SLAC-TN-04-057
September 2004

Analysis of Luminosity Data in BaBar

Rebekah D. Schiller
Office of Science, Student Undergraduate Laboratory Internship (SULI)
Oglethorpe University
Stanford Linear Accelerator Center, Menlo Park, California

August 2004

Prepared in partial fulfillment of the requirements of the Office of Science, Department of Energy's Student Undergraduate Laboratory Internship under the direction of Steffen Luitz and Matthias Wittgen in the Computing Services (SCS) division at the Stanford Linear Accelerator Center

Participant: Rebekah D. Schiller

Research Advisors:

Steffen Luitz

Matthias Wittgen

Work supported in part by Department of Energy contract DE-AC02-76SF00515.

Contents

List of Figures

Abstract

Luminosity is a value describing the number of interactions between particles when their respective beams collide. The BaBar (B and B-bar) ambient database and the Oracle server contain archived measurements of instantaneous luminosity from two separate detectors. It is important to understand these data to describe the performance of the B-Factory. Extracting these data, the more reliable PEP-II (Positron Electron Two) luminosity detector can be calibrated to the data collected by the more accurate BaBar L3 (Level 3) detector. Using the ROOT programming language and standard BaBar tools for data extraction, graphs and statistics are generated. Some logistical errors contained in the logbook are also corrected. These programs help aid in understanding both what is happening within the B-Factory, as well as the correlations between the interrelationship of detector data and the information recorded in the logbook.

1 Introduction

In the BaBar detector, it is useful to know how many interactions there are between electrons and positrons when their respective beams collide. Knowing this number, referred to as the luminosity, can help in understanding the experiments implemented in BaBar and in making them more efficient. In addition, luminosity readings help describe the performance of the B-Factory. Luminosity can be calculated using the following equations.

$$\frac{dN}{dt} = \sigma \cdot L \tag{1}$$

$$N = \sigma \cdot \int L dt \tag{2}$$

These equations describe the relationship between N, the number of occurrences of a particular reaction in a given interaction, σ, the cross section of the reaction, and L, the instantaneous luminosity. $\int L dt$ is the integrated luminosity, which is the sum of all of the instantaneous luminosity measurements in a given time interval multiplied by that time interval. During any experiment, the luminosity is measured constantly to find N.

There are two detectors in the B-Factory that detect the instantaneous luminosity: the PEP-II (Positron Electron Project Two) luminometer and the L3 (Level Three) BaBar luminosity detector. The data from these detectors are converted to integrated luminosity measurements at different rates. The PEP-II integrated luminosity measurements are reset every shift; the L3 measurements are reset every run. A new shift begins every eight hours, whereas a run may be an arbitrary amount of time, never longer than one hour. The PEP-II detector is always on, and therefore reliable, yet it needs to be calibrated against the more accurate BaBar detector about once a week to give useful numbers. The L3 detector, on the other hand, may have gaps in information. These gaps could stem from data that never was counted, data that was detected, but not recorded for one reason or another, or dead time, which occurs when the BaBar Data Acquisition System (DAQ) is too slow at recording the fast-paced reactions.

These detectors send their information to two separate databases. The Oracle server is organized both by shift and by run. This contains some raw data in addition to calculated statistics, such as the L3 recorded luminosity corrected by the dead time, BaBar recorded integrated luminosity, or ratios between the L3 information and the PEP-II information. The logbook is the interface from which most people see what data were recorded in the Oracle server; it is available on the Internet. The next database is the BaBar ambient database, which contains archived measurements from both of the detectors.

This current data storage system can sometimes become problematic. For instance, the BaBar recorded integrated luminosity data is recorded only once per run, which may span a shift change. The logbook records the numbers per shift, which means some of a shift's data can be recorded in the logbook as the previous shift's data. This makes the information in the logbook inaccurate.

A correlation that can be useful to understand is how the highest luminosity achieved during a specific run affects the total luminosity for that run. One theory is that the higher the peak luminosity is, the lower the overall luminosity during that run will be. To check this, the data contained in the logbook must first be more precise.

