1. Motivation for Predicting Luminosity Behavior ... 3
2. Luminosity Models ... 3
   a. Simple Exponential Model ... 4
   b. Modified Exponential Model ... 6
   c. Simple Inverse Time to Power Model ... 7
   d. Modified Inverse Time to Power Model ... 8
   e. Chi Square Test of Merit ... 9
3. Building a Spreadsheet Tool ... 10
   a. Spreadsheet Setup ... 10
   b. External Data Spreadsheets ... 13
      i. SuperTable II ... 13
      ii. Lumberjack Data ... 13
   c. Luminosity Predictor Spreadsheet: How to Analyze the Data ... 14
      i. Input Data Workbook (Procedure for using the spreadsheet) ... 15
   d. Luminosity Predictor Spreadsheet: Workbooks in Detail ... 23
      i. LBOE Workbook ... 23
      ii. Lumberjack Data import ... 26
      iii. Input Data for the Luminosity Models ... 29
      iv. The Luminosity Models ... 30
      v. Error Sums ... 34
      vi. Summaries ... 35
      vii. Fits over time ... 36
      viii. Plots ... 38
   e. VBA Scripts ... 39
      i. Clear All Data Script (Ctrl-Shift-C) ... 39
      ii. Analyze the Data Script (Ctrl-Shift-A) ... 39
      iii. Archive (Write) the Data Script (Ctrl-Shift-W) ... 40
   a. Archive Data Files ... 41
      i. Individual store{####}-fit-data.xls files ... 41
      ii. Store-fit-summary.xls file ... 42
4. How luminosity fits improve over time ... 43
   a. Simple Exponential Fit (Equation 1) ... 43
   b. Modified Exponential Fit (Equation 2) ... 46
   c. Inverse Time to Power Fit (Equation 3) ... 49
   d. Modified Inverse Time to Power Fit (Equation 3) ... 52
5. Beginning of Store ... 55
   a. Simple Exponential Fit (Equation 1) ... 55
   b. Modified Exponential Fit (Equation 2) ... 56
   c. Inverse Time to Power Fit (Equation 3) ... 58
   d. Modified Inverse Time to Power Fit (Equation 4) ... 59
6. End of Store ... 61
   a. Simple Exponential Fit (Equation 1) ... 61
   b. Modified Exponential Fit (Equation 2) ... 62
   c. Inverse Time to Power Fit (Equation 3) ... 63
   d. Modified Inverse Time to Power Fit (Equation 4) ... 66
7. Comparing Fit Numbers for Top 5 Stores ... 67

a. Simple Exponential Fit (Equation 1) 68
b. Modified Exponential Fit (Equation 2)..... 68
c. Inverse Time to Power Fit (Equation 3) 70
d. Modified Inverse Time to Power Fit (Equation 4)..... 71
8. Conclusions..... 73
9. References and Useful Sources..... 73

1. Motivation for Predicting Luminosity Behavior

When planning Collider store turn around times, it would be beneficial to have a tool that could be used anytime in the store to predict the Luminosity behavior later in that store.

We have three existing tools that can help us determine the Luminosity behavior of a store. First, there are models of the Tevatron luminosity[1-4]. Tevatron experts have models that closely predict the luminosity behavior of a store given a few constants including the initial luminosity and luminosity lifetime. If we could determine the correct values for the above mentioned constants early on in the store, we could use a luminosity model to predict the Luminosity behavior over the entire store. Second, we have the SuperTable. Experts examine the luminosity data during the first few hours of each store and calculate an initial luminosity and luminosity lifetime of the store based on a simple exponential fit. These numbers are placed in the SuperTable, and can be easily retrieved in an Excel Spreadsheet. These values provide early feedback as to the initial health of the store. We will see that a simple exponential fit using the SuperTable values does not provide a good long term prediction of the Luminosity behavior of the store, but does provide a good starting point. Last, we have the datalogger. The luminosity readings are datalogged for each store. We can easily export this data to an Excel spreadsheet and plot how the Luminosity has progressed anytime during the store.

The goal of this exercise is to build an Excel spreadsheet to help predict the Luminosity behavior of a store. We will construct the spreadsheet so that can be used at anytime during a store. The spreadsheet would use existing Luminosity Models to calculate Luminosity behavior. The initial guesses at the initial luminosity and luminosity lifetimes would be gathered from the SuperTable. The Luminosity Model constants would then be fit to the Lumberjack data for that store. As the store progressed, the tool could be used repetitively to get better and better Luminosity behavior predictions. When the store is finished, we could then examine how accurate the predictions from the various models matched the actual luminosity data.

2. Luminosity Models

We will look at four basic Luminosity Models. The constants for each of these fits are calculated at the end of each store by Elliot McCrory[2-4] and are displayed online at http://mccrory.fnal.gov/tevatronDecayFits/. Future additions of this website will also calculate the fits at different times during the store.

In this exercise, we will make a tool to complete the same calculations. There are two primary differences between this tool and the webpage mentioned above. This tool will complete the calculations on demand, as opposed to the webpage

which collects data at pre-determined times. The webpage fits about 200 points over the entire store, and often cuts a significant number of these points away. My tool will be designed to collect lumberjack data at a sample rate op up to four times a minute, giving us more than an order of magnitude more data points to work with.

For each fit, we will show example store data. I have chosen store 4639, which holds our record for integrated luminosity. The store was long-lived and thus provides a large data sample for us to analyze.

a. Simple Exponential Model

The Simple Exponential fit is what is used to create the luminosity lifetime numbers posted in the SuperTable and is given by Equation (1)

$$L(t) = L_0 e^{-\left(\frac{t}{\tau}\right)} \qquad (1)$$

where L(t) is the Luminosity at time t, L_0 is the initial luminosity, t is the time, and τ is the luminosity lifetime.

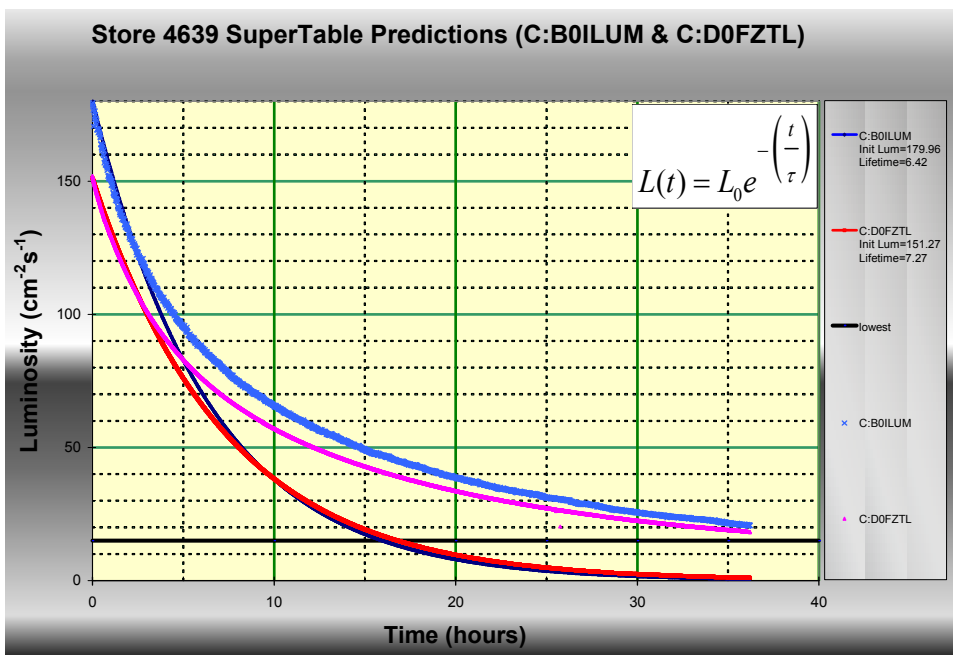


Figure 2-1: The medium blue and pink traces are the luminosity over time for CDF and D0 for store 4369. The dark blue and red traces are luminosity distributions for CDF and D0 calculated from lifetime and initial luminosity data in the SuperTable II .

Figure 2-1 shows the exponential curve and the lumberjack data for store 4369. The x-axis is time in hours from the beginning of the store, and the y-axis is store luminosity. The dark blue and red curves are the exponential curves for CDF and D0. The curves are generated using Equation (1) with the initial luminosity and luminosity lifetime numbers from the SuperTable II. The medium blue and pink traces are the luminosity data for CDF and D0 collected from the lumberjack. If the luminosity truly followed the exponential expression in Equation (1) then we would expect the CDF predicted (red line) and actual (pink) to be aligned, and the D0 predicted (dark blue) and actual (blue) to be aligned. We see in the first few hours of the store there is very good agreement between the exponential curve and the lumberjack data; however, as we look later in the store the data soon diverges. If we were to use the exponential fit with the SuperTable numbers to predict the Luminosity later in the store, we would not make a very good prediction.

What happens if we modify the values of initial luminosity and luminosity lifetime in Equation (1) to try to make the experimental and real data better match? The results are shown in Figure 2-2.

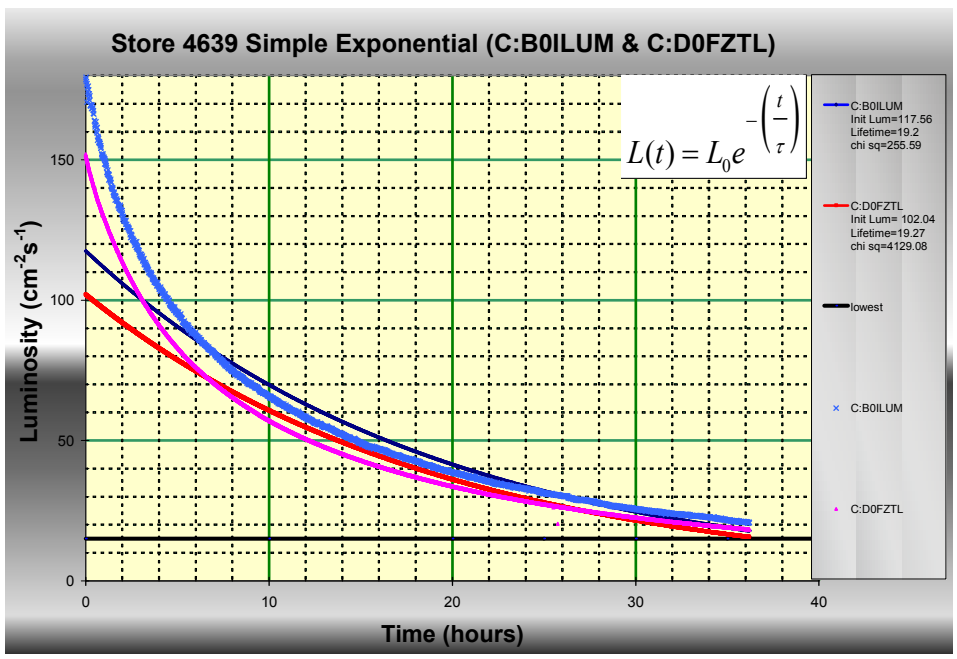


Figure 2-2: Here we attempt to match the luminosity data with the Simple Exponential model of Equation (1). We modify the initial luminosity and luminosity lifetime constants in the equation to attempt to make the best fit.

Figure 2-2 shows that the luminosity data and the curve from Equation (1). Using the Excel solver, we were not able to find values for the initial luminosity and luminosity lifetime that would

make the curves generated by Equation (1) match the lumberjack data. We will find better results in our next model.

b. Modified Exponential Model

The second fit is a modification of the exponential fit given in Equation (1). We still use the exponential fit, but assume that the lifetime in the denominator varies with time. We add a constant multiplied by the time raised to another constant to the initial lifetime. The result is shown in Equation (2).

$$L(t) = L_0 e^{-\left(\frac{t}{\tau + \mu t^{\alpha}}\right)} \qquad \text{Equation (2)}$$

where L(t) is the Luminosity at time t, L_0 is the initial luminosity, t is the time, τ is the luminosity lifetime, μ is a positive constant and α is a positive constant.

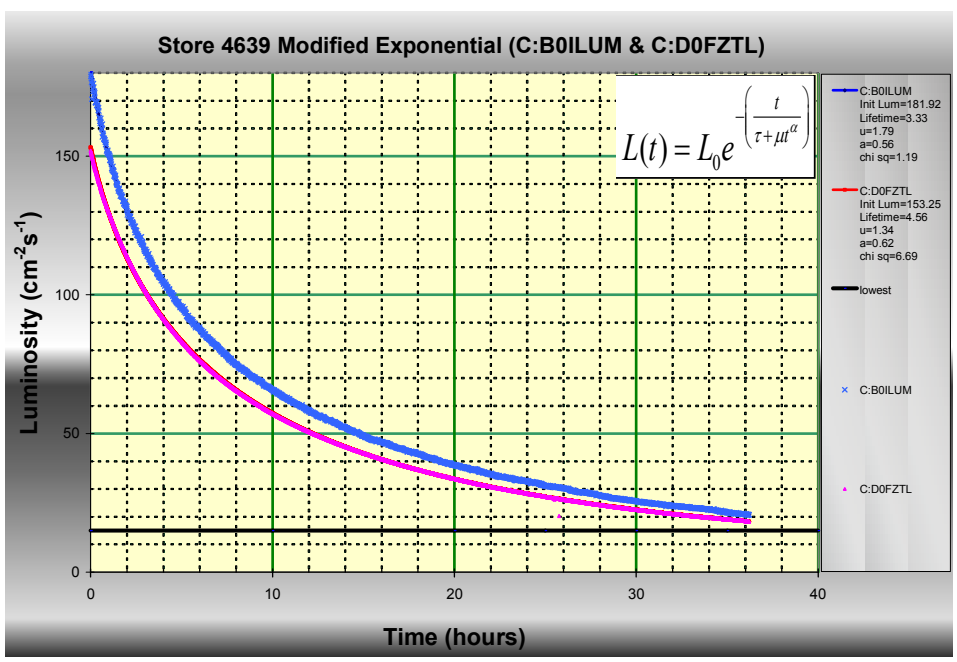


Figure 2-3: The constants in the modified exponential fit of Equation (2) were modified with the Excel Solver to obtain a very good match with the luminosity data from Store 4369.

Figure 2-3 shows data from Store 4639. We used the Excel Solver to modify the four constants in Equation (2) to match the lumberjack data. Unlike the simple exponential fit, our modified exponential fit gives us very good agreement between the curves and the lumberjack data.

c. Simple Inverse Time to Power Model

Another model, found in "Recycler-Only Operations Luminosity Projections" by Dave McGinnis[1] provides a luminosity fit with only three constants. This equation is given in Equation (3)

$$L(t) = \frac{L_0}{\left(1 + \frac{t}{\mu\tau}\right)^{\mu}} \qquad \text{Equation (3)}$$

where L(t) is the Luminosity at time t, L_0 is the initial luminosity, t is the time, τ is the luminosity lifetime, and μ is a positive constant.

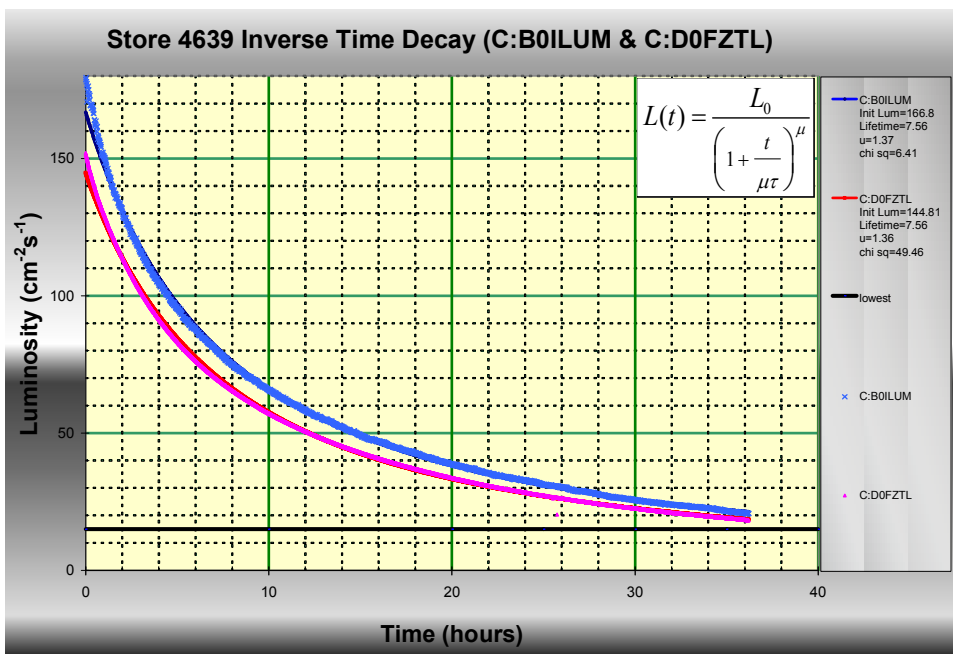
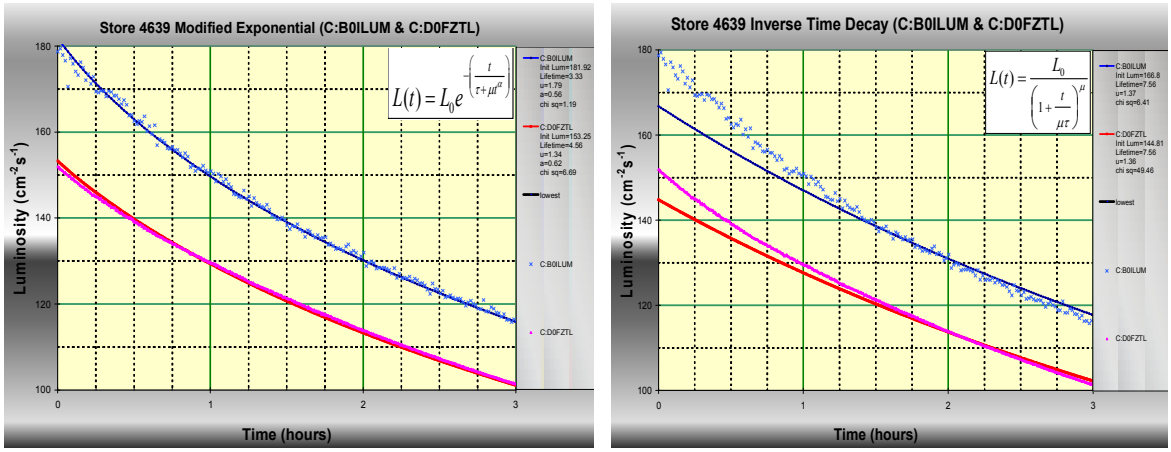


Figure 2-4: Modifying the constants in equation (3) we were able to obtain a fairly good fit for the Store 4369 luminosity data. Careful inspection of the graph shows that the model is least accurate in the first few hours of the store.

Figure 2-4 shows data from Store 4639. We used the Excel Solver to modify the three constants in Equation (3) to match the lumberjack data. This fit works very well; however, a closer inspection shows that it is not quite as accurate at the beginning of the store as Equation (2). We will take a closer look.



Equation (2)
Modified Exponential Fit

Equation (3)
Inverse Time Decay Fit

Figure 2-5: Comparing how well Equations (2) and (3) fit the data from Store 4639 during the first few hours of the store.

Figure 2-5 shows the same data plotted in Figures 3 and 4, only blown up to look at the first three hours of store 4639. This is an attempt to show the relative accuracy of the two fits at the beginning of the store. We can see that during the first hour and a half of the store, Equation (3) does not fit the data as well as Equation (2).

d. Modified Inverse Time to Power Model

Our fourth fit is a modification of the fit given in Equation (3). We will assume that the exponent varies with time. We add a constant multiplied by the time to the initial exponent. The result is shown in Equation (4).

$$L(t) = \frac{L_0}{\left(1+\frac{t}{\mu\tau}\right)^{\mu+t\alpha}} \qquad \text{Equation (4)}$$

where L(t) is the Luminosity at time t, L_0 is the initial luminosity, t is the time, τ is the luminosity lifetime, μ is a positive constant and α is a positive constant.

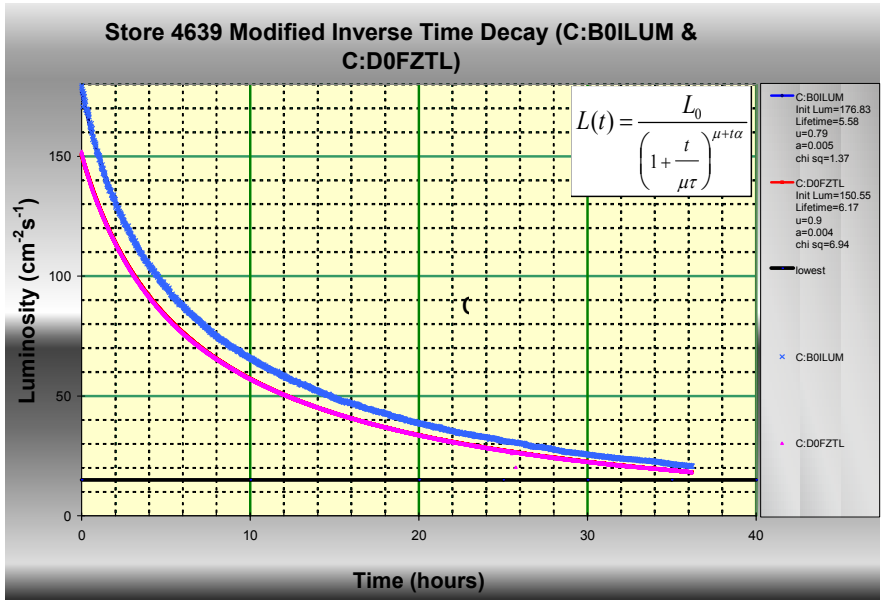


Figure 2-6: The constants in the modified time decay fit of Equation (3) were modified with the Excel Solver to obtain a very good match with the luminosity data from Store 4369.

Figure 2-6 shows data from Store 4639. We used the Excel Solver to modify the four constants in Equation (4) to match the lumberjack data. This fit works very well and fits the live data better at the beginning and end of the store than does Equation (3). We will see that Equations (2) and (4) appear to give the best results.

e. Chi Square Test of Merit

When comparing how well each of our models fit the actual luminosity data, it would be helpful to calculate a statistical value representing the quality of the fit. We have chosen a chi square fit as shown in Equation (5)

$$\chi^2 = \left(\frac{1}{n-C}\right)\sum_{i=1}^{n}\frac{(M_i - L(x_i))^2}{\sigma^2} \qquad \text{Equation (4)}$$

where n is the number of lumberjack sample points, C is the number of constants in our luminosity model, M_i is the measured luminosity , L(x) is the calculated luminosity, σ is the sigma of the measured luminosity.

So to calculate our χ^2 value we will simply subtract the calculated luminosity (from our luminosity models) from our measured luminosity (from the lumberjack data) and square that number. At each point, we will

divide by the σ^2 of the luminosity measurement, where σ is the half height of the error bar of the measured value. We then sum this value over each of our luminosity readings and divide by the total number of data points less the number of constants in our model.

For this exercise, we will assume that σ values quoted by Elliot McCrory[1], which is 0.006*(Measured Luminosity)$_{CDF}$ and 0.0015*(Measured Luminosity)$_{D0}$ for CDF and D0 respectively. If the σ values are incorrect, the χ^2 value will be scaled incorrectly, but we will still get a relative comparison between the fits. The smaller the χ^2 value, the better the fit.

3. Building a Spreadsheet Tool

So far we found that three of our fits work well, two of which work extremely well. However, we have only examined data from one store. We will need to verify that the fits behave similarly for other stores. In addition, we have only fit the data using all of the lumberjack data after the store has been completed. We have not yet covered how well our equations predict luminosity behavior when only given a limited amount of luminosity data. For example, if we are four hours into a store, can we predict what the luminosity will be at 30 hours into the store? How about if we are six hours into the store? Eight hours? How do our luminosity model constants change as we get more and more lumberjack data? Also, how do the constants in our fits change from store to store? Do they always have similar values, or do they change a lot from store to store? Our goal is to build a tool with Excel to help us answer these questions.

a. Spreadsheet Setup

The default Excel spreadsheet configuration provided in the AD drive image does not have all of the features that we will need enabled. In order to run the spreadsheet we will need to enable the analysis toolpak and solver. Go to Tools -> Add-ins. Check the boxes next to “Solver Add-in,” “Analysis ToolPak,” and “Analysis ToolPak - VBA.”

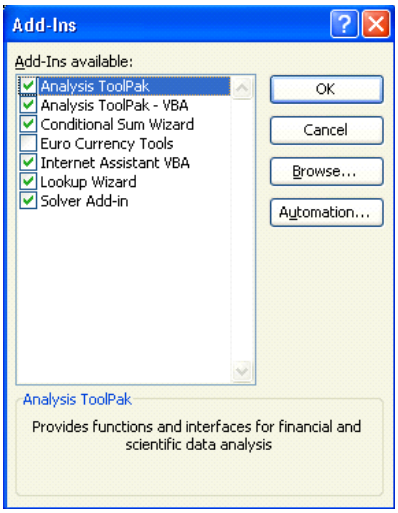


Figure 3-1: Enable the Solver and Analysis Toolpak.

We next need to verify that our security settings allow us to run Macros. Go to Tools -> Macros -> Security. Select "Medium" and click ok.

Figure 3-2: Set the macro security setting to medium. This allows the user to choose if macros are enabled at the time that a spreadsheet with macros is opened. Be careful! Only enable macros on spreadsheets that you are absolutely sure of their source. Macros are a popular way to spread viruses on Windows computers.

We will also need to setup the VBA editor to run scripts with Solver. With out this step, we would have to run the solver manually to do our analysis. Go to Tools>Macro>Visual Basic Editor. The visual basic editor will open. On the Visual Basic Editor, use Tools>References. A dialog box of references will open. Select Solver.

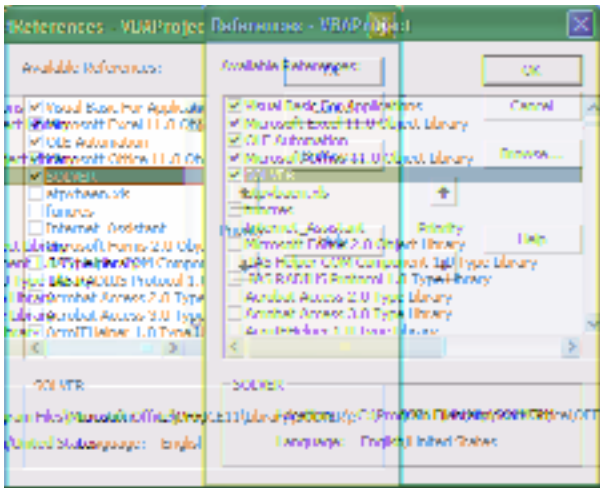


Figure 3-3: Allowing the Solver to run inside of VBA.

Excel has a great feature that automatically saves your work every 10 minutes. This feature helps the user recover their edits when Excel crashes unexpectedly. Unfortunately, this feature can interfere with our data analysis that we will run from VBA scripts. In order to maximize the resources during our data analysis, we turn off the "Save AutoRecover" feature before we run the data analysis VBA script. Go to Tools -> Options -> Save Tab and uncheck the box next to "AutoRecover." We do not want to forget to turn this feature back on after the data analysis is complete, since the "AutoRecover" feature is very useful.

Figure 3-4: Turn off the "AutoRecover" feature when running an Excel data analysis. Turn "AutoRecover" back on when the data analysis is complete**.**

Excel should now be configured with all of the settings that we need to analyze our Collider luminosity data!

b. External Data Spreadsheets

We will want to call data from two external spreadsheets.

i. SuperTable II

An Excel version of the SuperTable is readily available for Windows users at \\daesrv\java_engines\files\SupertableExport. Copy this file to the same directory as our master Excel spreadsheet with the filename **new_supertableII.xls**.

One of the Excel lookup functions that we will use requires that the SuperTable spreadsheet have the store numbers listed in ascending order. By default the SuperTable spreadsheet is sorted by store number, but in descending order. Open the SuperTable spreadsheet, then sort all of the data by store number in ascending order, and save the file as an Excel Spreadsheet.

Update the SuperTable as necessary to ensure that you have the data required for the stores that you want to analyze.

ii. Lumberjack Data

We next need to gather the Lumberjack data for the store we are interested in looking at. From Acnet D44 we can start a luminosity plot by going to Users -> Brian Drendel and then Recall -> ShotSetup. The default dataloggers and sample rates for the luminosity readings for this plot are:

- **C:B0ILIM**: .CDF sampled at a 1 minute rate.
- **C:D0FZTL**: .DZero sampled at a 15 second rate.

