Contribution ID: **114** Type: **Oral Presentation**

# Proficiency Test Data Interpretation and Data Rejection

*Monday, 2 May 2022 16:50 (20 minutes)*

The interpretation, evaluation and visualisation of radioactivity proficiency test exercises (PTEs) are presented in this paper.
Data submitted as part of NPL radioactivity measurement PTEs between 1989 and 2018 represents a useful resource for gauging current capabilities for radioactivity measurement since it includes ~14 000 individual reported measurements; since this programme was started in 1989 the submitted data have been of varying quality.
The extant data has been analysed, using the steps laid out in this paper; the data are tested via the ζ-test, z-test and a relative uncertainty test, the R-test. The outcomes from the three tests may be used to assign an overall score to a particular laboratory's reported result for a given measurement. Assigning priority to the tests in this order; ζ-test > z-tests > R-test, with the z- and R-tests subdivided into pass or fail at the lower limit and pass and fail at the upper limit, each test returns a simple binary result. These are combined to generate the overall score; in addition, the individual test results may be combined to give a performance comparator.
The z-test and R-test rely on an assessment of the quality of reported data. Results reported without an uncertainty are excluded from further analysis. Furthermore, data that are 'obviously'wrong (possibly a 'blunder' , according to ISO 13528:2015) are identified. What may constitute 'obviously wrong'(termed an 'improbable deviation'in this paper) may be debated, but for the purposes of this work, these are data that are quite clearly far away from the assigned value, and are identified using an objective treatment. and outliers are identified in the remaining data. Outlier identification is an area of much debate, and differing methods for identifying outliers have been employed, such as Dixon's Q outlier test; Grubbs outlier test; Chauvenet's Criterion. For this work the use of Peirce's Criterion was selected; the advantage of using Peirce's Criterion is that it is based on finding the most probable distribution of data by rejecting one or more contributing values, and then using the culled data to determine overall performance, median relative uncertainty and acceptability limits on relative uncertainties.
Apart from the common visualisations of proficiency test data, an additional graph is used. This is an adaption of the Naji- and Pom- plots, and presents the data by using the relative deviation as the abscissa and the relative reported uncertainty as the ordinate. This plot is bounded by the limits of the ζ-test, z-test and R-test, resulting in a defined area for results that are acceptable, according to the outcomes of the three tests.
The use and interpretation of these tests and scores will be described in detail in the paper and illustrated with a selection of data from past NPL proficiency tests.

**Primary authors:** Dr HARMS, Arvic (CTBTO); JEROME, Simon (Norges miljø- og biovitenskapelige universitet)

**Presenter:** JEROME, Simon (Norges miljø- og biovitenskapelige universitet)

**Session Classification:** Quality and Intercomparisons

**Track Classification:** Quality and Intercomparisons