Experiment 0

Treatment of Real Experimental Data: statistical methods and least squares fitting

Objective .

To develop proficiency at employing the least-squares curve fitting method to fit real experimental data.

Introduction .

As scientist we observe a physical phenomenon, and then propose a theory to explain and predict the behavior of that phenomenon. This theory will often come in the form of a mathematical model or expression. In this lab you will implement a least-squares technique to perform ‘curve fitting’ on real experimental data. The purpose of least-squares ‘curve fitting’ is to match your proposed theoretical mathematical model to the measured experimental data as closely as possible by finding the correct numerical value of the function’s variables.

The Least-Squares Technique: Minimizing Chi-Squared

Let’s start with some experimental data. Figure 1a shows a series of data points, values of y as a function x. The underlying phenomena that the data represents, that is our mathematical model, would predict a linear relationship between x and y. Notice even though the relationship between x and y should be linear the data contains some noise and is not in a perfect straight line. Figure 1b shows the same series of data with the ‘best fit’ line (in red) passing through the data. Notice that the ‘best fit’ line does not intersect with any one data point yet, it represent the “best” approximation of the data’s trend.

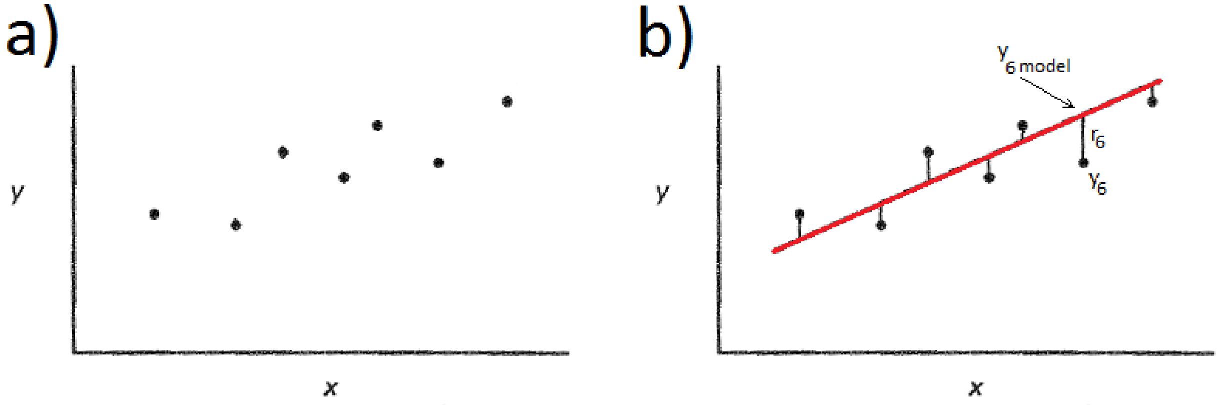


Figure 1. A straight line is assumed to fit seven data points. r_6 is the “residual” of the 6th data point; the vertical line show the residual of the data point. This figure was taken from: Experiments in Physical Chemistry, Halpern, Arthur M., page 18

How do we determine what the ‘best fit’ is (Figure 1b)? The fit is best if it minimizes the value of Chi-squared. Chi (also known as the residual) is the difference between the measured y value and the model y value. That is:

$$r_i = y_i - y_{model} \qquad \text{equ. 1}$$

For any one observed ‘y_i’ value there is a corresponding ‘y_m’ value for the model function or ‘best fit’ line. We want the ‘y_m’ value to match as closely as possible the corresponding ‘y_i’ value. However,

To find the ‘best fit’ function we must calculate the residual for each point. Notice that the magnitude of some residuals will be positive and some will be negative. Before we sum the residual values we must square them so that each residual contributes in a positive fashion to the overall error. Therefore Chi-Squared is defined as the sum of the squared residuals, or:

$$sum\ of\ the\ squared\ residuals = \chi^2 = \sum_{i=1}^{n} \frac{r_i^2}{\sigma_i^2} \qquad \text{equ. 2}$$

where r_i is the residual for data point i, and σ_i (sigma) is the variance for the data point y_i. The ‘best fit’ function will be the one that that minimizes all the values of Chi-squared (χ^2).

It is important to include weighting factors (or variance) in least-squared analysis if the uncertainty in you dependent variables are not constant. This might happen if you collect data using two different methods (each with its own unique uncertainty) and combine that data in one least-squares analysis. In general this will not be an issue in this class (that is sigma is the same value for all your data points), but you should be aware of this possibility.

For least-square simulations of non-linear functions it is important that your ‘initial guesses’ for your parameter values are close to the true values. The ‘error surface’ (that is the sum of Chi-squared (χ^2)) for non-linear functions may contain both local and global minima. If you plot the sum of Chi-squared (χ^2) as a function of all possible parameter values you may end up with a plot like Figure 2.

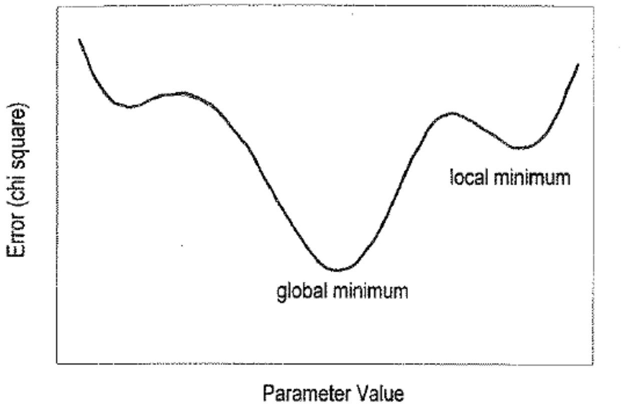


Figure 2. Cross section of a χ^2 surface showing a local minimum in addition to a global minimum. An initial guess that starts the search algorithm at or near the local minimum will probably return a parameter that is not truly, or globally, optimized.

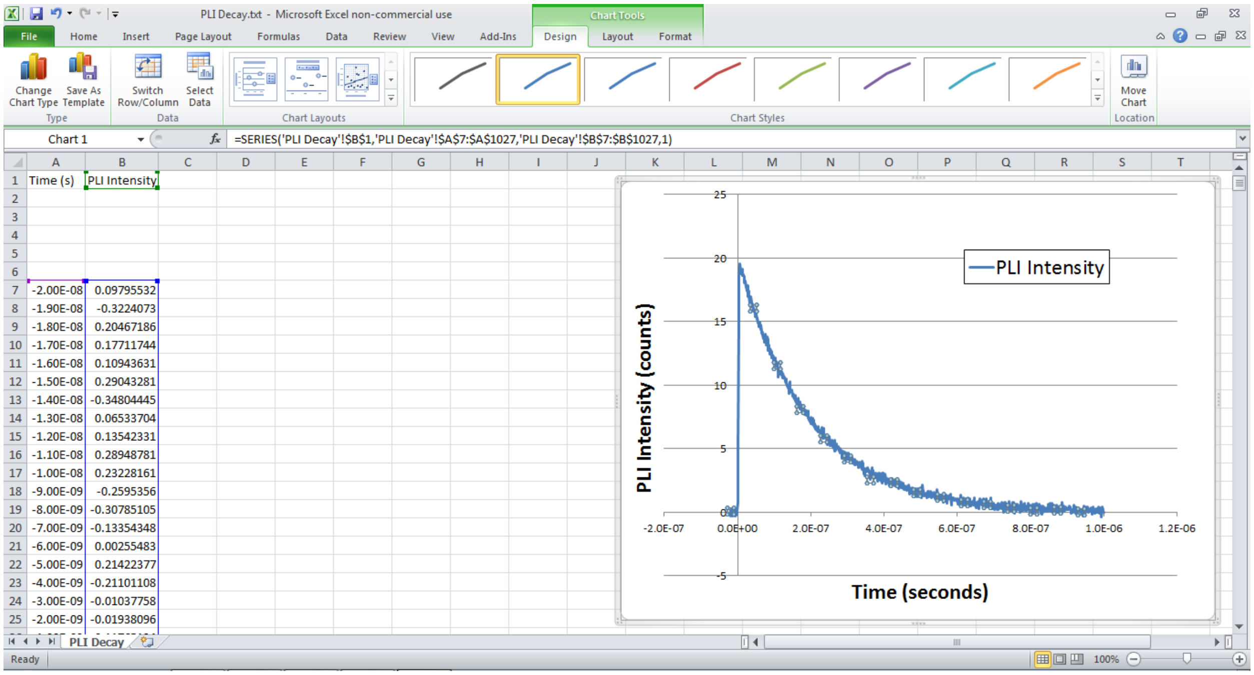
Now we will use practice implementing least-squares 'curve fitting' to match a proposed mathematical model with experimental data.

Experimental Apparatus .