We next plot the data from the store in question. Once the plot has been made, we export the data to an Excel spreadsheet using the following steps.

- Select Export Data.
- We only want to export the Luminosity parameters (top two choices). De-select the others and click ok.
- Change the time format for both luminosity parameters from “Lumberjack format” to “hours.”
- Select “Excel File”
- Use the name **shot{four digit shot number}.xls**.

The data has been exported, but we still need to make a local copy.

- Open a web browser to **http://www-ad.fnal.gov/excel** or **http://adcon.fnal.gov/pic/capture/excel**, depending on if your D44 instance was run from a Linux console or a VMS console.
- Right-click on desired file and save it in the same directory as the Luminosity Predictor spreadsheet. Use the name **shot{four digit shot number}.xls**.

We now need to cleanup the data file before we can analyze it.

- Open the **shot{four digit shot number}.xls** file (luminosity-predictor.xls should remain closed).
- We want our Luminosity data to start exactly when the luminosity readings show their initial luminosity values. This will give us the best fit later on. D0 data and CDF data should be done separately.
- CDF: Select any data in the first two columns starting with cells A2 & B2 down to where the luminosity signal comes online. Select Edit->Delete and then select to "shift the cells up."
- D0: Select any data in the first two columns starting with cells C2 & D2 down to where the luminosity signal comes to full value. Make sure to include in your selection the early luminosity data where the luminosity is not at full value. Select Edit -> Delete and then select "shift the cells up."
- Repeat the above two steps for any zero or bad luminosity data at the bottom of the list (When deleting the cells, select to shift cells up)
- Also scan the file for any bad luminosity data and remove those cells (When deleting the cells, select to shift cells up)
- Save the file as **shot{four digit shot number}.xls** as an Excel workbook.

The data is now in a format ready to be analyzed. Repeat the above procedure for each store that you want to examine.

c. Luminosity Predictor Spreadsheet: How to Analyze the Data

The above section concentrated on getting our data formatted in Excel spreadsheets. We will not modify those spreadsheets. We have a separate Excel Spreadsheet that has the tools built in to analyze that data. We call that spreadsheet **Luminosity-Predictor-Plus.xls**

i. Input Data Workbook (Procedure for using the spreadsheet)

By default, all workbooks in the Luminosity Predictor spreadsheet are protected. Most user interaction with the workbook will occur in the "*InputData*" workbook. There are interactive buttons connected to VBA scripts that complete a majority of the tasks. To use this spreadsheet, start at the top and work your way down. We will start by selecting a store number, opening the external SuperTable and Lumberjack spreadsheets, verifying the data inside of the external spreadsheets, and then analyzing the data.

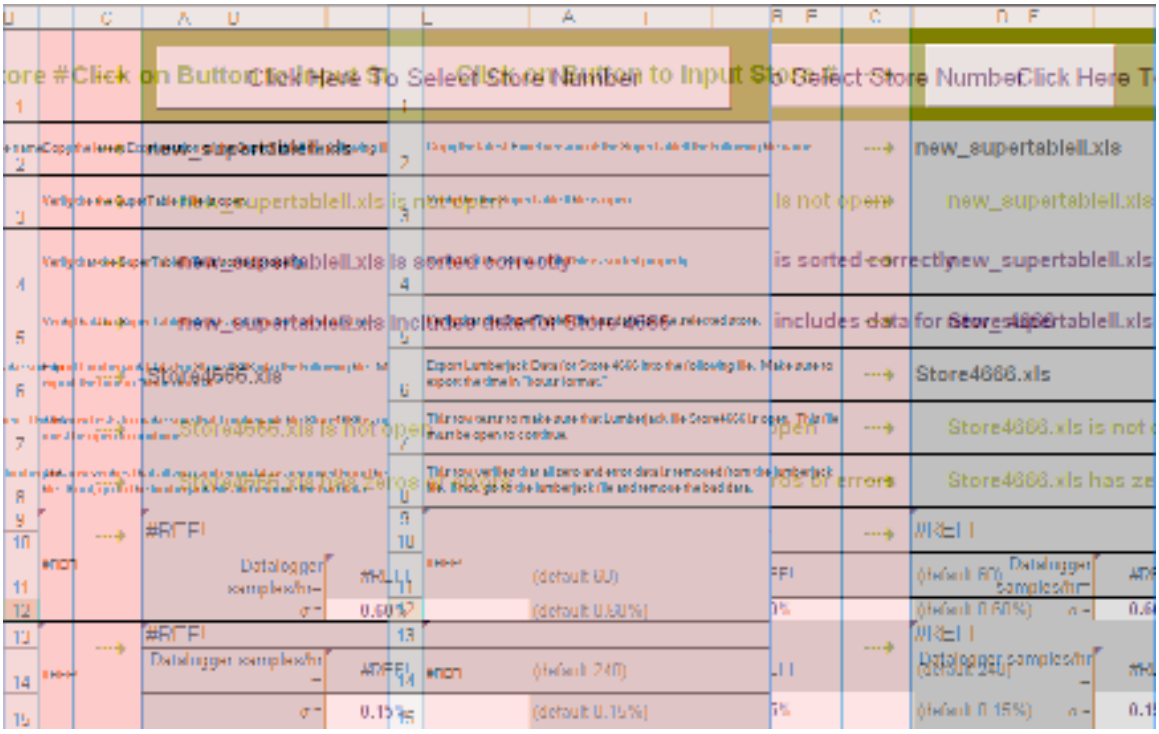


Figure 3-5: The Luminosity Predictor Spreadsheet opened to the "InputData" workbook. We start at the top. Click on the interactive button in Cell D1 to choose a Collider store.

Start by clicking on the interactive button in cell D1.

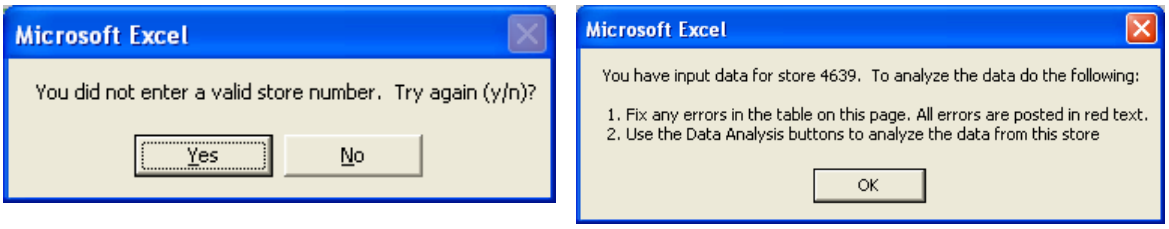


Figure 3-6: Clicking on the interactive button to choose the store number, there are a number of message boxes that the user may encounter. The first message box (upper left) asks the user to input the desired store number. In this example, we want to analyze store 4639, so we enter the store number (upper right) and then click OK. The VBA script has some error checking, so that if we type a store number not recognized by the VBA script, we get an error (lower left). If we chose a number that is a possible store number, we receive a message box (lower right) with some simple instructions on how to continue.

As shown in Figure 3-6, we are greeted with a message box asking us to input our desired store number. Type the desired store number and then click OK. In this example, we will type 4369 to analyze store 4369. There is built in error handling if we try to enter a store number outside of the range of the current store numbers. If we chose a valid store number, we get prompted with another message box, as shown in Figure 3-6, providing simple instructions on how to continue. After reading the instructions click OK. We will now walk through the steps needed to complete our data analysis.

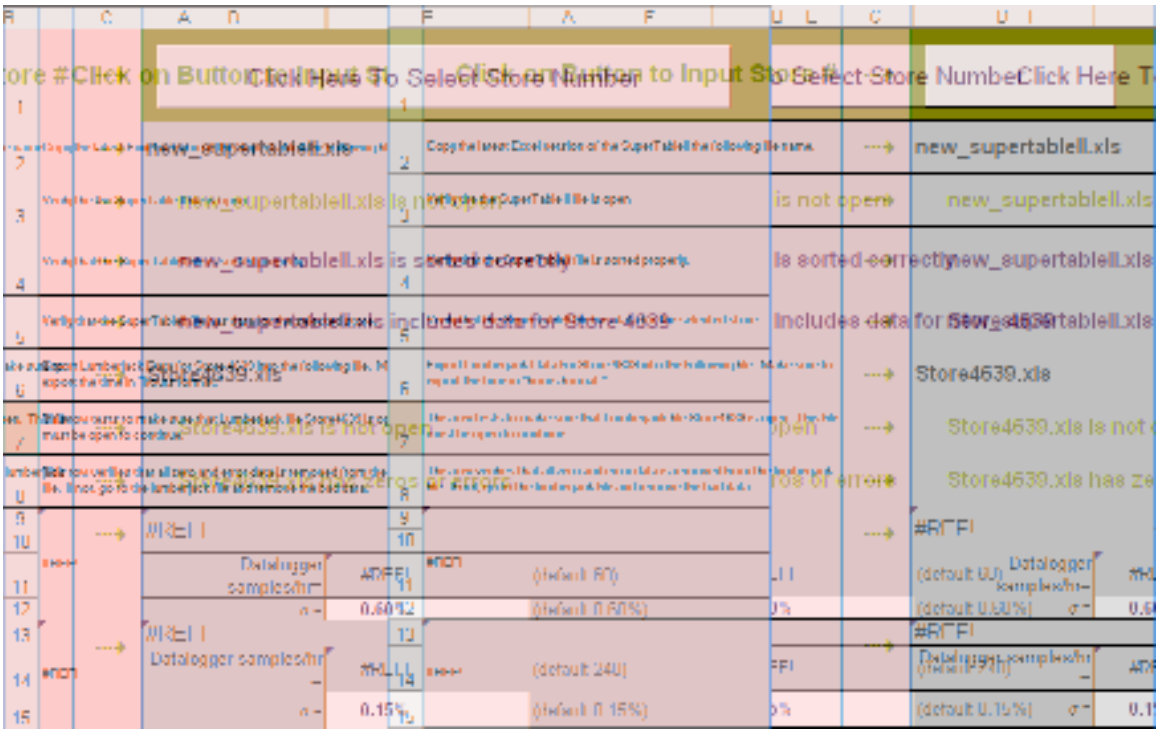


Figure 3-7: Once the desired store number is entered, we need to collect the data for this store from the external SuperTable and Lumberjack spreadsheets. Cell range D2:F15 provide feedback on our external

spreadsheets. Fields that are not in the desired state are displayed in red text. We start with D2 and work our way down.

Cell D2 reminds the user the name of our SuperTable II spreadsheet. If a filename was used, it will not work with the Luminosity Predictor spreadsheet. Cell D3 checks to see if the SuperTable II spreadsheet is open and has conditional formatting to notify us as to its status. In Figure 3-7, we see that the SuperTable II file is not open. The user should now open the file.

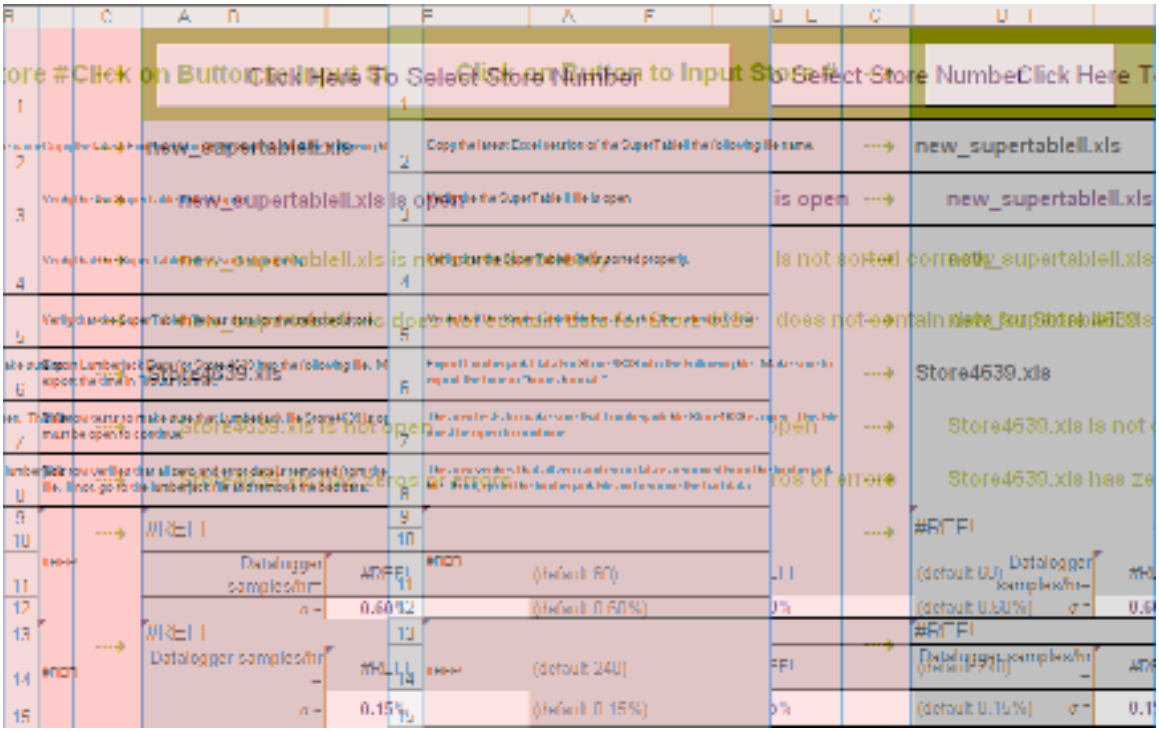


Figure 3-8: We have selected to analyze Store 4639. Cell D3 shows that the SuperTableII spreadsheet is open, but Cell D4 shows that it is not sorted correctly.

Figure 3-8 shows the status after we open the SuperTable II file. Cell D4 checks to ensure that the SuperTable has data sorted by store number in ascending order. One of our Excel lookup functions that we will user later requires that the data be sorted in this manner. In this example, the data is not sorted correctly. The user should now sort the data.

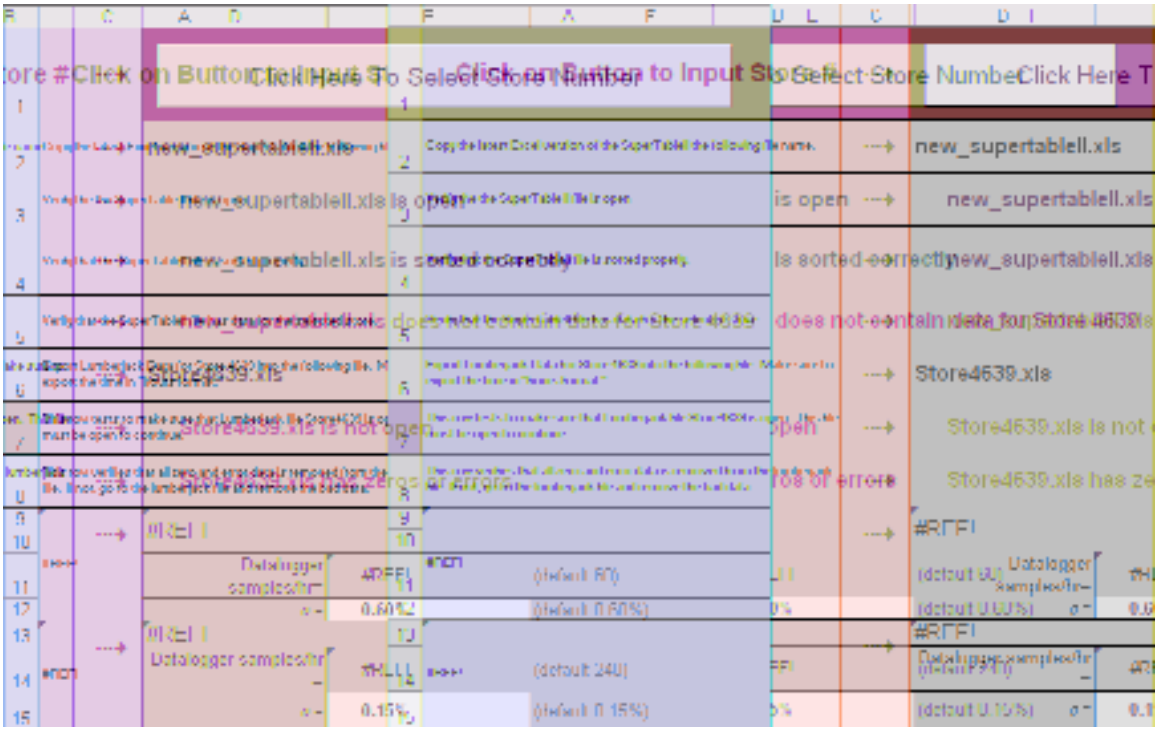


Figure 3-9: Cell D3 shows that the SuperTable spreadsheet is open, and Cell D4 shows that it is sorted properly. However, Cell D5 shows that there is no SuperTable II data for Store 4639 in our spreadsheet. We will either need to change store numbers, or replace the SuperTable II spreadsheet.

Figure 3-9 shows us that the SuperTable II file is open and sorted correctly; however Cell D5 shows that the file does not have data for the selected store. The two most likely causes of this problem are that we have selected a store number that does not exist, or our SuperTable II spreadsheet is old or corrupt. Store 4639 is a valid store number, so we replace our SuperTable II spreadsheet with the latest version from \\daesrv\java_engines\files\SupertableExport\. This file is readily available from the user desktop; however, it is not accessible via wireless or from home without a Controls VPN connection.

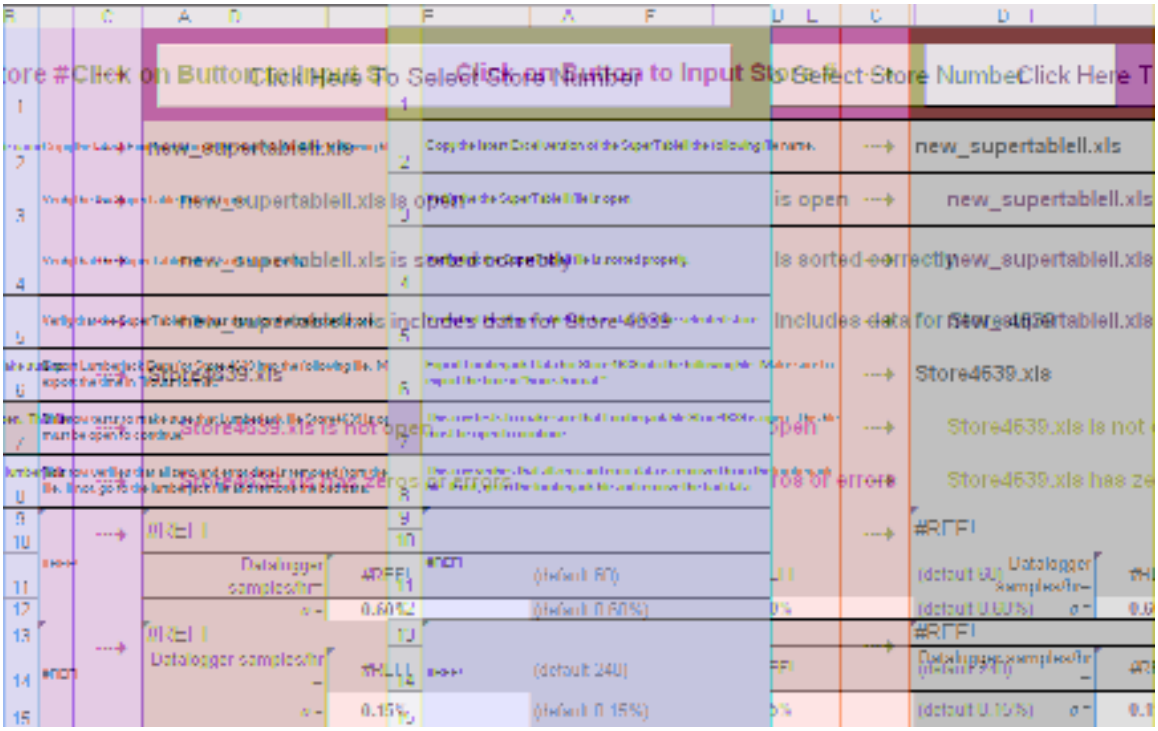


Figure 3-10: The SuperTable II spreadsheet is open, is sorted properly and has data for Store 4639. We will next need to open our Lumberjack data file for Store 4639.

.

Figure 3-10 shows the results of obtaining the latest SuperTable II spreadsheet and having it sorted by store number in ascending order. Cell D3 shows that our SuperTable II spreadsheet is open, cell D4 shows that it is sorted properly, and cell D5 shows us that it contains data for Store 4639. We next turn our attention to the Lumberjack data for Store 4639. Cell D6 provides the file name that the Luminosity Predictor is looking for. Follow the earlier given directions on generating the Excel file from D44.

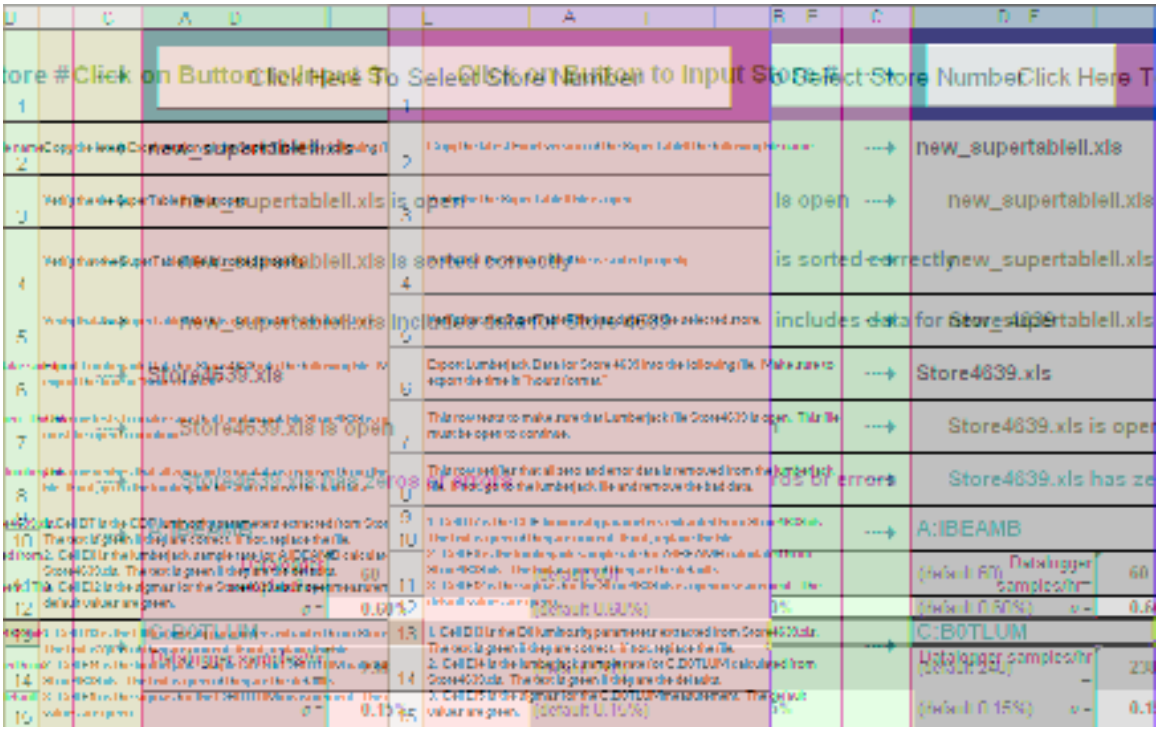


Figure 3-11: Cell D7 shows that we have opened our Lumberjack spreadsheet; however, cells D9 and D13 shows that we had exported the wrong data.

Figure 3-11 shows that we have opened the Store4639.xls spreadsheet; however, Cells 9 and 10 show that we had exported the wrong parameters to our spreadsheet. We must now go back and recreate our Store4639.xls file from our earlier instructions.

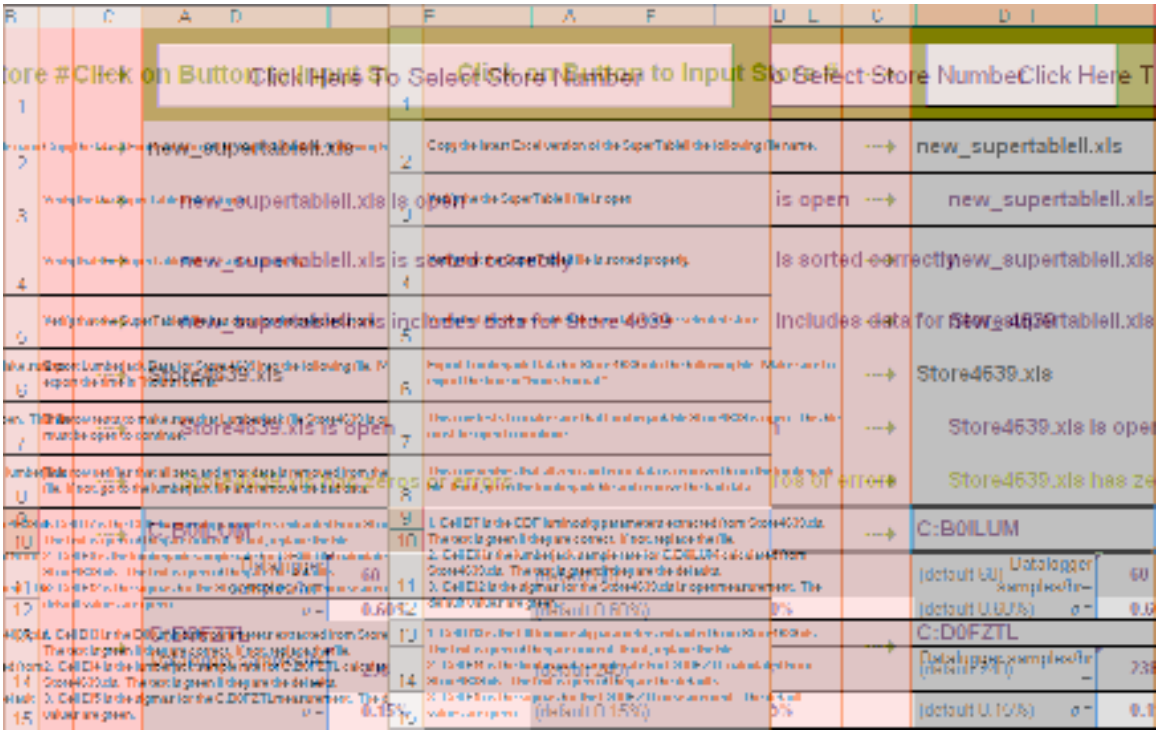


Figure 3-12: Cell D7 shows that we have our Lumberjack spreadsheet open for Store 4639, and Cells D9 and D13 show that we have exported the correct devices; however, Cell D8 still shows an error. The most likely cause are the zero Luminosity values at the beginning and end of the store.

In Figure 3-12, Cells D7 and D8 show that we have our Lumberjack spreadsheet open for Store 4639 and we have exported the correct parameters; however, Cell C8 shows that there are some problems with the data. This most likely cause of this problem is the zero luminosity readings at the starting or end of a store. Follow the directions given earlier to trim the errant data from our file.

It should be noted that cells E11 and E14 look at the Lumberjack data file and calculate the sample rate from entries in the time column. The default values are listed and cells will be posted in green if they are equal to those values. The spreadsheet was built so that if you export the lumberjack data from other dataloggers sampled at different rates, the spreadsheet will automatically adjust. Also note that cells C12 and C15 are the sigma of the luminosity reading. This is the half height of the error bar on the reading. At this point, we are using sigma values provided by Elliot McCrory[2]. These numbers are important in that they impact the scaling of our χ^2 quality of fit test.

