item by both the facility operator and international inspectors, current empirical (top-down) UQ is described for estimating operator and inspector systematic and random error variance components. A Bayesian alternative is introduced that easily accommodates constraints on variance components, and is more robust than current top-down methods to the underlying measurement error distributions.

1. Introduction

In metrology, uncertainty is a term that characterizes the dispersion of estimates of a true quantity known as the measurand. In applications of NDA used in nuclear nonproliferation, the measurand is typically the amount of special nuclear material (SNM, such as U or Pu) in an item. NDA uses calibration and modelling to infer SNM mass on the basis of radiation particles, such as neutrons and gammas emitted by the item and registered by detectors. An example used throughout this paper is a common NDA technique consisting of a gamma detector that can be calibrated to measure the ^{235}U enrichment (atom percent of ^{235}U/U in an item) by applying the enrichment meter principle (EMP, see Section 3). For any NDA technique, one can take a first-principles physics-based or "bottom-up" approach to UQ by considering each key step and assumption of the particular method. Because most NDA methods rely on a calibration step, and because calibration uncertainty has not been adequately quantified for NDA applications, this paper focuses on calibration as an example of bottom-up UQ. The second focus is a new Bayesian approach to "top-down" UQ. Top-down UQ compares measured values for selected items to their corresponding nominal values, or to other measurements on the same items (Section 4).

A well-known guide for bottom-up UQ in metrology is the Guide to the Expression of Uncertainty in Measurement [20]. The GUM also briefly mentions top-down UQ in the context of applying analysis of variance to data from inter-laboratory comparisons, in which multiple measurements are made on the same or similar items by multiple laboratories. Although the GUM is useful, it is being revised because it has known limitations [2]. For example, the GUM uses Bayesian and frequentist concepts without clearly distinguishing them [17,37,5], and provides very little technical guidance regarding calibration as an example of bottom-up UQ, or regarding top-down UQ.

In the NDA applications considered, the facility operator declares

* Corresponding author.
E-mail address: t.burr@iaea.org (T. Burr).

http://dx.doi.org/10.1016/j.chemolab.2016.10.007
Received 22 June 2016; Received in revised form 5 October 2016; Accepted 8 October 2016
Available online 14 October 2016
0169-7439/ Published by Elsevier B.V.

the SNM mass in each item. Some of those items are randomly selected for verification measurement by inspectors, who often use NDA. This is a challenging NDA application, because the detector is brought to the facility where ambient conditions can vary over time, and because the items to be assayed are often heterogeneous in some way. Because of such challenges, "dark uncertainty" [33]; Wuester et al., 2016) can be large, as is evident whenever bottom-up UQ predicts smaller uncertainty than is observed in top-down UQ [4]. Here, the term "uncertainty" is used in a general sense, and is defined by context; for example, the uncertainty of an assay method is often defined as the reproducibility standard deviation as estimated in an inter-laboratory comparison [15]; ISO 21748:20, 2010; [35]. Verification measurements are regarded as a special case of an inter-laboratory evaluation (Sections 2 and 4).

For top-down UQ applied to SNM measurements of the same item by both the operator and the inspector, this paper describes current approaches to estimate operator and inspector systematic and random error variance components separately. Systematic errors and random error components must be separated because their mode of propagation can be different, depending on the context and end use. Currently, random error variance estimates (from paired data) are based on the Grubb's estimator, or variations of the Grubb's estimator, which was originally developed by Grubb's to separately estimate random error variance of each of two methods applied to each of several items, without repetition of measurement by either method [19,26]. These applications of Grubb's estimators are illustrated, and a simple but effective Bayesian alternative to the Grubb's estimator is introduced, so that parameter constraints and prior information regarding the relative magnitudes of variance components can be exploited to improve top-down UQ.

This paper is organized as follows. Section 2 gives background on UQ for NDA. Section 3 describes the EMP and illustrates why simulation is necessary for improved bottom-up UQ for calibration data. Section 4 reviews currently-used top-down UQ and introduces a new Bayesian option that applies approximate Bayesian computation. Section 5 is a discussion and summary.

2. Background

The need to improve UQ for NDA measurements is recognized [4,5,7,8], partly from observing larger uncertainty in top-down UQ applied to paired measurements from the operator and inspector than currently-used bottom-up UQ predicts. Statistical perspectives of these observations are described in the next sub-sections.

2.1. Bottom-up UQ

For bottom-up UQ, the GUM assumes the measured value is produced in a manner that can be expressed using a measurand equation or algorithm that relates input quantities (data collected during the measurement process and relevant fundamental nuclear data used, for example, in attenuation corrections) to the output (the final measurement value).The GUM's main technical tool is a first-order Taylor approximation applied to the measurand equation

$$Y = f(X_1, X_2, \ldots, X_N), \tag{1}$$

which relates input quantities $X_1, X_2, \ldots, X_N$ to the measurand Y. The input quantities can include estimates of other measurands, or of calibration parameters, so Eq. (1) is quite general. The estimated variance of each X, σ_i^2, and any covariances, $\sigma_i\sigma_j\rho_{i,j}$, between pairs of Xs are then propagated using the Taylor approximation to obtain $\sigma_Y^2 \approx \sum_{i=1}^{N}\left(\frac{\partial f}{\partial x_i}\right)^2\sigma_i^2 + 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j}\sigma_i\sigma_j\rho_{i,j}$ (or using simulation if the Taylor approximation is not sufficiently accurate) to estimate the variance in Y, σ_Y^2.