To achieve greater exactness, we must be able to interpolate how much of the shift's data is recorded into the former shift, and incorporate this into all of the calculations. It would also be helpful to put up some kind of warning in the logbook when data has not been recorded from the L3 detector, and how the numbers were estimated.

In this project, information will be extracted from the two databases to analyze what is happening within the B-Factory. These data can be used to find correlations between the detector data, calculate various statistics in order to calibrate the detectors, and create improvements in the logbook. Programs written in the ROOT programming language will be used to calculate these statistics and create histograms of certain data to obtain a better understanding of luminosity measurements. For more technical information about BaBar see Reference [1].

2 Materials and Methods

A number of computer programs were written for this project. The programs are generally constructed to have the user input a shift ID to see a histogram on the computer screen, along with some data. A shift ID is the identifying number associated with any given shift, and is constructed in the following manner: it is one nine digit integer with no spaces, consisting of the year, followed by the numerical, two-digit month, the two-digit day, and the shift number. The shift number can either be 0, signifying the “owl” shift (midnight to 8:00), 1, the “day” shift (8:00 to 16:00), or 2, the “swing” shift (16:00 to midnight). This shift ID number is used to obtain information from either database.

In the first program, both databases are accessed for data pertaining to luminosity. The program will loop over all the shifts in a given period of time, and create a tree saved to file containing data such as the instantaneous luminosity recorded in both the Ambient database and the logbook, the integrated luminosity from both databases, the ratios of these data, the integrated luminosity calculated from the sum of the instantaneous luminosity recorded by the PEP-II detector, and the shift IDs. Each branch of the tree has one entry for every shift looped over. This tree is then accessed in another program. It prints nine separate histograms in one window. Each histogram is constructed with one entry per bin, and is therefore a timeline of data history.

In the “Instantaneous Luminosity” histogram, each bin reads out the calculated sum of all of the PEP-II detected instantaneous luminosity in that shift. In the “BaBar vs. Intlum” histogram, each bin contains the ratio of L3 detected integrated luminosity per shift and the same data gathered by the PEP-II detector. The “L3 Recorded Intlum” histogram shows the L3 detected integrated luminosity in a shift as recorded in the logbook. The “PEP 2 Luminosity” histogram is the sum of all instantaneous luminosity data in a shift as recorded in the logbook. The “Integrated Luminosity” histogram shows the sum of the total integrated luminosity for every shift. The “BaBar-recorded Integrated Luminosity” histogram contains the sum of the L3 integrated luminosity per shift. The “Logbook PEP/BaBar Luminosity

Ratio" histogram contains the calibration factor recorded in the logbook between the two detectors. In the "BaBar Recorded Instantaneous Luminosity" histogram, each bin contains the sum of the L3 recorded instantaneous luminosity for that shift. The "PEP/BaBar integrated Luminosity Ratio" histogram shows the program's calculation of what the calibration factor between detectors should be, using the data recorded in the logbook. See Appendix B, Figure 3 for an example of the output window from this program.

Whereas the latter program mentioned must be run after a tree is built using the first program, the next program is self-contained. The user types the start and stop shift IDs for the program at a prompt. The program then loops over all the shifts in this interval, reading data from the Oracle database of the PEP-II instantaneous luminosity and beam information, and creates a tree with this information. The total luminosity per shift is calculated and the highest luminosity reading detected during that shift is found. The program then creates a multi-layered, two-dimensional histogram of this information and color codes each shift according to the month in which it occurs. This histogram also includes a line on which the maximum achievable value of total luminosity for every possible peak luminosity rests. The PEP-II detector data is used here because of its more accurate measurements. In Appendix B, Figures 1 and 2 contain two examples of this output.

