Luminosity Decay Fits:
How Much Data Are Enough?

Elliott McCrory
June 27, 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:

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

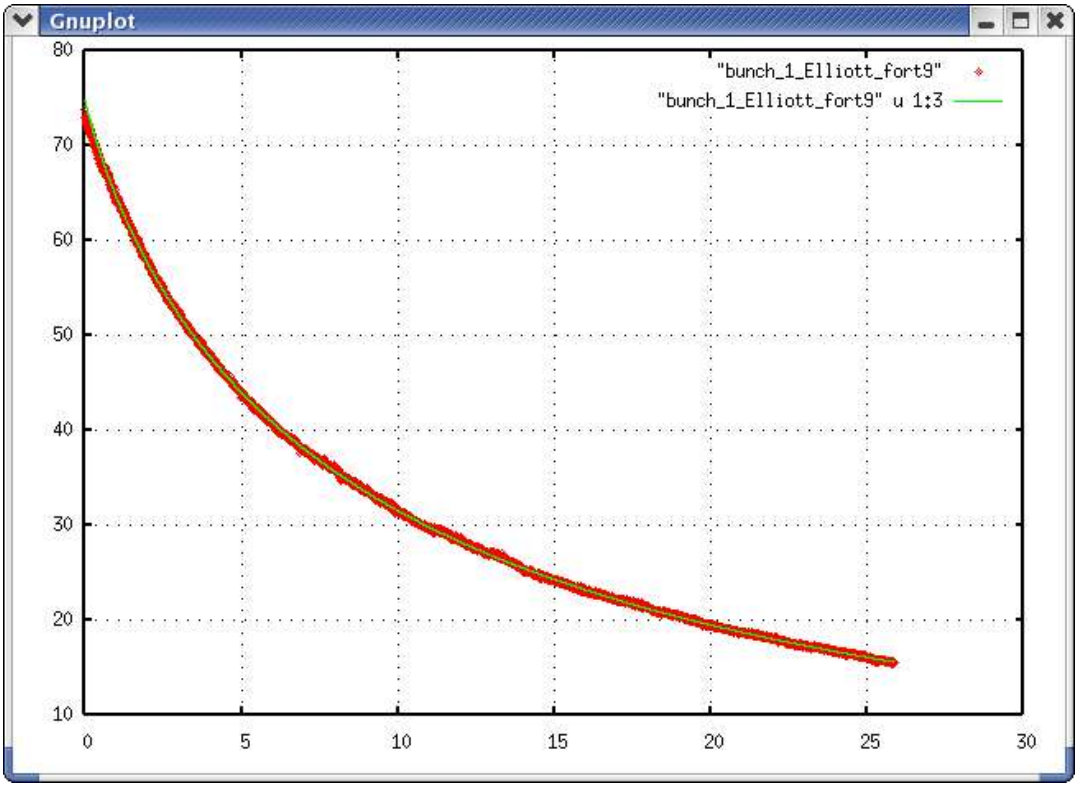

Illustration 1, Store 4797, 1Hz data with fit

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	1085019	11.66	74.7	5.37	1.52	0.61
2	46521	542690.9	11.67	74.71	5.37	1.52	0.61
4	23261	271194.1	11.66	74.7	5.38	1.52	0.61
8	11631	135314.7	11.64	74.7	5.37	1.52	0.61
16	5816	67710.95	11.65	74.71	5.37	1.52	0.61
32	2909	33408.22	11.5	74.69	5.38	1.51	0.61
64	1455	16571.99	11.42	74.7	5.4	1.5	0.61
128	728	8343.19	11.52	74.65	5.43	1.49	0.62
256	365	4149.07	11.49	74.67	5.43	1.49	0.62
512	183	2158.21	12.06	74.7	5.41	1.5	0.62
1024	92	1014.85	11.53	74.9	5.34	1.51	0.61
2048	47	510.28	11.87	74.86	5.37	1.5	0.61
4096	24	172.12	8.61	74.74	5.24	1.6	0.6
8192	13	57.52	6.39	73.99	5.36	1.66	0.59
16384	7	9.97	3.32	73.84	4.7	2.15	0.52

Table 1, Fit Results for All Intervals

(Note: The ChiSQ/DOF value of 11 indicates that the error bars in this fit are about sqrt(11) too small. The numerical values of these fits are unaffected by this

choice.)

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.

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.

