# Luminosity Decay Fits: How Much Data Are Enough?

Elliott McCrory
June 30, 2006

## Introduction

The SuperTable[i] uses 1Hz data from the datalogger to fit the luminosity decay in the first two hours of a store in the Tevatron. We also have been doing more careful fits to the entire luminosity decay for the entire store[ii]. It seems apparent that 1Hz data were not necessary to obtain useful and accurate fits. Since SDA stores luminosity data during a store with a period of 10 minutes, the more detailed fits, which used SDA data, used the 10-minute interval for the source of the data. These fits seemed to be identical to the more copious data used in the SuperTable fits. We have not attempted until now to quantify how much data are actually necessary to make an acceptable fit.

## The Data

Store 4797 has been chosen because (1) it is new with respect to the writing of this document, (2) it is a long store, over 25 hours, (3) there are no discontinuous jumps in the luminosity data, and (4) it ended normally. Illustration 1 shows this store (CDF luminosity data) with the results of a fit to the form used successfully elsewhere:

$$L = L_0 e^{-t/tau}, Where: Tau = T_0 + C_1 * t^{-C_2}$$

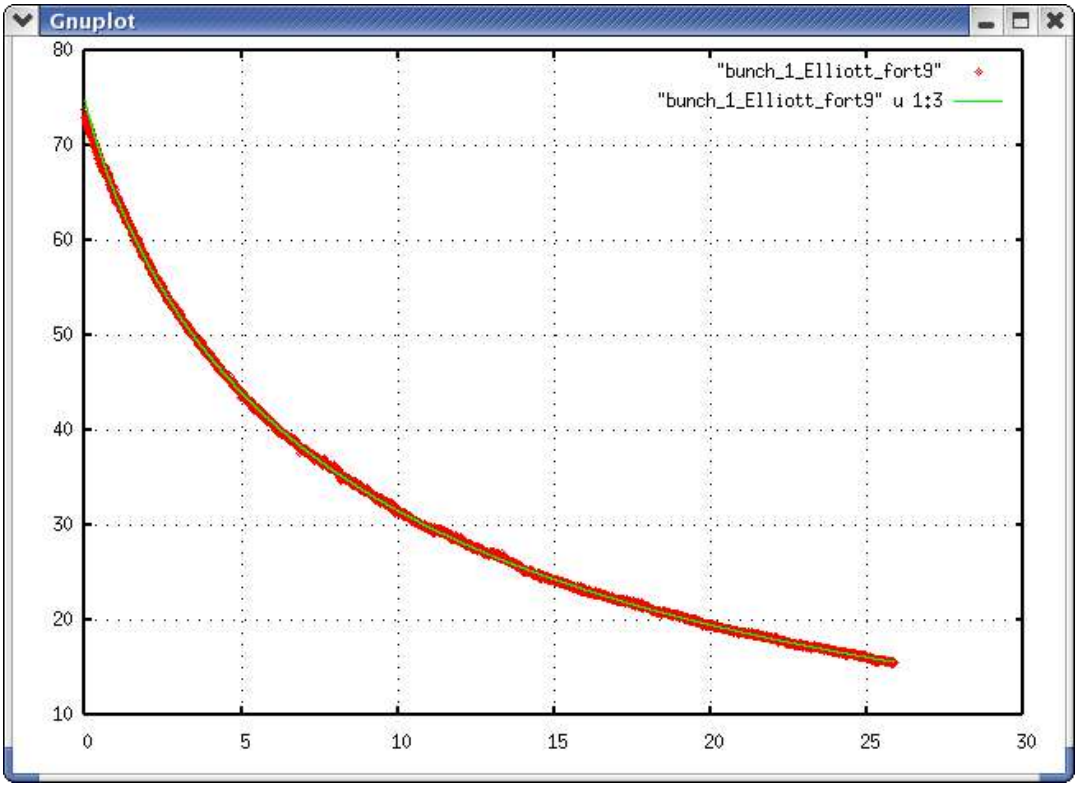


*Illustration 1, Store 4797, 1Hz data with fit*

The X-axis is time, in hours, and the Y-axis is the luminosity at CDF in the units E30 1/[cm**2 sec]. The numeric results of this (and all the other fits in this note) are shown in Table 1.

