CIRMS 15th Annual Meeting

23 October 2006

# Measurement Uncertainty Assessment & Reports:

# Practical Considerations

R. Collé

Radioactivity Group
NIST Physics Laboratory

# Objective / Goal of this Talk:

Nothing new –

no methods; no techniques; no propaganda

Just some insight –

things to consider & think about

Practical considerations

# Practical consideration : want realistic uncertainties

(like the 3 bears: not too big, not too small, "just right"

Constraint (whether good, bad or indifferent)

Ground rules (boundary conditions) :

Must follow Internationally-accepted (ISO, et alia) GUM + NIST guidance

Guidance based on 5 simple **recommendations** of a 1980 BIPM Working Group and CIPM .....

which led to an 8-step **procedure** in the "Guide"

Assume you know it

Here's a short version

1. Express math relation between input $X_i$ & output $Y$ *(i.e., MODEL)*
2. Get values of $X_i$
3. Evaluate "standard uncertainty" $u(X_i)$ on each by type A or type B evaluations – square of these assumed to act like "variances" $u(X_i)^2$
4. Evaluate any "covariances" $u(X_i,X_j)$ associated with $X$ values -- *Oops, a small problem here when reporting)*
5. Calculate $Y$ from the math function
6. Determine the "combined standard uncertainty" $u_c(Y)$ – assumed to be positive square root of total "variance" obtained by summing all variance $u(X_i)^2$ and covariance $u(X_i,X_j)$ components using "Propagation of Error Law"
7. If need higher level of confidence, multiply $u_c(Y)$ by coverage factor $k$ to get $U$ – *whether you know what it means or not*
8. Report Y and U or $u_c(Y)$ -- AND list all uncertainty components & how evaluated

# One underlying Principle of the Guidance (often violated):

Thou shalt not take “STATISTICS” in vain

# Most frequently misunderstood concept

Type A

&

Type B

Who believes that a good job on uncertainties is being done?

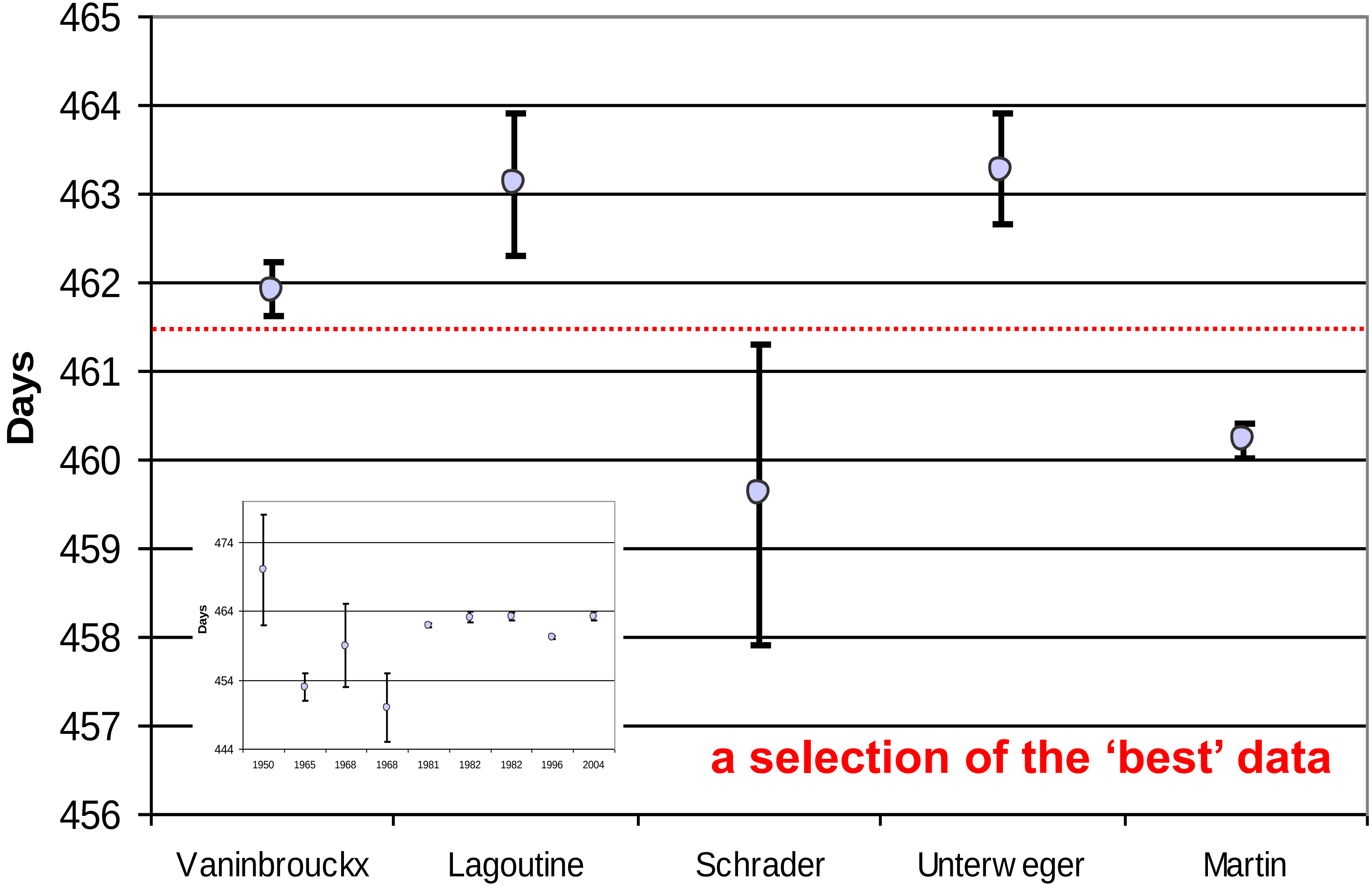
“Just right”

