When the model doesn't cover reality

Stefaan Pommé

Joint Research Centre
Institute for Reference Materials and Measurement

Uncertainties workshop
BIPM, 15 June 2015

The prime objective of JRC-IRMM is…

… **to build *confidence in measurements* and ensure their comparability.**

- Method development and validation
- Validated data
- Reference measurements
- Production of reference materials
- Inter-laboratory comparisons
- Training, knowledge transfer

'Once measured, accepted everywhere.'

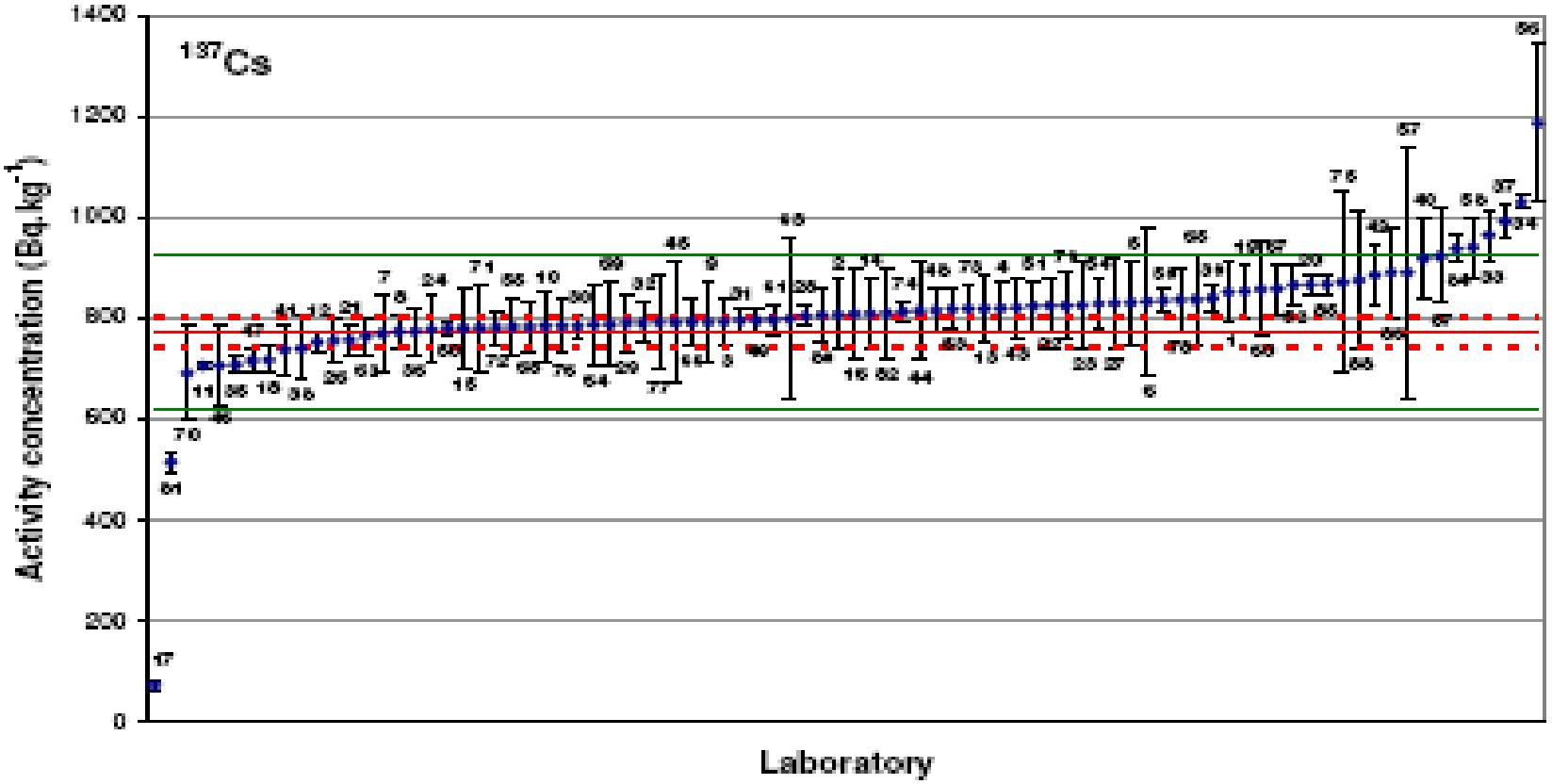
European Commission

Example :

