Calibration – What Changed, Why, and What's Next?

Richard Burrows

Calibration curve

A mathematical model for translating an instrumental response into an amount or concentration of a target analyte

Models used in Environmental Analysis

Average response factor

Linear regression

Weighted or unweighted

Quadratic regression

Weighted or unweighted

Measuring calibration quality

Average RF: % Relative Standard Deviation, RSD

Regression curves:

Correlation coefficient / Coefficient of determination

And now:

Relative Standard Error, RSE

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

A few questions



What is this RSE?

Why do we need it?

Are we done?

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

For Average Response Factor, RSE = RSD

Let's analyze 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

Average response (or calibration) factor



$$RF = \frac{A_x}{C_x}$$

Amount	Response	**RF**
2.00	38345	19173
5.00	104587	20917
10.00	211363	21136
20.00	432675	21634
40.00	871485	21787
80.00	1483247	18541
120.00	2084890	17374

Average response (or calibration) factor



$$RF_{AVE} = \sum \frac{RF_i}{n}$$

$$RF_{AVE} = \frac{(19173 + 20917 + 21136 + 21634 + 21787 + 18541 + 17374)}{7} = 20080$$

$$Y = 20080x$$

Important:
Each point has the same weight in determining the RF_{AVE}, regardless of concentration

Evaluating the Average RF calibration



$$Std\ Dev = \frac{\sqrt{\sum(RF_i - RF_{AVE})^2}}{n-1}$$

$$RSD\% = \frac{Std\ Dev}{RF_{AVE}} \times 100$$

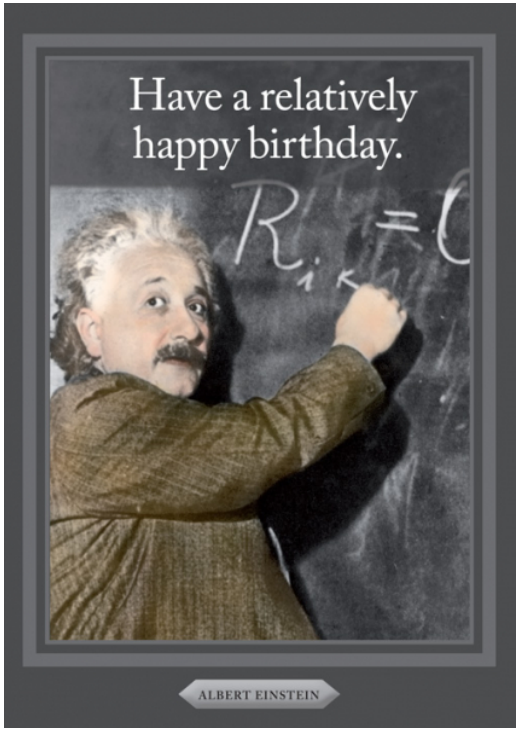
$$RSD\% = 8.5\%$$

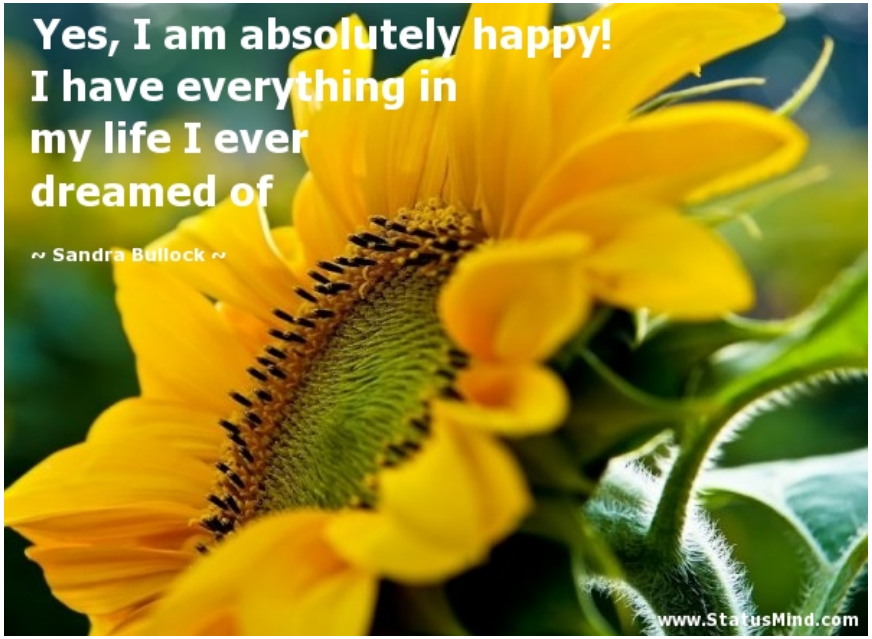
Important:
RSD is relative – the error at each point is a percentage of the concentration.
An error of 10ppb at 100ppb has the same weight as an error of 0.1ppb at 1.0 ppb.

