MTR SOFTWARE
77 Carlton Street, Suite 808
Toronto ON Canada
M5B 2J7
http://www.fitall.com/
Tel: (416) 596 - 1499
email: support@fitall.com

The Importance of Weighting

Purpose:

The purpose of this document is to:

1. Provide some background that explains why you should consider using weighting when doing a regression analysis.

2. Provide examples, using ***FitAll***, that illustrate the importance of weighting.

Background:

The objective of a regression analysis is to draw a line, which may be a straight line or a curved one, through a series of data points that minimizes the difference between the calculated values and the actual, measured values and report the values of the adjustable parameters that were required to obtain the "best fit".

The most common criterion used to determine the "best" fit is to minimize the sum of the squares of the difference between the observed and calculated values.

If no weighting, or more properly equal weighting, is used when doing the regression analysis the value of each data point is assumed to be as reliable as any of the other data points. This assumption may or may not be true.

Measuring Data Reliability:

There are several ways to measure or estimate the relative reliability of the data points.

1. A typical method, but not as commonly used as it should be, of determining the relative reliability of the measured values of the data points is to do multiple experiments and measurements of each data point and then average the values and calculate the standard deviation of the average value.

   A variation of this is to repeat each experiment several times and analyze all of the data from all of the experiments together.

2. A commonly used method assumes that the expected error of a measured value is directly proportional to its absolute value.

3. A less commonly used method is to recognize that most measurement equipment, such as voltmeters, have a measurement precision that is directly proportional to its full-scale range setting.

   For example, for an instrument that has a stated precision of its full-scale setting of 1%, the precision of a measured value between 1 and 9.99 would be ±0.1 and between 10 and 99.99 would be ±1.

   Note:
   In the "olden" days it was necessary to manually change the full-scale range setting of the measurement instrument. More recently; that is, more than a few years before either you or I were born, most measurement instruments automatically change their full-scale range setting in response to the input that is being measured.

4. Last, but not least, it is important to recognize that all of the most common regression analysis programs, ***FitAll*** included, assume that the measurement errors follow a Gaussian (Normal) distribution.

   If you are doing a "counting" experiment, such as counting the number of particles that are emitted by a decaying radionuclide, it is important to note that measurement errors in such data follow a Poisson, rather than a Gaussian, distribution and the regression analysis should use a weighting factor of 1/Y, in which Y is the measured count.

When using weighting during a regression analysis, the absolute values of the weights are not important. What is important is the *relative* values of the weights, because the weighting factors are used to measure the *relative* reliability of the measured data point values.

Transforming the Fitting Function and the Data:

In the really, really olden days; that is, before the advent of computers, it was common practice to transform an intrinsically non-linear equation that described the data to a linear form that could be fit using linear regression instead of non-linear regression.

This was done to simplify the calculations. Hey, even simple calculations can be quite time consuming when one is using pencil and paper!
[If you are confused, check Wikipedia or some other authoritative source, such as a grandparent, for a description of pencils and paper].

Unfortunately, this practice has continued to this day, even though it is no longer necessary. Most importantly, more often than not the weighting factors are not transformed and weighting is not used. The result of this is that the regression results will be biased.

Example 1: Transforming a Non-linear function to a Linear Function:

Purpose:

The purpose of this example is to illustrate how transforming a function may lead to results that are not what they should be unless weighting is used.

The differences observed in this example are not particularly large, but still illustrate the point.

Background:

First-order exponential decay functions describe many different types of physical processes, such as chemical reactions, the decay of radio nuclides, the charging and discharging of a capacitor in a resistor-capacitor electrical circuit and some simple diffusion processes.

The simplest form of the equation is: $Y = P1 * e^{-P2*X}$, in which P1 and P2 are the parameters that are to be determined.

This equation is non-linear in its parameters, P1 and P2.
That is, it is not of the form Y = P0 + P1 * X1 + P2 * X2 + ….