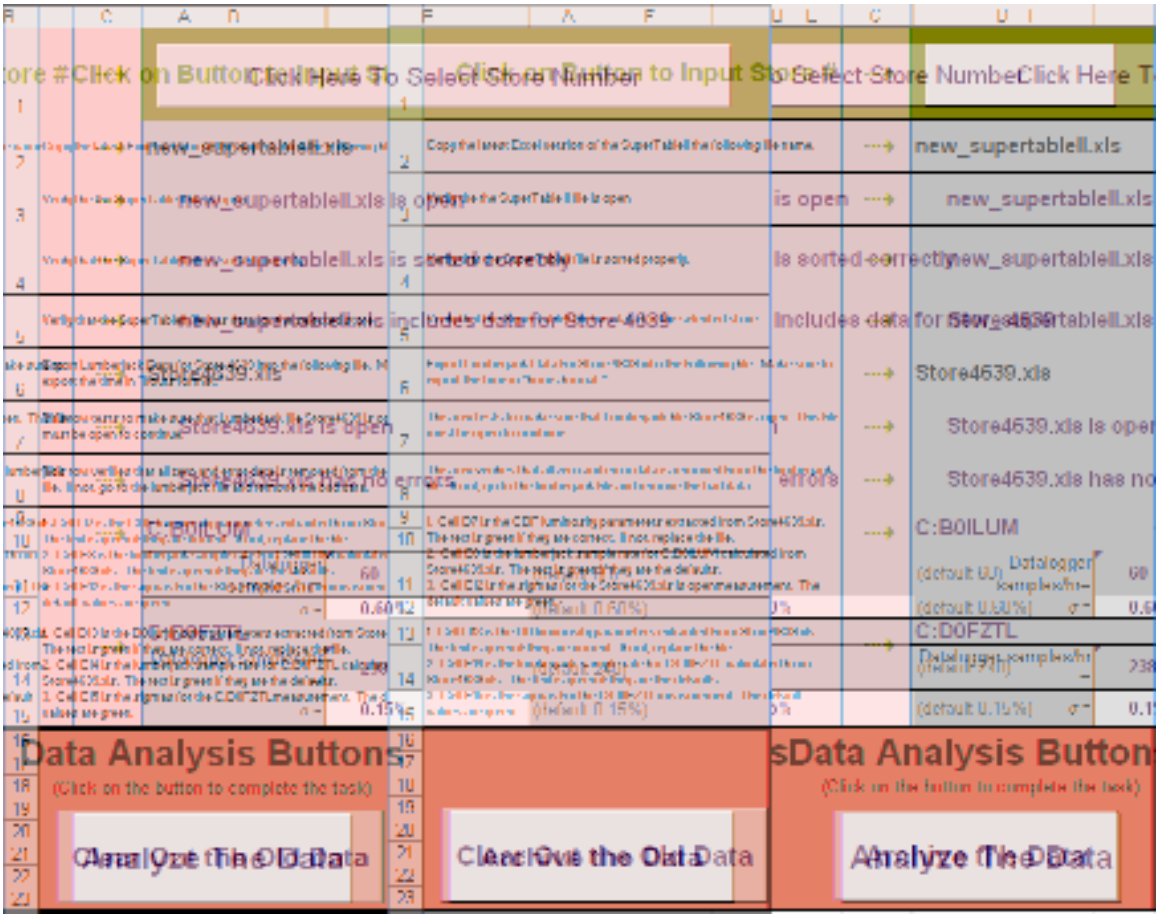


Figure 3-13: All cells in the range D1:F15 are green, which means that we have both a valid SuperTable II and Lumberjack spreadsheet open for Store 4639. We can now analyze the data from this store. In the box starting at cell A16 are interactive buttons. These buttons point to VBA scripts which do all of the data manipulation and analysis.

Cell range D2:D5 in Figure 3-13 shows that the SuperTable II spreadsheet open, is sorted correctly, and contains data from Store 4639. Cell range D6:F16 show us the Lumberjack spreadsheet is open, with the correct luminosity parameters and the data has no zero or error values. We are now ready to analyze the data. In the box starting in cell A16, there are three buttons. These buttons are attached to VBA scripts that do all of the data manipulation and analysis. Simply click on the button to complete the task assigned to that button.

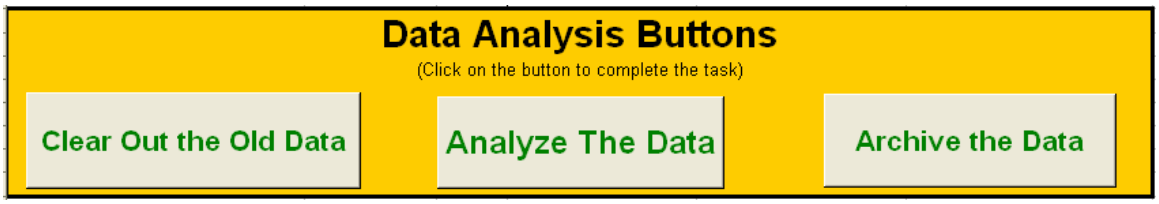


Figure 3-14: The data analysis buttons provide shortcuts to completing all of the necessary data analysis tasks.

The three interactive buttons shown in Figure 3-14 complete the following tasks. More details on the precise steps that each VBA script executes can be found in Section 4 later in this document.

- **Clear out the old data**: Runs a VBA script that clears all calculated values from cells that may be leftover from previous data analysis runs. This script is run anytime we change which store we want to analyze.
- **Analyze the data**: Runs a VBA script that analyzes the data. This is an interactive script that interfaces the user asking the user over how much lumberjack data to analyze the store and which Tevatron model to use for analysis. This script can be run repeatedly until all of the desired analysis is completed on a store.
- **Archive the data**: Runs a VBA script to archive the analyzed data. Once we have completed our analysis on a store and want to move on, we archive the data to two Excel spreadsheets: One with all of the analyzed data from this store and one with a selection portion of the analyzed data

d. Luminosity Predictor Spreadsheet: Workbooks in Detail

In the last section, we covered the procedure of how to complete a round of data analysis using the "*InputData*" workbook in the Luminosity Predictor spreadsheet. Inside of the Luminosity Predictor spreadsheet are a number of workbooks that compete all of the number crunching. We will now discuss the functions of each of these workbooks.

i. LBOE Workbook

Many of the miscellaneous functions needed for the Luminosity Predictor spreadsheet are handled in the "*LBOE*" workbook (LBOE is an old acronym borrowed from AD\Controls that stands for "little bit of everything"). Cell range B12:E16 contains initial luminosity, luminosity lifetime and store duration numbers that are

imported from the SuperTable spreadsheet. Anytime we need any of these numbers in this spreadsheet, we point back to these cells. This is done so that if our source of these parameters ever changes, we only have to edit this location. The title of this table contains the store number obtained from Cell D1. The title changes automatically when the user inputs a new store number.

SuperTableII Data for 4639		
	C:B0ILUM	C:D0FZTL
L_0 = Luminosity (e30) =	179.96	151.27
τ = Lifetime (hr) =	6.42	7.27
Store Length in hours =	37.01	

Figure 3-15: Cell Range B12:E16 displays of the data that is imported from the SuperTable II.

The SuperTable II numbers in Figure 3-15 are obtained by doing a VLOOKUP of data in our SuperTable II spreadsheet. The VLOOKUP command looks for the store number in the first column of the SuperTable II spreadsheet. Once the store number is found, the function collects the data for that row in columns #7 (store duration), #13 (SDA CDF initial luminosity), #14 (SDA D0 initial luminosity), #23 (CDF luminosity lifetime), and #24 (D0 luminosity lifetime). When the store number is changed from the interactive button in cell D1, the VLOOKUP function automatically updates the data in this table to match the new store number.

The VLOOKUP function requires that the first column of the SuperTable II spreadsheet be sorted in ascending order. This explains why we sorted the file earlier.

Lumberjack data from Store4639.xls		
	C:B0ILUM	C:D0FZTL
Valid Data Points (Last Row) =	2171	8837
Lumberjack Estimated Store Hours =	35.02	36.82

Figure 3-16: Cell Range B18:B21 examines the lumberjack data file. It determines the number of valid data points and calculates an estimated time of available data.

Figure 3-16 shows cell range B18:B21 on the "*LBOE*" Workbook. Here we look at the lumberjack luminosity data. The Excel count function is used to count the number of valid luminosity data points in the lumberjack file for this store. In this example, the CDF and D0 lumberjack data were sampled at different rates. This explains why there are more data points for D0 than there are for CDF. Using the calculated sample rate from "*InputData*" workbook cells E11 and E14, we calculate an estimated time of store data that we have to analyze. If all of the data is good, these

times should be close to the store duration number in cell C5 of the "*LBOE*" workbook, which was obtained from the SuperTable II.

# of Lumberjack Data Points at different times			
		C:B0ILUM	C:D0FZTL
Datalogger samples/hour		60	240
1 hours	1	62	242
2 hours	2	122	482
4 hours	4	242	962
6 hours	6	362	1442
8 hours	8	482	1922
10 hours	10	602	2402
12 hours	12	722	2882
14 hours	14	842	3362
16 hours	16	962	3842
18 hours	18	1082	4322
20 hours	20	1202	4802
22 hours	22	1322	5282
24 hours	24	1442	5762
26 hours	26	1562	6242
28 hours	28	1682	6722
30 hours	30	1802	7202
All Data	EOS	1822	7314

Figure 3-17: The cell range A12:D31 in the "*LBOE*" workbook contains the last row of lumberjack luminosity data at different slices in time.

Cell range A12:D31 calculates how many datalogger data points exist at various hour breakpoints. This data is used to build the cell names that correspond to different time slices of the data.

	Cell locations on our Error Workbooks at					
	Simple Exponential		Modified Exponential		Inverse Time Decay	
Hours	C:B0ILUM	C:D0FZTL	C:B0ILUM	C:D0FZTL	C:B0ILUM	C:D0FZTL
1 hour	SimpleFitE62	SimpleFitG240	ModSimpleFitE62	ModSimpleFitG240	1-1FitE62	1-1FitG240
2 hours	SimpleFitE122	SimpleFitG478	ModSimpleFitE122	ModSimpleFitG478	1-1FitE122	1-1FitG478
4 hours	SimpleFitE242	SimpleFitG954	ModSimpleFitE242	ModSimpleFitG954	1-1FitE242	1-1FitG954
6 hours	SimpleFitE362	SimpleFitG1430	ModSimpleFitE362	ModSimpleFitG1430	1-1FitE362	1-1FitG1430
8 hours	SimpleFitE482	SimpleFitG1906	ModSimpleFitE482	ModSimpleFitG1906	1-1FitE482	1-1FitG1906
10 hours	SimpleFitE602	SimpleFitG2382	ModSimpleFitE602	ModSimpleFitG2382	1-1FitE602	1-1FitG2382
12 hours	SimpleFitE722	SimpleFitG2858	ModSimpleFitE722	ModSimpleFitG2858	1-1FitE722	1-1FitG2858
14 hours	SimpleFitE842	SimpleFitG3334	ModSimpleFitE842	ModSimpleFitG3334	1-1FitE842	1-1FitG3334
16 hours	SimpleFitE962	SimpleFitG3810	ModSimpleFitE962	ModSimpleFitG3810	1-1FitE962	1-1FitG3810
18 hours	SimpleFitE1082	SimpleFitG4286	ModSimpleFitE1082	ModSimpleFitG4286	1-1FitE1082	1-1FitG4286
20 hours	SimpleFitE1202	SimpleFitG4762	ModSimpleFitE1202	ModSimpleFitG4762	1-1FitE1202	1-1FitG4762
22 hours	SimpleFitE1322	SimpleFitG5238	ModSimpleFitE1322	ModSimpleFitG5238	1-1FitE1322	1-1FitG5238
24 hours	SimpleFitE1442	SimpleFitG5714	ModSimpleFitE1442	ModSimpleFitG5714	1-1FitE1442	1-1FitG5714
26 hours	SimpleFitE1562	SimpleFitG6190	ModSimpleFitE1562	ModSimpleFitG6190	1-1FitE1562	1-1FitG6190
28 hours	SimpleFitE1682	SimpleFitG6666	ModSimpleFitE1682	ModSimpleFitG6666	1-1FitE1682	1-1FitG6666
30 hours	SimpleFitE1802	SimpleFitG7142	ModSimpleFitE1802	ModSimpleFitG7142	1-1FitE1802	1-1FitG7142
All Data	SimpleFitE2171	SimpleFitG8837	ModSimpleFitE2171	ModSimpleFitG8837	1-1FitE2171	1-1FitG8837

Figure 3-18: Cell range A33:K52 contains the cell names at various slices in time for the workbooks that we will use to calculate the difference between our luminosity curves and the lumberjack data. Due to space limitations, not all of the columns in this spreadsheet are shown.

Cell Range A33:K52 contains the cell names for our data fit spreadsheets at the time slices specified in Figure 3-7. We will need these cell names later to calculate errors between the predicted and actual data at various hour breakpoints.

Plot Labels
Store 4581 SuperTablell Numbers (C:B0ILUM & C:D0FZTL)
Store 4581 Simple Exponential (C:B0ILUM & C:D0FZTL)
Store 4581 Modified Exponential (C:B0ILUM & C:D0FZTL)
Store 4581 Inverse Time to Power (C:B0ILUM & C:D0FZTL)
Store 4581 Modified Inverse Time to Power (C:B0ILUM & C:D0FZTL)
Store 4581 C:B0ILUM Simple Exponential Fit at various store times
Store 4581 C:B0ILUM Modified Exponential Fit at various store times
Store 4581 C:B0ILUM Inverse Time to Power Fit at various store times
Store 4581 C:B0ILUM Modified Inverse Time to Power Fit at various store times
Store 4581 C:D0FZTL Simple Exponential Fit at various store times
Store 4581 C:D0FZTL Modified Exponential Fit at various store times
Store 4581 C:D0FZTL Inverse Time to Power Fit at various store times
Store 4581 C:D0FZTL Modified Inverse Time to Power Fit at various store times

Figure 3-19: Plot labels are generated based on data in the Luminosity Predictor spreadsheet. If we change the store number or other parameters the plot labels will automatically adjust.

The labels for our plots are concatenated by colleting data from various cells in the spreadsheet. The resulting plot titles are output to cell range A54:E67 in the "*LBOE*" workbook. If we change the store number, name of the fit, or luminosity parameters, the plot labels automatically adjust.

# of constants in each fit				
	Simple Exponential	Modified Exponential	Inverse Time Decay	Modified Inverse Time Decay
# Constants	2	4	3	4

Figure 3-20: Cell range B84:F86 shows the number of constants in luminosity model.

The last data displayed on the "*LBOE*" workbook are the number of constants for each available luminosity model. We will need these numbers to help calculate our χ^2 test of merit between our luminosity curve and our lumberjack data. These numbers are manually entered.

ii. Lumberjack Data import

As discussed earlier, Lumberjack data is imported from another Excel Spreadsheet named **Store{*Store Number*}.xls**, where "Store Number" is obtained through the interactive button in the "*InputData*" workbook cell D1. When analyzing the data from this store, we do not modify the original **Store{*Store Number*}.xls** file. Instead, we mirror the data and manipulate it inside of the Luminosity Predictor spreadsheet.

	A	B	C	D
1	Time	C:B0ILUM	Time	C:D0FZTL
2	3.50028	183.409	3.49	158.642
3	3.51722	180.527	3.49417	158.733
4	3.53389	183.267	3.49833	158.685
5	3.55056	179.874	3.5025	158.105
6	3.56722	180.185	3.50667	157.913
7	3.58389	179.617	3.51083	157.893
8	3.60056	178.165	3.515	157.653
9	3.6175	177.936	3.51917	157.615

Figure 3-21: Lumberjack data for store 4639 is stored in Store4639.xls.

Store{*Store Number*}.xls has four columns of data that we need to import that represent Luminosity/Time pairs for both experiments. We will complete this task using the "*ReformatLumberjack*" workbook in our Luminosity Predictor spreadsheet.

	A	B	C	D	E
1	Counter	Generated cell names for lumberjack file Store4639.xls			
2	1	store4639.xls!A1	store4639.xls!B1	store4639.xls!C1	store4639.xls!D1
3	2	store4639.xls!A2	store4639.xls!B2	store4639.xls!C2	store4639.xls!D2
4	3	store4639.xls!A3	store4639.xls!B3	store4639.xls!C3	store4639.xls!D3
5	4	store4639.xls!A4	store4639.xls!B4	store4639.xls!C4	store4639.xls!D4
6	5	store4639.xls!A5	store4639.xls!B5	store4639.xls!C5	store4639.xls!D5
7	6	store4639.xls!A6	store4639.xls!B6	store4639.xls!C6	store4639.xls!D6
8	7	store4639.xls!A7	store4639.xls!B7	store4639.xls!C7	store4639.xls!D7
9	8	store4639.xls!A8	store4639.xls!B8	store4639.xls!C8	store4639.xls!D8
10	9	store4639.xls!A9	store4639.xls!B9	store4639.xls!C9	store4639.xls!D9
11	10	store4639.xls!A10	store4639.xls!B10	store4639.xls!C10	store4639.xls!D10
12	11	store4639.xls!A11	store4639.xls!B11	store4639.xls!C11	store4639.xls!D11

Figure 3-22: Columns A through E of the "*Reformat Lumberjack*" workbook.

The A Column contains a counter to help construct names in the B, D, F, and H Columns. The B, D, F, and H columns in the "*Reformat Lumberjack*" workbook construct the file and cell names for the A, B, C, and D Columns of **Store{*Store Number*}.xls**. These names change depending on what store number we have entered using the interactive button in cell D1 of the "*InputData*" workbook. If we change the store number, the names in these cells automatically change.

	F	G	H	I	J	K	L	M
1	Raw Time/Luminosity Pairs for Store4639				Recalculated Time/Luminosity Pairs for Store4639			
2	Time	C:B0ILUM	Time	C:D0FZTL	Time	C:B0ILUM	Time	C:D0FZTL
3	19.1969	178.701	19.1889	151.935	0.00	178.70	0.00	151.94
4	19.2136	179.427	19.1931	151.917	0.02	179.43	0.00	151.92
5	19.2303	177.914	19.1972	151.563	0.03	177.91	0.01	151.56
6	19.2469	176.607	19.2014	151.749	0.05	176.61	0.01	151.75
7	19.2636	170.58	19.2056	151.357	0.07	170.58	0.02	151.36
8	19.2803	177.148	19.21	151.266	0.08	177.15	0.02	151.27
9	19.2972	175.901	19.2142	151.179	0.10	175.90	0.03	151.18
10	19.3139	174.091	19.2183	151.007	0.12	174.09	0.03	151.01
11	19.3306	173.243	19.2225	151.041	0.13	173.24	0.03	151.04

Figure 3-23: Columns F through M of the "*ReformatLumberjack"* workbook.

The F, G, H, and I Columns of the "*Reformat Lumberjack*" workbook use the Excel INDIRECT command with the values in columns B, C, D and E. This provides a mirror of the **Store{*Store Number*}.xls** A, B, C, and D Columns. Again, if we change the desired store number from the interactive button in cell D1 of the "*InputData*" workbook, these cells change to look at the spreadsheet from that store.

Columns F through I gives us time/luminosity pairs, but they are not quite ready for data analysis. Note that the times in columns F and I are in hours starting at the hour that the store began. We instead want to construct our time columns to be the number of hours since the store started. We use the following IF statement to construct the time columns.

=IF(AND(ExtractLumberjack!C3>0, NOT(ExtractLumberjack!C3=" "),NOT(ExtractLumberjack!C3=" "")), ExtractLumberjack!C3-ExtractLumberjack!C3, NA())

The verification that a cell does not have 7 or 8 blank spaces was added after it was discovered that sometimes the Lumberjack data has cells with 7 or 8 blank spaces where no real data exists. If this check is not added, some of my later calculations are give errors since they do not know how to handle the empty spaces.

We now have successfully imported our lumberjack store data. Columns J through M of the "*Reformat Lumberjack*" workbook are now time/luminosity pairs for CDF and D0 with the time columns starting at zero. Now that we have the luminosity data for a specified store, we will try to fit our Collider luminosity model equations to the data to see how well we can make them agree.

iii. Input Data for the Luminosity Models

We now turn our attention to building curves for the four luminosity models that were outlined in Equations (1) through (4) of Section 2 in this document.

We first create a location for the constants in our four models. We do this on the "*FitNumbers*" Workbook.

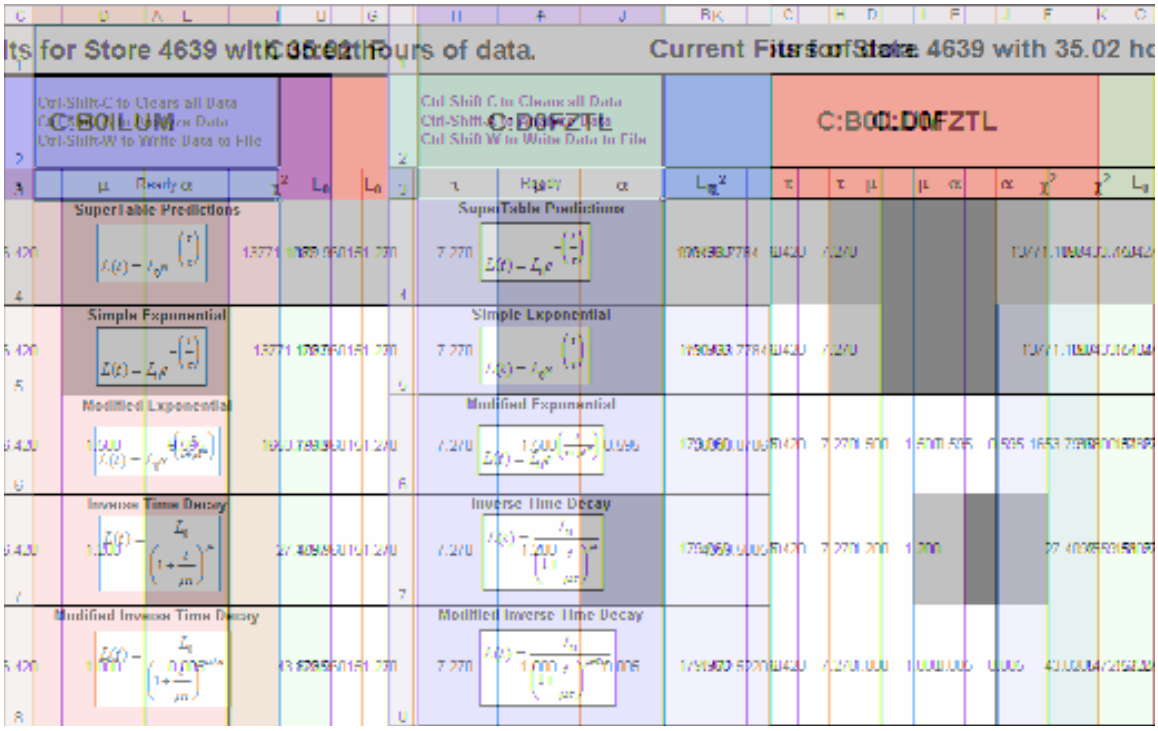


Figure 3-24: The "FitNumbers" workbook in the Luminosity Predictor spreadsheet.

Columns B through F are the constants for CDF and columns G through K are the same for D0. Recall from Equations (1) through (4), that we had.

- L_0 = Initial Luminosity
- τ = Luminosity Lifetime
- μ = constant (if applicable)
- α = constant (if applicable)
- χ^2 = chi square test of merit (We will see later these cells have a blue background because they are mirrored from another location in this spreadsheet).

Each row of the "FitNumbers" workbook contains constants from a different Collider luminosity model.

- **Row 4**: SuperTable II numbers that were derived from the Simple Exponential Fit of Equation (1). These cell backgrounds are colored grey since we do not change these values.
- **Row 5**: Simple Exponential Fit of Equation (1)
- **Row 6**: Modified Exponential Fit of Equation (2)
- **Row 7**: Time Decay Model of Equation (3)
- **Row 8**: Modified Time Decay Model of Equation (4)

It is important to note that we never manually change anything in this workbook. The interactive buttons in the "*InputData*" workbook point to VBA macros that complete all of the data analysis for us. The analysis script will automatically minimize the χ^2 value for each model by varying the constants for that model. For example, if we chose to analyze the Modified Exponential luminosity model for CDF, the script would vary the parameters in the "*FitNumbers*" workbook cell range B6:E6 to minimize the χ^2 value in "*FitNumbers*" workbook cell F6.

iv. The Luminosity Models

There are four separate workbooks dedicated to calculating the CDF and D0 luminosity/time pairs for each of the luminosity models given in Equations (1) through (4), with one workbook dedicated to each luminosity model. Recall, that the "*ReformatLumberjack*" workbook contains the CDF and D0 luminosity/time pairs from the lumberjack data. For each model, we calculate the CDF and D0 luminosity at each time value given in the "*ReformatLumberjack*" workbook along with the constants from the "*FitNumbers*" workbook. The result is columns A through E in each of our model workbooks contain the CDF and D0 time/luminosity pairs from the model equation at all of the same times as the time/luminosity pairs gathered from the lumberjack data. We can then compare the luminosity values predicted by the model against the luminosity values from the lumberjack data. Columns E and F of our model workbooks contain the square of the differences between the measured CDF and D0 luminosity and the CDF and D0 luminosity calculated from our model, divided by the square of the sigma of our measurement. Equations (5) and the data from these columns will be used to determine our χ^2 quality of fit test. We will now examine the workbooks from each luminosity model.

1. SuperTable II Prediction:

We start by plugging in the initial luminosity and luminosity lifetime numbers into our simple exponential decay model that was given in Equation (1). This is done in the "*SuperTable Predictions*" Workbook.

	A	B	C	D	E	F
1	**Store 4639 SuperTable Predictions** $L(t) = L_0 e^{-\left(\frac{t}{\tau}\right)}$				**Measured - Predicted** $\frac{(M_i - L(x_i))^2}{\sigma^2}$	
2	hours from start of store	C:B0ILUM Init Lum=179.96 Lifetime=6.42	hours from start of store	C:D0FZTL Init Lum=151.27 Lifetime=7.27	C:B0ILUM Errors $\frac{(M_i - L(x_i))^2}{\sigma^2}$	C:D0FZTL Errors $\frac{(M_i - L(x_i))^2}{\sigma^2}$
3	0.00	179.960	0.000	151.270	1.379	8.514
4	0.02	179.492	0.004	151.183	0.004	10.386
5	0.03	179.026	0.008	151.097	1.086	4.194
6	0.05	178.564	0.012	151.010	3.410	10.537
7	0.07	178.100	0.017	150.923	53.985	3.656
8	0.08	177.637	0.021	150.832	0.212	3.665
9	0.10	177.170	0.025	150.744	1.446	3.671
10	0.12	176.710	0.029	150.659	6.287	2.354

Figure 3-25: The "SuperTable-Prediction" workbook calculates time/luminosity pairs in columns A through D. The times are a mirror of the times in our lumberjack date file. Columns E and F show the square of the differences between the measured and predicted divided by the square of the sigma of our measurement. These columns will be used to determine our χ^2 merit of fit.

This method was determined not to be very useful, but is still completed for comparison sake.

2. Simple Exponential Fit

The "*SimpleFit*" Workbook uses Equation (1) and the corresponding input parameters from the "*FitNumbers*" Workbook to create Luminosity/Time pairs. The "*FitNumbers*" values are modified via a script to optimize the agreement between the lumberjack luminosity values and the calculated luminosity values.

	A	B	C	D	E	F
1	Simple Exponential $L(t) = L_0 e^{-\left(\frac{t}{\tau}\right)}$				Measured - Predicted $\frac{(M_i - L(x_i))^2}{\sigma^2}$	
2	hours from store start	C:B0ILUM Init Lum=117.56 Lifetime=19.2 chi sq=255.59	hours from store start	C:D0FZTL Init Lum= 102.04 Lifetime=19.27 chi sq=4129.08	C:B0ILUM Errors $\frac{(M_i - L(x_i))^2}{\sigma^2}$	C:D0FZTL Errors $\frac{(M_i - L(x_i))^2}{\sigma^2}$
3	0.00	117.5560979	0.000	102.042966	3252.101	47925.253
4	0.02	117.4536781	0.004	102.0207126	3313.818	47944.765
5	0.03	117.3517471	0.008	101.9990034	3218.706	47529.553
6	0.05	117.2503156	0.012	101.9767696	3137.769	47812.341
7	0.07	117.1483617	0.017	101.9545406	2725.450	47348.868
8	0.08	117.0464963	0.021	101.9312583	3197.390	47275.962
9	0.10	116.9435012	0.025	101.9090392	3120.606	47206.193
10	0.12	116.841814	0.029	101.8873538	3003.886	47025.539
11	0.13	116.7402152	0.034	101.8651443	2954.785	47112.009

Figure 3-26: The "*SimpleFit*" workbook calculates values for the Simple Exponential luminosity model. Columns A and C mirror the time values in our lumberjack file. Columns B and D are the calculated CDF and D0 luminosity based on the simple Exponential Model given in Equation (1) with the constants from the "*FitNumbers*" workbook. Columns E and F calculate an error between the measured and calculated luminosity numbers. These errors are used to calculate our χ^2 quality of fit test.

3. Modified Exponential Fit

The "*ModSimpleFit*" Workbook uses Equation (2) and the corresponding input parameters from the "*FitNumbers*" Workbook to create Luminosity/Time pairs. The "*FitNumbers*" values are modified via a script to optimize the agreement between the lumberjack luminosity values and the calculated luminosity values.

