The trials and tribulations of certifying geological reference materials using the GeoPT proficiency testing programme.

Philip J. Potts

Emeritus Professor of Geoanalysis,
Faculty of Science, Technology, Engineering and Mathematics, The Open University, Walton Hall, Milton Keynes, MK7 6AA, UK.

phil.j.potts@gmail.com

Outline

- Why are reference materials important in geochemical research.
- Earlier approaches to the characterization of geological reference materials.
- GeoPT proficiency testing programme.
- Opportunities for certification of RMs provided by ISO Guide 35.
- The way in which GeoPT complies with these certification requirements.

Other considerations

ISO/REMCO Guide 35 is currently under revision (ISO TC334).

Issues in the interpretation of documentation.

A potted history of geochemical research

Period	Geochemical research	Technical innovation	Use to which data was put
1920s-	Major elements	Classical then rapid methods	Classification of rock types.
1960s-			
1970s-			
1980s-			
1990s-			
2000s-			

A potted history of geochemical research

Period	Geochemical research	Technical innovation	Use to which data was put
1920s-	Major elements	Classical then rapid methods	Classification of rock types.
1960s-	Trace element geochemistry	XRF	Understanding geological processes.
1970s-			
1980s-			
1990s-			
2000s-			

A potted history of geochemical research

Period	Geochemical research	Technical innovation	Use to which data was put
1920s-	Major elements	Classical then rapid methods	Classification of rock types.
1960s-	Trace element geochemistry	XRF	Understanding geological processes.
1970s-	REE geochemistry	INAA (Ge(Li) detectors)	Geological processes linked to plate tectonics.
1980s-			
1990s-			
2000s-			

A potted history of geochemical research

Period	Geochemical research	Technical innovation	Use to which data was put
1920s-	Major elements	Classical then rapid methods	Classification of rock types.
1960s-	Trace element geochemistry	XRF	Understanding geological processes.
1970s-	REE geochemistry	INAA (Ge(Li) detectors)	Geological processes linked to plate tectonics.
1980s-	Radiogenic isotope geochemistry	TIMS	Age of geological processes.
1990s-			
2000s-			

A potted history of geochemical research

Period	Geochemical research	Technical innovation	Use to which data was put
1920s-	Major elements	Classical then rapid methods	Classification of rock types.
1960s-	Trace element geochemistry	XRF	Understanding geological processes.
1970s-	REE geochemistry	INAA (Ge(Li) detectors)	Geological processes linked to plate tectonics.
1980s-	Radiogenic isotope geochemistry	TIMS	Age of geological processes.
1990s-	Stable isotope geochemistry	High resolution ICP-MS	Climate and environmental change over geological time scales.
2000s-			

A potted history of geochemical research

Period	Geochemical research	Technical innovation	Use to which data was put
1920s-	Major elements	Classical then rapid methods	Classification of rock types.
1960s-	Trace element geochemistry	XRF	Understanding geological processes.
1970s-	REE geochemistry	INAA (Ge(Li) detectors)	Geological processes linked to plate tectonics.
1980s-	Radiogenic isotope geochemistry	TIMS	Age of geological processes.
1990s-	Stable isotope geochemistry	High resolution ICP-MS	Climate and environmental change over geological time scales.
2000s-	All above in an individual mineral scale	Microbeam techniques (LA-ICP-MS, SIMS)	Assessment of the above phenomena on a mineral scale.

Elements involved in geochemical research

Major elements	**SiO_2, TiO_2, Al_2O_3, Fe_2O_3T, MnO, MgO, CaO, Na_2O, K_2O, P_2O_5**
Trace elements (XRF)	Ni, Cu, Zn, Ga, Rb, Sr, Y, Nb, Zr, Ba, Pb
REE (INAA)	La, Ce, Nd, Sm, Eu, Gd, Tb, Tm, Yb, Lu (+ Ta, Hf, Th)
Radiogenic isotopes	Nd/Sm, Rb/Sr, U/Th/Pb
Stable isotopes	H, C, N, O, S, Fe, Cu, Zn, Mo
Modern lab (routine use of ICP-MS, XRF and ICP-AES)	SiO_2, TiO_2, Al_2O_3, Fe_2O_3T, MnO, MgO, CaO, Na_2O, K_2O, P_2O_5, Ag, Ba, Be, Bi, Cd, Ce, Co, Cr, Cs, Cu, Dy, Er, Eu, Ga, Gd, Ge, Hf, Ho, La, Li, Lu, Mo, Nb, Nd, Ni, Pb, Pr, Rb, Sb, Sc, Sm, Sn, Sr, Ta, Tb, Th, Tl, Tm, U, V, W, Y, Yb, Zn and Zr.

Development of geological reference materials

G1 (Granite) and W1 (Diabase).

One of the most important early investigations (1940s-1950s) in the characterization of geological reference materials was undertaken by the US Geological Survey/MIT/Carnegie Institute

Reason: To develop a well characterized RMs for the calibration of dc arc OES.

Programme: Samples distributed and 34 laboratories from 10 countries who reported results (mainly by classical methods).

Results of interlaboratory study (Fairburn *et al.*, USGS Bulletin 980, 1951):

- Significant interlaboratory bias.
- "The normative variation was so extreme that in the absence of direct evidence to the contrary, there would be little reason to suppose that all the analyses submitted were of the same two rocks" (Chayes, 1951).

See also AMC Technical Brief No.26 (2006) - How good were analysts in the good old days before instrumentation?

What followed the G1/W1 study?

- The need for geological reference materials recognized.
- Indeed, because of the importance of ***matrix matching (especially the mineralogy)***, ideally a RM for every known rock type.
- The benefits of participation in ***interlaboratory trials*** recognized (no PT schemes yet!).
- There then followed a ***free for all*** (unregulated) period in which many geological reference materials were issued.
- ***Early contributors:*** US Geological Survey *(Father Flanagan),* Geological Survey of Japan *(Terashima and others),* Centre de Recherches Pétrographiques et Géochimiques, France *(Raj Govindaraju)*, National Institute for Metallurgy, South Africa *(T.W. Steele),* Geological Survey of Canada *(Sidney Abbey)*

During this period (1970s-1990s), how were geological reference materials characterized?

(i) Using a limited amount of data from the ***survey's own laboratories***.

(ii) Data from ***selected labs*** invited (or paid) to participate.

(iii) Data from any lab who wished to contribute (***no assessment of competence***).

(iv) In at least one case, new (potential) RMs were issued with guide values, but no properly assessed reference values at all.

(v) At the same time, there developed a practice of ***compiling data published in the research literature*** to derive reference values. These compilations were often undertaken by independent researchers (not the RM producer).

There are difficulties in demonstrating current ISO-REMCO compliance with some of these approaches!

But, ISO REMCO was not established until 1975 and the first (?) edition of ISO-REMCO Guide 35 (Certification of reference materials – General and statistical principles) was not published until 1989.

ISO-REMCO – Committee on reference materials – TC334

- **Established in 1975.**
- **To establish concepts, terms and definitions related to reference materials**.
- To specify the basic characteristics of reference materials as required by their intended use.
- To propose actions on reference materials required to support other ISO activities.
- To **prepare guidelines** for ISO technical committees when dealing with reference material issues.
- To communicate with other international organizations on reference material matters.
- To advise the ISO Technical Management Board (TMB) on reference material issues.

- **ISO REMCO was an advisory committee that published Guides but is now a formal ISO Technical Committee with the authority to develop new standards.**

Why all this interest in Certified Reference Materials?

Definitions from VIM3 (under review):

5.13 (6.13)
reference material
RM
material, sufficiently homogeneous and stable with reference to specified properties, which has been established to be fit for its intended use in **measurement** or in examination of **nominal properties**

5.14 (6.14)
certified reference material
CRM
reference material, accompanied by documentation issued by an **authoritative body** and providing one or more specified property values with associated ***uncertainties*** and ***traceabilities***, using valid procedures

Therefore, a CRM represents a higher metrological standard because property values must be accompanied by **uncertainties** and a statement of **traceability.**

2.26 (3.9)
measurement uncertainty
non-negative parameter characterizing the dispersion of the **quantity values** being attributed to a **measurand**, based on the information used

2.41 (6.10)
metrological traceability
property of a **measurement result** whereby the result can be related to a reference through a documented unbroken chain of **calibrations**, each contributing to the **measurement uncertainty**

ISO-REMCO Guide 35 – Use of Proficiency Testing for Certification Characterisation Studies

ISO Guide 35:2017 (Reference materials – Guidance for characterization and assessment of homogeneity and stability).

The 2017 Guide accepted (for the first time) that proficiency testing could in some circumstances by used in certification studies.

After listing all the disadvantages of doing so:

"Nonetheless, it can be useful to consider ***combining characterization studies for RMs with other studies to save costs****, providing that due care is taken to avoid the principal disadvantages and that certain conditions are met. A.3.2 provides guidance on the principal disadvantages; clause A.3.3 provides conditions for combination of such studies with RM characterization".*

ISO-REMCO Guide 35 – Use of Proficiency Testing for Certification Characterisation Studies

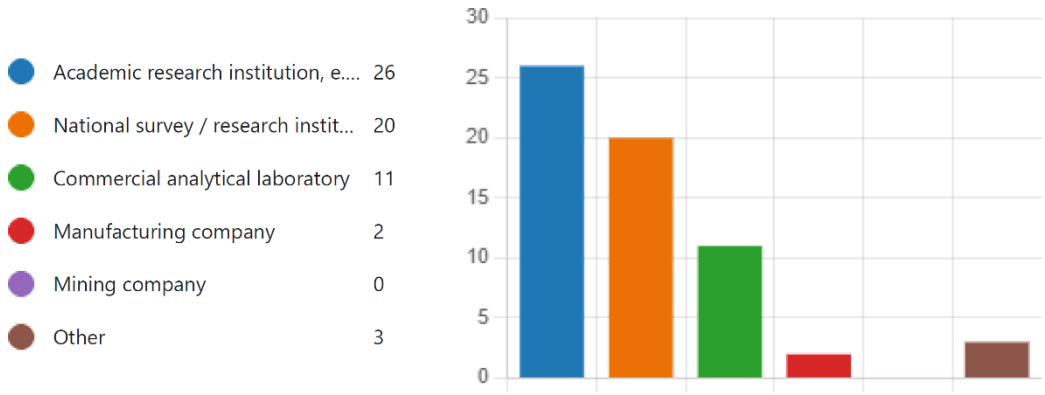
- Are there opportunities to use the GeoPT Proficiency Testing Programme for the certification of geological reference materials?