This equation can be transformed to one that is linear in its (new) parameters by taking the natural logarithm of both sides of the equation to get: (Ln|Y|) = (Ln|P1|) - P2 * X.

The estimated error (precision) of Ln|Y| is given by: ∂(Ln|Y|) = -∂Y/Y.

What the above means is that if we plot Ln|Y| vs. X and draw a straight line through the data the slope of the line will be equal to -P2 and the intercept will be equal to Ln|P1|.

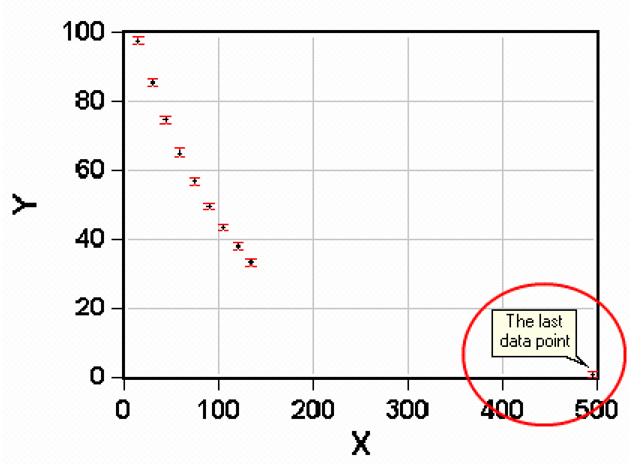
The Data:

<u>The Original Data:</u>

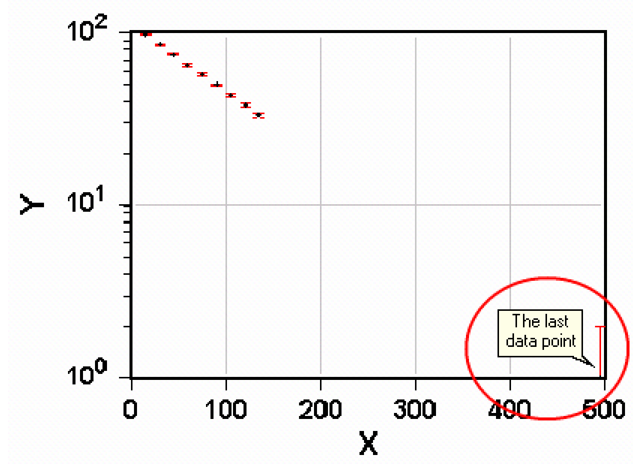
Pt#	X	Y	SigmaY
1	15	97.9	1.0
2	30	85.5	1.0
3	45	74.7	1.0
4	60	65.3	1.0
5	75	57.0	1.0
6	90	49.9	1.0
7	105	43.6	1.0
8	120	38.1	1.0
9	135	33.3	1.0
10	495	1.0	1.0

Note:
All of the estimated measurement errors in the Y-values, SigmaY, are equal.

The data graph, including the error bars, looks like this:



or

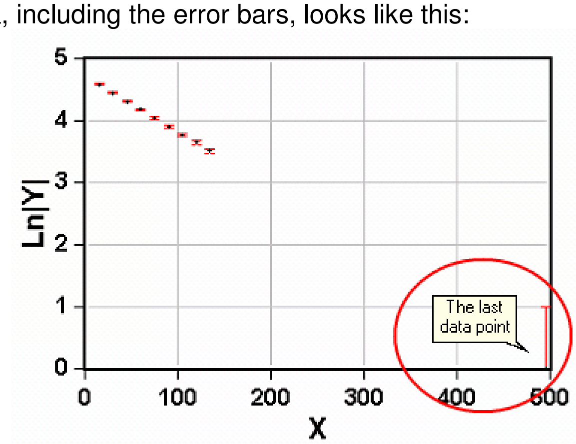


Note:
When plotted using a logarithmic scale the error bar for the last data point appears much larger than the error bars for the other data points.