For this experiment you will need a computer with a spread sheet application (like Excel) with and 'optimization algorithm' like the 'Solver Tool' for Excel. You will need to install the Solver add-in if you do not already have it.

Data Analysis .

Download the "PLI Decay.txt" file from this week's folder. This file should contain a column of x data and a column of y data. Plot the data in you preferred spreadsheet application program (like Excel). Label the x and y axes appropriately



=SERIES('PLI Decay'!$B$1,'PLI Decay'!$A$7:$A$1027,'PLI Decay'!$B$7:$B$1027,1)

	A	B
1	Time (s)	PLI Intensity
2		
3		
4		
5		
6		
7	-2.00E-08	0.09795532
8	-1.90E-08	-0.3224073
9	-1.80E-08	0.20467186
10	-1.70E-08	0.17711744
11	-1.60E-08	0.10943631
12	-1.50E-08	0.29043281
13	-1.40E-08	-0.34804445
14	-1.30E-08	0.06533704
15	-1.20E-08	0.13542331
16	-1.10E-08	0.28948781
17	-1.00E-08	0.23228161
18	-9.00E-09	-0.2595356
19	-8.00E-09	-0.30785105
20	-7.00E-09	-0.13354348
21	-6.00E-09	0.00255483
22	-5.00E-09	0.21422377
23	-4.00E-09	-0.21101108
24	-3.00E-09	-0.01037758
25	-2.00E-09	-0.01938096

The data shown here represents the photo luminescence decay of some molecule. In this experiment the sample was illuminated with a pulsed laser and the photoluminescence emission from the sample was monitored as a function of time (the laser pulse width in this example would be ~1 nano second). You could conduct an experiment like this with one of those 'glowing star' decorations for bedroom ceilings. If you were in a dark room with the star then turned the lights on for a few seconds, then turned them off you would see the stars glow in the dark. The intensity of the stars 'glow' would decay much as the data above does, exponentially with time.

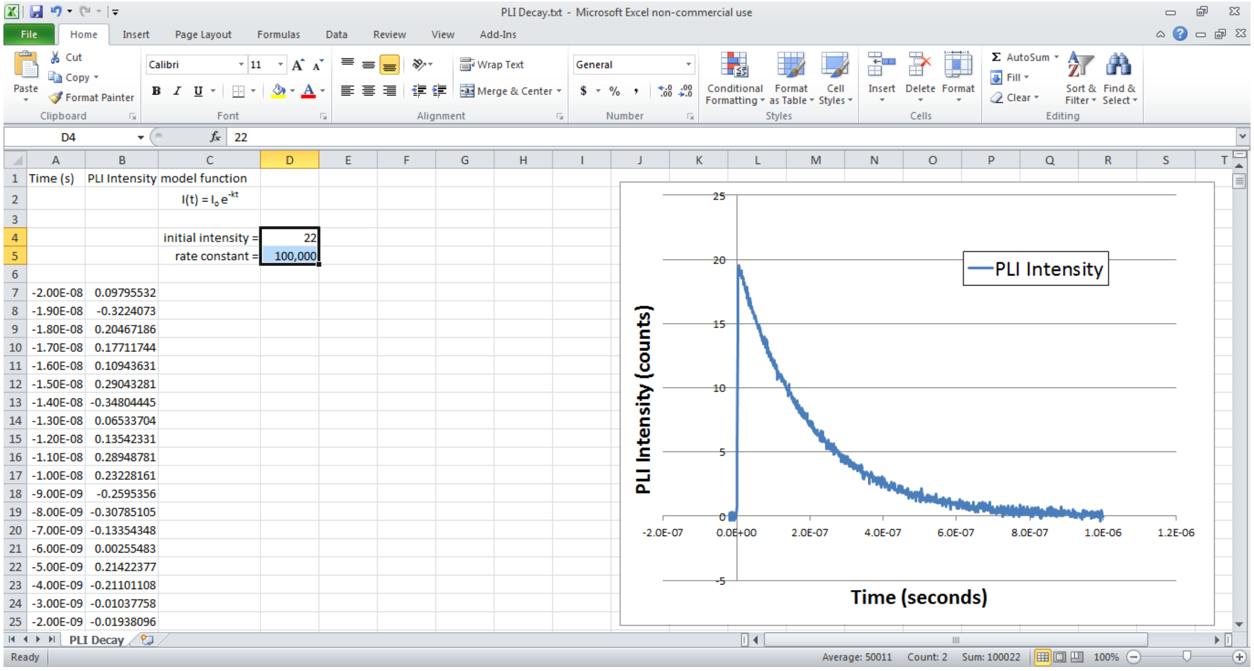
We hypothesize that the PLI intensity is decaying expediently with time (that is a first order reaction). We need to propose a model function to simulate the data. The equation for an expediential decay is:

$$I(t) = I_o e^{-kt} \qquad \text{equ. 3}$$

Where I(t) is the intensity at some time t, I_o is the initial intensity, and k is the rate constant.

We can see from the graph above that the initial intensity is ~20, but it is difficult to estimate the rate constant just by inspecting the graph. Our two variable in this least-squares fit will be the initial intensity and the rate constant.

In you spread sheet make another column for the 'model function' and type the model function in one of the boxes to help keep track of the variables. In two separate boxes type the labels 'initial intensity =' and 'rate constant ='. In the box to the right of thee entries place some initial guesses for the values. I am guessing 22 counts and 100,000 s^{-1}.

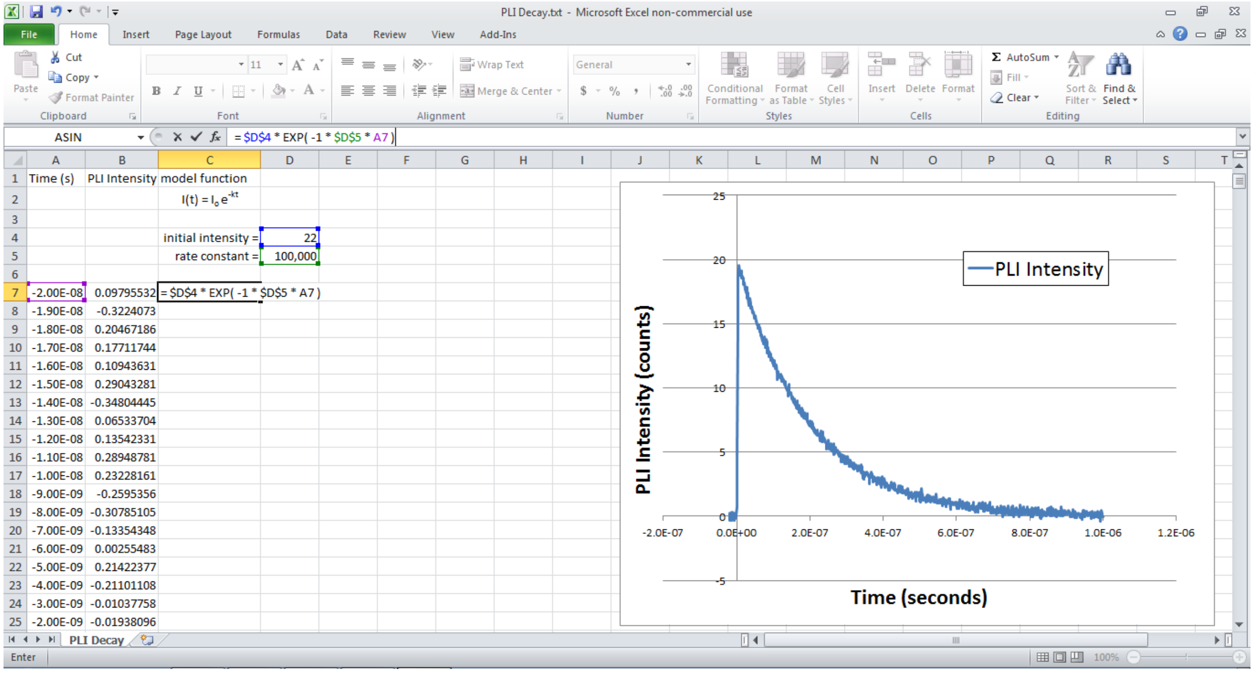


D4 f_x 22

	A	B	C	D
1	Time (s)	PLI Intensity	model function	
2			$I(t) = I_o e^{-kt}$	
3				
4			initial intensity =	22
5			rate constant =	100,000
6				
7	-2.00E-08	0.09795532		
8	-1.90E-08	-0.3224073		
9	-1.80E-08	0.20467186		
10	-1.70E-08	0.17711744		
11	-1.60E-08	0.10943631		
12	-1.50E-08	0.29043281		
13	-1.40E-08	-0.34804445		
14	-1.30E-08	0.06533704		
15	-1.20E-08	0.13542331		
16	-1.10E-08	0.28948781		
17	-1.00E-08	0.23228161		
18	-9.00E-09	-0.2595356		
19	-8.00E-09	-0.30785105		
20	-7.00E-09	-0.13354348		
21	-6.00E-09	0.00255483		
22	-5.00E-09	0.21422377		
23	-4.00E-09	-0.21101108		
24	-3.00E-09	-0.01037758		
25	-2.00E-09	-0.01938096		

Average: 50011 Count: 2 Sum: 100022 100%

We will now calculate our ‘model function’ using our ‘initial guesses’ of 25 and 100,000.