GeoPT Proficiency Testing Programme

- Established in 1994 and now operated by the International Association of Geoanalysts
- Designed for laboratories that **routinely measure silicate rocks** for the major elements and a wide range of trace elements.
- No requirement for laboratories to demonstrate **competence** before participating.
- Laboratories are sent test materials **twice a year**.
- Typically, over **100 laboratories** world-wide participate in each round.
- Reports from previous rounds are available on the IAG web site. ***(https://www.geoanalyst.org/geopt-previous-rounds/).***

What do we know of GeoPT participants?

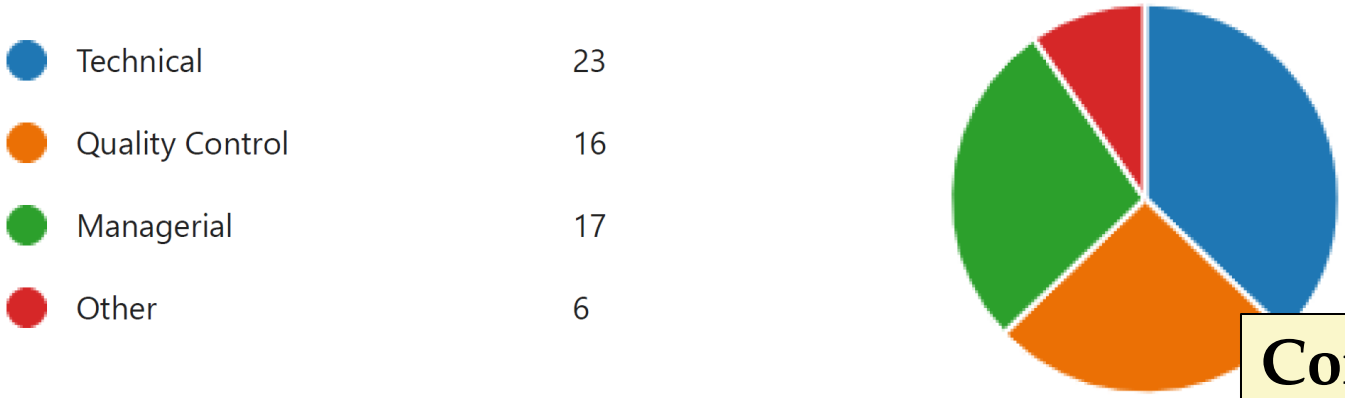
1. Please indicate which of the following categories best describes the function of your organisation. (0 point)



Type of establishment.

GeoPT contact.

2. Please indicate your role in the organisation that you represent. (0 point)

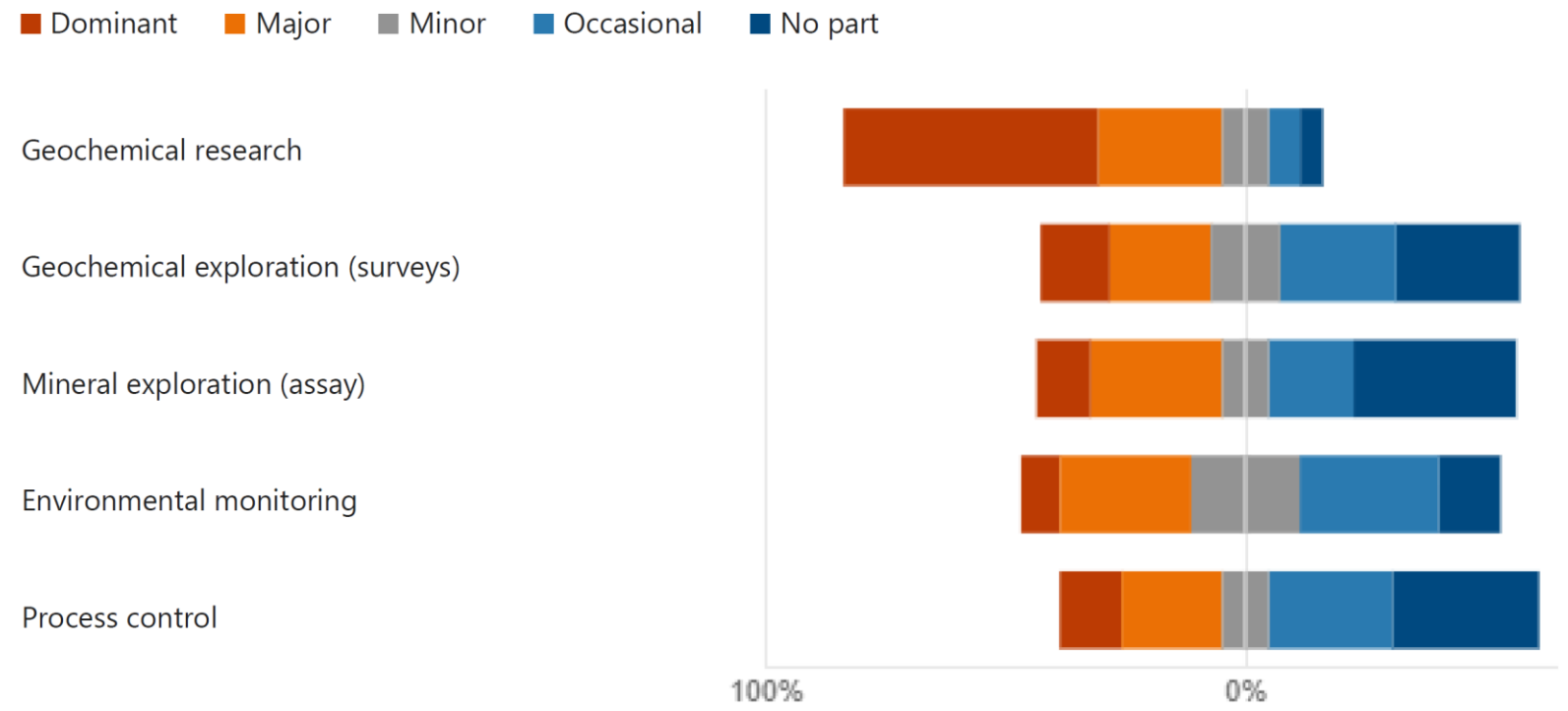


Contacts role on organisation

Details of GeoPT participants

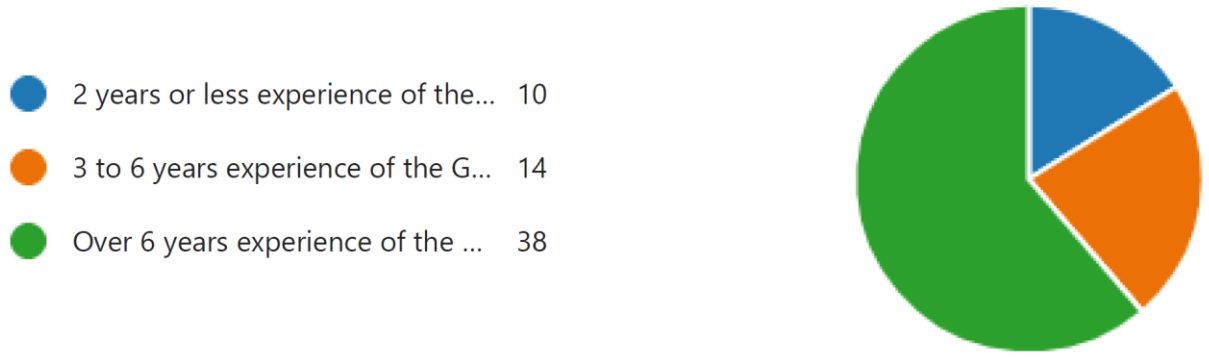
Main analytical activity of organization.

3. Please indicate the extent to which the following activities form part of your analytical workload. (0 point)



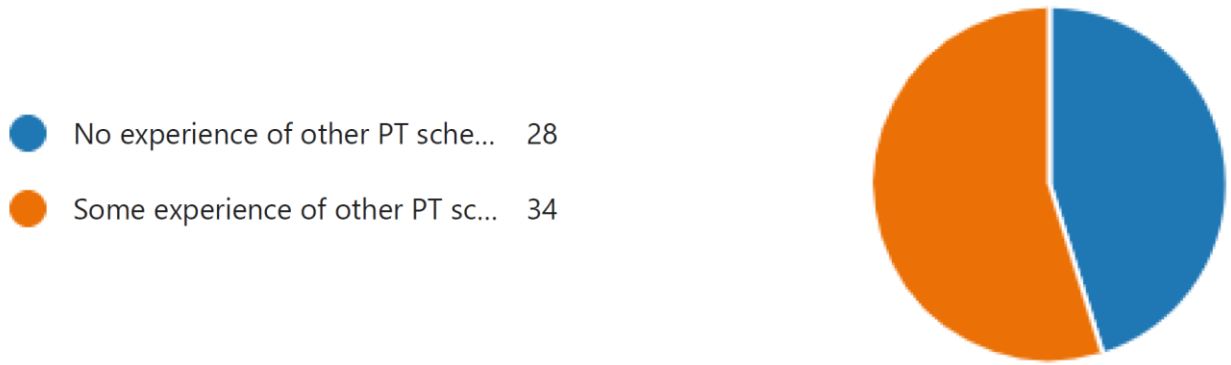
Details of GeoPT participants

4. Please indicate the length of your experience of the Geo*PT* scheme . (0 point)



Number of years participating in GeoPT.