	A	B	C	D	E	F
1	Modified Exponential $L(t) = L_0 e^{-\left(\frac{t}{\tau + \mu t^{\alpha}}\right)}$				Measured - Predicted $\frac{(M_i - L(x_i))^2}{\sigma^2}$	
2	from store start	C:B0ILUM Init Lum=181.92 Lifetime=3.33 u=1.79 a=0.56 chi sq=1.19	from store start	C:D0FZTL Init Lum=153.25 Lifetime=4.56 u=1.34 a=0.62 chi sq=6.69	C:B0ILUM Errors $\frac{(M_i - L(x_i))^2}{\sigma^2}$	C:D0FZTL Errors $\frac{(M_i - L(x_i))^2}{\sigma^2}$
3	0.00	181.9231475	0.000	153.249986	9.031	33.292
4	0.02	181.0608569	0.004	153.110169	2.303	27.416
5	0.03	180.2435268	0.008	152.9752061	4.762	38.586
6	0.05	179.4605122	0.012	152.8381039	7.252	22.893
7	0.07	178.6965439	0.017	152.7020009	62.890	35.096
8	0.08	177.9529163	0.021	152.5603765	0.573	32.543
9	0.10	177.2186242	0.025	152.4260308	1.559	30.240

Figure 3-27: The "*ModSimpleFit*" workbook calculates values for the Modified Exponential luminosity model of Equation (2). Columns A and C mirror the time values in our lumberjack file. Columns B and D are the calculated CDF and D0 luminosity based on the Modified Exponential Model given in Equation

(2) with the constants from the "*FitNumbers*" workbook. Columns E and F calculate an error between the measured and calculated luminosity numbers. These errors are used to calculate our χ^2 quality of fit test.

4. Inverse Time to Power Fit

The "*t-1Fit*" Workbook uses Equation (3) and the corresponding input parameters from the "*FitNumbers*" Workbook to create Luminosity/Time pairs. The "*FitNumbers*" values are modified via a script to optimize the agreement between the lumberjack luminosity values and the calculated luminosity values.

	A	B	C	D	E	F
1	Inverse Time Decay $L(t) = \frac{L_0}{\left(1+\frac{t}{\mu\tau}\right)^{\mu}}$				Measured - Predicted $\frac{(M_i - L(x_i))^2}{\sigma^2}$	
2	hours from store start	C:B0ILUM Init Lum=166.8 Lifetime=7.56 u=1.37 chi sq=6.41	hours from store start	C:D0FZTL Init Lum=144.81 Lifetime=7.56 u=1.36 chi sq=49.46	C:B0ILUM Errors $\frac{(M_i - L(x_i))^2}{\sigma^2}$	C:D0FZTL Errors $\frac{(M_i - L(x_i))^2}{\sigma^2}$
3	0.00	166.7953217	0.000	144.8132942	123.297	976.495
4	0.02	166.4277995	0.004	144.7328816	145.799	993.921
5	0.03	166.0616745	0.008	144.6544582	123.278	923.430
6	0.05	165.6991192	0.012	144.5741985	105.965	993.541
7	0.07	165.335758	0.017	144.4940159	26.255	913.773
8	0.08	164.9737717	0.021	144.4100978	131.192	912.986
9	0.10	164.6088424	0.025	144.330073	114.476	912.178
10	0.12	164.2496	0.029	144.2520277	88.768	889.345
11	0.13	163.8917104	0.034	144.1721547	80.934	919.169
12	0.15	163.5372974	0.038	144.0923583	144.389	888.627
13	0.17	163.1820837	0.042	144.0145357	126.259	845.655

Figure 3-28: The "*t-1Fit*" workbook calculates values for the Time Decay luminosity model of Equation (3). Columns A and C mirror the time values in our lumberjack file. Columns B and D are the calculated CDF and D0 luminosity based on the Time Decay Model given in Equation (3) with the constants from the "*FitNumbers*" workbook. Columns E and F calculate an error between the measured and calculated luminosity numbers. These errors are used to calculate our χ^2 quality of fit test.

5. Modified Inverse Time to Power Fit

The "*t-1ModFit*" Workbook uses Equation (4) and the corresponding input parameters from the "*FitNumbers*" Workbook to create Luminosity/Time pairs. The "*FitNumbers*" values are modified via a script to optimize the agreement between the lumberjack luminosity values and the calculated luminosity values.

	A	B	C	D	E	F
1	**Modified Inverse Time Decay** $L(t)=\frac{L_0}{\left(1+\frac{t}{\mu\tau}\right)^{\mu+1/a}}$				**Measured - Predicted** $\frac{(M_i-L(x_i))^2}{\sigma^2}$	
2	hours from store start	C:B0ILUM Init Lum=176.83 Lifetime=5.58 u=0.79 a=0.005 chi sq=1.37	hours from store start	C:D0FZTL Init Lum=150.55 Lifetime=6.17 u=0.9 a=0.004 chi sq=6.94	C:B0ILUM Errors $\frac{(M_i-L(x_i))^2}{\sigma^2}$	C:D0FZTL Errors $\frac{(M_i-L(x_i))^2}{\sigma^2}$
3	0.00	176.8345416	0.000	150.550231	3.030	36.920
4	0.02	176.3068004	0.004	150.4477428	8.400	41.572
5	0.03	175.7825153	0.008	150.3478331	3.987	28.570
6	0.05	175.2647593	0.012	150.2456278	1.605	43.621
7	0.07	174.7472591	0.017	150.1435652	16.578	28.566
8	0.08	174.2331092	0.021	150.0367952	7.521	29.348
9	0.10	173.7161773	0.025	149.9350239	4.285	30.093

Figure 3-29: The "*t-1ModFit*" workbook calculates values for the Modified Time Decay luminosity model of Equation (4). Columns A and C mirror the time values in our lumberjack file. Columns B and D are the calculated CDF and D0 luminosity based on the Modified Time Decay Model given in Equation (4) with the constants from the "*FitNumbers*" workbook. Columns E and F calculate an error between the measured and calculated luminosity numbers. These errors are used to calculate our χ^2 quality of fit test.

v. Error Sums

We recall that our χ^2 quality of fit calculations in Equation (5) sums the error terms that we calculated in the workbook for each model and then divides that number by the number of data points minus the number of constants in the model equation. We have a separate χ^2 calculation for each luminosity model. We don't end there.

One of the goals of this project is to also look at how these fits project the luminosity with varying lengths of lumberjack data. For example, we may want to see how well we project the luminosity after the first two hours of data, then after four hours of data, etc. We will have a different χ^2 calculation for every time slice of data we want to examine. We don't want to have to manually trim our lumberjack data every time we want to look at a different time slice of data, so we'll make the Luminosity Predictor spreadsheet do the work for us.

Figure 3-30: The "*ErrorSums*" workbook makes all of the possible χ^2 calculations based on the numbers input to the "*FitNumbers*" workbook and the errors calculated for each time/luminosity pair in the model workbooks.

The "*ErrorSums*" workbook calculates the χ^2 quality of fit number for each fit at each possible time slice based on the numbers input in the "*FitNumbers*" workbook and the errors calculated in the luminosity model workbooks. A script automatically copies the correct χ^2 values depending on what fit the spreadsheet user is attempting to make.

vi. Summaries

After we complete our first set of fits at one time slice, we will want to save that data away so that we can move on to our next set of fits at a different time slice for the current store. The "D0_Results" and "CDF_Results" are used to store the data after we complete a fit.

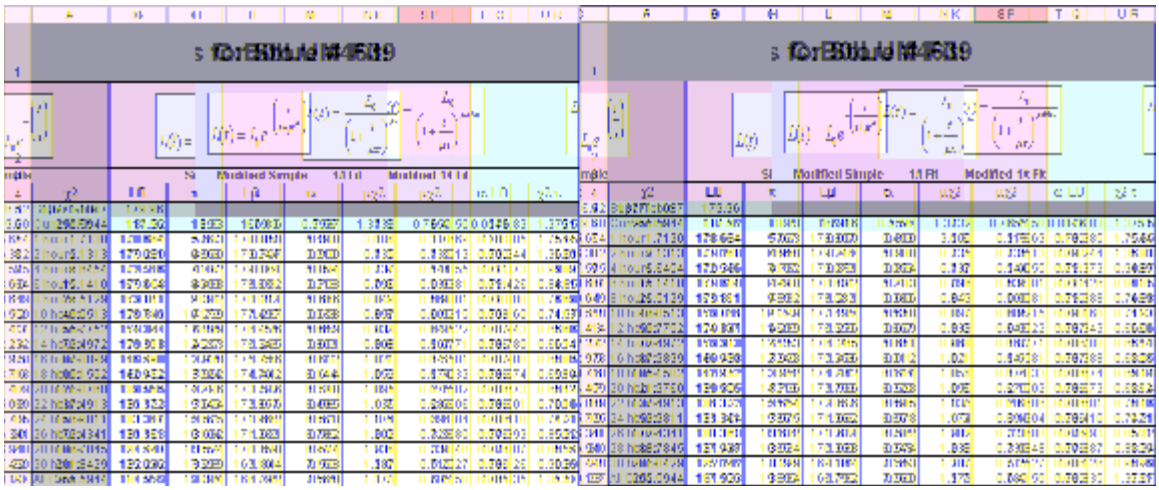


Figure 3-31: Summary of all CFD fits that were completed for Store 4639. If we chose not to complete certain fits for that store, those cells will be empty.

Figure 3-32: Summary of all D0 fits that were completed for Store 4639. If we chose not to complete certain fits for that store, those cells will be empty.

Figures 3-31 and 3-32 are summary spreadsheets that contain all of the fit model constants and χ^2 values at each time slice. Each row represents a different slice in time. For example, if we fit the store with 2 hours of lumberjack data, those fit values would be copied into row 8 of the spreadsheet. Row 6 is a mirror of whatever the current fit is on "*FitNumbers*" workbook. Row 23 is for all store data after the store has been completed.

vii. Fits over time

If data is fit over multiple time slices, it may be worthwhile to be able to plot the data for each of the fits over time to see how the predicted luminosity curves change over time. This could help us answer the following question? After how many hours of

lumberjack data would we expect to be able to make an accurate prediction of the luminosity value at the end of the store?

To add this functionality an additional workbook for each Tevatron luminosity model was added. The new workbooks take all of the values for that fit from the "*CDF_Results*" and "*D0_Results*" workbooks and create time/luminosity pairs for plotting. In this case the times are constructed from 0 to 40 hours at 0.05 hour increments. Figures 33 through 36 show the workbooks for our four luminosity models outlined in Equations (1) through (4).

	A	B	C	D	E	F	G	H	I
1	Simple Exponential fit at various times during the store $L(t) = L_0 e^{-\left(\frac{t}{\tau}\right)}$								
2	t (hours)	Predicted CDF L(t)	Predicted D0 L(t)	Predicted CDF L(t)	Predicted D0 L(t)	Predicted CDF L(t)	Predicted D0 L(t)	Predicted CDF L(t)	Predicted D0 L(t)
3	from store start	Current Fit Init Lum=118 Lifetime=19.2 chi sq=255.59	Current Fit Init Lum=102 Lifetime=19.27 chi sq=4343.43	1 hour Init Lum=179 Lifetime=5.65 chi sq=1.71	1 hour Init Lum=151 Lifetime=6.33 chi sq=2.7	2 hours Init Lum=177 Lifetime=6.3 chi sq=2.13	2 hours Init Lum=150 Lifetime=7.1 chi sq=9.42	4 hours Init Lum=173 Lifetime=7.57 chi sq=6.65	4 hours Init Lum=147 Lifetime=8.03 chi sq=39.33
4	0	117.5560979	102.042956	178.663777	151.0560774	177.0613086	149.844912	172.5461585	147.4716671
5	0.05	117.2503156	101.7784678	177.0906577	149.867117	175.6620305	148.793316	171.4109454	146.5565735
6	0.1	116.9453287	101.5146652	175.5313896	148.687515	174.2738106	147.7491	170.2832011	145.6471582
7	0.15	116.6411351	101.2515464	173.9858507	147.5171975	172.8965615	146.7122121	169.1628765	144.743386
8	0.2	116.3377328	100.9891095	172.4539202	146.3560917	171.5301966	145.682601	168.0499227	143.8452219

Figure 3-33: Time/Luminosity pairs from 0 to 40 hours based on the Equation (1) simple exponential fit values obtained by running the data fitting scripts in this spreadsheet. Parameters for the luminosity equation are take from the "*CDF_Results*" and "*D0_Results*" workbooks.

	A	B	C	D	E	F	G	H	I
1	Modified Exponential fit at various times during the store $L(t) = L_0 e^{-\left(\frac{t}{s+\mu t^a}\right)}$								
2	t (hours)	Predicted CDF L(t)	Predicted D0 L(t)	Predicted CDF L(t)	Predicted D0 L(t)	Predicted CDF L(t)	Predicted D0 L(t)	Predicted CDF L(t)	Predicted D0 L(t)
3	from store start	Current Fit Init Lum=182 Lifetime=3.33 u=1.79 a=0.56 chi sq=1.19	Current Fit Init Lum=153 Lifetime=4.56 u=1.34 a=0.62 chi sq=455.99	1 hour Init Lum=179 Lifetime=5.6 u=0.05 a=0.3 chi sq=1.78	1 hour Init Lum=152 Lifetime=5.05 u=1.25 a=0.86 chi sq=0.41	2 hours Init Lum=179 Lifetime=4.95 u=0.71 a=0.9 chi sq=1.37	2 hours Init Lum=152 Lifetime=3.23 u=3 a=0.3 chi sq=0.46	4 hours Init Lum=179 Lifetime=4.79 u=0.86 a=0.8 chi sq=1.01	4 hours Init Lum=152 Lifetime=4.7 u=1.62 a=0.48 chi sq=0.49
4	0	181.9231475	153.249986	178.6842912	151.9029727	179.2795953	152.1364258	179.3361761	151.7945806
5	0.05	179.4605122	151.6502299	177.1024886	150.4348652	177.4949394	150.4369808	177.5040379	150.3073797
6	0.1	177.2315103	150.1395266	175.5372987	149.0240092	175.7568568	148.9561492	175.7329032	148.9198503
7	0.15	175.1495501	148.692461	173.9873176	147.6627067	174.061494	147.5756328	174.0139251	147.5942796
8	0.2	173.1815512	147.298254	172.4520324	146.3463823	172.4063417	146.2648089	172.3422657	146.3173368

Figure 3-34: Time/Luminosity pairs from 0 to 40 hours based on the Equation (2) modified simple exponential fit values obtained by running the data fitting scripts in this spreadsheet. Parameters for the luminosity equation are take from the "*CDF_Results*" and "*D0_Results*" workbooks.

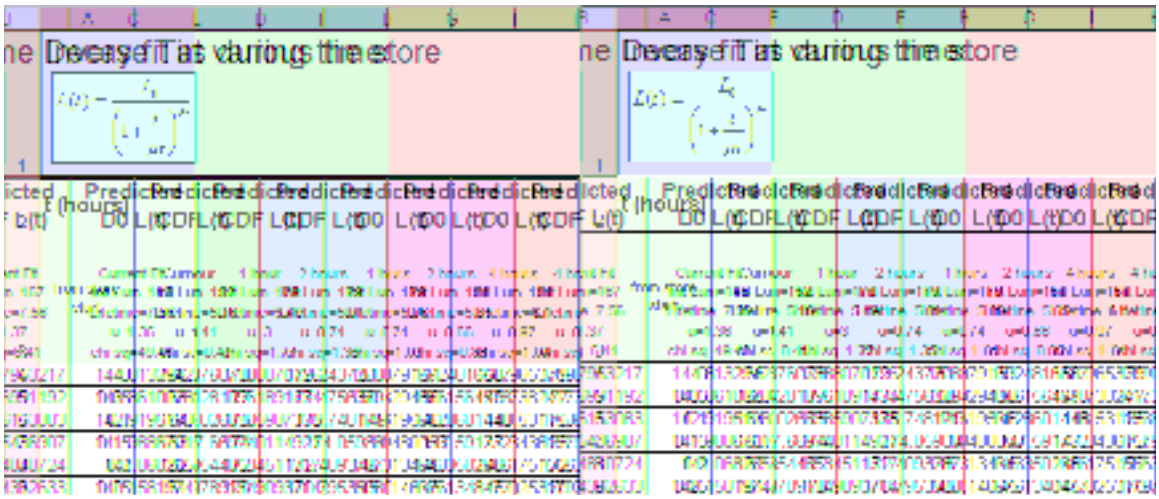


Figure 3-35: Time/Luminosity pairs from 0 to 40 hours based on the Equation (3)time decay fit values obtained by running the data fitting scripts in this spreadsheet. Parameters for the luminosity equation are take from the "*CDF_Results*" and "*D0_Results*" workbooks.

Figure 3-36: Time/Luminosity pairs from 0 to 40 hours based on the Equation (4) modified time decay fit values obtained by running the data fitting scripts in this spreadsheet. Parameters for the luminosity equation are take from the "*CDF_Results*" and "*D0_Results*" workbooks.

viii. Plots

We generate two different types of plots for each of our luminosity models. The first set of plots show the current store data and the current fits. The second set of data shows how the fit equations change over time. Both types of plots will be shown in the data results section below.

e. VBA Scripts

We have discussed all of the major internals of the Luminosity Predictor spreadsheet. There are portions of the spreadsheet that are complex. The goal is to make the analysis of the data as streamlined as possible, with as little user interaction as possible. To complete this goal, all data analysis occurs by running three VBA Scripts launched through interactive buttons in the "*InputData*" workbook. Below is an explanation of each of these VBA scripts.

i. Clear All Data Script (Ctrl-Shift-C)

If we are switching our analysis from one store to another store, we want to clear out any old data from other store before we begin. The Clear All Data script completes this task and is launched by clicking on the "*Clear Out the Old Data*" button in the "*InputData*" workbook, or by pressing the keyboard shortcut Ctrl-Shift-C. This script completes the following steps:

- Clears out all entries in the "*CDF_Results*" and "*D0_Results*" workbooks. We are starting over!
- Enters an initial guesses in the "*FitNumbers*" Workbook for each of constants for each of the models. The SuperTable II initial luminosity and luminosity lifetimes are used as the initial guess for those values in each of the fits.

Since this script uses the SuperTable II data, we want to be sure to have completed the steps outlined in Section 3.c.i before running this script.

ii. Analyze the Data Script (Ctrl-Shift-A)

All of the data analysis occurs with the Analyze Data Script, which is launched by pressing the "*Analyze the Data*" button in the "*InputData*" workbook, or by pressing the keyboard shortcut Ctrl-Shift-A. The spreadsheet completes the following tasks.

- Prompts the user with a message box asking how many hours of data to analyze.
    - Choices include 1, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 or 30 hours. This option cuts the lumberjack data to the number of hours that you specify.
    - We can also select all available data, which will make the fit with the entire contents of the

lumberjack file. When analyzing a store that is in progress, this will be the most likely option.
    - We can also select all data over all times. This option is intended to analyze a store after the fact. It loops through all data fits at every available time slice. Warning, this option normally takes on the order of 8 hours to complete.
- Next the user is prompted for which luminosity model to make the fit. Options include the Equation (1) simple exponential fit, the Equation (2) modified exponential fit, the Equation (3) time decay fit, the Equation (4) modified time decay fit, or all four fits. It normally takes on the order of 5 minutes to complete the fit on a single luminosity model for a single time slice and on the order of 20 minutes to complete the fit for all luminosity models for a single time slice.
- Based on the user input, the desired χ^2 calculations from the "*ErrorSums*" workbook are copied to the "*FitNumbers*" workbook.
- Based on the user input, the desired fits are completed for the selected luminosity models over the selected slice in time. This is completed by using the Excel Solver to minimize our χ^2 calculations by changing the model parameters in the "*FitNumbers*" workbook.
- The results of any data fit(s) completed are then copied over to the "*CDF_Results*" and "*D0_Results*" workbooks.

We can now examine our plots and re-run the script to analyze more data from this store. When using this script to look at a store that is still in progress, we periodically update our lumberjack data file with the latest data and re-run the analysis until we have enough lumberjack data to be confident of our fit.

iii. Archive (Write) the Data Script (Ctrl-Shift-W)

Once we have completed analysis on one store, we will want to save this data away before we move on to analyzing the next store. The Archive the Data Script completes this task and is called by pressing the "*Archive the Data*" button in the "*InputData*" workbook, or by pressing the Ctrl-Shift-W keyboard shortcut. This script does the following:

a. Archive the store data in a file dedicated to the current store.
    - Copies the "values" of each cell in each workbook of the Luminosity Predictor spreadsheet to the template store-fit-data.xls.

- Saves the file as **store####-fit-data.xls** (where #### is the store number).

b. Copy the end of store results to a spreadsheet containing the end of store results for all stores.
   - Opens the file store-fit-summary.xls.
   - Creates a new row.
   - Copies the end of store fit data from the "*CDF_Results*" and "*D0_Results*" workbooks into the newly created row.
   - Saves and closes the file.

a. Archive Data Files

When we archive our data set via or VBA script, we write to two Excel files. We will briefly mention the function of both files..

i. Individual store{####}-fit-data.xls files

We archive each set of store data in a spreadsheet named **store{####}-fit-data.xls**. The "value" of each cell from each workbook from the Luminosity Predictor spreadsheet is copied here. The workbook names mirror that of the Luminosity Predictor spreadsheet. The only difference is that we only copy the "values" of the cells. This means the equations and references are not copied, just the results. This allows us to remove the VBA scripts from our archive file and reduce the file size down from 50MB to 12MB. In addition, by removing the calculations, the spreadsheet opens must faster. The plots are left in place so that they can be easily examined later on.

If we want to backtrack and reanalyze the data for this store, or analyze it with more time slices, we can do so fairly easily. The "*InputData*" workbook of the Luminosity Predictor spreadsheet has an "*Import Data From Previously Analyzed Store*" button that calls a VBA script to open the archived in **store####-fit-data.xls** file and copy the "*CDF_Results*" and "*D0_Results*" workbook data into the same workbooks in Luminosity-Predictor.xls. The Luminosity Predictor is then ready to go.

ii. Store-fit-summary.xls file

We want to examine results across multiple stores, so there is also a store-fit-summary.xls spreadsheet. Each row in this spreadsheet contains the results from one store. The end of store fit numbers for both CDF and D0 using all four Tevatron luminosity models are included in this file. We can use this spreadsheet to plot store data from multiple stores.

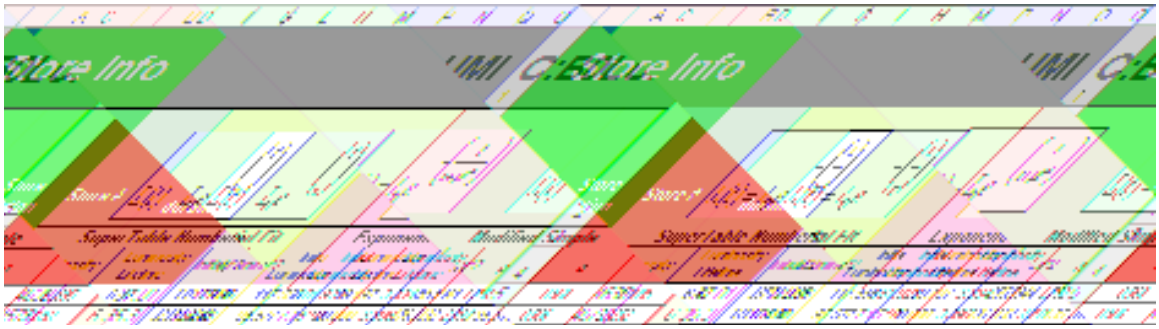


Figure 3-37: The summary spreadsheet that contains our fit data for our four luminosity models. Not all columns are displayed due to lack of space.

In addition, this spreadsheet has plots that allow us to compare how the model constants change with datalogger time sample size across multiple stores.

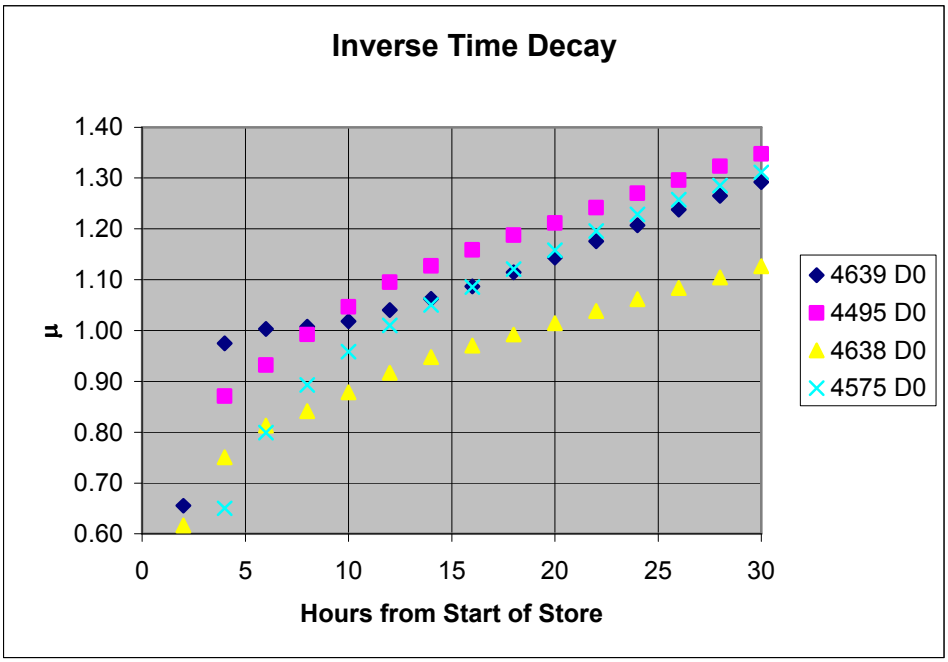


Figure 3-38

Figure 3-38 shows number of hours of luminosity data to fit on the x-axis and the constant μ from Equation (3) on the y-axis. We see that across multiple stores, the behavior of the fit for this parameter is consistent.

4. How luminosity fits improve over time

Now that we have built our spreadsheet tool, it is time to put it to work. This section graphically shows how the predictions from each of luminosity models given by Equations (1) – (4) fairs with different amounts of available lumberjack data. I took the first hour of lumberjack luminosity data, and used the Luminosity Predictor to make a predicted luminosity curve for the store. I then repeated with the first two, four, six, …, twenty-eight, and thirty hours of lumberjack luminosity data. I then plotted each of the predictions as well as the complete set of lumberjack luminosity data. The intent is to see how well the predicted curves follow the actual lumberjack data, and to see how well our predictions improve with an increased amount lumberjack data.

One might expect that the luminosity prediction for a store after only having the first hour of lumberjack luminosity data would not be as precise as a luminosity prediction made with multiple hours of Lumberjack data. If a Tevatron Luminosity Model equation is an accurate representation of luminosity, we would expect the fits to get better and better with more luminosity data, and we would expect be able to make a very good fit once we have all of the lumberjack luminosity data from the store. If we are to use the Luminosity Predictor to predict our luminosity behavior during a store, it would be worthwhile to understand how many hours of store data are needed to make a reasonable prediction of the end of store luminosity.

Again, I used Store 4639 for the plots. Other stores showed similar results.

a. Simple Exponential Fit (Equation 1)

The Simple exponential fit proves not be a good predictor of the luminosity behavior. Figure 4.1 shows the CDF lumberjack data for the entire store (thick blue line) and each of the predicted luminosity curves.