ASIN | = $D$4 * EXP(-1 * $D$5 * A7)

	A	B	C	D
1	Time (s)	PLI Intensity	model function	
2			$I(t) = I_o e^{-kt}$	
3				
4			initial intensity =	22
5			rate constant =	100,000
6				
7	-2.00E-08	0.09795532	= $D$4 * EXP(-1 * $D$5 * A7)	
8	-1.90E-08	-0.3224073		
9	-1.80E-08	0.20467186		
10	-1.70E-08	0.17711744		
11	-1.60E-08	0.10943631		
12	-1.50E-08	0.29043281		
13	-1.40E-08	-0.34804445		
14	-1.30E-08	0.06533704		
15	-1.20E-08	0.13542331		
16	-1.10E-08	0.28948781		
17	-1.00E-08	0.23228161		
18	-9.00E-09	-0.2595356		
19	-8.00E-09	-0.30785105		
20	-7.00E-09	-0.13354348		
21	-6.00E-09	0.00255483		
22	-5.00E-09	0.21422377		
23	-4.00E-09	-0.21101108		
24	-3.00E-09	-0.01037758		
25	-2.00E-09	-0.01938096		

PLI Decay

Graph your ‘model function’ on the same graph as the experimental data. Notice our initial guess were not very good; the rate constant is much too low and the initial amplitude is a little high.

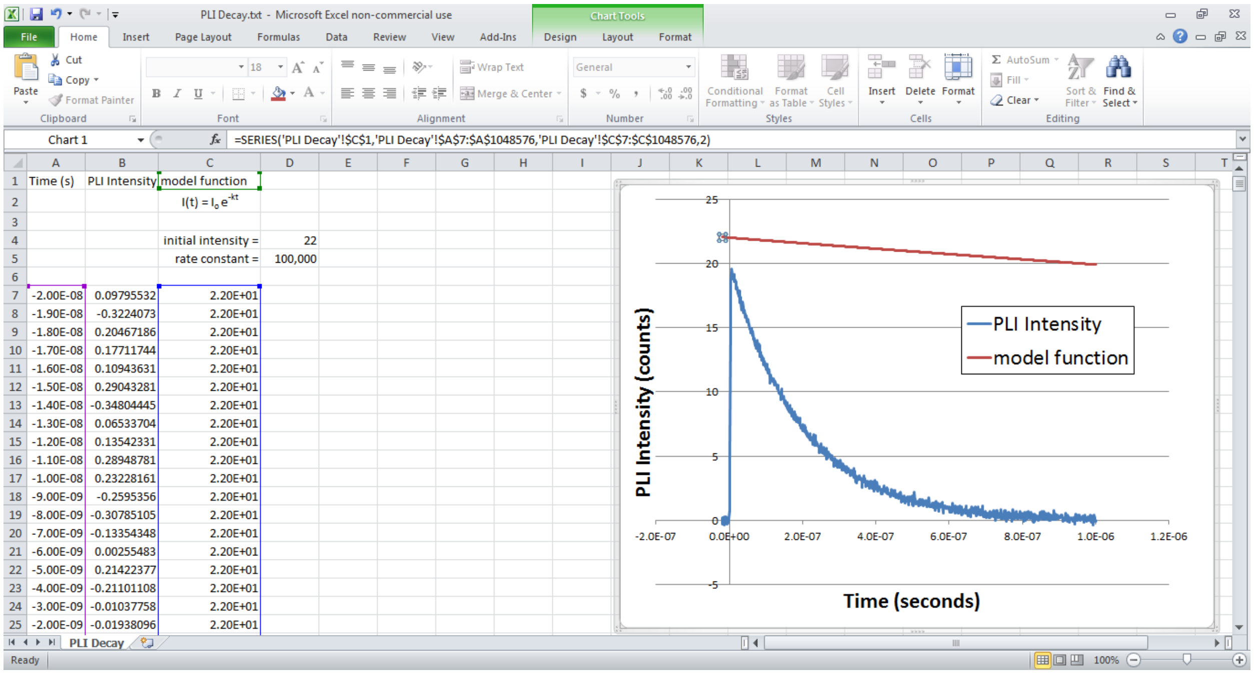
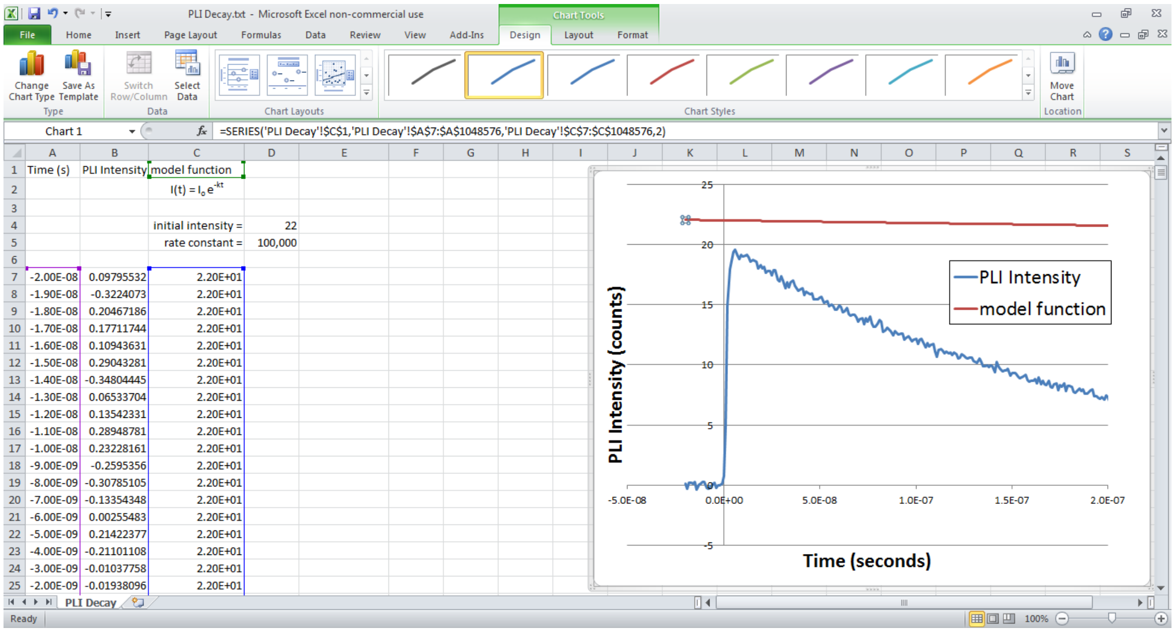


Chart 1 | =SERIES('PLI Decay'!$C$1,'PLI Decay'!$A$7:$A$1048576,'PLI Decay'!$C$7:$C$1048576,2)

	A	B	C	D
1	Time (s)	PLI Intensity	model function	
2			$I(t) = I_o e^{-kt}$	
3				
4			initial intensity =	22
5			rate constant =	100,000
6				
7	-2.00E-08	0.09795532	2.20E+01	
8	-1.90E-08	-0.3224073	2.20E+01	
9	-1.80E-08	0.20467186	2.20E+01	
10	-1.70E-08	0.17711744	2.20E+01	
11	-1.60E-08	0.10943631	2.20E+01	
12	-1.50E-08	0.29043281	2.20E+01	
13	-1.40E-08	-0.34804445	2.20E+01	
14	-1.30E-08	0.06533704	2.20E+01	
15	-1.20E-08	0.13542331	2.20E+01	
16	-1.10E-08	0.28948781	2.20E+01	
17	-1.00E-08	0.23228161	2.20E+01	
18	-9.00E-09	-0.2595356	2.20E+01	
19	-8.00E-09	-0.30785105	2.20E+01	
20	-7.00E-09	-0.13354348	2.20E+01	
21	-6.00E-09	0.00255483	2.20E+01	
22	-5.00E-09	0.21422377	2.20E+01	
23	-4.00E-09	-0.21101108	2.20E+01	
24	-3.00E-09	-0.01037758	2.20E+01	
25	-2.00E-09	-0.01938096	2.20E+01	

