

Environment Testing
TestAmerica

Calibration – Past, Present and Future

Richard Burrows

Fair Warning

I gave a similar presentation 2 years ago

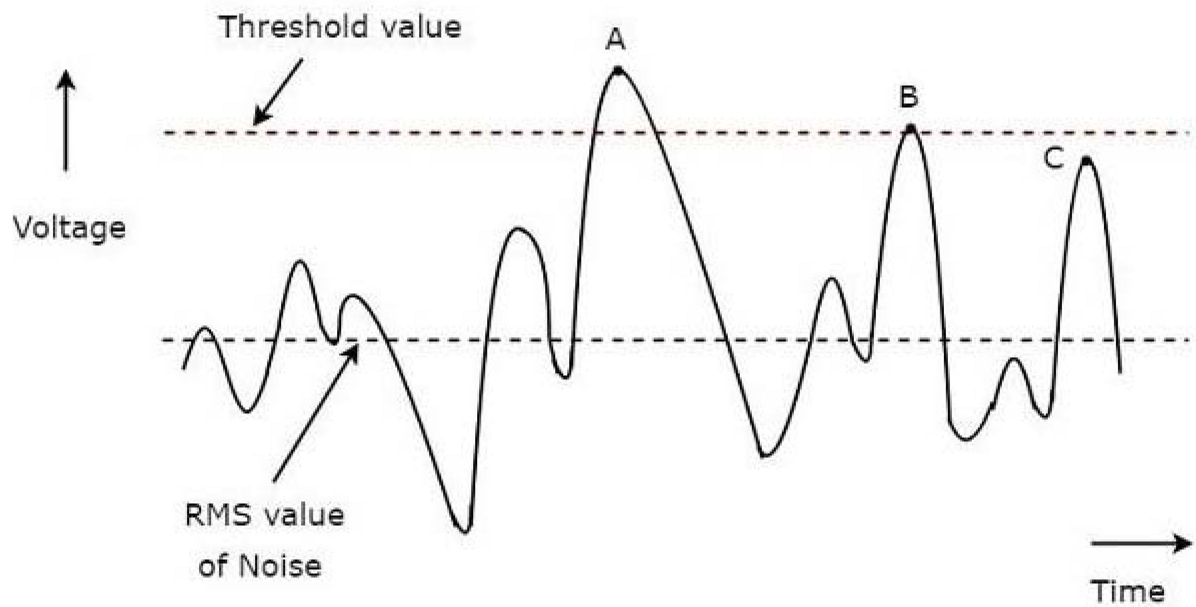
This issue adversely affects almost every method we perform

It leads to acceptance of poor data, and rejection of good data

We have not made much progress in the last two years

Most important property of an analytical method

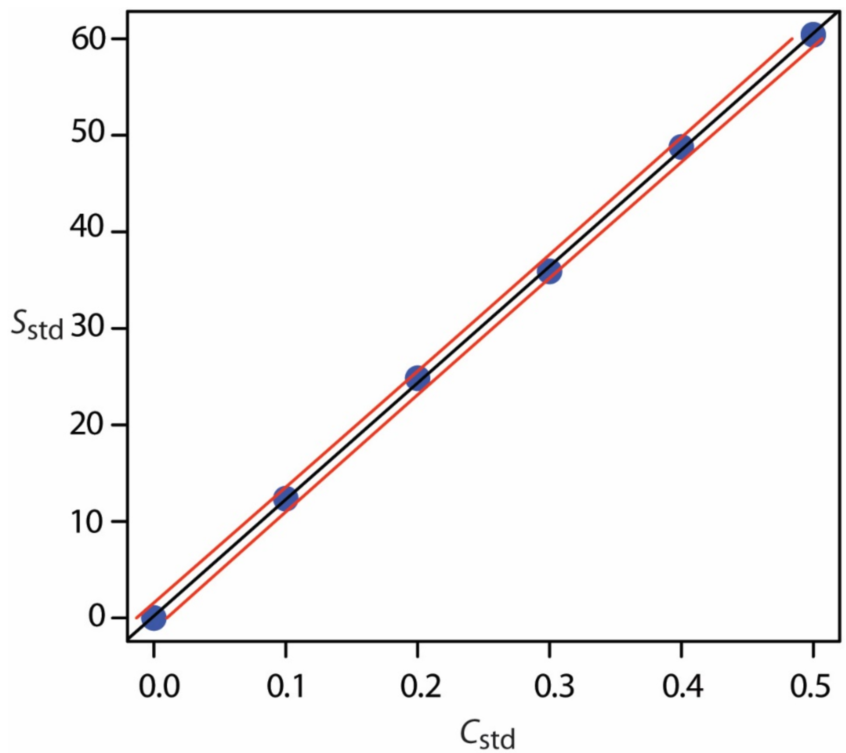
A detectable response that is proportional to the quantity of analyte



Second most important property of an analytical method

A calibration:

The ability to relate the size of the response to the quantity of analyte present



Measurement of a calibration quality

Once we have created a calibration, we need to be able to evaluate the quality of the calibration.

