

## GEOLOGICAL MATERIALS TESTING AND UNCERTAINTY CALCULATIONS: A SIMPLE GUM-BASED ALGORITHM

Xirouchakis D.<sup>1</sup> and Bouzinos A.<sup>2</sup>

<sup>1</sup> GeoTerra Ltd, Geomechanics & Quality Control Laboratory, 12 Anthrakorichon Street, 142 35  
Nea Ionia, Athens, Greece, dxirouch@gmail.com

<sup>2</sup> Energy Resources Consulting Pty Ltd, 3/55 Clarence St PO Box 54 Coorparoo, Qld 4151  
Australia, antonis.bouzinis@energyrc.com.au

### Abstract

*We have applied a simple GUM-based procedure to estimate the uncertainties of physical and mechanical properties in geological materials. First, we define the quantity to measure and decide whether we want to work with units or relative quantities. Subsequently, we calculate the repeatability standard deviation ( $s_r$ ) and the standard uncertainty. If we have proficiency test data or use certified reference materials, we use them to estimate the laboratory bias, the reproducibility standard deviation ( $s_R$ ) and the reproducibility standard uncertainty. We also make sure that we know or have estimated the standard uncertainty of the instruments that we use in the measurements. The latter is typically taken from the instrument calibration or precision statement. We estimate the standard uncertainty of the reference materials and the standard uncertainty of the laboratory bias. The final two steps include the calculation of (1) the laboratory standard uncertainty uncorrected for bias and corrected for bias, and (2) the laboratory expanded uncertainty at the 95% confidence limit.*

**Keywords:** uncertainty, repeatability, reproducibility, vitrinite reflectance, specific gravity.

### Περίληψη

*Εφαρμόσαμε μια απλή διαδικασία υπολογισμού αβεβαιοτήτων φυσικών και μηχανικών παραμέτρων γεωλογικών υλικών με βάση τον οδηγό υπολογισμού αβεβαιοτήτων του Διεθνούς Γραφείου Μέτρων & Σταθμών. Αρχικά, ορίζουμε την ποσότητα προς μέτρηση και αποφασίζουμε αν θέλουμε να χρησιμοποιήσουμε μονάδες ή σχετικά μεγέθη. Ακολούθως, υπολογίζουμε την τυπική απόκλιση της επαναληψιμότητας ( $s_r$ ) και την τυπική αβεβαιότητα. Επίσης, χρησιμοποιούμε δεδομένα δοκιμών από διεργαστηριακά ή υλικά αναφοράς ώστε εκτιμηθεί η συστηματική απόκλιση του εργαστηρίου, η τυπική απόκλιση της αναπαραγωγιμότητας ( $s_R$ ) και η τυπική αβεβαιότητα της αναπαραγωγιμότητας. Εκτιμούμε ή υπολογίζουμε την τυπική αβεβαιότητα των οργάνων που χρησιμοποιούμε στις μετρήσεις, την τυπική αβεβαιότητα του υλικού αναφοράς και την τυπική αβεβαιότητα της συστηματικής απόκλισης του εργαστηρίου. Τα τελευταία βήματα περιλαμβάνουν τον υπολογισμό (1) της συνολικής τυπικής αβεβαιότητας του εργαστηρίου με ή χωρίς διόρθωση για την συστηματική απόκλιση και (2) την διευρυμένη αβεβαιότητα του εργαστηρίου σε επίπεδο εμπιστοσύνης 95%.*

**Λέξεις κλειδιά:** αβεβαιότητα, επαναληψιμότητα, αναπαραγωγιμότητα, ανακλαστικότητα βιτρίνη, ειδικό βάρος.

## 1. Introduction

Understanding and evaluating measurement uncertainty is important to fully exploit laboratory results. The basic step involves the estimation of uncertainty either for a single operator or within a laboratory, or among laboratories. A measurement technique should (1) distinguish between error and uncertainty; (2) recognize that all measurements have uncertainty; (3) identify types of error, sources of error, and how to detect or minimize errors; and (4) estimate, describe, and express uncertainty in measurements and calculations.