What do we want to minimize?

Average RF minimizes % error

We want to minimize error, but do we want to minimize absolute error or relative error?





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

Average RF minimizes the relative error
%RSD measures the average relative error

Which is what we want!

Are we done?

Let’s go back to our curve

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

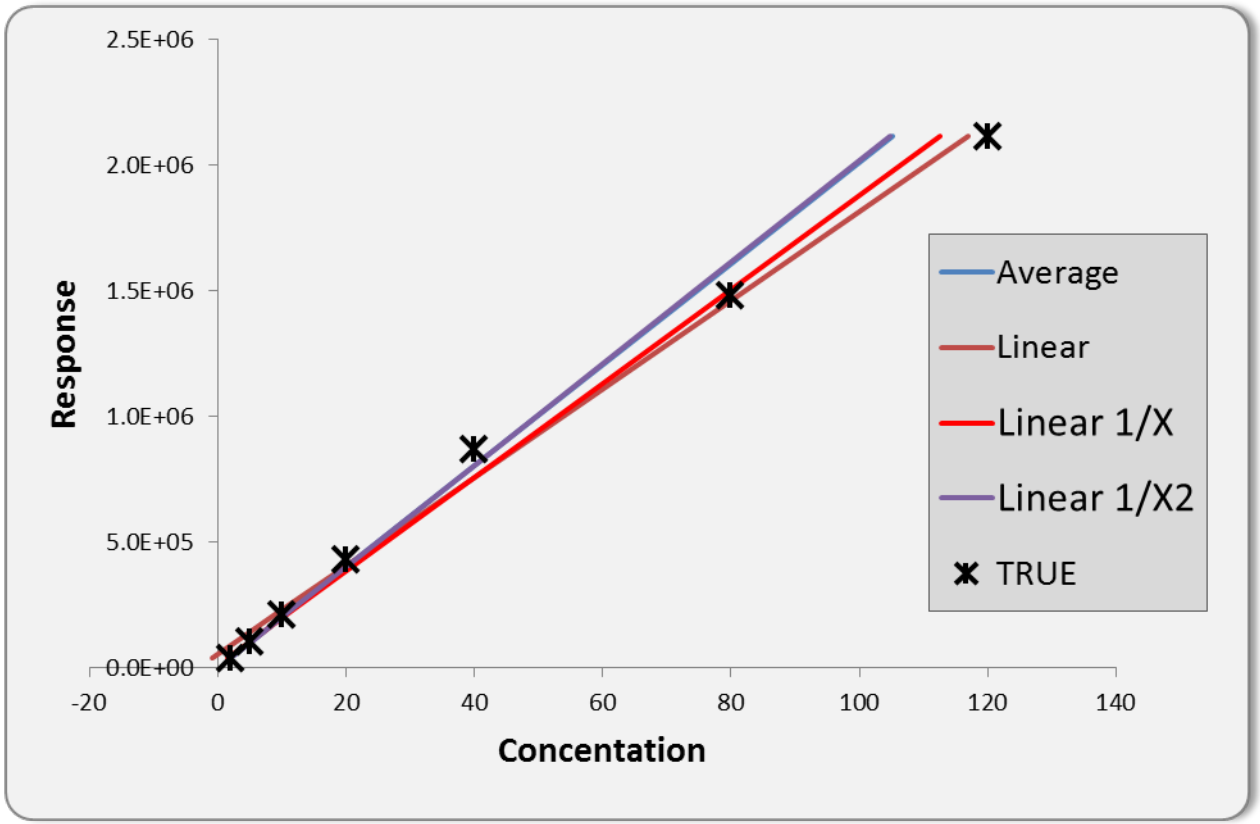
Acceptance criteria

