Chemometrics and Intelligent Laboratory Systems 141 (2015) 94–99

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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

A Bayesian approach to an interlaboratory comparison

George S. Jackson [a], Paul Muzikar [a,*], Brent Goehring [b]

[a] PRIME Laboratory, Purdue University, West Lafayette, IN 47907, United States
[b] Department of Earth and Environmental Sciences, Tulane University, 70118, United States

ARTICLE INFO

Article history:
Received 27 October 2014
Received in revised form 16 December 2014
Accepted 18 December 2014
Available online 27 December 2014

Keywords:
Bayesian
Interlaboratory comparison
Accelerator mass spectrometry

ABSTRACT

Interlaboratory comparisons are an important check of the quality of a measurement technique. In this paper we examine the accelerator mass spectrometry (AMS) measurement of ^{41}Ca, an unstable isotope of calcium that has emerged as a valuable tracer for a variety of studies. We use a Bayesian framework to explore the quality and consistency of the AMS measurements made by Lawrence Livermore National Laboratory (LLNL) and the Purdue Rare Isotope Measurement Laboratory (PRIME Lab). This framework should be generalizable to other interlaboratory comparisons. The laboratories measured 47 samples, with each lab measuring an aliquot of each sample. The Bayesian approach allowed us to derive a probability distribution for four parameters reflecting the quality of the data, and to then address the following questions: (1) Are the results from the two labs consistent? (2) Are the uncertainties quoted by the two labs reasonable? We find that any consistent offset between the two labs is negligible, and that the uncertainties may be slightly underestimated.

1. Introduction

Interlaboratory comparisons, in which aliquots of the same samples are measured at different laboratories and the results compared, serve many important purposes [1–3]. These purposes range from routine quality control to investigating problems arising in challenging new techniques. In a recent paper [4] we discussed such a comparison, in which PRIME Lab (the Purdue Rare Isotope Measurement Laboratory) and LLNL (Lawrence Livermore National Laboratory) performed ^{41}Ca accelerator mass spectrometry (AMS) [5] measurements on a series of samples. These samples were not special, standard material; rather they were unknowns that had already been chemically prepared and measured at PRIME Lab for other studies [6,7]. Thus, the results are representative of the actual quality of the measurements being done at both labs. In our paper [4], the results from the two labs were compared in a variety of ways; one particular method was grounded in a Bayesian approach. In this new paper we focus on the Bayesian approach [8–10], expanding on our original discussion. We analyze more aspects of the data, and explore more possibilities that are inherent in the Bayesian methodology. We hope in this paper to both provide a more in-depth look at our particular data set, and more generally, to provide ideas useful in a wider context.

^{41}Ca is an unstable isotope of calcium, with a half-life of 1.03×10^5 years. This half-life is short enough to make ^{41}Ca rare in the environment, but long enough to make decay counting impractical. It is an ideal candidate for AMS studies, since AMS has an extremely low detection limit [5]. Its AMS measurement is reported as the following ratio: $R = {}^{41}\mathrm{Ca}/\mathrm{Ca}$, where the numerator is the number of ^{41}Ca atoms in the sample, and the denominator is the total number of the stable isotopes of calcium (mainly ^{40}Ca) in the sample.

^{41}Ca has emerged as a valuable tool for a variety of studies. For example, ^{41}Ca has been used to glean information about pre-atmospheric meteorite dimensions [11,12]. For biomedical studies, the low detection limit and the long half-life mean that a single, safe dose can be given to an animal or a human subject, and then research can be conducted for the biological lifetime of the subject without another dose ever being administered [13]. In the 1990's it was suggested that ^{41}Ca could be used as a tracer to elucidate calcium metabolism in humans [14,15], and this has now emerged as the most important application of this nuclide [16–19]. For example, clinical trials have studied the efficacy of both commercial anti-osteoporosis drugs and of botanical treatments [6,7]. The increasing use of ^{41}Ca in these biomedical studies adds importance to our interlaboratory comparison.