The Guide to the Expression of Uncertainty (GUM) published by the Joint Committees for Guides in Metrology establishes the general rules for evaluating and expressing uncertainty in measurements (BIPM, 2008). For further explanation or clarifications to definitions and detailed procedures the reader may refer to the excellent publications by NIST, EURACHEM, NordTest, UKAS, and ASTM. The GUM defines uncertainty as the parameter that is associated with the result of a measurement and characterizes the dispersion of the values that could be attributed to the measurand. In addition, the GUM recognizes two types of uncertainty components: (1) type A that are evaluated using statistical analysis of measurement series, and (2) type B that are evaluated using other than statistical analysis of measurement series.

For type A uncertainty components, the standard uncertainty is

$$u = \frac{s}{\sqrt{n}}, \quad (1)$$

where  $s$  is the standard deviation and  $n$  the number of measurements used to calculate the standard deviation. Examples of type B components are the resolution of a thermometer and the tolerance of an instrument. Type B components are commonly described by rectangular-, trigonal-, and U-type distributions. To estimate the total uncertainty associated with the quantity  $Y = f(X_i)$ , the GUM combines the type A and B components of the uncertainty of the measurand  $y = f(x_i)$ , which is designated as  $u_c(y)$  (BIPM, 2008).

Assuming normal (Gaussian) probability distribution for  $y$  and its combined standard uncertainty  $u_c(y)$ , the quantity  $Y$  is greater than or equal to  $y - u_c(y)$  and is less than or equal to  $y + u_c(y)$  at an approximate level of confidence of 68%, which is expressed as  $Y = y \pm u_c(y)$ . The combined standard uncertainty  $u_c$  is used to express the uncertainty of many measurement results. Nonetheless, it is often required to express the uncertainty at the 95% or 99% confidence level. This is called the expanded uncertainty  $U$  and is obtained by multiplying  $u_c(y)$  by a coverage factor  $k$  ( $k = 2$  for 95% confidence and  $k = 3$  for 99% confidence), thus  $U = k \cdot u_c(y)$  and  $Y = y \pm U$ .

## 2. Materials and Methods

We have used the GUM (BIPM, 2008) and the NordTest TR 537 report (Magnusson et al., 2003) to formulate a simple algorithm for calculating the uncertainties of physical and mechanical parameters that are commonly determined in testing laboratories of geological materials by using standard test methods. We introduce the relevant parameters and basic equations that we have used under the assumption of normally distributed data and, subsequently, we give two examples of our adopted approach.

The agreement between results obtained with the same method on identical test or reference material under the same conditions (job done by one person, in the same laboratory, with the same equipment, at the same time or with only a short time interval) is the best precision a laboratory can obtain. This is called the within-batch precision or single-operator repeatability. The repeatability standard deviation is estimated by

$$s_r = \sqrt{\frac{\sum_i (x_i - \mu)^2}{n-1}} \quad \text{and} \quad (2)$$

$$\mu = \frac{\sum_i x_i}{n}, \quad (3)$$

where  $n$  represents the number of measurements  $x_i$  (e.g., Mandel, 1964). In the case of combining several ( $k$ ) series of measurements performed under similar conditions, which is also common in testing laboratories, the repeatability standard deviation is estimated with Eq. (4)

$$s_r = s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + (n_3 - 1)s_3^2 + \dots + (n_k - 1)s_k^2}{(n_1 - 1) + (n_2 - 1) + (n_3 - 1) + \dots + (n_k - 1)}}. \quad (4)$$

The repeatability limit  $r$  is calculated by using  $r = (1.96 \times \sqrt{2}) \times s_r = 2.8 \times s_r$ . This means that the difference between two measurements made under repeatability conditions is unlikely to exceed the repeatability limit  $r$  (with 95% confidence).

The within-laboratory repeatability is a measure of the agreement between results obtained with the same method on identical test material under different conditions (execution by different persons, with the same or different equipment, in the same laboratory, at different times) is a more realistic type of precision for a method over a longer span of time when conditions are more variable than defined for repeatability. The measure is the standard deviation of these results  $s_L$ , which is also called the between-batch precision. The within-laboratory repeatability limit  $R_L$  is calculated by  $R_L = 2.8 \times s_L$ ; that is, two measurements of the same material from the same laboratory should not exceed  $R_L$ . The within-laboratory repeatability standard deviation is estimated as above.