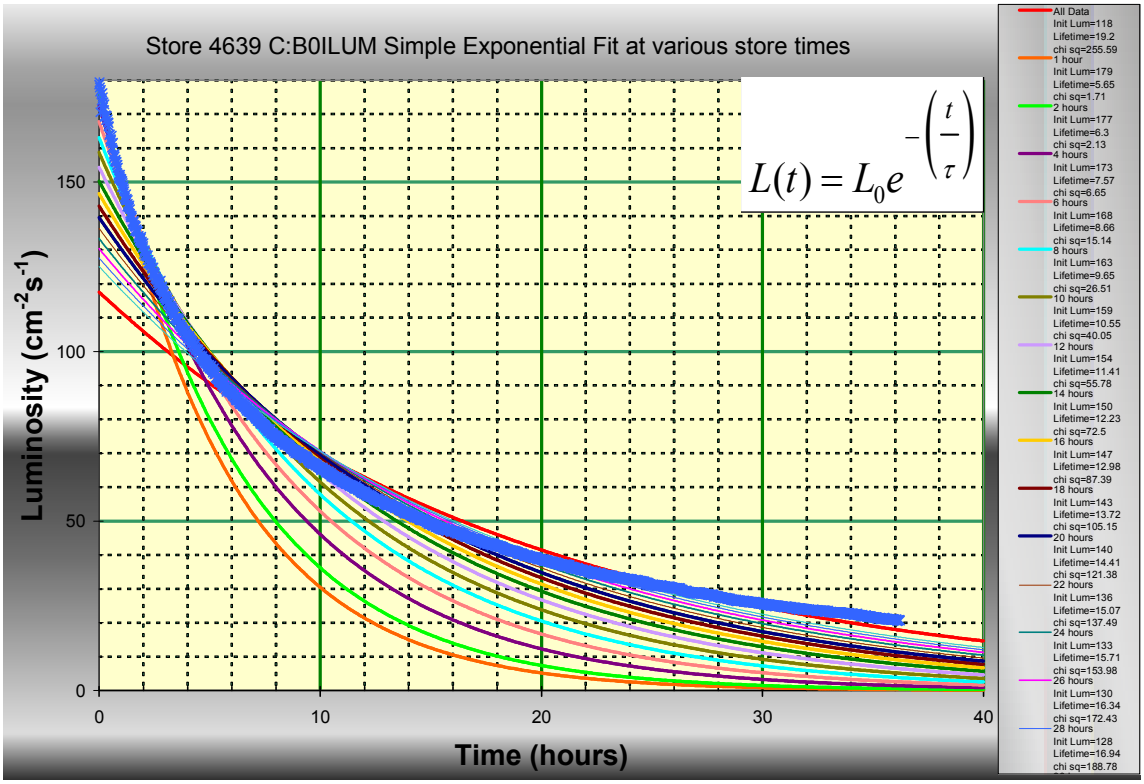
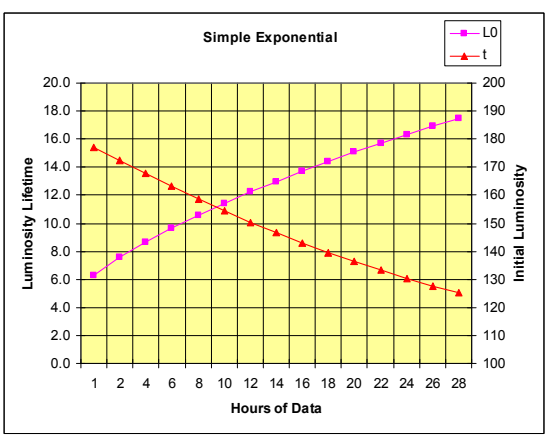
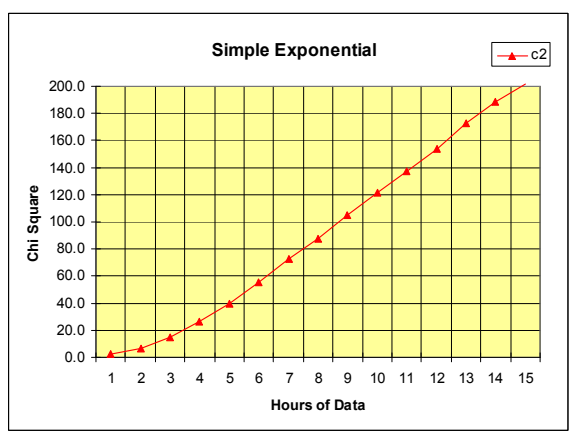


Figure 4-1: Simple Exponential fit of CDF Luminosity data for store 4639 taken with varying samples of lumberjack data.

It is clear that the fits never match the store data. The orange curve shows the fit after one hour of Lumberjack data and the red curve shows the fit using the lumberjack data for the entire store. The thicker blue line is the lumberjack data for the store. The other curves are fits with varying amounts of lumberjack data. We see that the early predictions match the lumberjack data at the beginning of the store, but are very far off at the end of the store. The later predictions are closer at the end of the store, but are further away at the start of the store. Overall, the simple exponential fit is not a good model of Luminosity behavior.



CDF Initial Luminosity and Lifetime



Chi Square

Figure 4-2: These plots show how the Simple Exponential fit (Equation 1) predictions change over the number of hours of lumberjack luminosity data used to make the prediction. We are looking at the CDF data for Store 4639. The x-axis in both plots is the number of hours of luminosity data from the lumberjack (1= first hour of data, 2= first two hours of data, etc…). The plot on the left shows the Initial Luminosity and Luminosity Lifetime for each sample of luminosity data. The plot on the right shows the chi square value for those fits.

Figure 4-2 shows how the two constants in our fit change as we increase the amount of Lumberjack luminosity data. We see that neither of our constants reaches a stable value and the chi square value gets larger as you increase the amount of lumberjack data. In fact the chi square value exceeds 200.0 when fitting 30 hours of luminosity data. This shows that the simple luminosity fit (Equation 1) is not a good model for our luminosity behavior. The results for the D0 Luminosity fits showed similar results.

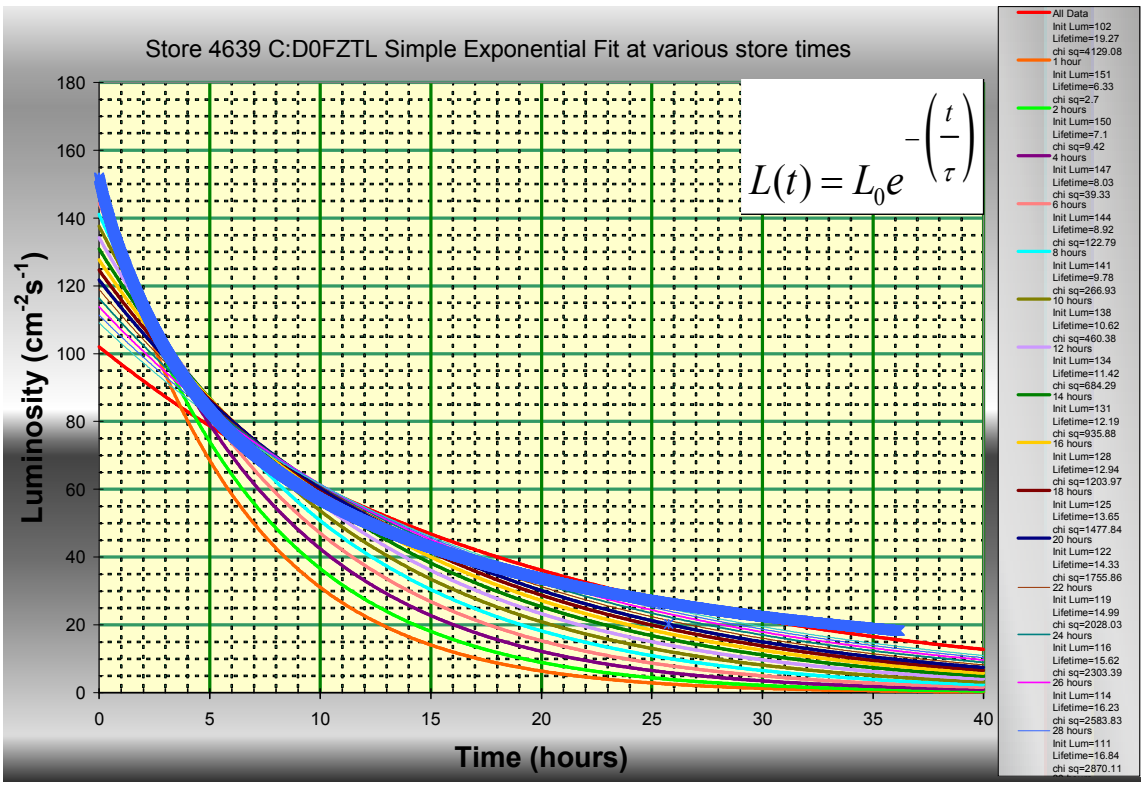
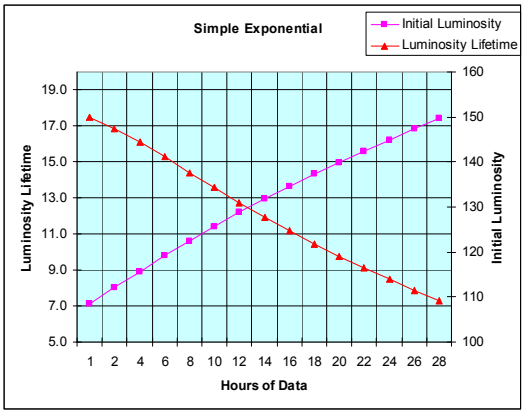
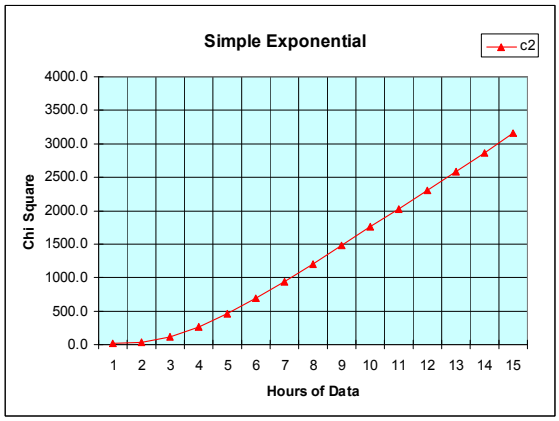


Figure 4-3: Simple Exponential fit of D0 Luminosity data for store 4639 taken with varying samples of lumberjack data.

Figure 4.3 shows the simple exponential fit predictions for D0 luminosity for Store 4639. Similar to the CDF fits, the simple exponential fit proves to be a poor predictor of luminosity behavior.



D0 Initial Luminosity and Lifetime



Chi Square

Figure 4-4: These plots show how the Simple Exponential fit (Equation 1) predictions change over the number of hours of lumberjack luminosity data used to make the prediction. We are looking at the D0 data for Store 4639. The x-axis in both plots is the number of hours of luminosity data from the lumberjack (1= first hour of data, 2= first two hours of data, etc…). The plot on the left shows the Initial Luminosity and Luminosity Lifetime for each sample of luminosity data. The plot on the right shows the chi square value for those fits.

Figure 4-4 shows how the two constants in our fit change as we increase the amount of Lumberjack luminosity data. We see that neither of our constants reaches a stable value and the chi square value gets larger as you increase the amount of lumberjack data. In fact the chi square value exceeds 3000.0 when fitting 30 hours of luminosity data. This shows that the simple luminosity fit (Equation 1) is not a good model for our luminosity behavior

b. Modified Exponential Fit (Equation 2)

The modified exponential fit proves to be a very good predictor of the luminosity behavior. Figure 4.5 shows the CDF lumberjack data for the entire store (thick blue line) and each of the predicted luminosity curves.

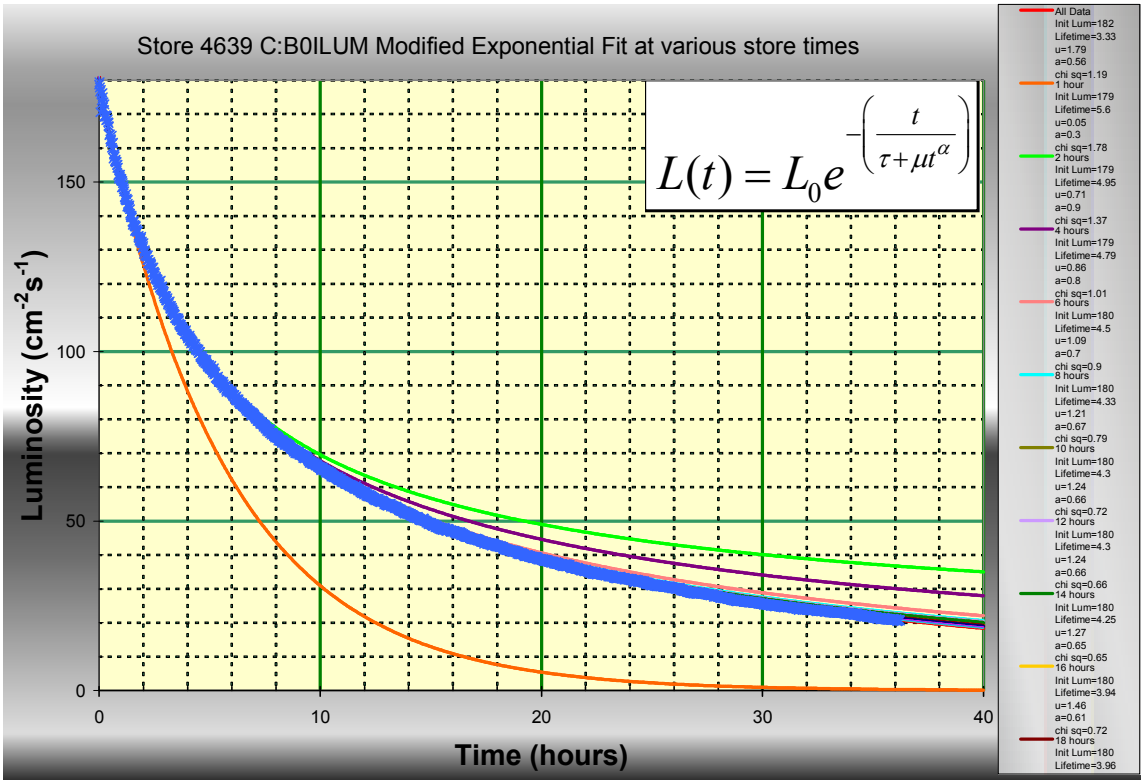
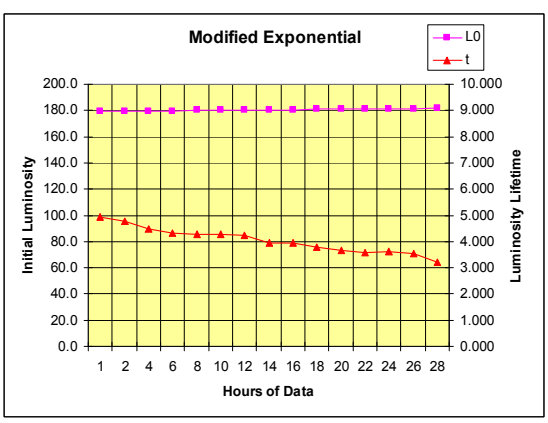
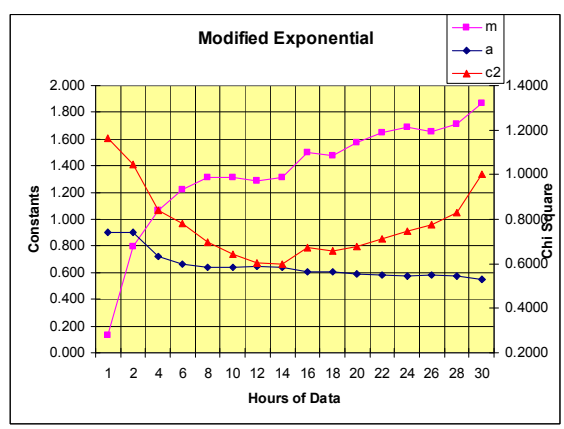


Figure 4-5: Modified Exponential fit of CDF Luminosity data for store 4639 taken with varying samples of lumberjack data.

The thick blue trace is the lumberjack data for the entire store. The orange trace is the prediction after the first hour of lumberjack data, the bright green trace is the prediction after the first two hours of lumberjack data, the maroon trace is the prediction after the first four hours of lumberjack data, and the peach colored trace is the prediction after the first six hours of lumberjack data. Once we get past about eight hours of lumberjack data, the predicted curves and real data start to match fairly well.



CDF Initial Luminosity and Lifetime



Constants and Chi Square

Figure 4-6: These plots show how the Modified Exponential fit (Equation 2) predictions change over the number of hours of lumberjack luminosity data used to make the prediction. We are looking at the CDF data for Store 4639. The x-axis in both plots is the number of hours of luminosity data from the

lumberjack (1= first hour of data, 2= first two hours of data, etc…). The plot on the left shows the Initial Luminosity and Luminosity Lifetime for each sample of luminosity data. The plot on the right shows the chi square value for those fits.

The results in Figure 4-6 show some interesting features. We see that over the first six hours of lumberjack data, the model constants are changing from fit to fit. This tells us that our model is not adequate at that point. Between 8 and 14 hours of lumberjack data the model appears to be starting to give repeatable results. The values of our constants not changing much from fit to fit at this point, with the luminosity lifetime ~4.2 hours, α ~1.3, and μ ~ .64. χ^2 values for those fits are under well 1.0 indicating that this model is doing a good job of fitting the data. Between the 14 and 16 hour point, we see an interesting feature in the fits. The constants start to change by small amounts from fit to fit. χ^2 values remain under 1.0, so the fit is still good. It is possible that we have a few bad data points later on in the store, throwing off the numbers a little. Or maybe, this luminosity model is not as accurate in later portions of the store. These are questions that would be of interest to address. Later we will check the model behavior across different stores to see if this is representative behavior for this model.

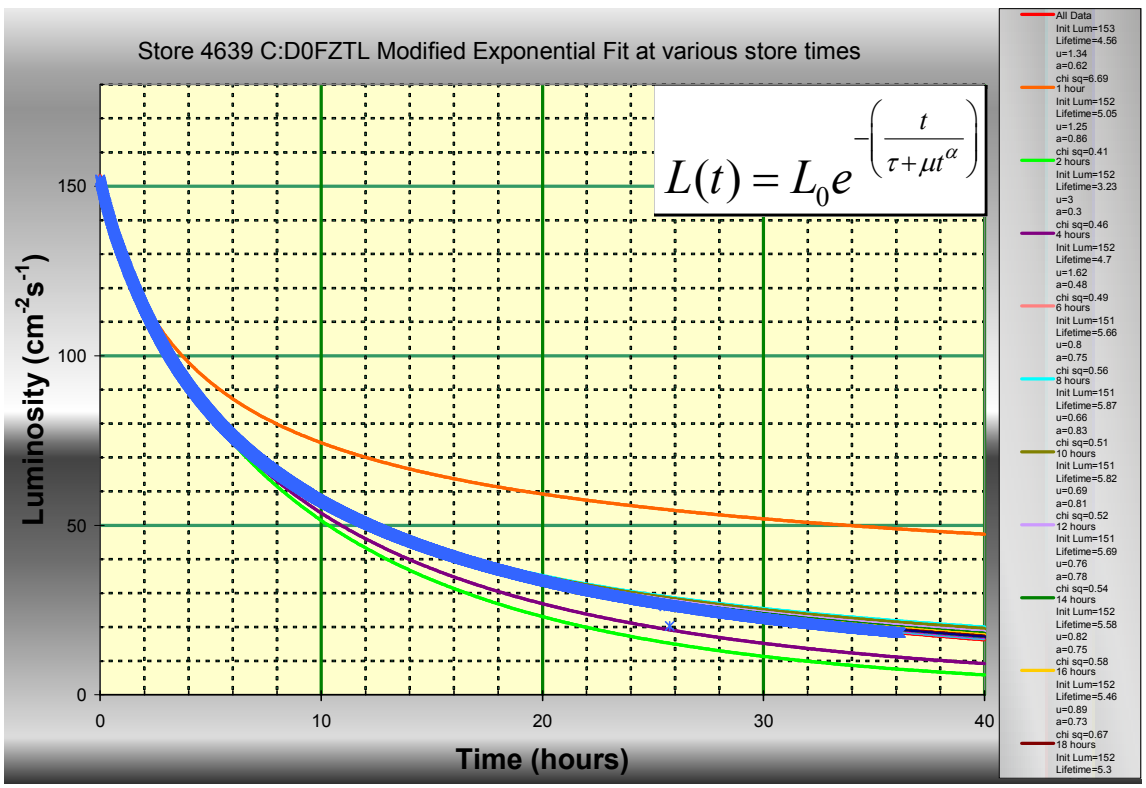
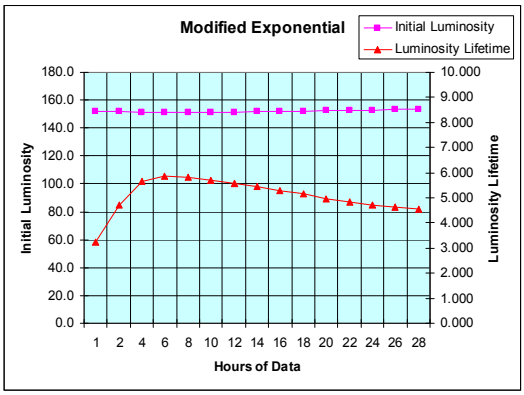


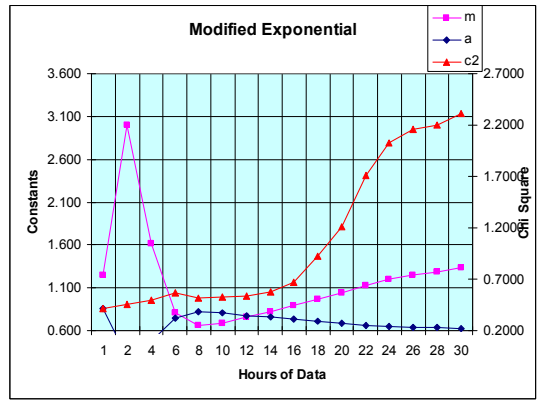
Figure 4-7: Modified Exponential fit of D0 Luminosity data for store 4639 taken with varying samples of lumberjack data.

Again, the thicker blue line is the lumberjack data and the thinner lines are fits made with varying amounts of lumberjack data. We can see that after about 8

hours of lumberjack data, our predicted curves start to match the actual data fairly well.



D0 Initial Luminosity and Lifetime



Constants and Chi Square

Figure 4-8: These plots show how the Modified Exponential fit (Equation 2) predictions change over the number of hours of lumberjack luminosity data used to make the prediction. We are looking at the D0 data for Store 4639. The x-axis in both plots is the number of hours of luminosity data from the lumberjack (1= first hour of data, 2= first two hours of data, etc…). The plot on the left shows the Initial Luminosity and Luminosity Lifetime for each sample of luminosity data. The plot on the right shows the chi square value for those fits.

Figure 4.8 shows the Modified Exponential fit model constants calculated with varying amounts of lumberjack data. Between 8 and 14 hours of lumberjack data the model appears to be starting to give repeatable results. The values of our constants not changing much from fit to fit at this point, with the luminosity lifetime ~5.6 hours, α ~0.8, and μ ~ .55. χ^2 values for those fits are under well 1.0 indicating that this model is doing a good job of fitting the data. Between the 14 and 16 hour point, we see an interesting feature in the fits. The constants start to change and our χ^2 values slowly get worse from fit to fit. Our χ^2 values peaks at just over 2.0. It is possible that we have a few bad data points later on in the store, throwing off the numbers a little. Or maybe, this luminosity model is not as accurate in later portions of the store. These are questions that would be of interest to address. Later we will check the model behavior across different stores to see if this is representative behavior for this model.

c. *Inverse Time to Power Fit (Equation 3)*

The inverse time to power fit proves to be a good predictor of the luminosity behavior. Figure 4.9 shows the CDF lumberjack data for the entire store (thick blue line) and each of the predicted luminosity curves.

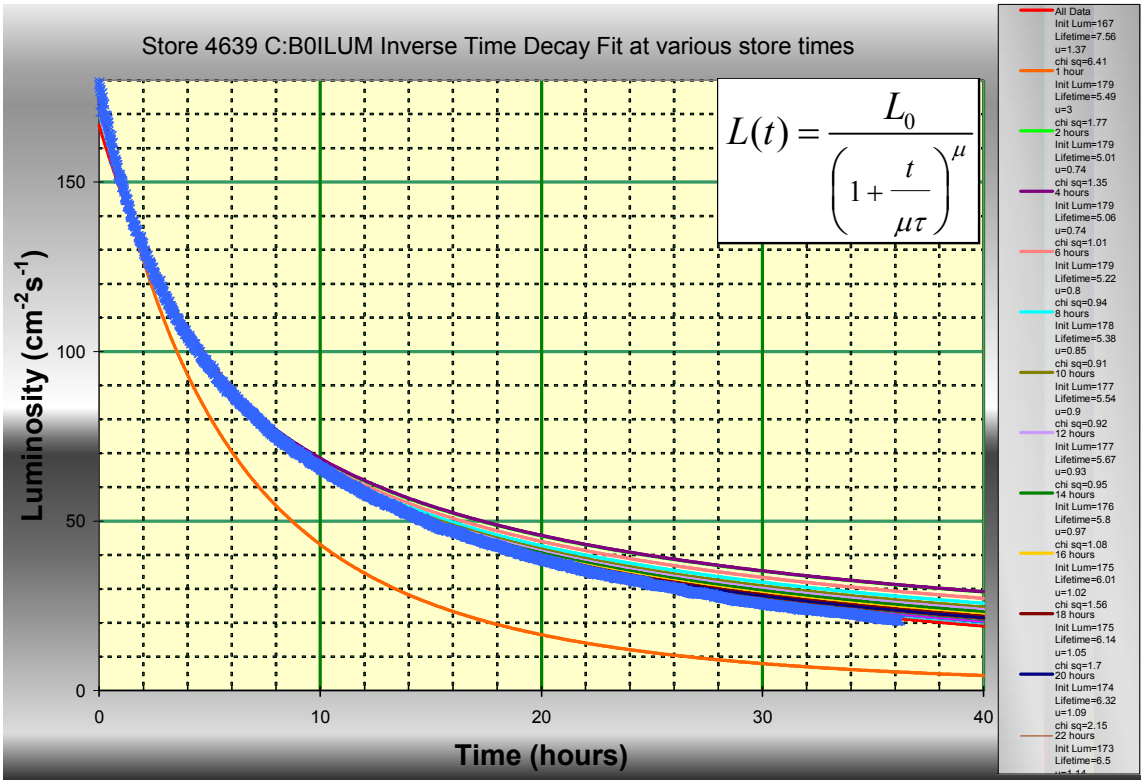
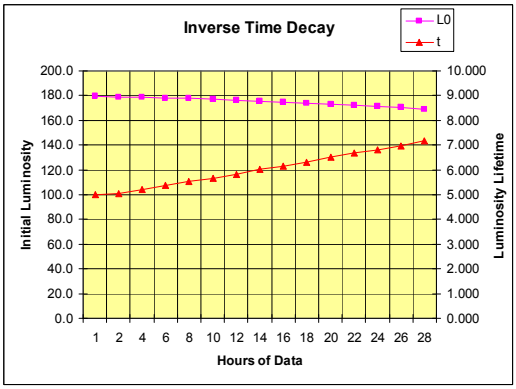
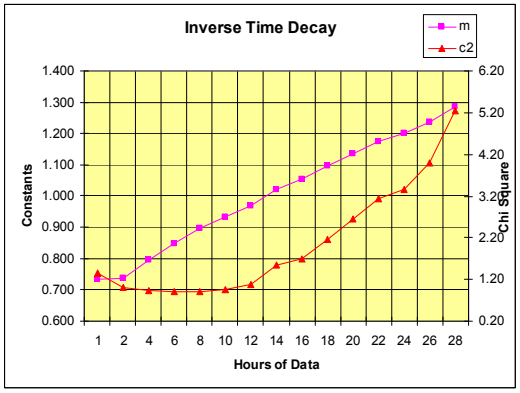


Figure 4-9: Inverse Time to Power Fit of CDF Luminosity data for store 4639 taken with varying samples of lumberjack data.

Again, the thicker blue line is the lumberjack data and the thinner lines are fits made using the model with varying amounts of lumberjack data. The orange trace is after only one hour of lumberjack data. We can see that this trace does not represent the lumberjack data very well. After about four to six hours of lumberjack data, we begin to see a pattern. The predicted curve always over predicts the luminosity, but with each subsequent fit, the fit gets closer to predicting the end of store values. Examination of multiple stores, shows that this behavior is very repeatable.



CDF Initial Luminosity and Lifetime



Constants and Chi Square

Figure 4-10: These plots show how the Inverse Time fit (Equation 3) predictions change over the number of hours of lumberjack luminosity data used to make the prediction. We are looking at the CDF data for

Store 4639. The x-axis in both plots is the number of hours of luminosity data from the lumberjack (1= first hour of data, 2= first two hours of data, etc…). The plot on the left shows the Initial Luminosity and Luminosity Lifetime for each sample of luminosity data. The plot on the right shows the μ constant and chi square value for those fits.

Figure 4-10 shows that the constants for the inverse time fit never stabilize at a given value. The luminosity lifetime increases and the constant μ increases with each subsequent fit. Up to the 12 hour mark in the store the χ^2 values for those remain well under 1.0, but increasing get worse as we get more and more lumberjack data. The final χ^2 value is over 5.0. We will later see that the gradual increase in χ^2 values in this case is due to the fit not being able to fit the beginning store data. As we get more and more lumberjack data allowing us to get closer predictions to the end of store luminosity, we can no longer match the data at the very beginning of the store.