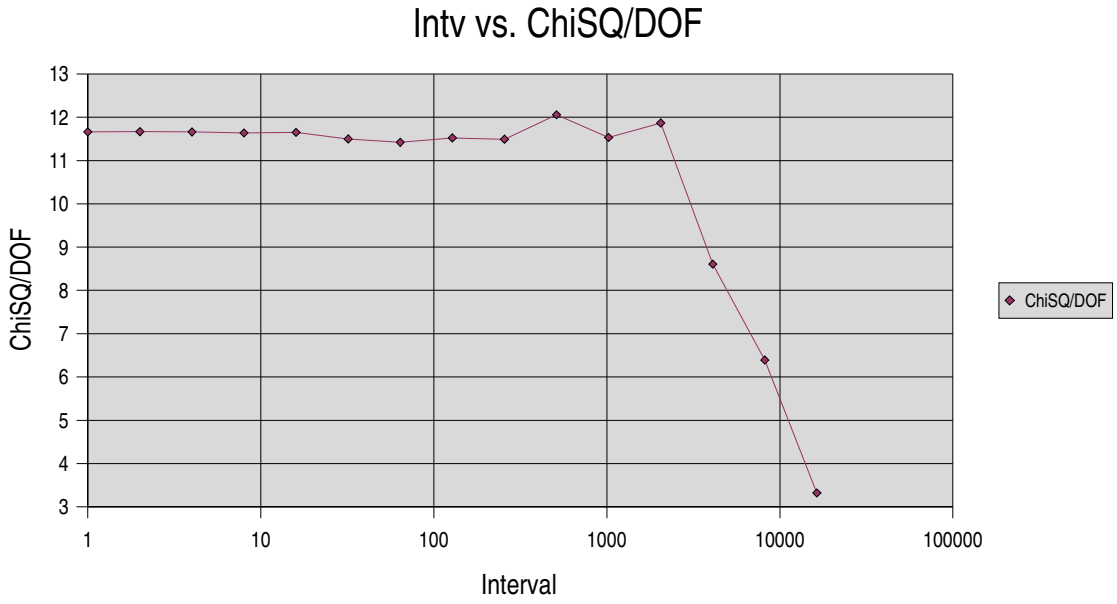

Illustration 2, Interval versus Chi-squared per Degree of Freedom

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:

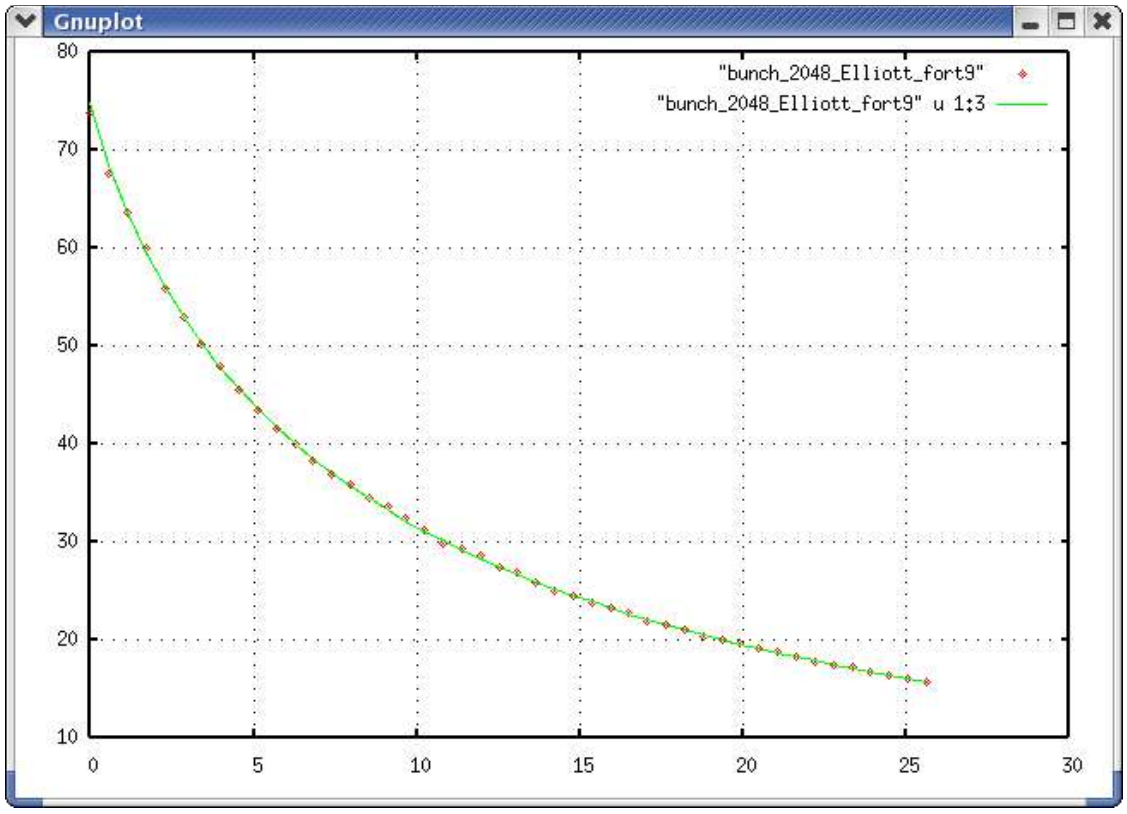

Illustration 3, Data and fit for the 2048 Interval

Discussion

An interesting feature of the original data is shown in Illustration 4, blown up 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 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 be sure to reject bad points. This analysis did not reject any points—it was a clean store.

It is not completely clear if this store is typical or extraordinary with respect to this analysis. No other stores after the 2006 shutdown yet satisfy the four criteria listed above.

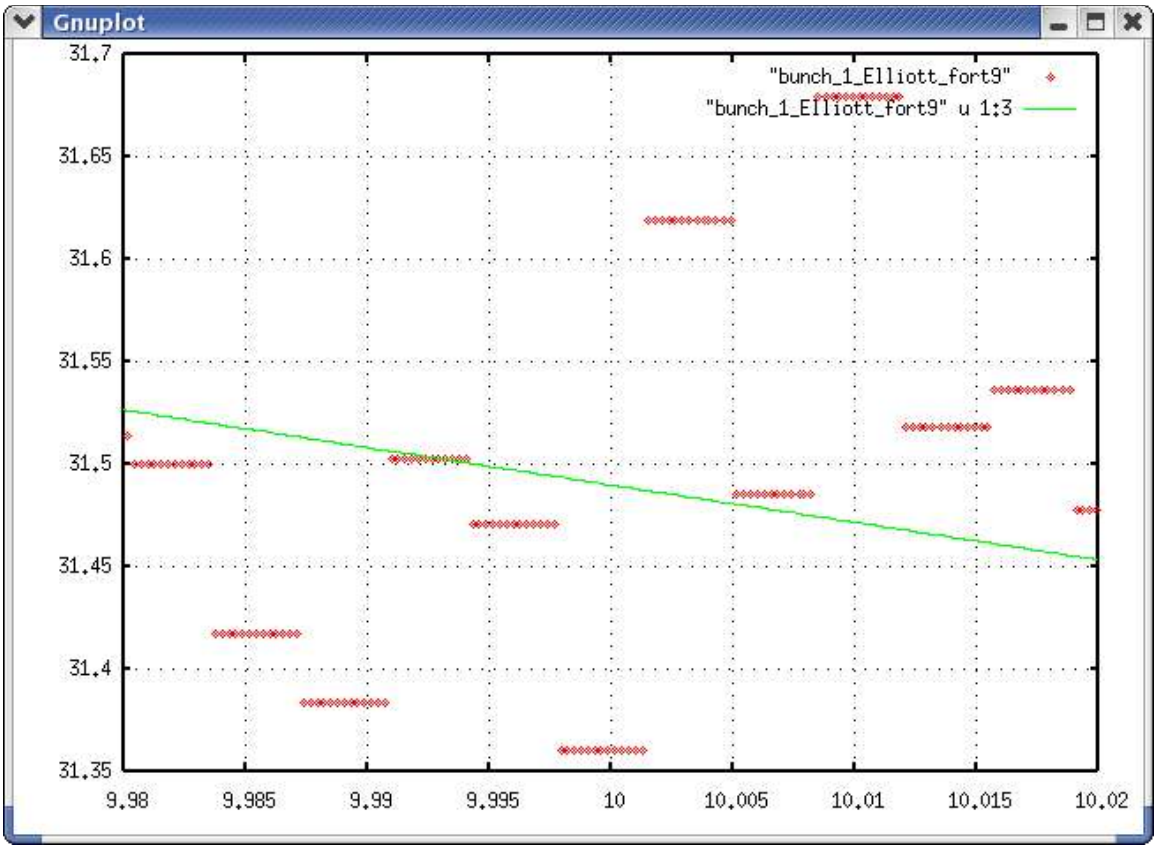

Illustration 4, Blow-up of the full, 1Hz data around the 10 hour mark

Conclusion

The quantity of data necessary to get a good fit to the luminosity decay from the CDF data on a typical store should be greater than about one point every 30 minutes. SDA data, one point every 10 minutes, is equivalent to 1Hz data.

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