Not too big

Not too small

# Half-life of $^{109}Cd$

Thanks to Pommé for data; *Appl. Radiat. Isot*. **64**, 1412 (2006)



# WHY ?

It’s the Model,

Stu ! ! d

# What is a *Measurement Model*?

*Given as :*

*"Formalization of the relationship between input and output quantities involved in a measurement and how the uncertainty in an input quantity leads to an uncertainty in the output quantity"*

Doesn't help much for making realistic uncertainty assessment

$$A = \frac{\left(\dfrac{C}{t} - B\right)}{m\varepsilon \exp\left[\dfrac{-\ln(2)T}{T_{1/2}}\right]}$$

$A =$ massic activity

$C$ = number counts

$t$ = measurement livetime

$B$ = background count rate

$m$ = source mass

$\varepsilon =$ detection efficiency

$T$ = reference time interval

$T_{1/2} =$ half-life

Uncertainties $\xrightarrow{\text{ESTIMATE}}$ Errors

(components of inaccuracy)

IF measurement THEN model

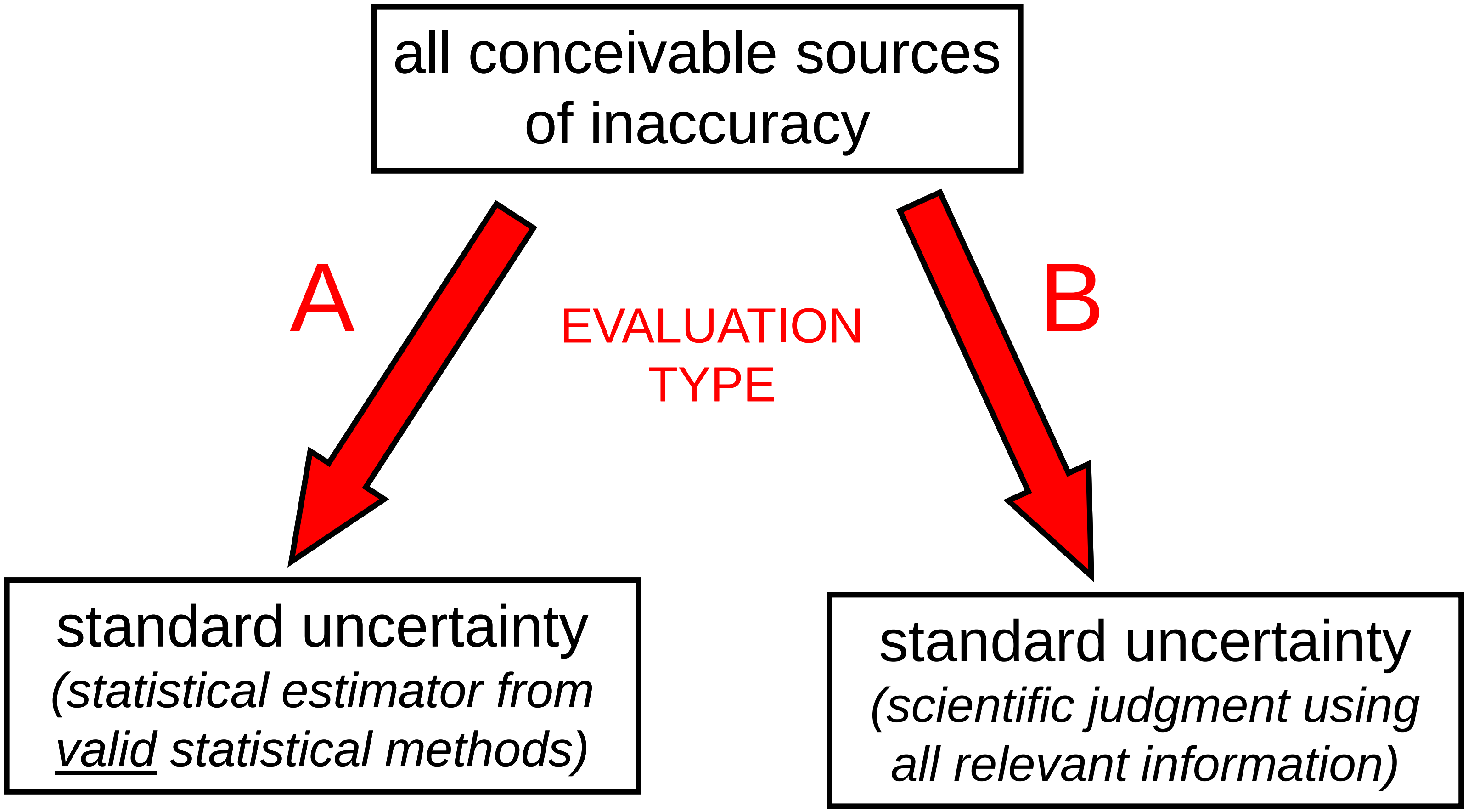
The model is often exactly what you did –
knowingly or unknowingly