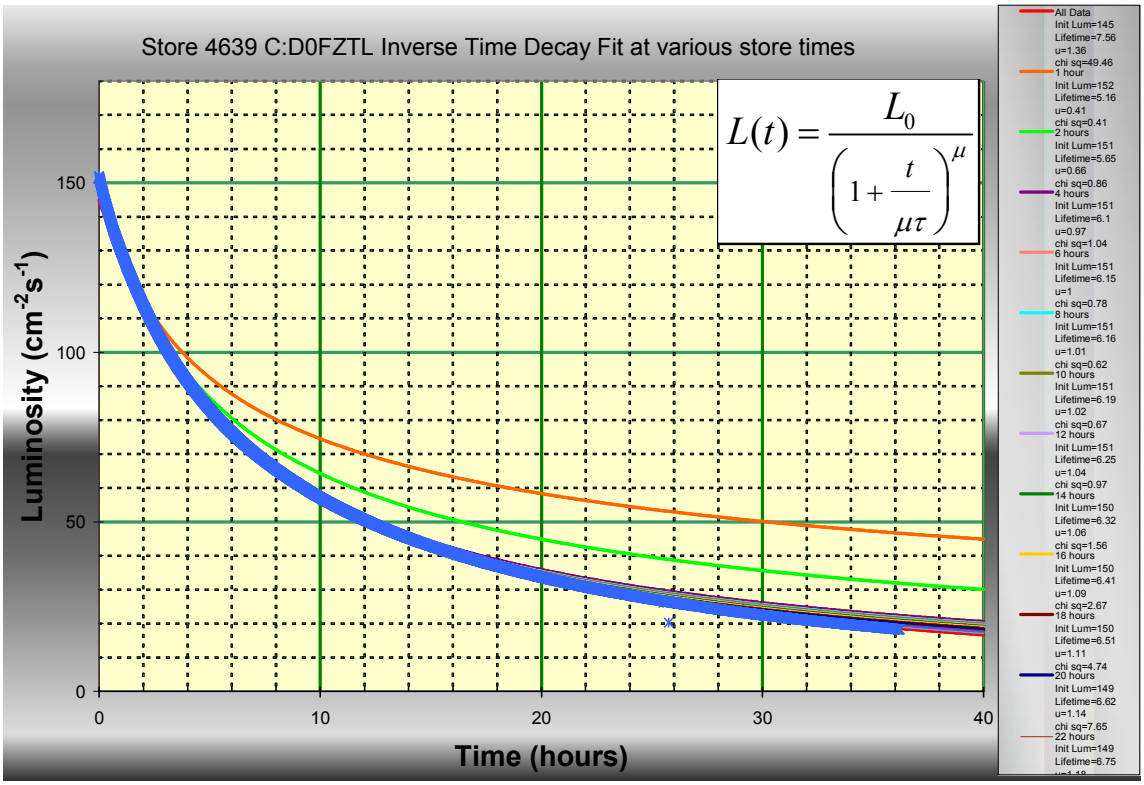
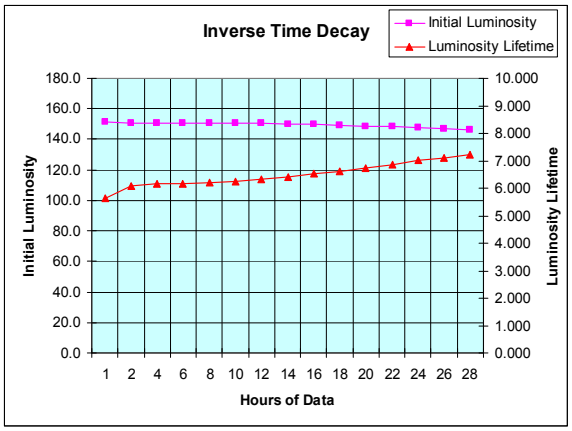
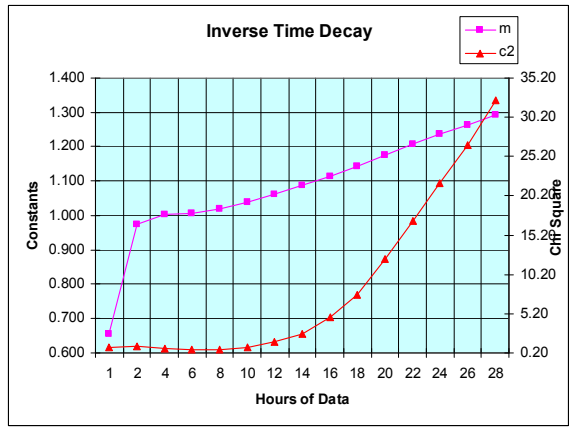


Figure 4-11: Inverse Time to Power Fit of D0 Luminosity data for store 4639 taken with varying samples of lumberjack data.

The D0 data is consistent with what we already saw in the CDF case. The first few hours of store data yields predictions that are not representative of the luminosity data. Once we get four to six hours of lumberjack data, our pattern emerges. We over predict the luminosity, but get better and better predictions of the end of store luminosity with each subsequent fit.



D0 Initial Luminosity and Lifetime



Constants and Chi Square

Figure 4-12: These plots show how the Inverse Time fit (Equation 3) predictions change over the number of hours of lumberjack luminosity data used to make the prediction. We are looking at the D0 data for Store 4639. The x-axis in both plots is the number of hours of luminosity data from the lumberjack (1= first hour of data, 2= first two hours of data, etc…). The plot on the left shows the Initial Luminosity and Luminosity Lifetime for each sample of luminosity data. The plot on the right shows the μ constant and chi square value for those fits.

Figure 4-12 shows that the constants for the inverse time fit never stabilize at a given value. The luminosity lifetime increases and the constant μ increases with each subsequent fit. Up to the 12 hour mark in the store the χ^2 values for those remain well under 1.0, but increasing get worse as we get more and more lumberjack data. The final χ^2 value is over 30.0. We will later see that the gradual increase in χ^2 values in this case is due to the fit not being able to fit the beginning store data. As we get more and more lumberjack data allowing us to get closer predictions to the end of store luminosity, we can no longer match the data at the very beginning of the store.

d. Modified Inverse Time to Power Fit (Equation 3)

The modified inverse time to power fit proves to be a very good predictor of the luminosity behavior. Figure 4.13 shows the CDF lumberjack data for the entire store (thick blue line) and each of the predicted luminosity curves.

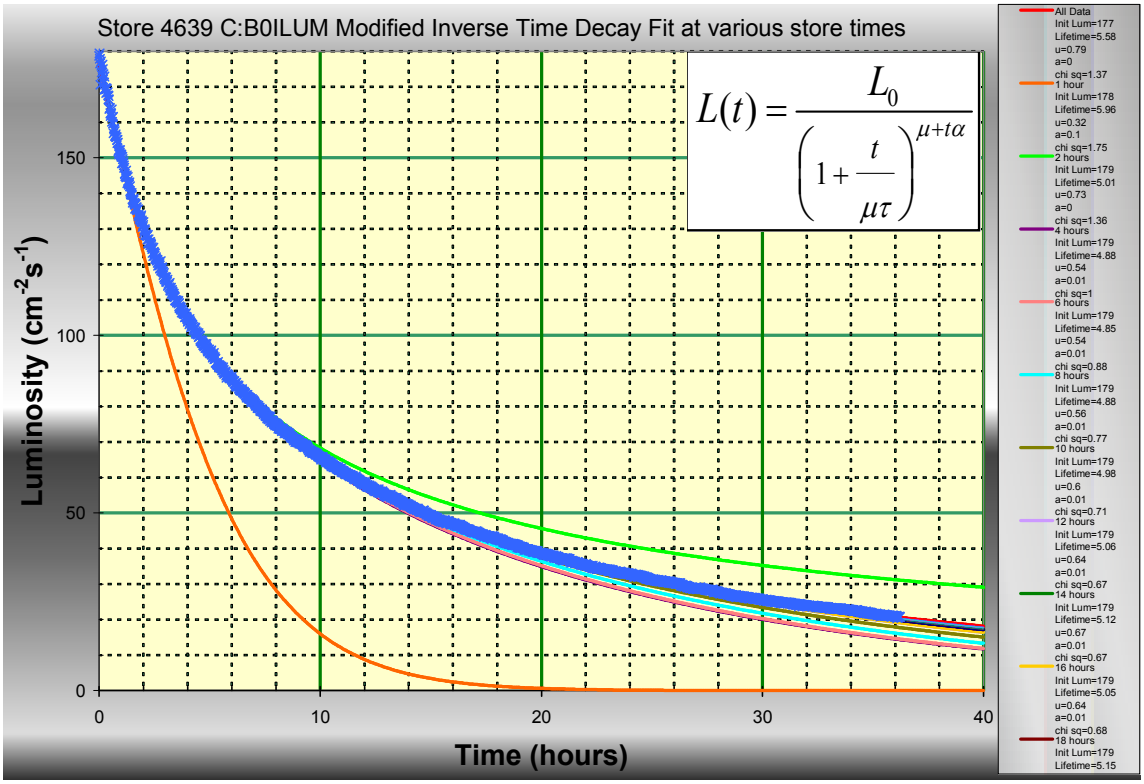
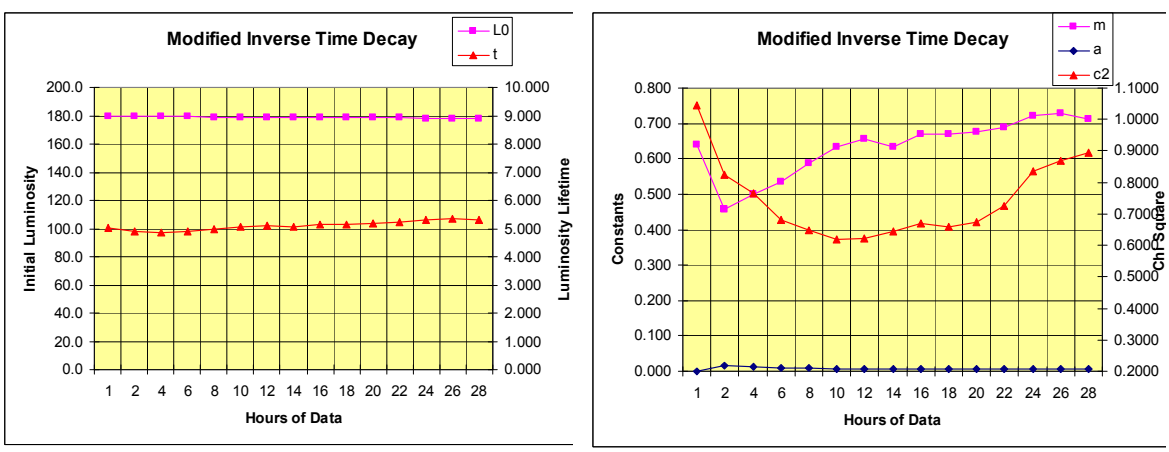


Figure 4-13: Modified Inverse Time to Power Fit of CDF Luminosity data for store 4639 taken with varying samples of lumberjack data.

In Figure 4-13 the thicker blue line is the lumberjack data and the thinner lines are fits made using the model with varying amounts of lumberjack data. The orange trace is after only one hour of lumberjack data. We can see that this trace does not represent the lumberjack data very well. After about four to six hours of lumberjack data, we begin to see a pattern. The predicted curve always under predicts the luminosity, but with each subsequent fit, the fit gets closer to predicting the end of store values. Examination of multiple stores, shows that this behavior is very repeatable.



CDF Initial Luminosity and Lifetime

Constants and Chi Square

Figure 4-14: These plots show how the Modified Inverse Time fit (Equation 4) predictions change over the number of hours of lumberjack luminosity data used to make the prediction. We are looking at the

CDF data for Store 4639. The x-axis in both plots is the number of hours of luminosity data from the lumberjack (1= first hour of data, 2= first two hours of data, etc…). The plot on the left shows the Initial Luminosity and Luminosity Lifetime for each sample of luminosity data. The plot on the right shows the μ constant and chi square value for those fits.

Figure 4-14 shows that between 8 and 20 hours of lumberjack data the model appears to be starting to give repeatable results. The values of our constants have a slight upward trend, but do not changing much from fit to fit at this point, with the luminosity lifetime ~5 hours, α ~0.07, and μ ~ .6. χ^2 values for those fits are under well 1.0 indicating that this model is doing a good job of fitting the data. Starting at around the 22 hour mark, the χ^2 values start to rise, but remain under 1.0, so the fit is still good. It is possible that we have a few bad data points later on in the store, throwing off the numbers a little. Or maybe, this luminosity model is not as accurate in later portions of the store. These are questions that would be of interest to address. Later we will check the model behavior across different stores to see if this is representative behavior for this model.

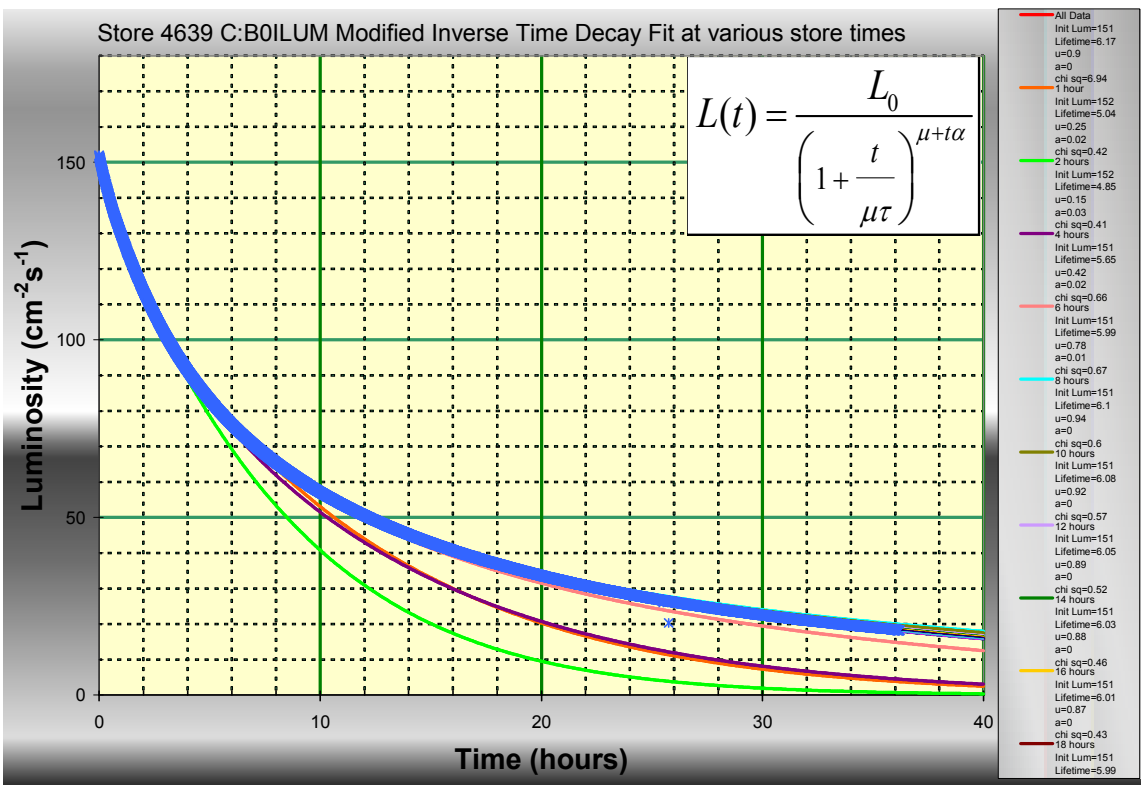
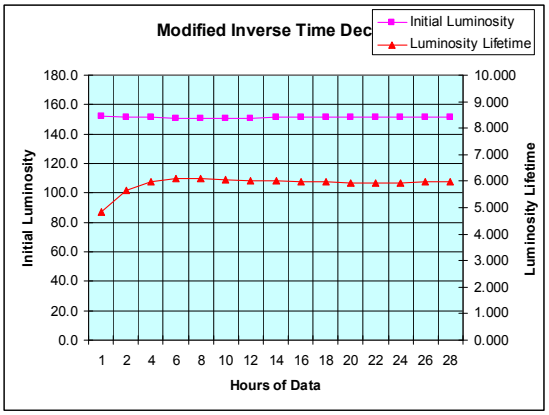
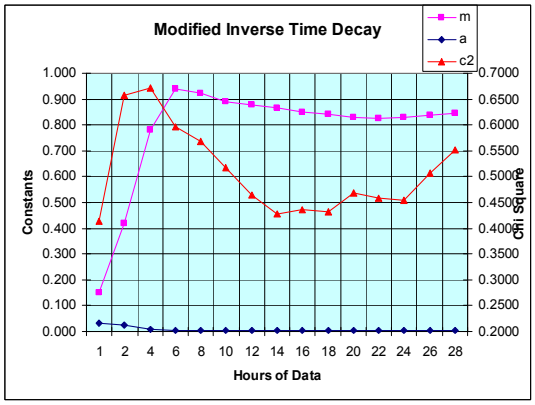


Figure 4-15: Modified Inverse Time to Power Fit of D0 Luminosity data for store 4639 taken with varying samples of lumberjack data.

Figure 4-15 shows that we have similar results for D0 luminosity predictions. In this case, it takes a full 8 hours before the prediction starts matching the store data.



D0 Initial Luminosity and Lifetime



Constants and Chi Square

Figure 4-16: These plots show how the Modified Inverse Time fit (Equation 4) predictions change over the number of hours of lumberjack luminosity data used to make the prediction. We are looking at the D0 data for Store 4639. The x-axis in both plots is the number of hours of luminosity data from the lumberjack (1= first hour of data, 2= first two hours of data, etc…). The plot on the left shows the Initial Luminosity and Luminosity Lifetime for each sample of luminosity data. The plot on the right shows the μ constant and chi square value for those fits.

Figure 4-16 shows that after eight hours until the end of store, the model constants for the D0 luminosity start giving repeatable results. The values of our do not changing much from fit to fit at this point, with the luminosity lifetime ~6 hours, α ~0.04, and μ ~ .85. χ^2 values for those fits are under well 1.0 indicating that this model is doing a good job of fitting the data.

We have used our Luminosity Predictor tool to predict luminosity profiles from varying amounts of lumberjack data for Store 4639. We found that the simple exponential fit did very poorly and cannot be used in any way to predict luminosity. We found that the inverse time fit, did ok but always over-predicted the luminosity. We found that the modified exponential and modified inverse time fits did best.

5. Beginning of Store

We have seen that the calculated constants from our Tevatron luminosity models change as we include more and more lumberjack data. We also saw hints that all but one of our models match up well with most of the data, but some do better matching the data at both the beginning and end of the store. We will again look at the data from Store 4639, but will zoom in to take a closer look at the match between the fits and our lumberjack data during the first two hours of the store.

a. Simple Exponential Fit (Equation 1)

The simple exponential fit proves to be a very poor predictor of the luminosity behavior. Figure 5-1 shows the CDF lumberjack data for the entire store and each of the predicted luminosity curves.

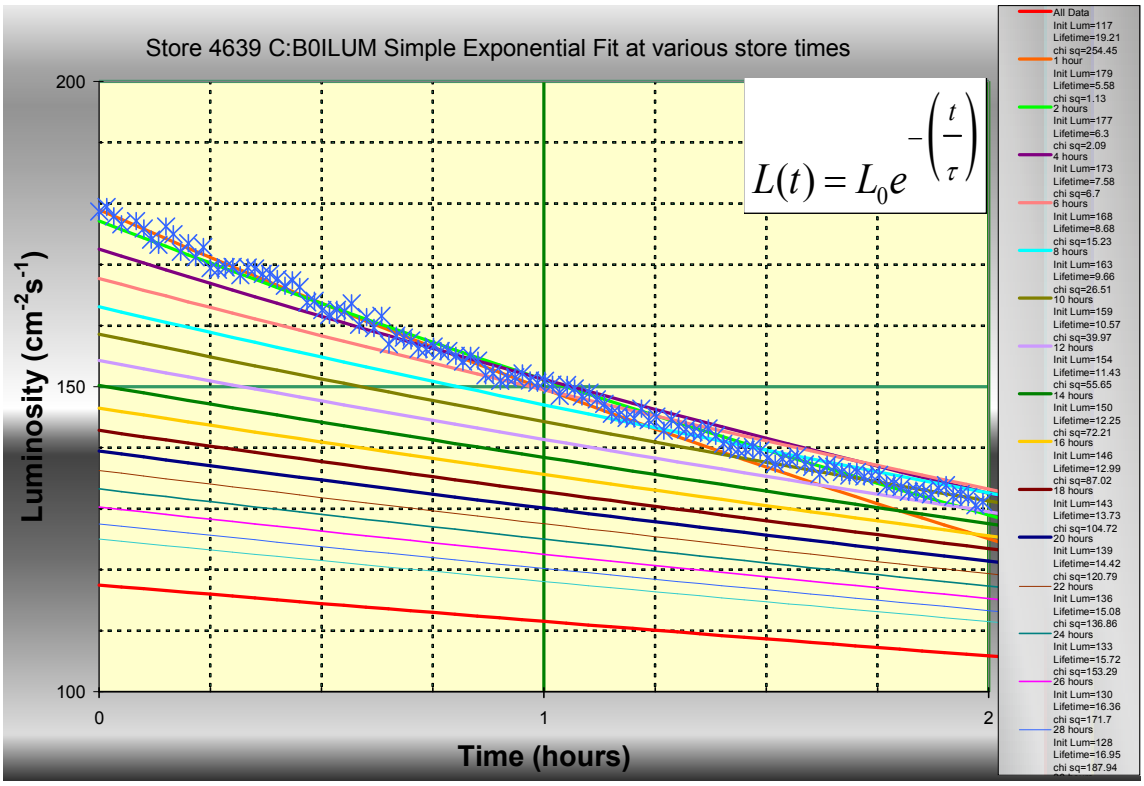


Figure 5-1: Here we compare the lumberjack luminosity data with the Simple Exponential model of Equation (1) over the first two hours of Store 4639. The blue "**x**"s are the lumberjack data, and the lines are the predicted curves given different quantities of lumberjack data.

The Simple Exponential model of Equation 1 makes a poor model of the luminosity behavior of the store. The blue "**x**"s represent the lumberjack data for the store. The orange line is taken after only one hour of lumberjack data. This curve can be made to match the first hour of the store, but as we will see later, does a very poor job of predicting the luminosity at the end of the store. Each successive curve (two hours of lumberjack data, four hours of lumberjack data, etc…) match up less and less with the beginning of the store data. Notice that the end of store luminosity fit has the initial luminosity around 120 x 10^{30} cm^{-1} s^{-1}, when the actual initial luminosity was closer to 180 x 10^{30} cm^{-1} s^{-1}. This is a 33% difference, showing that this fit is not a good representation of lumberjack data.

b. Modified Exponential Fit (Equation 2)

The modified exponential fit proves to be a very good predictor of the luminosity behavior. Figure 5.2 shows the CDF lumberjack data for the entire store and each of the predicted luminosity curves.

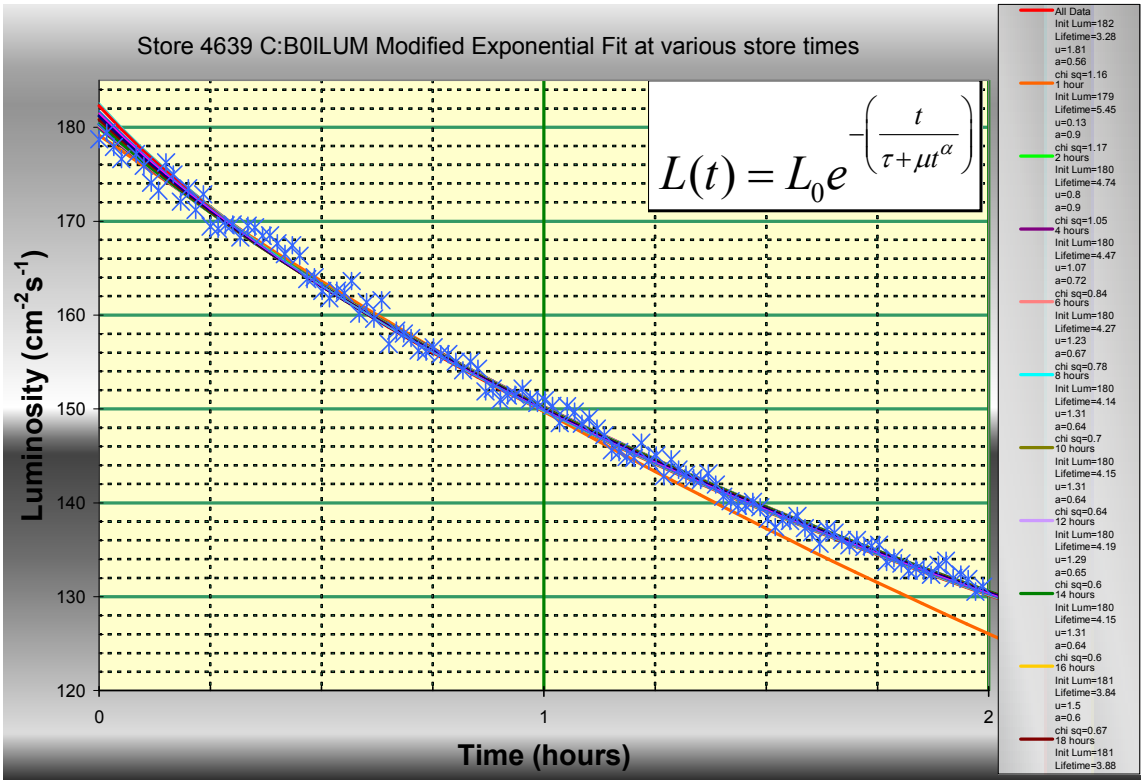


Figure 5-2: Here we compare the lumberjack luminosity data with the Modified Exponential model of Equation (2) over the first two hours of Store 4639. The blue "**x**"s are the lumberjack data, and the lines are the predicted curves given different quantities of lumberjack data.

Figure 5-2 shows that the Modified Exponential model of Equation 2 makes a good model of the luminosity behavior of the store. The blue "**x**"s represent the lumberjack data for the store. The orange line is taken after only one hour of lumberjack data. This is the only curve that does not fit the data well. Each of the other curves (two hours of lumberjack data, four hours of lumberjack data, etc…) appear to match the first two hours of the store very well. There appears to be an interesting area in the first 15 minutes of the data. We will now blow up on that data to take a closer look.

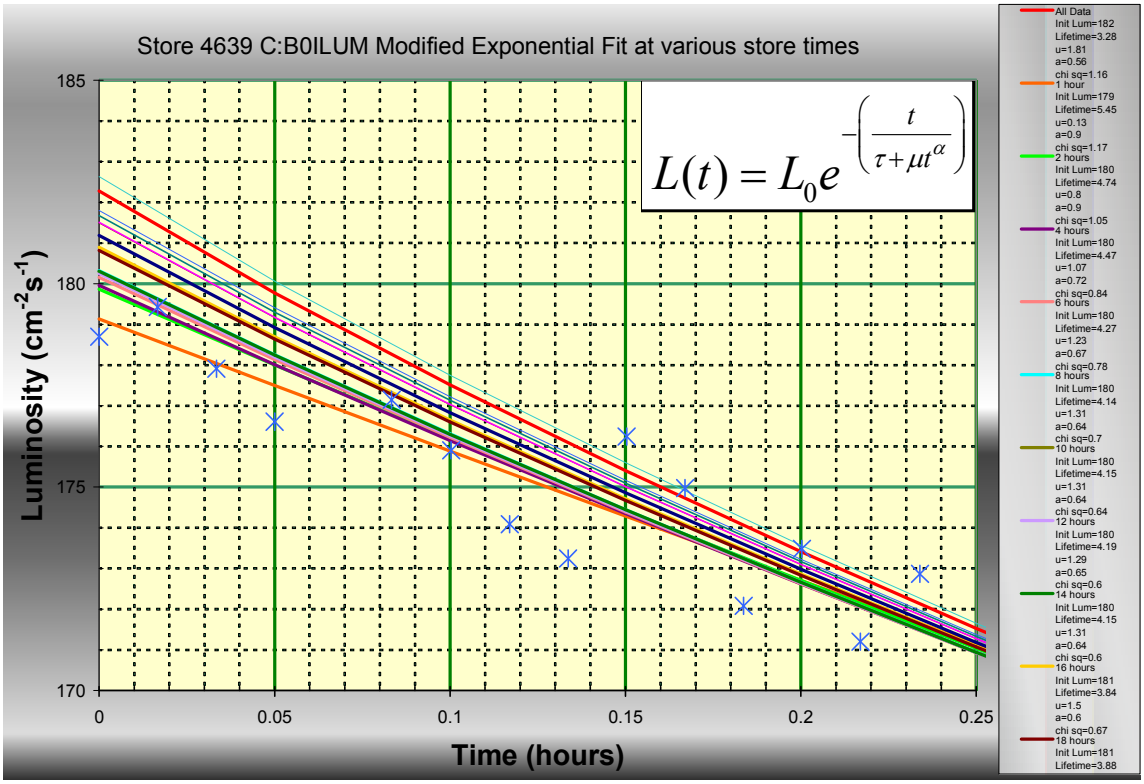


Figure 5-3: Here we compare the lumberjack luminosity data with the Modified Exponential model of Equation (2) over the first fifteen minutes of Store 4639. The blue "**x**"s are the lumberjack data, and the lines are the predicted curves given different quantities of lumberjack data.

