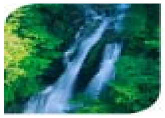
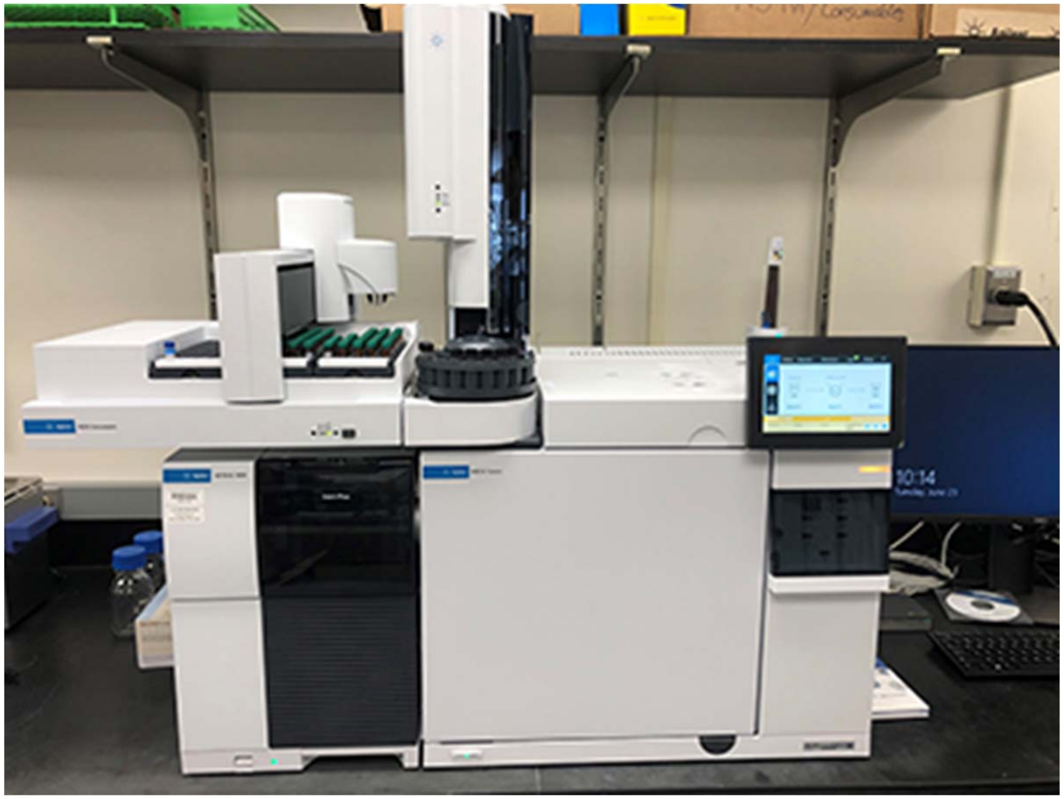
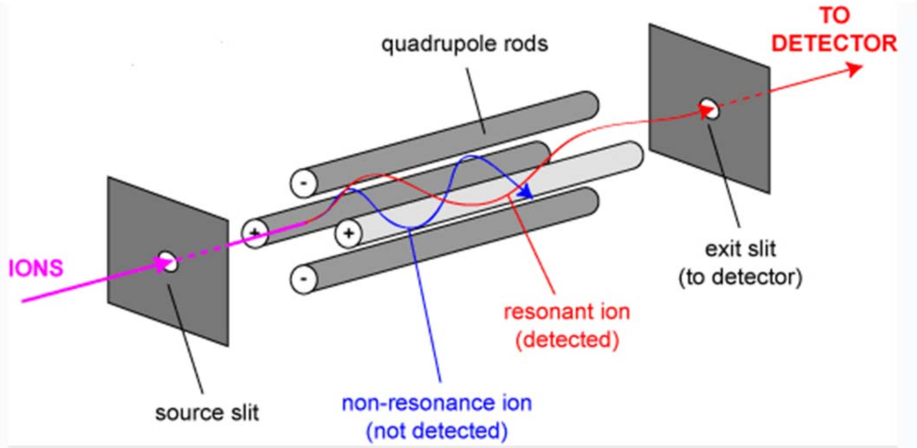
Modern Mass Spectrometers and the Correlation Coefficient – Are they compatible?