One more program is useful to look at in conjunction with the last. It first calculates the length, in seconds, of a hypothetical shift whose instantaneous luminosity is constant at the peak luminosity measurement for a real shift, and whose total luminosity is equal to the total luminosity of that real shift. This calculation is done by computing the integrated luminosity of a shift, and dividing it by the peak luminosity of that shift. The length of this hypothetical shift is then divided by the number of seconds in a real shift (28,800). This number represents how efficient each shift is, with 1 being the most efficient, and zero being the worst. The program next creates a histogram with one entry per bin. Examples of the output from this program are found in Appendix B, Figures 4 and 5.

All of the aforementioned programs take into account the phenomenon referred to in the introduction where a run may span a shift change. This is corrected by first creating a new tree for every shift ID, containing data collected during that shift and the previous one. Then, the programs loop over all runs contained in a given shift to find which, if any, overlap into the current shift from the one before it. If the run spans a shift change, the amount of overlap time is calculated, and the fraction of the overlap which occurs in the current shift is found. This fraction is multiplied by the value of the relevant data, estimating the data collected in the shift at hand. This is likewise done for the other part of the run, and included in the data for the previous shift.

3 Results

As is easily seen in the histograms in Appendix B, Figures 1 and 2, the average luminosity achieved in one shift has grown significantly over the last two runs. In Fig. 1, the October data were all zeroes, because data was not collected during this time. As the months go by, the average peak luminosity goes up considerably. The beams were not run in July and August as well, and therefore these data are also all zeroes. In Fig. 2, the peak luminosity per shift is pushed higher and higher for every month. In August, the data are all zeroes, following the above reasoning. During the months of May, June, and July, the average peak luminosity is about the same, and the total luminosity varies to about the same extent. From April until July, the data seems to lay closer to the maximum efficiency line.

Both histograms in Figures 4 and 5 do not seem to have much of a noticeable trend. Fig. 4 seems to have a maximum around .9, while Fig. 5 has ratios very close to, and sometimes greater than, one. In Fig. 4, the data occurring during the non-data taking shifts are zero.

The nine histograms printed together in Fig. 3 were constructed with data from the shifts occurring in April and May of this year. The data contained in the "Instantaneous Luminosity" histogram seems to fluctuate during these two months. Although the ratio in the "BaBar vs. Intlum" histogram is near one most of the time, there are a few shifts when this data is large or small. During these months, most of the data contained in the "L3 Recorded Intlum" histogram is between 0 and 250 pb^{-1} (inverse picobarns). This histogram also contains some very high and very low data. The data contained in the "PEP 2 Luminosity" histogram is nearly the same as the data in the "Instantaneous Luminosity" histogram, the "BaBar-recorded Integrated Luminosity" histogram, the "BaBar Recorded Instantaneous Luminosity" histogram, and the "Integrated Luminosity" histogram. The data in the "Logbook PEP/BaBar Luminosity Ratio" histogram fluctuates around 1.13 except for one entry, which is very high. The data contained in the "PEP/BaBar integrated Luminosity Ratio" histogram fluctuates between around .5 to 1.5, with some scattered zeroes and one very large reading.

4 Discussion and Conclusion

We decided to create histograms from the data collected during the dates chosen for non-arbitrary reasons. What is known as run 3 spanned from October of 2002 to August of 2003. Run 4 began in November of 2003 and ended this month. During these long runs the accelerator is on almost all of the time. Two months out of the year the accelerator is down for maintenance and repairs. Also, as of March 2004, SLAC began to use the double trickle injection method of maintaining the PEP-II beams. Before this, data could not be collected during the injection of the PEP-II High Energy Ring (HER) and Low Energy Ring (LER), when electrons or positrons were injected into their respective rings at one time, and not again until the beams were low. Using the double trickle injection method, the rings are kept constant by continuously injecting particles. This method improves the luminosity of experiments greatly, as the histograms in Appendix B show.

In Fig. 1, the zero data from October shows that the accelerator was still not running from the scheduled maintenance break at this time. The increase of both peak luminosity and total luminosity as time goes on shows the increased efficiency in how the experiments were run. The zero data from July and August show that a break in experiments occurred during these months as well.