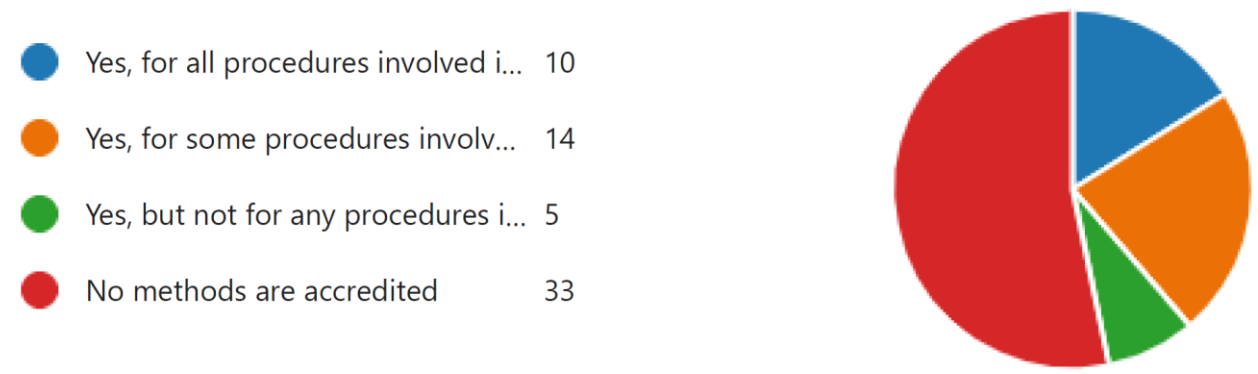
5. Please indicate any experience you may have of other PT schemes. (0 point)



Experience of other GeoPT schemes.

Details of GeoPT participants

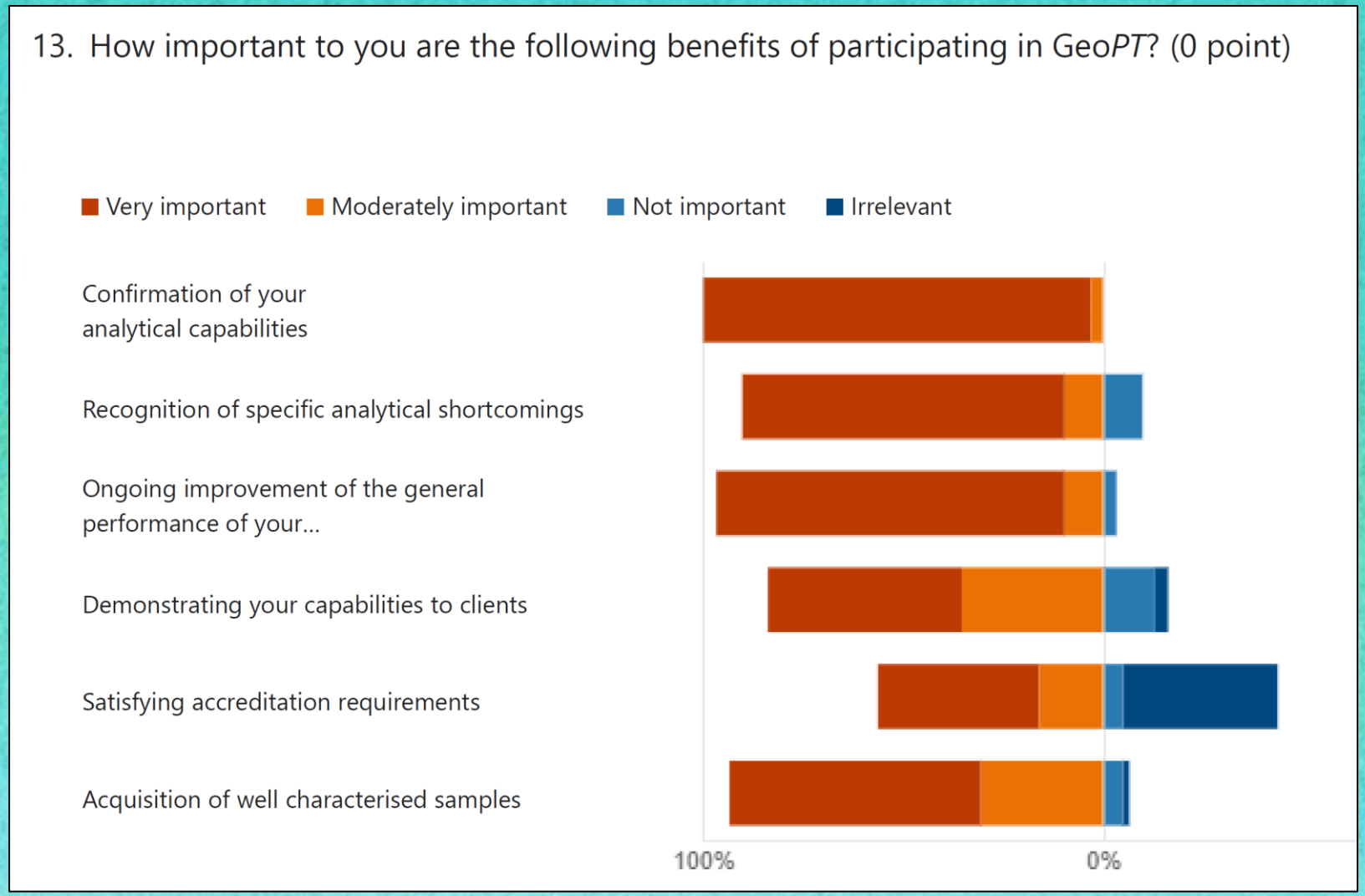
7. Please tell us if your analytical laboratory is accredited? (0 point)



Is your lab accredited?

Accreditation key words.

Details of GeoPT participants



Benefits of participating in GeoPT.

What sort of material is distributed to GeoPT participants?

Finely ground silicate rock powder (at least 125 x 40 g).

Options:

(1) An established RM or CRM *(Ideal but rarely possible – commercial cost could be more than £15,000 per round).*

(2) -

(3) -

What sort of material is distributed to GeoPT participants?

Finely ground silicate rock powder (at least 125 x 40 g).

Options:

(1) An established RM or CRM *(Ideal but rarely possible – commercial cost could be more than £15,000 per round).*

(2) A material prepared from high purity chemical reagents *(Although such a material might match the chemical composition of a silicate rock, it cannot match the mineralogy).*

(3) -

What sort of material is distributed to GeoPT participants?

Finely ground silicate rock powder (at least 125 x 40 g).

Options:

(1) An established RM or CRM *(Ideal but rarely possible – commercial cost could be more than £15,000 per round).*

(2) A material prepared from high purity chemical reagents *(Although such a material might match the chemical composition of a silicate rock, it cannot match the mineralogy).*

(3) A representative silicate rock, the composition of which is assessed from the results submitted by participating laboratories (Only practical option).

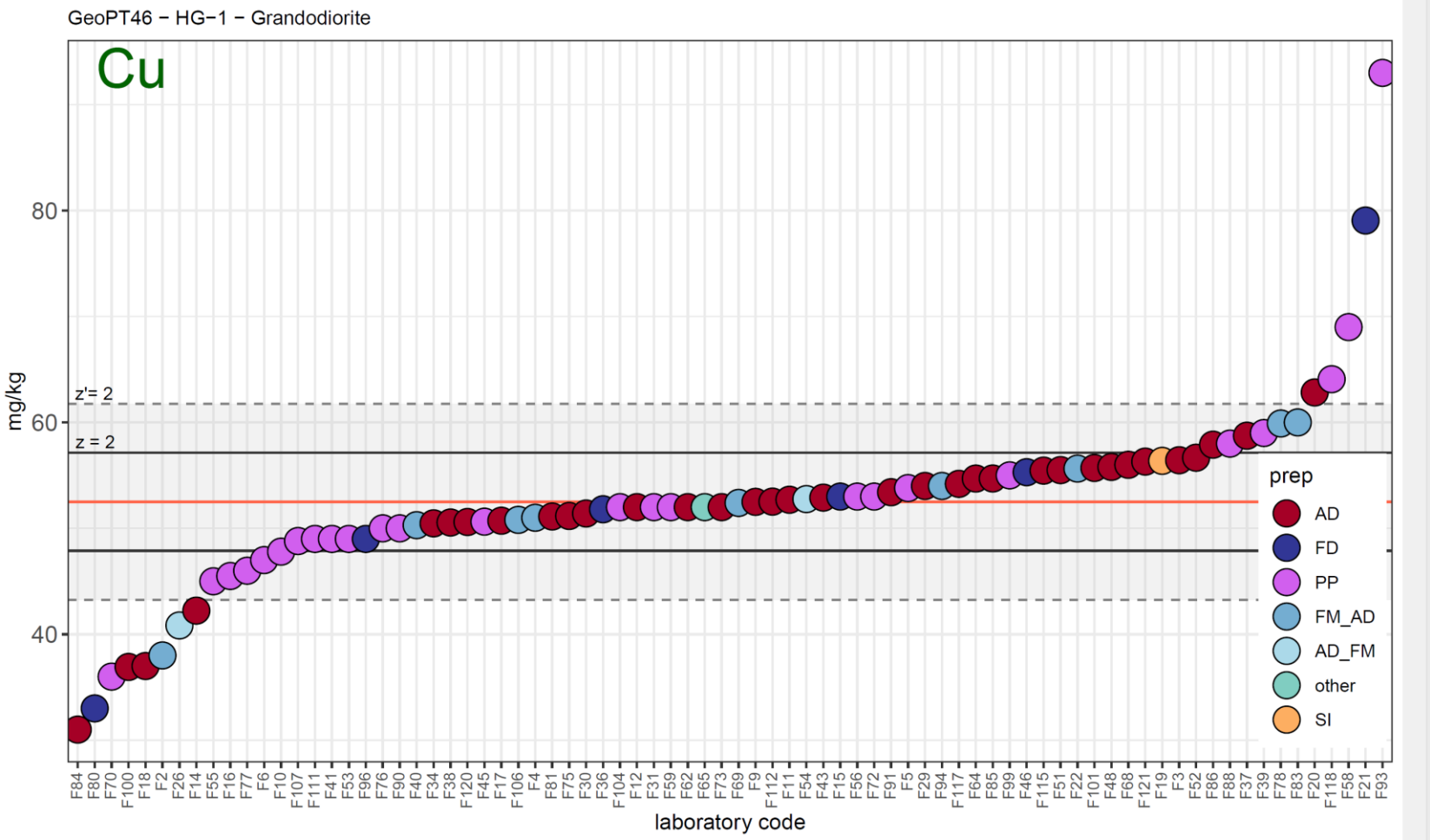
Assessing the composition from the results submitted by participating laboratories.