The measure that we use should:

Guide us towards selecting the best type of calibration for the data

- **Linear, curvilinear, weighted, unweighted, forced zero, etc.**

Tell us how effective the selected calibration fit is at translating the instrument response to the amount of analyte

Current types

Type of calibration typically used in environmental analysis include:

- **Linear**
- **Quadratic (curvilinear)**
- **Unweighted**
- **Weighted by the reciprocal of concentration or concentration squared**
- **Average response factor (a special case of weighted and forced through zero)**

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

Let’s look at a calibration



Amount	Response
2.00	38345
5.00	104587
10.00	211363
20.00	432675
40.00	871485
80.00	1483247
120.00	2084890

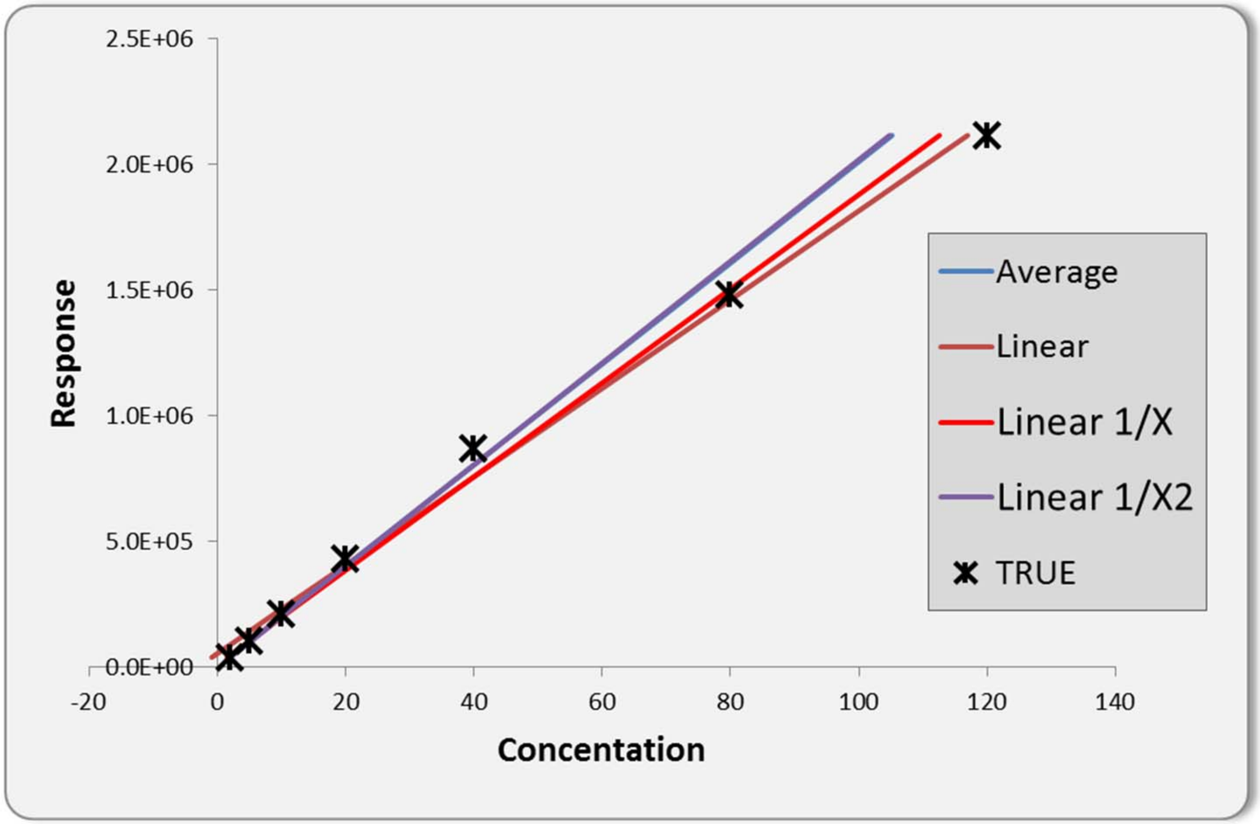
Acceptance criteria

- Criteria
  - Coeff. Determination = $r^2 > 0.990$

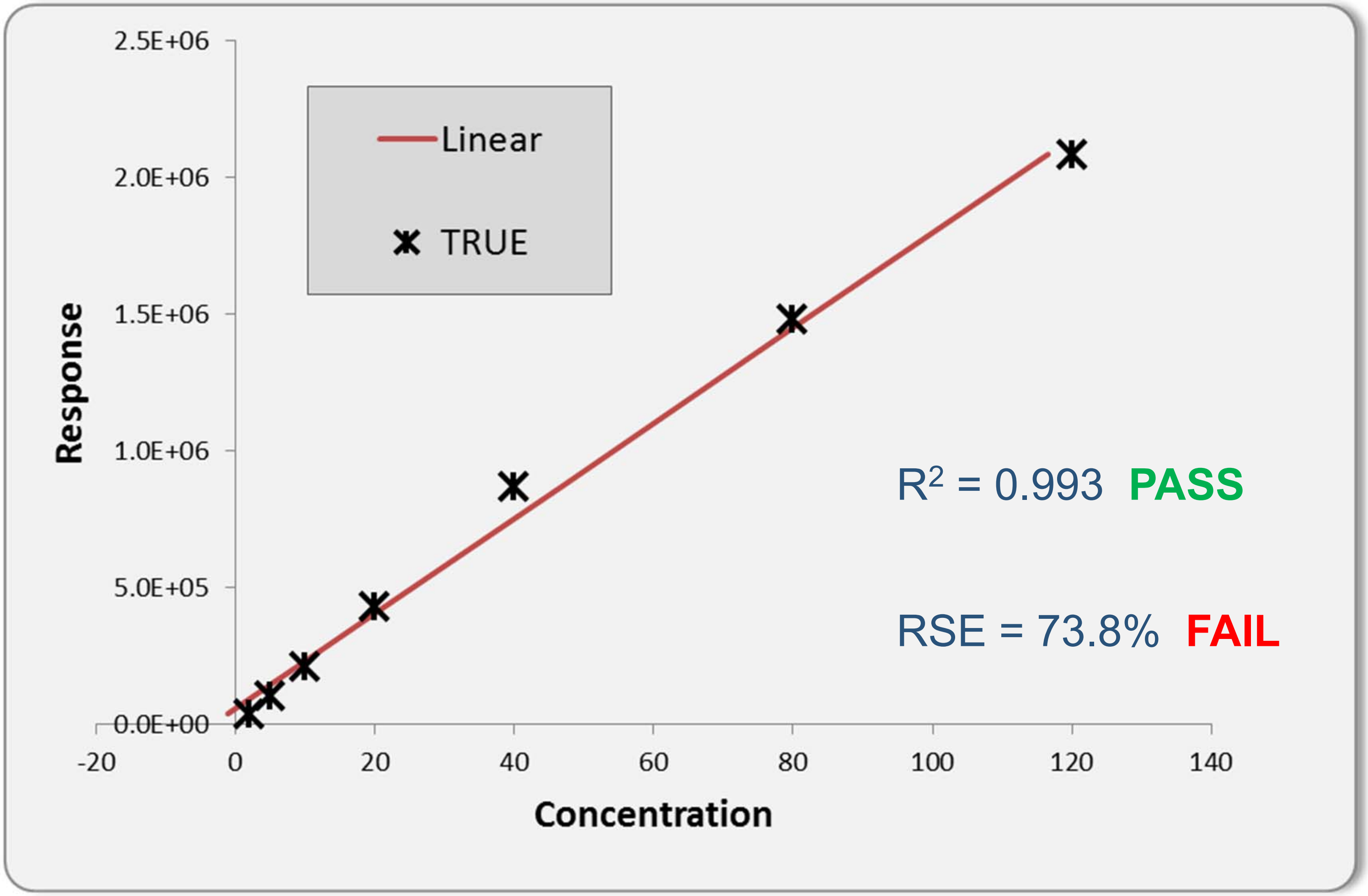
    And
  - RSE $\leq$ 20%

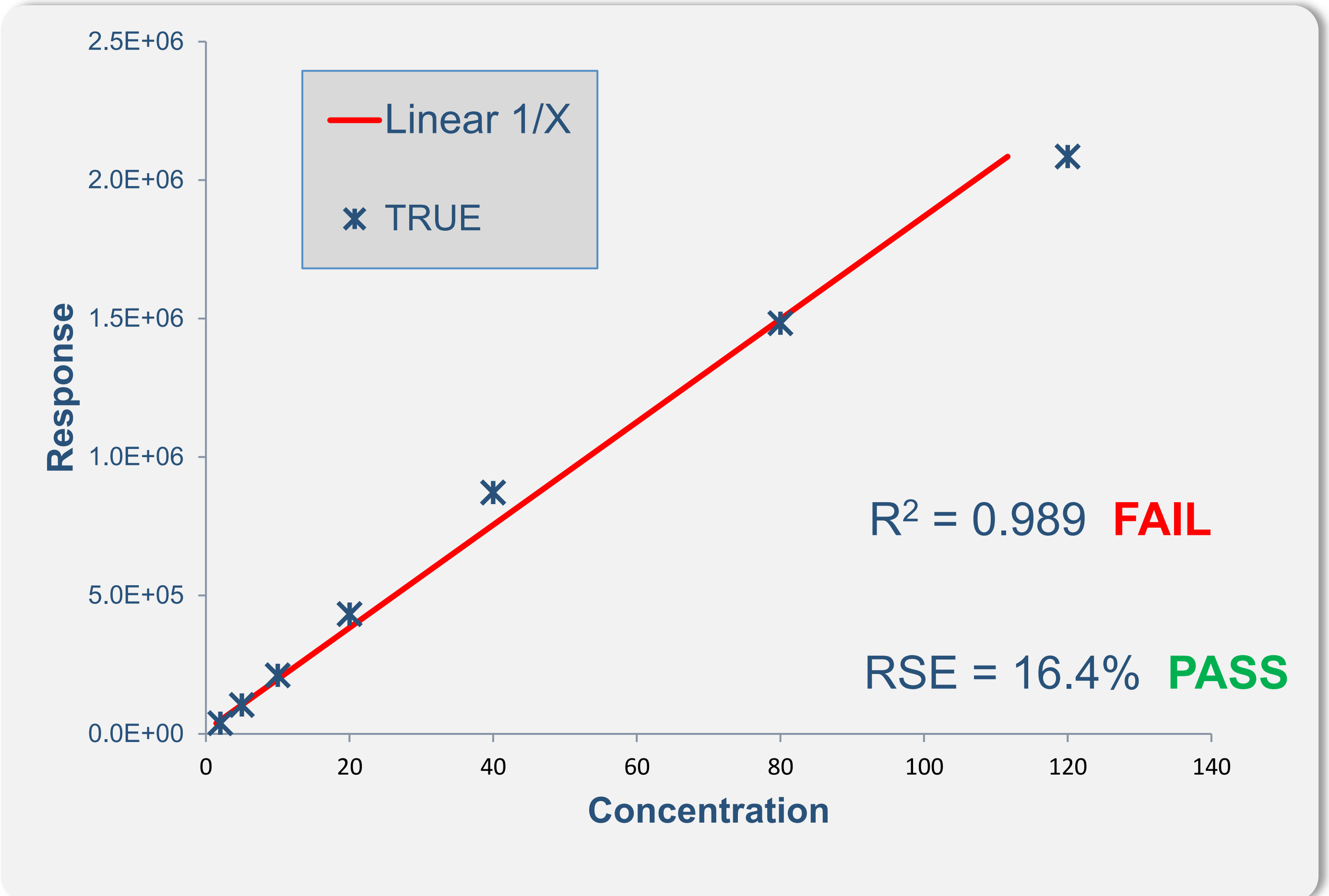
    Or
  - Mid point relative error < 20%
  - Low point relative error < 30%