Zooming in on the interesting feature in the first fifteen minutes of the store we see that the predicted curves have slightly different initial luminosity numbers. Throwing out the first hour fit which does not match the data very well, our range of initial luminosities on all of our fits from two hours of luminosity data to 30 hours of luminosity data is 180 x 10^{30} cm^{-1} s^{-1} to 182.5 x 10^{30} cm^{-1} s^{-1}. This is approximately a 1.4% difference. Compared that to our assumed measurement error of +/- 0.6%, showing that most of this difference is in the assumed error of the measurement.

c. *Inverse Time to Power Fit (Equation 3)*

The inverse time to power fit proves to be a good predictor of the luminosity behavior. Figure 5.3 shows the CDF lumberjack data for the entire and each of the predicted luminosity curves.

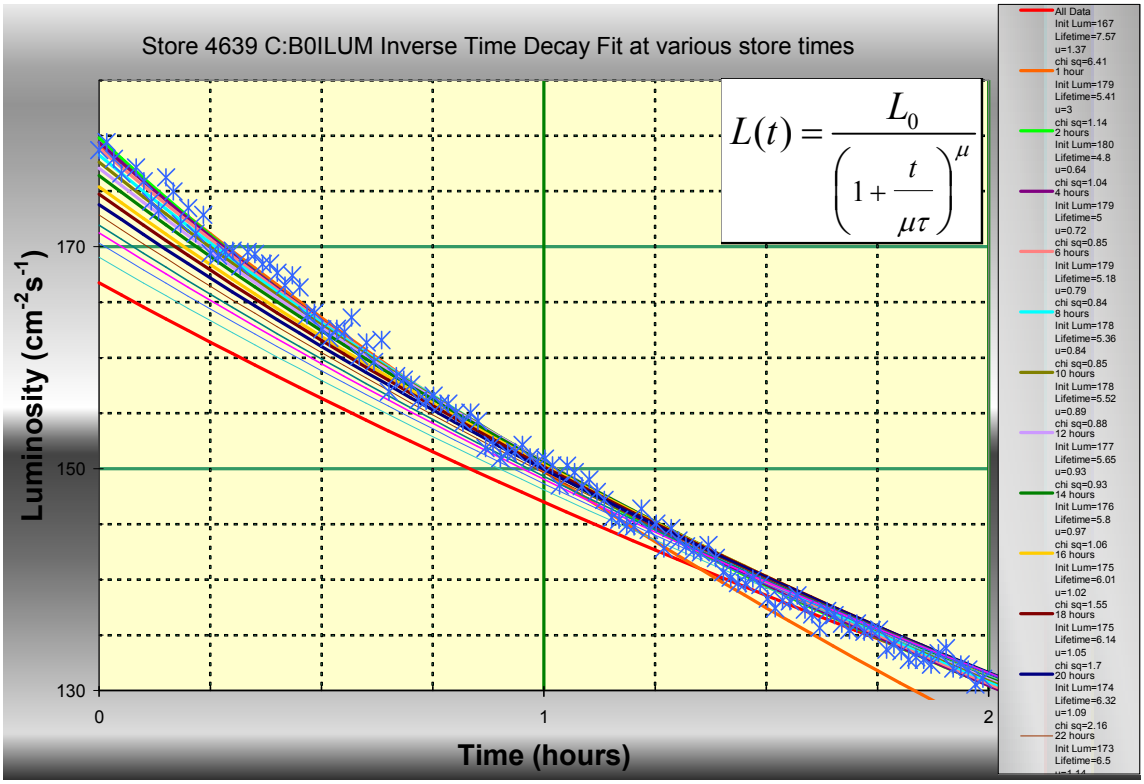


Figure 5-4: Here we compare the lumberjack luminosity data with the Inverse Time model of Equation (3) over the first two hours of Store 4639. The blue "**x**"s are the lumberjack data, and the lines are the predicted curves given different quantities of lumberjack data.

Figure 5-4 shows that the Inverse Time model of Equation 3 has trouble matching the first two hours of store data. The blue "**x**"s represent the lumberjack data for the store. The orange line is taken after only one hour of lumberjack data. The other curves show an interesting pattern. The green curve is taken with two hours of lumberjack data and appears to closely match the initial luminosity. Each successive curve (two hours of lumberjack data, four hours of lumberjack data, etc..) has the initial luminosity lower and lower. The range is 179 x 10^{30} cm^{-1} s^{-1} for the two hours of data fit to 167 x 10^{30} cm^{-1} s^{-1} for the 36 hours of lumberjack data. This is a 6.7% range.

d. Modified Inverse Time to Power Fit (Equation 4)

The modified inverse time to power fit proves to be a good predictor of the luminosity behavior. Figure 5-5shows the CDF lumberjack data for the entire store and each of the predicted luminosity curves.

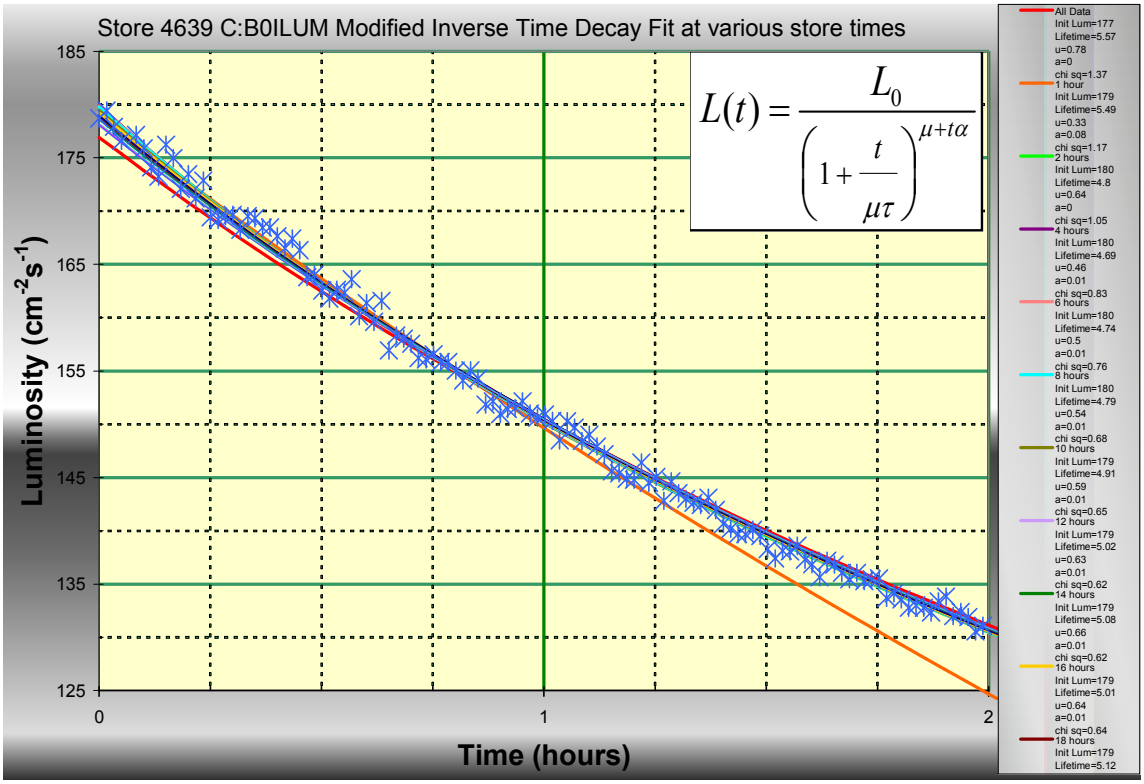


Figure 5-5: Here we compare the lumberjack luminosity data with the Modified Inverse Time model of Equation (4) over the first two hours of Store 4639. The blue "**x**"s are the lumberjack data, and the lines are the predicted curves given different quantities of lumberjack data.

Figure 5-5 shows that the Modified Inverse Time model of Equation 4 provides a good representation of luminosity behavior at the beginning of the store. The blue "**x**"s represent the lumberjack data for the store. The orange line is taken after only one hour of lumberjack data. The other curves appear to fit the data well. The green curve is taken with two hours of lumberjack data and appears to closely match the initial luminosity. The range of predicted initial luminosities is 177 x 10^{30} cm^{-1} s^{-1} to 180 x 10^{30} cm^{-1} s^{-1}. This is a 1.6% range, which can be compared to the +/- 0.6% assumed error in the measurement.

In this section we looked at each of our four luminosity model predictions for varying amounts of lumberjack data. We see that the Simple Exponential fit does not match the data at all. The Inverse Time fit matches the data to the 6% level. The Modified Exponential and Modified Inverse Time fits were both equally impressive with a range of only about 1.5%. Of course, getting the data to fit the start of store luminosity is easy. To put these fits to the test, we need to see how well they predict the end of store luminosities.

6. End of Store

The bottom line is we want to see how well our Luminosity models predict the end of store luminosity numbers given varying amounts of initial lumberjack data. This section examines the predictions for the last two hours of the store for each Tevatron model. Again, we use Store 4639.

a. Simple Exponential Fit (Equation 1)

The simple exponential fit proves to be a very poor predictor of the luminosity behavior. Figure 6-1 shows the CDF lumberjack data for the entire store and each of the predicted luminosity curves.

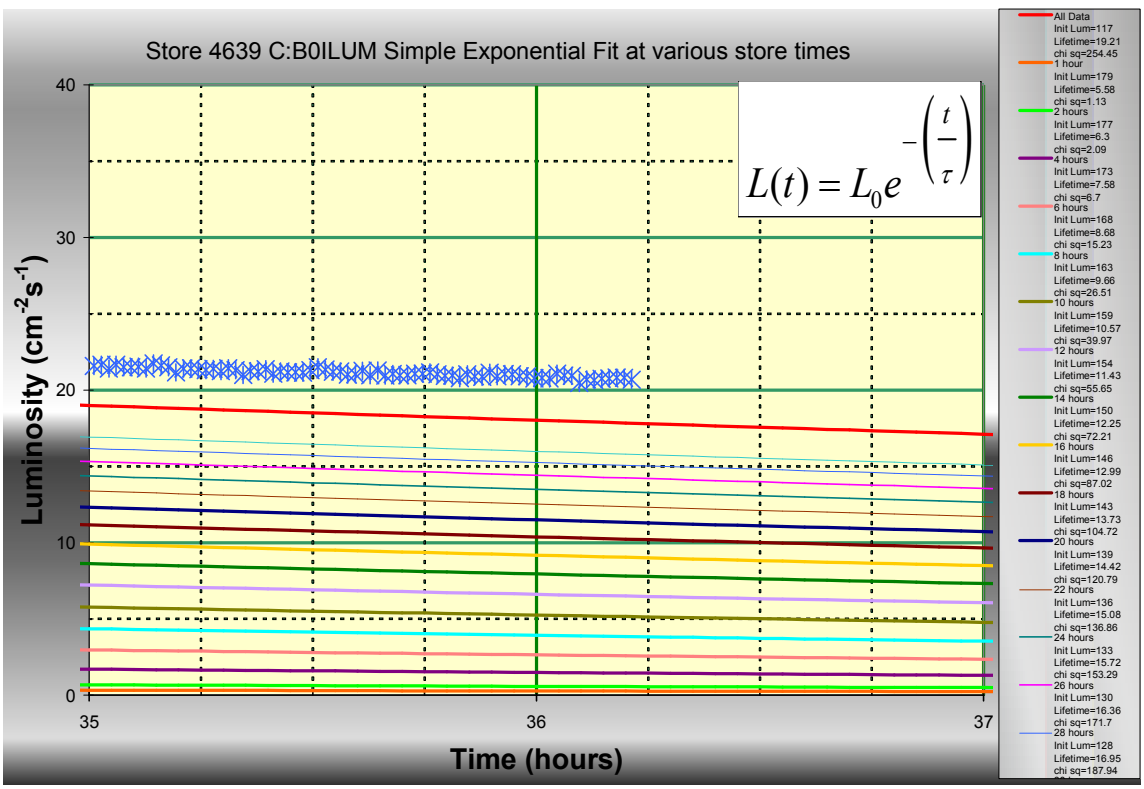


Figure 6-1: Here we compare the lumberjack luminosity data with the Simple Exponential model of Equation (1) over the last hour of Store 4639. The blue "**x**"s are the lumberjack data, and the lines are the predicted curves given different quantities of lumberjack data.

Figure 6-1 clearly shows that the Simple Exponential model is a poor predictor of end of store luminosity. The red trace, which uses all of the lumberjack data for the entire store, shows the luminosity 7% lower than measured at the end of store. Recall this trace was 33% lower than measured at injection. The fits at the end of store time get worse and worse with less and less lumberjack data. As an example, after the first eight hours of lumberjack data, the curve predicts an end of store luminosity of 3.9 x 10^{30} cm^{-1} s^{-1}. Compare that to the measured end of

store luminosity 20.7 x 10^{30} cm^{-1} s^{-1}, giving an error in the prediction of 81%. Overall, the Simple Exponential model is not very useful in predicting the end of store luminosity.

b. Modified Exponential Fit (Equation 2)

The modified exponential fit proves to be a very good predictor of the luminosity behavior. Figure 6-2 shows the CDF lumberjack data for the entire and each of the predicted luminosity curves.

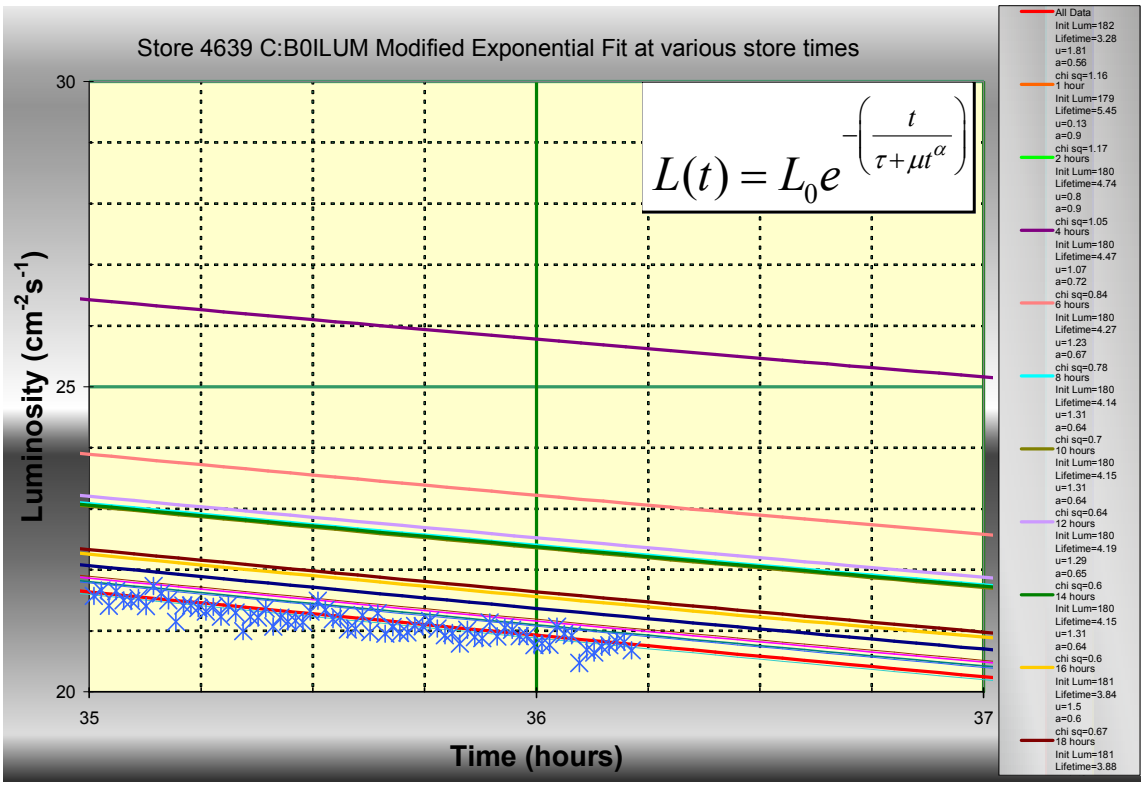


Figure 6-2: Here we compare the lumberjack luminosity data with the Modified Exponential model of Equation (2) over the last hour of Store 4639. The blue "**x**"s are the lumberjack data, and the lines are the predicted curves given different quantities of lumberjack data.

Figure 6-2 shows the end of store luminosity predictions given varying amounts of lumberjack data. The fits with one and two hours of lumberjack data are poor enough that they are not on-scale. The red trace, which uses all 36 hours of lumberjack data for the entire store, shows a luminosity curve that matches very well at the end of store time. The general trend is the model predicts the luminosity too high, but gets better and better with more and more lumberjack data. Let's look a little closer.

Modified Exponential Model End of Store Predictions		
Lumberjack Data (Hours)	Predicted - Actual (e30)	Error (%)
4	4.97	24.03%
6	2.42	11.70%
8	1.57	7.59%
10	1.57	7.59%
12	1.68	8.12%
14	1.56	7.54%
16	0.71	3.43%
18	0.79	3.82%
20	0.52	2.51%
22	0.32	1.55%
24	0.32	1.55%
26	0.22	1.06%
28	0.02	0.10%
30	0.02	0.10%

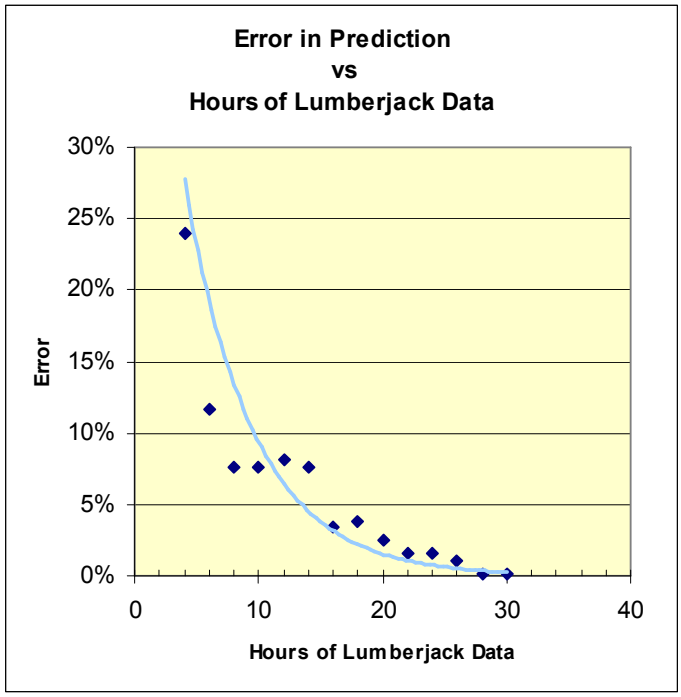


Figure 6-3: This table compares the lumberjack end of store luminosity with the Modified Exponential model prediction given varying amounts of lumberjack data.

Figure 6-3 looks at the data in Figure 6-2 and shows the difference between the predicted and measured luminosities given different amounts of lumberjack data. After the first six hours of lumberjack data, our prediction is within ~2.4 x 10^{30} cm^{-1} s^{-1}, or ~11.7%. After eight to fourteen hours of lumberjack data, we are within ~1.5 x 10^{30} cm^{-1} s^{-1}, or ~7.5%. Once we have more than sixteen hours of lumberjack data our predictions are less than 1.0 x 10^{30} cm^{-1} s^{-1} different. Overall, it appears that after about eight hours of lumberjack data we can start using this model to make a rough luminosity prediction out as far as 36 hours. For our purposes, this is reasonable. Fits that deviate from the data by less than 1% would require at least 28 hours of store data.

c. *Inverse Time to Power Fit (Equation 3)*

The inverse time to power fit proves to be a good predictor of the luminosity behavior. Figure 6-4 shows the CDF lumberjack data for the entire store and each of the predicted luminosity curves.

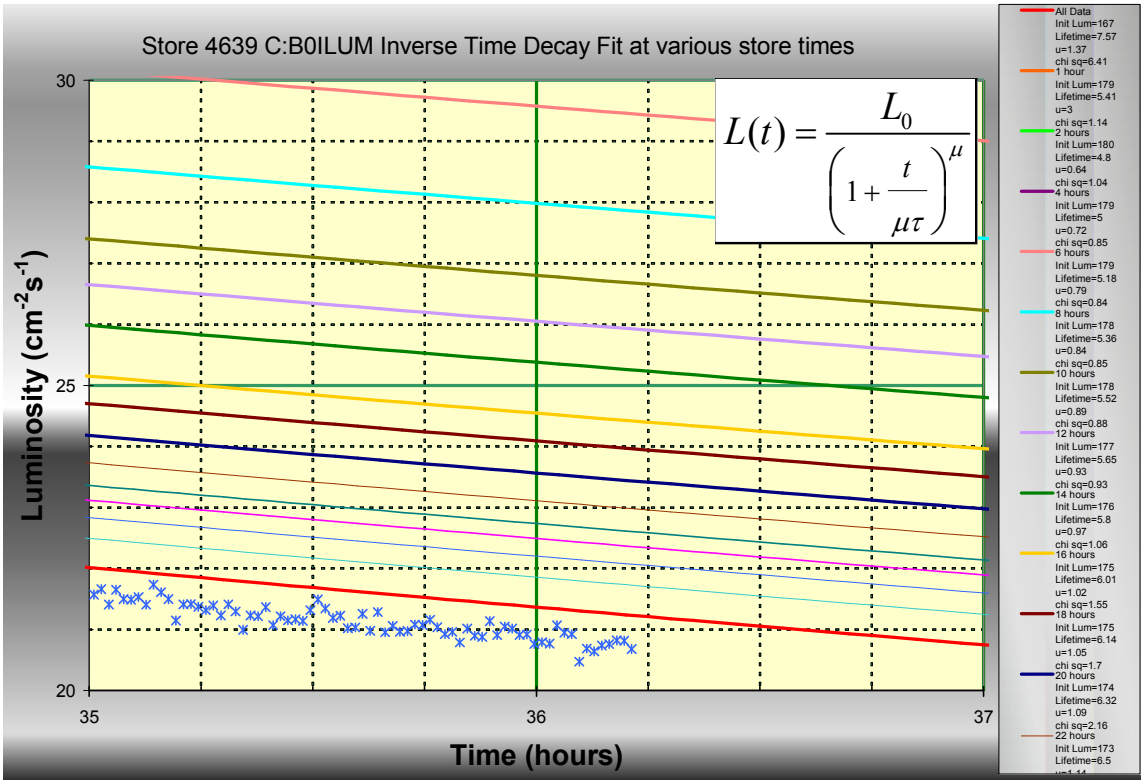


Figure 6-4: Here we compare the lumberjack luminosity data with the Inverse Time model of Equation (3) over the last hour of Store 4639. The blue "**x**"s are the lumberjack data, and the lines are the predicted curves given different quantities of lumberjack data.

Figure 6-4 shows the end of store luminosity predictions given varying amounts of lumberjack data. The fits with one, two and four hours of lumberjack data are poor enough that they are not on-scale. This plot shows that the end of store luminosity prediction for this model are not as close as the predictions from the Modified Exponential fit shown in Figure 6-2. In Figure 6-4, the red trace, which uses all 36 hours of lumberjack data for the entire store, shows a luminosity curve that matches fairly well at the end of store time. The general trend is the model predicts the luminosity too high, but gets better and better with more and more lumberjack data. Let's look a little closer.

Inverse Time Model End of Store Predictions		
Lumberjack Data (Hours)	Predicted - Actual (e30)	Error (%)
6	8.78	42.46%
8	7.19	34.77%
10	5.99	28.97%
12	5.25	25.39%
14	4.58	22.15%
16	3.75	18.13%
18	3.29	15.91%
20	2.77	13.39%
22	2.31	11.17%
24	1.94	9.38%
26	1.69	8.17%
28	1.39	6.72%
30	1.05	5.08%

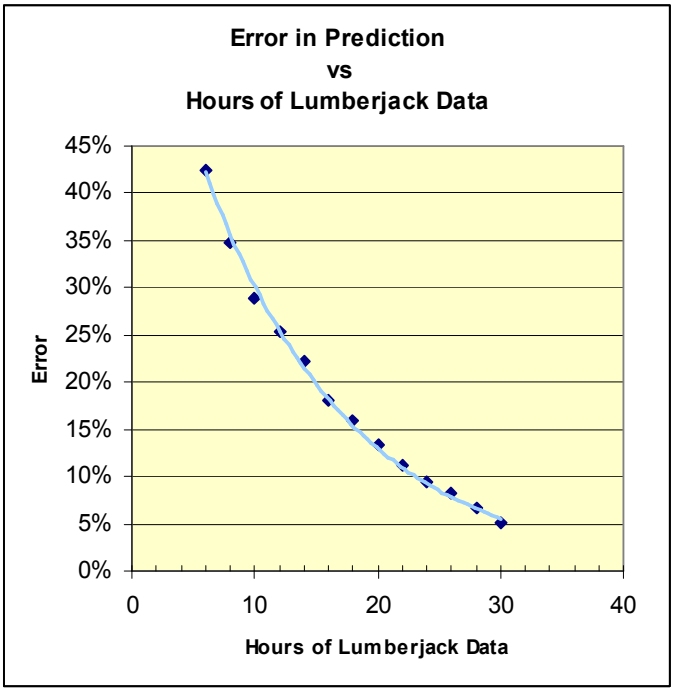


Figure 6-5: This table compares the lumberjack end of store luminosity with the Inverse Time model prediction given varying amounts of lumberjack data.

Figure 6-5 looks at the data in Figure 6-4 and shows the difference between the predicted and measured luminosities given different amounts of lumberjack data. In the early stages, this fit gives a very poor prediction as to the luminosity at the end of the store; however, the predictions do get better with more lumberjack data. It is not until we have over 22 hours of data, are we able to get an end of store luminosity prediction within 10% for a 36 hour store. Even after 30 hours of lumberjack data, we still can only predict the end of store luminosity within 5%. Using the fit, as I have implemented it, is a poor predictor of the end of store luminosity.

Before we completely abandon this model for our luminosity predictions, there are some interesting features. After examining multiple stores, the errors in the predictions follow the same pattern. With little luminosity data, the fit is good at the beginning of the store, but poor at the end of the store. With many hours of luminosity data, the fit gets better at the end of the store, at the expense of getting worse at the beginning of the store. Later in this document, we will examine how the constants in the models change with varying amounts of lumberjack data across multiple stores. It may be of interest to see if we could modify the early store predictions based on how our fit consistently errors in the same direction to get a better prediction. Let's save that thought until later in this document.

d. Modified Inverse Time to Power Fit (Equation 4)

The modified inverse time to power fit proves to be a good predictor of the luminosity behavior. Figure 6-6 shows the CDF lumberjack data for the entire store and each of the predicted luminosity curves.