In Fig. 2, it is easily seen that the peak luminosity starts at a level analogous to that of the run 3 data, and steadily increases until the run ends at the end of July. The double-trickle injection method does not seem to affect the peak luminosity significantly, but the probability that a run will be more efficient goes up after this method was implemented. This is evident in the fact that the data for these months are closer to the maximum efficiency line. The values that are greater than one point to problems with the raw luminosity data during these shifts.

The two histograms in Figures 4 and 5 show that, although there are many shifts whose down beam time offset the good beam time, making the values vary greatly, the maximum efficiency went up over time. This is consistent with the histograms from the other programs. The zero data at the beginning and end of Fig. 4 is again from the scheduled shutdown of the accelerator. It is not yet known why some of the values in Figure 5 are greater than one.

The nine histograms printed together in Fig. 3 were constructed with data from the chosen shifts because these months show the luminosity data directly after the double-trickle injection method was implemented. The "Instantaneous Luminosity" histogram shows the sum of all of the instantaneous luminosity readings for every shift. Because this number is not corrected for efficiency, a very low value could mean either many low luminosity readings or normal readings spaced with a high amount of dead time during the shift. The "BaBar vs. Intlum" histogram reads, as is expected, near unity at all shifts. When the data is either large or small, this means that one detector is reading very different information from the other. This could be caused by either a high statistical error developed from very little data being recorded, or somehow the data entries never being recorded in the logbook. In the "L3 Recorded Intlum" histogram, the negative and very high data point to bookkeeping errors in the logbook as well. Many of the histograms constructed in this program should look much like each other because they are calculating the same data from different detectors.

This explains why the “PEP 2 Luminosity” histogram looks the same as the “Instantaneous Luminosity”, the “Integrated Luminosity” histogram, the “BaBar Recorded Instantaneous Luminosity” histogram, and the “BaBar-recorded Integrated Luminosity” histogram. Because these are related, we can know that the logbook recorded the same data that the program calculated from raw data, and is a way to compare data and statistics. This also shows how accurate the logbook is.

Because the data is exactly what is recorded in the logbook, the one anomalous entry in the “Logbook PEP/BaBar Luminosity Ratio” histogram is probably caused by another logbook error. The “PEP/BaBar integrated Luminosity Ratio” histogram is expected to look close to the “Logbook PEP/BaBar Luminosity Ratio” histogram. The fact that the two look different probably means that the logbook’s calibration factors are inaccurate. A zero reading in this histogram means that the logbook shows that the L3 recorded integrated luminosity for that particular shift is zero. The very large readings could possibly come from a logbook error.

The programs written for this internship are useful for public relations and understanding methods of data collection. The nine histograms in the same window show information that is practical for promoting the understanding of the threefold relationship between the two detection devices and the logbook. Comprehending the way the detectors work together and their specific limitations can aid in the understanding of the data recorded, as well as helping in the possible design of new detectors. The histograms involving the efficiency of the luminosity can help scientists understand how to continuously be profitable at collecting as much data as possible in a shorter amount of time. These histograms can also be used to show the impressive value of how the much luminosity in the B-Factory has increased in the past. The method in which the integrated luminosity statistics were calculated is an improvement on the way the integrated luminosity is calculated in the logbook, because it takes into account split runs. This means that the logbook can become more accurate.

References

[1] B. Aubert *et al.* [BABAR Collaboration], "The BaBar detector," Nucl. Instrum. Meth. A **479** (2002) 1 [arXiv:hep-ex/0105044].

[2] "ROOT User's Guide v.4.08,"(2004, July 2).[Online] Available: http://root.cern.ch/root/doc/RootDoc.html

[3] Griffiths, D.F. and D.J. Higham, *Learning LaTeX*, Society for Industrial and Applied Mathematics, Philadelphia, 1997.