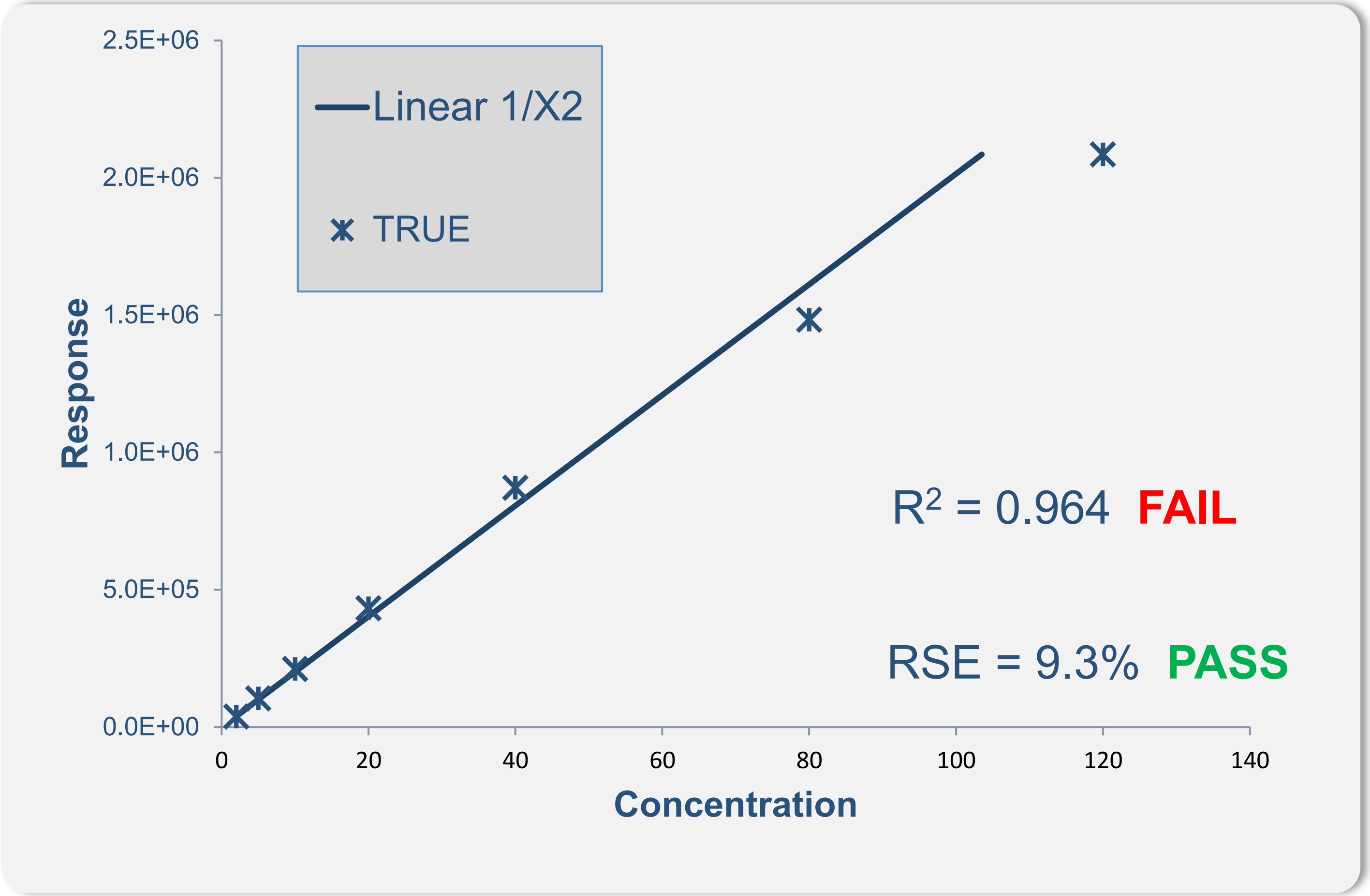
Plot different curve fits



Linear regression calibration



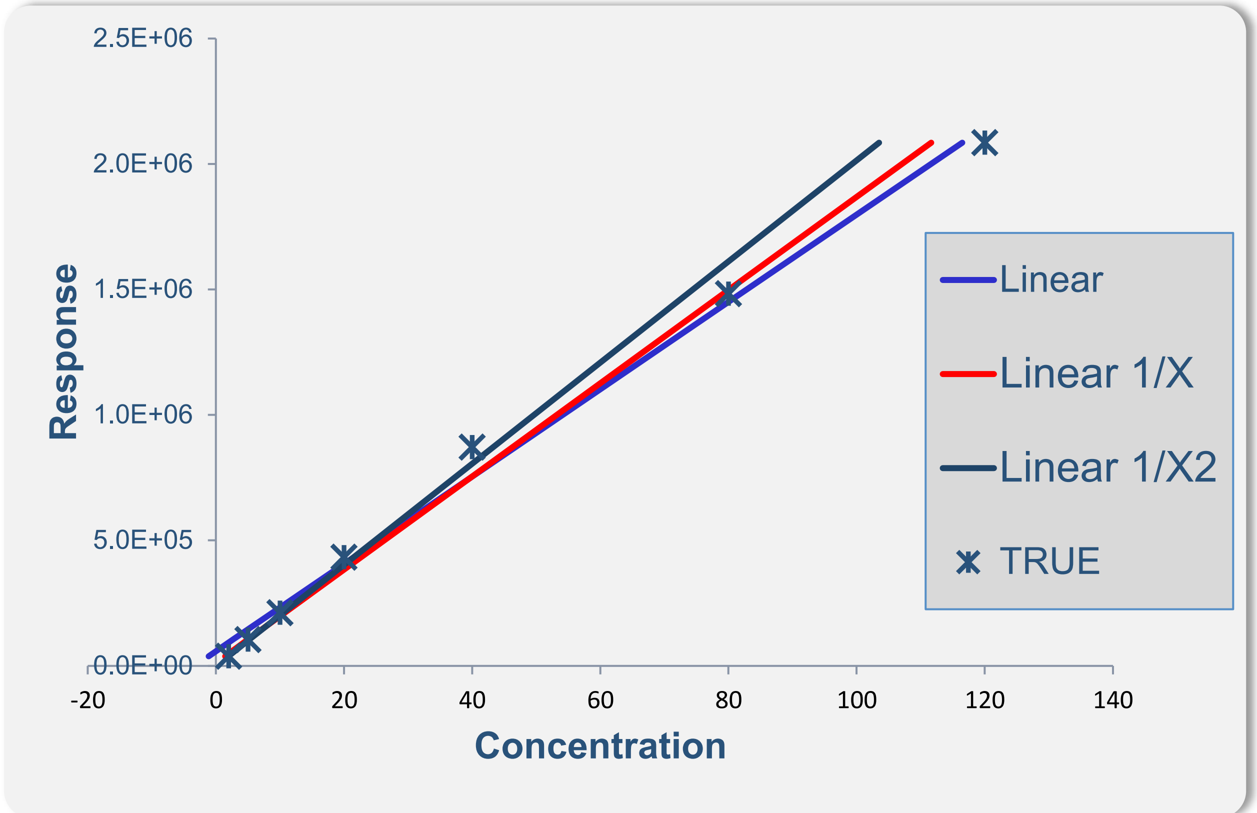

Linear 1/X
TRUE
2.5E+06
2.0E+06
1.5E+06
1.0E+06
5.0E+05
0.0E+00
Response
0
20
40
60
80
100
120