| Interval | Num Points | Chisq | ChiSQ/DOF | L0 | Tau | C1 | C2 |
|---|---|---|---|---|---|---|---|
| 1 | 93040 | 120557.7 | 1.3 | 74.7 | 5.37 | 1.52 | 0.61 |
| 2 | 46521 | 60298.99 | 1.3 | 74.71 | 5.37 | 1.52 | 0.61 |
| 4 | 23261 | 30132.67 | 1.3 | 74.7 | 5.38 | 1.52 | 0.61 |
| 8 | 11631 | 15034.97 | 1.29 | 74.7 | 5.37 | 1.52 | 0.61 |
| 16 | 5816 | 7523.44 | 1.29 | 74.71 | 5.37 | 1.52 | 0.61 |
| 32 | 2909 | 3712.03 | 1.28 | 74.69 | 5.38 | 1.51 | 0.61 |
| 64 | 1455 | 1841.33 | 1.27 | 74.7 | 5.4 | 1.5 | 0.61 |
| 128 | 728 | 927.02 | 1.28 | 74.65 | 5.43 | 1.49 | 0.62 |
| 256 | 365 | 461.01 | 1.28 | 74.67 | 5.43 | 1.49 | 0.62 |
| 512 | 183 | 239.8 | 1.34 | 74.7 | 5.41 | 1.5 | 0.62 |
| 1024 | 92 | 112.76 | 1.28 | 74.9 | 5.34 | 1.51 | 0.61 |
| 2048 | 47 | 56.7 | 1.32 | 74.86 | 5.37 | 1.5 | 0.61 |
| 4096 | 24 | 19.12 | 0.96 | 74.74 | 5.24 | 1.6 | 0.6 |
| 8192 | 13 | 6.39 | 0.71 | 73.99 | 5.36 | 1.66 | 0.59 |
| 16384 | 7 | 1.11 | 0.37 | 73.84 | 4.7 | 2.15 | 0.52 |

*Table 1, All Fit Results for Store 4797.*

## Reducing the Number Of Points

Several other fits to the data have been performed using the same functional form, but with fewer points, see Table 1. In all cases, the point at time=0 is retained, and the number of points taken for the fit is reduced by the factor in the "Interval" column by skipping this number of points. Since the source data is 1Hz, the frequency of the data in the fit decreases accordingly.

