Determination of KCRV, u(KCRV) & DoE

Final Proposal by KCWG to CCRI(II)

Stefaan Pommé

John Keightley

CCRI(II) Meeting, BIPM,
14-16 May 2013

Estimators for mean

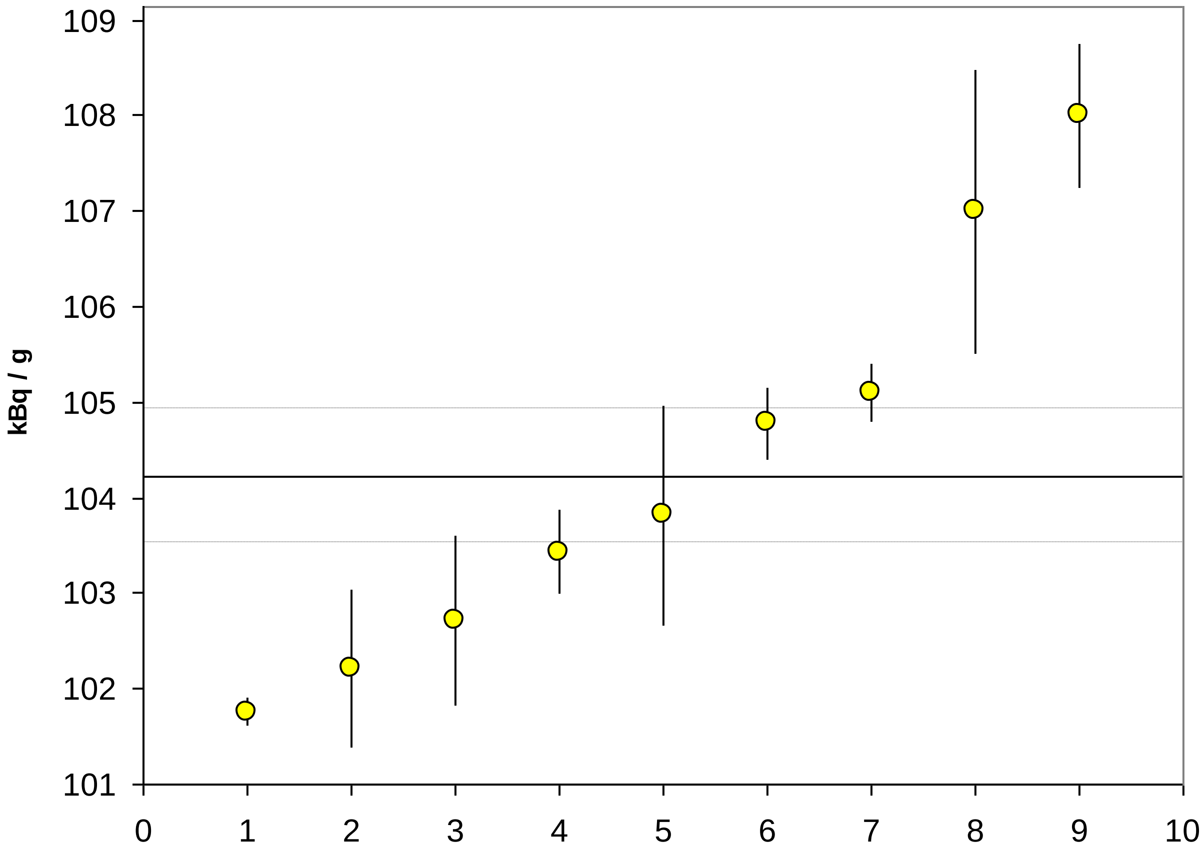
arithmetic
weighted
Mandel-Paule
Power Moderated Mean

Arithmetic mean

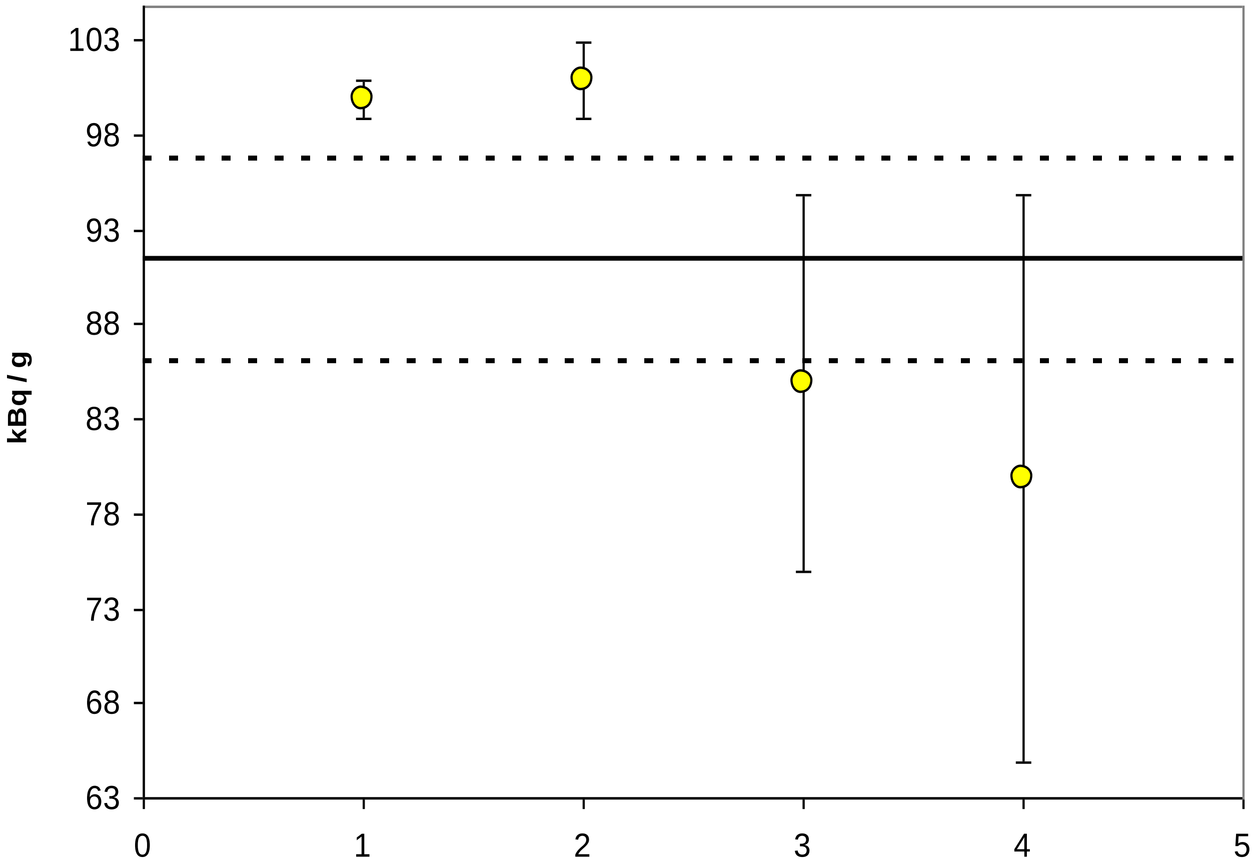
$$x_{\text{ref}} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$u^2(x_{\text{ref}}) = \sum_{i=1}^{N}(x_i - x_{\text{ref}})^2 / (N-1).$$