140
Concentration
R² = 0.989 FAIL
RSE = 16.4% PASS



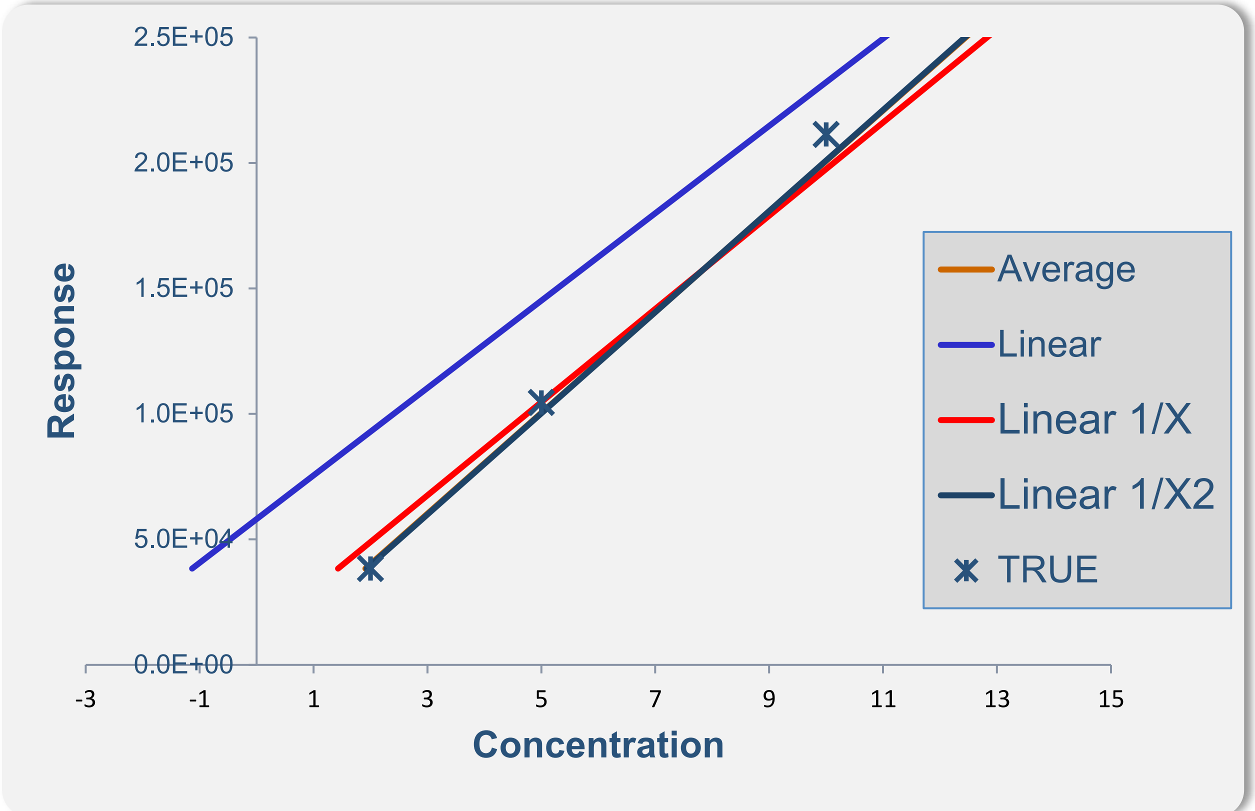
Linear 1/X2
TRUE
Response
Concentration
2.5E+06
2.0E+06
1.5E+06
1.0E+06
5.0E+05
0.0E+00
0
20
40
60
80
100
120
140
R^2 = 0.964 FAIL
RSE = 9.3% PASS