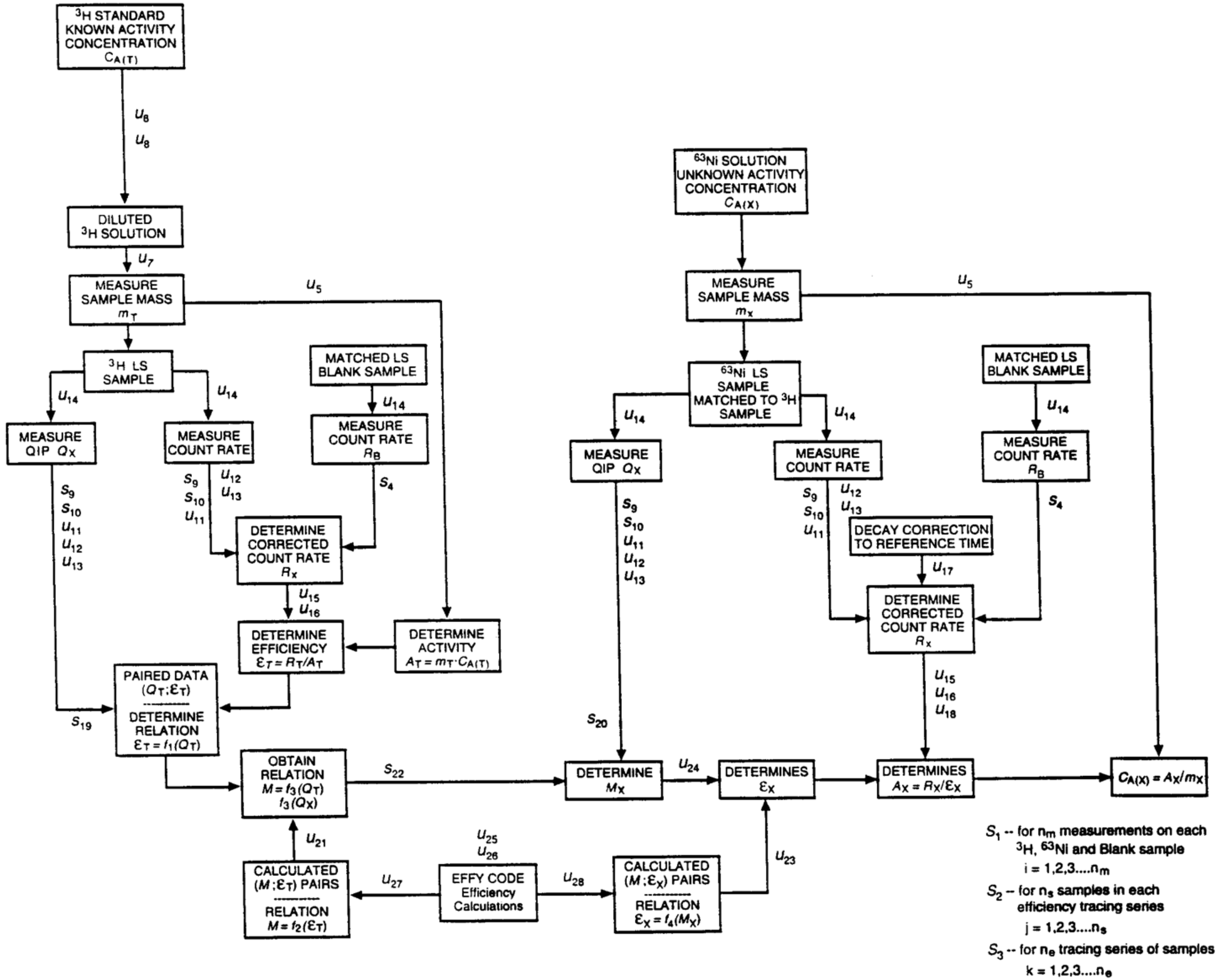
Look at experiment

Look at data

Think

Look at Experiment (measurement method) -- THINK





3H STANDARD KNOWN ACTIVITY CONCENTRATION CA(T)
u8
u8
DILUTED 3H SOLUTION
u7
MEASURE SAMPLE MASS mT
u5
3H LS SAMPLE
u14
u14
MEASURE QIP QX
MEASURE COUNT RATE
MATCHED LS BLANK SAMPLE
u14
MEASURE COUNT RATE RB
s9 s10 u11 u12 u13
s9 s10 u11
u12 u13
s4
DETERMINE CORRECTED COUNT RATE RX
u15 u16
DETERMINE EFFICIENCY εT = RT/AT
DETERMINE ACTIVITY AT = mT·CA(T)
PAIRED DATA (QT;εT)
DETERMINE RELATION εT = f1(QT)
s19
OBTAIN RELATION M = f3(QT) f3(QX)
s22
u21
CALCULATED (M;εT) PAIRS
RELATION M = f2(εT)
u27
u25 u26
EFFY CODE Efficiency Calculations
u28
CALCULATED (M;εX) PAIRS
RELATION εX = f4(MX)
u23
63Ni SOLUTION UNKNOWN ACTIVITY CONCENTRATION CA(X)
MEASURE SAMPLE MASS mx
u5
63Ni LS SAMPLE MATCHED TO 3H SAMPLE
u14
u14
MEASURE QIP QX
MEASURE COUNT RATE
MATCHED LS BLANK SAMPLE
u14
MEASURE COUNT RATE RB
s9 s10 u11 u12 u13
s9 s10 u11
u12 u13
s4
DECAY CORRECTION TO REFERENCE TIME
u17
DETERMINE CORRECTED COUNT RATE RX
u15 u16 u18
s20
DETERMINE MX
u24
DETERMINES εX
DETERMINES AX = RX/εX
CA(X) = AX/mX
S1 -- for nm measurements on each 3H, 63Ni and Blank sample
i = 1,2,3....nm
S2 -- for ns samples in each efficiency tracing series
j = 1,2,3....ns
S3 -- for ne tracing series of samples
k = 1,2,3....ne