Good: ignore wrong unc



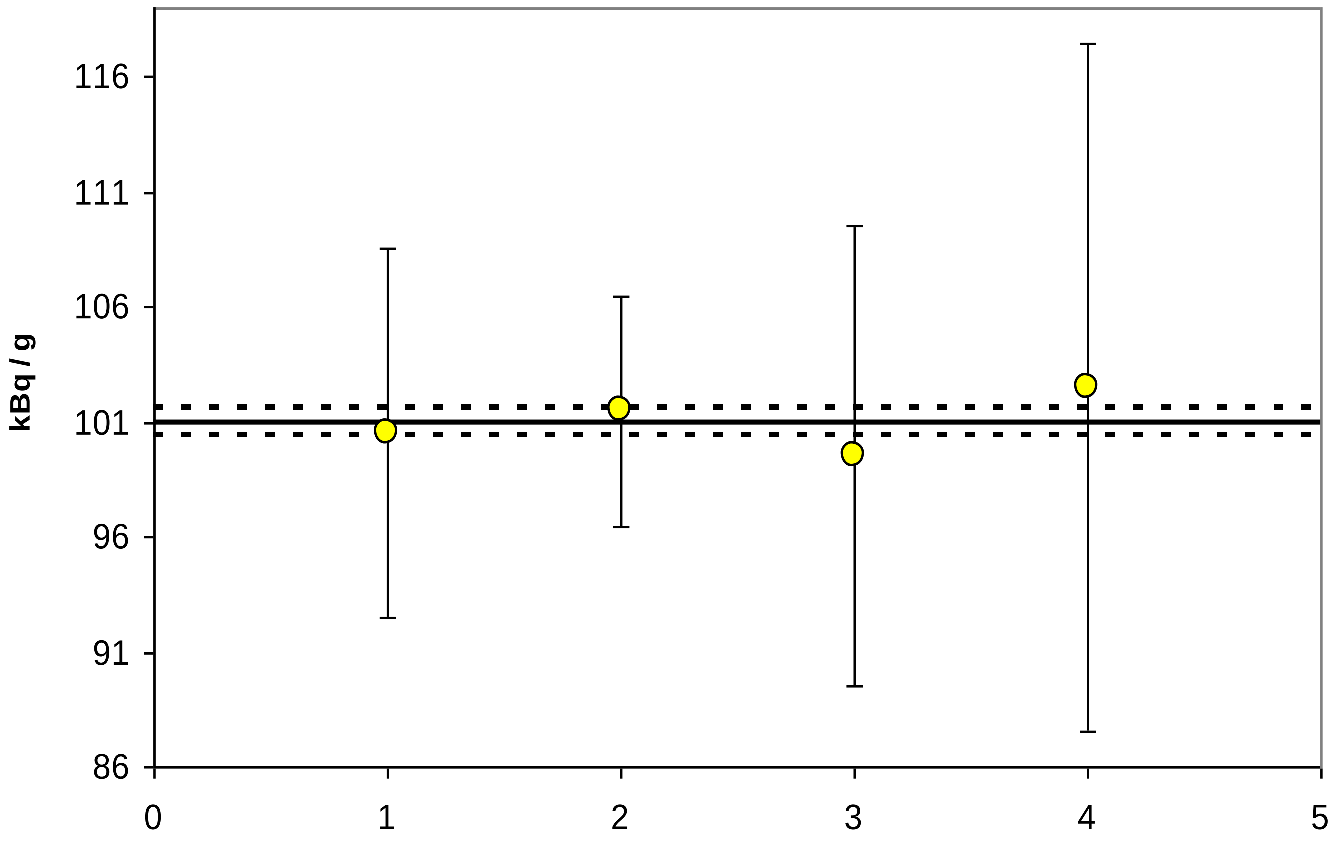
Bad: inefficient



biased mean

large uncertainty
= not "efficient"

Bad: low sample variance



low uncertainty

Remedy against low unc

Calculate uncertainty from maximum of

- propagated sum of stated uncertainties
- sample variance

$$u(\bar{x}) = \max\left(\frac{1}{N}\sqrt{\sum_{i=1}^{N} u_i^2}, \sqrt{\sum_{i=1}^{N} \frac{(x_i - \bar{x})^2}{N(N-1)}} \right)$$

Arithmetic mean

Best solution is close to *arithmetic mean* if

- measurement uncertainty contains no useful information
- magnitude error on uncertainty is >2x larger than magnitude due to metrological reasons

= the “best” with “bad uncertainty data”

= inefficient with “consistent data”