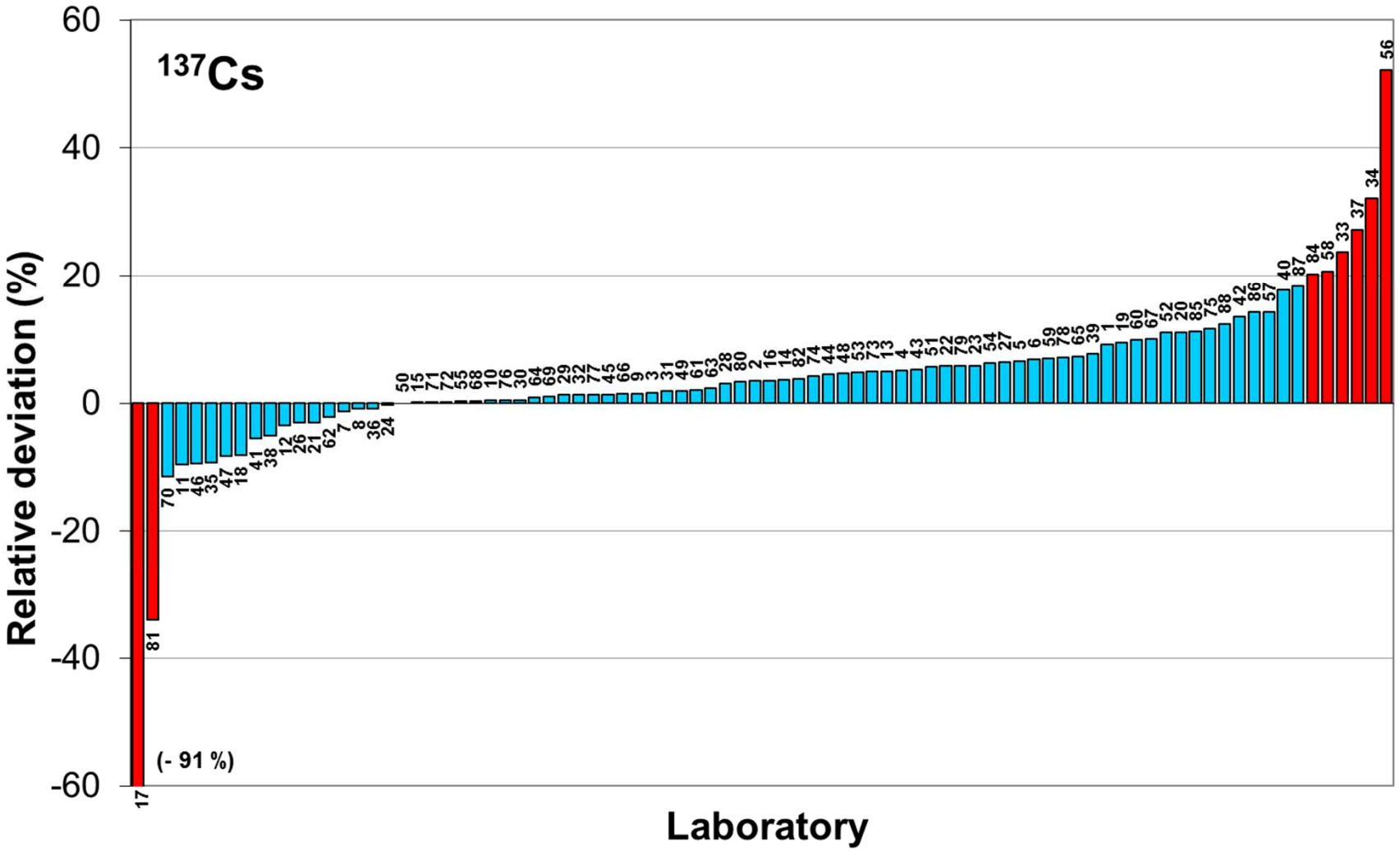
Proficiency test results

Measurement of Radioactivity ^{137}Cs in bilberry

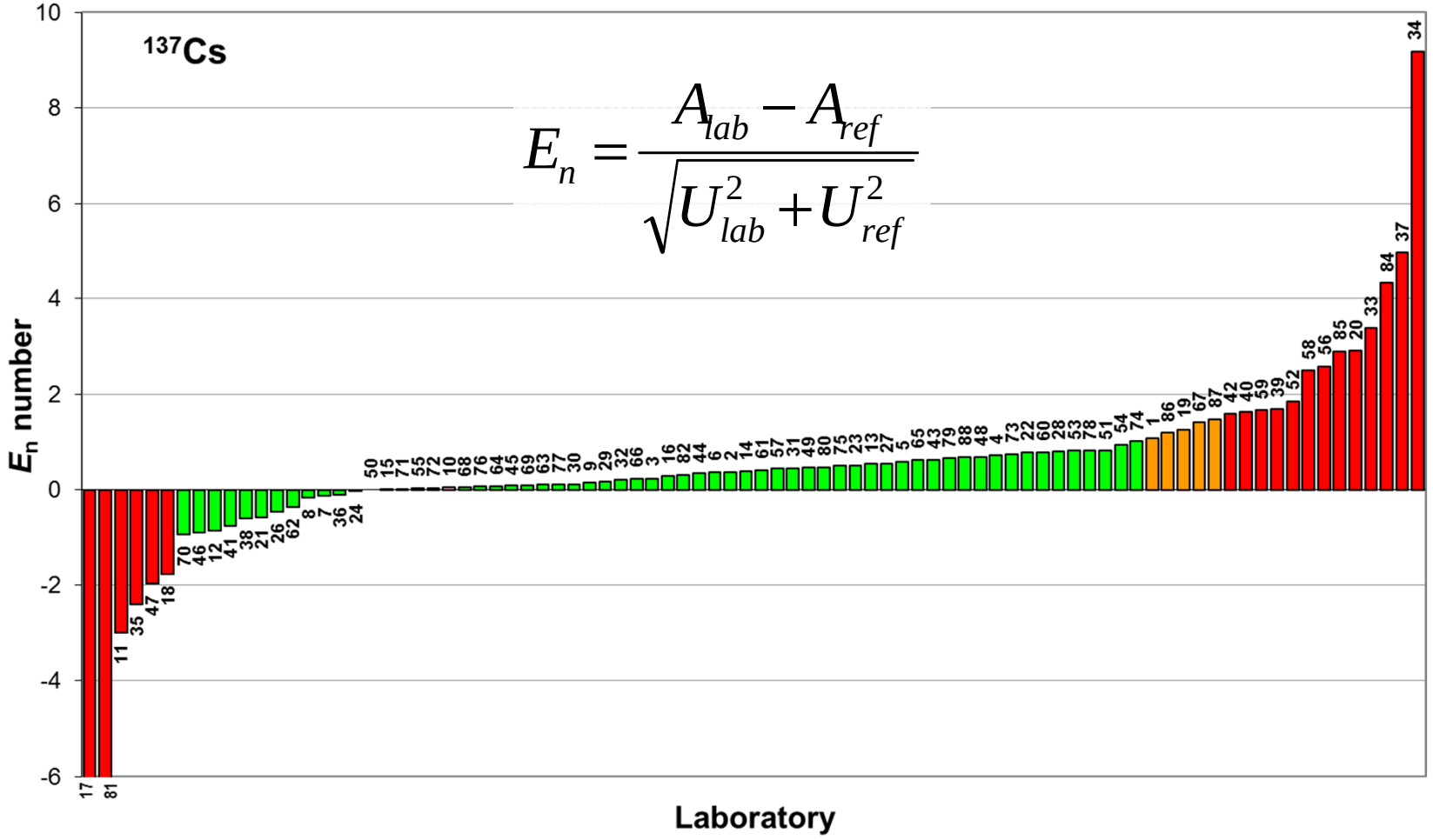
Proficiency test



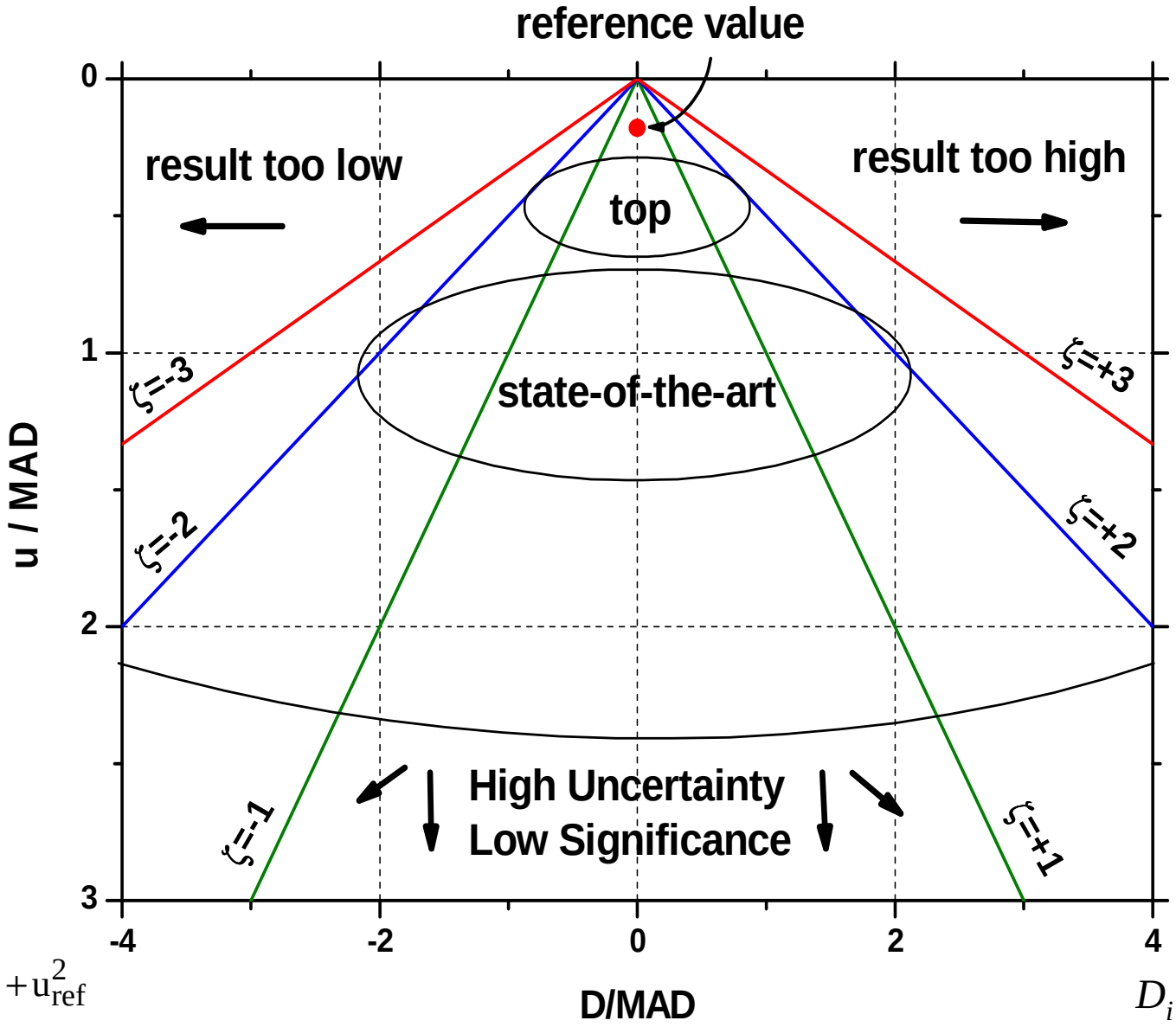
Relative deviation



E_n-score


137Cs
$E_n = \frac{A_{lab} - A_{ref}}{\sqrt{U_{lab}^2 + U_{ref}^2}}$
En number
Laboratory
10
8
6
4
2
0
-2
-4
-6
17 81 11 35 47 18 70 46 12 41 38 21 26 62 8 7 36 1 24 50 15 71 55 72 10 68 76 64 45 69 63 77 30 9 29 32 66 3 16 82 44 6 2 14 61 57 31 49 80 75 23 13 27 5 65 43 79 88 48 4 73 22 60 28 53 78 51 54 74 1 86 19 67 87 42 40 59 39 52 58 56 85 20 33 84 37 34