PLI Decay

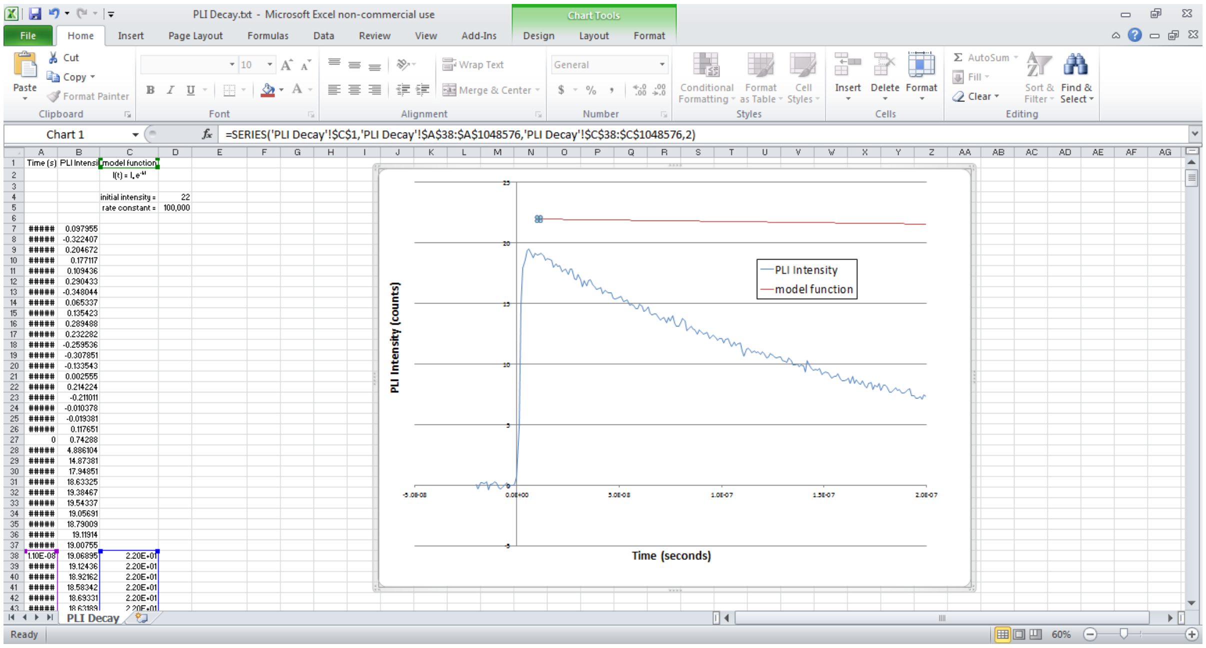
This is real data, therefore it contains evidence of the experimental technique used to acquire it. This is an important point because you should not use (that is simulate) all the data that you gather in an experiment. In this example you can discard the data from the negative time domain, and you can disregard the small rise time for the PLI Intensity signal near time zero. Remember that this sample was pulsed with a laser; the laser pulse is not infinity sharp. There is some time domain in which the sample is still in the process of being excited by the laser pulse. Notice the ‘zoomed-in’ screen shot below, the derivative of the signal between zero and the first few nanoseconds changes. This is an experimental artifact. You can restrict you ‘model function’ to only fit data that is reporting on the behavior you want to simulate.



=SERIES('PLI Decay'!$C$1,'PLI Decay'!$A$7:$A$1048576,'PLI Decay'!$C$7:$C$1048576,2)

	A	B	C	D
1	Time (s)	PLI Intensity	model function	
2			$I(t) = I_o e^{-kt}$	
3				
4			initial intensity =	22
5			rate constant =	100,000
6				
7	-2.00E-08	0.09795532	2.20E+01	
8	-1.90E-08	-0.3224073	2.20E+01	
9	-1.80E-08	0.20467186	2.20E+01	
10	-1.70E-08	0.17711744	2.20E+01	
11	-1.60E-08	0.10943631	2.20E+01	
12	-1.50E-08	0.29043281	2.20E+01	
13	-1.40E-08	-0.34804445	2.20E+01	
14	-1.30E-08	0.06533704	2.20E+01	
15	-1.20E-08	0.13542331	2.20E+01	
16	-1.10E-08	0.28948781	2.20E+01	
17	-1.00E-08	0.23228161	2.20E+01	
18	-9.00E-09	-0.2595356	2.20E+01	
19	-8.00E-09	-0.30785105	2.20E+01	
20	-7.00E-09	-0.13354348	2.20E+01	
21	-6.00E-09	0.00255483	2.20E+01	
22	-5.00E-09	0.21422377	2.20E+01	
23	-4.00E-09	-0.21101108	2.20E+01	
24	-3.00E-09	-0.01037758	2.20E+01	
25	-2.00E-09	-0.01938096	2.20E+01	

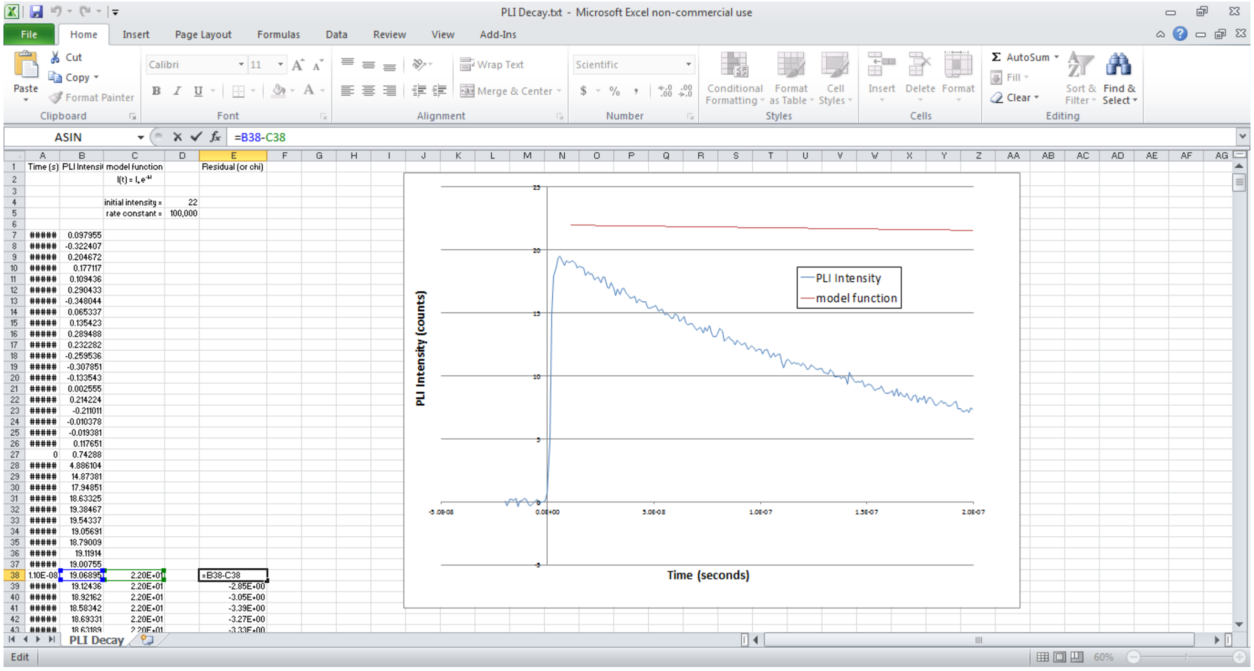
Constrict the domain of the ‘model function’ to a region that reflects the behavior you are actually trying to simulate. This is an important point, in real world experiments YOU will need to determine what data is ‘fair’ to apply your model to.