The GUM does not clearly distinguish between Bayesian and nonBayesian methods. However, Eq. (1) is expressed in a convenient form for application of first-order Taylor series and for depicting that the estimated value Y of the measurand is a random variable, regardless of whether the left side of Eq. (1) is expressed as Y or as $\hat{\mu}_Y$. The hat notation is a convention for denoting an estimator, so $\hat{\mu}_Y$ is a random variable, and μ_Y (which is also denoted as y_T, where the "T" denotes the true value) denotes the unknown true value of the measurand. In some cases, there could be a function, not necessarily with the same function $f()$ as in Eq. (1), relating the true inputs to the true outputs $y_T = g(x_{1,T}, x_{2,T}, \ldots, x_{N,T})$. In a nonBayesian framework, the true quantities are regarded as fixed and unknown constants. In a Bayesian framework, even the true quantities are regarded as random variables. The GUM also does not describe calibration in much detail; in the statistical literature on calibration, the true response μ_Y is nearly always written as being a function of the true measurand value μ_X. Therefore, Section 3 assumes a simple model, $\mu_Y = \beta_0 + \beta_1\mu_X$ and the roles of X and Y are reversed from that in Eq. (1).

2.2. Calibration UQ as an example of bottom-up UQ

Following calibration of the EMP using reference materials having nominal values of enrichment of ^{235}U (known to within relatively small uncertainty), Eq. (1) can be expressed as $\hat{\mu}_Y = \hat{\beta}_O + \hat{\beta}_1 X_N$, where $\hat{\mu}_Y$ is the estimated enrichment, $\hat{\beta}_0$ and $\hat{\beta}_1$ are parameters estimated from calibration data, X_N is the net gamma count rate in the spectral region of interest, and three uncertain inputs in mapping to Eq. (1) are $X_1 = \hat{\beta}_0$, $X_2 = \hat{\beta}_1$, and $X_3 = X_N$. The estimates $\hat{\beta}_0$ and $\hat{\beta}_1$ will vary in predictable ways across repeats of the calibration.

In Section 3, the convention in most statistical literature to reverse the roles of X and Y from that in Eq. (1) is followed, so X denotes the quantity to be inferred (the measurand value here) and Y denotes the detected radiation count rate. Then, in the case of reverse regression (see Section 3), Eq. (1), for a different $f()$ from the $f()$ in Eq. (1), can be expressed as $X = f(Y_1, Y_2, \ldots, Y_N) = \hat{\alpha}_0 + \hat{\alpha}_1 Y$, identifying $Y_1 = \hat{\beta}_0$, $Y_2 = \hat{\beta}_1$, and $Y_3 = Y$. Following calibration on training data consisting of n (x_i, y_i) pairs (lower case denotes observed value of random variables), the three "input quantities" $Y_1 = \hat{\beta}_0$, $Y_2 = \hat{\beta}_1$, and $Y_3 = Y$ have variances and covariances that can be estimated. However, in most applications of calibration in NDA, accurate estimation of these variances and covariances requires simulation because analytical approximations have been shown to be inadequate (Section 3). Section 3 illustrates that extensions to standard regression results applied to the special case of Eq. (1), $\hat{\mu}_Y = \hat{\beta}_O + \hat{\beta}_1 X_N$ for bottom-up UQ are not sufficiently accurate, and so simulation is used to perform bottom-up UQ, including assessment of item-specific bias and long-term bias.

Elster [17] notes that the GUM provides almost no quantitative guidance on bottom-up UQ for calibration (but more guidance is planned in a future GUM version). Toward using GUM's Eq. (1) in simple calibration examples, note that Eq. (1) assumes that $f()$ is known and deterministic, and that the X's are random variables, possibly having nonzero correlations. This implies that Y is a random variable. Hence, the GUM implicitly takes a Bayesian approach, which is now recognized [17], with a known probability density if the X's have known probability densities, or a known variance if only the variances of the X's and covariances between the X pairs are known. However, in some applications, $f()$ is not perfectly known, or there is some modeling error not captured in $f()$, so one could include an error term in Eq. (1). It is case-specific whether including such an error term is needed. In the case of simple linear calibration in Section 3, the GUM assumption that $f()$ is known exactly is reasonable, and expressing Eq. (1) as $X = f(Y_1, Y_2, \ldots, Y_N) = \hat{\alpha}_0 + \hat{\alpha}_1 Y$, indicates how the estimate X is computed, and how to assign both "systematic" and "random" error variances to X (Section 3). For example, and to introduce notation used in top-down UQ (Section 2.3), one could express the estimate as $X = \mu_X + S + R$, where μ_X denotes the true value of the measurand, S

Download English Version:

https://daneshyari.com/en/article/5132362

Download Persian Version:

https://daneshyari.com/article/5132362

Daneshyari.com