A Acknowledgments

I would like to thank the U.S. Department of Energy (DOE), Office of Science, for creating, funding, and organizing this Science Undergraduate Laboratory Internship (SULI), and the Stanford Linear Accelerator Center (SLAC) for hosting this opportunity. While working here, mentors Matthias Wittgen and Steffen Luitz taught me the ROOT programming language, use of the writing tool LaTeX, better C++ skills, and more about computer programming structure and design. Thank you to all the lecturers who introduced me to the most current scientific problems and the physics behind particle accelerators, and strengthened my interest and resolve to attend graduate school in Physics. I extend my gratitude also to Helen Quinn, manager of the SULI program and Science Education Director at SLAC, Roberto Vega, SULI program director at SLAC, and Pauline Wethington, SULI program coordinator at SLAC, for working to provide such a wonderful learning experience for all those involved.

B Figures

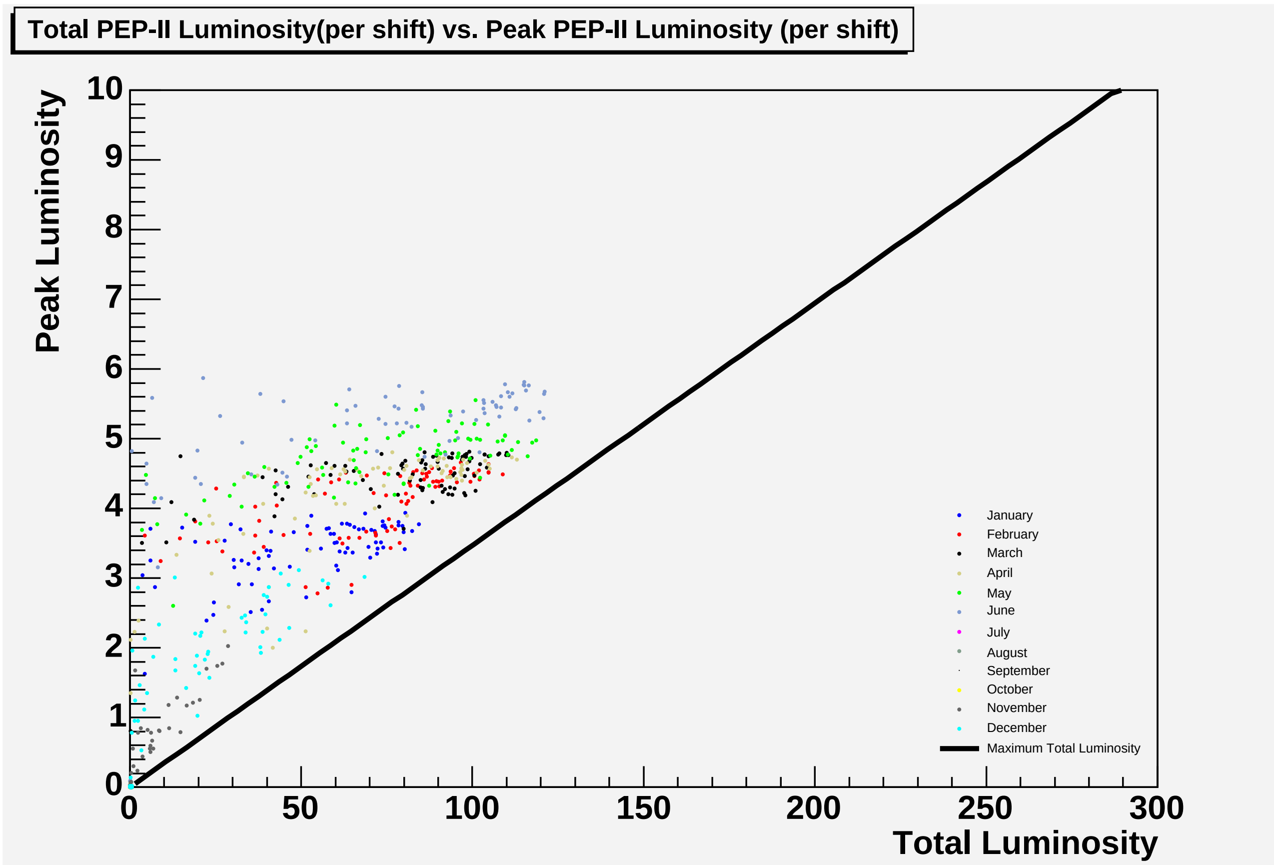

Figure 1: Run 3 data: Peak Luminosity vs. Total Luminosity per Shift

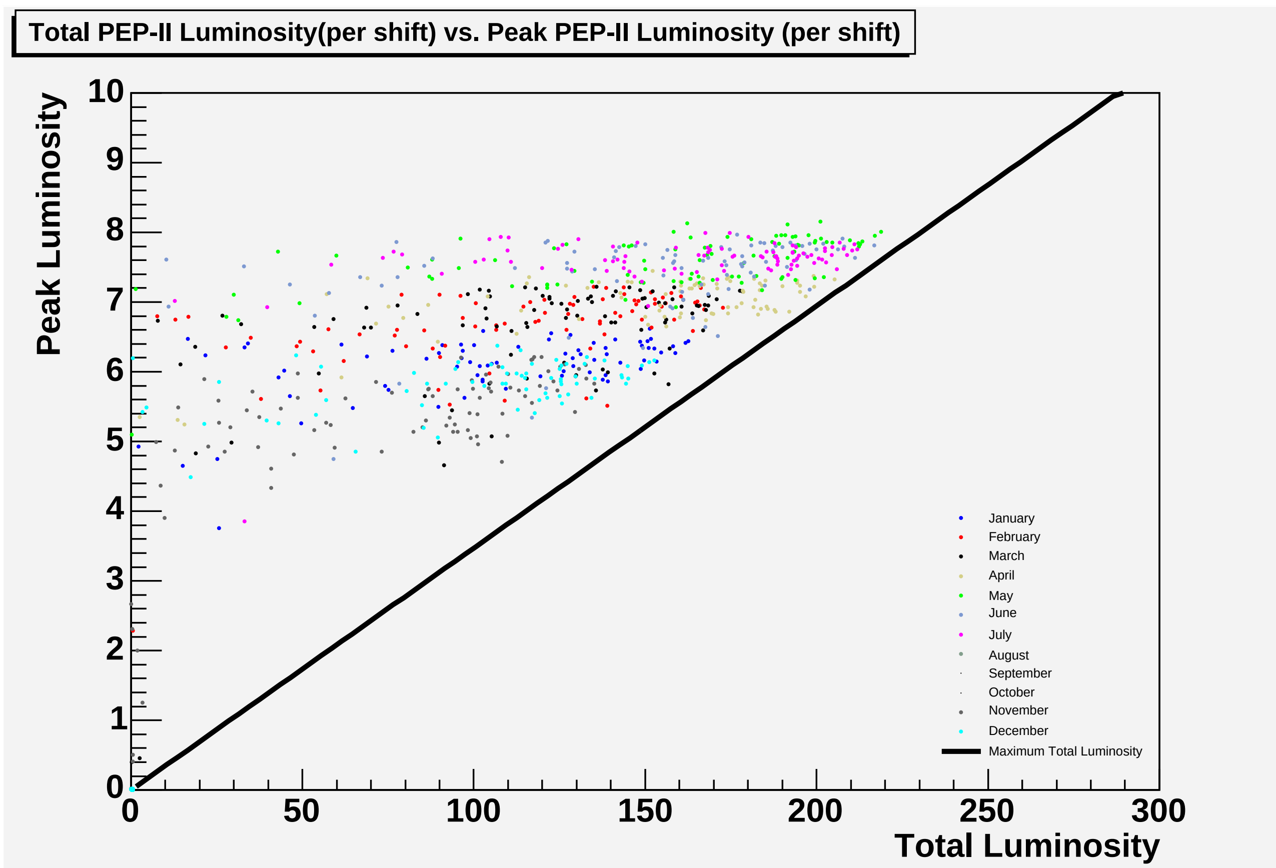

Figure 2: Run 4 data: Peak Luminosity vs. Total Luminosity per Shift

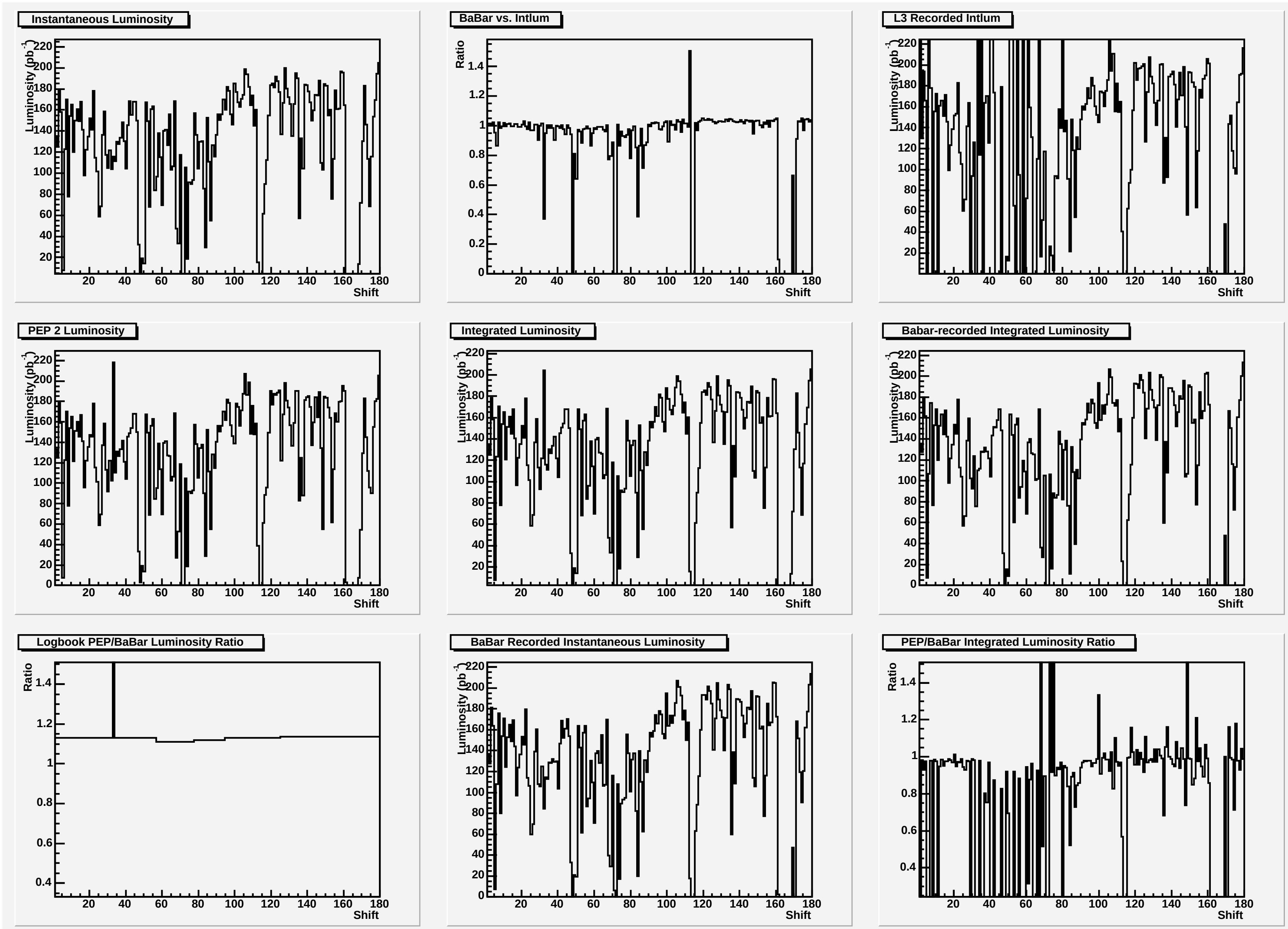

Figure 3: Nine histograms printed in one window. The data pertains to the correlations between detectors and the logbook.