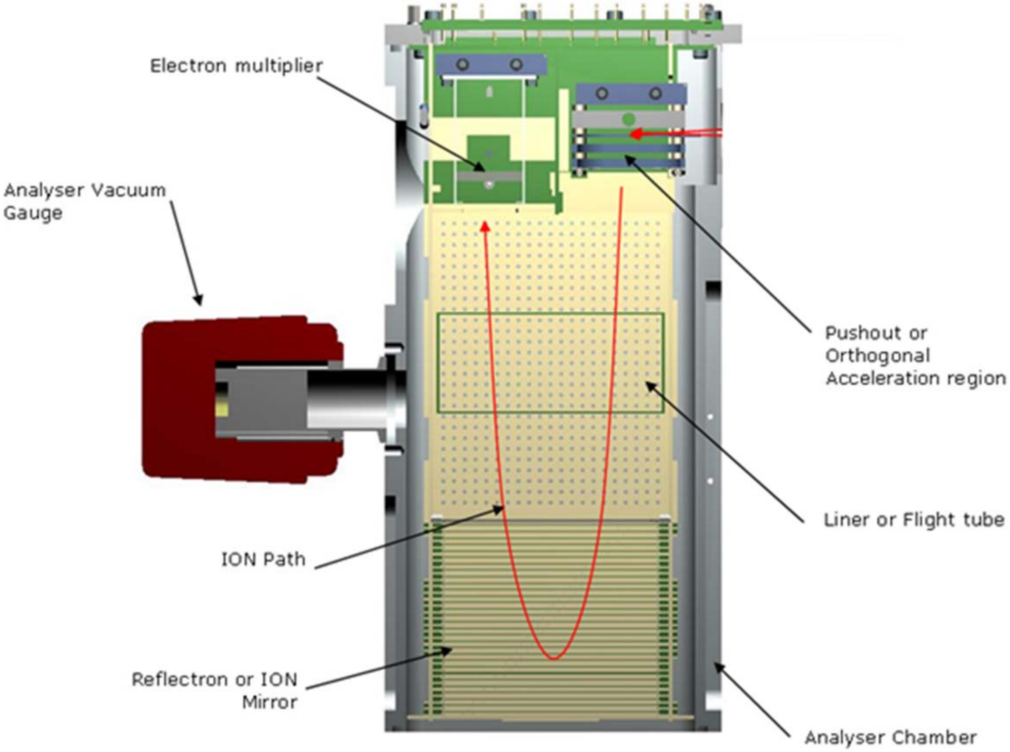
Richard Burrows

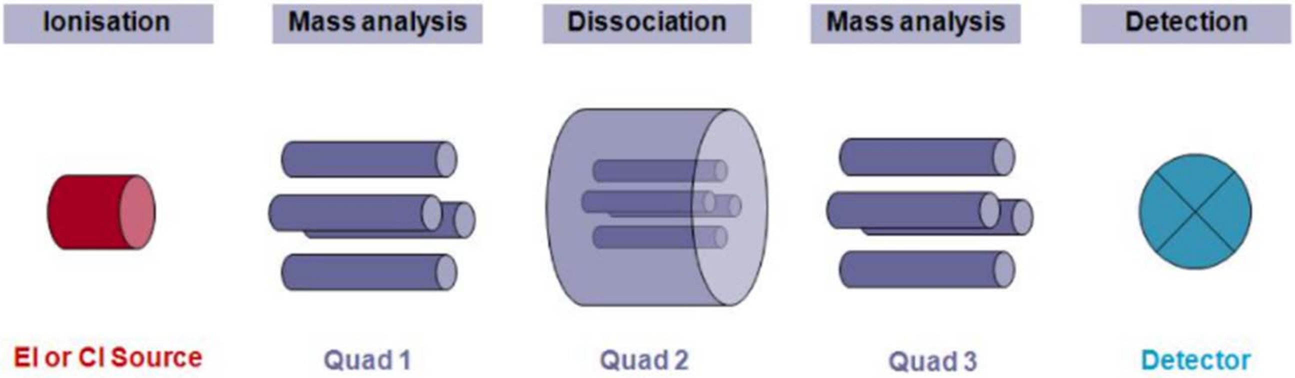





TO
DETECTOR
quadrupole rods
IONS
exit slit
(to detector)
resonant ion
(detected)
source slit
non-resonance ion
(not detected)