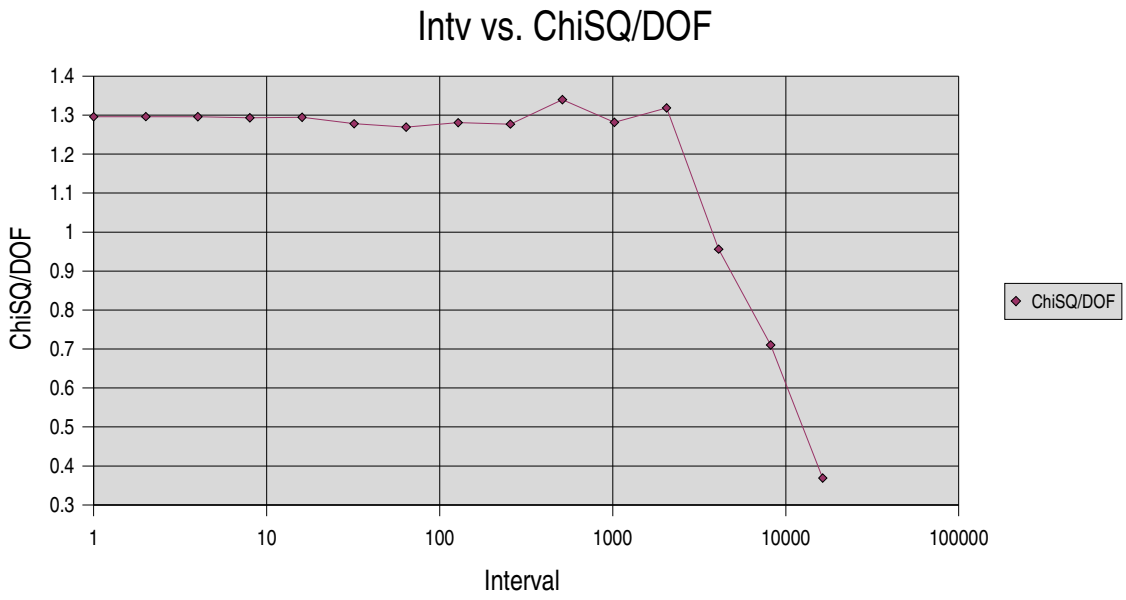


*Illustration 2, Interval versus Chi-squared per Degree of Freedom for Store 4797*

The standard measure of the goodness of fit is the Chi-squared per degree of freedom, "ChiSQ/DOF". This quantity is plotted here in Illustration 2 for each fit.

The quality of the fit does not change until after the 2048 interval, or one point every 34 minutes. Here in Illustration 3 is a plot of that fit result:

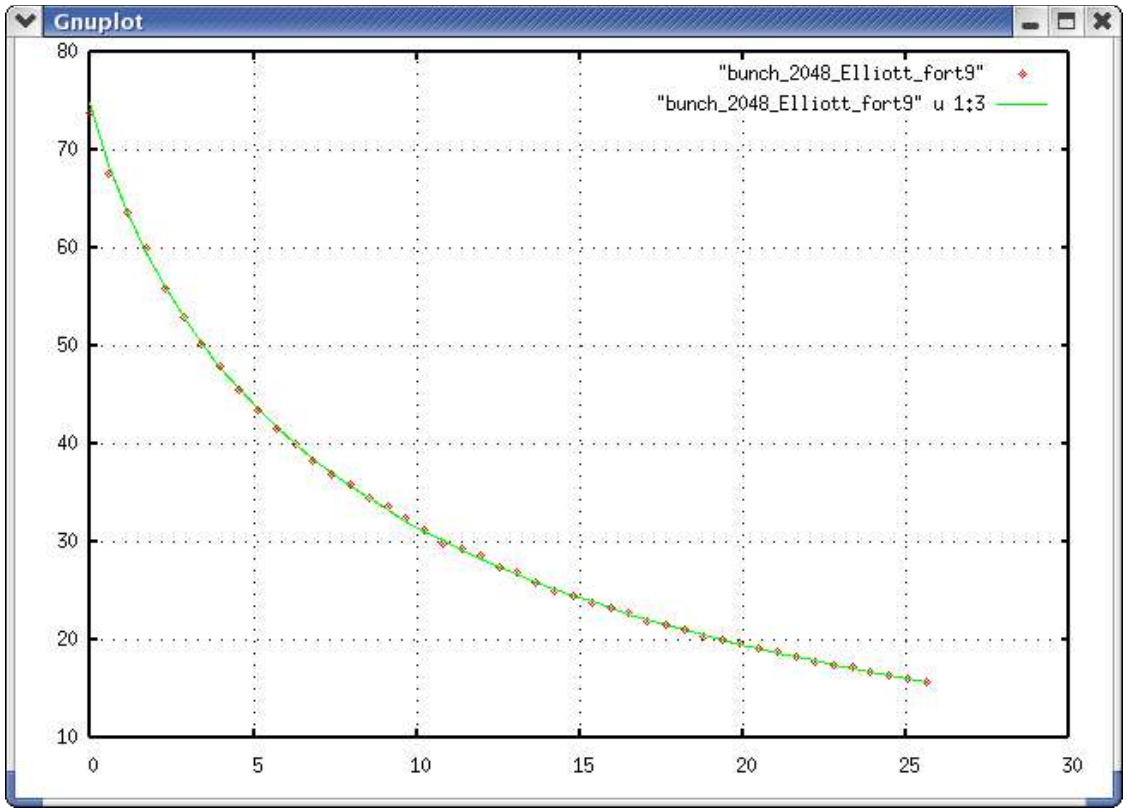


*Illustration 3, Data and fit for the 2048 Interval for Store 4797*

## Analyzing Store 4808

Store #4808 also satisfies the criteria established earlier, so identical fits versus the interval are shown in Table 2.