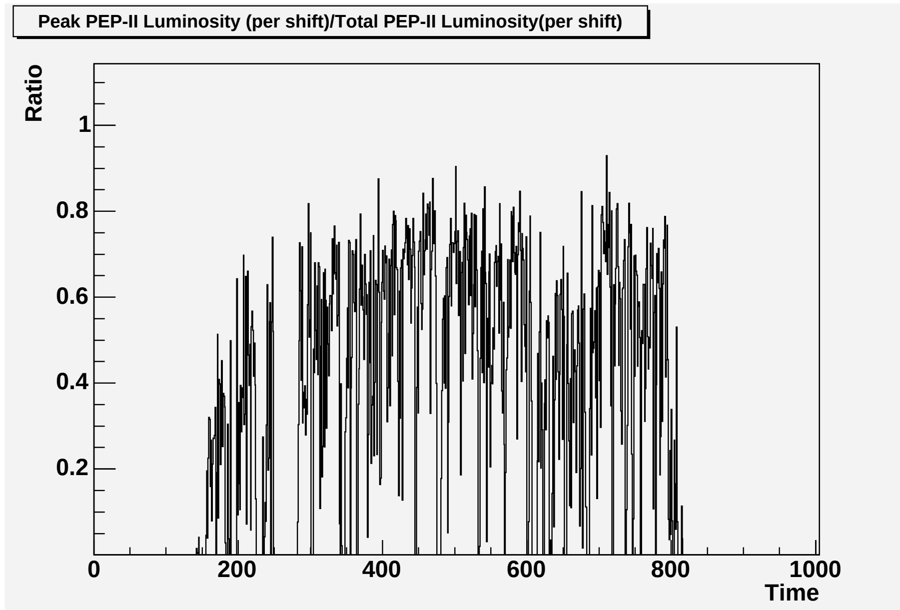

Figure 4: Run 3 data: Percentage of Ideal Luminosity

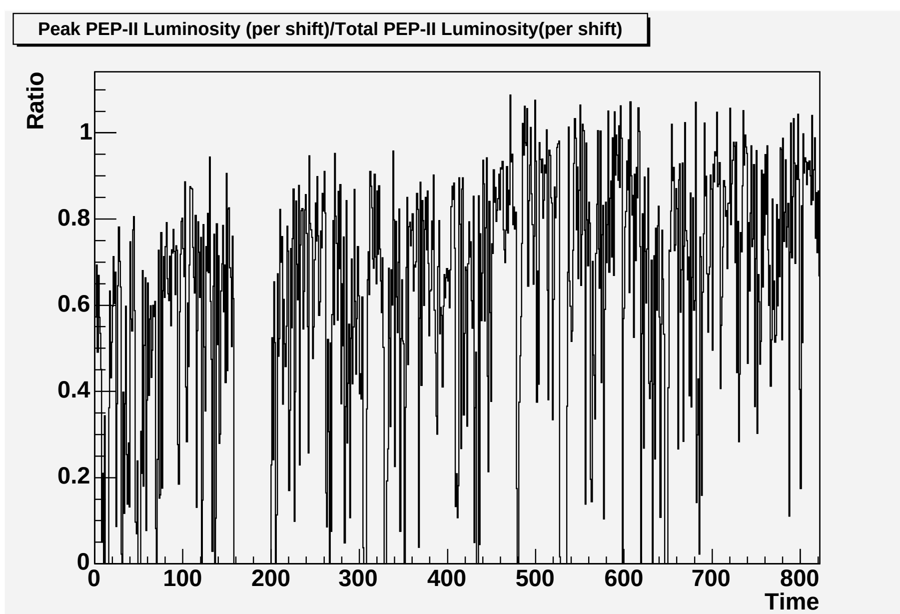

Figure 5: Run 4 data: Percentage of Ideal Luminosity