Time of Flight





Ionisation | **Mass analysis** | **Dissociation** | **Mass analysis** | **Detection**

EI or CI Source

Quad 1

Quad 2

Quad 3

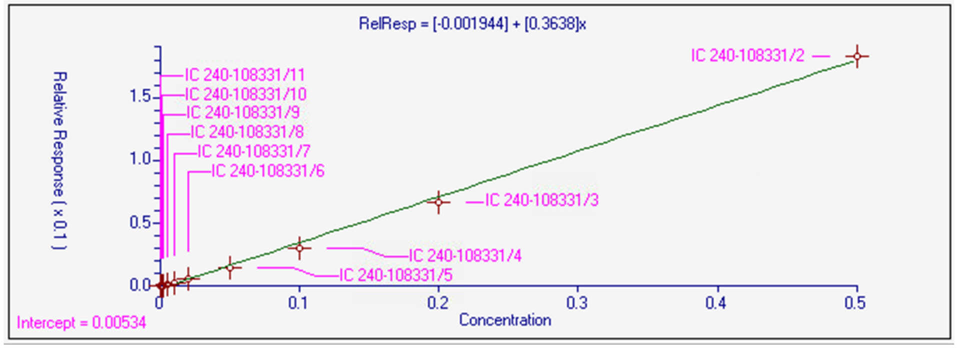
Detector

Why use advanced mass spectrometers?

- More sensitive
- Greater working range
- Better selectivity

Faster Cheaper Better
Choose all three!

Hexadecane, 0.5-500 ng/mL



Linear Unweighted

Coefficient of determination: 0.998

Low point error 1109%

Bottom 5 points all have error > 30% (RSE 213%)

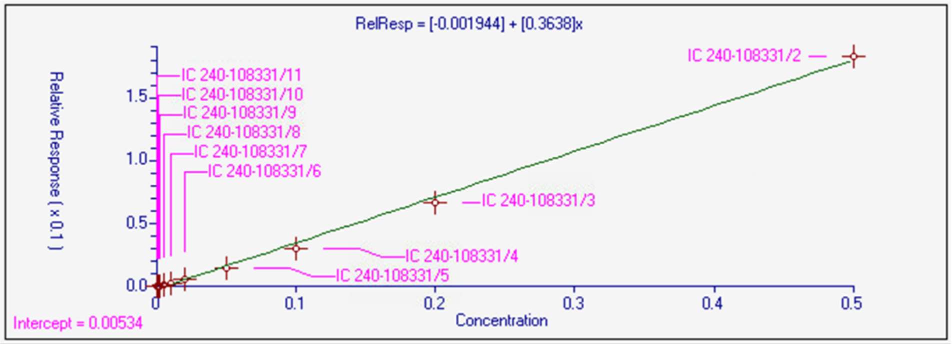
Quadratic unweighted

Coefficient of determination 1.000

Low point error 326%

Bottom 3 points all have error> 50% (RSE 134%)

Hexadecane, 0.5-500 ng/mL



Linear Weighted

Coefficient of determination: 0.963

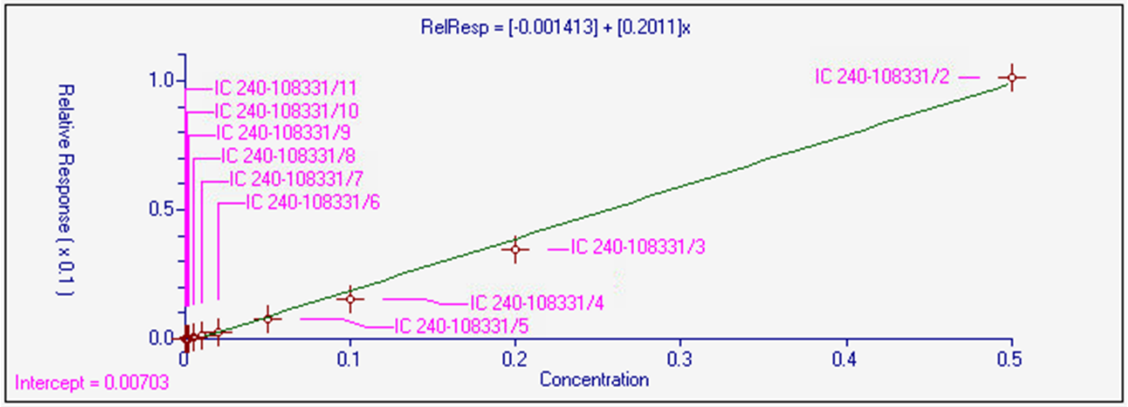
All points error < 30% (RSE 18.5%)