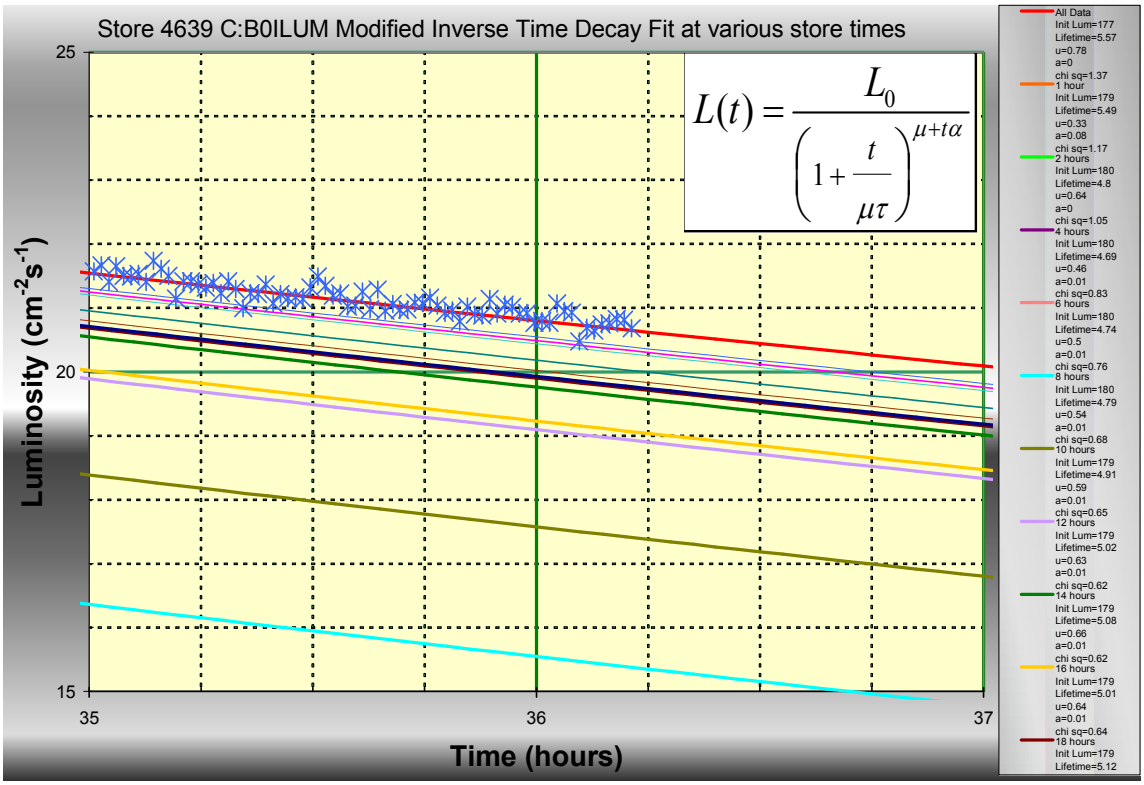


Figure 6-6: Here we compare the lumberjack luminosity data with the Modified Inverse Time model of Equation (4) over the last hour of Store 4639. The blue "**x**"s are the lumberjack data, and the lines are the predicted curves given different quantities of lumberjack data.

Figure 6-6 shows the end of store luminosity predictions given varying amounts of lumberjack data. The fits with the first six hours of lumberjack data are poor enough that they are not on-scale. The red trace, which uses all 36 hours of lumberjack data for the entire store, shows a luminosity curve that matches very well at the end of store time. The general trend is the model predicts the luminosity too low, but gets better and better with more and more lumberjack data. Let's look a little closer.

Modified Inverse Time Model End of Store Predictions		
Lumberjack Data (Hours)	Predicted - Actual (e30)	Error (%)
8	-5.29	25.58%
10	-3.26	15.76%
12	-1.74	8.41%
14	-1.07	5.17%
16	-1.64	7.93%
18	-0.9	4.35%
20	-0.89	4.30%
22	-0.81	3.92%
24	-0.65	3.14%
26	-0.39	1.89%
28	-0.34	1.64%
30	-0.28	1.35%

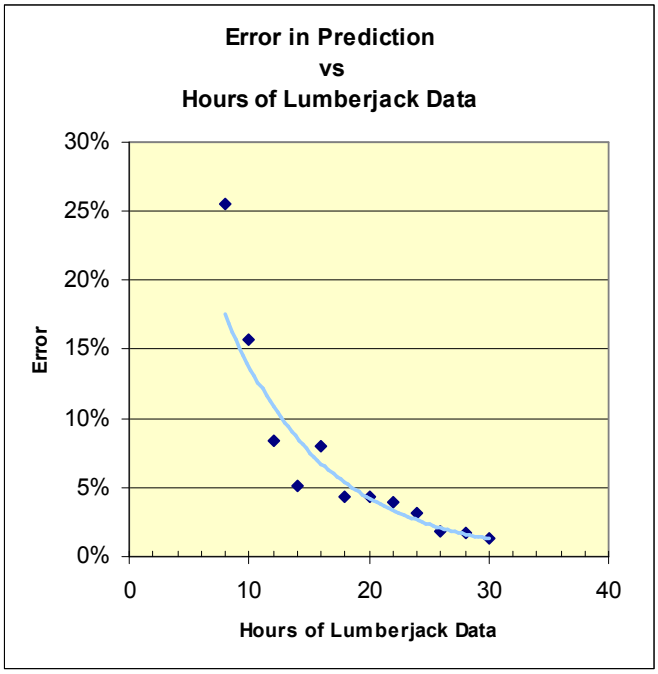


Figure 6-7: This table compares the lumberjack end of store luminosity with the Modified Exponential model prediction given varying amounts of lumberjack data.

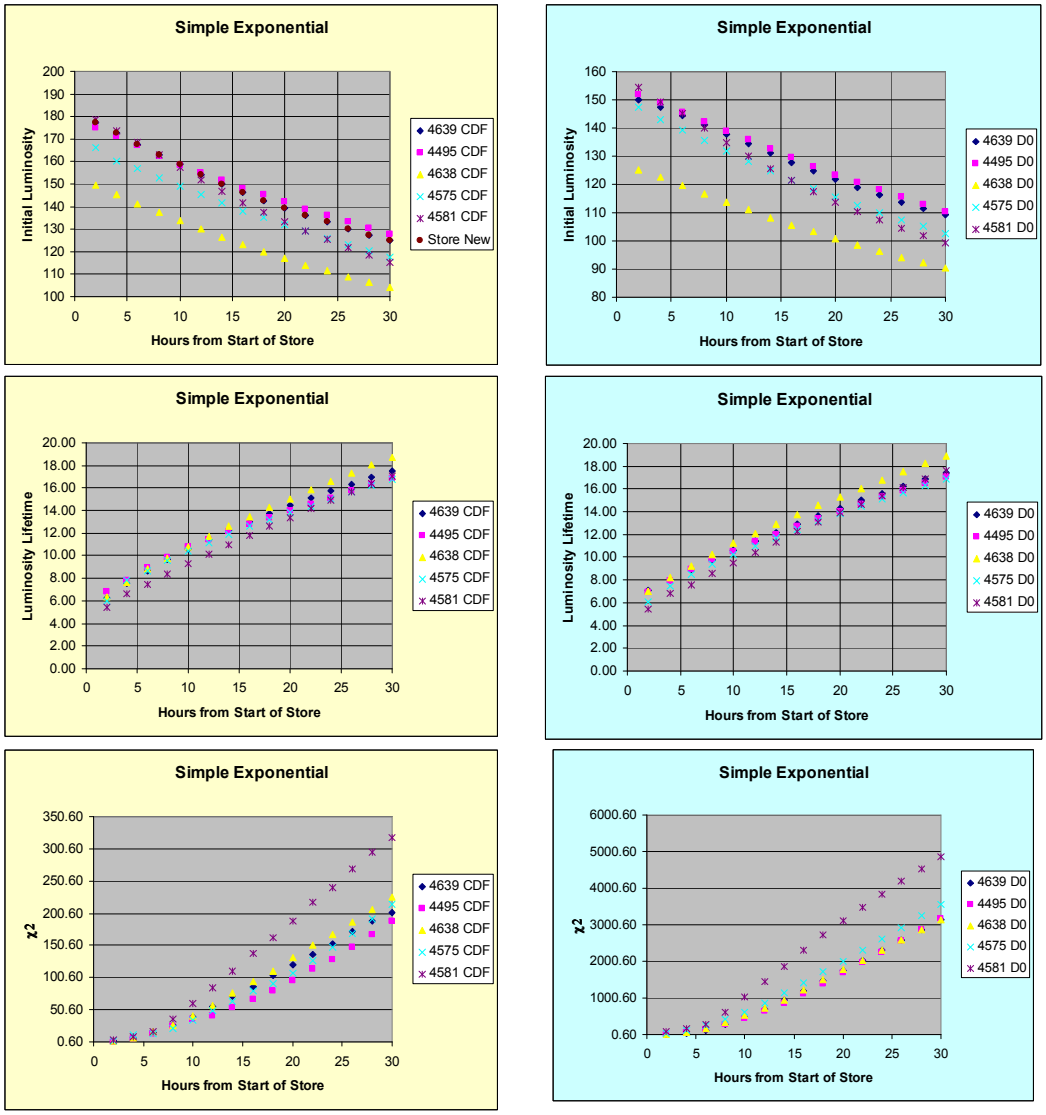
Figure 6-7 looks at the data in Figure 6-6 and shows the difference between the predicted and measured luminosities given different amounts of lumberjack data. After the first six hours of lumberjack data, our prediction is within ~3.25 x 10^{30} cm^{-1} s^{-1}, or ~16%. After twelve to fourteen hours of lumberjack data, we are within ~1.5 x 10^{30} cm^{-1} s^{-1}, or ~8%. Once we have more than eighteen hours of lumberjack data our predictions are less than 1.0 x 10^{30} cm^{-1} s^{-1} different. Overall, it appears that after about ten or twelve hours of lumberjack data we can start using this model to make a rough luminosity prediction out as far as 36 hours. For our purposes, the time to get to a predictable luminosity is a little long, but once there the predictions are good.

7. Comparing Fit Numbers for Top 5 Stores

Now that we have examined the data from Store 4639 in depth, it would be interesting to see how luminosity models change across multiple stores. For this section, plots of each of the model constants were constructed for the top 5 delivered luminosity stores (Store 4639, Store 4495, Store 4638, Store 4575, and Store 4581). In addition the complete set of fits was also run on the next five best delivered luminosity stores (Store 4574, Store 4573, Store 4473, Store 4477, and Store 4560) and can be found in the AD documents database along with this document.

a. Simple Exponential Fit (Equation 1)

The simple exponential fit proves to be a very poor predictor of the luminosity behavior. Figure aaa shows the CDF lumberjack data for the entire store (thick blue line) and each of the predicted luminosity curves.

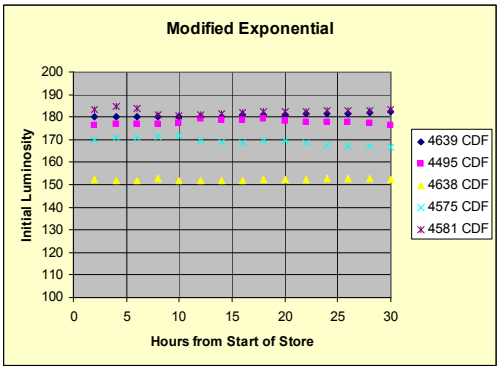


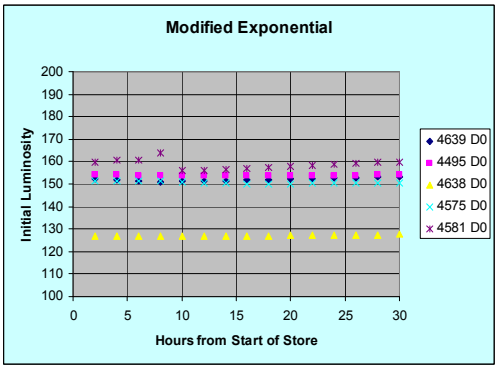
CDF D0

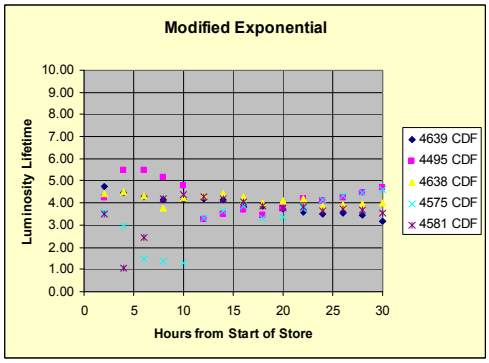
Figure 7-1

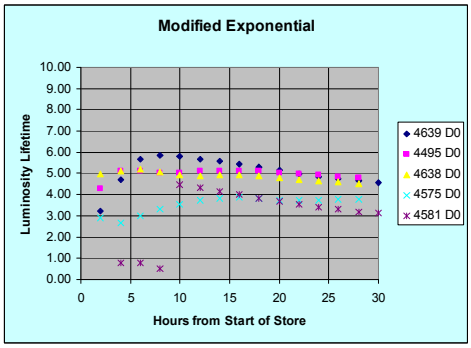
b. Modified Exponential Fit (Equation 2)

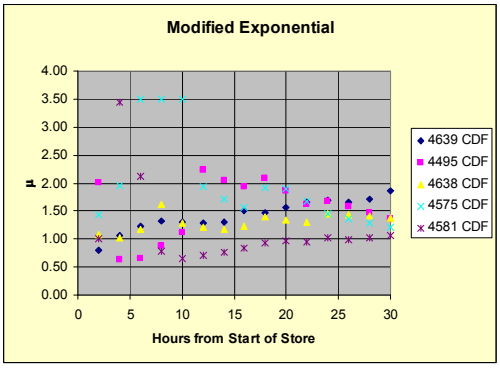
The modified exponential fit proves to be a very good predictor of the luminosity behavior. Figure aaa shows the CDF lumberjack data for the entire store (thick blue line) and each of the predicted luminosity curves.

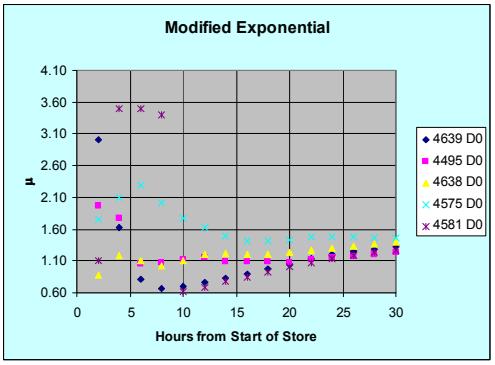

Modified Exponential
Initial Luminosity
Hours from Start of Store
4639 CDF
4495 CDF
4638 CDF
4575 CDF
4581 CDF

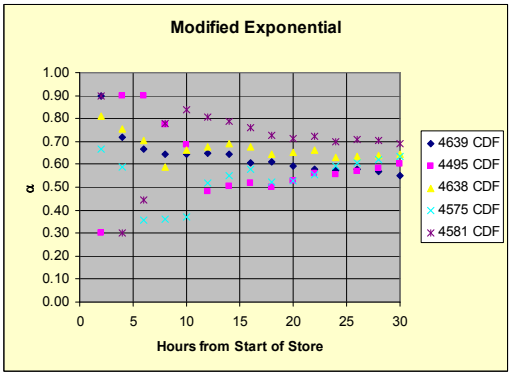

Modified Exponential
Initial Luminosity
Hours from Start of Store
4639 D0
4495 D0
4638 D0
4575 D0
4581 D0

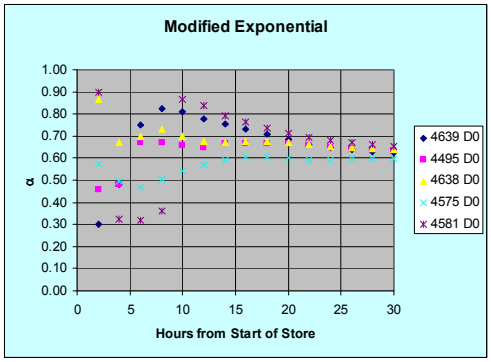

Modified Exponential
Luminosity Lifetime
Hours from Start of Store
4639 CDF
4495 CDF
4638 CDF
4575 CDF
4581 CDF

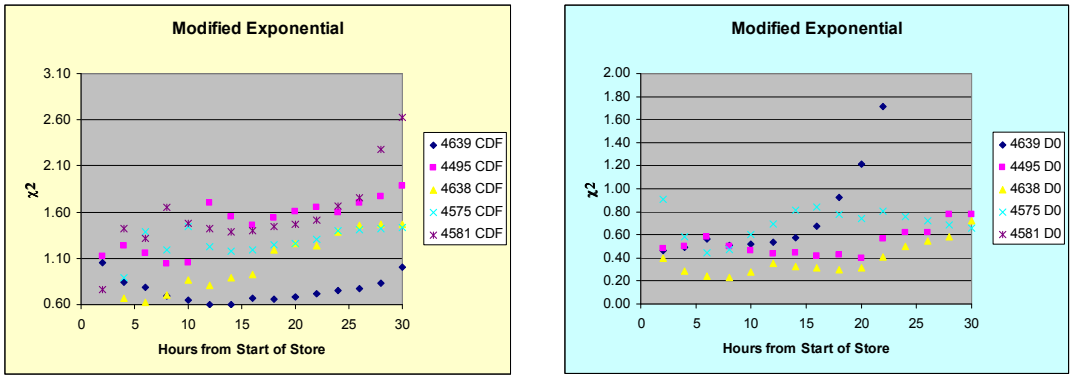

Modified Exponential
Luminosity Lifetime
Hours from Start of Store
4639 D0
4495 D0
4638 D0
4575 D0
4581 D0


Modified Exponential
μ
Hours from Start of Store
4639 CDF
4495 CDF
4638 CDF
4575 CDF
4581 CDF


Modified Exponential
μ
Hours from Start of Store
4639 D0
4495 D0
4638 D0
4575 D0
4581 D0


Modified Exponential
α
Hours from Start of Store
4639 CDF
4495 CDF
4638 CDF
4575 CDF
4581 CDF


Modified Exponential
α
Hours from Start of Store
4639 D0
4495 D0
4638 D0
4575 D0
4581 D0

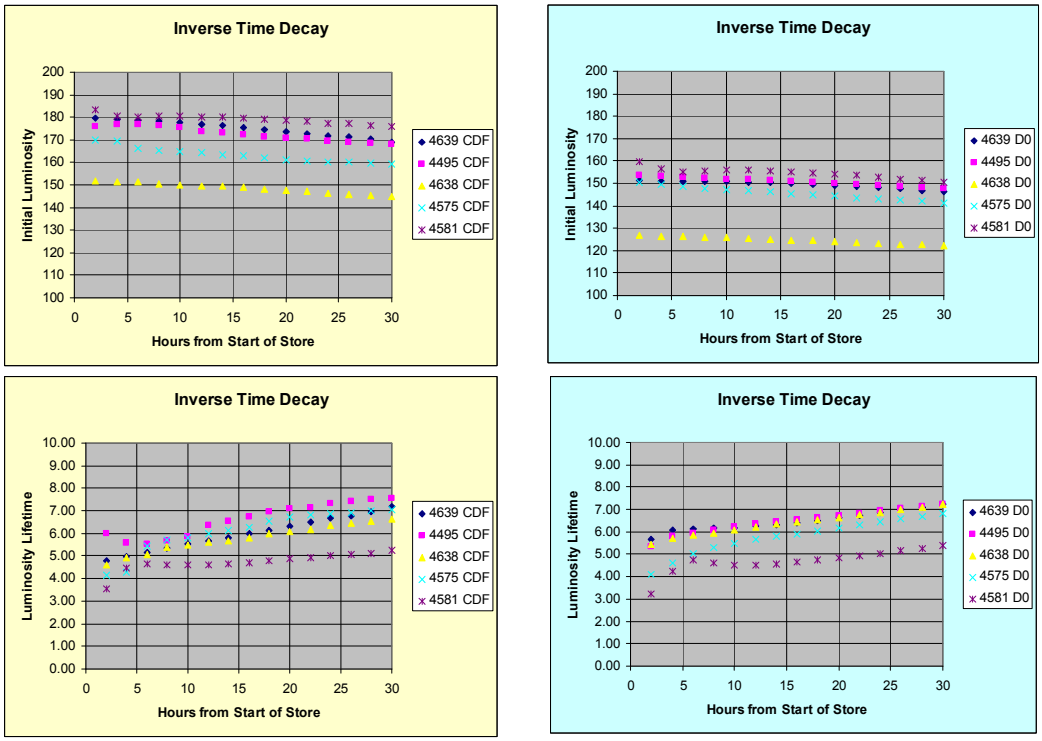


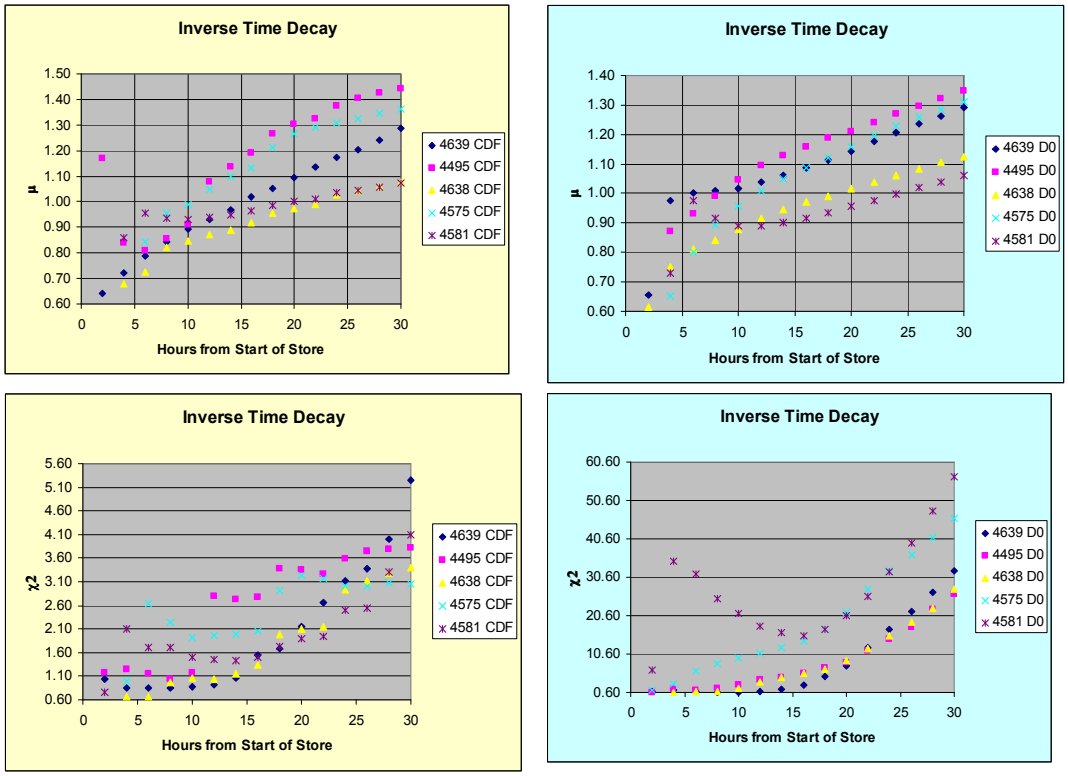
CDF D0

Figure 7-2

c. *Inverse Time to Power Fit (Equation 3)*

The inverse time to power fit proves to be a good predictor of the luminosity behavior. Figure aaa shows the CDF lumberjack data for the entire store (thick blue line) and each of the predicted luminosity curves.



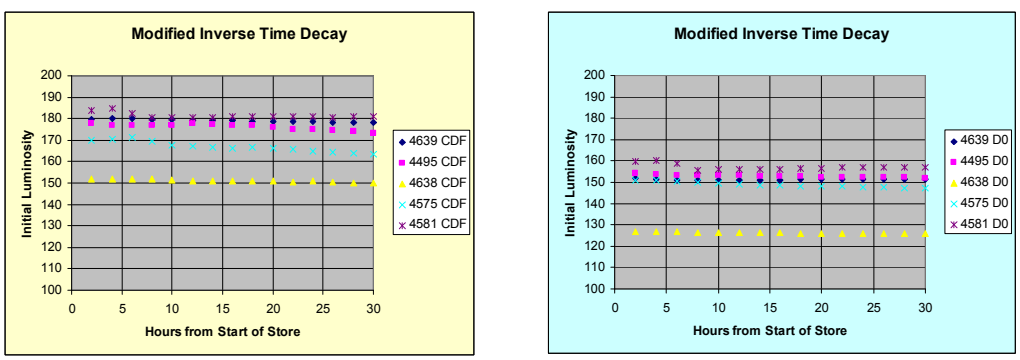


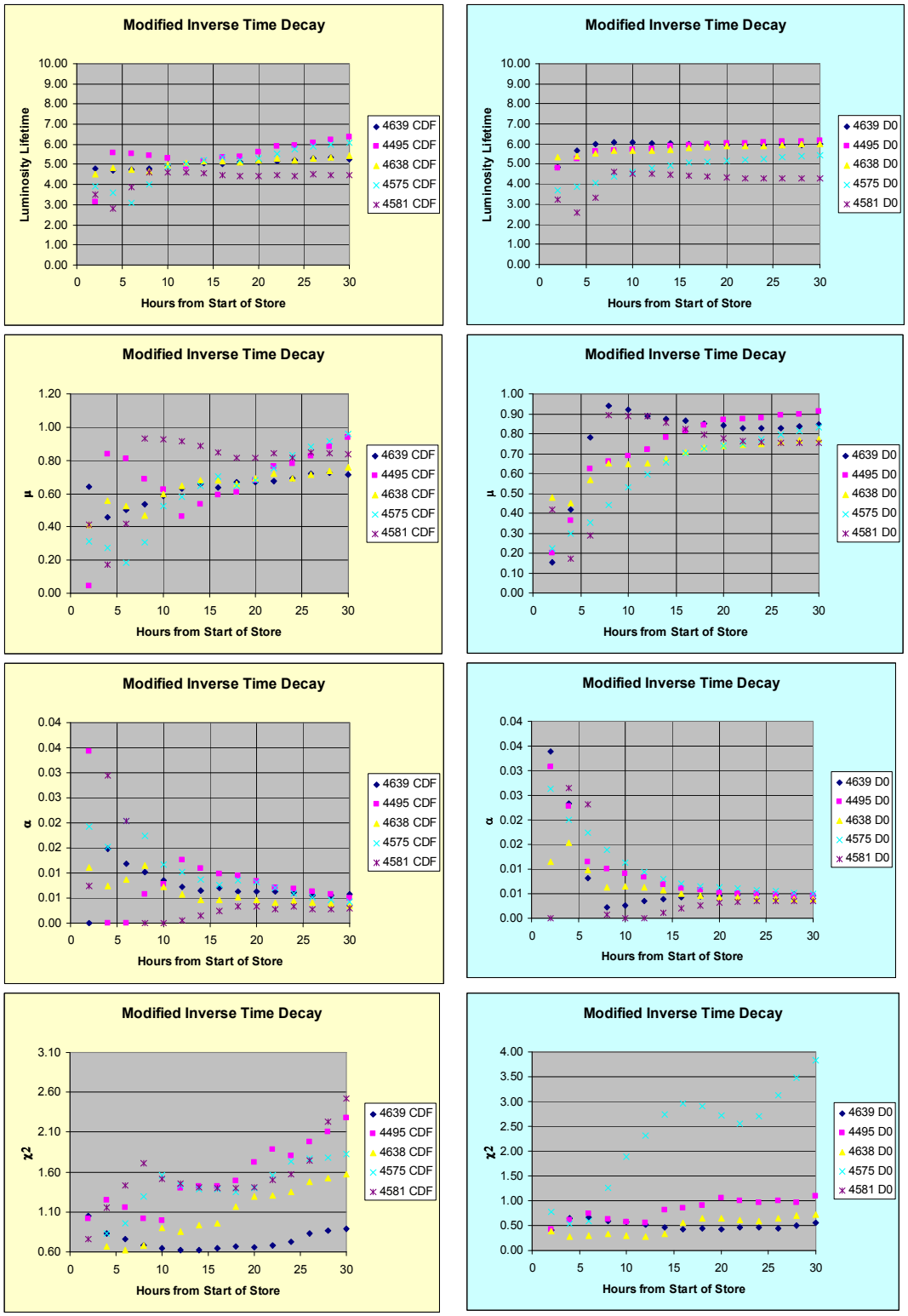
CDF D0

Figure 7-3

d. Modified Inverse Time to Power Fit (Equation 4)

The modified inverse time to power fit proves to be a good predictor of the luminosity behavior. Figure aaa shows the CDF lumberjack data for the entire store (thick blue line) and each of the predicted luminosity curves.





CDF D0

Figure 7-4

8. Conclusions

What did we learn?

9. References and Useful Sources

1. McGinnis, Dave. *Recycler-Only Operations Luminosity Projections*. Beams Document Database #2022. <http://beamdocs.fnal.gov/cgi-bin/public/DocDB/ShowDocument?docid=2022>.
2. McCrory, Elliot. *A Monte Carlo Model of the Tevatron*, Beams Document Database #829. <https://beamdocs.fnal.gov/AD-private/DocDB/ShowDocument?docid=829>.
3. McCrory, Elliot. *Tevatron Decay Fits*. Home page. 2006. <http://mccrory.fnal.gov/tevatronDecayFits/key.html>.
4. McCrory, Elliot. *Fitting the Luminosity Decay: Fits and Correlations*. Beam Document Database #1305. < https://beamdocs.fnal.gov/AD-private/DocDB/ShowDocument?docid=1305>
5. Roman, Steven. *Writing Excel Macros with VBA*, Ed. Warren T. Reich. 4 vols. O'Reily Publishers, 2002.
6. Liengme, Bernard V. *Microsoft Excel 2003 For Scientists and Engineers*. Third Edition. Elsevier Butterworth Heinemann, 2004.
7. Walkenbach, John. *Microsoft Excel 2000 Power Programming with VBA*. Hungry Minds, Inc. 1999.