Reproducibility is the agreement between results obtained with the same method on identical test or reference material under different conditions (execution by different persons, in different laboratories, with different equipment and at different times). The measure of reproducibility  $R$  is the standard deviation of these results  $s_R$  and is defined by  $R = 2.8 \times s_R$ ; that is, the difference between the measurements of two different laboratories under reproducibility conditions should not exceed  $R$  (with 95% confidence). The reproducibility standard deviation is typically derived from interlaboratory (*IL*) measurements and proficiency tests (*PT*), and is given by the agency that is responsible for these tests. Nonetheless, it can be readily calculated providing that all the measurements from all the laboratories are known. From the above it is anticipated that  $s_R > s_L \geq s_r$ .

To calculate the laboratory uncertainty, we first assume uncorrelated uncertainties, then define the quantity  $y$  to measure, and finally decide whether we want to work with units or relative quantities (e.g.,  $s$  or  $\%s = 100 \times s/\mu$ ). The latter is preferred. Furthermore, in the case a laboratory uses reference materials (RF), or participates in proficiency tests, the bias, which is a measure of nonrandom error, is estimated. Second, we calculate the single-operator ( $s_r$ ) or the within-laboratory repeatability standard deviation ( $s_L$ ), and the repeatability standard uncertainty

$$u_r = \frac{s_r}{\sqrt{n}}, \quad \text{or} \quad (5)$$

$$u_r = \frac{s_L}{\sqrt{n}}.$$

Third, we estimate the standard uncertainty associated with all the laboratory apparatus that we use in the measurements

$$u_{instruments} = \sqrt{\sum_i u_{instrument}^2} \quad (6)$$

Fourth, from proficiency tests, or reference materials, we estimate the standard uncertainty of the reference material(s)

$$u_{RF} = \frac{S_R}{\sqrt{n}} \quad (7)$$

and the standard uncertainty of the laboratory bias

$$u_{bias} = \sqrt{RMS_{bias}^2 + u_{RF}^2} \quad (8)$$

$$\text{where } RMS_{bias} = \sqrt{\frac{\sum_i (Y_i - Y_{ref})^2}{n}} \quad ,$$

or the percent relative  $RMS_{bias}$

$$pRRMS_{bias} = 100 \cdot \sqrt{\frac{\sum_i \left(\frac{Y_i - Y_{ref}}{Y_{ref}}\right)^2}{n}} \quad .$$

(9)

The fifth step is the calculation of the bias-uncorrected laboratory standard uncertainty

$$u_{Lab, bias-uncorrected} = \sqrt{(u_r^2 + u_{instruments}^2)} \quad ,$$

(10)

or bias-corrected laboratory standard uncertainty

$$u_{Lab, bias-corrected} = \sqrt{(u_r^2 + u_{instruments}^2 + u_{bias}^2)} \quad .$$

(11)

Finally, we calculate the expanded laboratory uncertainty, either bias-uncorrected or bias-corrected,

$$U_{Lab} = k \times u_{Lab}$$

(12)

at the 95% confidence limit. The factor  $k$  depends on the degrees of freedom ( $\nu = n - 1$ ), (Mandel, 1964).

### 3. Examples

#### 3.1. Vitrinite Reflectance

In this example, the laboratory (Energy Resources Consulting) estimates the single-operator and within-laboratory uncertainty using a reference material (ICCP Single Coal Accreditation

Program). The test involves the determination of the vitrinite reflectance according to ISO 7404-5. Six different samples are measured by three laboratory scientists, and the data are listed in Table 1.

**Table 1 – Vitrinite reflectance data measured on coal samples.**

Sample	User	N	$\mu$	$s$	% $s$	$s_p$	% $s_p$	$\mu_{RF}$	$SRF$	% $SRF$
#1	A	100	0.64	0.047	7.3	0.04	6.3	0.64	0.04	6.6
	B	100	0.63	0.050	7.9					
	C	100	0.63	0.005	0.8					
#2	A	100	0.67	0.042	6.2	0.05	7.0	0.67	0.04	5.9
	B	100	0.67	0.045	6.7					
	C	100	0.67	0.053	7.9					
#3	A	100	0.89	0.054	6.1	0.05	5.8	0.88	0.06	7.3
	B	100	0.91	0.045	4.9					
	C	100	0.90	0.058	6.4					
#4	A	100	1.17	0.050	4.2	0.05	4.1	1.15	0.10	9.0
	B	100	1.20	0.033	2.8					
	C	100	1.18	0.058	4.9					
#5	A	100	0.98	0.043	4.4	0.04	4.1	0.95	0.05	5.1
	B	100	0.97	0.032	3.3					
	C	100	0.98	0.045	4.6					
#6	A	100	0.91	0.055	6.0	0.06	6.2	0.88	0.06	6.3
	B	100	0.90	0.047	5.2					
	C	100	0.91	0.066	7.2					