The Transformed Data:

Pt#	X	Ln|Y|	SigmaLn|Y| = (SigmaY)/Y
1	15	4.5839	0.0102
2	30	4.4485	0.0117
3	45	4.3135	0.0134
4	60	4.1790	0.0153
5	75	4.0431	0.0175
6	90	3.9100	0.0200
7	105	3.7751	0.0229
8	120	3.6402	0.0262
9	135	3.5056	0.0300
10	495	0.0000	1.0000

A graph of the transformed data, including the error bars, looks like this:

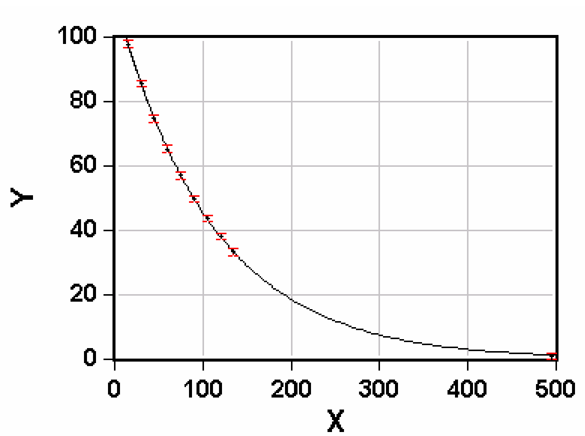


The Fits:

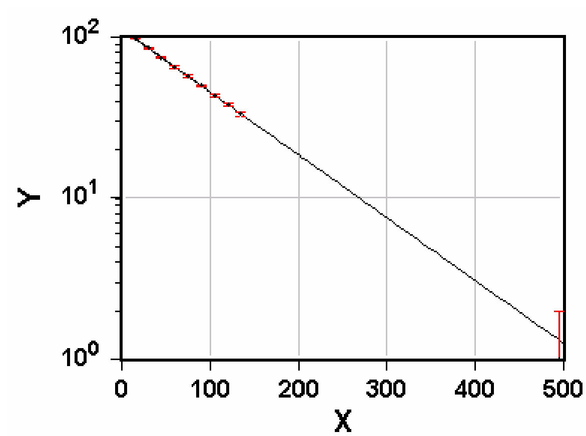
Fit of the Original Data with Equal Weighting:

Function:	0001..1st order exponential + bkgrnd: Y=P1*exp(-P2*K1*X) + Sum{A[i]*X^(i)}
Analysis Method:	Non-Linear Least Squares (nlls)
Analysis Range:	1 to 10 of 10
Weighted as:	**1**
Variance:	0.0130275448919275
Std. Dev. of Fit:	0.114138270934544
Iterations:	4

Parameter	Value	Std. Dev.	RSD /%
1	1.120E+0002	1.192E-0001	0.11
2	**8.992E-0003**	**1.753E-0005**	0.19



or



Notes:

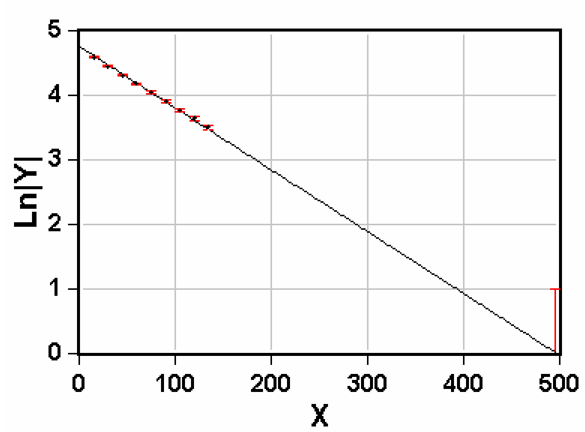
1. Equal weighting was used because the measured errors in the Y-values are the same; that is, one.