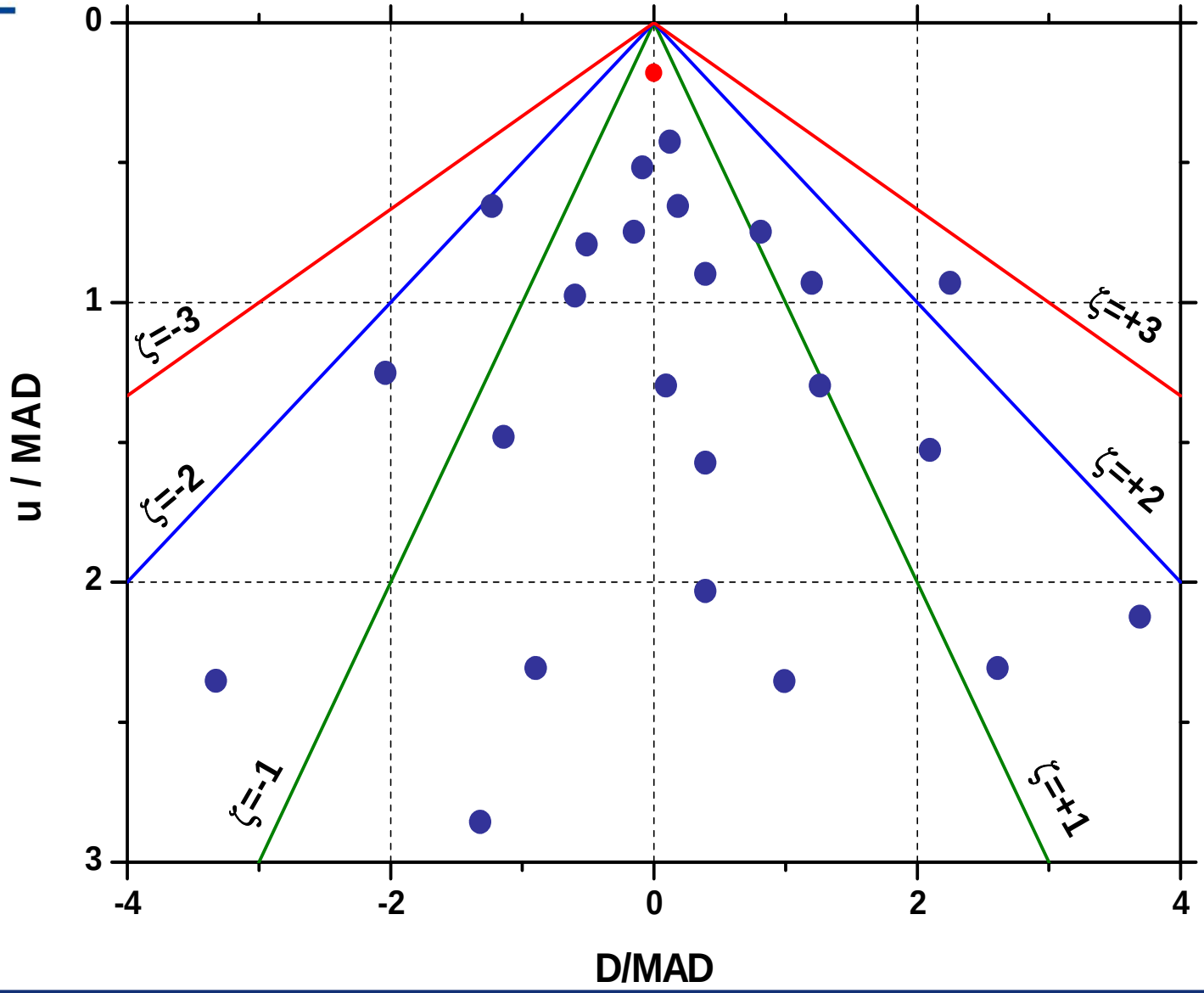
PomPlot



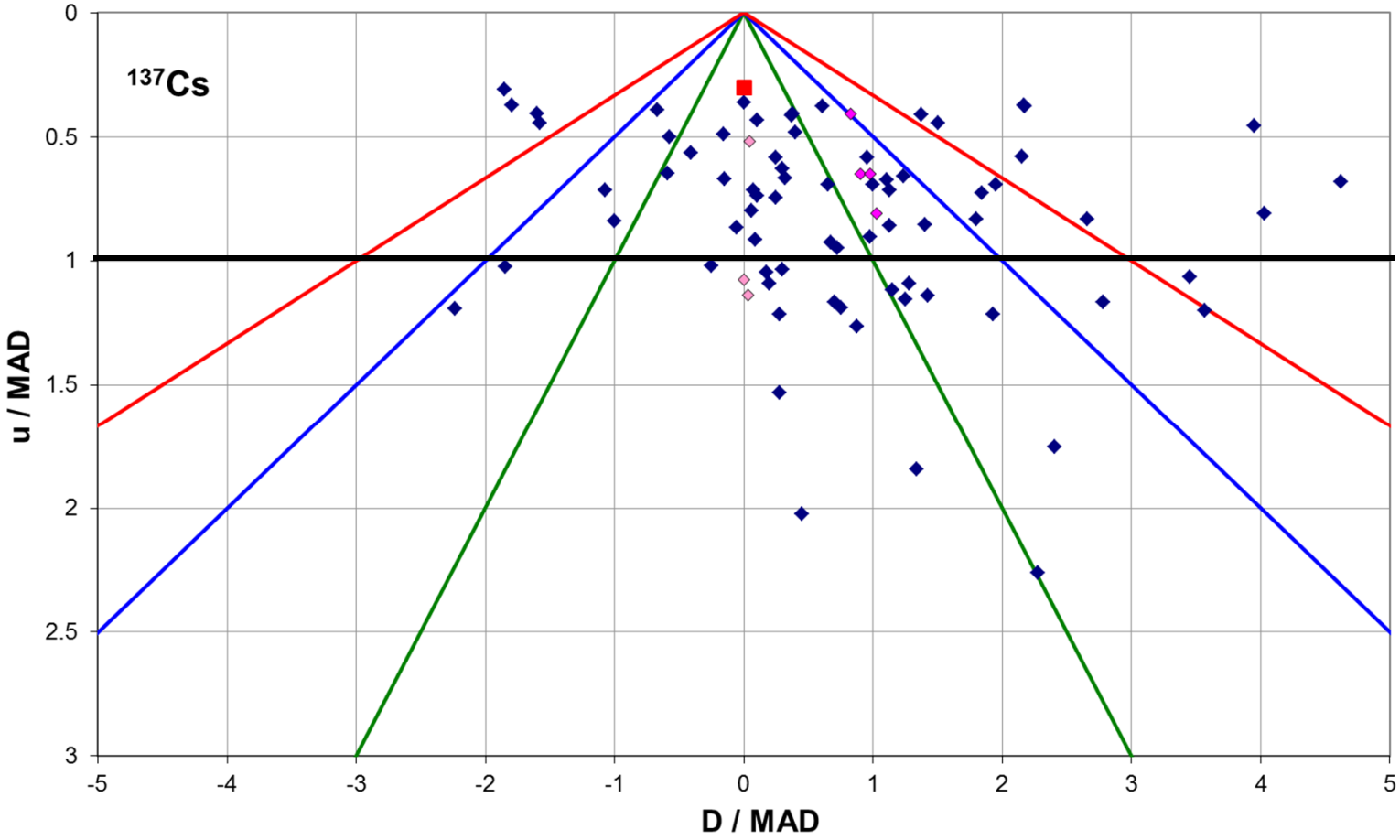
$u_i^2 = u_{lab,i}^2 + u_{ref}^2$

$D_i = A_{lab,i} - A_{ref}$

‘Normal’ distribution of data

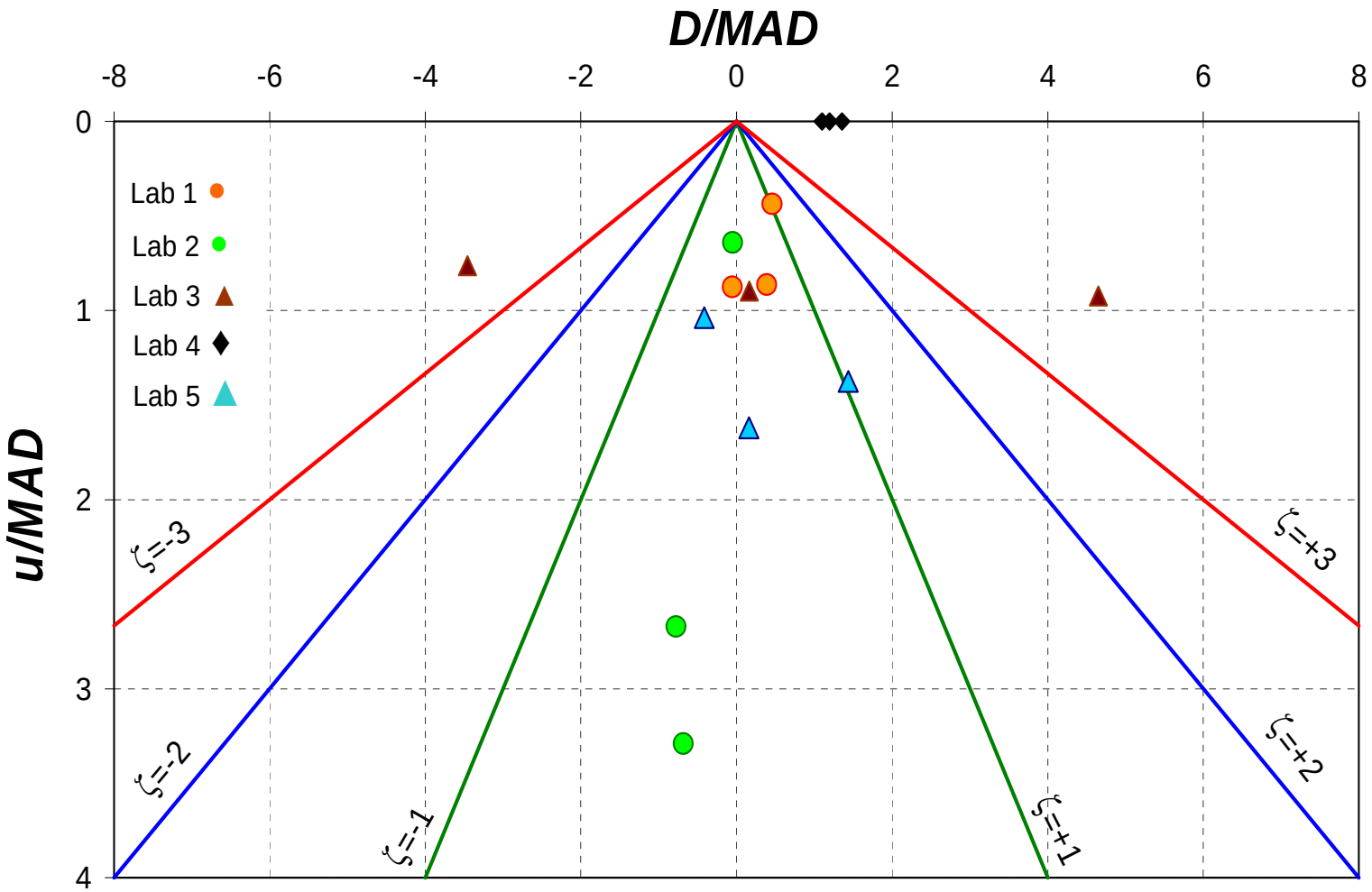


Cs-137 in bilberry



Accred Qual Assur 12, 623

Combined PomPlots for individual labs



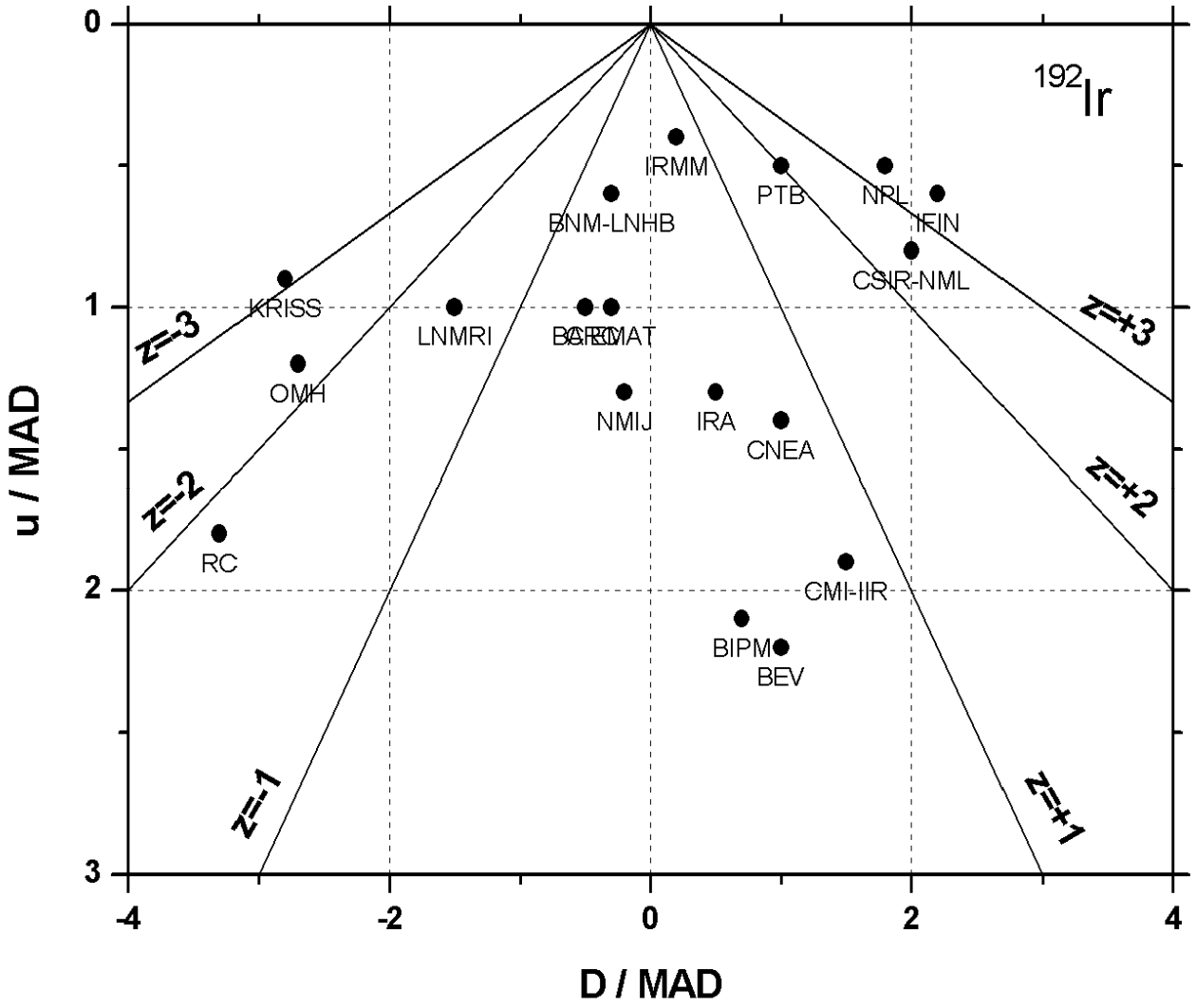
European Commission

Do NMIs control uncertainty better?