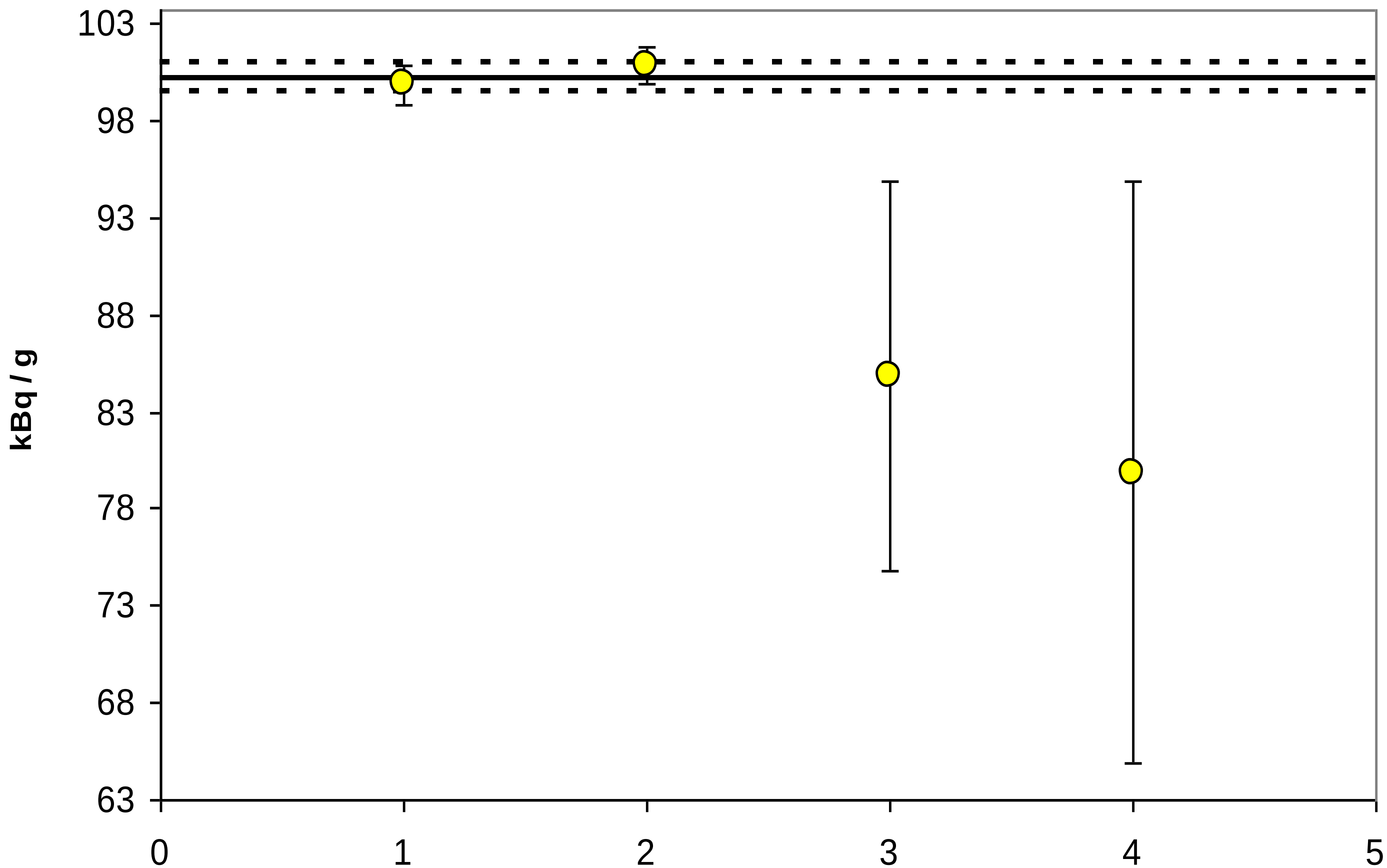
Weighted mean

$$x_{\text{ref}} = u^2(x_{\text{ref}}) \sum_{i=1}^{N} \frac{x_i}{u_i^2}$$

$$\frac{1}{u^2(x_{\text{ref}})} = \sum_{i=1}^{N} \frac{1}{u_i^2}$$

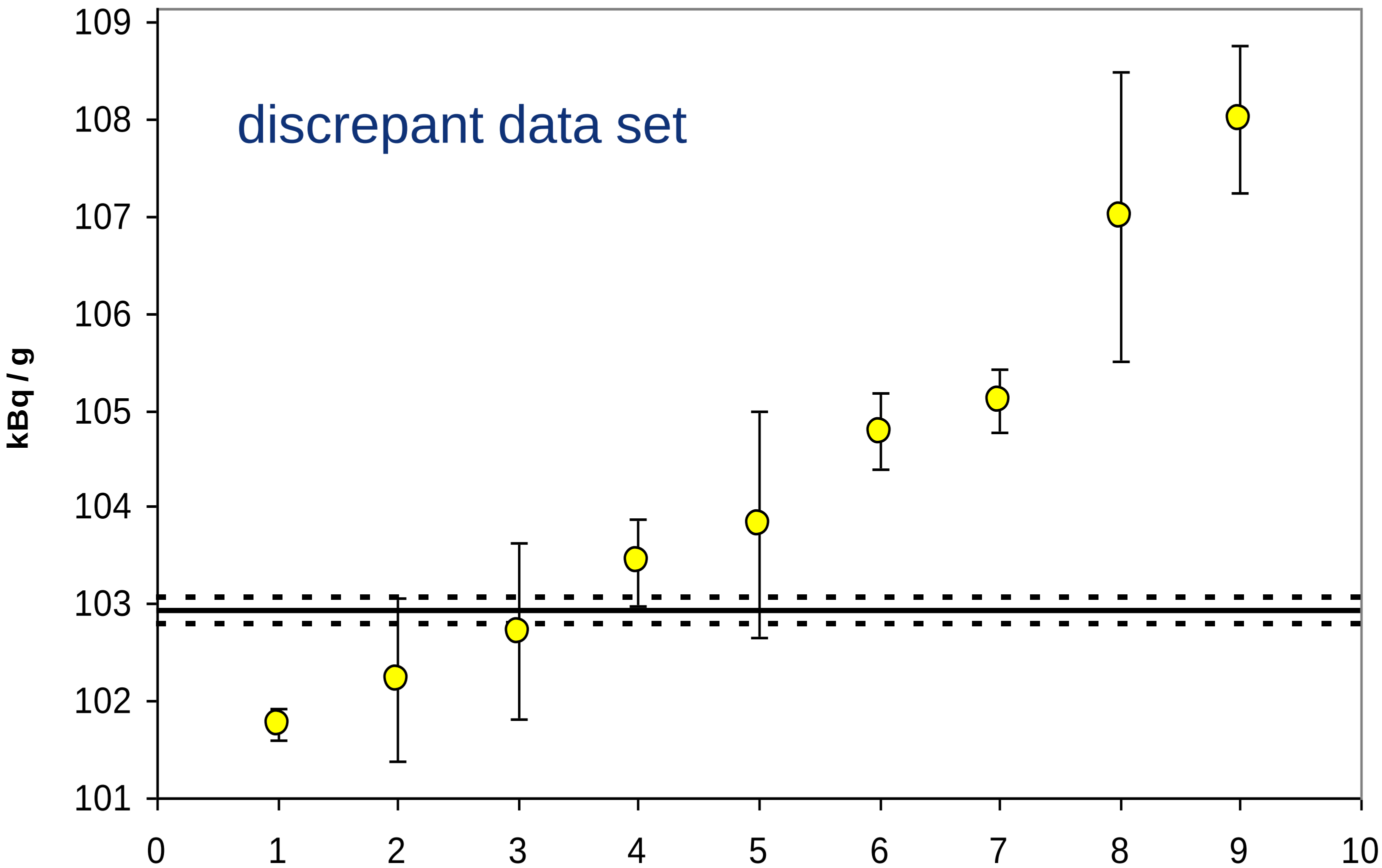
statistical weight = reciprocal variance associated with x_i