Quadratic weighted

Coefficient of determination 0.986

All points error < 25% (RSE 13.2%)

2,4,5 Trichlorophenol, 0.5-500 ng/mL



Linear Unweighted

Coefficient of determination: 0.996

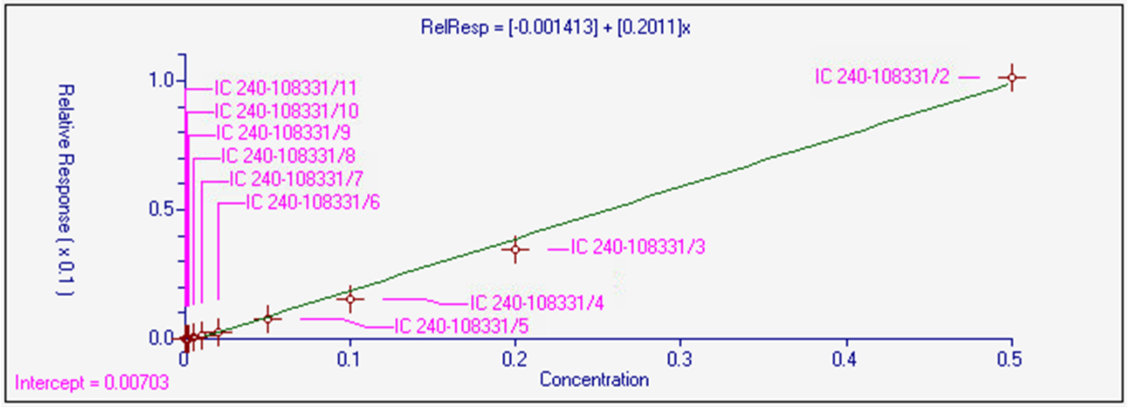
Low point error 1335%, bottom 4 points all > 85% error (RSE 535%)

Quadratic unweighted

Coefficient of determination 1.000

Low point error 220%, bottom 3 points > 50% error (RSE 90.4%)

2,4,5 Trichlorophenol, 0.5-500 ng/mL



Linear weighted

Coefficient of determination: 0.958

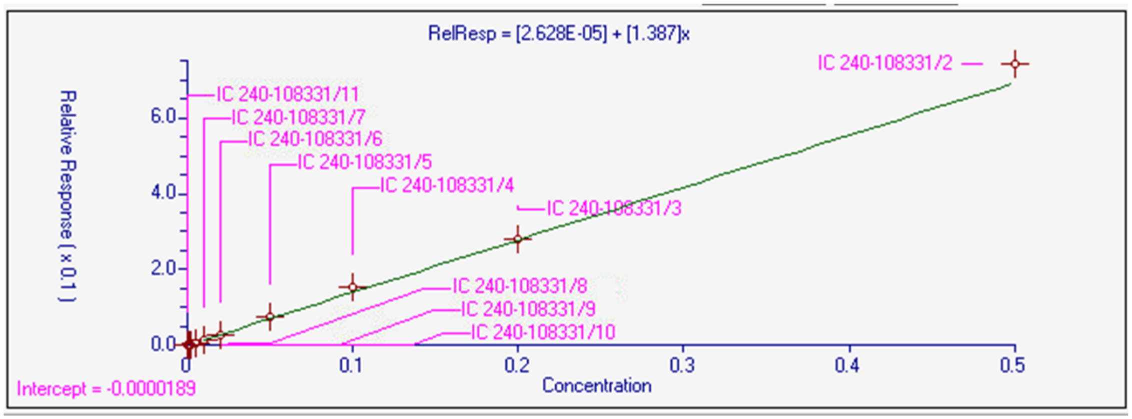
All points < 30% error (RSE 19.8%)

Quadratic weighted

Coefficient of determination 0.985

All points < 22% error (RSE 13.8%)