- As a consequence, data distributions must be assessed to derive **consensus values** that are judged to be the best estimates of the true composition.
- In GeoPT, practical experience has led to procedures based on **robust statistical procedures** with the use of **expert judgement** to assess the quality of data distributions.
- Precision targets (the standard deviation for proficiency testing) is set by a modified form of the Horwitz function.
- *(see AMC Technical Brief No.17 (2004) – The amazing Horwitz function).*
- Experience has also led to circumstances where **analytical bias** can be detected in contributed data distributions.

Assessing GeoPT data distributions

Example of ‘good’ GeoPT data distributions

Data plotted in increasing order of magnitude as the difference between individual values and the consensus value.



- **Majority of data is within the inner tram lines (z=+/-2).**
- **Mix of procedures contribute to this section.**
- **Clearly defined point of inflection corresponding to the consensus value.**
- **No distracting tails.**
- **Estimators converge.**

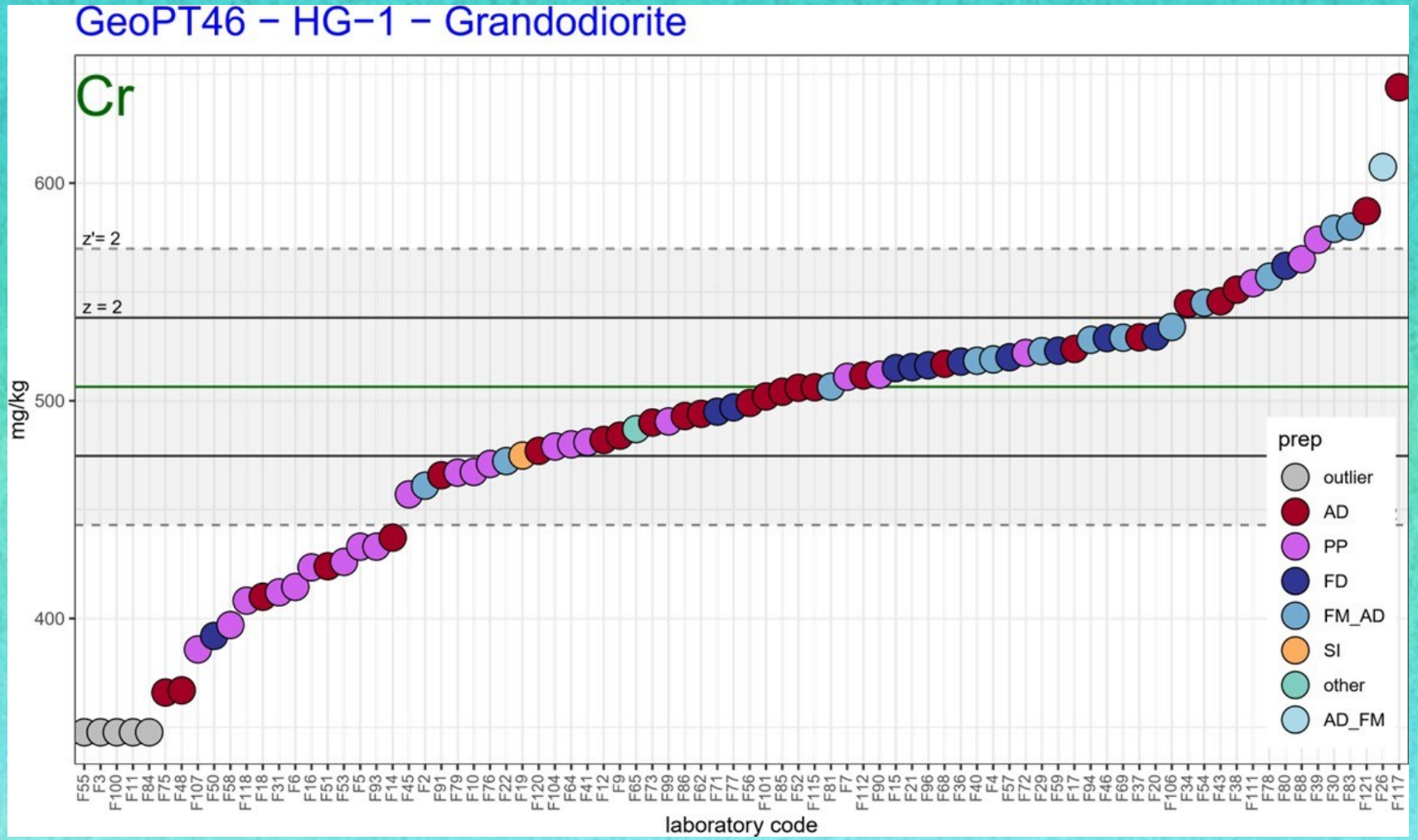
AD – acid digestion
FD – fusion disk (XRF)
PP – powder pellet (XRF)
FM/AD – fusion followed by acid digestion.
AD/FM – Acid digestion followed by fusion of residue.
Si - sintering

Sample preparation

estimators

analyte	n	median	robust	Rmode mean	Mike's mode	Mike's mode boot
Cu	79	52.4	52.509	52.725	52.511	52.201

Examples of distributions judged to be not sufficiently good to allocate an ‘assigned’ value.