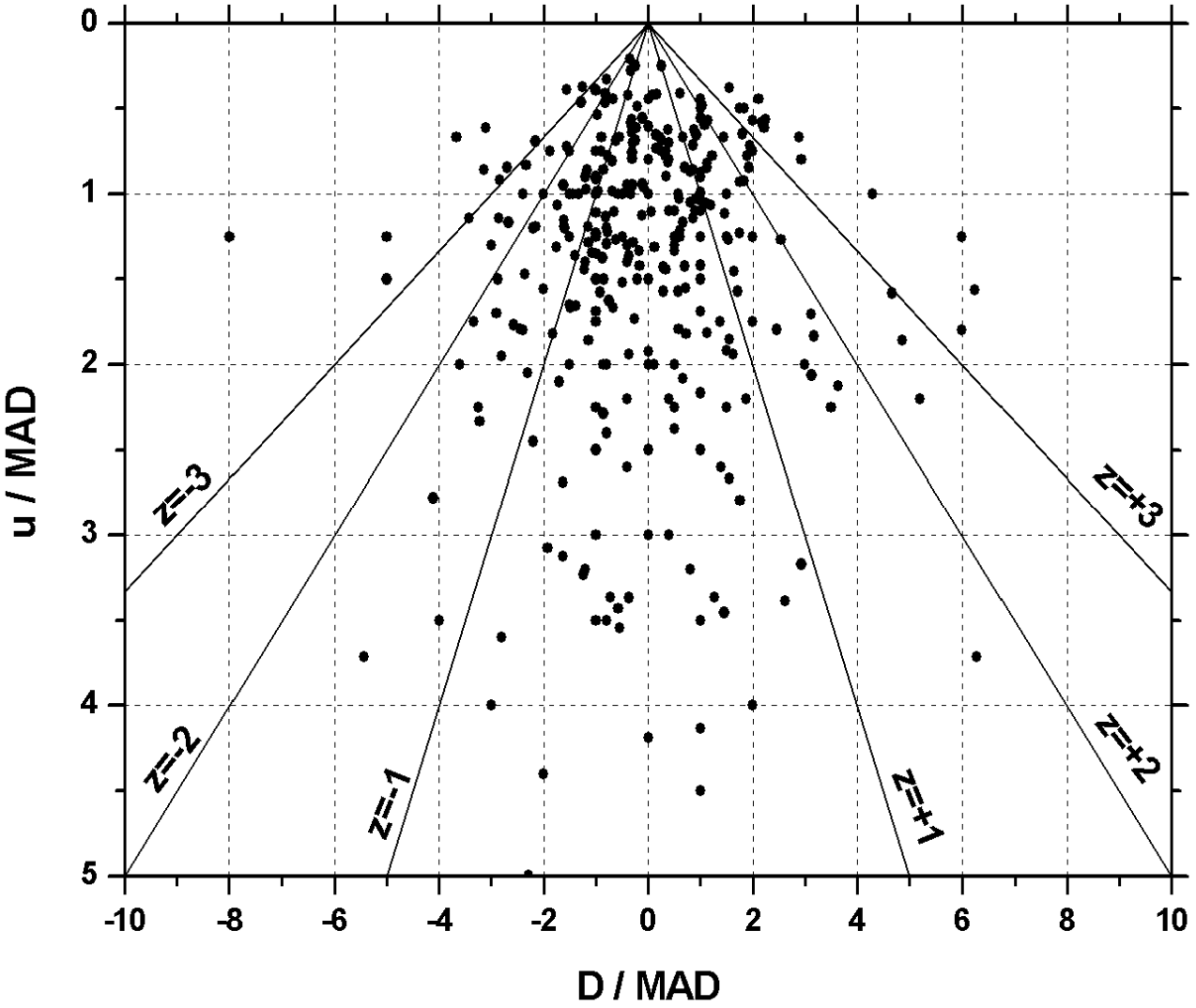
CCRI(II) Key Comparison data

Activity concentration in a solution

Pomplot of 1 key comparison

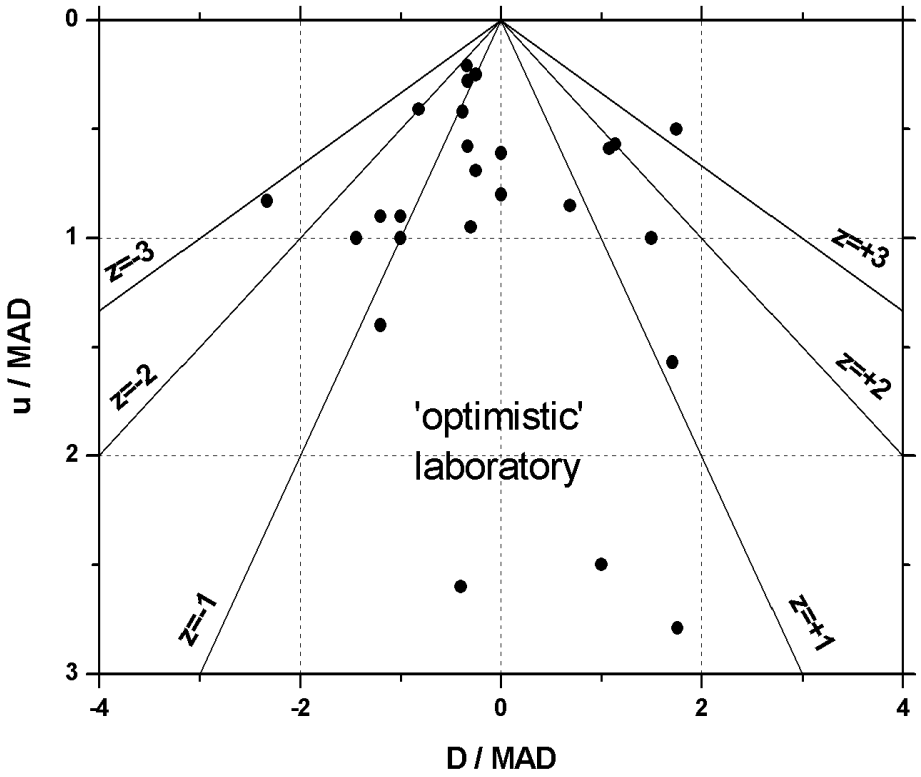


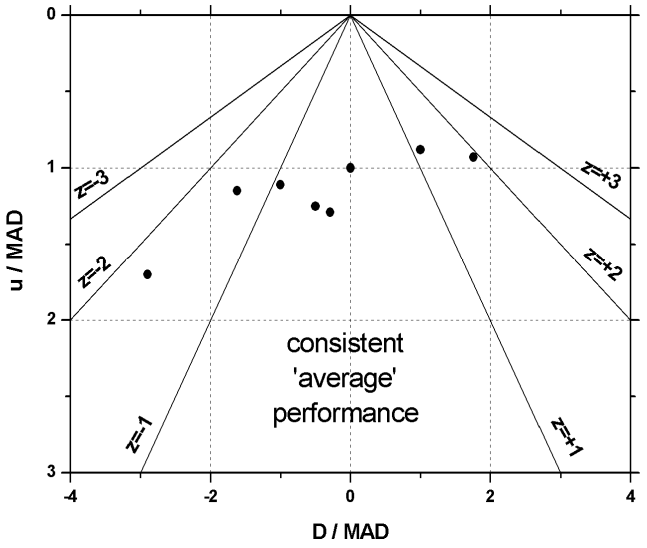
Pomplot of all KCs in KCDB

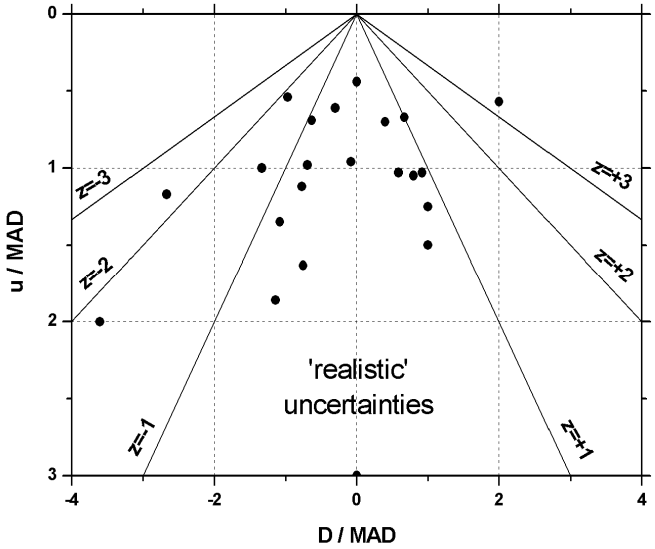


Appl. Rad. Isot. 64, 1158

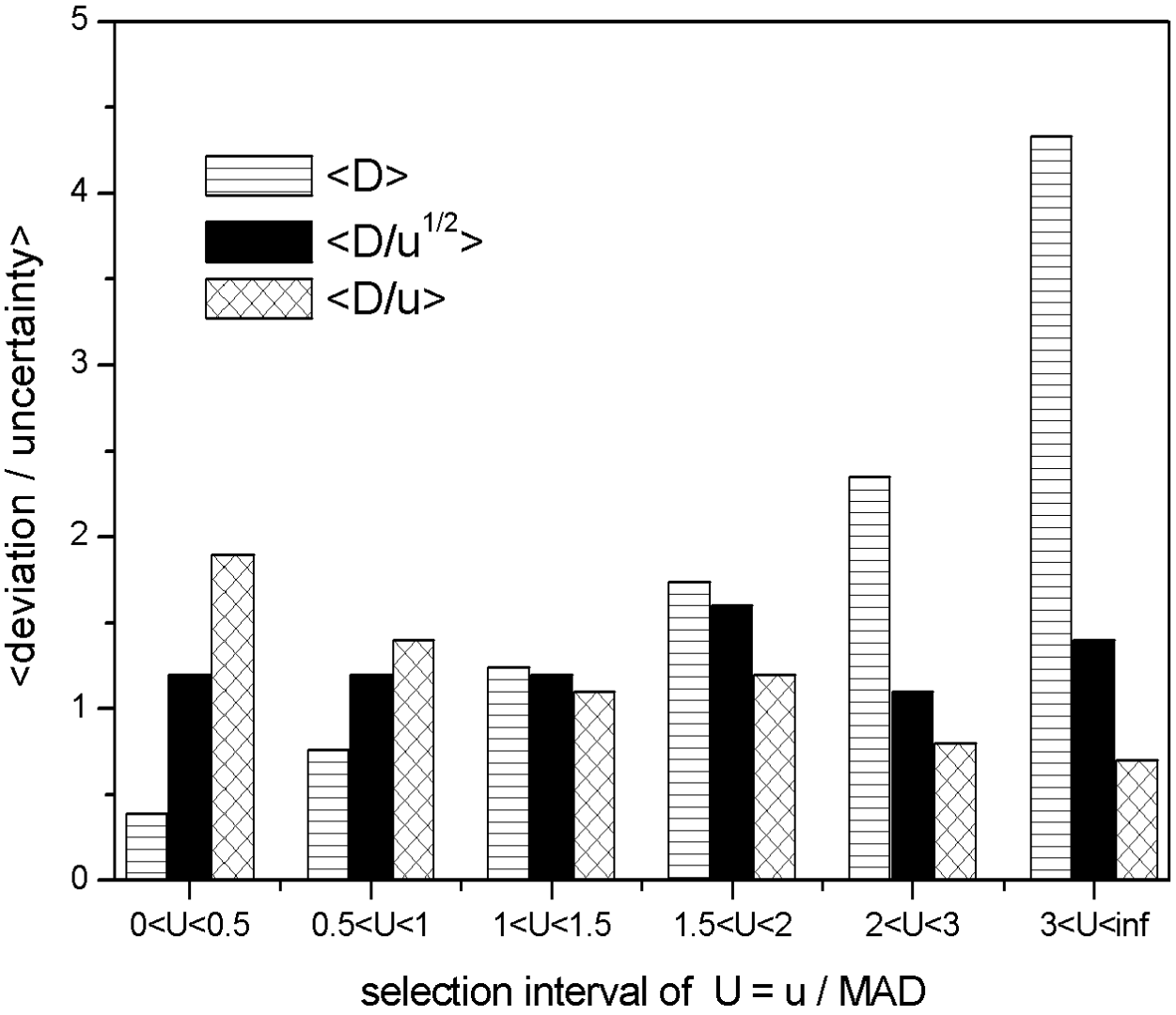
'Fingerprint' of laboratories



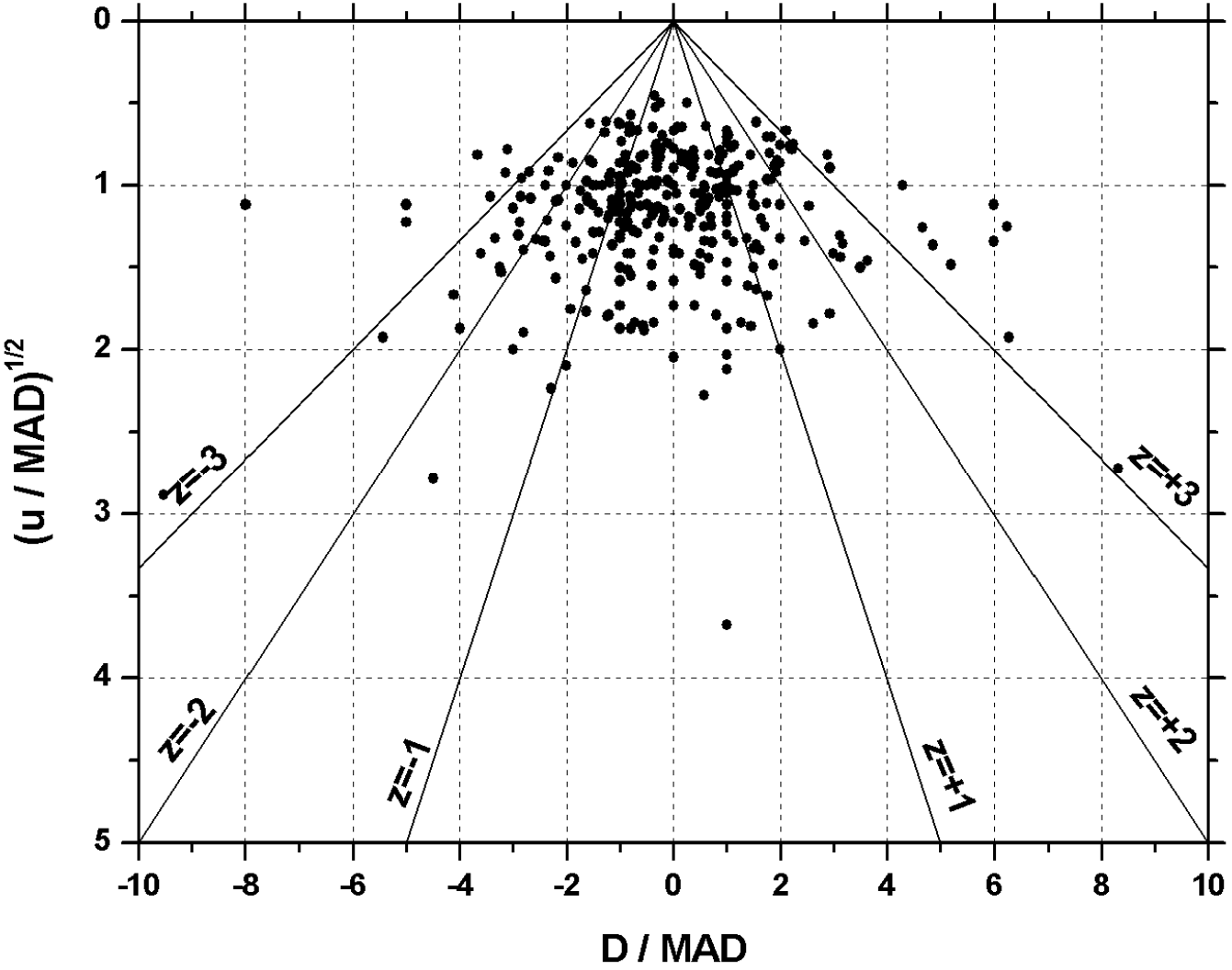




KCDB CCRI(II) data: D proportional to sqrt(u)



PomPlot of data with sqrt(u) -> looks more 'normal'



IC uncertainty evaluation

Metrologia **51** (2014) *Tech. Suppl.* 06018

International Comparison CCRI(II)-S7 on the analysis of Uncertainty Budgets for $4\pi\beta\gamma$ Coincidence Counting

Claude J. Bailat (IRA), John Keightley (NPL), Youcef Nedjadi (IRA), Li Mo (ANSTO), Guy Ratel and Carine Michotte (BIPM), Miguel Roteta (CIEMAT), Maria Sahagia and Anamaria C. Wätjen (IFIN-HH), Ming-Chen Yuan (INER), Jong Man Lee, Tae Soon Park, K.B. Lee, Sang Han Lee and Pil Jae Oh (KRISS), Akira Iwahara (LNMRI/IRD), Yasushi Sato, Yasuhiro Unno and Akira Yunoki (NMIJ), Wilfrid M. van Wyngaardt, Joline Lubbe, Martin J. van Staden and Bruce R.S. Simpson (NMISA), Andrey V. Zanevsky (VNIIM), and François O. Bochud (IRA)

Result from extrapolation

Metrologia **51** (2014) *Tech. Suppl.* 06018

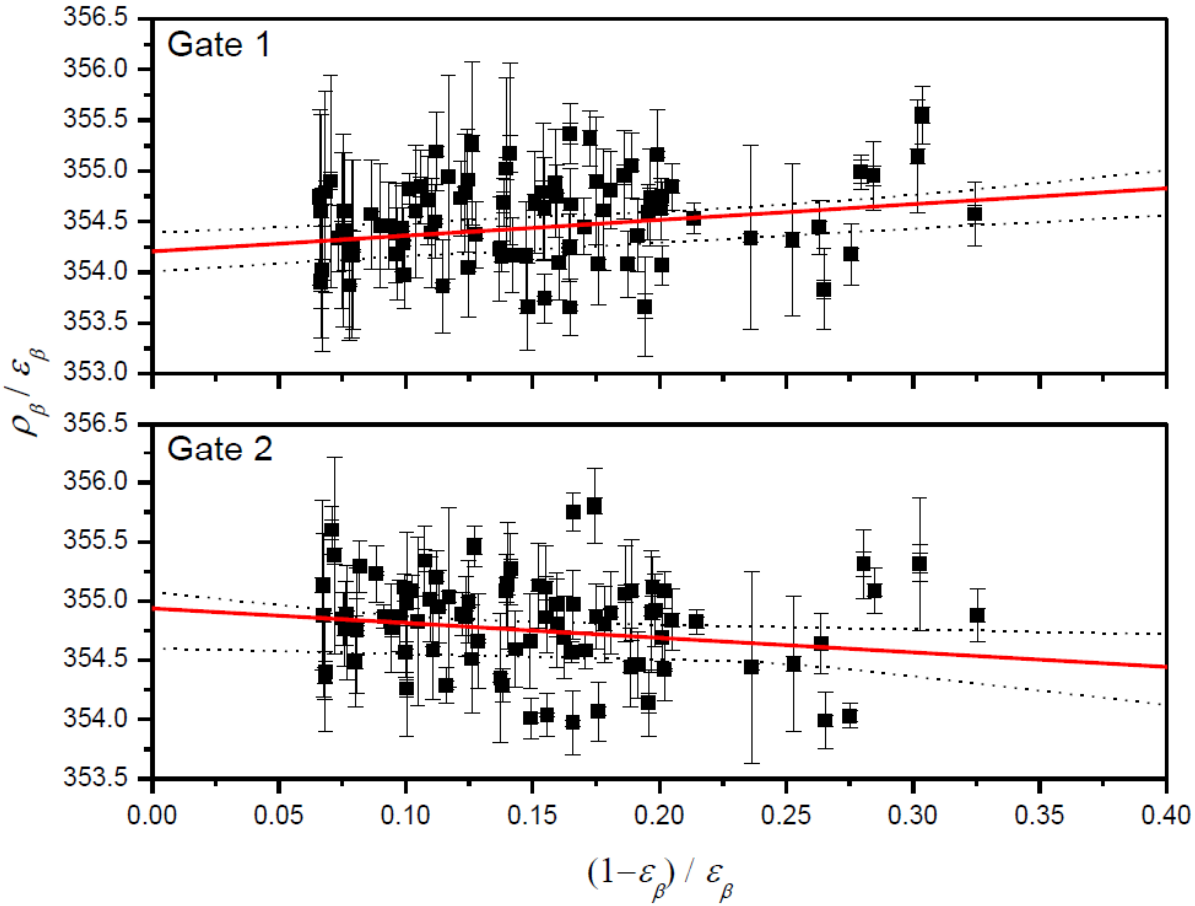


Figure 7: Plot of the extrapolation results for gates 1 and 2 superimposed on the provided experimental data. It shows the average extrapolation results with the minimum and maximum obtained by the participants.

Analysis of 1 data set

Metrologia **51** (2014) *Tech. Suppl.* 06018

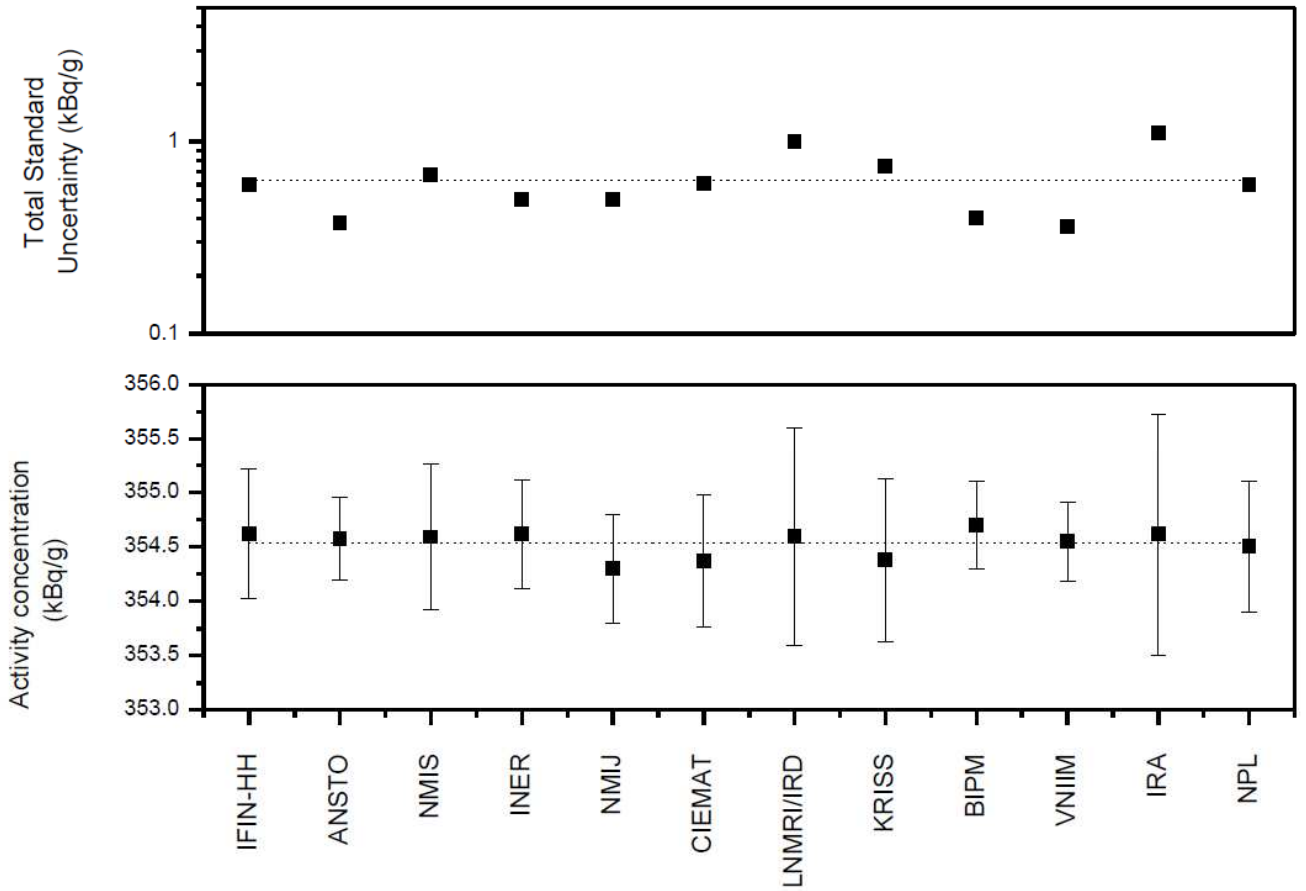


Figure 2: Plot of the final result values obtained by each laboratory.