Chrysene, 0.5-500 ng/mL



Linear unweighted

Coefficient of determination: 0.999

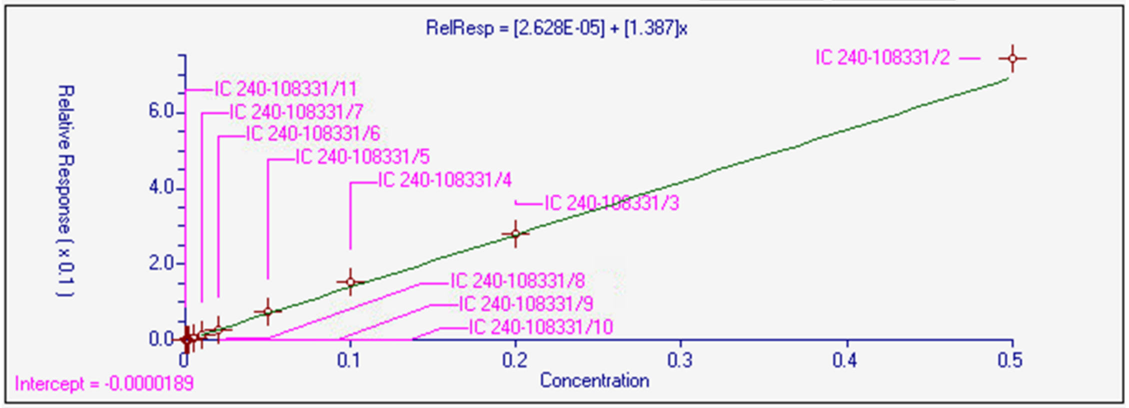
Low point 166% error (RSE 62.4%)

Quadratic unweighted

Coefficient of determination 1.000

Low point error 142% (RSE 68.7%)

Chrysene, 0.5-500 ng/mL



Linear weighted

Coefficient of determination: 0.985

All points < 20% error (RSE 11.7%)

Quadratic weighted

Coefficient of determination 0.987

All points < 20% error(RSE 12%)

What kind of weighting is best for Environmental Analysis

What kind of weighting is best for environmental analysis?

- No weighting?
- 1/Concentration weighting?
- 1/(Concentration)^2 weighting?

What are we trying to achieve?

What do we want from a calibration?

Accurate translation of response into concentration

> **What do we mean by “accurate”?**

Minimize the <u>relative</u> error at all points in the working range

Assumption – we are equally interested in accuracy at all points in the working range

What weighting assigns equal weight to each calibration point?

1/(Concentation)^2

Why?

The regression minimizes the sum of the squares of the absolute residuals

$$S_{min} = \sum_{i=1}^{n} r_i^2$$

What about Average RF?

Average RF = $1/(\text{Concentration})^2$ weighted regression forced through zero

Good choice, unless

- Best line does not go through zero

Or

- Response is curvilinear

Best calibration weighting

Average

Unless does not go through zero, in which case

1/(Concentration)^2 weighted linear regression

Unless curvilinear response, in which case

1/(Concentration)^2 weighted quadratic regression

What do our calibration measures offer?

- Correlation coefficient and coefficient of determination
  - Unweighted almost always has a “better” number than 1/(Concentration) weighted
  - 1/(Concentration) weighted almost always has a “better” number than 1/(Concentration)^2 weighted

For almost any calibration, the correlation coefficient and coefficient of determination almost always lead us in the direction of choosing the wrong calibration

Applies to quadratic as well as linear regressions

What do our calibration measures offer?

- Relative Standard Error (RSE)
  - Unweighted almost always has a worse number than 1/Concentration weighted
  - 1/Concentration weighted usually has a worse number than 1/(Concentration $)^2$ weighted

For almost any calibration, the RSE almost always leads us in the direction of choosing the correct calibration

Calibration measures

What if we were routinely using a measure for calibration quality that:

- **Guided us towards using the worst possible type of calibration fit (one that creates large errors in the amount of analyte)**
- **For the same data set, told us that curve fits with very large errors were good, and curve fits with much smaller errors were bad**

That would not be a good thing, right?

Measures

Unfortunately, that is exactly what we do.

For most of our methods!!