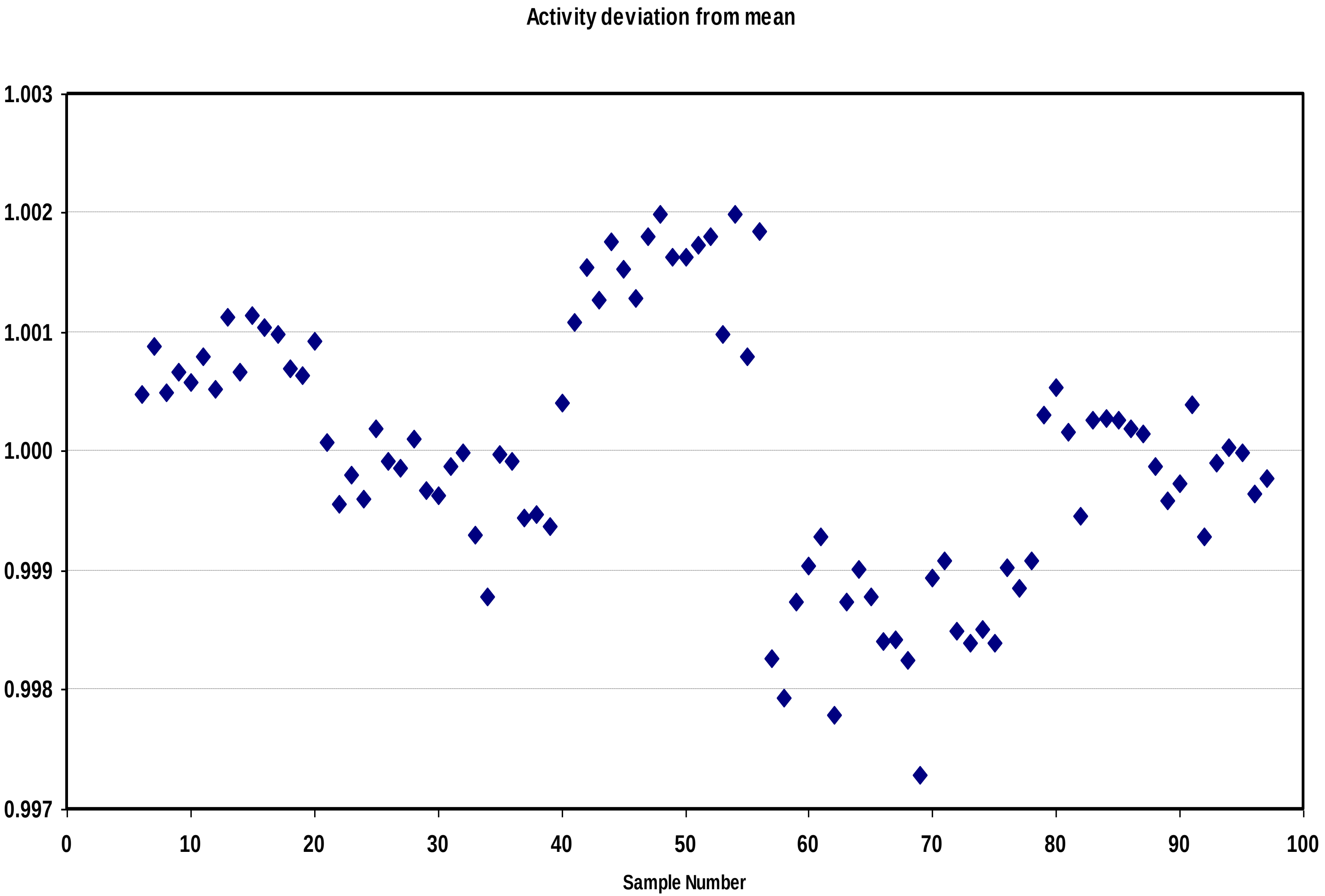
## Look at Data -- THINK

How measurement is performed influences
data treatment

How measurement is performed influences
what effects might be seen

How measurement is performed influences
what uncertainties might be considered &
assessed

Oh yes, might think about doing it backward in designing experiment
to get realistic uncertainty