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

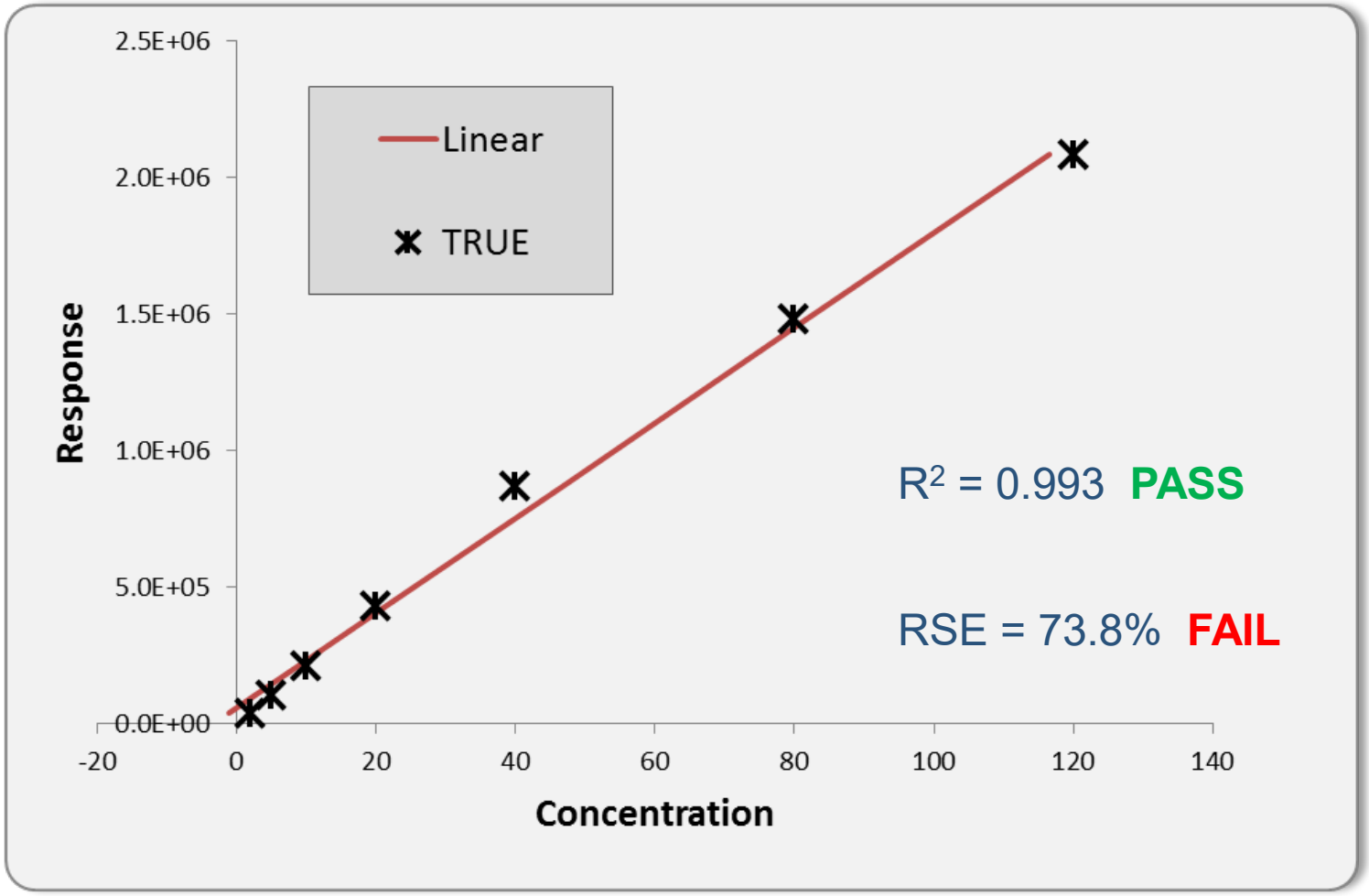
  And
  - RSE $\leq$ 20%

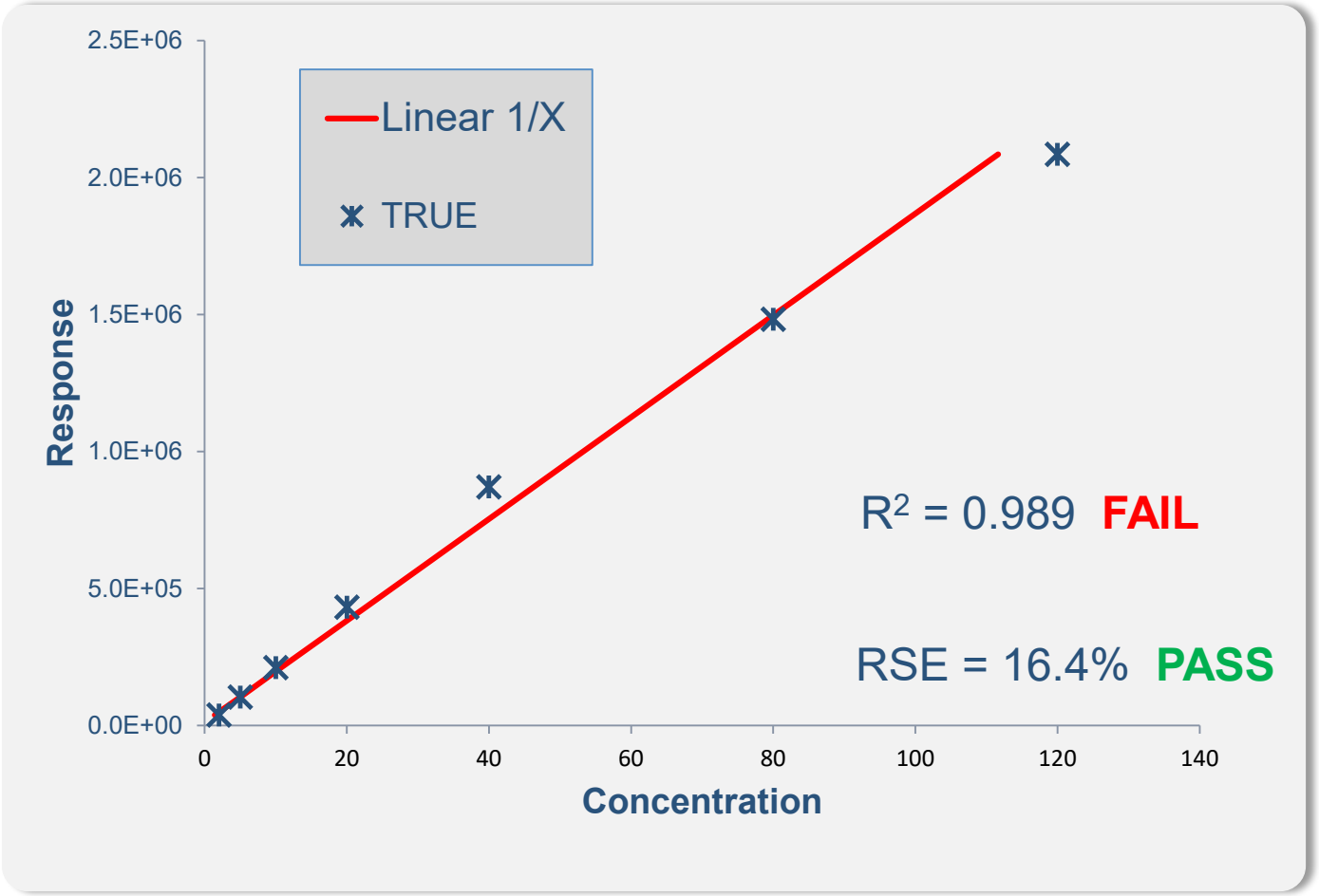
  Or
  - Mid point relative error < 20%
  - Low point relative error < 30%