2. The value of parameter 2 is (8.99 ± 0.018) x 10^{-3}.

Fit of the Transformed Data with Equal Weighting:

Function:	0008..Multiple Linear: Yi = P0 + Sum{P[j]*Xij}
Analysis Method:	Linear Least Squares (lls)
Analysis Range:	1 to 10 of 10
Weighted as:	**1**
Variance:	0.000650923466870696
Std. Dev. of Fit:	0.0255132018153484

Parameter	Value	Std. Dev.	RSD /%
1	4.761E+0000	1.081E-0002	0.23
2	**-9.579E-0003**	**6.147E-0005**	0.64



Notes:

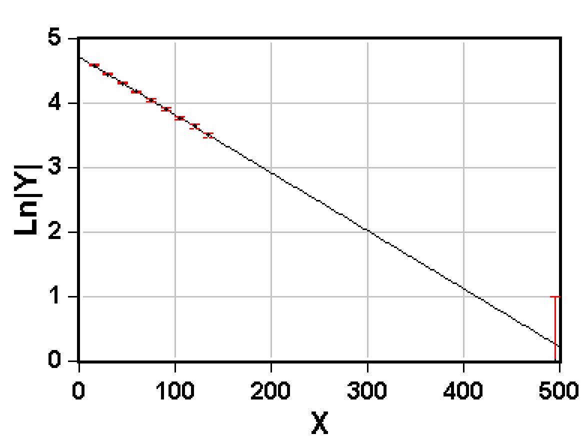
1. Equal weighting was used, because this is often done in such circumstances, to illustrate that the results differ significantly from those obtained using the original data and the function that describes that data.

2. The value of parameter 2 is **(9.58 ± 0.061) x 10^{-3}**, which is significantly different from that obtained in the first fit, **(8.99 ± 0.018) x 10^{-3}**.

Fit of the Transformed Data WITH Weighting:

Function:	0008..Multiple Linear: Yi = P0 + Sum{P[j]*Xij}
Analysis Method:	Linear Least Squares (lls)
Analysis Range:	1 to 10 of 10
Weighted as:	**1/(SigmaLn|Y|)^2**
Variance:	2.76293037573828E-006
Std. Dev. of Fit:	0.00166220647807012

Parameter	Value	Std. Dev.	RSD /%
1	4.718E+0000	9.429E-0004	0.02
2	**-8.991E-0003**	**1.553E-0005**	0.17



Notes:

1. The fit was weighted using the transformed weighting factors.

2. The value of parameter 2 is **(8.99 ± 0.016) x 10^{-3}**, which is virtually identical to that obtained in the first fit, **(8.99 ± 0.018) x 10^{-3}**.

Example 2:Need for Weighting When Data Spans a Large Range:

Purpose:

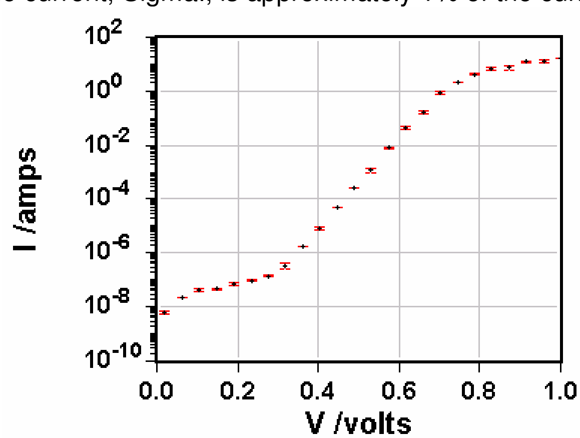
The purpose of this example is to illustrate that the use of equal weighting of data, in which the Y-values span several orders of magnitude, can result in misleading results.