| Interval | Num Points | Chisq | ChiSQ/DOF | L0 | Tau | C1 | C2 |
|---|---|---|---|---|---|---|---|
| 1 | 100001 | 183851.5 | 1.84 | 125.91 | 4.7 | 1.61 | 0.59 |
| 2 | 33335 | 61161.96 | 1.83 | 125.92 | 4.69 | 1.61 | 0.59 |
| 4 | 20001 | 36932.31 | 1.85 | 125.91 | 4.7 | 1.61 | 0.59 |
| 8 | 11113 | 20265.27 | 1.82 | 125.96 | 4.68 | 1.62 | 0.59 |
| 16 | 5884 | 11036.65 | 1.88 | 125.92 | 4.7 | 1.61 | 0.59 |
| 32 | 3032 | 5345.52 | 1.77 | 125.78 | 4.74 | 1.6 | 0.59 |
| 64 | 1540 | 2809.4 | 1.83 | 125.78 | 4.74 | 1.6 | 0.59 |
| 128 | 777 | 1540.89 | 1.99 | 125.73 | 4.73 | 1.61 | 0.59 |
| 256 | 391 | 658.7 | 1.7 | 125.31 | 4.9 | 1.54 | 0.6 |
| 512 | 196 | 366.46 | 1.91 | 125.6 | 4.76 | 1.59 | 0.59 |
| 1024 | 99 | 154.13 | 1.62 | 126.43 | 4.55 | 1.65 | 0.58 |
| 2048 | 50 | 71.01 | 1.54 | 126.03 | 4.73 | 1.57 | 0.6 |
| 4096 | 26 | 39.93 | 1.82 | 125.32 | 5.09 | 1.42 | 0.62 |
| 8192 | 14 | 12.37 | 1.24 | 124.77 | 4.89 | 1.58 | 0.59 |
| 16384 | 8 | 3.3 | 0.83 | 124.64 | 4.99 | 1.52 | 0.6 |

*Table 2, Fit results for Store 4808*

Similar, but not identical results are obtained. In particular, the ChiSQ/DOF versus interval chart, Illustration 4, is not as uniform out to the 2048 interval as in Store 4797.