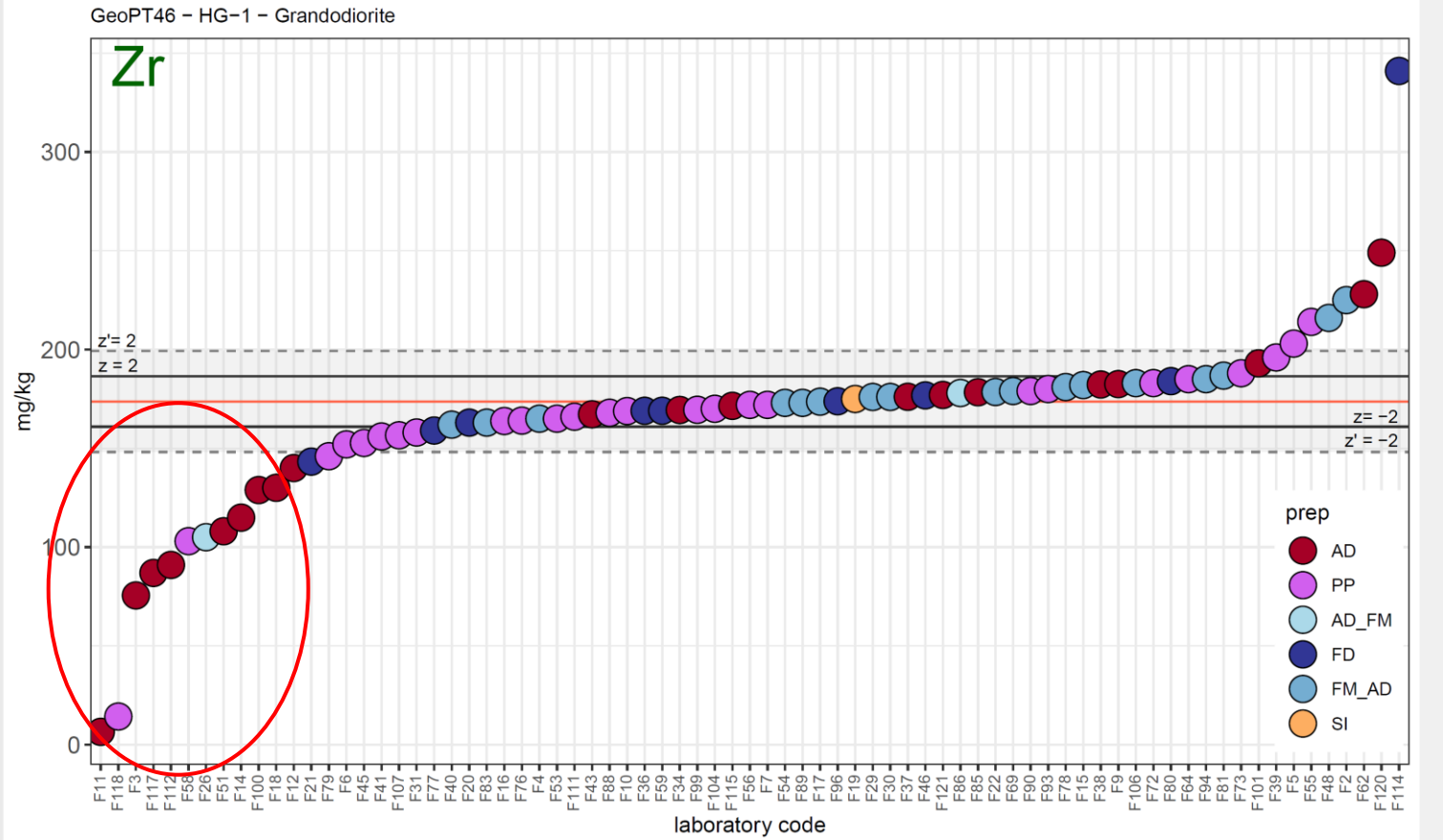


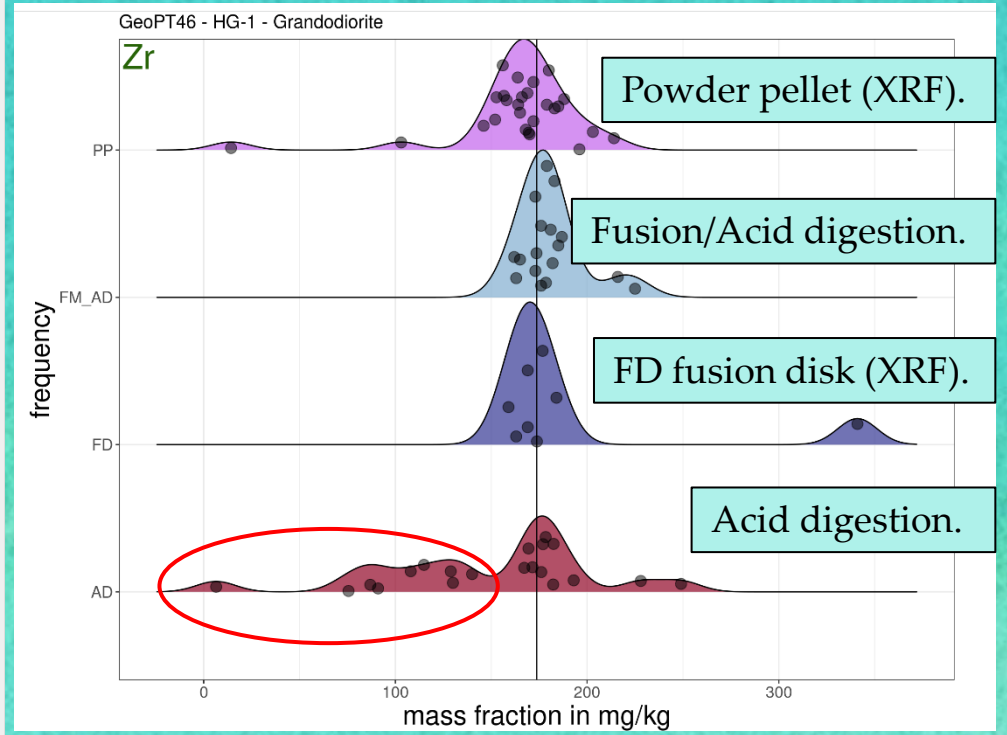
- **Data distribution fails the Horwitz precision target test.**
- **Poorly defined point of inflection.**
- **Consensus value only suitable as an information value.**

AD – acid digestion.
PP – powder pellet (XRF).
FD – fusion disk (XRF).
FM/AD – fusion followed by acid digestion.
Si – sintering.
AD/FM – Acid digestion followed by fusion of residue.

Sample preparation

Examples of distributions judged to demonstrate analytical bias





estimators

analyte	n	median	robust	Rmode mean	Mike's mode	Mike's mode boot
Zr	75	172	169.23	172.75	173.7	174.28

Rigorous criteria for allocating ‘assigned value’ status to a GeoPT consensus value.

- Sufficient laboratories **(15 or more)** had contributed data to the region of the distribution from which the consensus value was evaluated.
- **Visual assessment** gave confidence that a substantial proportion of the results distribution was symmetrically disposed about the consensus.
- The **ratio of the uncertainty** in the location estimate to the target precision was an acceptably small value.
- An evaluation of measurement results by procedure – including both methods of measurement and sample preparation – indicated no significant **procedural bias** among the results from which the consensus was derived.

Some issues on how GeoPT complies with ISO/REMCO Guide 35 as a certification scheme'

"Decide, before the start of a study, ***which subset of data from specified laboratories will be used for value assignment.***"

Note that there is no requirement for laboratories to demonstrate their competence to participate in the GeoPT proficiency testing scheme.

Evaluation of competence was undertaken on GeoPT round 39

Test Material: Syenite SyMP-1

Co-analysed CRM: nepheline syenite CGL 006 (LNS).

102 laboratories contributed to SyMP-1

Decide, before the start of a study, ***which subset of data from specified laboratories will be used for value assignment.***

	Laboratory selection criteria	Rationale for selection criteria
1	Use data from labs that participated in at **least three of the last four GeoPT proficiency testing rounds**.	Regular participation means lab has effective quality control procedures and has had an opportunity to respond to z-score feedback.
2		
3		
4		
5		

Decide, before the start of a study, ***which subset of data from specified laboratories will be used for value assignment.***

	Laboratory selection criteria	Rationale for selection criteria
1	Use data from labs that participated in at least three of the last four GeoPT proficiency testing rounds.	Regular participation means lab has effective quality control procedures and has had an opportunity to respond to z-score feedback.
2	Use data from labs that reported **at least 70% 'satisfactory'** according to the GeoPT z-score criterion whereby $-2 < z < 2$.	Competent laboratories should consistently contribute satisfactory data.
3		
4		
5		

Decide, before the start of a study, ***which subset of data from specified laboratories will be used for value assignment.***

	Laboratory selection criteria	Rationale for selection criteria
1	Use data from labs that participated in at least three of the last four GeoPT proficiency testing rounds.	Regular participation means lab has effective quality control procedures and has had an opportunity to respond to z-score feedback.
2	Use data from labs that reported at least 70% 'satisfactory' according to the GeoPT z-score criterion whereby $-2 < z < 2$.	Competent laboratories should consistently contribute satisfactory data.
3	Use data from labs that reported **at least 70% 'satisfactory'** measurement results (according to the GeoPT z-score criterion whereby $-2 < z < 2$) **on the co-analysed CRM**.	Competent laboratories should consistently contribute satisfactory data. In this case the evaluation is based on the CRM results.
4		
5		

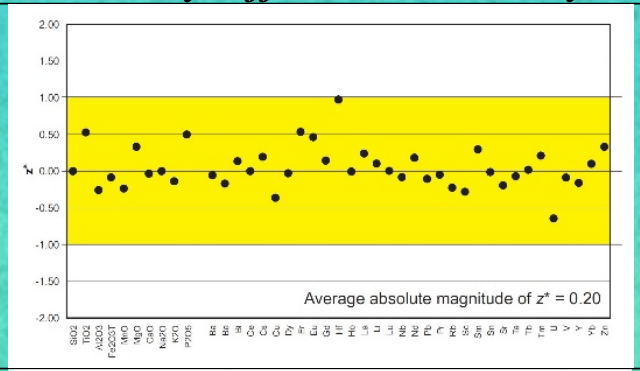
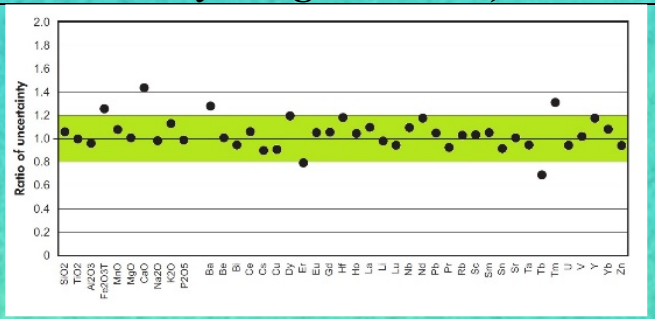
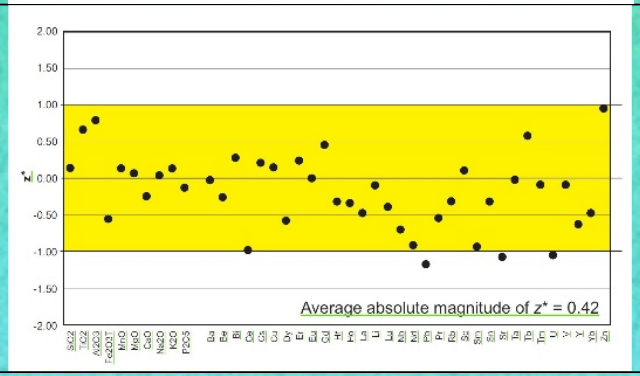
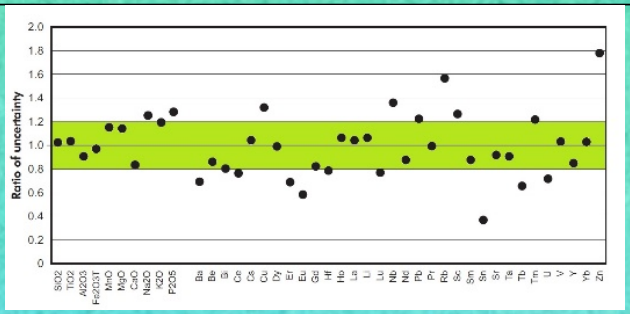
Decide, before the start of a study, ***which subset of data from specified laboratories will be used for value assignment.***