=SERIES('PLI Decay'!$C$1,'PLI Decay'!$A$38:$A$1048576,'PLI Decay'!$C$38:$C$1048576,2)

	A	B	C	D
1	Time (s)	PLI Intensi	model function	
2			$I(t) = I_o e^{-kt}$	
3				
4			initial intensity =	22
5			rate constant =	100,000
6				
7	#####	0.097955		
8	#####	-0.322407		
9	#####	0.204672		
10	#####	0.177117		
11	#####	0.109436		
12	#####	0.290433		
13	#####	-0.348044		
14	#####	0.065337		
15	#####	0.135423		
16	#####	0.289488		
17	#####	0.232282		
18	#####	-0.259536		
19	#####	-0.307851		
20	#####	-0.133543		
21	#####	0.002555		
22	#####	0.214224		
23	#####	-0.211011		
24	#####	-0.010378		
25	#####	-0.019381		
26	#####	0.117651		
27	0	0.74288		
28	#####	4.886104		
29	#####	14.87381		
30	#####	17.94851		
31	#####	18.63325		
32	#####	19.38467		
33	#####	19.54337		
34	#####	19.05691		
35	#####	18.79009		
36	#####	19.11914		
37	#####	19.00755		
38	1.10E-08	19.06895	2.20E+01	
39	#####	19.12436	2.20E+01	
40	#####	18.92162	2.20E+01	
41	#####	18.58342	2.20E+01	
42	#####	18.69331	2.20E+01	
43	#####	18.63189	2.20E+01	

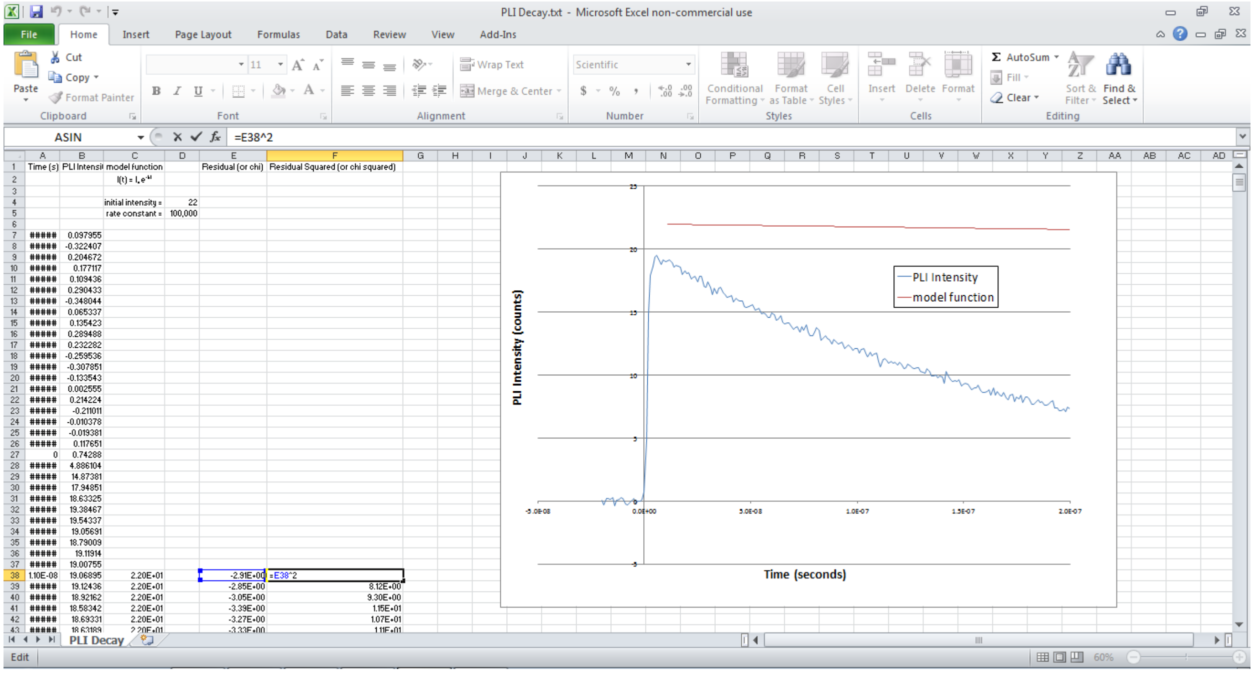
Next we will make a column and calculate the residuals from our model function (see equation 1).



=B38-C38

	A	B	C	D	E	F
1	Time (s)	PLI Intensi	model function		Residual (or chi)	
2			$I(t) = I_0e^{-kt}$			
3						
4			initial intensity =	22		
5			rate constant =	100,000		
6						
7	#####	0.097955				
8	#####	-0.322407				
9	#####	0.204672				
10	#####	0.177117				
11	#####	0.109436				
12	#####	0.290433				
13	#####	-0.348044				
14	#####	0.065337				
15	#####	0.135423				
16	#####	0.289488				
17	#####	0.232282				
18	#####	-0.259536				
19	#####	-0.307851				
20	#####	-0.133543				
21	#####	0.002555				
22	#####	0.214224				
23	#####	-0.211011				
24	#####	-0.010378				
25	#####	-0.019381				
26	#####	0.117651				
27	0	0.74288				
28	#####	4.886104				
29	#####	14.87381				
30	#####	17.94851				
31	#####	18.63325				
32	#####	19.38467				
33	#####	19.54337				
34	#####	19.05691				
35	#####	18.79009				
36	#####	19.11914				
37	#####	19.00755				
38	1.10E-08	19.06895	2.20E+01		=B38-C38	
39	#####	19.12436	2.20E+01		-2.85E+00	
40	#####	18.92162	2.20E+01		-3.05E+00	
41	#####	18.58342	2.20E+01		-3.39E+00	
42	#####	18.69331	2.20E+01		-3.27E+00	
43	#####	18.63189	2.20E+01		-3.33E+00	

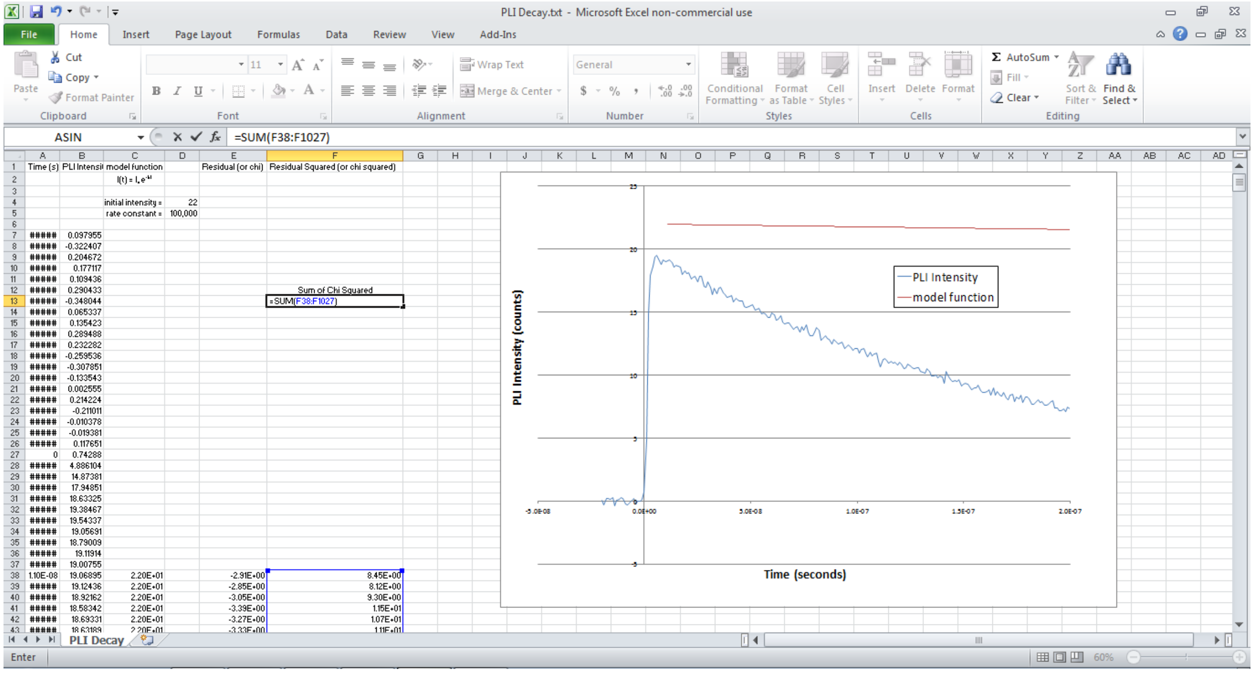
Now we will square the value of the calculated residuals, giving us chi squared for each y_i.