The measures are the correlation coefficient and the coefficient of determination

r = Correlation coefficient

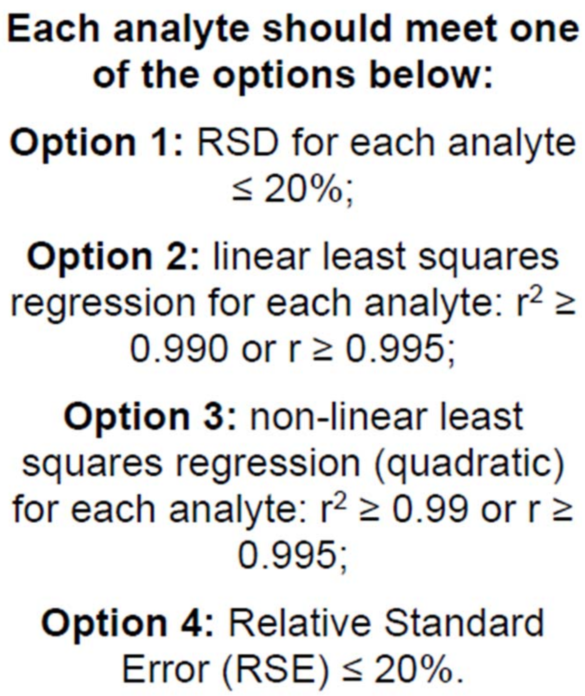
r^2 = Coefficient of determination

Not just environmental analysis –
pervasive problem in analytical chemistry
in general

Progress so far…

- RSE added to Method 8000 and 600 series
- RSE added to TNI standards
- Relative error added to 8000 series
- Relative Error added to TNI standards

DOD data validation guidelines



Each analyte should meet one of the options below:

Option 1: RSD for each analyte ≤ 20%;

Option 2: linear least squares regression for each analyte: $r^2 \geq 0.990$ or $r \geq 0.995$;

Option 3: non-linear least squares regression (quadratic) for each analyte: $r^2 \geq 0.99$ or $r \geq 0.995$;

Option 4: Relative Standard Error (RSE) ≤ 20%.

DOD Data Validation Guidelines do not require assessment of correlation coefficient or coefficient of determination

RSE Status

- RSE adoption should be relatively straightforward because:
  - For the average RF calibration RSE = RSD
  - RSE essentially just allows RSD to be applied to all types of curves, instead of just Average RF
- However:
- Virtually unused
  - May increase after 2016 standards are adopted
  - Needs to be incorporated into major manufacturer instrument software
  - Needs removal of correlation coefficient option??
  - Needs champions

Correlation coefficient

IUPAC, 1998
Guidelines for Calibration in Analytical Chemistry

The correlation coefficient, which is a measure of two random variables, has no meaning in calibration because the values x are not random quantities

Correlation Coefficient

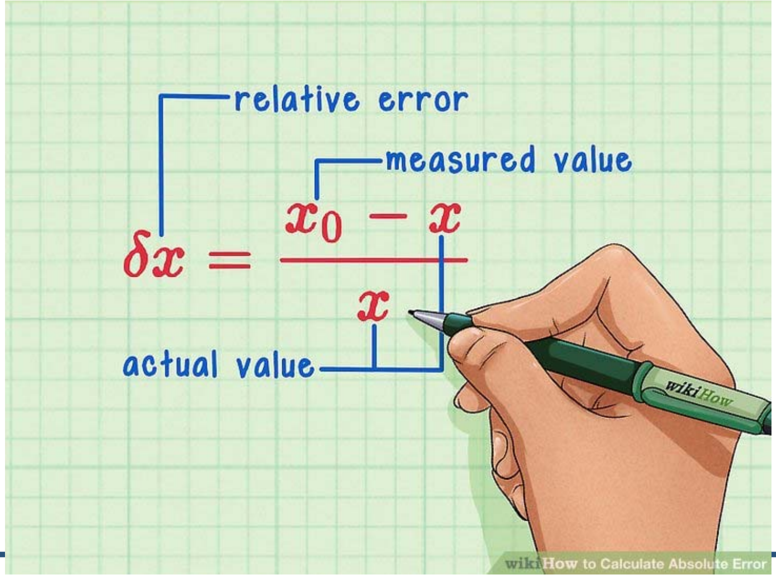
For most applications, and calibration curves in particular, the correlation coefficient must be regarded as a relic of the past

- **Meier and Zund, Statistical Methods in Analytical Chemistry, 2000**

So what is important?

•Measuring relative error

•Do we already have measures of relative error in EPA methods?