Uncertainty budgets

Metrologia **51** (2014) *Tech. Suppl.* 06018

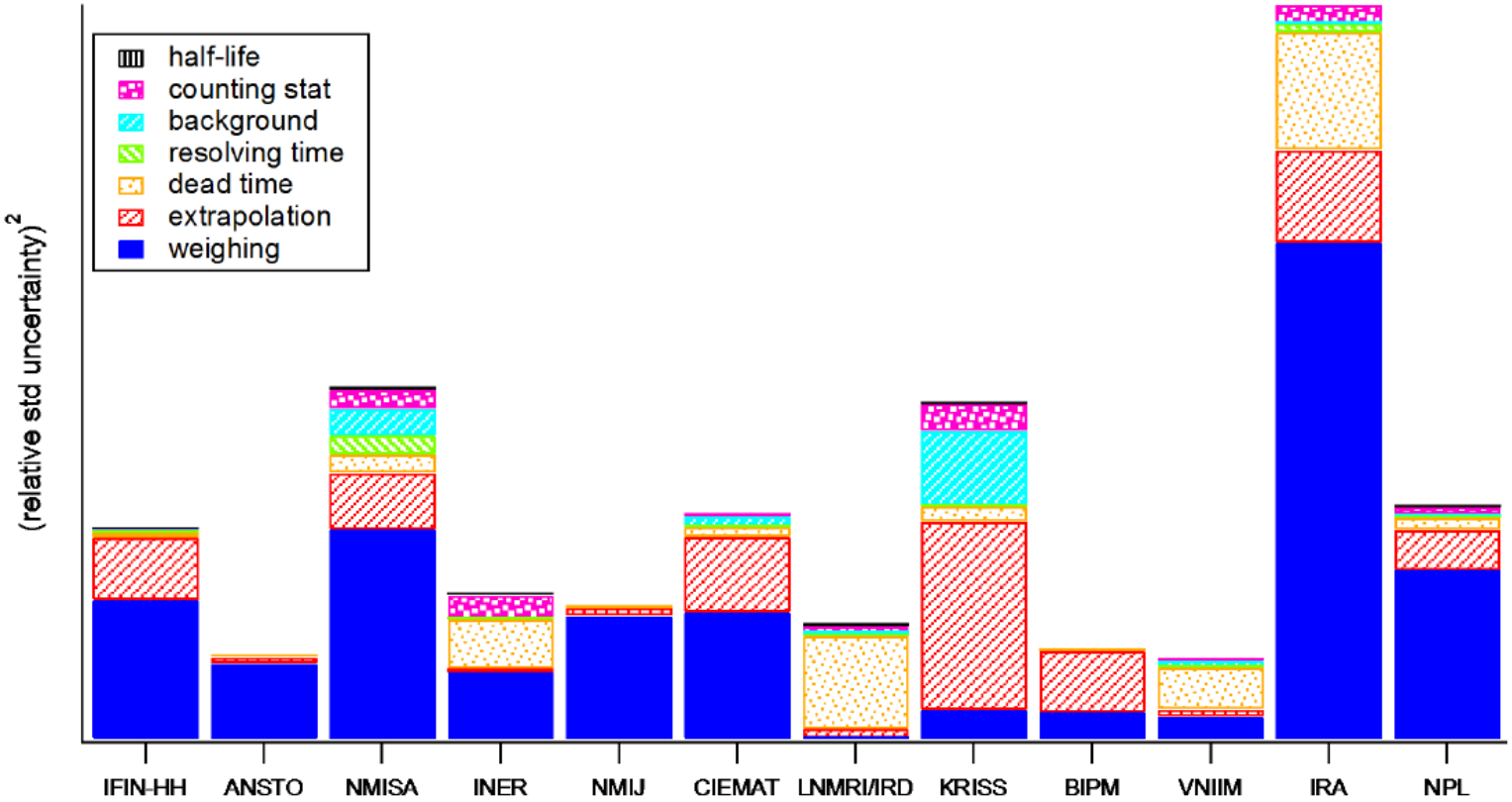


Figure 3: Square of the relative-uncertainty contributions for each laboratory.

spread on uncertainties

Metrologia **51** (2014) *Tech. Suppl.* 06018

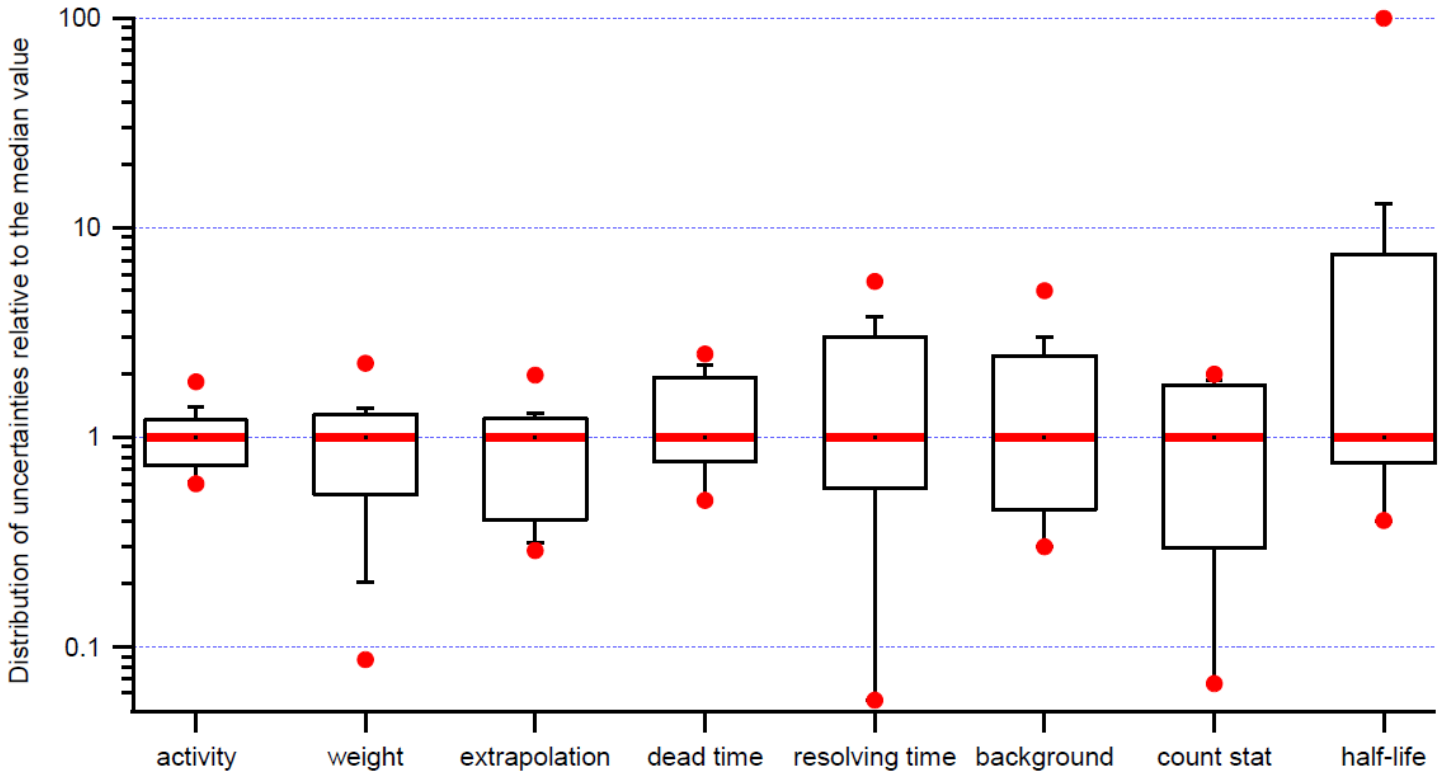


Figure 6: Box-plot showing the spread of the uncertainties estimations. The most left box-plot shows the total uncertainty estimation. The next ones present the different contributions of the uncertainty budget. Each uncertainty value has been normalized by its median value. The central horizontal bold line is the median (set equal to 1). The boxes represent the 25th and the 75th percentiles. The whiskers represent the 10th and the 90th percentiles. The full circles are the points out of the 10th – 90th percentile-interval.

Known - Unknown

"[…] as we know, there are known knowns; there are things we know we know.

We also know there are known unknowns; that is to say we know there are some things we do not know.

But there are also unknown unknowns – the ones we don't know we don't know."