The technical challenges of AMS ^{41}Ca measurements mostly stem from the low ion source currents that are produced when CaF_2 is used as the target material, and from the large number of interfering ions that are present in the detector when measuring ^{41}Ca. The easily prepared compound CaF_2 is the preferred target material for studies that require high sample throughput.

The data set we will consider is a set of measurements on 47 samples, in which R ranged from about 10^{-10} to about 36×10^{-10}. (The data are presented in [4]). PRIME Lab and LLNL separately measured an aliquot of each sample, and reported an R value and an uncertainty for each sample. As already mentioned, these samples were not standards, so the true values are not known. We will use a Bayesian approach to address the following questions: (1) Are the results from the two laboratories consistent? (2) Are the uncertainties quoted by the

* Corresponding author. Tel.: +1 765 494 3046.
E-mail address: pmuzikar@purdue.edu (P. Muzikar).

http://dx.doi.org/10.1016/j.chemolab.2014.12.006
0169-7439/

two labs reasonable? Since the samples were all unknowns, the answer to a third question, ((3) how close are the results to the true values?), involves quantitative input that goes beyond the data; this input is incorporated in the prior distribution, as we discuss.

2. Theoretical development

2.1. Bayesian approach

In this section, we develop the formalism we will use. We enlarge on the Bayesian discussion in our previous paper [4], again following the approach presented by Dose [10] (see also [8,9]). A closely related problem arises when a group of independent laboratories measure the same quantity, and we want to combine the results to determine the best estimate. For approaches related to ours, see [20–22].

Let $[x_i, \sigma_{Li}]$ be the LLNL data points and their associated uncertainties for sample i, and $[y_i, \sigma_{Pi}]$ be the PRIME Lab results for sample i. So, for example, LLNL asserts that for sample i the probability density for the true value has an average equal to x_i and a standard deviation of σ_{Li}. We will define four parameters, and use a Bayesian approach to formulate a probability distribution for them. We will use, throughout this paper, the notation $p(A|B)$, which means the probability of A, given B.

First, we allow for a consistent offset of $-a$ of the LLNL samples from the true value, such that $x_i + a$ is the true value. We define a similar consistent offset for the PRIME Lab samples, such that $y_i + b$ is the true value. We will see what the data allow us to say about these offsets. We note that other choices are possible; for example, we could assume that the LLNL offset is proportional to x_i, so that the true value is $x_i + \gamma x_i$, where γ is the same for all of the LLNL samples. In this paper we will focus our study on the offsets a and b.

Next, we also allow for the fact that the labs may have misestimated their uncertainties. A simple model for this is that $v\sigma_{Li}$ and $w\sigma_{Pi}$ are the true uncertainties of LLNL and PRIME Lab. This defines the two non-negative parameters v and w. For example, if $v < 1$, this means that LLNL is overestimating the uncertainty it assigns to its results, while if $v > 1$ the uncertainties are being underestimated.

So, we have defined four parameters which characterize the quality of the data; a and b tell us how accurate the measurements are, while v and w tell us if the quoted uncertainties are correct. We should emphasize that our approach is not based on a more fundamental, detailed model of the experimental situation. Rather, these four parameters are a simple way to allow for a range of errors in the reported results. As mentioned above, other choices are possible; we hope that this paper makes clear how alternative models could be handled in a Bayesian spirit.

In the Bayesian approach, we do not simply use the data to estimate the values of these parameters. Instead, the data are used to generate a probability distribution (since these are continuous parameters, it is in fact a probability density function, or pdf) for these parameters. This is a more informative result, since from the complete pdf we can compute all averages (such as $< a >, < a^2 >, \ldots$), generate contour plots, and calculate probability distributions for any subset of our parameters.