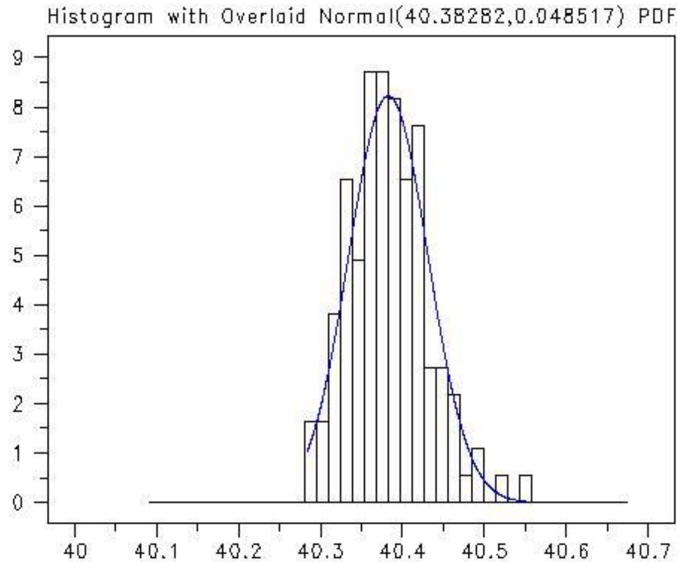


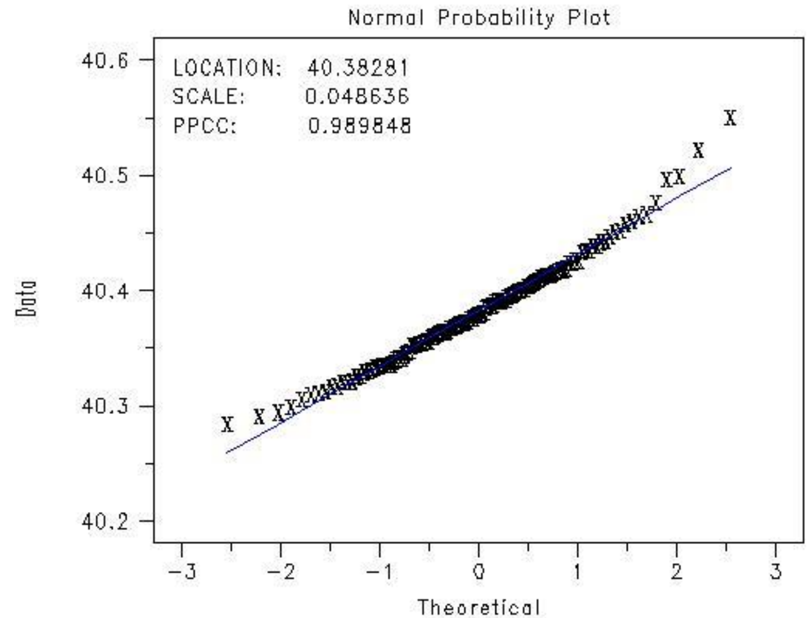

Obviously something going on here

Really nice data here

| Sample | 55 | 60 | 68 | 71 | 78 | 82 | 96 |
|---|---|---|---|---|---|---|---|
| | 40.320 | 40.360 | 40.362 | 40.379 | 40.418 | 40.367 | 40.359 |
| | 40.371 | 40.365 | 40.404 | 40.317 | 40.419 | 40.433 | 40.284 |
| | 40.310 | 40.335 | 40.326 | 40.363 | 40.373 | 40.385 | 40.391 |
| | 40.341 | 40.329 | 40.390 | 40.367 | 40.392 | 40.372 | 40.316 |
| | 40.409 | 40.413 | 40.321 | 40.375 | 40.341 | 40.522 | 40.464 |
| | 40.457 | 40.424 | 40.401 | 40.311 | 40.338 | 40.418 | 40.396 |
| | 40.333 | 40.435 | 40.364 | 40.331 | 40.397 | 40.400 | 40.412 |
| | 40.417 | 40.344 | 40.376 | 40.414 | 40.312 | 40.299 | 40.476 |
| | 40.364 | 40.386 | 40.404 | 40.375 | 40.447 | 40.413 | 40.416 |
| | 40.396 | 40.380 | 40.353 | 40.394 | 40.330 | 40.353 | 40.401 |
| | 40.442 | 40.426 | 40.354 | 40.452 | 40.294 | 40.451 | 40.369 |
| | 40.325 | 40.356 | 40.346 | 40.386 | 40.340 | 40.356 | 40.335 |
| | 40.392 | 40.371 | 40.354 | 40.426 | 40.335 | 40.424 | 40.398 |
| | 40.336 | 40.334 | 40.406 | 40.291 | 40.439 | 40.413 | 40.352 |
| | 40.391 | 40.373 | 40.411 | 40.320 | 40.361 | 40.387 | 40.439 |
| | 40.400 | 40.466 | 40.364 | 40.379 | 40.391 | 40.443 | 40.355 |
| | 40.369 | 40.381 | 40.499 | 40.306 | 40.347 | 40.408 | 40.435 |
| | 40.460 | 40.550 | 40.388 | 40.405 | 40.378 | 40.370 | 40.496 |
| **mean** | **40.380** | **40.390** | **40.379** | **40.366** | **40.370** | **40.401** | **40.394** |
| **sd** | **0.047** | **0.055** | **0.041** | **0.045** | **0.044** | **0.048** | **0.056** |