Donald Rumsfeld

Exponential decay

Known known

radioactive decay is a random process

=> activity decreases exponentially with time

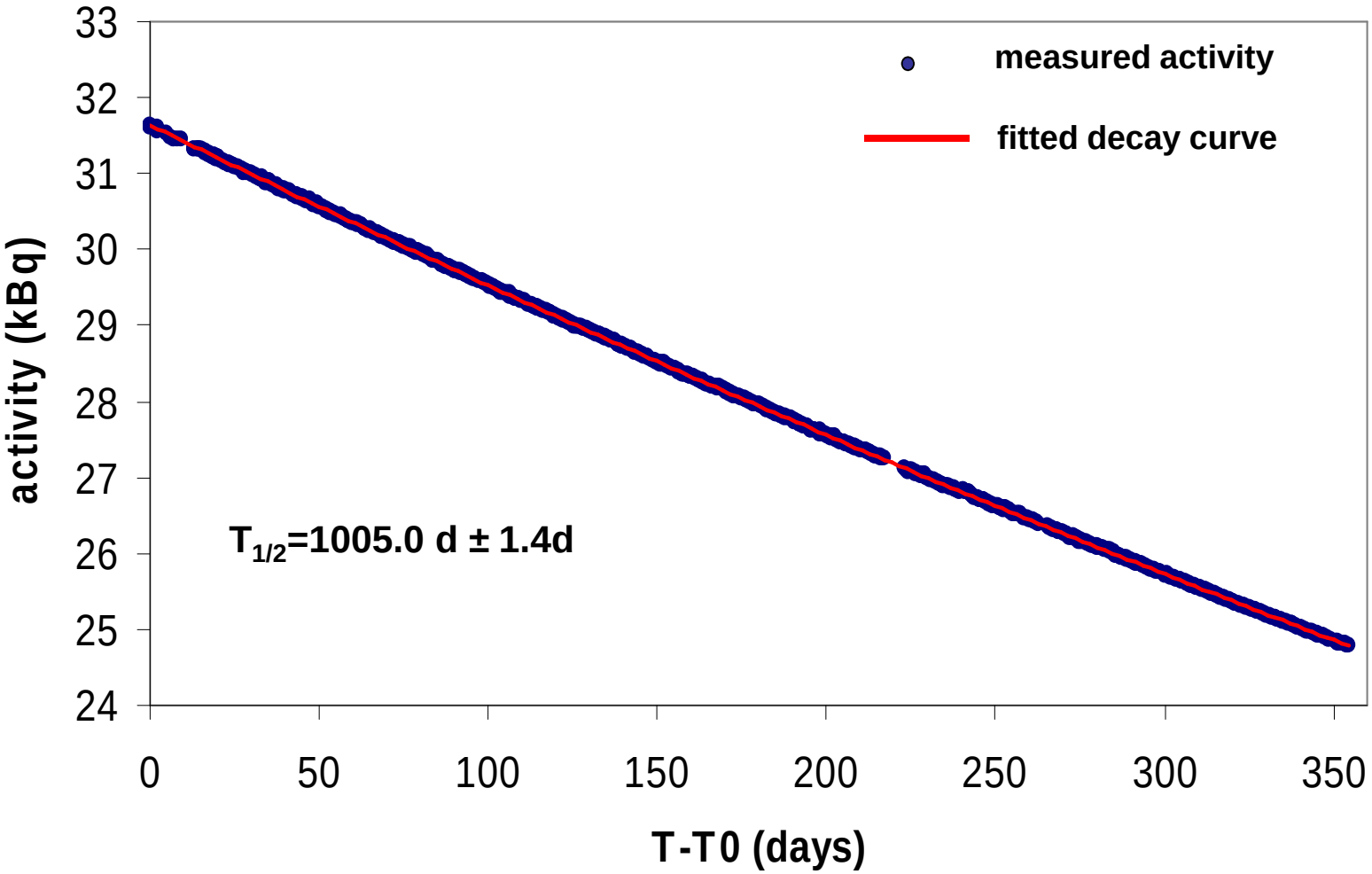
Known unknown

Poisson process with statistical variations

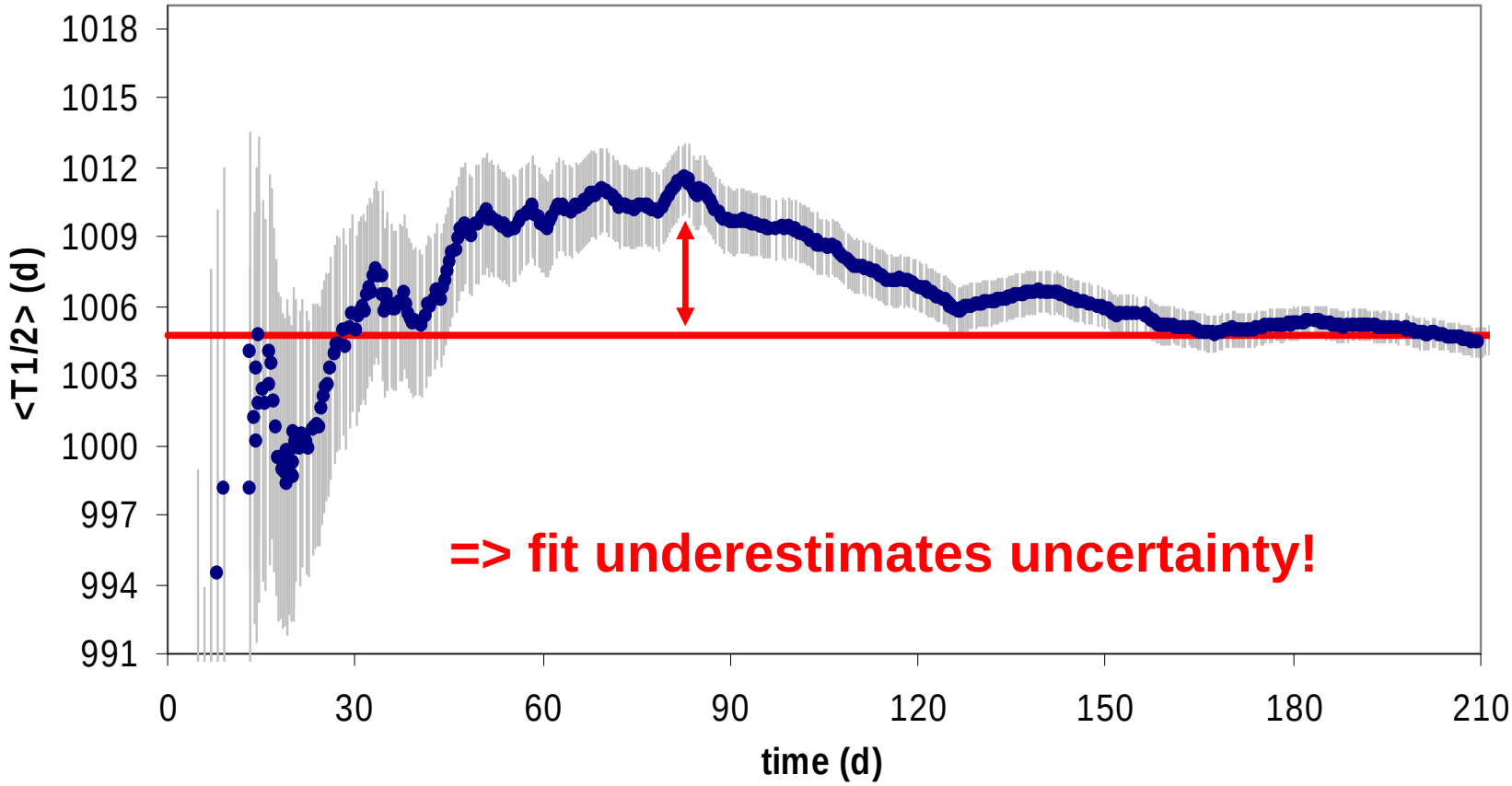
Unknown unknown

Variation in count rate due to instability of measurement equipment

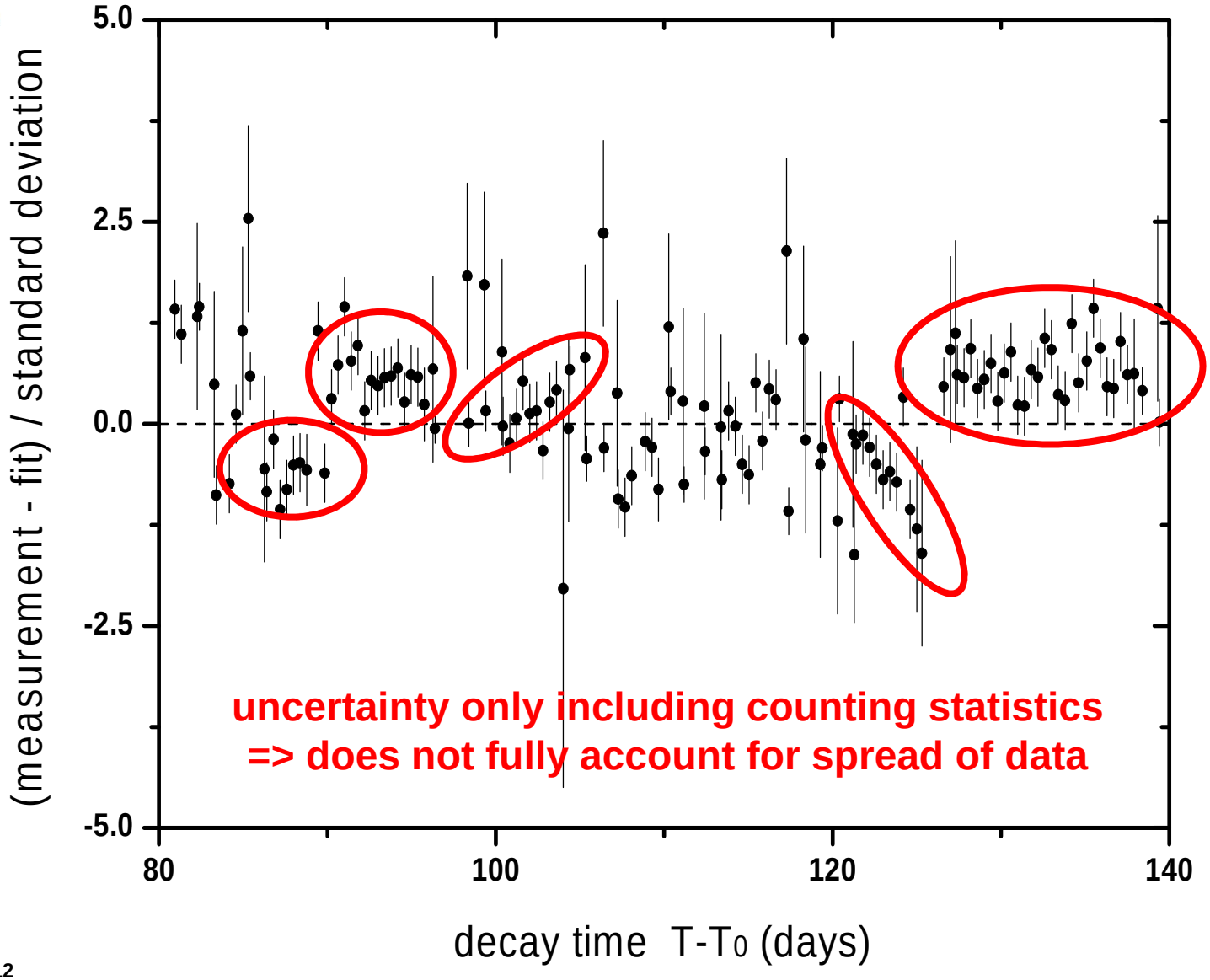
Example: decay of ^{55}Fe



Half-life and uncertainty from least-squares fit



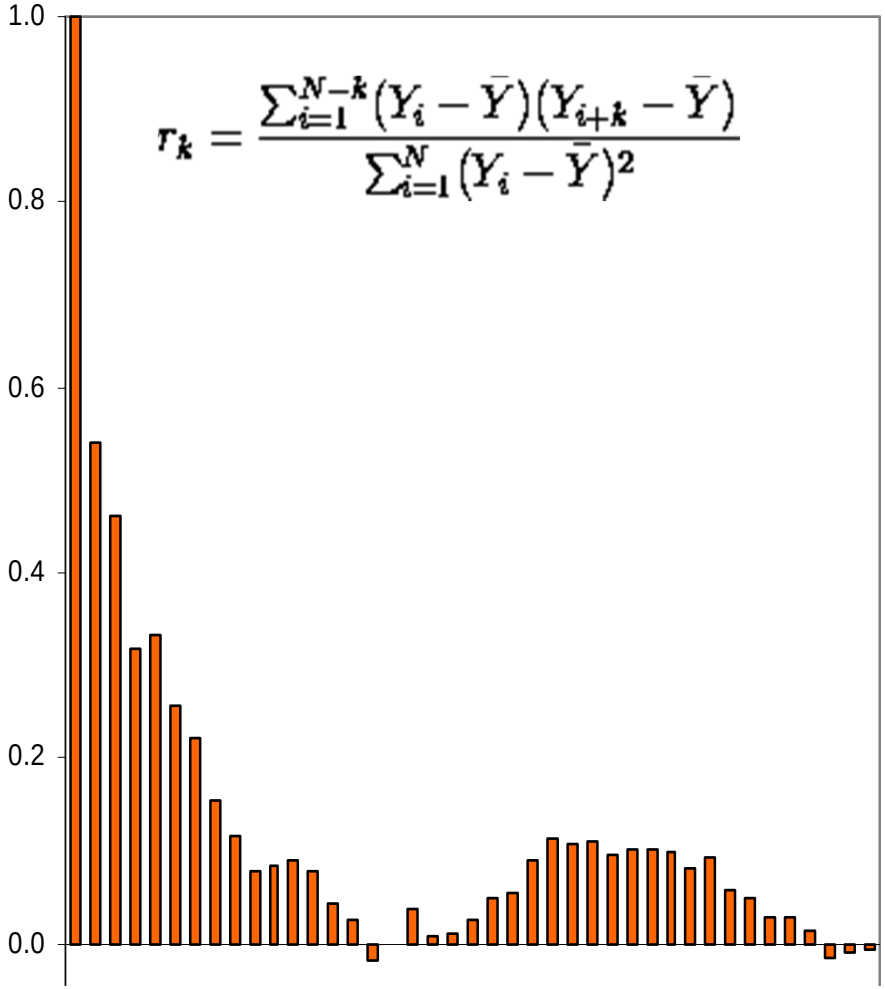
Residuals ^{55}Fe: blow up



Appl. Rad. Isot. 64, 1412

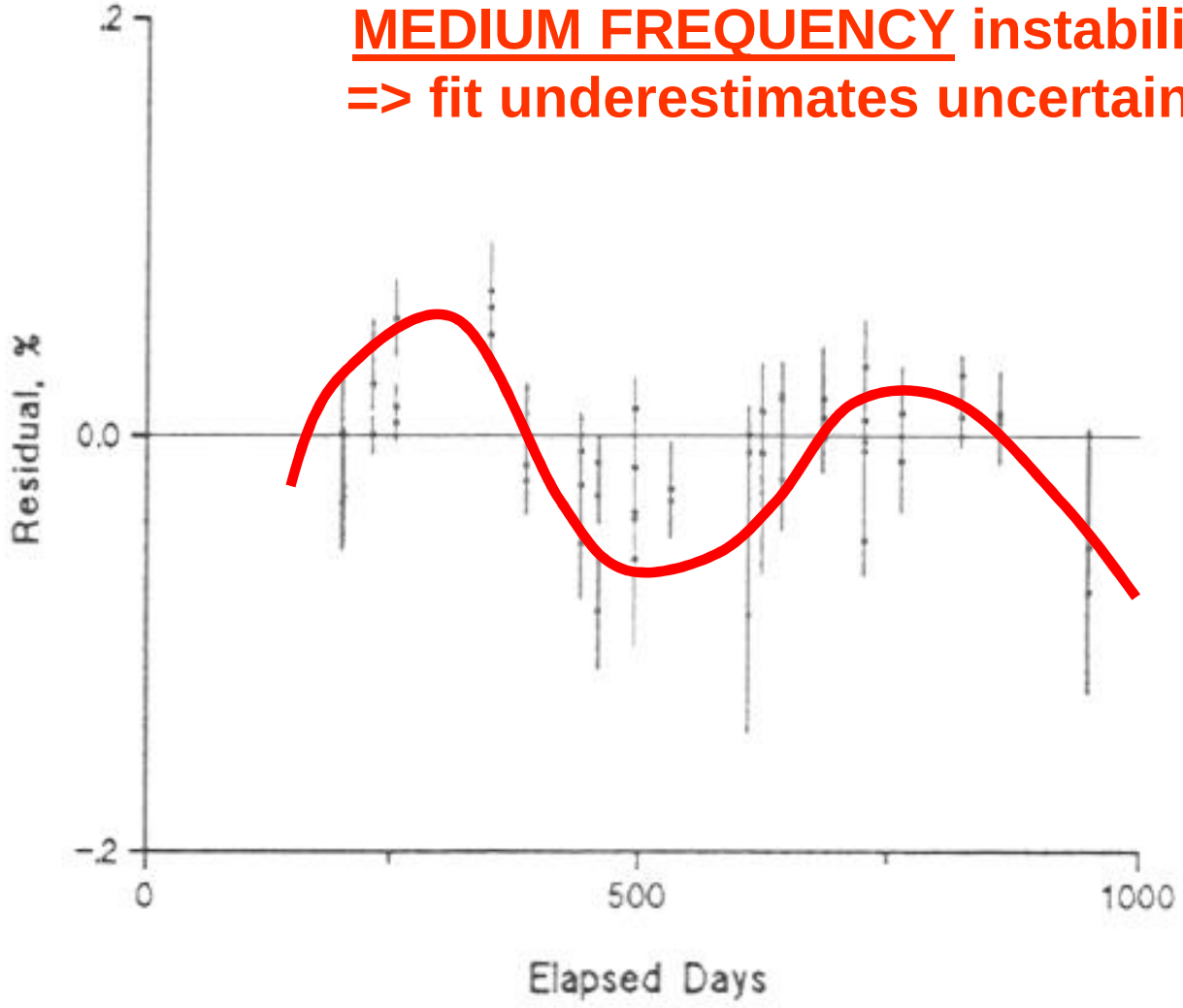
Autocorrelated data => not stochastic

autocorrelation plot of the residuals



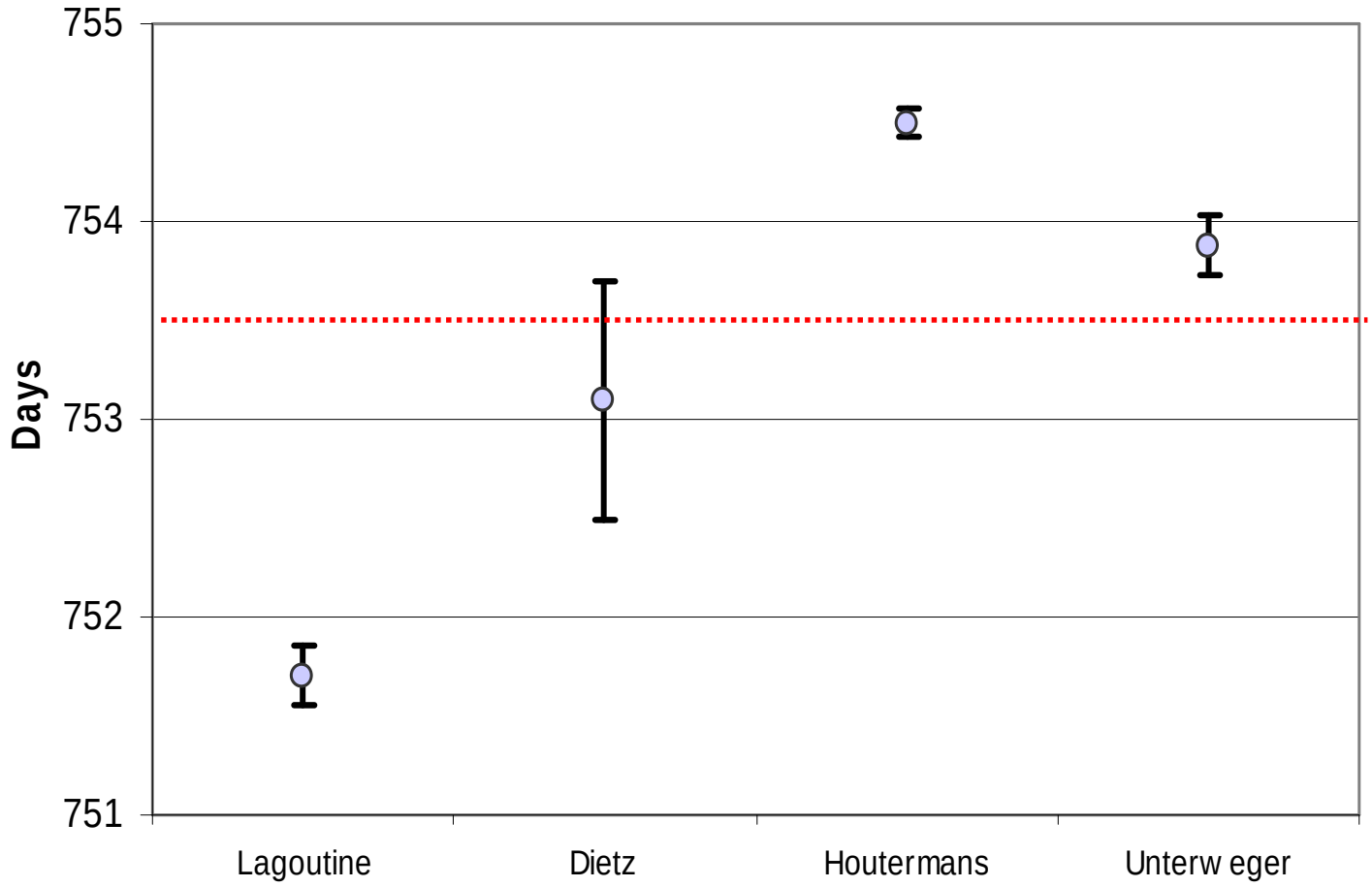
Residuals Cs-134

MEDIUM FREQUENCY instability
=> fit underestimates uncertainty



NIM A390 (1997) 267-273

Half-life of ^{134}Cs



Consequences

*Due to lack of statistical control, **speculations** arose that "half-lives are not constant": depend on solar distance, temperature, chemical state, etc.*