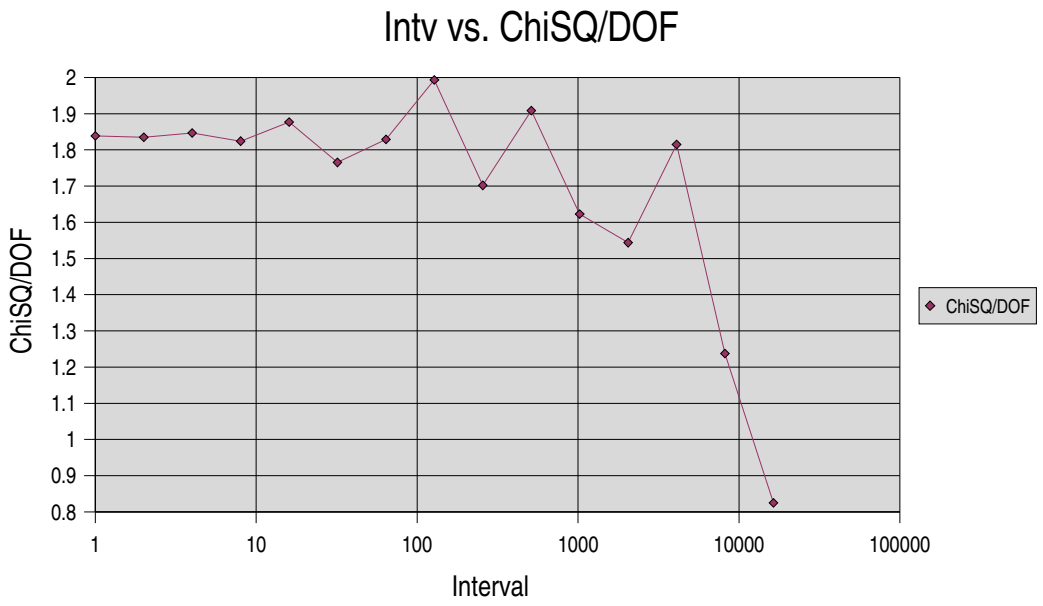


*Illustration 4, ChiSQ/DOF for store 4808 fits*

The ChiSQ/DOF does not drop off as clearly as for store 4797, but it looks as if the 2048 interval is still a good, believable fit.

## Comments on the Error Bars

The uncertainty on each luminosity measurement is assumed to be 0.15 E30. The fitting program, Minuit, estimates the uncertainty of the values of each fit result. The fractional error in each fit number with respect to the interval in question, as calculated by Minuit, is shown in Illustration 5. This uncertainty increases approximately as the logarithm of the fit interval. But note that the uncertainty at 2048 of the initial luminosity parameter is below 1% and the others are between 2 and 10%.