# From eFits

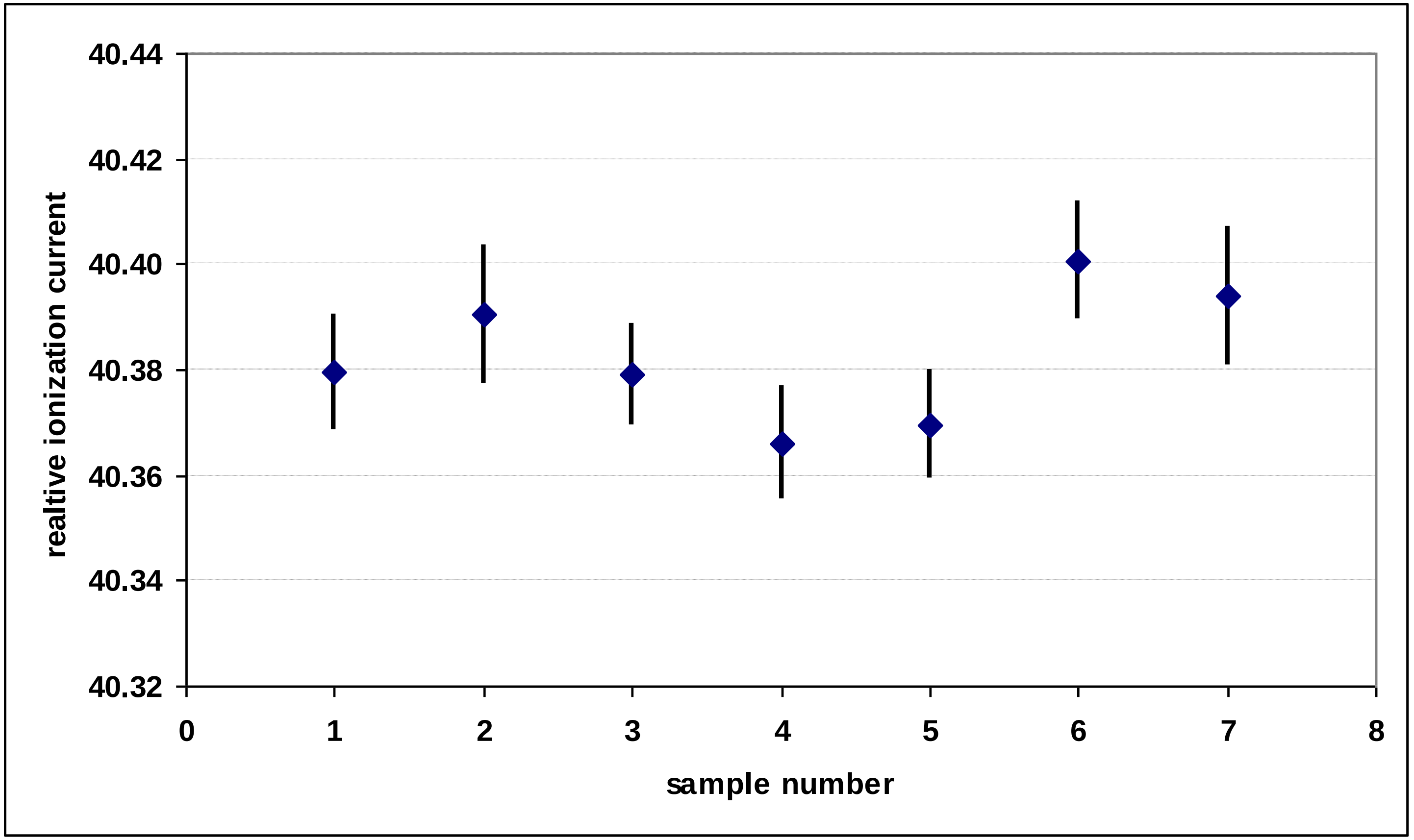



| | |
|---|---|
| Anderson-Darling test: | at 95 % level, data do come from normal dist. |
| Wilk-Shapiro test: | at 95 % and 99 %, accept normality assumption |
| Normal PPCC test | accept normality at 95 % and 99 % |

Nice data ...

What is the sdm?



| mean (18) | sd (18) | sdm (18) |
|---|---|---|
| 40.3796 | 0.0466 | 0.0110 |
| 40.3904 | 0.0553 | 0.0130 |
| 40.3791 | 0.0406 | 0.0096 |
| 40.3662 | 0.0451 | 0.0106 |
| 40.3696 | 0.0437 | 0.0103 |
| 40.4008 | 0.0480 | 0.0113 |
| 40.3941 | 0.0561 | 0.0132 |

0.0110
typical sdm

within variability

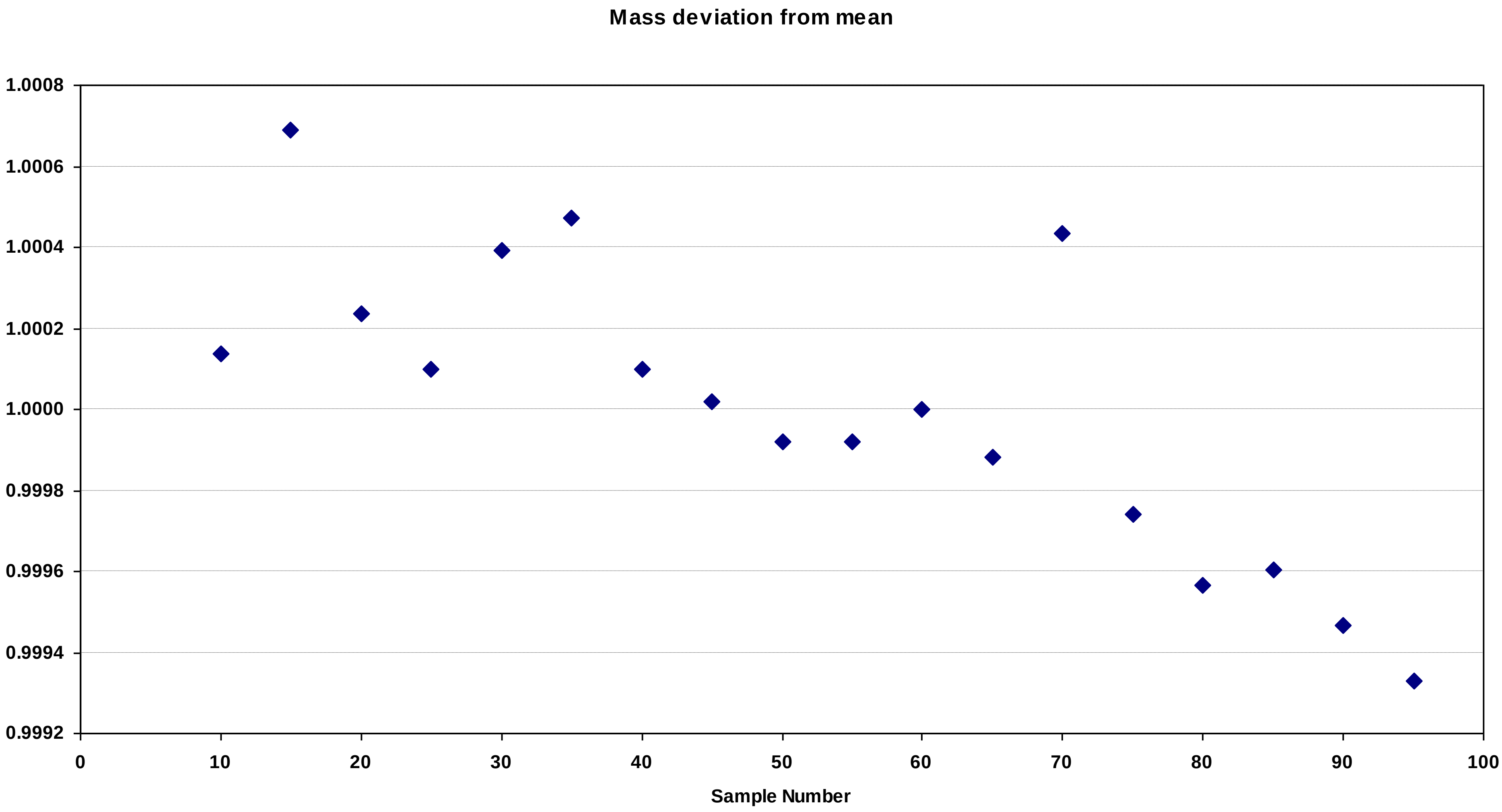
| | |
|---|---|
| mean (7) | 40.3828 |
| sd (7) | 0.0128 |