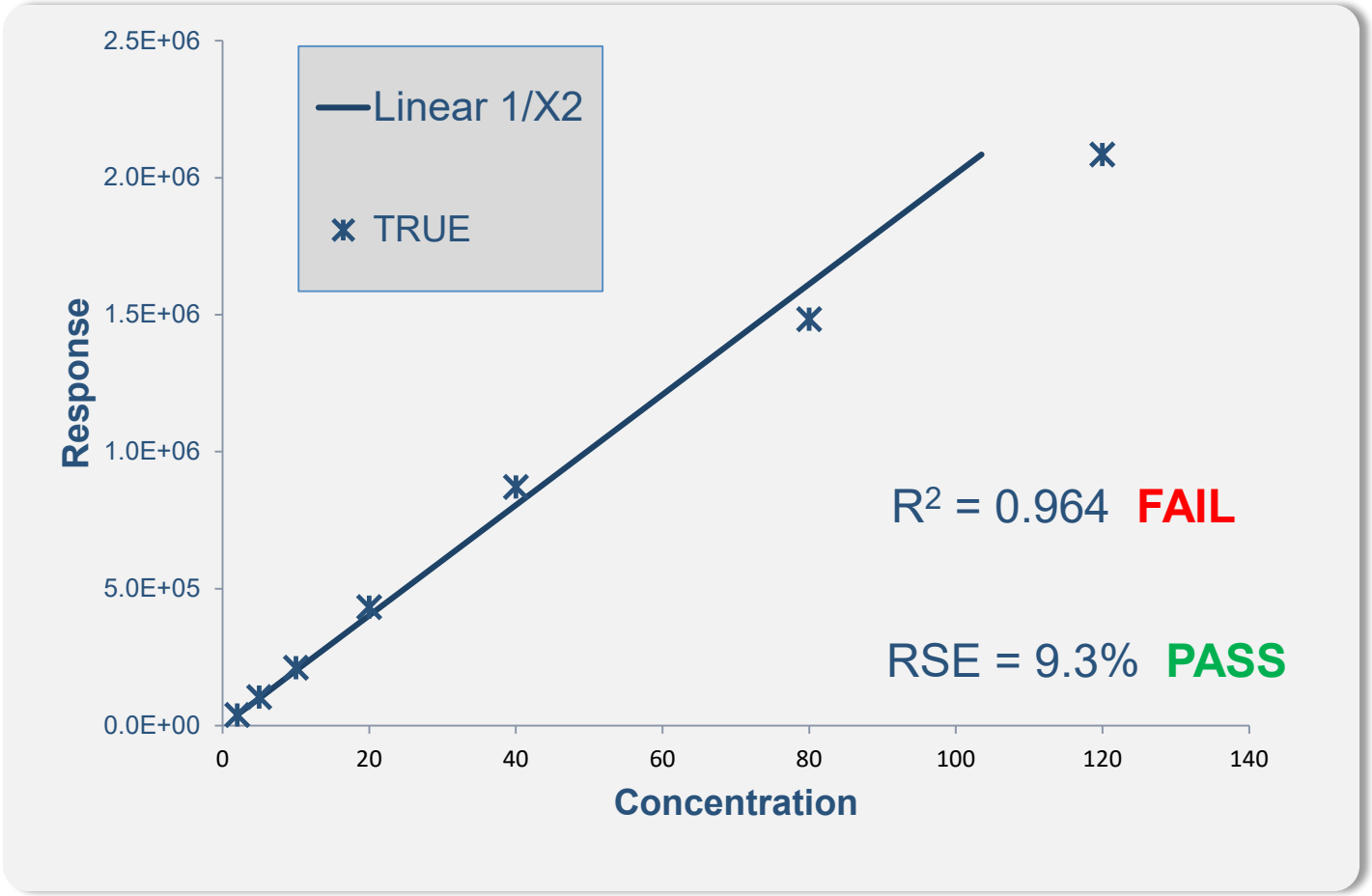
Let's analyze a calibration



Linear regression calibration

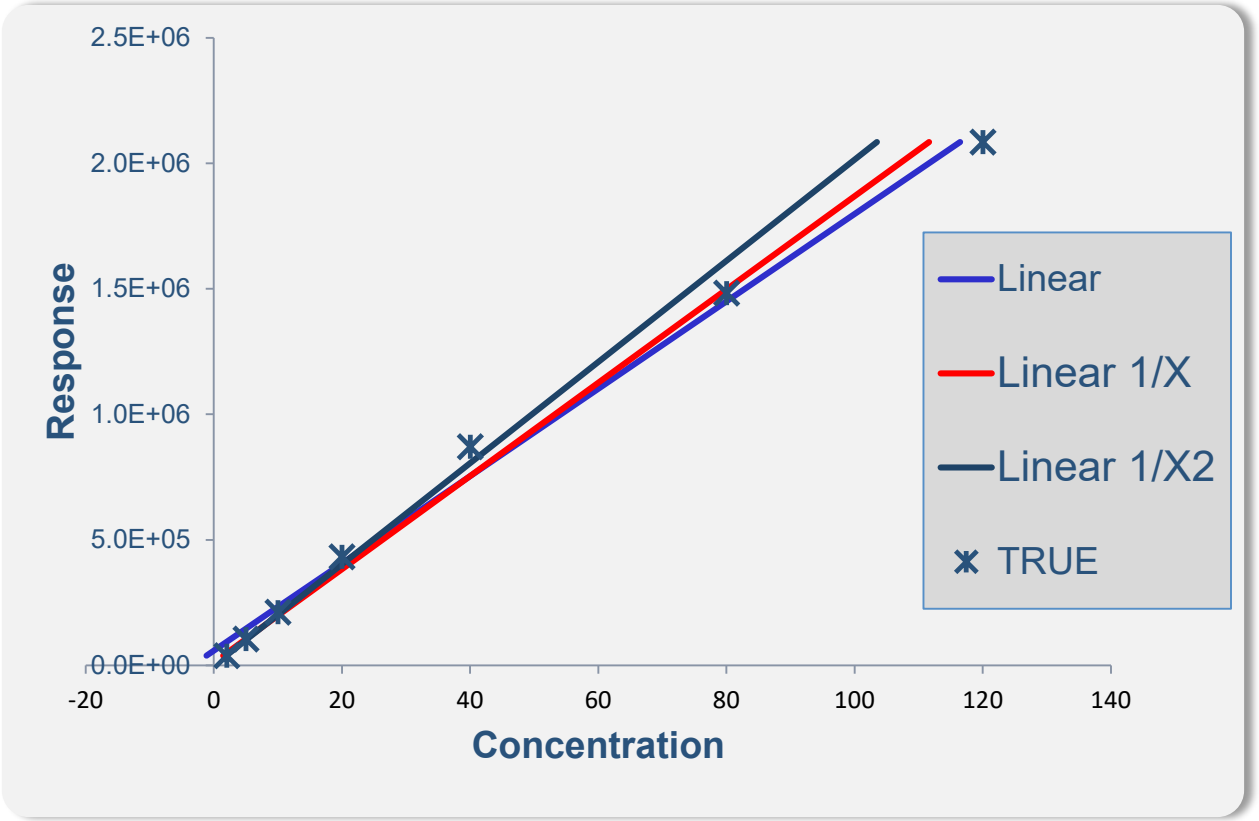


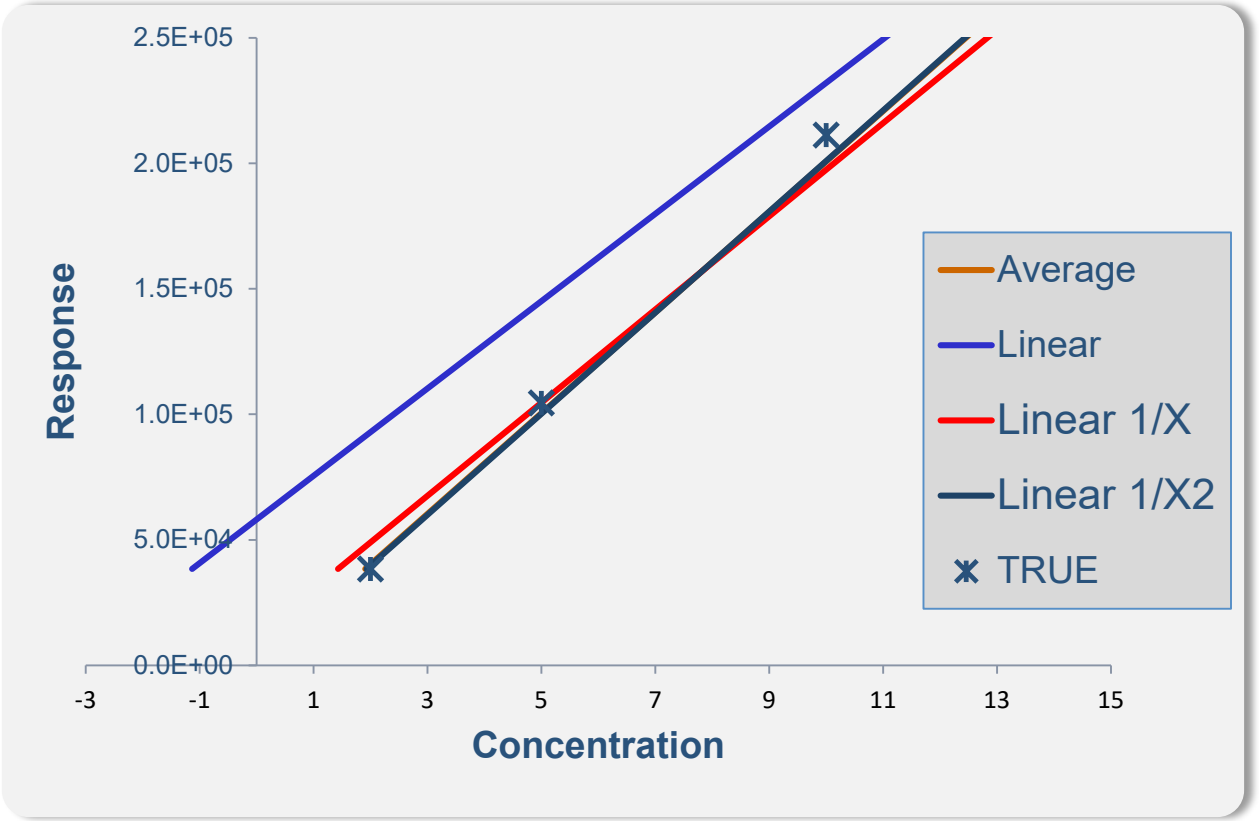

2.5E+06
2.0E+06
1.5E+06
1.0E+06
5.0E+05
0.0E+00
Response
Linear 1/X
TRUE
R² = 0.989 FAIL
RSE = 16.4% PASS
0
20
40
60
80
100
120
140
Concentration