Good: efficient



efficient

Bad: sensitive to error



biased mean
low uncertainty

Weighted mean

European Commission

Best solution is close to *weighted mean* if

- measurement uncertainties are correct
- ‘value’ and ‘uncertainty’ outliers are excluded

= the “best” with “perfect data”

= sensitive to “low uncertainty” outliers

Mandel-Paule mean

$$x_{\text{ref}} = u^2(x_{\text{ref}}) \sum_{i=1}^{N} \frac{x_i}{u_i^2 + s^2},$$

$$\frac{1}{u^2(x_{\text{ref}})} = \sum_{i=1}^{N} \frac{1}{u_i^2 + s^2},$$

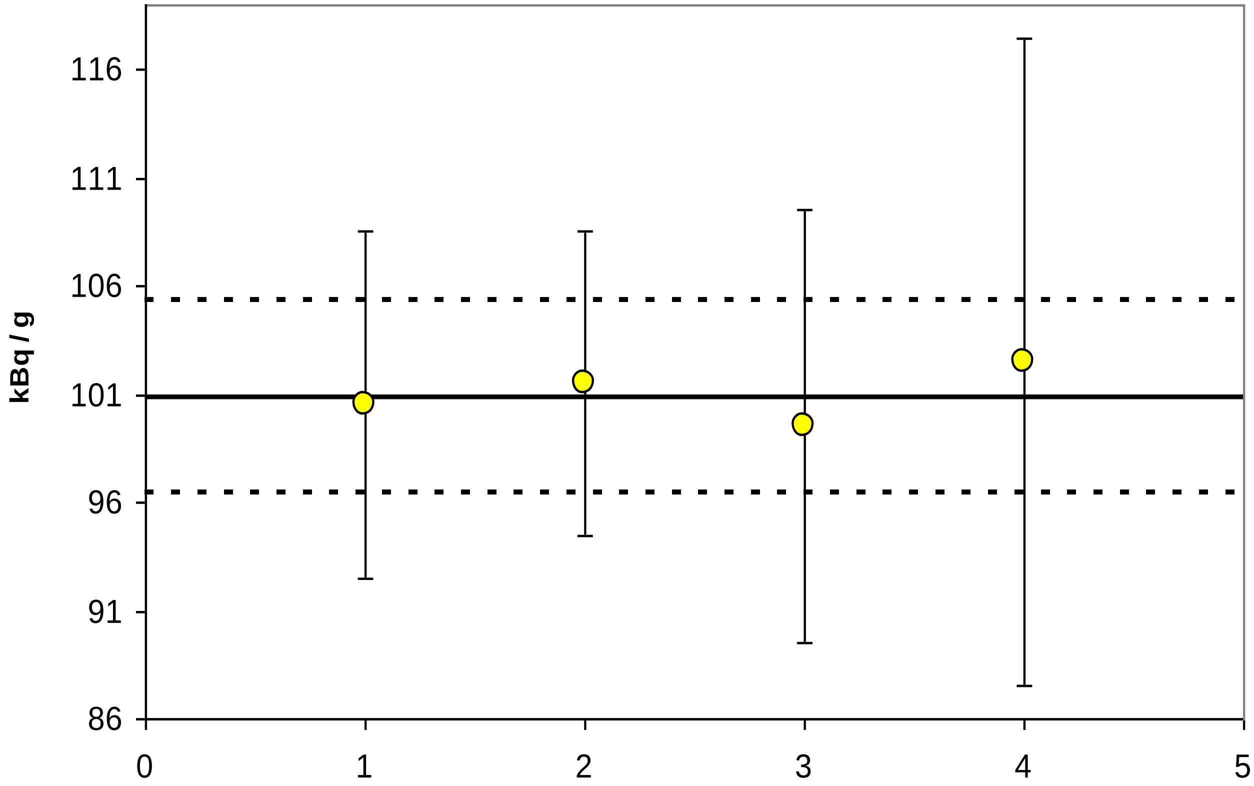
s^2 is artificially added “interlaboratory” variance to make reduced chi = 1

$$\tilde{\chi}_{\text{obs}} = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} \frac{(x_i - x_{\text{ref}})^2}{u_i^2}}.$$