between variability

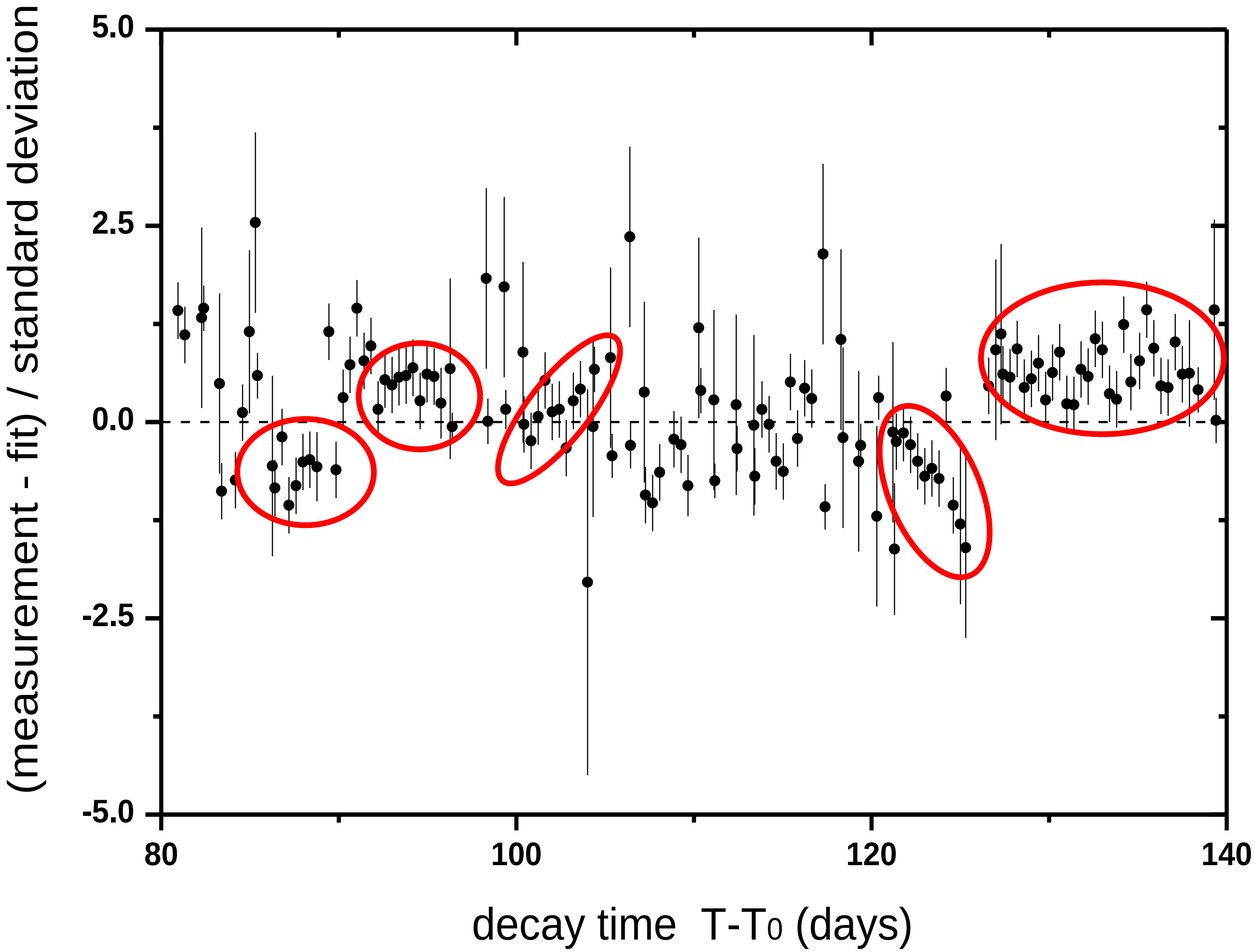
| | |
|---|---|
| mean (126) | 40.3828 |
| sd (126) | 0.0485 |
| sdm (126) | 0.0043 |

$$s = (0.0110^2 + 0.0128^2)^{1/2} = 0.017$$

Something going on here too !