The Data:

Pt#	V /volts	I /amps	SigmaI
1	2.041E-0002	6.064E-0009	7.398E-0010
2	6.300E-0002	2.170E-0008	6.877E-0010
3	1.056E-0001	4.008E-0008	4.816E-0009
4	1.482E-0001	4.620E-0008	3.544E-0009
5	1.908E-0001	7.123E-0008	5.739E-0009
6	2.334E-0001	9.653E-0008	8.473E-0009
7	2.760E-0001	1.390E-0007	8.403E-0009
8	3.185E-0001	3.378E-0007	6.774E-0008
9	3.611E-0001	1.707E-0006	2.869E-0008
10	4.037E-0001	7.852E-0006	1.002E-0006
11	4.463E-0001	4.879E-0005	1.583E-0006
12	4.889E-0001	2.661E-0004	1.193E-0005
13	5.315E-0001	1.171E-0003	1.991E-0004
14	5.741E-0001	8.023E-0003	6.525E-0004
15	6.167E-0001	4.358E-0002	4.474E-0003
16	6.593E-0001	1.694E-0001	2.469E-0002
17	7.019E-0001	8.550E-0001	8.759E-0002
18	7.445E-0001	2.146E+0000	8.879E-0002
19	7.870E-0001	4.133E+0000	1.741E-0001
20	8.296E-0001	7.006E+0000	7.743E-0001
21	8.722E-0001	7.362E+0000	1.359E+0000
22	9.148E-0001	1.222E+0001	8.715E-0001
23	9.574E-0001	1.280E+0001	1.262E+0000
24	1.000E+0000	1.705E+0001	2.163E-0001

Notes:

1. The current, I, is the Y-value and ranges over ten orders of magnitude.
2. The error in measuring in the current, SigmaI, is approximately 1% of the current's value.



The Fits:

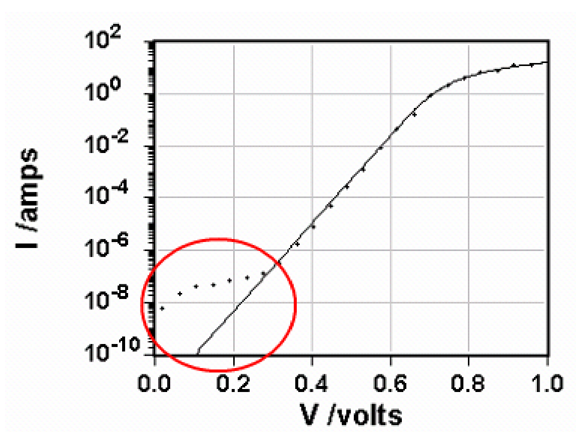
The equation that describes the data in this example is rather complex; however, its complexity is not relevant to this example. Suffice it to say that the current, I, is the Y-value and the voltage, V, is the X-value in an equation of the general form: Y = f(X).

The Fit: Using Equal Weighting:

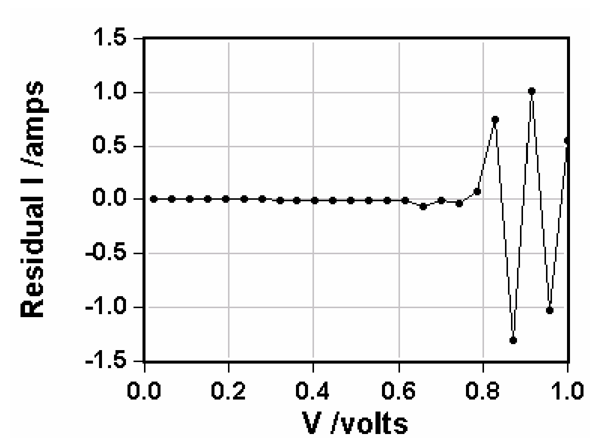
1.

Function:	0501..Dark I-V: Ideal Y = P2*{exp[K1*(X-Y*P3)/K2]-1} + (X-Y*P3)/P1
Analysis Method:	Non-Linear Least Squares (nlls)
Analysis Range:	1 to 24 of 24
Weighted as:	**1**
Variance:	0.221816114848999
Std. Dev. of Fit:	**0.470973581901362**
Iterations:	26

Parameter	Name	Value	Std. Dev.	RSD /%
1	Rsh /Ω	**6.000E+0014**	**8.281E+0019**	13801977.76
2	Io /A	1.898E-0012	2.328E-0019	0.00
3	Rs /Ω	1.422E-0002	2.861E-0010	0.00
Constant				
1	q/k /C°K/J	1.160E+0004		
2	t /°C	2.500E+0001		



and



Notes:

1. The fit graph clearly shows that the calculated line is nowhere near the measured values at low voltages.

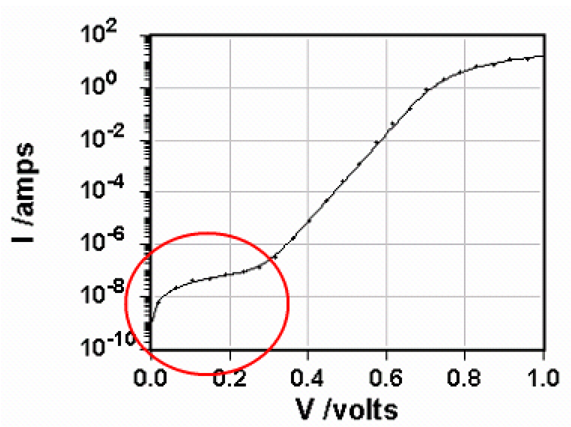
2. From the residuals graph it is difficult, actually impossible, to tell that for low voltages the calculated values are not randomly distributed about the zero line.

3. The standard deviation of parameter 1, Rsh, is much larger than the value of the parameter itself. – Clearly, something is amiss.

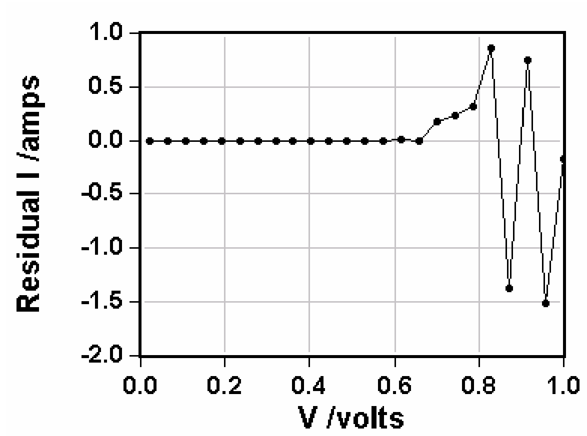
The Fit: Weighting the analysis as $1/(\text{SigmaY})^2$, in which SigmaY is the measurement error in the Y-value:

Function:	0501..Dark I-V: Ideal Y = P2*{exp[K1*(X-Y*P3)/K2]-1} + (X-Y*P3)/P1
Analysis Method:	Non-Linear Least Squares (nlls)
Analysis Range:	1 to 24 of 24
Weighted as:	**1/SigmaY^2**
Variance:	1.5177320546999E-017
Std. Dev. of Fit:	**3.89580807368626E-009**
Iterations:	7

Parameter	Name	Value	Std. Dev.	RSD /%
1	Rsh /Ω	**2.996E+0006**	**7.509E+0004**	2.51
2	Io /A	1.308E-0012	1.837E-0014	1.40
3	Rs /Ω	1.299E-0002	1.741E-0004	1.34
Constant				
1	q/k /C°K/J	1.160E+0004		
2	t /°C	2.500E+0001		



and



Notes:

1. In the fit graph, the calculated curve **does** go through the data points, even at low voltages.

2. The residuals appear to be randomly distributed around the zero line, although, because of the graph's scale, it is rather difficult to see this for the low voltage values.

3. The standard deviation of the fit, $3.89 * 10^{-9}$, is much smaller than that, 0.47, obtained when equal weighting was used.

Conclusions:

1. When doing a regression analysis it is important to consider weighting the data.

2. Ideally, the weighting factors used should be measured values.

3. If measured values of the standard deviation of each the Y-values are not available one should carefully consider how the data was collected and apply an appropriate weighting scheme.

4. If the function that describes the data is transformed from one form to another, it is important to apply the appropriate transformation to the estimated errors and use weighting when fitting the transformed data to the transformed function.