Mandel-Paule mean

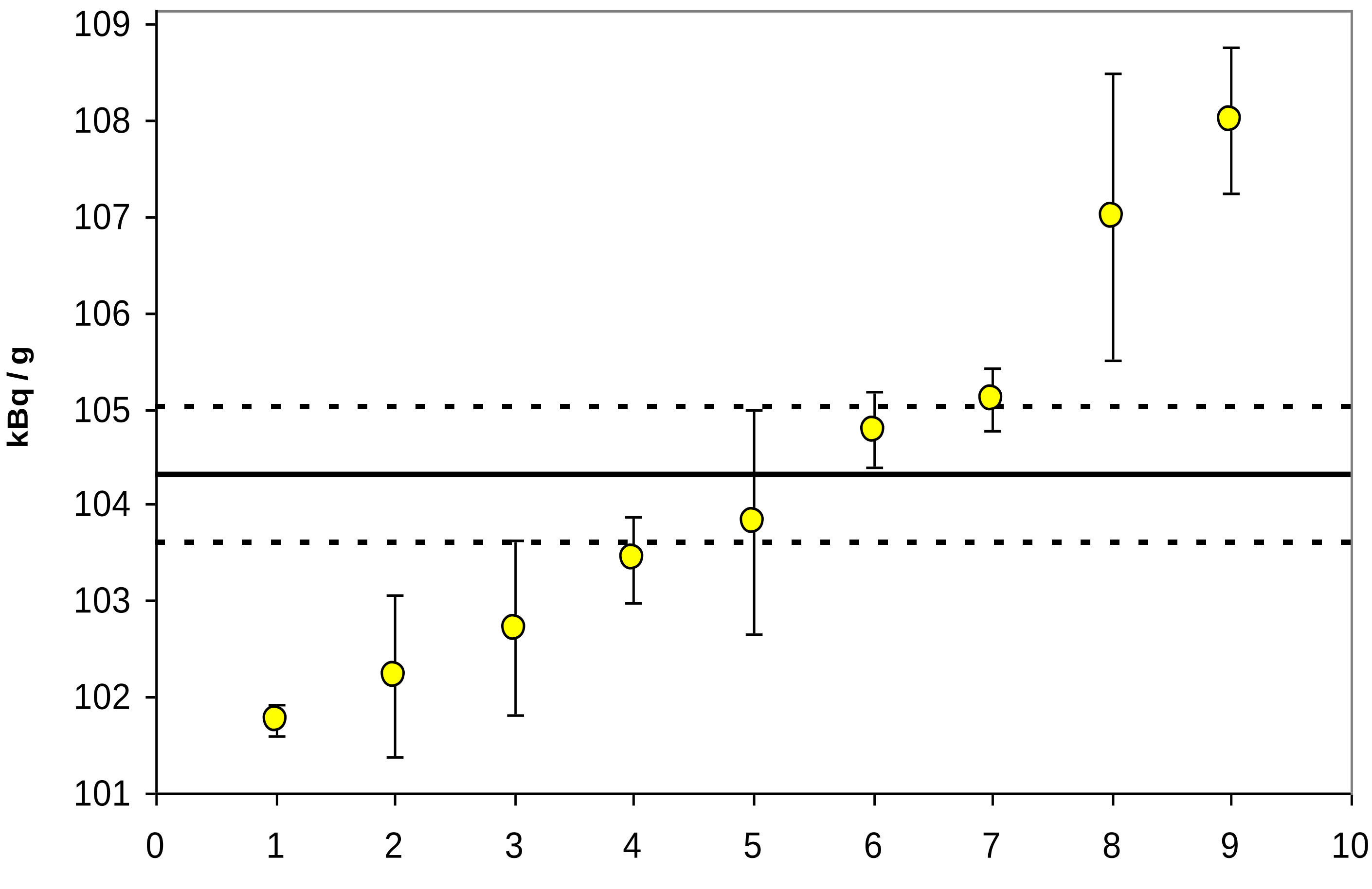
European Commission

= weighted mean for a consistent data set



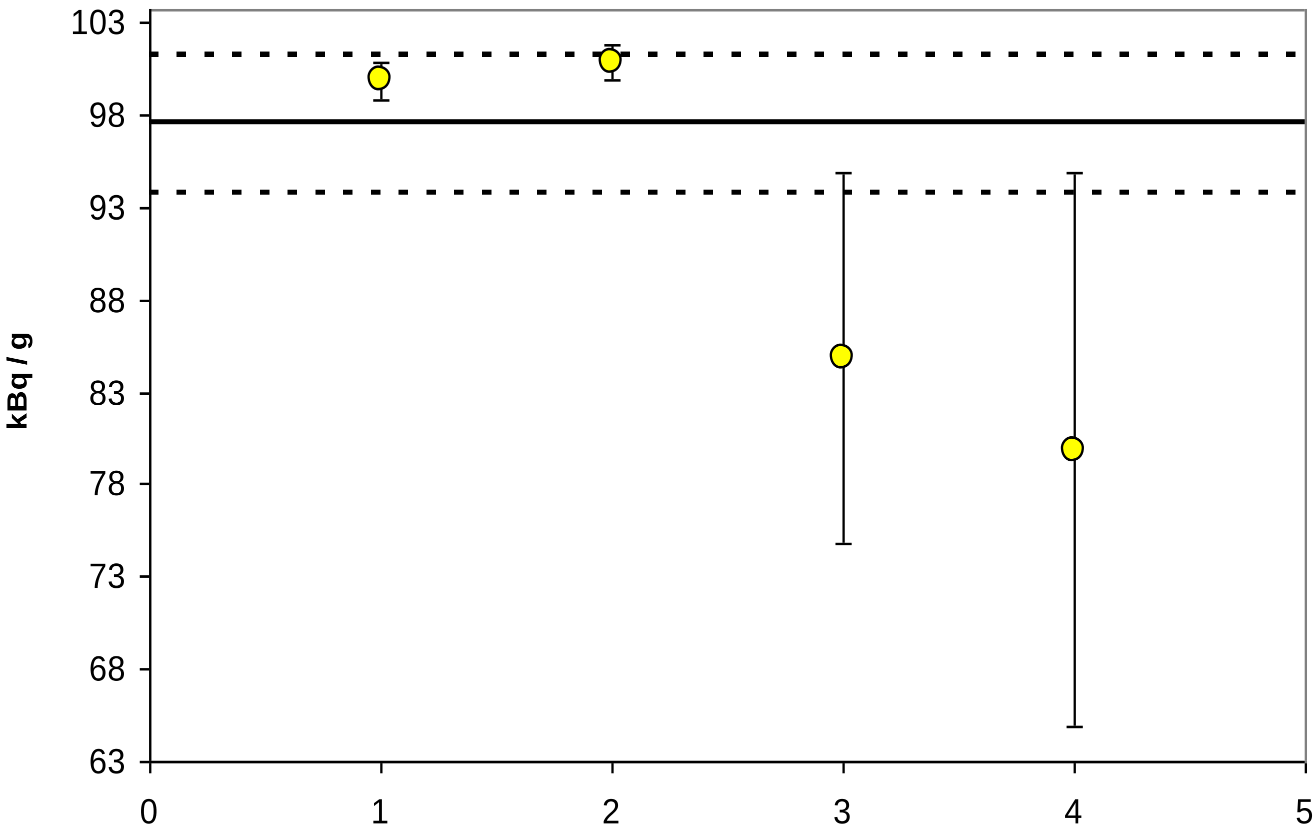
Mandel-Paule mean

≈ arithmetic mean for an extremely discrepant data set



Mandel-Paule mean

= intermediate between arithmetic and weighted mean for a slightly discrepant data set



Mandel-Paule mean

Best solution is close to Mandel-Paule mean if

- measurement uncertainties are informative
- ‘value’ and ‘uncertainty’ outliers are symmetric

= one of the “best” with “imperfect data”
if no tendency to underestimate uncertainty