=E38^2

	A	B	C	D	E	F
1	Time (s)	PLI Intensi	model function		Residual (or chi)	Residual Squared (or chi squared)
2			$I(t) = I_0e^{-kt}$			
3						
4			initial intensity =	22		
5			rate constant =	100,000		
6						
7	#####	0.097955				
8	#####	-0.322407				
9	#####	0.204672				
10	#####	0.177117				
11	#####	0.109436				
12	#####	0.290433				
13	#####	-0.348044				
14	#####	0.065337				
15	#####	0.135423				
16	#####	0.289488				
17	#####	0.232282				
18	#####	-0.259536				
19	#####	-0.307851				
20	#####	-0.133543				
21	#####	0.002555				
22	#####	0.214224				
23	#####	-0.211011				
24	#####	-0.010378				
25	#####	-0.019381				
26	#####	0.117651				
27	0	0.74288				
28	#####	4.886104				
29	#####	14.87381				
30	#####	17.94851				
31	#####	18.63325				
32	#####	19.38467				
33	#####	19.54337				
34	#####	19.05691				
35	#####	18.79009				
36	#####	19.11914				
37	#####	19.00755				
38	1.10E-08	19.06895	2.20E+01		-2.91E+00	=E38^2
39	#####	19.12436	2.20E+01		-2.85E+00	8.12E+00
40	#####	18.92162	2.20E+01		-3.05E+00	9.30E+00
41	#####	18.58342	2.20E+01		-3.39E+00	1.15E+01
42	#####	18.69331	2.20E+01		-3.27E+00	1.07E+01
43	#####	18.63189	2.20E+01		-3.33E+00	1.11E+01

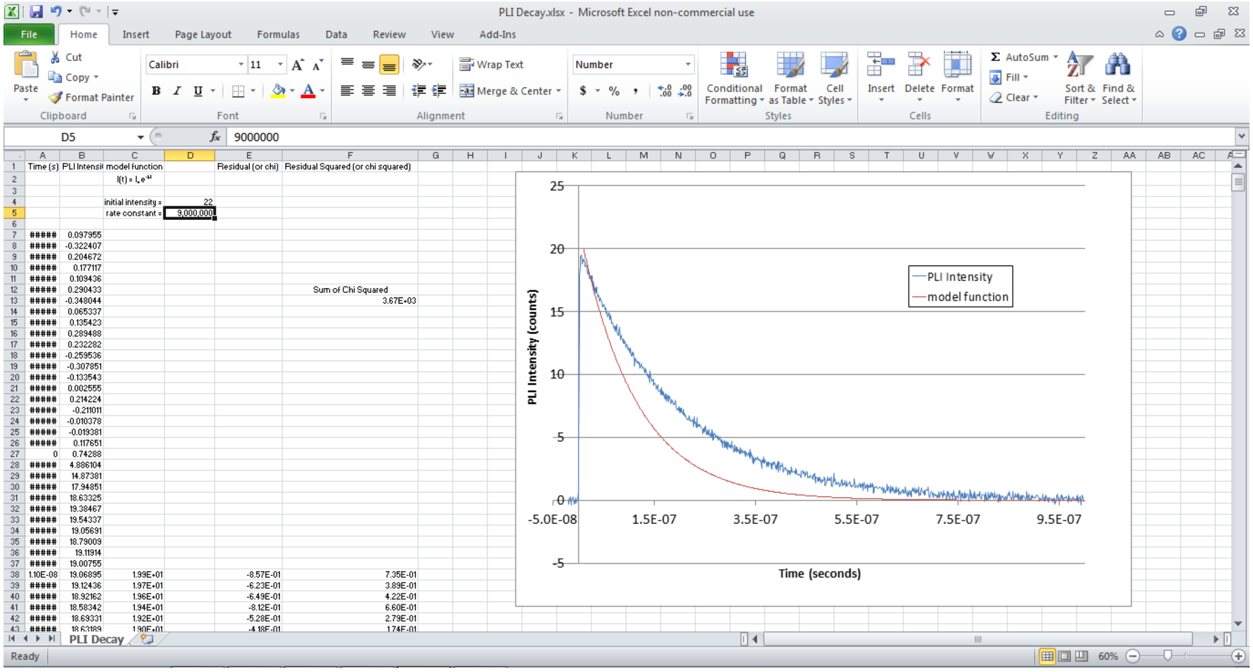
We now sum all the squared residuals (see equation 2).



ASIN =SUM(F38:F1027)

	A	B	C	D	E	F	G
1	Time (s)	PLI Intensit	model function		Residual (or chi)	Residual Squared (or chi squared)	
2			$I(t) = I_o e^{-kt}$				
3							
4			initial intensity =	22			
5			rate constant =	100,000			
6							
7	#####	0.097955					
8	#####	-0.322407					
9	#####	0.204672					
10	#####	0.177117					
11	#####	0.109436					
12	#####	0.290433				Sum of Chi Squared	
13	#####	-0.348044				=SUM(F38:F1027)	
14	#####	0.065337					
15	#####	0.135423					
16	#####	0.289488					
17	#####	0.232282					
18	#####	-0.259536					
19	#####	-0.307851					
20	#####	-0.133543					
21	#####	0.002555					
22	#####	0.214224					
23	#####	-0.211011					
24	#####	-0.010378					
25	#####	-0.019381					
26	#####	0.117651					
27	0	0.74288					
28	#####	4.886104					
29	#####	14.87381					
30	#####	17.94851					
31	#####	18.63325					
32	#####	19.38467					
33	#####	19.54337					
34	#####	19.05691					
35	#####	18.79009					
36	#####	19.11914					
37	#####	19.00755					
38	1.10E-08	19.06895	2.20E+01		-2.91E+00	8.45E+00	
39	#####	19.12436	2.20E+01		-2.85E+00	8.12E+00	
40	#####	18.92162	2.20E+01		-3.05E+00	9.30E+00	
41	#####	18.58342	2.20E+01		-3.39E+00	1.15E+01	
42	#####	18.69331	2.20E+01		-3.27E+00	1.07E+01	
43	#####	18.63189	2.20E+01		-3.33E+00	1.11E+01	

Notice with a rate constant of 100,000 s^{-1} and an initial intensity of 22 we have a chi squared value of 3.10E+05. Change one of the parameters and see what happens to chi squared. I changed the rate constant to 9,000,000 and the chi squared value decreased to 3.67E+03. You can see from the screen shot (below) that the larger rate constant gives a better fit to the experimental data.

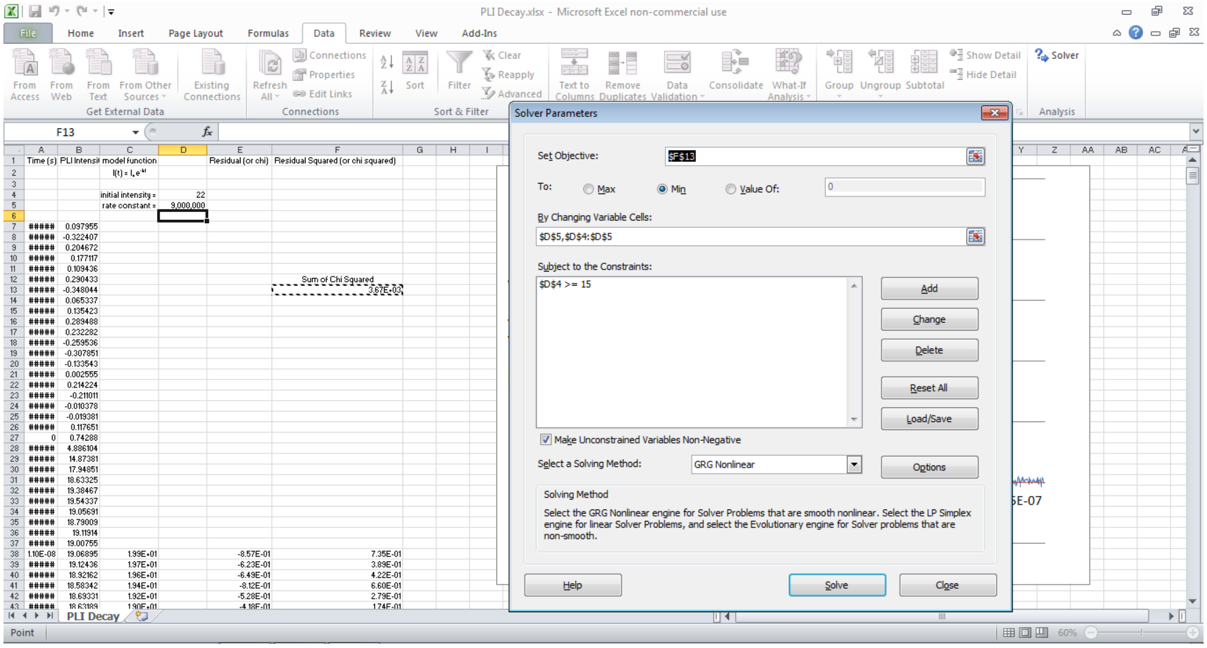


D5 9000000

	A	B	C	D	E	F	G
1	Time (s)	PLI Intensit	model function		Residual (or chi)	Residual Squared (or chi squared)	
2			$I(t) = I_o e^{-kt}$				
3							
4			initial intensity =	22			
5			rate constant =	9,000,000			
6							
7	#####	0.097955					
8	#####	-0.322407					
9	#####	0.204672					
10	#####	0.177117					
11	#####	0.109436					
12	#####	0.290433				Sum of Chi Squared	
13	#####	-0.348044				3.67E+03	
14	#####	0.065337					
15	#####	0.135423					
16	#####	0.289488					
17	#####	0.232282					
18	#####	-0.259536					
19	#####	-0.307851					
20	#####	-0.133543					
21	#####	0.002555					
22	#####	0.214224					
23	#####	-0.211011					
24	#####	-0.010378					
25	#####	-0.019381					
26	#####	0.117651					
27	0	0.74288					
28	#####	4.886104					
29	#####	14.87381					
30	#####	17.94851					
31	#####	18.63325					
32	#####	19.38467					
33	#####	19.54337					
34	#####	19.05691					
35	#####	18.79009					
36	#####	19.11914					
37	#####	19.00755					
38	1.10E-08	19.06895	1.99E+01		-8.57E-01	7.35E-01	
39	#####	19.12436	1.97E+01		-6.23E-01	3.89E-01	
40	#####	18.92162	1.96E+01		-6.49E-01	4.22E-01	
41	#####	18.58342	1.94E+01		-8.12E-01	6.60E-01	
42	#####	18.69331	1.92E+01		-5.28E-01	2.79E-01	
43	#####	18.63189	1.90E+01		-4.18E-01	1.74E-01	

Now it would be tedious to spend all day systematically changing the initial intensity and rate constant values until you calculated the smallest value of Chi-squared (χ^2). Minimizing Chi-squared (χ^2) by exploring many initial intensity and rate constant values is an ideal task for a computer algorithm. Go to the "Data" tab under the "Analysis" section and open the "Solver".