Method 524.4

- Linear or quadratic regression may be used
- Calibration points ≤ MRL must calculate within 50% of true value (Relative Error)
- Calibration points above the MRL must calculate within 30% of true value (Relative Error)
- **No correlation coefficient or coefficient of determination!**

Alternative to RSE is measuring relative error at each point, or at key points (for example at the mid point and the low point

Relative error (Method 524)

Using relative error of each point is less desirable than RSE, but it is good:

- Measures what is important, relative error
- Consistent with TNI standards
- Consistent for different curve fits

8270E Relative Error

Average curve fit – RSD (Relative Error)

Linear or quadratic regression

- Has RSE option (Relative Error)
- Recalc at low point 50%, other points 30% (Should) (Relative Error)
- Consistent with method 524
- Unfortunately includes correlation coefficient and coefficient of determination

Just drop r and r^2!!

SW846 guidance

The intent of Section 11.5.4 of Method 8000D is consistent with the title, ‘Acceptance criteria independent of calibration model’. This section provides Relative Standard Error (RSE) or % Error across a range of individual calibration standards as stand-alone alternatives to other measures of calibration fit.

In other words, SW-846 does not require use of correlation coefficient or coefficient of determination

624.1 Relative Error

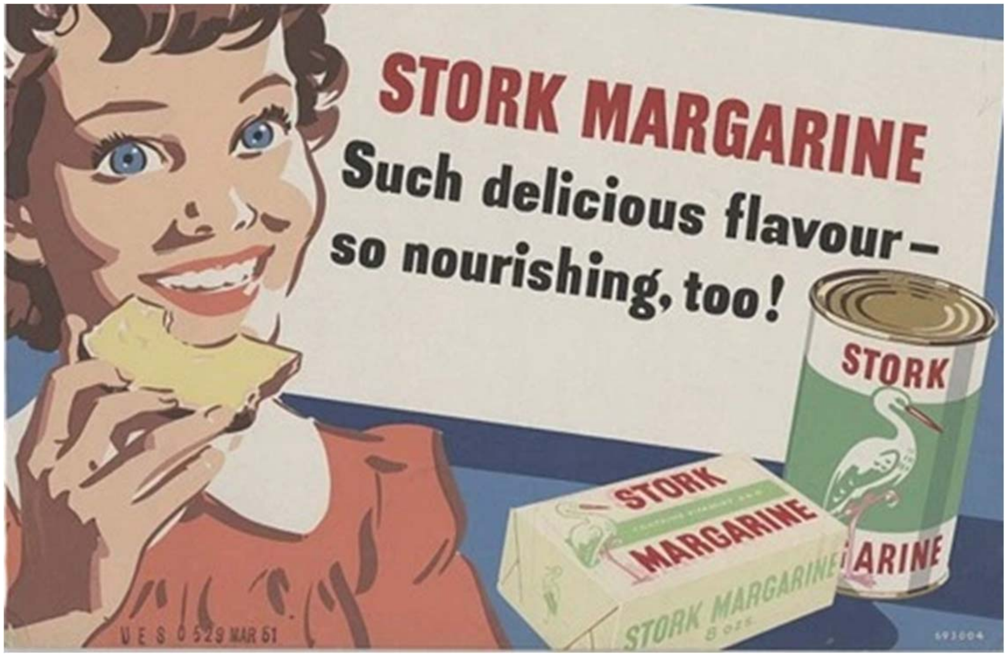
Average curve fit – RSD

Linear or quadratic regression

- Has RSE option (Relative Error)
- No recalc
- Unfortunately includes coefficient of determination

Just drop r^2 and add Recalc!

Margarine


STORK MARGARINE
Such delicious flavour –
so nourishing, too!
STORK
STORK
MARGARINE
STORK MARGARINE
8 ozs.
ARINE
V E S 0529 MAR 51
693004

Radioactive face cream

Arsenic based face cream

DR. MOUNTEBANK'S

REGAIN YOUTHFUL PALLOR

REMOVES BLEMISHES PERMANENTLY

ARSENICAL LOTION

SAFE ~ ABSOLUTELY HARMLESS ~ EXPENSIVELY PERFUMED

RESTORES A LOVELY COMPLEXION

Since 1807

Mdme. Esme Ennui finds Dr. Mountebank's Arsenical Lotion most excellent

From PARIS

What Next?

- Calibration is the most critical part of an analytical method

- We must have good measures of calibration quality

Therefore, we have to get rid of the correlation coefficient and coefficient of determination