Power Moderated Mean - PMM

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

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

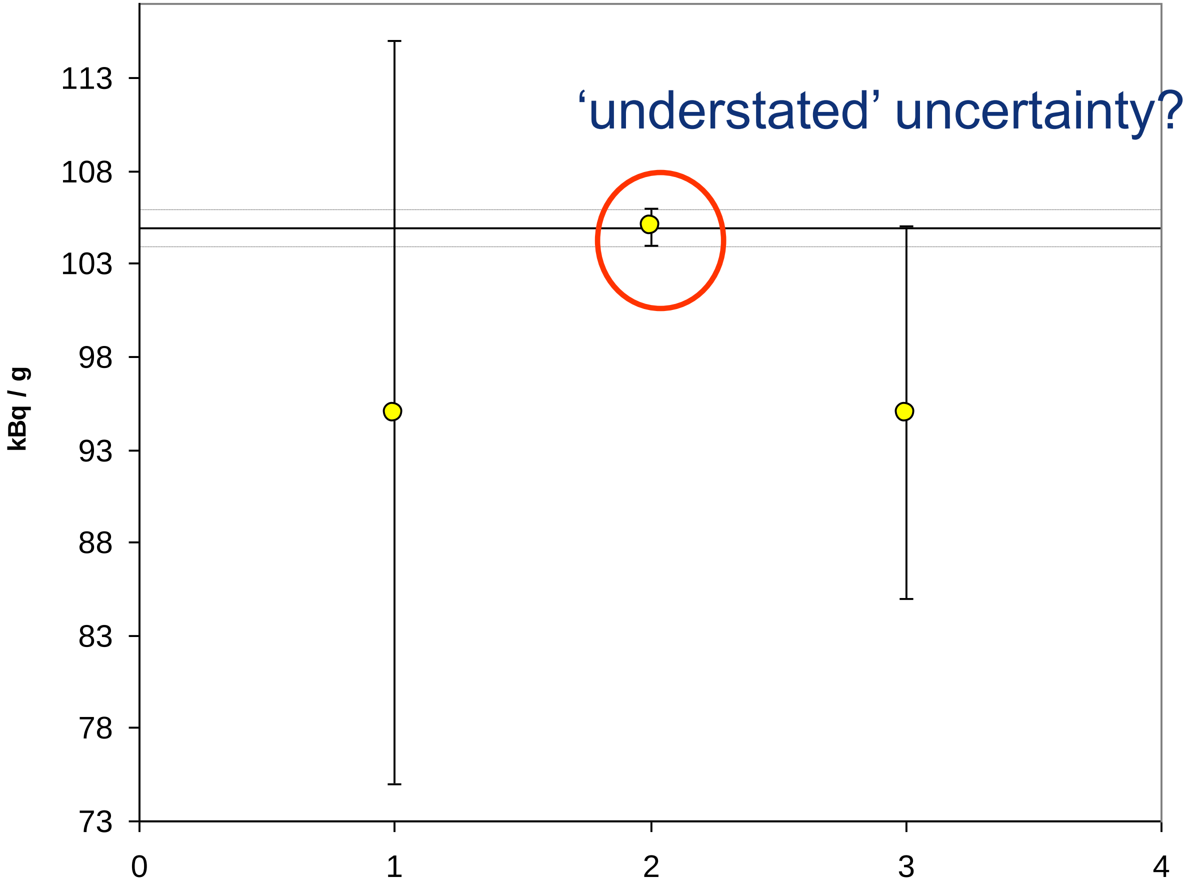
$$S = \sqrt{N \cdot \max\left(u^2(\bar{x}), u^2(x_{\mathrm{mp}})\right)}$$

S is a typical uncertainty per datum (max arithmetic or M-P unc)

$0<\alpha<2$ = power reflects level of trust in uncertainties

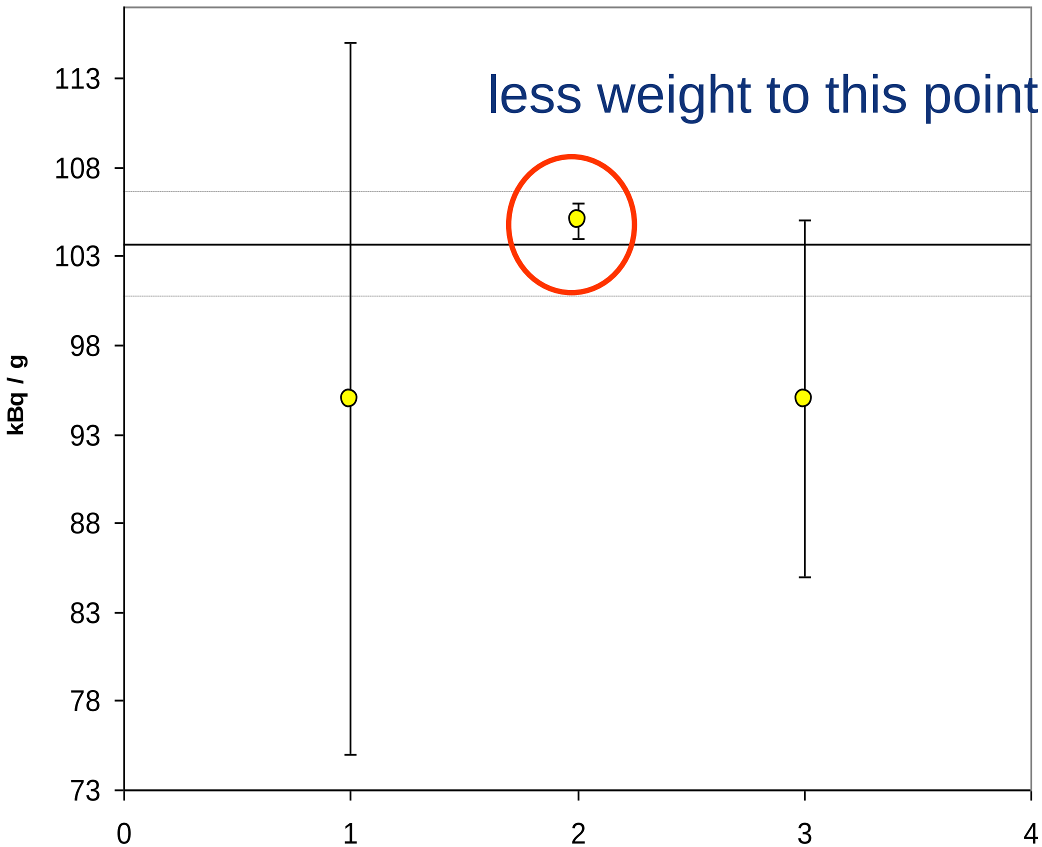
PMM: α=2

= Mandel-Paule mean



PMM: $\alpha=1$

= closer to arithmetic mean



Choice of power α

power	reliability of uncertainties
$\alpha = 0$	uninformative uncertainties (arithmetic mean)
$\alpha = 0$	uncertainty variation due to error at least twice the variation due to metrological reasons (arithmetic mean)
$\alpha = 2\text{-}3/N$	informative uncertainties with a tendency of being rather underestimated than overestimated (intermediately weighted mean)
$\alpha = 2$	informative uncertainties with a modest error; no specific trend of underestimation (Mandel-Paule mean)
$\alpha = 2$	accurately known uncertainties, consistent data (weighted mean)