The within-laboratory standard deviation and standard uncertainty are calculated as follows:

number of group measurements  $N = 6$ ;

number of measurements per group  $n = 100$ ;

repeatability standard deviation (pooled)  $s_L = 5.7\%$ ;

repeatability limit  $r = 2.8 \cdot s_L = 15.9\%$ ;

standard uncertainty  $u_L = \frac{s_L}{\sqrt{N}} = 2.3\%$ .

The microscope standard uncertainty is estimated from the daily calibration log and the nominal reflectance of Y-Al garnet. Thus, the instrument standard uncertainty is

$$u_{instruments} = \frac{s_{instrument}}{\sqrt{N}} = 0.3\%.$$

The reference material standard uncertainty is

number of samples  $N = 6$ ;

standard uncertainty  $u_{RF} = \frac{S_{RF}}{\sqrt{N}} = 2.7\%$ .

The standard uncertainty of the laboratory bias is the square root of the sum of the squares of  $pRRMS$  and  $u_{RF}$ ,

$pRRMS = 2.5\%$ ;

standard uncertainty of reference material (global average)  $u_{RF} = 2.7\%$ ;

standard uncertainty of bias  $u_{bias} = 2.9\%$ .

Finally, the expanded laboratory uncertainty  $U$  at the 95% confidence level is

bias-uncorrected standard uncertainty  $u_{Lab, bias-uncorrected} = 2.3\%$ ;

bias-corrected standard uncertainty  $u_{Lab, bias-corrected} = 3.4\%$ ;

degrees of freedom  $\nu = N - 1 = 5$ ;

coverage factor  $k = 2.571$ ;

expanded uncertainty  $U_{Lab} = k \cdot u_{Lab, bias-corrected} = 9.5\%$ .

### 3.2. Soil Specific Gravity

In this example, the laboratory (GeoTerra) participates in proficiency tests (AASHTO Materials Reference Laboratory) that involves the determination of soil specific gravity according to ASTM D 854 and the results are from a period of six years. The details are given in Table 2. Each test comprises two soil samples that are analyzed by different members of the laboratory staff and the results are sent to the PT provider for evaluation. Several hundred laboratories around the globe take part in the test and the large number of the participating laboratories ( $N > 800$ ) ensures that the global average for each A–B pair can be safely considered as the reference value.

**Table 2 – Laboratory and proficiency test data for soil specific gravity.**

Sample	Laboratory			Global			
	$G_s$	$s_r$	$\%s_r$	$G_s$	$S_R$	$\%S_R$	N
A	2.694	0.004	0.1	2.683	0.0318	1.2	803
B	2.696			2.680	0.0336	1.3	
A	2.703	0.004	0.1	2.658	0.0380	1.4	809
B	2.696			2.657	0.0367	1.4	
A	2.685	0.002	0.1	2.704	0.0349	1.3	849
B	2.682			2.699	0.0336	1.2	
A	2.726	0.0002	0.01	2.733	0.0417	1.5	829
B	2.687			2.694	0.0341	1.3	
A	2.664	0.016	0.6	2.681	0.0350	1.3	862
B	2.627			2.667	0.0333	1.2	
A	2.636	0.004	0.2	2.654	0.0326	1.2	881
B	2.629			2.653	0.0338	1.3	

The within-laboratory standard deviation and standard uncertainty are calculated as follows:

$$\begin{aligned} &\text{number of group measurements } N = 6; \\ &\text{number of measurements per group } n = 2; \\ &\text{repeatability standard deviation (pooled) } S_L = 0.3\%; \\ &\text{repeatability limit } r = 2.8 \cdot S_L = 0.8\%; \\ &\text{standard uncertainty } u_L = \frac{S_L}{\sqrt{N}} = 0.1\%. \end{aligned}$$