2.5E+06
2.0E+06
1.5E+06
1.0E+06
5.0E+05
0.0E+00
Response
Linear 1/X2
TRUE
0
20
40
60
80
100
120
140
Concentration
R2 = 0.964 FAIL
RSE = 9.3% PASS



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


2.5E+06
2.0E+06
1.5E+06
1.0E+06
5.0E+05
0.0E+00
Response
-20
0
20
40
60
80
100
120
140
Concentration
Linear
Linear 1/X
Linear 1/X2
TRUE

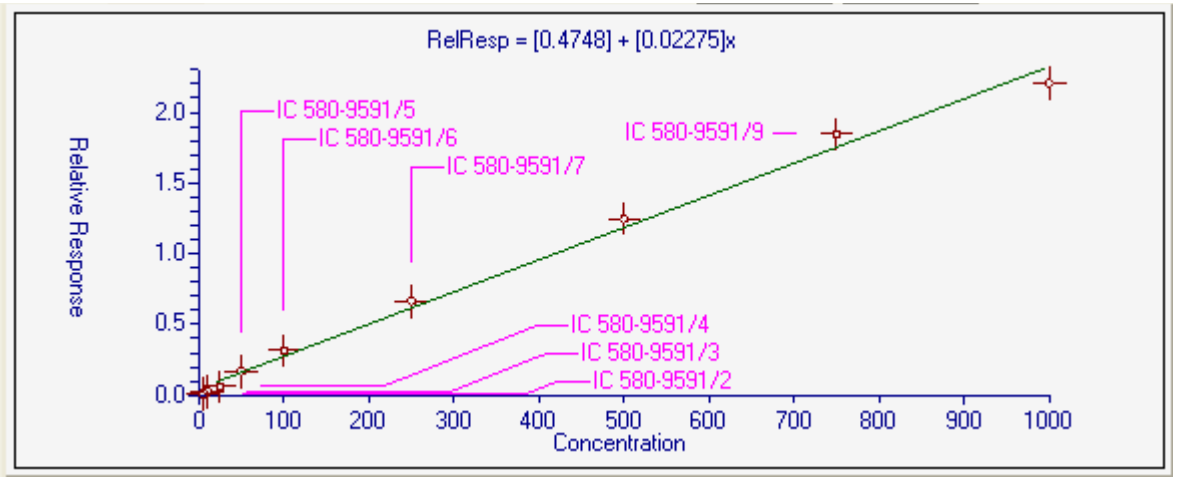

2.5E+05
2.0E+05
1.5E+05
1.0E+05
5.0E+04
0.0E+00
Response
-3
-1
1
3
5
7
9
11
13
15
Concentration
Average
Linear
Linear 1/X
Linear 1/X2
TRUE

Correlation coefficient gets better

Curve quality gets worse

Calibration issues





$r = 0.997$, $r^2 = 0.994$

RSE = 179%

Calibration Standard Levels

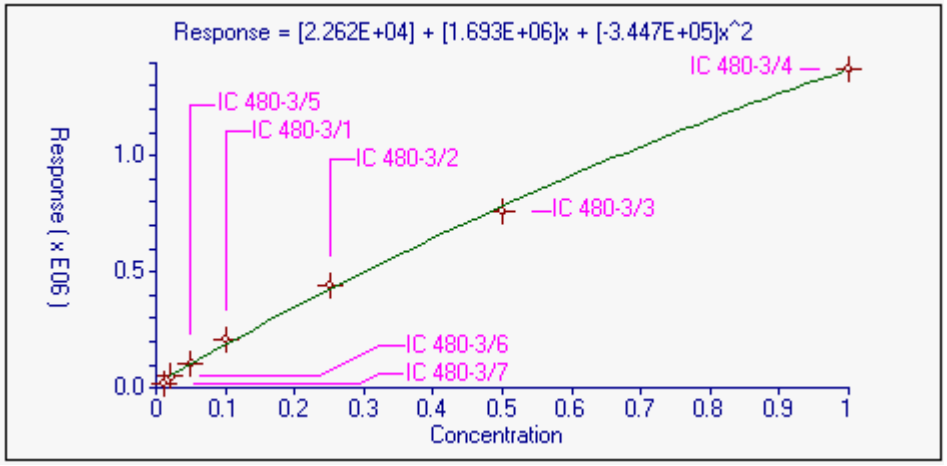
Level	Used	Amount	Area	ISArea	%Error
IC 580-9591/2	☑	5	1348	618332	421.63
IC 580-9591/3	☑	10	3250	647316	198.43
IC 580-9591/4	☑	25	7697	646400	78.87
IC 580-9591/5	☑	50	23729	700099	7.13
IC 580-9591/6	☑	100	47131	748204	17.47
IC 580-9591/7	☑	250	111297	833662	8.93
IC 580-9591/8	☑	500	229185	917698	5.52
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!