Set the objective as your Chi-squared (χ^2), select "min" (for minimization of the objective). You can add constraints to your parameters. I added the constraint that the initial intensity value could not be less than 15.



F13

	A	B	C	D	E	F
1	Time (s)	PLI Intensi	model function		Residual (or chi)	Residual Squared (or chi squared)
2			$I(t) = I_o e^{-kt}$			
3						
4			initial intensity =	22		
5			rate constant =	9,000,000		
6						
7	#####	0.097955				
8	#####	-0.322407				
9	#####	0.204672				
10	#####	0.177117				
11	#####	0.109436				
12	#####	0.290433				Sum of Chi Squared
13	#####	-0.348044				3.67E+03
14	#####	0.065337				
15	#####	0.135423				
16	#####	0.289488				
17	#####	0.232282				
18	#####	-0.259536				
19	#####	-0.307851				
20	#####	-0.133543				
21	#####	0.002555				
22	#####	0.214224				
23	#####	-0.211011				
24	#####	-0.010378				
25	#####	-0.019381				
26	#####	0.117651				
27	0	0.74288				
28	#####	4.886104				
29	#####	14.87381				
30	#####	17.94851				
31	#####	18.63325				
32	#####	19.38467				
33	#####	19.54337				
34	#####	19.05691				
35	#####	18.79009				
36	#####	19.11914				
37	#####	19.00755				
38	1.10E-08	19.06895	1.99E+01		-8.57E-01	7.35E-01
39	#####	19.12436	1.97E+01		-6.23E-01	3.89E-01
40	#####	18.92162	1.96E+01		-6.49E-01	4.22E-01
41	#####	18.58342	1.94E+01		-8.12E-01	6.60E-01
42	#####	18.69331	1.92E+01		-5.28E-01	2.79E-01
43	#####	18.63189	1.90E+01		-4.18E-01	1.74E-01

Solver Parameters

Set Objective: $F$13

To: ○ Max ● Min ○ Value Of: 0

By Changing Variable Cells:
$D$5,$D$4:$D$5

Subject to the Constraints:
$D$4 >= 15

Add | Change | Delete | Reset All | Load/Save

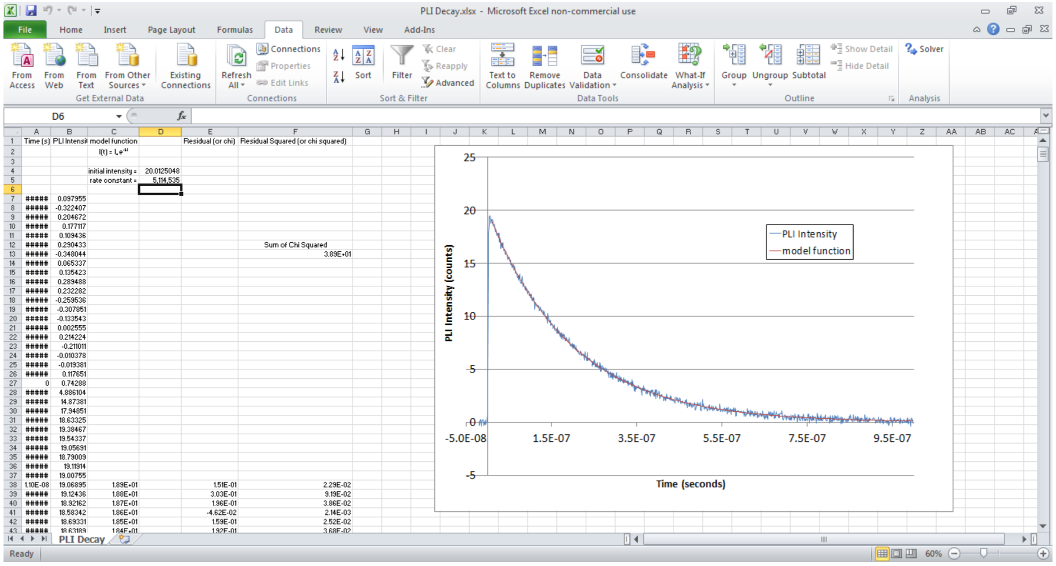
☑ Make Unconstrained Variables Non-Negative

Select a Solving Method: GRG Nonlinear | Options

Solving Method
Select the GRG Nonlinear engine for Solver Problems that are smooth nonlinear. Select the LP Simplex engine for linear Solver Problems, and select the Evolutionary engine for Solver problems that are non-smooth.

Help | Solve | Close

Click on "solve"

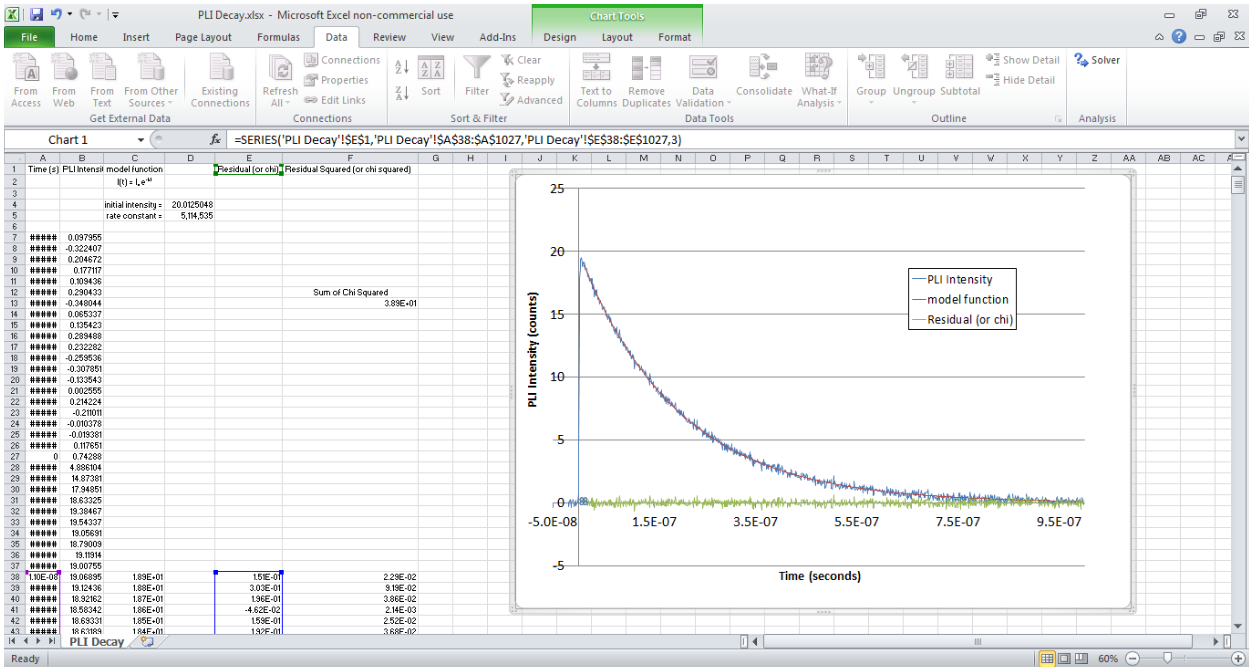


D6

	A	B	C	D	E	F
1	Time (s)	PLI Intensi	model function		Residual (or chi)	Residual Squared (or chi squared)
2			$I(t) = I_o e^{-kt}$			
3						
4			initial intensity =	20.0125048		
5			rate constant =	5,114,535		
6						
7	#####	0.097955				
8	#####	-0.322407				
9	#####	0.204672				
10	#####	0.177117				
11	#####	0.109436				
12	#####	0.290433				Sum of Chi Squared
13	#####	-0.348044				3.89E+01
14	#####	0.065337				
15	#####	0.135423				
16	#####	0.289488				
17	#####	0.232282				
18	#####	-0.259536				
19	#####	-0.307851				
20	#####	-0.133543				
21	#####	0.002555				
22	#####	0.214224				
23	#####	-0.211011				
24	#####	-0.010378				
25	#####	-0.019381				
26	#####	0.117651				
27	0	0.74288				
28	#####	4.886104				
29	#####	14.87381				
30	#####	17.94851				
31	#####	18.63325				
32	#####	19.38467				
33	#####	19.54337				
34	#####	19.05691				
35	#####	18.79009				
36	#####	19.11914				
37	#####	19.00755				
38	1.10E-08	19.06895	1.89E+01		1.51E-01	2.29E-02
39	#####	19.12436	1.88E+01		3.03E-01	9.19E-02
40	#####	18.92162	1.87E+01		1.96E-01	3.86E-02
41	#####	18.58342	1.86E+01		-4.62E-02	2.14E-03
42	#####	18.69331	1.85E+01		1.59E-01	2.52E-02
43	#####	18.63189	1.84E+01		1.92E-01	3.68E-02

The best fit rate constant is 5.12E+06 and the initial intensity is 20.