To derive our basic equation, we start with Bayes theorem:

$$p(a, b, v, w|data)p(data) = p(data|a, b, v, w)p(a, b, v, w) \tag{1}$$

We first discuss the meaning of each of the four probability densities in the previous equation. $p(a, b, v, w|data)$ is the joint probability density for our four unknowns (a, b, v, w) given the data. This function is the result we want. $p(data)$ is the probability of acquiring the data; it does not depend on the four unknowns, and can be dropped if we are willing to normalize our result. Thus, we write

$$p(a, b, v, w|data) = cp(data|a, b, v, w)p(a, b, v, w) \tag{2}$$

The value of c will be determined by requiring that

$$\int da \int db \int dv \int dw p(a, b, v, w|data) = 1 \tag{3}$$

Next we discuss $p(data|a, b, v, w)$; this is the probability of obtaining our data, given the values of a, b, v, w, and is also known as the likelihood. We must use our knowledge of the measurement process to construct an explicit expression for this likelihood.

Finally, we come to the pdf $p(a, b, v, w)$. This is called the prior, and embodies our information about the values of these parameters before we have made our measurements. Thus, a nice way to view Eq. (2) is as follows. The prior, $p(a, b, v, w)$ represents our knowledge of the four parameters before we collected the AMS data. This knowledge is then updated through the likelihood, in order to produce $p(a, b, v, w|data)$, which represents our improved state of knowledge.

2.2. Derivation of the likelihood function

To derive the likelihood function, we start with

$$p(data|a, b, v, w) = \prod_i p(x_i, y_i|a, b, v, w) \tag{4}$$

where $\{x_i\}$ are the R values measured at LLNL, and $\{y_i\}$ are the R values measured at PRIME Lab. Then, for each sample we may write

$$p(x_i, y_i|a, b, v, w) = \int dz_i p(x_i, y_i|z_i, a, b, v, w)p(z_i|a, b, v, w) \tag{5}$$

where z_i is the (unknown) true R value for sample i.

We must now make several more definite assumptions. We take the measurements to have a Gaussian distribution about the true value, with the shifts a and b included:

$$p(x_i, y_i|z_i, a, b, v, w) \propto \left(\frac{exp\left(-\frac{(x_i + a - z_i)^2}{2v^2\sigma_{Li}^2} \right)}{v\sigma_{Li}} \right) \left(\frac{exp\left(-\frac{(y_i + b - z_i)^2}{2w^2\sigma_{Pi}^2} \right)}{w\sigma_{Pi}} \right) \tag{6}$$

Note that the parameters v and w each appear in two places, since the overall normalization of each Gaussian factor involves its uncertainty. Numerical factors that do not involve the parameters a, b, v and w may be omitted, since we will normalize at the end. The Gaussians are a reasonable choice, and are convenient for evaluating the integral. We take the conditional probability $p(z_i|a, b, v, w)$ to be a constant, since the true value z_i should not depend on these parameters, and will have a very broad range of possible values.

We may then evaluate the integral in Eq. (5) to obtain:

$$p(x_i, y_i|a, b, v, w) \propto \left(v^2\sigma_{Li}^2 + w^2\sigma_{Pi}^2 \right)^{-1/2} exp\left(-\frac{(x_i - y_i + a - b)^2}{2(v^2\sigma_{Li}^2 + w^2\sigma_{Pi}^2)} \right) \tag{7}$$

We note that the likelihood involves a and b only in the combination $a - b$. Thus, instead of using a and b, it is instructive to make a change of variables:

$$r = a - b \qquad Q = (a + b)/2 \tag{8}$$

We then write our final form for the likelihood as

$$p(data|r, Q, v, w)) = \prod_i \left(v^2\sigma_{Li}^2 + w^2\sigma_{Pi}^2 \right)^{-1/2} exp\left(-\frac{(x_i - y_i + r)^2}{2(v^2\sigma_{Li}^2 + w^2\sigma_{Pi}^2)} \right) \tag{9}$$