	Laboratory selection criteria	Rationale for selection criteria
1	Use data from labs that participated in at least three of the last four GeoPT proficiency testing rounds.	Regular participation means lab has effective quality control procedures and has had an opportunity to respond to z-score feedback.
2	Use data from labs that reported at least 70% 'satisfactory' according to the GeoPT z-score criterion whereby $-2 < z < 2$.	Competent laboratories should consistently contribute satisfactory data.
3	Use data from labs that reported at least 70% 'satisfactory' measurement results (according to the GeoPT z-score criterion whereby $-2 < z < 2$) on the co-analysed CRM.	Competent laboratories should consistently contribute satisfactory data. In this case the evaluation is based on the CRM results.
4	Use data from laboratories that **reported satisfactory data** (according to the GeoPT z-score criterion whereby $-2 < z < 2$) **on the co-analysed CRM**, restricting measurands to only those that gave such satisfactory data.	If measurement results on the CRM are satisfactory, it is likely that those on the test sample will be as well.
5		

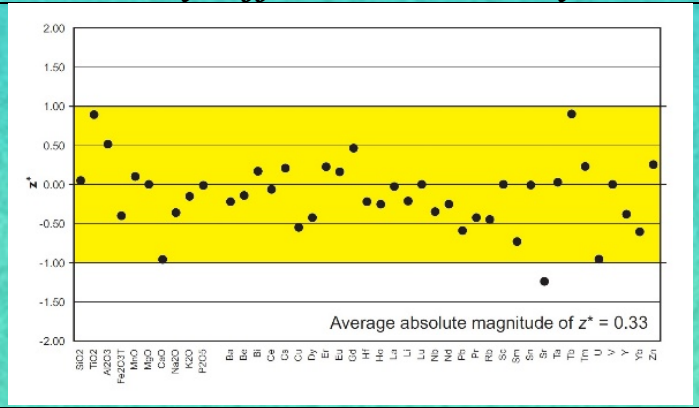
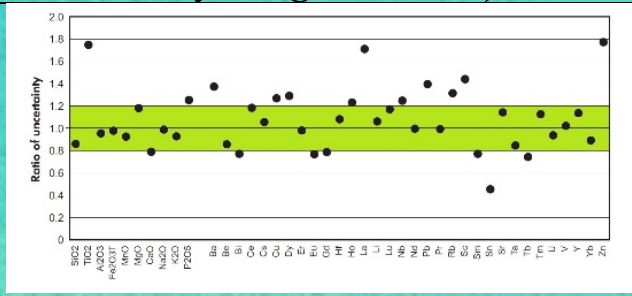
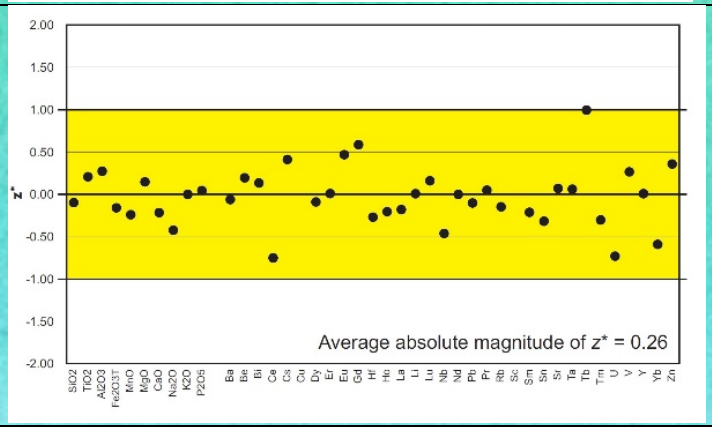
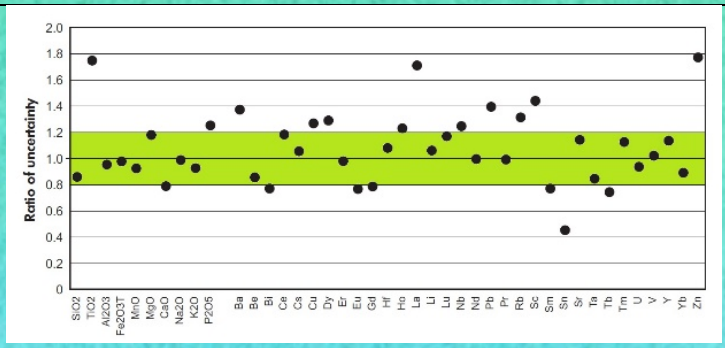
Decide, before the start of a study, ***which subset of data from specified laboratories will be used for value assignment.***

	Laboratory selection criteria	Rationale for selection criteria
1	Use data from labs that participated in at least three of the last four GeoPT proficiency testing rounds.	Regular participation means lab has effective quality control procedures and has had an opportunity to respond to z-score feedback.
2	Use data from labs that reported at least 70% 'satisfactory' according to the GeoPT z-score criterion whereby $-2 < z < 2$.	Competent laboratories should consistently contribute satisfactory data.
3	Use data from labs that reported at least 70% 'satisfactory' measurement results (according to the GeoPT z-score criterion whereby $-2 < z < 2$) on the co-analysed CRM.	Competent laboratories should consistently contribute satisfactory data. In this case the evaluation is based on the CRM results.
4	Use data from laboratories that reported satisfactory data (according to the GeoPT z-score criterion whereby $-2 < z < 2$) on the co-analysed CRM, restricting measurands to only those that gave such satisfactory data.	If measurement results on the CRM are satisfactory, it is likely that those on the test sample will be as well.
5	No selection criteria – **use the full GeoPT data set** for the assessment of consensus values.	The rationale for this approach is that when measurement results are assessed from a large number of laboratories operating independently of one another, errors will cancel out.

Decide, before the start of a study, ***which subset of data from specified laboratories will be used for value assignment (Criteria 1 and 2).***

	Laboratory selection criteria	**Agreement between consensus values when compared to unfiltered data *(not statistically different within the yellow band)***	**Agreement between uncertainties when ratioed to unfiltered data *(+/-20% indicated by the green band)***
1	Use data from labs that participated in at least three of the last four GeoPT proficiency testing rounds. 83 LABS complied		
2	Use data from labs that reported at least 70% ‘satisfactory’ according to the GeoPT z-score criterion whereby -2 < z < 2. 38 LABS complied		

Decide, before the start of a study, ***which subset of data from specified laboratories will be used for value assignment (Criteria 3 and 4).***

	Laboratory selection criteria	**Agreement between consensus values when compared to unfiltered data *(not statistically different within the yellow band)***	**Agreement between uncertainties when ratioed to unfiltered data *(+/-20% indicated by the green band)***
3	Use data from labs that reported at least 70% 'satisfactory' measurement results (according to the GeoPT z-score criterion whereby -2 < z < 2) on the co-analysed CRM. 40 LABS complied		
4	Use data from laboratories that reported satisfactory data (according to the GeoPT z-score criterion whereby -2 < z < 2) on the co-analysed CRM, restricting measurands to only those that gave such satisfactory data. Between 13 and 37 LABS complied		

Decide, before the start of a study, ***which subset of data from specified laboratories will be used for value assignment.***

Conclusion:

No advantage was observed in using a sub-set of labs participating in GeoPT round 39, compared with evaluating consensus values and uncertainties from the full data set.

Vol. 43 — N° 2 06 19 p. 217–229

GEOSTANDARDS and GEOANALYTICAL RESEARCH

An Evaluation of Methods for Assessing the Competence of Laboratories Based on Performance in the Geo*PT* Proficiency Testing Scheme

Philip J. **Potts*** and Peter C. **Webb**

Faculty of Science, Technology, Engineering and Mathematics, The Open University, Walton Hall, Milton Keynes, MK7 6AA, UK
* Corresponding author. e-mail: phil.j.potts@gmail.com

The original certification protocol, published by the International Association of Geoanalysts in 2003, specified that the competence of laboratories selected as competent to contribute certification measurements should be evaluated from their performance in the Geo*PT* proficiency testing programme. Round 39 of the IAG Geo*PT* proficiency testing programme provided an opportunity to examine four methods of evaluating laboratory competence based largely on the use of proficiency testing *z*-scores as performance indicators. This opportunity arose because two test materials were co-analysed by participating laboratories in this round: a syenite, SyMP-1, supplied by the USGS, and an established CRM, the nepheline syenite, CGL 006. The performance of laboratories was assessed in four ways; in each case, consensus values and their uncertainties as derived from selective data sets of competent laboratories were compared with results derived from the routine Geo*PT* data assessment, involving all submitted measurements. An overall comparison of results showed no significant statistical differences in either consensus values or uncertainties between these data sets. This conclusion was unexpected and calls into question the widely held assumption that 'better' consensus data would be obtained from a subset of laboratories judged to be competent on the basis of proficiency testing performance indicators.

Keywords: reference material certification, laboratory competence, laboratory selection, Geo*PT*, proficiency testing, USGS SyMP-1, syenite, CGL 006.

Received 21 Jan 18 – Accepted 06 Nov 18

Geostandards and Geoanalytical Research (2019), 43, 217-229.

As an alternative to demonstrate the competence of participating laboratories, is it possible to demonstrate the competence of the proficiency testing scheme itself?

Specifically, by comparing GeoPT consensus values with certified values on the small number of occasions when Certified Reference Materials have been used as GeoPT test materials.

