Chapter 3

The Data Analysis

3.1 Introduction

This analysis is based upon the cosmic ray induced muons seen by MACRO. The job of turning the volumes of raw streamer tube hits into muons projected back onto the sky has kept many people and computers quite busy over the years. The data used for this analysis were drawn from a long period of operation including many different detector configurations. These configurations ranged from the very first data taken by MACRO in February 1989, when only the lower half of a single supermodule was operational, to the current full MACRO. Thus, the data from different modes of operation were difficult to integrate into a coherent whole.

3.2 Data Sets Used

For ease of processing, the data were broken into nine subsets. These are divided partly based upon different detector configurations and partly into convenient sized chunks of data. The data sets are summarized in Table 2.

Dataset Name	Start Date	Start Run #	Stop Date	End Run #	Total Hours	MACRO Uptime	% Livetime	Muons Analyzed
SM1	27Feb89	68	18Nov91	4168	23,856.71	15,308.79	64.2%	5,712,208
Interim	18Nov91	4169	11Dec92	5520	9,347.90	7,182.33	76.8%	5,392,038
6 Month	11Dec92	5521	13Jun93	6329	4,407.90	3,731.83	84.7%	3,152,833
Newdat	13Jun93	6330	27Jan94	6977	5,456.32	4,570.85	83.8%	3,589,980
94_1a	27Jan94	6878	29Apr94	7472	3,058.83	1,253.61	41.0%	987,077
94_1b	29Apr94	7473	15Jul94	7907	1,846.72	1,543.25	83.6%	1,503,016
94_2	15Jul94	7908	01Jan95	9019	4,068.12	3,594.92	88.4%	3,637,927
95_1	01Jan95	9020	30Aug95	10796	5,784.08	4,756.76	82.2%	4,938,183
95_2	30Aug95	10797	01Jan96	11512	2,979.39	2,435.54	81.7%	2,514,884
Totals:	27Feb89	68	01Jan96	11512	60,805.97	44,377.87	73.0%	31,428,146

Table 1: Summary of the data used in this analysis.

The "SM1" set was data taken when the only operational scintillator was in the first supermodule. Streamer tubes in acquisition ranged from one to four supermodules, all only in the lower half of the detector. With the decommissioning of the original Supermodule One scintillator in November 1991, the “Interim” dataset started. This lasted until all six lower supermodules came on line in December 1992 with between four and six supermodules having active streamer tubes. However, there was no active scintillator.

December 1992 began a six month period when the stability of the data taking was emphasized over further construction of the detector. The "Six Month Run" featured the lower parts of all six supermodules in acquisition of high quality data with all systems active. At this period’s conclusion in June 1993, the existing detector was tuned, and construction of the "Attico", or top half of MACRO, began. This dataset was called "Newdat", because at the time that set was named, it was the newest data.

The detector construction and renovation proceeded apace in the following years, so the datasets were organized chronologically into six month chunks (94_1, 94_2, 95_1, and 95_2). The attico streamer tubes came online in the middle of the first such dataset, so 94_1 was subdivided into 94_1a and 94_1b, to account for the difference in detector acceptance. Further inclusion of data in this analysis ceased at the end of 1995, in order to write this document.

3.3 Data Handling

The first pass analysis of the data for muon tracks is performed automatically. After a data run (usually lasting four to eight hours) is finished, the newly created output data file is copied over the network from the acquisition computer in the tunnel to a disk farm in the external laboratory. A program (based upon the DREAM general MACRO data reduction package) is then run to do the basic track reconstruction of the event. This program writes the muon events to a summary file, called the Muon Astronomy Data Summary Tape (“MuADST" or "DST" for short). These files contain the reconstructed tracking data, as well as the raw streamer tube and scintillator hit information. Such summary files are much smaller (by a factor of ten) than the complete MACRO data files because the large, cumbersome data such as monopole waveforms and gravitational collapse triggers are excluded. To further reduce the size of the DST's, they are written as FORTRAN unformatted files. This avoids invoking the hideous overhead of the ZEBRA banks used in the standard data format.

The complete MACRO dataset was also spun off to magnetic tapes in both TA90 and Exabyte formats. These tapes were copied and sent to all the MACRO institutions. This analysis used these tapes for two secondary purposes. First, sections of the data stream missing from the automated DST-making procedure were reconstructed. Second, selected sections of the data were examined in greater detail for debugging and cut selection purposes.