Can calculate R for Average



	Average	1/X	$1/X^2$	Unweighted
2	-4.52%			
5	4.17%			
10	5.26%			
20	7.74%			
40	8.50%			
80	-7.67%			
120	-13.48%			
RSD	**8.5%**			
R^2	**0.983**			

RSD = **PASS**
Coeff Determination = **FAIL**

To get ridiculous for a moment



Conc	Resp
1	0.00
2	0.00
3	0.00
4	0.00
5	0.00
10	0.00
100	117
slope	0.81564
corr	0.99679
int	4.16667

Because in **absolute** terms, nothing is far off

Correlation coefficient

“Very common mistakes in the analytical calibration process are the use of correlation and or determination coefficients…

Evaluation of analytical calibration based on least squares linear regression for instrumental Techniques, Francisco Raposo, TrAC 77, Match 2016, Pages 167-185

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**

Correlation Coefficient

“One practice that should be discouraged is the use of the correlation coefficient as a means of evaluating goodness of fit of linear models”

- **Van Arendonk and Skogerboe, Anal. Chem. 53, 1981, 2349-2350**

Taylor, Statistical Techniques for Data Analysis, 1990

"The author has seen cases where a correlation coefficient of 0.997 was believed to be a better fit than 0.996 of a 5 point calibration curve. One can even find requirements in quality assurance plans to recalibrate if the correlation coefficient is less than 0.995!"

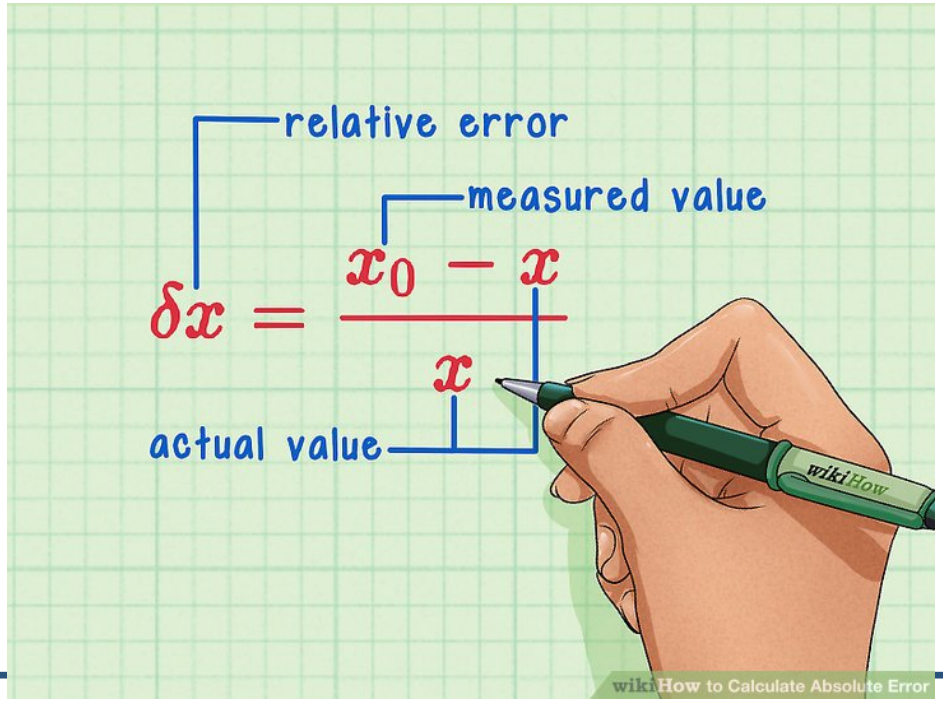
RSE Status

- 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

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!

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?

- Raise your hand if you think calibration is an important part of an analytical chemical method

- Keep it raised if you think that a good measure of whether a calibration is acceptable is important

- Keep it raised if you think a bad measure of calibration acceptability is a problem


THEN--AS IF AN IRRESISTIBLE FORCE WERE MEETING AN IMMOVABLE OBJECT...
KA-
POW!!