GeoPT versus CRM comparison

GeoPT round	Origin	Name	Rock type	Certified by
GeoPT9	IAG (International Association of Geoanalysts)	OU-6	Penrhyn Slate	IAG Certification Protocol (Competent labs selected by PT and asked to submit certification quality data)
GeoPT12	MGL (Central Geological Laboratory, Mongolia)	GAS	Serpentine	IAG Certification Protocol
GeoPT14	MGL (Central Geological Laboratory, Mongolia)	OShBO	Alkaline Granite	IAG Certification Protocol
GeoPT16	NIST	SRM 688	Basalt	NIST standard procedures (mainly primary methods)
GeoPT25	Instituto di Geociencias (Campinas, Brazil)	BRP-1	Basalt	In collaboration with the US Geological Survey
GeoPT27	MGL (Central Geological Laboratory, Mongolia)	AND	Andesite	Accredited procedures, CGL, Mongolia

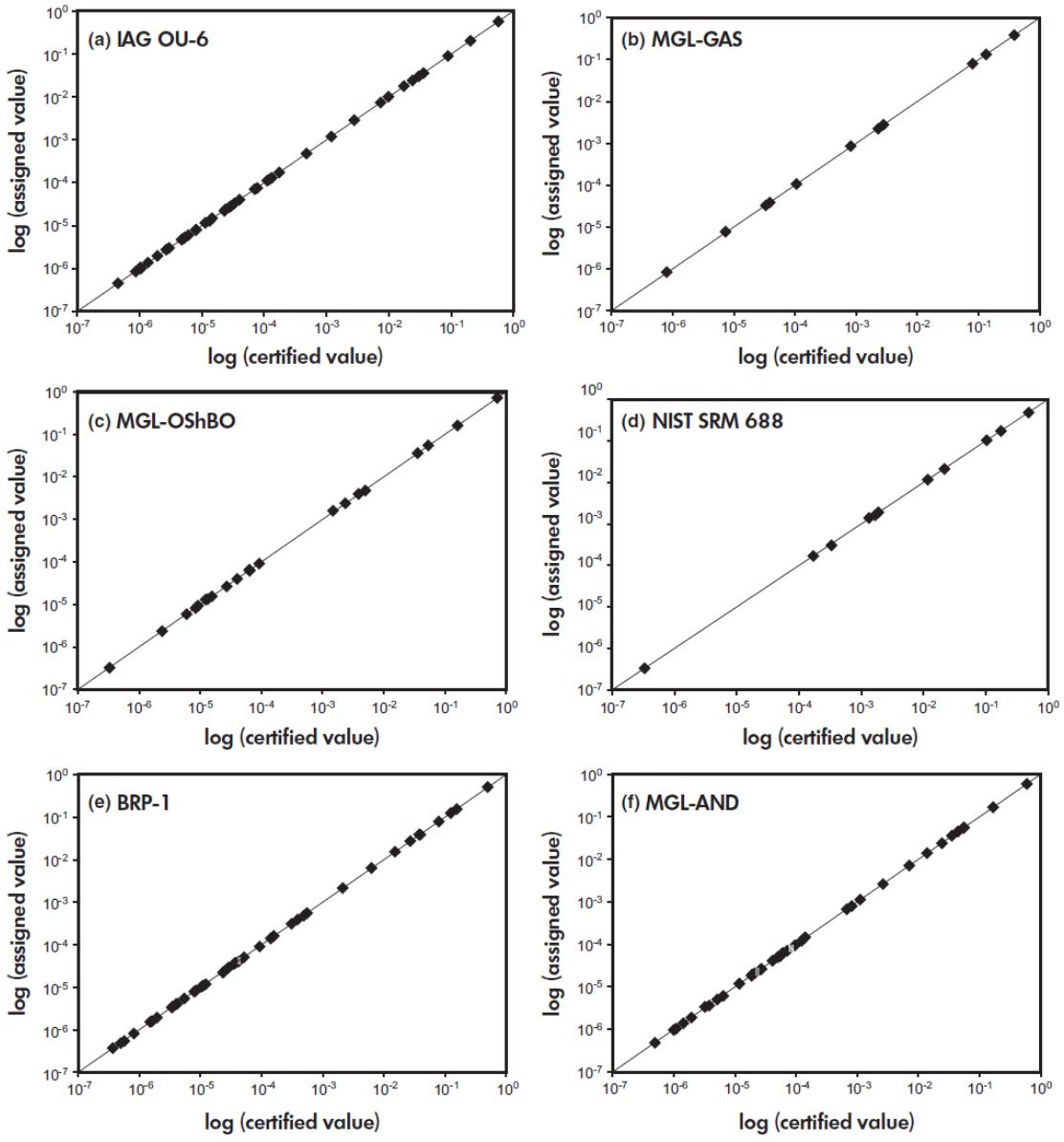


Figure 1. Comparison of certified and assigned values with lines of equality for (a) IAG OU-6, (b) MGL-GAS, (c) MGL-OShBO, (d) NIST SRM 688, (e) BRP-1, (f) MGL-AND.

GeoPT versus CRM comparison

X axis: Certified value.

Y axis: GeoPT consensus value.

Major and trace elements plotted on a **log scale.**

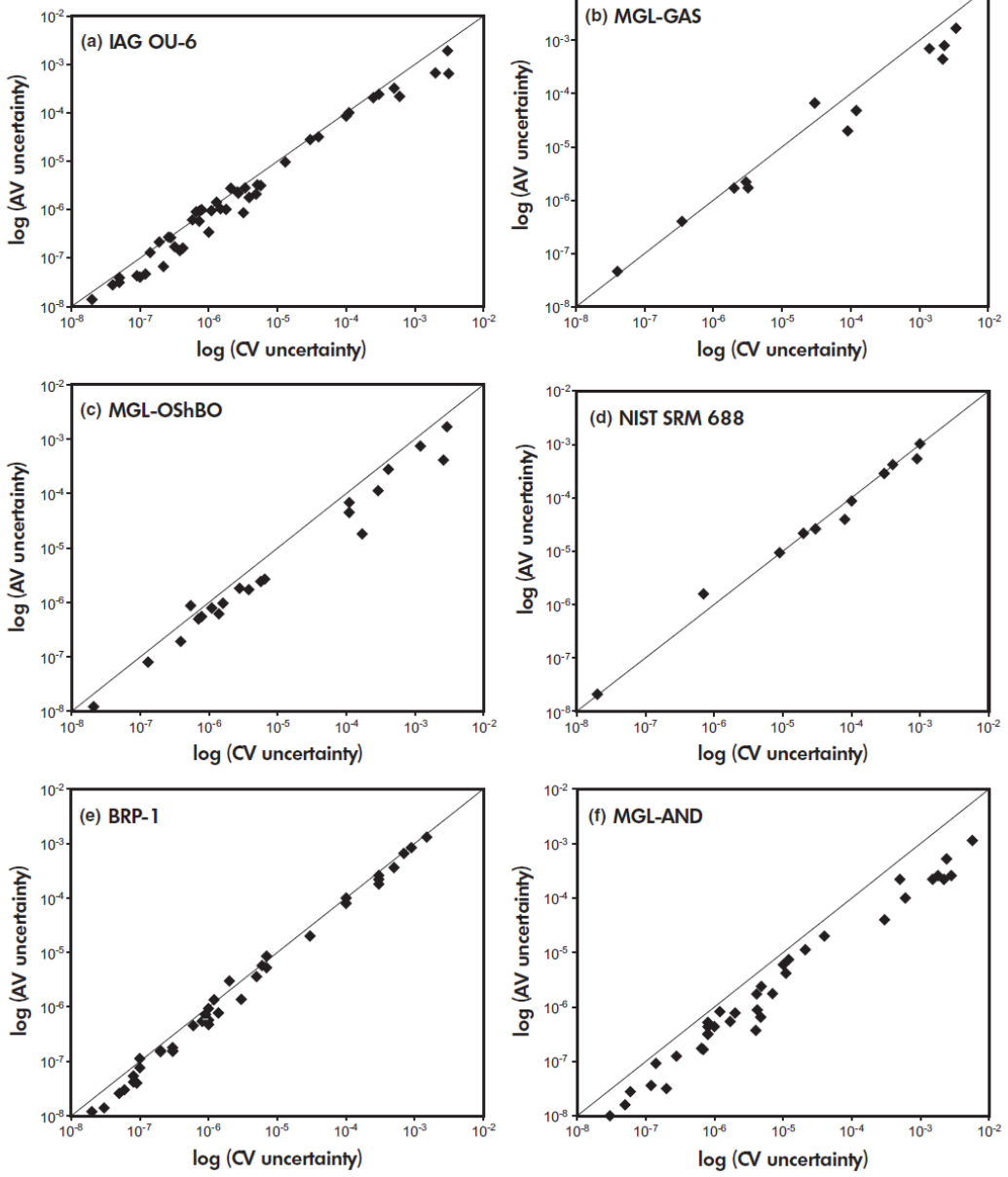


Figure 2. Comparison of uncertainties on assigned values and on certified values showing the lines of equality for (a) IAG OU-6, (b) MGL-GAS, (c) MGL-OShBO, (d) NIST SRM 688, (e) BRP-1, (f) MGL-AND.

GeoPT versus CRM comparison

X axis: Uncertainty in certified value.

Y axis: GeoPT uncertainty in GeoPT consensus value.

Major and trace uncertainty data plotted on a **log scale.**

GeoPT versus CRM comparison - Conclusions

- A statistical analysis of these data showed that the difference between individual assigned and certified values **seldom reached significance**, except for a few values in two of the six CRMs (MGL-OShBO and NIST SRM 688).
- More importantly, when differences were **scaled** according to the respective GeoPT fitness-for-purpose criteria, there was no suggestion of differences of a consequential magnitude.
- GeoPT assigned values are a **reliable estimate** of the true quantity value derived from the rigorous certification process and are fit for purpose for use in the GeoPT scheme.

Vol. 39 – N° 4 12 15 P. 407 – 417

GEOSTANDARDS and GEOANALYTICAL RESEARCH

The Reliability of Assigned Values from the Geo*PT* Proficiency Testing Programme from an Evaluation of Data for Six Test Materials that have been Characterised as Certified Reference Materials

Philip J. **Potts** (1)*, Michael **Thompson** (2) and Peter C. **Webb** (1)

(1) Faculty of Science, The Open University, Walton Hall, Milton Keynes, MK7 6AA, UK
(2) School of Biological and Chemical Sciences, Birkbeck University of London, Malet Street, London, WC1E 7HX, UK
* Corresponding author. e-mail: philjpotts@gmail.com

Assigned values derived from the Geo*PT* proficiency testing programme were compared with certified values for six certified reference materials that have been used as test materials in the Geo*PT* programme. Statistical analysis showed that there were few significant differences between these sets of data and that these differences had no significant impact on the Geo*PT* assessment when fitness-for-purpose criteria were taken into account.