=> proper uncertainty evaluations needed to refute these claims

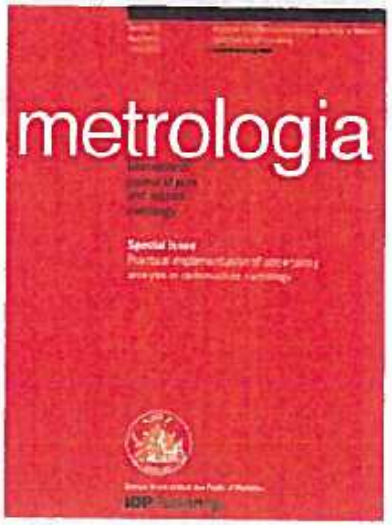
Other examples

Incorrect decay data (e.g. simplified modelling of beta spectra) leads to erroneous detection efficiency calculations

Use of computer codes (e.g. transport simulations) alienates metrologist from the underlying physics/chemistry

Spectral fitting underestimates uncertainty due to imperfection of analytical model to exactly reproduce reality

Metrologia 52 (June 2015!)



Uncertainties in Radionuclide Metrology

GUEST EDITORS

Lisa Karam, NIST, Radiation Physics Division, Gaithersburg, MD, USA [Chairman, CCRI(II)], e-mail lisa.karam@nist.gov

John Keightley, NPL Radioactivity Group, Teddington, Middlesex, UK [Chairman, KCWG(II)], e-mail john.keightley@npl.co.uk

José María Los Arcos, BIPM Head Ionizing Radiation Department, Sèvres CEDEX, France [Head, Ionizing Radiation Department], e-mail jm.losarcos@bipm.org

Detailed uncertainty analysis for each technique underpinning SI-unit becquerel

Power-Moderated Mean

CCRI(II)/13-37

2013 CCRI(II) decision on the evaluation of the Key Comparison Reference Values

1. The key comparison reference value

In May 2013 the CCRI(II) decided to no longer calculate the key comparison reference value (KCRV) by using an unweighted mean but rather by using the power-moderated weighted mean [1]. This type of weighted mean is similar to a Mandel-Paule mean in that the NMIs' uncertainties may be increased until the reduced chi-squared value is one. In addition, it allows for a power smaller than two in the weighting factor. Therefore, all relevant key comparison results can be selected for the KCRV with the following provisions:

Power Moderated Mean - PMM

Metrologia 52, S200

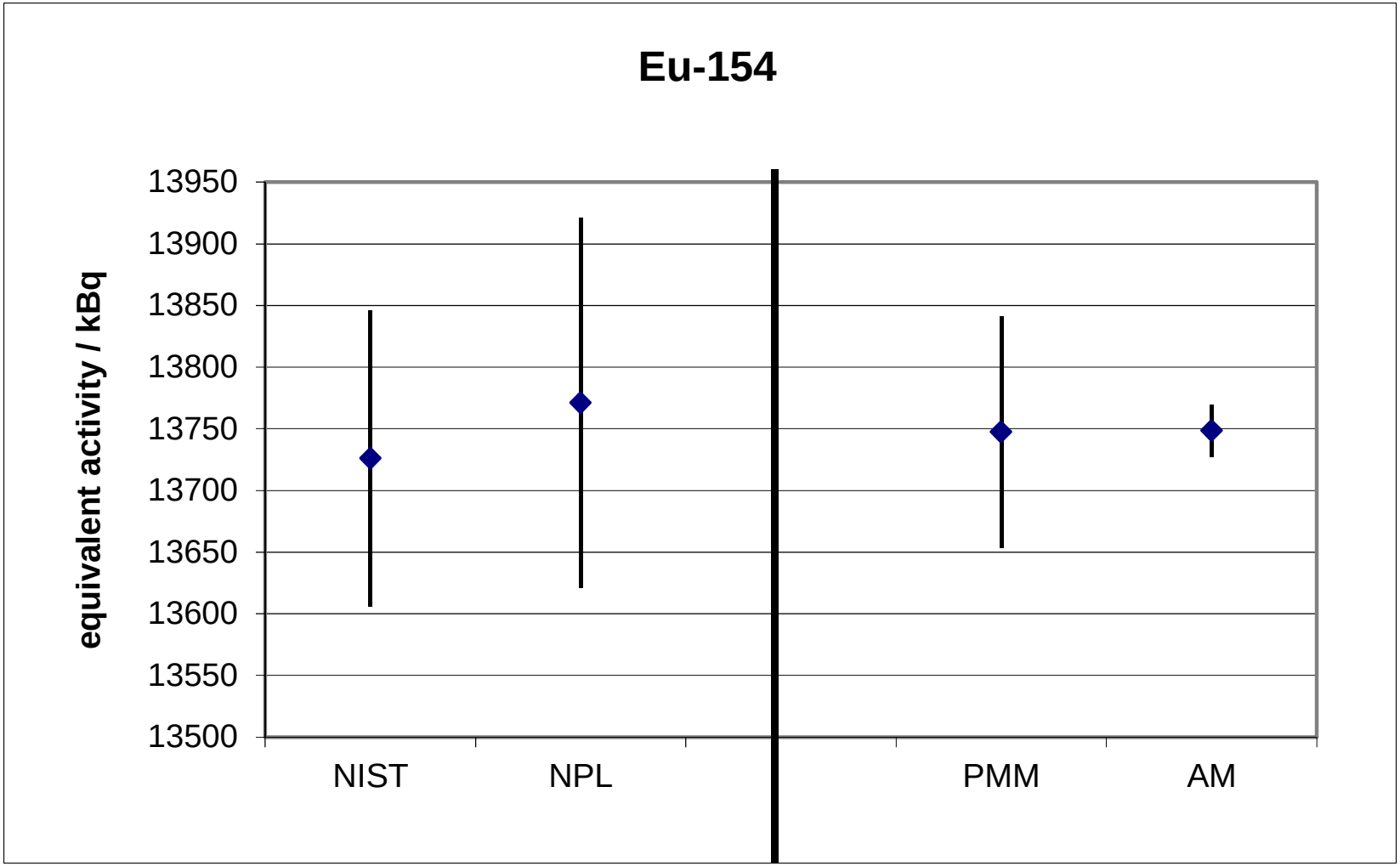
$$x_{\text{ref}} = \sum_{i=1}^{N} w_i x_i$$

$$w_i = u^2(x_{\text{ref}}) \left[\left(\sqrt{u_i^2 + s^2} \right)^{\alpha} S^{2-\alpha} \right]^{-1}$$

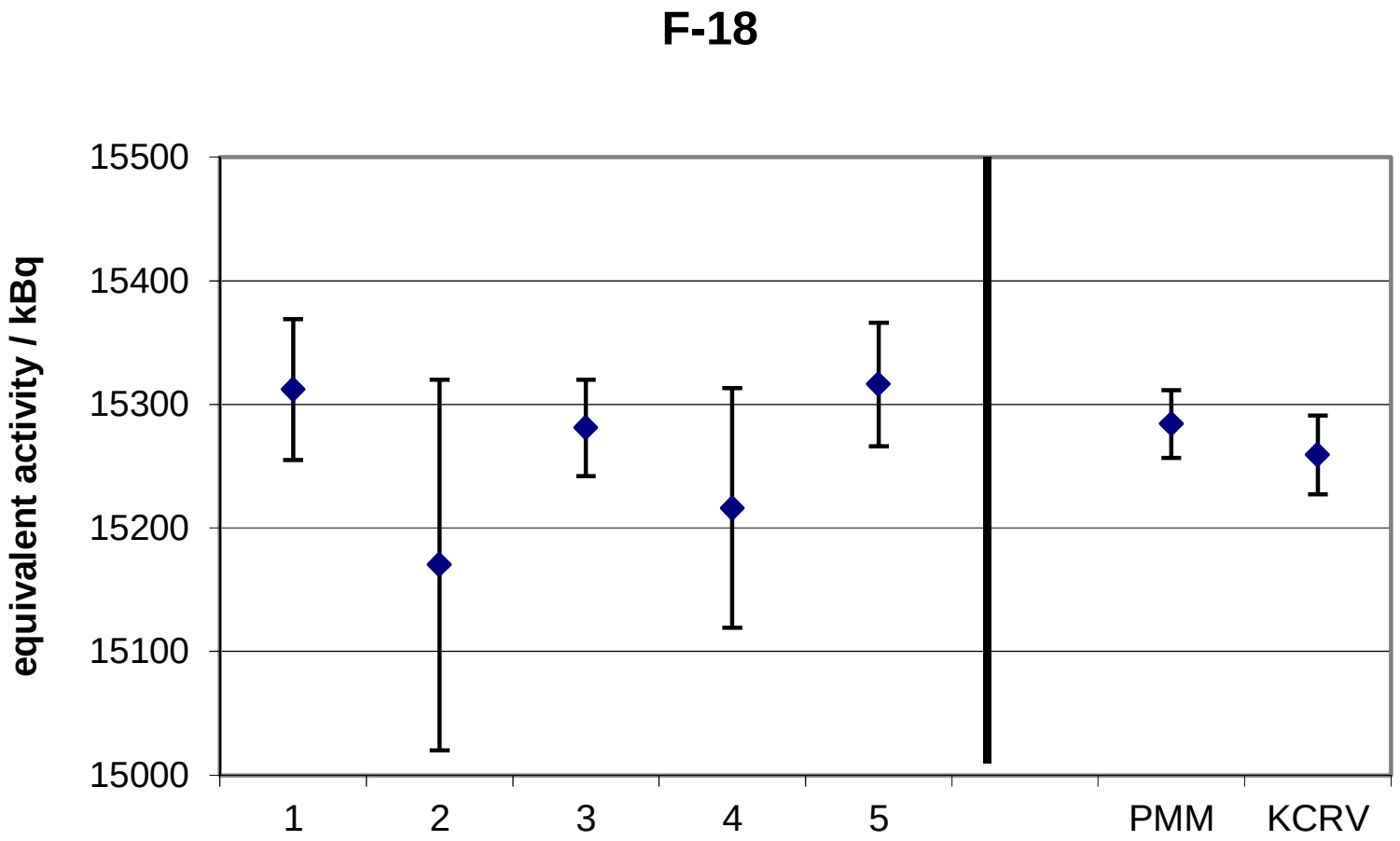
$$S = \sqrt{N \cdot \max(u^2(\bar{x}), u^2(x_{\text{mp}}))}$$

S is a typical uncertainty per datum (max arithmetic or M-P unc)
$0<\alpha<2$ = power reflects level of trust in uncertainties

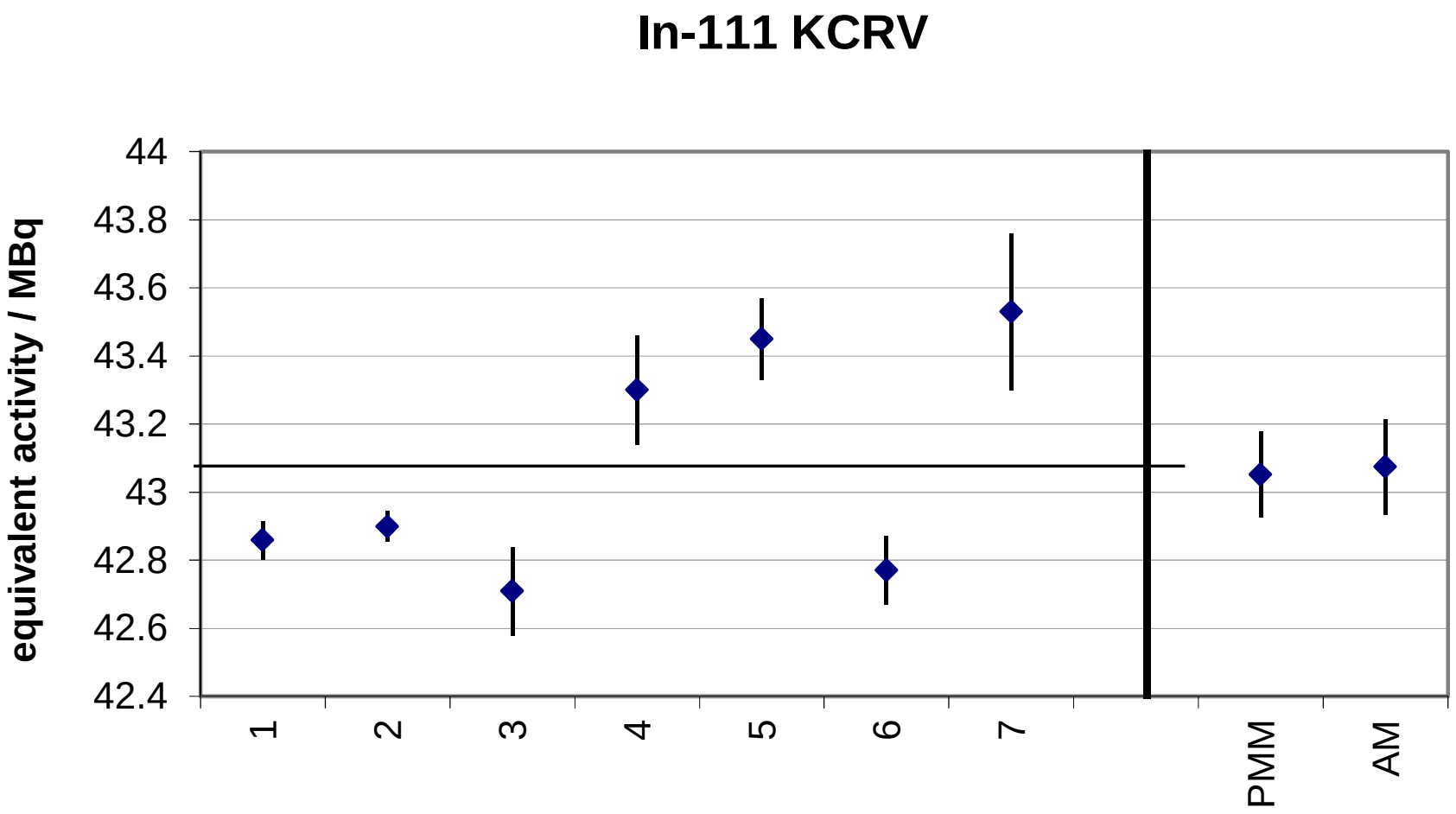
PMM gives realistic uncertainty for consistent data set