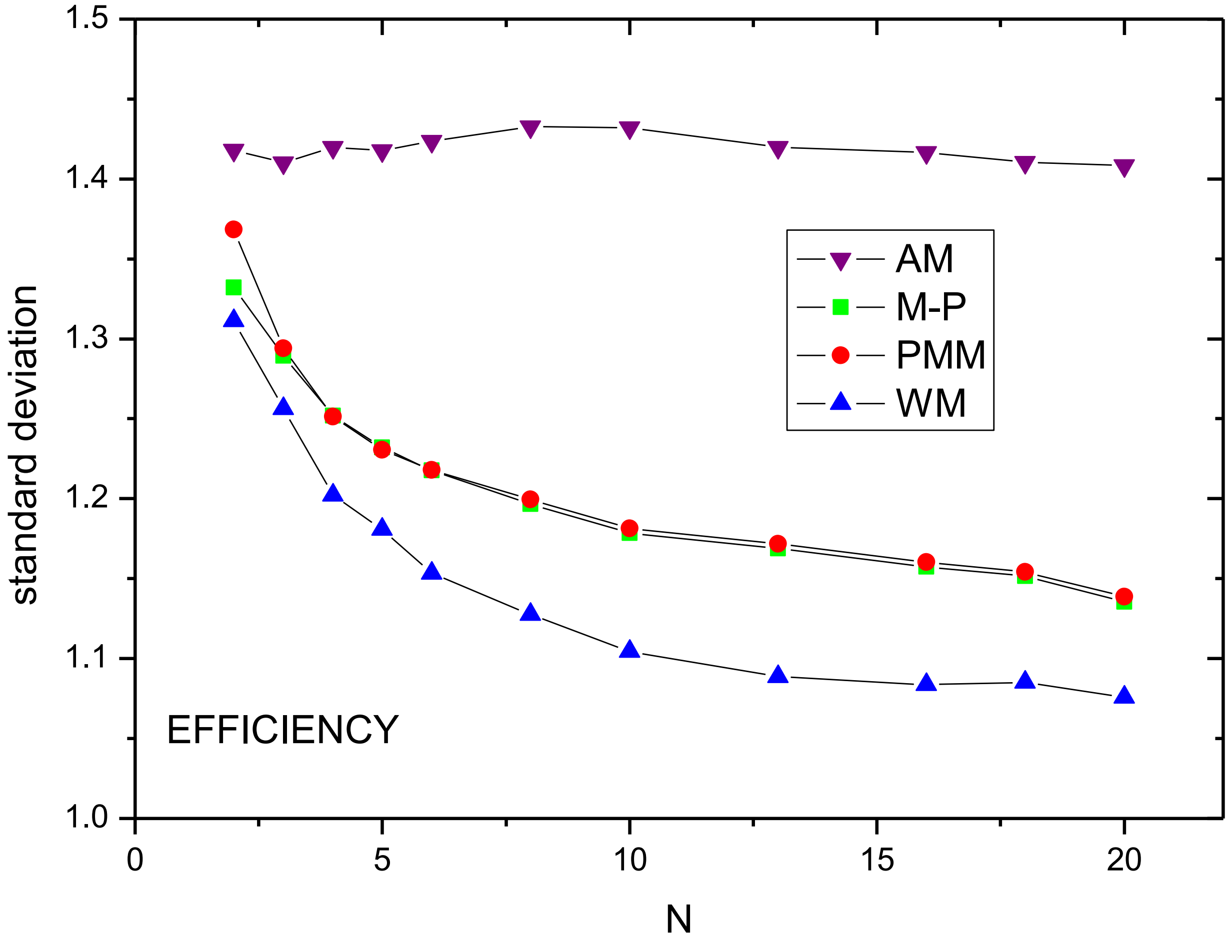
Test of Estimators by Computer Simulation

arithmetic
weighted
Mandel-Paule
Power Moderated Mean

Efficiency for discrepant data

European Commission

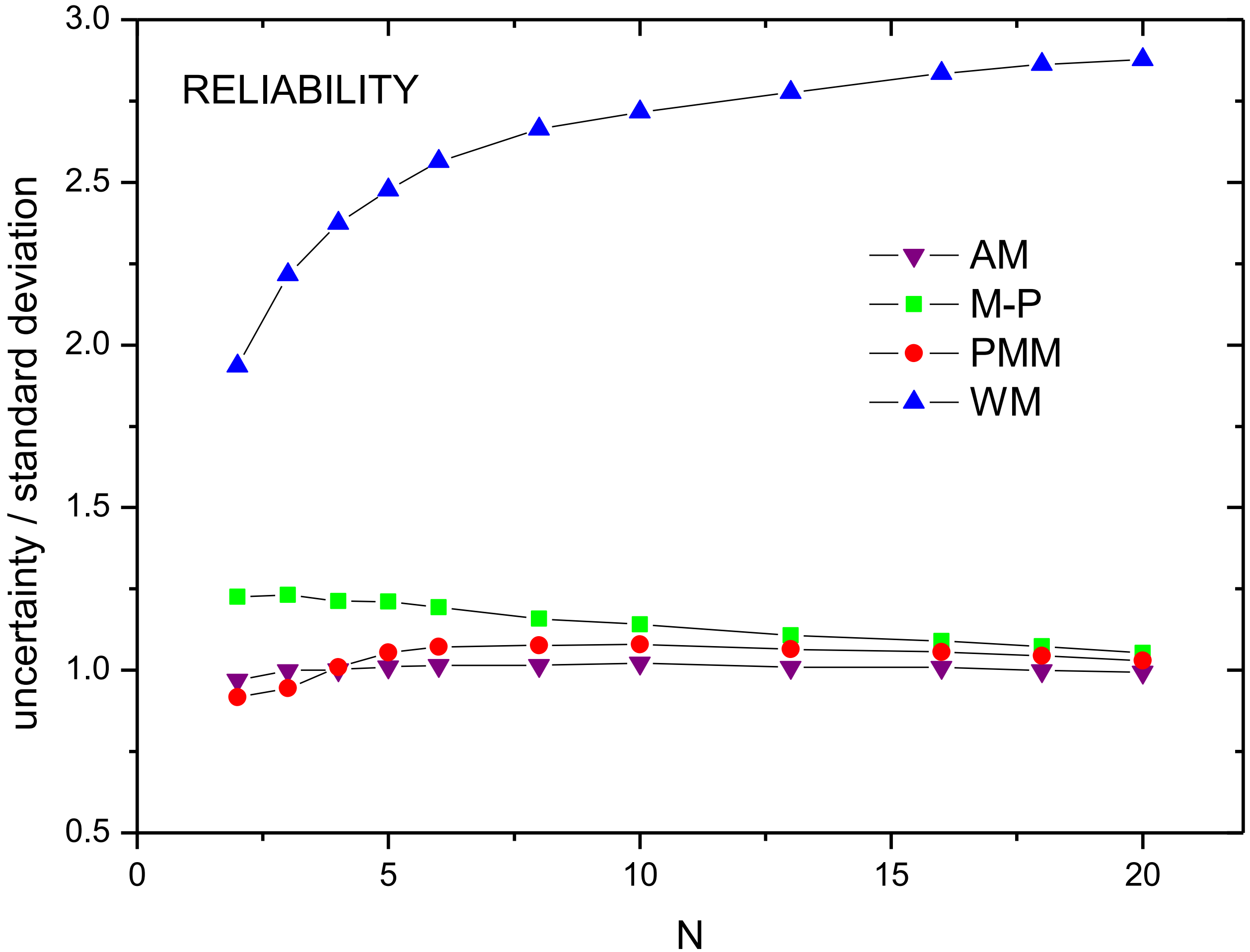
arithmetic < PMM, M-P < weighted



Reliability of uncertainty

European Commission

weighted << M-P < PMM < arithmetic



Power Moderated Mean

Best solution is close to PMM if

- measurement uncertainties are informative
- uncertainties tend to be understated
- data seem consistent but are not