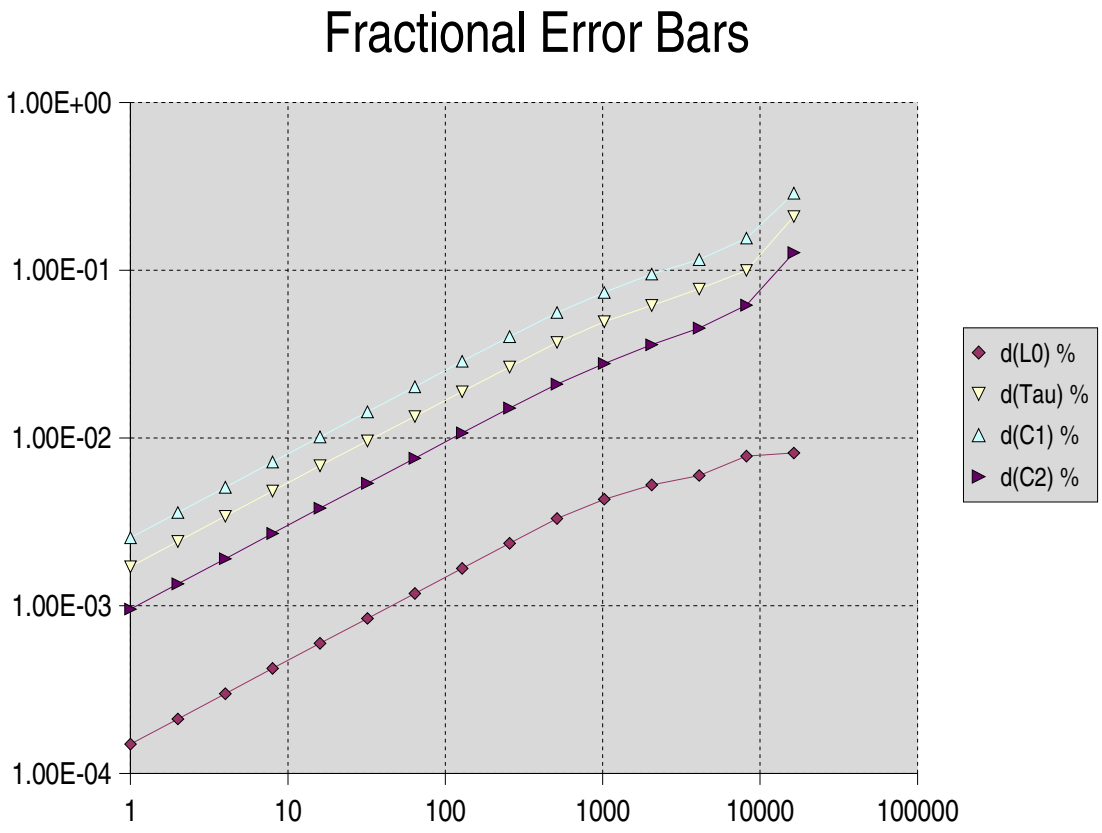


*Illustration 5, Uncertainty of fit result numbers for Store 4797*

The fractional error bars for Store 4808 are shown in Illustration 6.

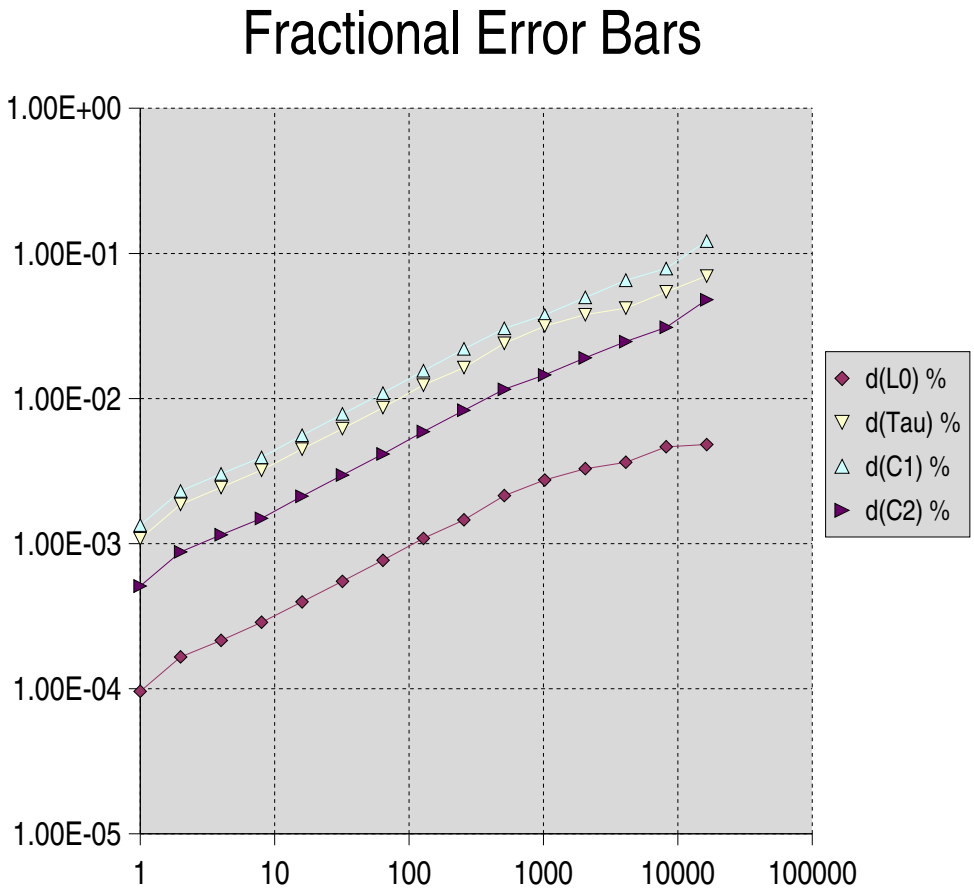


*Illustration 6, Fraction error bars for Store 4808 fits*

## Discussion

An interesting feature of the original data is shown in Illustration 7, blown up of the CDF luminosity of this store around hour 10. Clearly, the CDF data is not 1Hz—it looks as if there are between 10 and 14 points with the same value. Therefore one does not need more than 1 point every 10 seconds. (The D0 luminosity readout, C:D0FZTL, changes once every 5 seconds.)