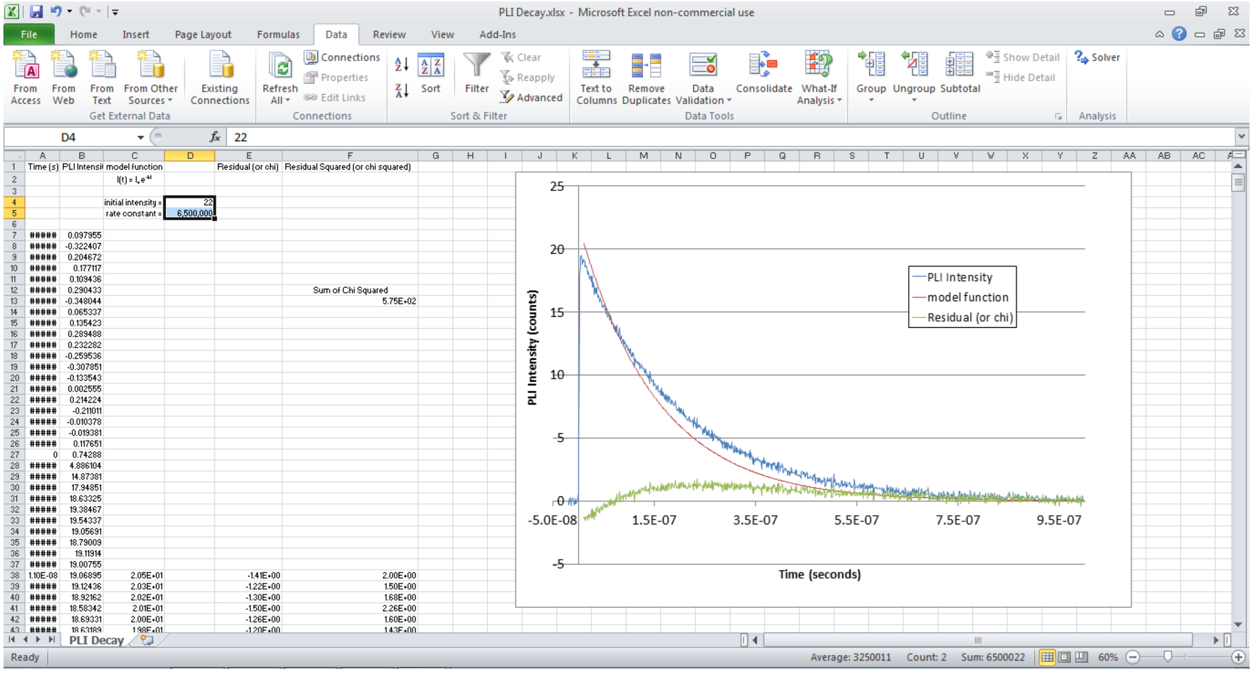
The fit looks prey good, however we should plot the residuals of the fit just to make sure that there is no systematic variation in the residuals. If the residual plot does not randomly oscillate around zero then there is something wrong with the fitting procedure or the proposed model function.



=SERIES('PLI Decay'!$E$1,'PLI Decay'!$A$38:$A$1027,'PLI Decay'!$E$38:$E$1027,3)

	A	B	C	D	E	F
1	Time (s)	PLI Intensit	model function		Residual (or chi)	Residual Squared (or chi squared)
2			$I(t) = I_o e^{-kt}$			
3						
4			initial intensity =	20.0125048		
5			rate constant =	5,114,535		
6						
7	#####	0.097955				
8	#####	-0.322407				
9	#####	0.204672				
10	#####	0.177117				
11	#####	0.109436				
12	#####	0.290433				Sum of Chi Squared
13	#####	-0.348044				3.89E+01
14	#####	0.065337				
15	#####	0.135423				
16	#####	0.289488				
17	#####	0.232282				
18	#####	-0.259536				
19	#####	-0.307851				
20	#####	-0.133543				
21	#####	0.002555				
22	#####	0.214224				
23	#####	-0.211011				
24	#####	-0.010378				
25	#####	-0.019381				
26	#####	0.117651				
27	0	0.74288				
28	#####	4.886104				
29	#####	14.87381				
30	#####	17.94851				
31	#####	18.63325				
32	#####	19.38467				
33	#####	19.54337				
34	#####	19.05691				
35	#####	18.79009				
36	#####	19.11914				
37	#####	19.00755				
38	1.10E-08	19.06895	1.89E+01		1.51E-01	2.29E-02
39	#####	19.12436	1.88E+01		3.03E-01	9.19E-02
40	#####	18.92162	1.87E+01		1.96E-01	3.86E-02
41	#####	18.58342	1.86E+01		-4.62E-02	2.14E-03
42	#####	18.69331	1.85E+01		1.59E-01	2.52E-02
43	#####	18.63189	1.84E+01		1.92E-01	3.68E-02

Now change the rate constant value to an incorrect number. You can see that the model function does not match the experimental data very well and that the residual plot does not randomly oscillate around zero.



22

	A	B	C	D	E	F
1	Time (s)	PLI Intensit	model function		Residual (or chi)	Residual Squared (or chi squared)
2			$I(t) = I_o e^{-kt}$			
3						
4			initial intensity =	22		
5			rate constant =	6,500,000		
6						
7	#####	0.097955				
8	#####	-0.322407				
9	#####	0.204672				
10	#####	0.177117				
11	#####	0.109436				
12	#####	0.290433				Sum of Chi Squared
13	#####	-0.348044				5.75E+02
14	#####	0.065337				
15	#####	0.135423				
16	#####	0.289488				
17	#####	0.232282				
18	#####	-0.259536				
19	#####	-0.307851				
20	#####	-0.133543				
21	#####	0.002555				
22	#####	0.214224				
23	#####	-0.211011				
24	#####	-0.010378				
25	#####	-0.019381				
26	#####	0.117651				
27	0	0.74288				
28	#####	4.886104				
29	#####	14.87381				
30	#####	17.94851				
31	#####	18.63325				
32	#####	19.38467				
33	#####	19.54337				
34	#####	19.05691				
35	#####	18.79009				
36	#####	19.11914				
37	#####	19.00755				
38	1.10E-08	19.06895	2.05E+01		-1.41E+00	2.00E+00
39	#####	19.12436	2.03E+01		-1.22E+00	1.50E+00
40	#####	18.92162	2.02E+01		-1.30E+00	1.68E+00
41	#####	18.58342	2.01E+01		-1.50E+00	2.26E+00
42	#####	18.69331	2.00E+01		-1.26E+00	1.60E+00
43	#####	18.63189	1.98E+01		-1.20E+00	1.43E+00

The purpose of this lab was to help you understand and practice implementing least-squared curve fitting techniques. It is very likely that your future employer will have sophisticated least-squared curve fitting software specially designed for their work. The principles we used today will apply in that software as well, although it may be more difficult to determine what the program is actually doing.

In Physical Chemistry we will use OriginLab (http://www.originlab.com/) data analysis and graphing software for all lab reports. I will teach you how to use OriginLab during a special class session. OriginLab is a common software program in academic and industry research labs. OriginLab is capable of program high level least-squared curve fitting and creating publication quality graphs.

Questions and Further Thoughts .

1. What is a random vs. systemic error?
2. How would you know if your least-squares minimization program had found a local minimum or a global minimum on the Chi-squared (χ^2) surface?
3. What does it suggest if Chi-squared (χ^2) has been minimized and you have found the global minima using your model function, yet the 'residuals' plot shows systematic variation?