= one of the “very best” with “imperfect data”

= more realistic uncertainty than Mandel-Paule mean

= more adjustable to quality of data than M-P mean

Outlier identification

generally applicable method

Outlier identification

European Commission

CCRI(II) is the final arbiter regarding correcting or excluding any data from the calculation of the KCRV.

Statistical tools may be used to indicate data that are extreme.

= a way to protect the KCRV against erroneous data, data with understated uncertainty, extreme data asymmetrically disposed to the KRCV

= a way to lower the uncertainty on the KCRV

Outlier identification

$$|e_i| > ku(e_i), \quad e_i = x_i - x_{\text{ref}}$$

$$u^2(e_i) = u^2(x_{\text{ref}})(\frac{1}{w_i} - 1)$$

x_i included in mean

$$u^2(e_i) = u^2(x_{\text{ref}})(\frac{1}{w_i} + 1)$$

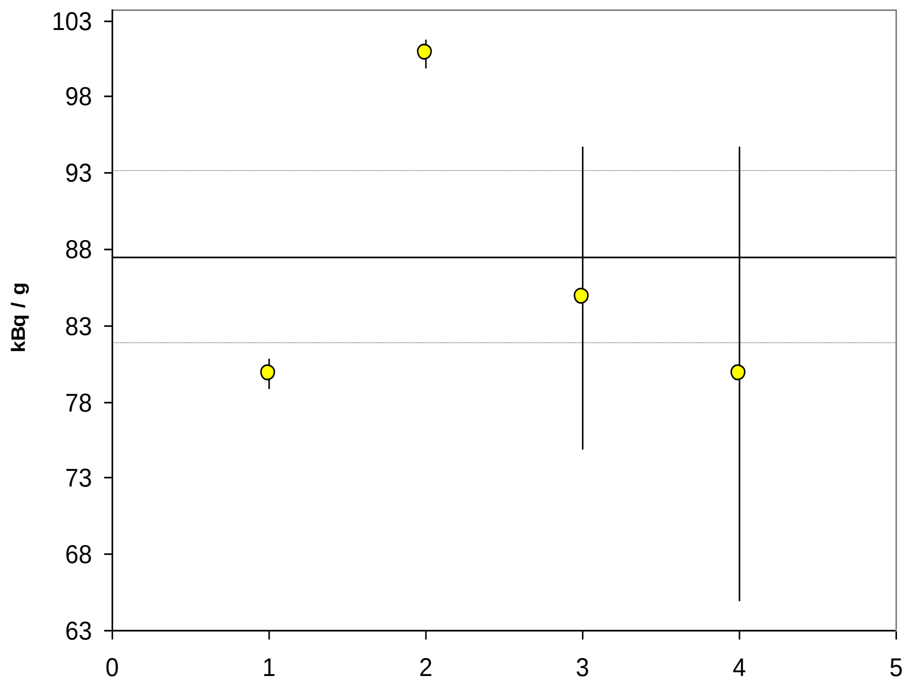
x_i excluded from mean

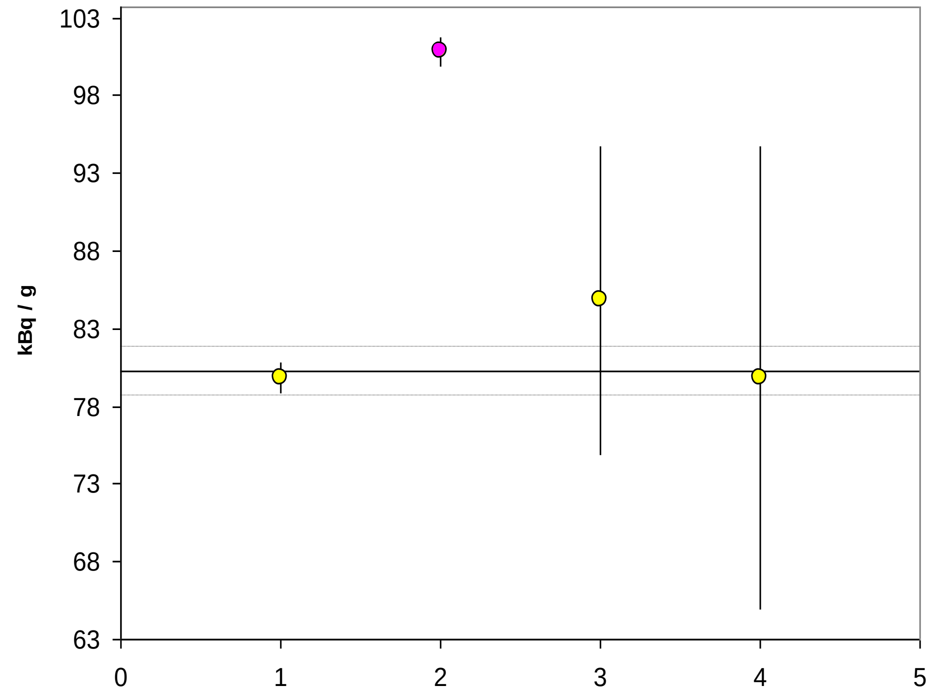
- valid for any type of mean, using normalised w_i
- default $k = 2.5$

Outlier or not?

Both options are possible within the method.

→ outlier rejection should be based on technical grounds





Degree of equivalence

generally applicable method

Degree of equivalence

$$d_i = x_i - x_{\mathrm{ref}}, \quad U(d_i) = 2u(d_i)$$

$$u^2(d_i) = (1-2w_i)u_i^2 + u^2(x_{\mathrm{ref}})$$ x_{i} included in mean

$$u^2(d_i) = u_i^2 + u^2(x_{\mathrm{ref}})$$ x_{i} excluded from mean

- valid for any type of mean

Conclusions

The **Power Moderated Mean** keeps a fine balance between efficiency and robustness, while providing also a reliable uncertainty.

Outlier identification and degrees of equivalence are readily obtained.