The problem with a smaller data set in this situation is the impact of a single bad point: this impact will be higher with fewer points. So if one chooses to perform the fit with fewer points, it is more important to reject bad points. This analysis did not reject any points.

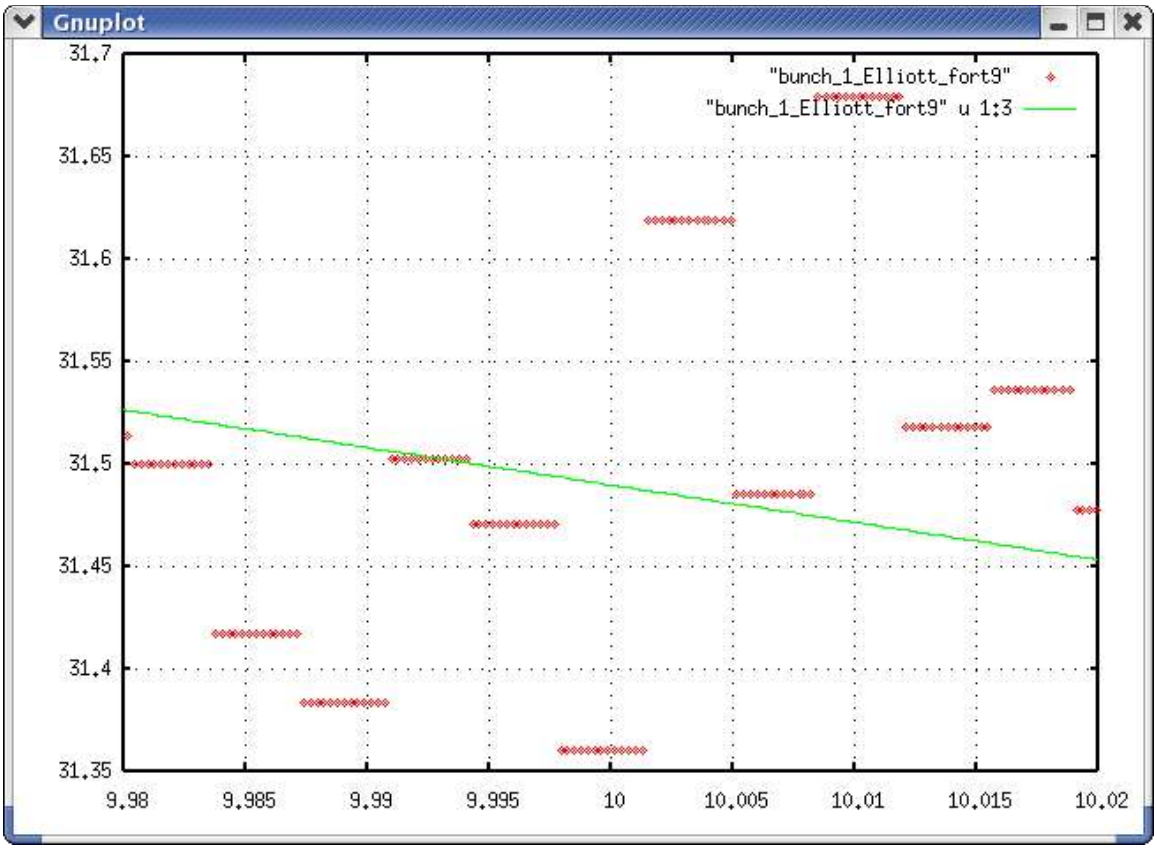


*Illustration 7, Blow-up of the full, 1Hz data around the 10 hour mark*

## Conclusion

The quantity of data necessary to get a good fit to the luminosity decay from the CDF data on a typical can be as small as about one point every 30 minutes. Therefore SDA data, one point every 10 minutes, is equivalent to 1Hz data. Some statistical certainty in the fit results is sacrificed by using the longer interval, but the resulting fit seems to be completely adequate.

i http://www-bd.fnal.gov/sda/supertable
ii http://tomato.fnal.gov/tevatronDecayFits