PMM is more efficient then AM



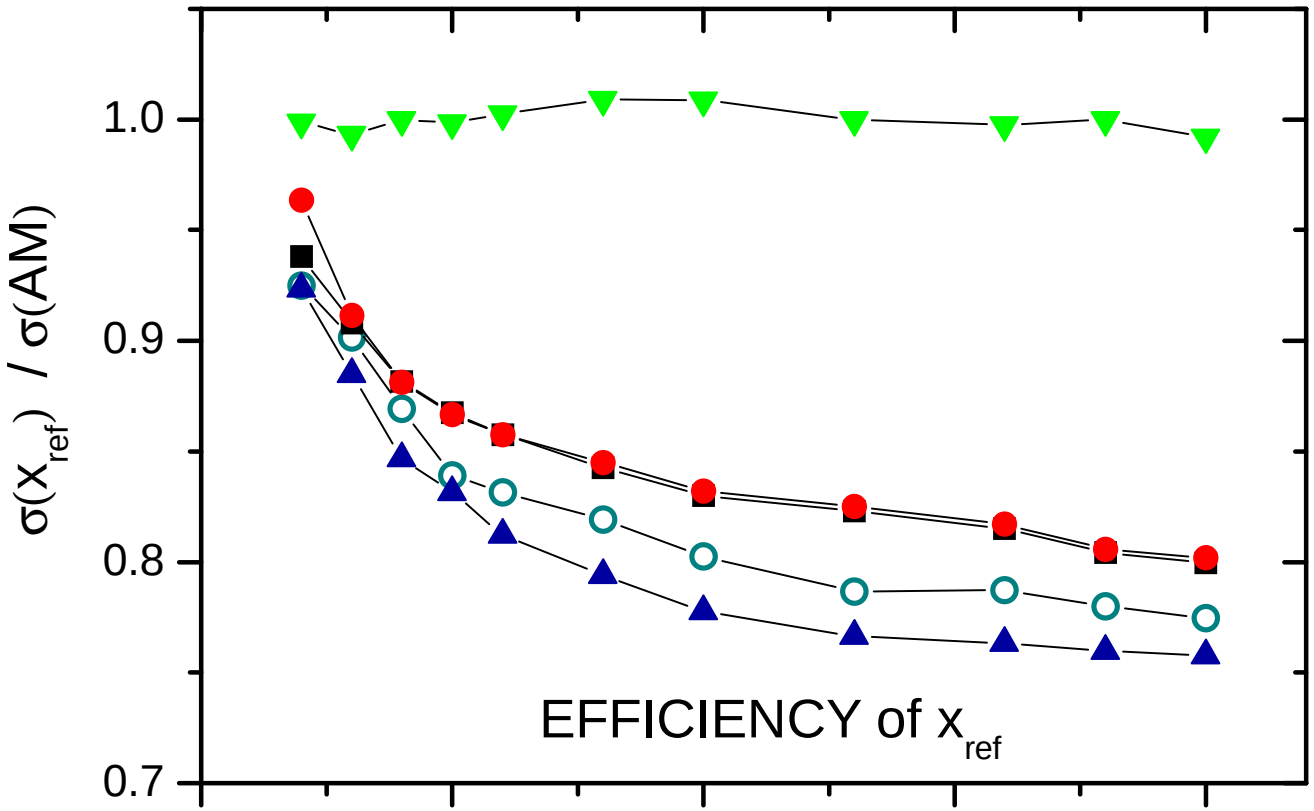
PMM close to arithmetic mean for discrepant data



Efficiency for discrepant data

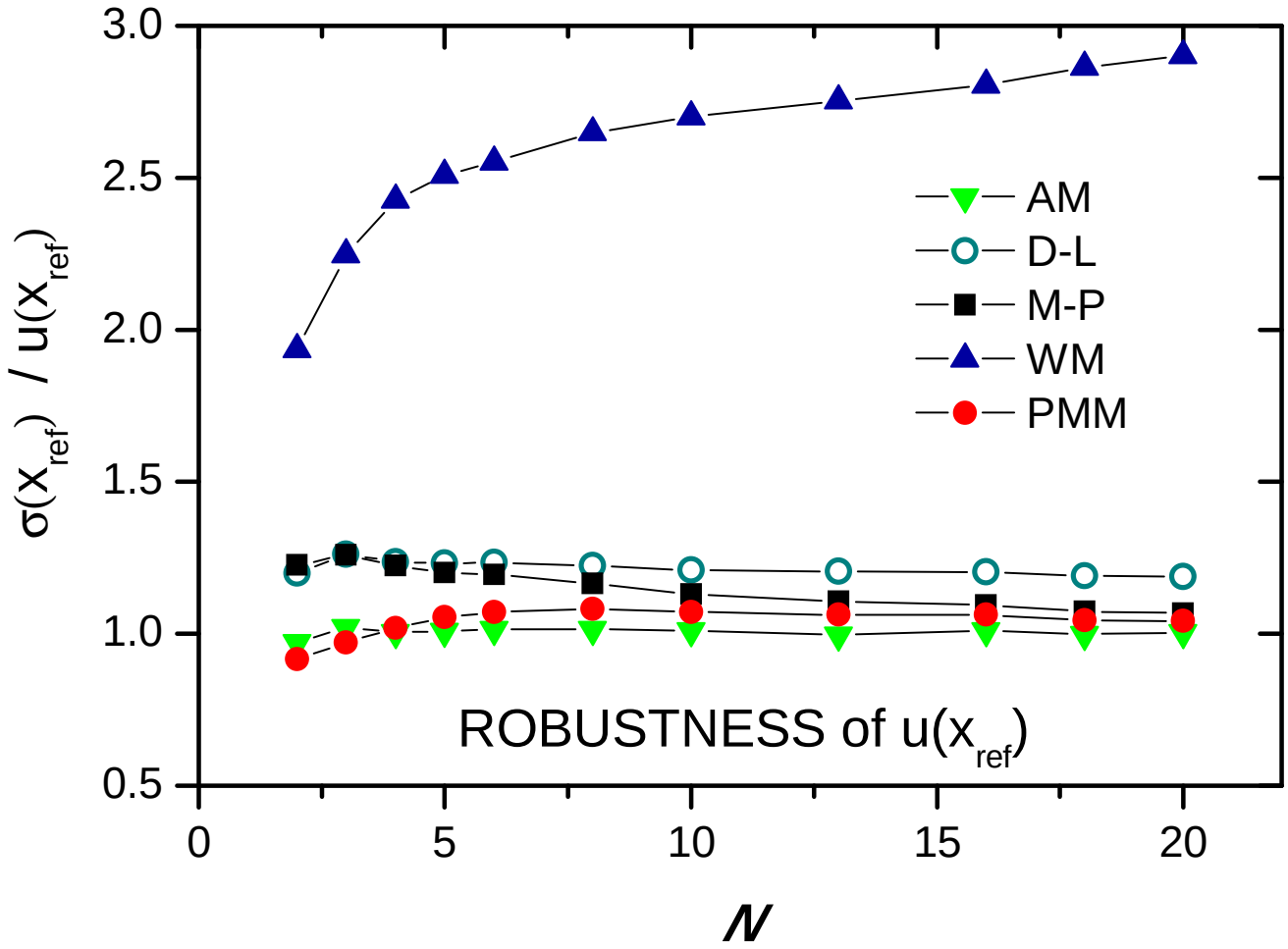
arithmetic < PMM, M-P < D-L < weighted

Too small but informative uncertainties u_i

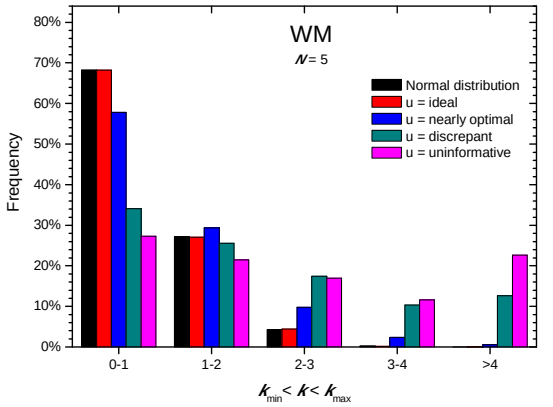


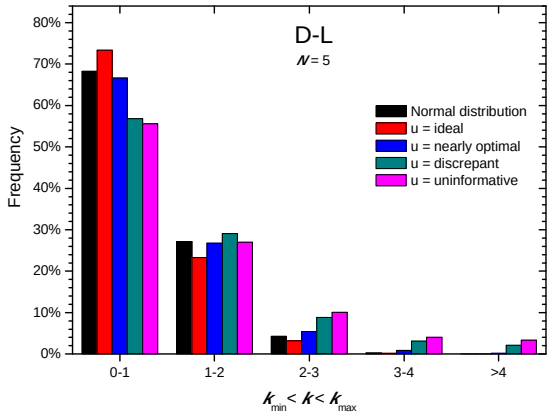
Robustness of uncertainty

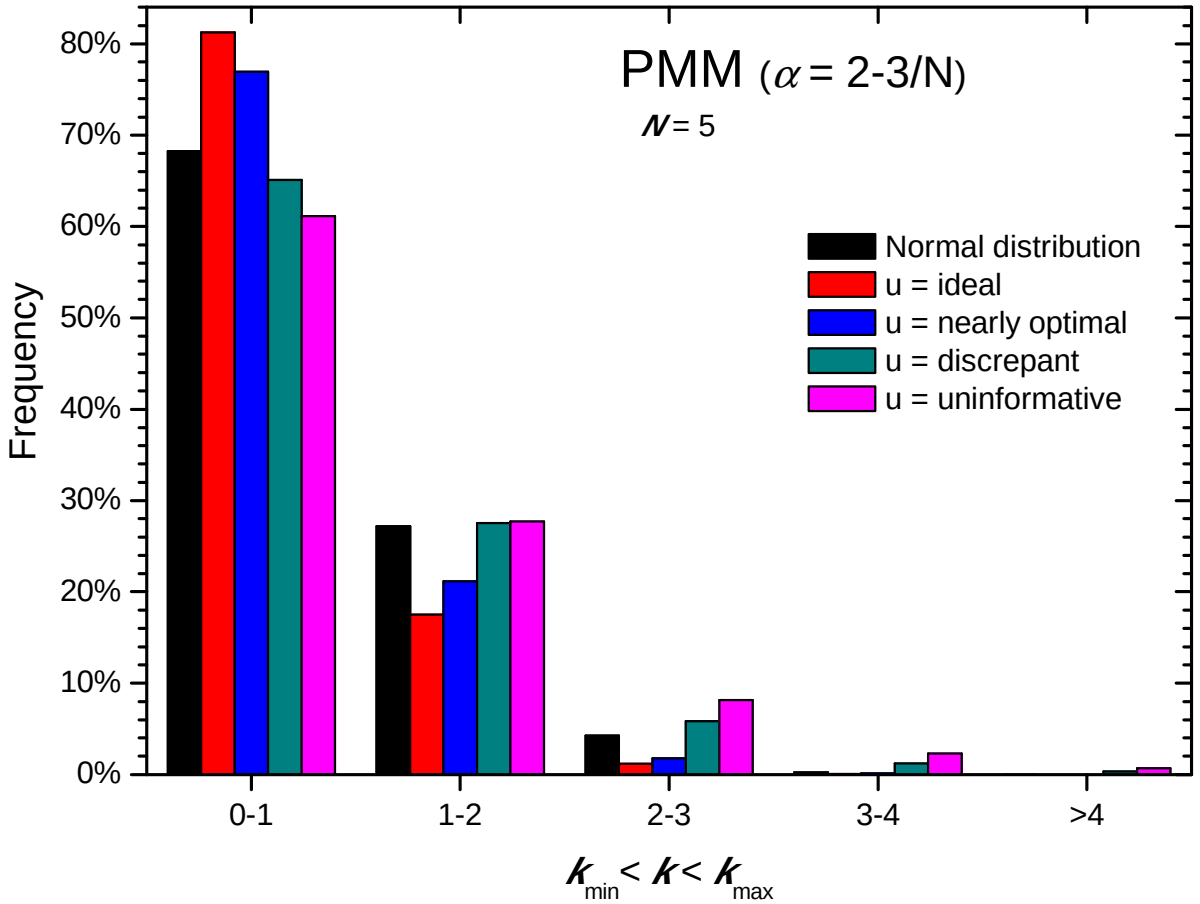
weighted << D-L < M-P < PMM < arithmetic



distribution around true value







Obtaining statistical control

Technical scrutiny: find unknown unknowns

Redundancy of methods: reveal systematic errors

Uncertainty monograph: pass on knowledge