*Les valeurs attribuées dérivées du programme de test d'aptitude Geo*PT* ont été comparés avec les valeurs certifiées pour six matériaux de référence certifiés qui ont été utilisés comme matériaux de test dans le programme Geo*PT*. L'analyse statistique a montré qu'il y avait peu de différences significatives entre ces ensembles de données, et que ces différences n'ont eu aucune incidence significative sur l'appréciation Geo*PT* quand les critères d'usage ont été pris en compte.*

Keywords: proficiency testing, assigned value, certified value, Geo*PT*, statistical analysis, fitness for purpose
Received 25 Sep 14 – Accepted 19 Mar 15

*Mots-clés : test de compétence, valeur attribuée, valeur certifiée, Geo*PT*, analyse statistique, aptitude à l'emploi.*

The Geo*PT* proficiency testing scheme is organised by the International Association of Geoanalysts (IAG) and has been in operation since 1995. The scheme is designed for the evaluation of the performance of laboratories involved in the routine analysis of silicate rocks and is undertaken in accordance with a published protocol (Thompson 2002), which was written to demonstrate compliance with the IUPAC international harmonised protocol (Thompson *et al.* 2006). The principal aim is to provide each participating laboratory with information to decide whether analytical performance is judged to be satisfactory or may be subject to unsuspected analytical bias. Full details of the scheme may be found in Thompson (2002), Potts *et al.* (2013) and the various Geo*PT* reports cited below.

Although this proficiency testing scheme is designed to provide quality control information to participating laboratories, it is occasionally used for the selection of laboratories judged to have satisfactory performance to contribute to the certification of geological reference materials using the IAG certification protocol (Kane *et al.* 2003). The IAG protocol was designed to establish good practice in the certification of geological reference materials by complying to the fullest extent possible with ISO Guide 35 (2006). Certified reference values and associated uncertainties obtained for CRMs produced according to the IAG protocol are essentially the most reliable values currently available, representing the current best estimates of locations of possible true composition values. Six samples, to date, have been both circulated as Geo*PT* proficiency testing samples and certified as reference materials. These samples are IAG OU-6 (Penrhyn slate), MGL-GAS (serpentinite), MGL-OShBO (alkaline granite), NIST SRM 688 (basalt), BRP-1 (basalt) and MGL-AND (andesite). Of these, NIST SRM 688, BRP-1 and MGL-AND were independently certified before being made available for circulation as Geo*PT* test samples and were distributed anonymously, so that Geo*PT* participating laboratories had no prior knowledge of the compositions.

doi: 10.1111/j.1751-908X.2015.00338.x

Geostandards and Geoanalytical Research (2015), 39, 407-417).

GeoPT and traceability

Values that are certified using the GeoPT certification protocol are primarily traceable to a rigorously assessed consensus values derived from this well characterized proficiency testing scheme itself, noting:

PROs

- Over **100 participating laboratories**.
- A rigorous assessment of consensus values from the distribution of submitted results is undertaken in GeoPT that has allowed the recognition of **method bias**.
- GeoPT consensus values agree with certified values when **CRMs** have been available as test materials.
- The GeoPT certification scheme requires the **co-distribution of an established matrix-matched CRM**.

CONs

- Traceability to a **consensus value** is controversial.
- **Lack of clarity** in ISO-REMCO Guide 35 in the use of proficiency testing schemes for certification creating uncertainty in some aspects of the present proposal.

However:

The last word on GeoPT and traceability?

Abraham Lincoln quote	Measurement science equivalent
You can fool all of the people some of the time.	Reporting measurement results with an incomplete uncertainty budget.

GeoPT and traceability – the last word?

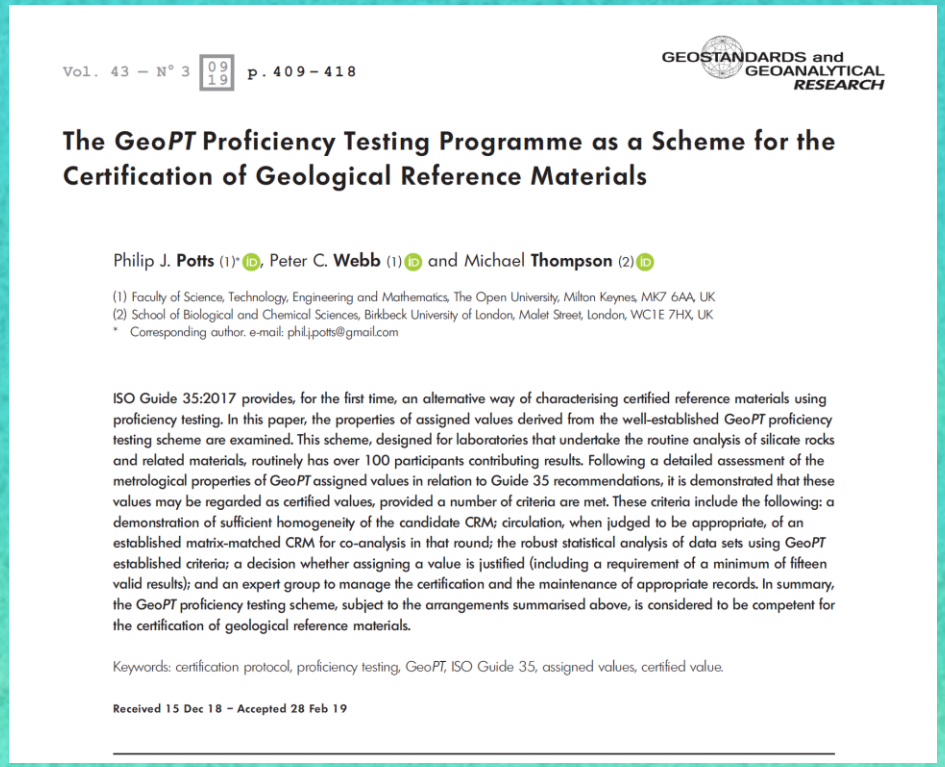
Abraham Lincoln quote	Measurement science equivalent
You can fool all of the people some of the time.	Reporting measurement results with an incomplete uncertainty budget.
You can fool some of the people all of the time.	A reference to those laboratories that do not participate in proficiency testing??

GeoPT and traceability – the last word?

Abraham Lincoln quote	Measurement science equivalent
You can fool all of the people some of the time.	Reporting measurement results with an incomplete uncertainty budget.
You can fool some of the people all of the time.	A reference to those laboratories that do not participate in proficiency testing??
But you cannot fool all of the people all of the time	"A properly developed consensus from a proficiency test is, for all practical purposes, interchangeable with a certified value derived from an interlaboratory comparison" (Thompson 2018)

And finally – A summary of GeoPT as a certification scheme

- The **candidate CRM** and an independently **certified CRM (matrix matched)** are circulated as test materials in the same GeoPT round.
- Consensus values are evaluated using **robust statistics** from the **full set of data** submitted by participating laboratories for both test materials.
- A comparison is made between certified and GeoPT consensus values for the established CRM to establish the **absence of significant bias**.
- Consensus values in the candidate CRM are considered to qualify as certified values if they comply with the **rigorous criteria as GeoPT 'assigned' values**.



Vol. 43 – N° 3 09 19 p. 409–418

GEOSTANDARDS and GEOANALYTICAL RESEARCH

The Geo*PT* Proficiency Testing Programme as a Scheme for the Certification of Geological Reference Materials

Philip J. **Potts** (1)*, Peter C. **Webb** (1) and Michael **Thompson** (2)

(1) Faculty of Science, Technology, Engineering and Mathematics, The Open University, Milton Keynes, MK7 6AA, UK
(2) School of Biological and Chemical Sciences, Birkbeck University of London, Malet Street, London, WC1E 7HX, UK
* Corresponding author. e-mail: phil.j.potts@gmail.com

ISO Guide 35:2017 provides, for the first time, an alternative way of characterising certified reference materials using proficiency testing. In this paper, the properties of assigned values derived from the well-established Geo*PT* proficiency testing scheme are examined. This scheme, designed for laboratories that undertake the routine analysis of silicate rocks and related materials, routinely has over 100 participants contributing results. Following a detailed assessment of the metrological properties of Geo*PT* assigned values in relation to Guide 35 recommendations, it is demonstrated that these values may be regarded as certified values, provided a number of criteria are met. These criteria include the following: a demonstration of sufficient homogeneity of the candidate CRM; circulation, when judged to be appropriate, of an established matrix-matched CRM for co-analysis in that round; the robust statistical analysis of data sets using Geo*PT* established criteria; a decision whether assigning a value is justified (including a requirement of a minimum of fifteen valid results); and an expert group to manage the certification and the maintenance of appropriate records. In summary, the Geo*PT* proficiency testing scheme, subject to the arrangements summarised above, is considered to be competent for the certification of geological reference materials.

Keywords: certification protocol, proficiency testing, Geo*PT*, ISO Guide 35, assigned values, certified value.

Received 15 Dec 18 – Accepted 28 Feb 19

Geostandards and Geoanalytical Research (2019), 43, 409-418

This scheme is currently being applied to the certification of a Meissen Granite.

With thanks to:

- Pete Webb, Mike Thompson and the late John Watson (pioneers of the GeoPT scheme).
- The many colleagues, especially within the International Association of Geoanalysts who have supported this work, including Thomas Meisel (data distribution display software).
- Analytical Methods Trust and the organisers of this meeting for this opportunity to present the 2022/3 Theobald Lecture.