	Average	1/X	$1/X^2$	Unweighted
2	-4.52%	-28.52%	-3.50%	**-156.7%**
5	4.17%	-0.07%	4.35%	**-46.54%**
10	5.26%	7.46%	5.17%	-11.91%
20	7.74%	13.32%	7.50%	7.64%
40	8.50%	15.73%	8.19%	16.86%
80	-7.67%	-0.95%	-7.95%	2.38%
120	-13.48%	-6.97%	-13.75%	-2.94%
RSE	**8.5%**	**16.4%**	**9.3%**	**73.8%**
R^2	**0.983**	**0.989**	**0.964**	**0.993**

Curve fit plots



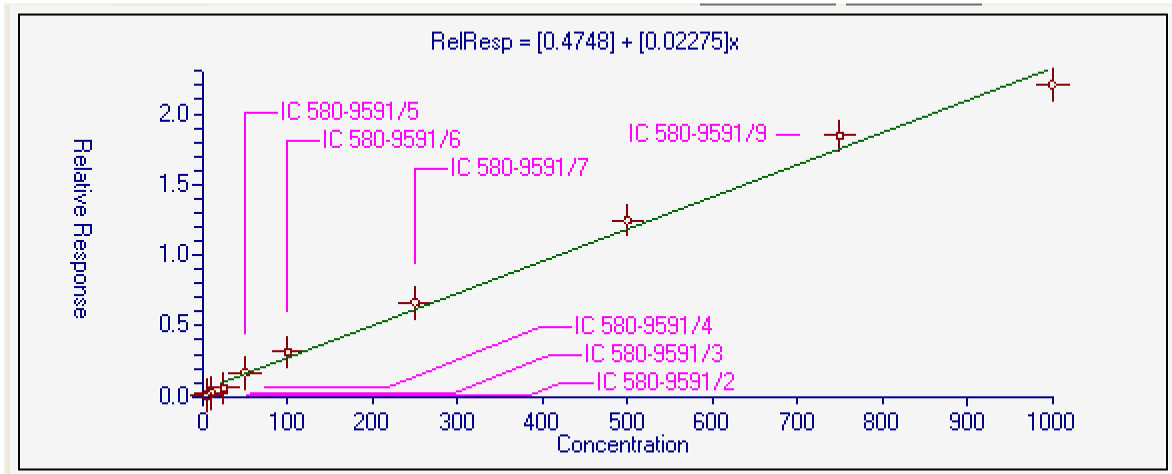
Plot at the low end



Correlation coefficient gets better

Curve quality gets worse

Calibration issues



r= 0.997, r^2 = 0.994

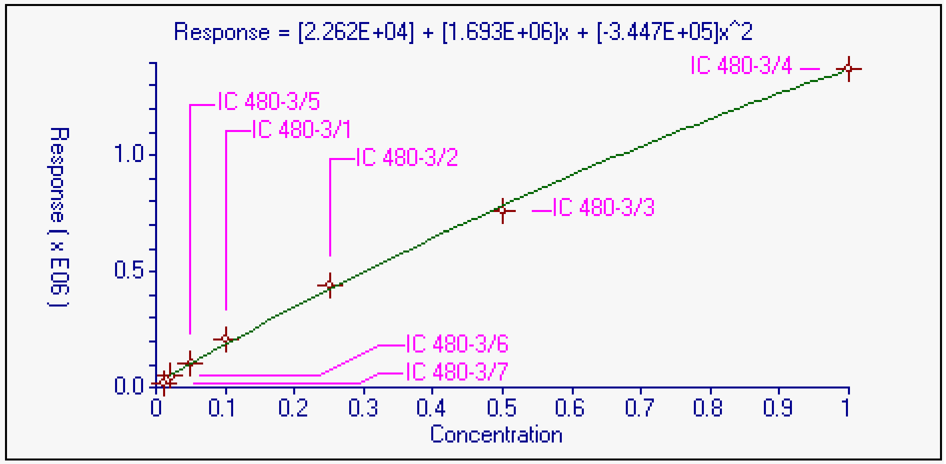
RSE = 179%

Calibration Standard Levels

Level	Used	Amount	Area	ISArea	%Error
IC 580-9591/2	✓	5	1348	618332	421.63
IC 580-9591/3	✓	10	3250	647316	198.43
IC 580-9591/4	✓	25	7697	646400	78.87
IC 580-9591/5	✓	50	23729	700099	7.13
IC 580-9591/6	✓	100	47131	748204	17.47
IC 580-9591/7	✓	250	111297	833662	8.93
IC 580-9591/8	✓	500	229185	917698	5.52
IC 580-9591/9	✓	750	371628	1005615	5.43
IC 580-9591/10	✓	1000	499631	1131444	5.11