The standard uncertainty of the instruments, in this case balance(s) and pycnometer(s) that conform to the requirements of the standard test method, is the combined uncertainties of the balance(s) and the pycnometer(s). For reasons that have to do with the daily use of the laboratory equipment, we prefer to estimate the uncertainty of instruments from tolerances than calibration records even if this results in higher standard uncertainty. The balance(s) for soil specific gravity according to ASTM D 854 should have a readability of 0.01 g and a basic tolerance  $\alpha$  equal to 0.1% (Class GP5, ASTM D 4753). The Class A 250 ml and 500 ml pycnometers have a specified tolerance of 0.15 ml and 0.25 ml, respectively, and the corresponding relative tolerance values are 0.06% and 0.05%. Consequently, the instrument standard uncertainty is calculated, assuming rectangular distribution,

$$\begin{aligned} &\text{standard uncertainty of balance } u_{balance} = \frac{a_{balance}}{\sqrt{3}} = 0.06\%; \\ &\text{standard uncertainty of pycnometer } u_{pycnometer} = \frac{\sqrt{a_{250ml}^2 + a_{500ml}^2}}{\sqrt{3}} = 0.05\%. \end{aligned}$$

The test method reproducibility standard deviation and uncertainty are similarly calculated

$$\begin{aligned} &\text{the number of proficiency tests } N = 6; \\ &\text{the number of measurements per proficiency test (average) } n = 839; \\ &\text{the reproducibility standard deviation (pooled) } S_R = 1.8\%; \\ &\text{the reproducibility limit } R = 2.8 \cdot S_R = 5.2\%; \\ &\text{the standard uncertainty } u_R = \frac{S_R}{\sqrt{N}} = 0.8\%. \end{aligned}$$

The laboratory bias standard uncertainty is then estimated

$$\begin{aligned} &\text{root mean square } RMS = 0.05\%; \\ &\text{standard uncertainty of reference material (global average) } u_{RF} = u_R = 0.06\%; \\ &\text{standard uncertainty of bias } u_{bias} = 0.08\%. \end{aligned}$$

Finally, the expanded laboratory uncertainty  $U$  at the 95% confidence level is

$$\begin{aligned} &\text{bias-uncorrected standard uncertainty } u_{Lab, bias-uncorrected} = 0.1\%; \\ &\text{bias-corrected standard uncertainty } u_{Lab, bias-corrected} = 0.2\%; \\ &\text{degrees of freedom } \nu = N - 1 = 5; \\ &\text{coverage factor } k = 2.571; \\ &\text{expanded uncertainty } U_{Lab} = k \cdot u_{Lab, bias-corrected} = 0.4\%. \end{aligned}$$

#### **4. Conclusion**

We have shown the use of a simple but sound procedure to calculate the standard and expanded uncertainty for typical tests of geological materials considering repeatability measurements, reference materials, and proficiency tests. In the future, we plan to expand the above described procedure to mineral chemistry and thermochemistry data.

#### **5. Acknowledgments**

We would like to thank all laboratory staff members who performed the tests.

#### **6. References**

- ASTM International.2010. Standard Test Methods for Specific Gravity of Soil Solids by Water Pycnometer, ASTM D 854, West Conshohocken, PA, 2007, DOI: 10.1520/D0854-10, 7 pp., [www.astm.org](http://www.astm.org).
- ASTM International.2007. Standard Guide for Evaluating , Selecting, and Specifying Balances and Standard Masses for use in Soil, Rock, and Construction Materials Testing, ASTM D 4753, West Conshohocken, PA, 2007, DOI: 10.1520/D4753-07, 5pp., [www.astm.org](http://www.astm.org).
- BIPM. 2008. Evaluation of measurement data—Guide to the expression of uncertainty in measurement, JCGM 100, 82 pp.
- Mandel J. 1964. The statistical analysis of experimental data, Dover Publications, Inc., New York, 410pp.
- Magnusson B., Näykki T., Hovind H., and Krysell M. 2011. Handbook for Calculation of Measurement Uncertainty in Environmental Laboratories, NT TECHN REPORT 537, 54pp. [www.nordtest.info](http://www.nordtest.info).
- ISO 7404-5 1994. Methods for the Petrographic Analysis of Bituminous Coal and Anthracite—Part 5: Method of Determining Microscopically the Reflectance of Vitrinite, *International Organization for Standardization*, Geneva, Switzerland, 12pp.