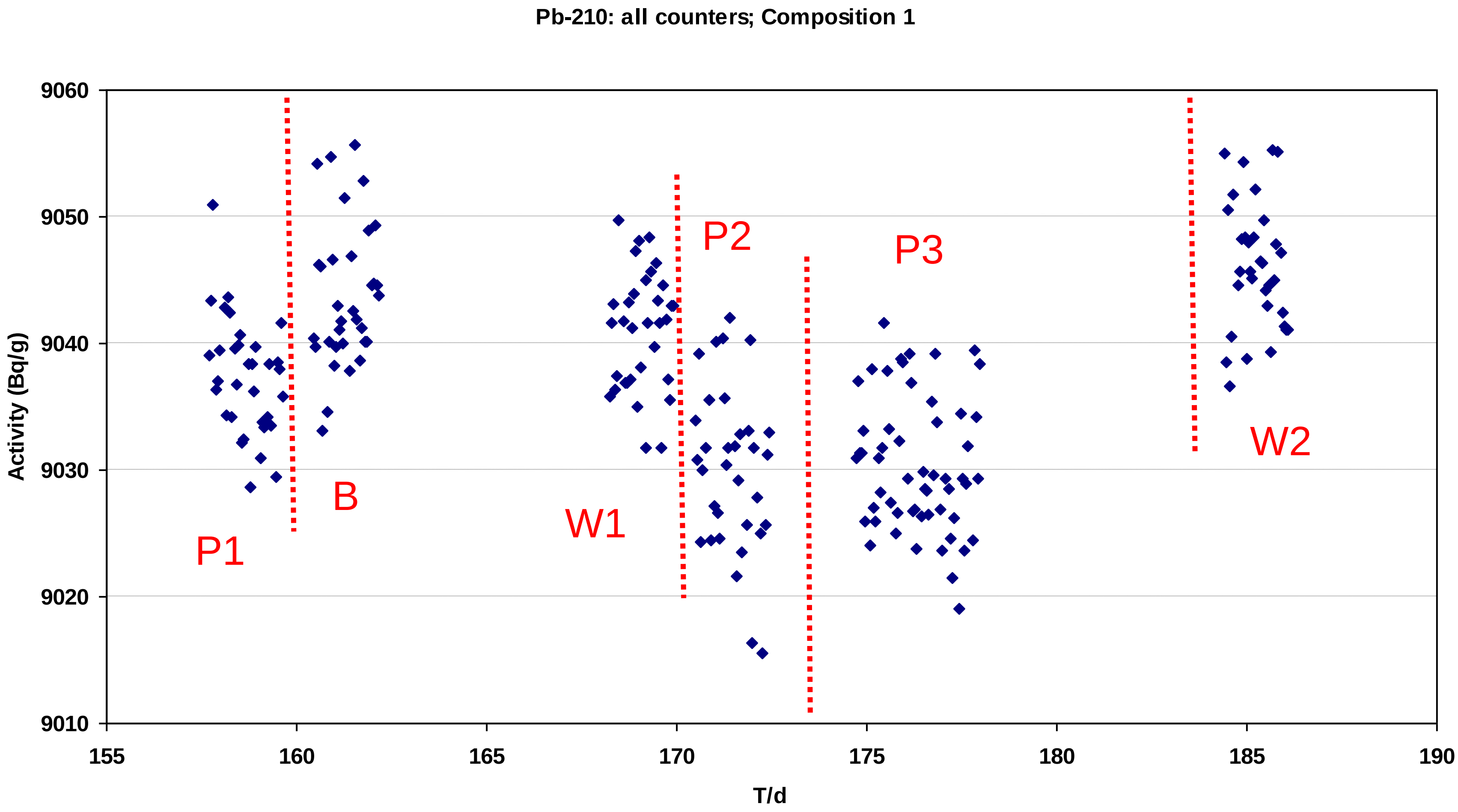


# $^{55}$Fe half-life data – residuals on fit




Thanks to Pommé for data

*Appl. Radiat. Isot*. **64**, 1412 (2006)


Pb-210: all counters; Composition 1
Activity (Bq/g)
9060
9050
9040
9030
9020
9010
155
160
165
170
175
180
185
190
T/d
P1
B
W1
P2
P3
W2

Often miss various instabilities & additional components of variance

Need to look at additional variables

Can divide instabilities according to frequency

**HIGH**

Counting statistics

Noise

Repeat measurement replication

Other short-term random effects

Etc....

**MEDIUM**

Day-to-day

“seasonal”

Between-measurement effects (same sample, different occasion)

Between-sample effects

Between-instrument/geometry effects

Short-lived impurity

Etc....

**LOW**

Deadtime

Background subtraction

Detector degradation

Reference source changes

Geometry changes

Etc....

Most of these not in the model’s formula !

# Lots of tools available

*(if have data from measurements)*

“statistics”

ANOVA

Correlation & autocorrelation analyses

Dataplot / eFit

*Etc ...*

***USE THEM***

*.. but not GUM software or spreadsheets*

# Main uncertainty reporting problems and issues

Loss of information

Meaning & interpretation

Incomplete assessments

OK to say negligible, but thought about it!

Correlations hard to put into component lists

# FIN.....

THANKS