421.63
198.43

Dalapon



$r^2 = 0.999$

RSE = 63%

Level	Used	Amoun	Area	%Error
IC 480-3/7	☑	0.01	22047	103.37
IC 480-3/6	☑	0.02	49262	21.07
IC 480-3/5	☑	0.05	106980	0.68
IC 480-3/1	☑	0.1	211249	14.05
IC 480-3/2	☑	0.25	442363	4.74
IC 480-3/3	☑	0.5	762496	3.04
IC 480-3/4	☑	1	1374873	0.38

%Error
103.37
21.07

ICPMS, 51V

	Blank offset	Unweighted	1/X	$1/X^2$	$1/SD^2$
1	1.23	-0.45	0.957	1.00	1.07
10	10.3	8.66	10.0	9.94	10.1
100	104.5	102.9	104.1	102	103.9
2000	1999	2000	1996	1963	1991
R	1.0000	1.0000	1.0000	1.0000	1.0000

So what is going on?

Regression

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

In an **unweighted** regression, we are minimizing the sum of the squares of the **absolute** values of the residuals

Correlation coefficient

$$r_{xy} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\ \sqrt{n\sum y_i^2 - (\sum y_i)^2}}.$$

The main takeaway is that the correlation coefficient is evaluating how far away from the expectation each point is – in **absolute** terms

Absolute or Relative?



Absolute error = 5					
True	1	5	20	50	100
Measured	-4 , 6	0 , 9	16 , 25	45 , 55	95 , 105

Relative error = 10%					
True	1	5	20	50	100
Measured	0.9 , 1.1	4.5 , 5.5	18 , 22	45 , 55	90 , 110

Unweighted regression minimizes the absolute error
Correlation coefficient evaluates absolute variance

Which is NOT what we want!

A few questions

What is this RSE?

$$\% RSE = 100 \times \sqrt{\sum_{i=1}^{n} \left[\frac{x_i' - x_i}{x_i}\right]^2 \Big/ (n-p)}$$

For Average Response Factor, RSE = RSD

	Average	1/X	$1/X^2$	Unweighted
2	-4.52%	-28.52%	-3.50%	**-156.7%**
5	4.17%	-0.07%	4.35%	**-46.54%**
10	5.26%	7.46%	5.17%	-11.91%
20	7.74%	13.32%	7.50%	7.64%
40	8.50%	15.73%	8.19%	16.86%
80	-7.67%	-0.95%	-7.95%	2.38%
120	-13.48%	-6.97%	-13.75%	-2.94%
RSE	**8.5%**	**16.4%**	**9.3%**	**73.8%**
R^2	**0.983**	**0.989**	**0.964**	**0.993**

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

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

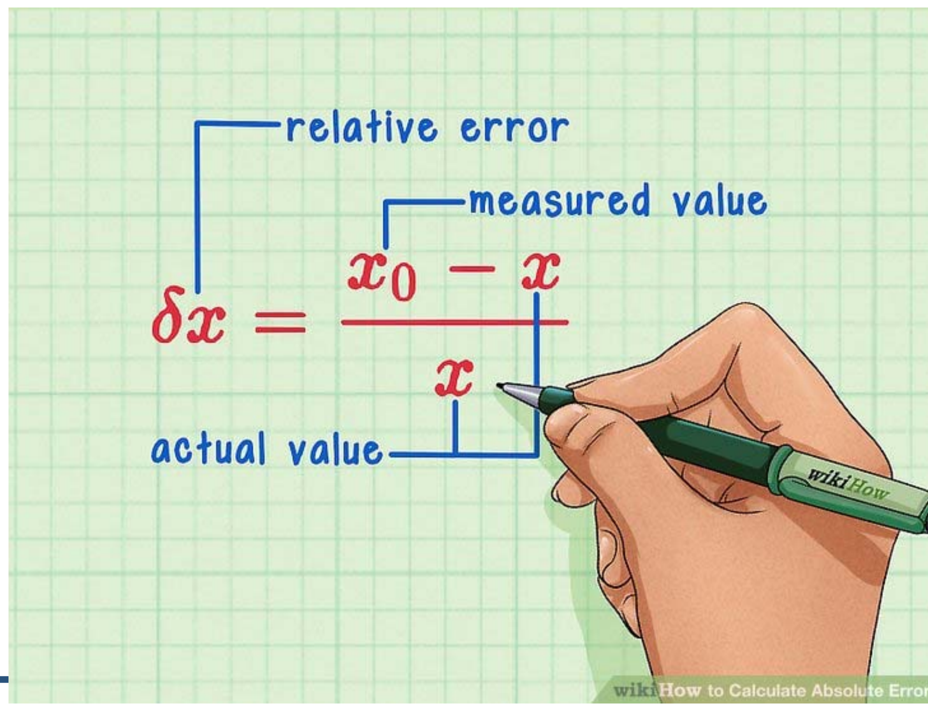
Correlation Coefficient

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**

•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!!

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!

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

It will not be easy…..