While the DST files are far less cumbersome than the complete data files, they are still far too large for more than a few months of data to reside on disk. Thus, a second set of summary files was made. These files contain only the events likely to be used in this analysis. To further reduce their size, only timestamps, multiplicity information and simple tracking information (the muon arrival angles and the hit streamer tube planes) were recorded. This minimal summary allowed the complete data set to be kept on disk rather than archived to tape, thereby enabling the comparatively easy processing and reprocessing of the data so necessary for good scientific work.

3.4 Cuts

The data were subjected to several cuts to filter out periods of bad data taking and instrumental noise. The zeroth cut, of course, was that the event had to be a fast particle track with at least one reconstructed track in both the strip and wire views.

Any other sort of event is simply not useful to an analysis looking for the arrival direction of the muons.

3.4.1 **Event Cuts**

The following cuts were used to eliminate a small number of complicated events. These were so difficult to reconstruct that there were serious doubts about the quality of the reconstructed track. Events were cut that had an excessive number of out-of-track hits (≥ 200) or reconstructed muon multiplicities differing by more than two in each view. Events with zenith angles of greater than 72° were cut too, since these few events were nearly parallel to the ten horizontal streamer tube planes doing the bulk of the tracking. This made for larger errors in the position of the muon hits upon those planes. Since a precise timestamp on the muon event is needed to reconstruct the arrival direction of the cosmic ray on the celestial sphere, events with bad timestamp data (mislatches of the atomic clock) were also cut.

3.4.2 **Run Cuts**

In addition to cutting individual events, some whole data runs were cut. This was done by examining the muon rate for each run. If a run rate was too high, this generally indicated a run in which a lot of instrumental noise produced false triggers. An excessively low run rate either was due to the detector's inefficiency for a time, or to the temporary death of part of the acquisition system. An additional cause of strange run rates were due to atomic clock problems that were not caught by the clock's self-check. Such problems produced odd run start and stop times, which lead to abnormal run length calculations, and thus bad rates.

To eliminate data runs with these problems, the rates of all the runs were calculated and compiled for each dataset. The mean rates (per active supermodule) and standard deviations of these mean rates were computed. Runs with a rate that fell three standard deviations from the mean were excluded from this analysis.

Finally, several runs were cut by hand. Run 3192 from the SM1 set had streamer tube noise, caused by gas system and high voltage problems, which was not at a high enough rate to be caught by the rate cut. However, that noise produced 230 noise events out of 892 total, that all resulted in muons reconstructed to come from the same point on the sky. Run 3730 had a similar problem on a smaller scale, but closer investigation of this run showed a bad combination of gas system problems, streamer tube noise and DAS problems. The dead time caused by the DAS problems offset the rate increases caused by the streamer tube noise, so that the overall run rate was within tolerances. Thus, runs 3727 through 3747 were cut to eliminate this noise.

3.5 Monte Carlo Backgrounds

To search for astronomical signals in MACRO's muon data, one needs the expected background for a given time and location on the sky. This background is the flux of muons expected from the known cosmic ray flux rather than an astronomical source. High energy cosmic rays arrive at the earth at random times from any direction.

However, there are many systematic factors to include in order to accurately calculate the background. MACRO's physical configuration has changed over time, as has its efficiency. The data collection has many gaps in it. These gaps were either small ones due to the end of one data run and the start of a new one, or large gaps due to maintenance and hardware failures. Convolved with these effects is the overall effect of the non-uniform rock overburden. The absolute flux from any given direction is strongly dependent upon the shape and composition of the mountain in that particular direction[1].

Luckily, the details of all these complicated effects do not need to be known to compute a background. The known space and time distributions can be combined using a Monte Carlo technique if the two distributions are assumed to be independant of each other. The distribution of observed muons in altitude and azimuth is known from the data itself. This directly includes the mountain information. Since this distribution is known for each geometrical change to MACRO itself, such changes in detector configuration can be accounted for. The arrival time distribution of the muons has been shown to be Poissonian[2]. The mean rate of muons is well known for each data run, thus downtime and overall efficiency can be explicitly accounted for by simulating each run individually.

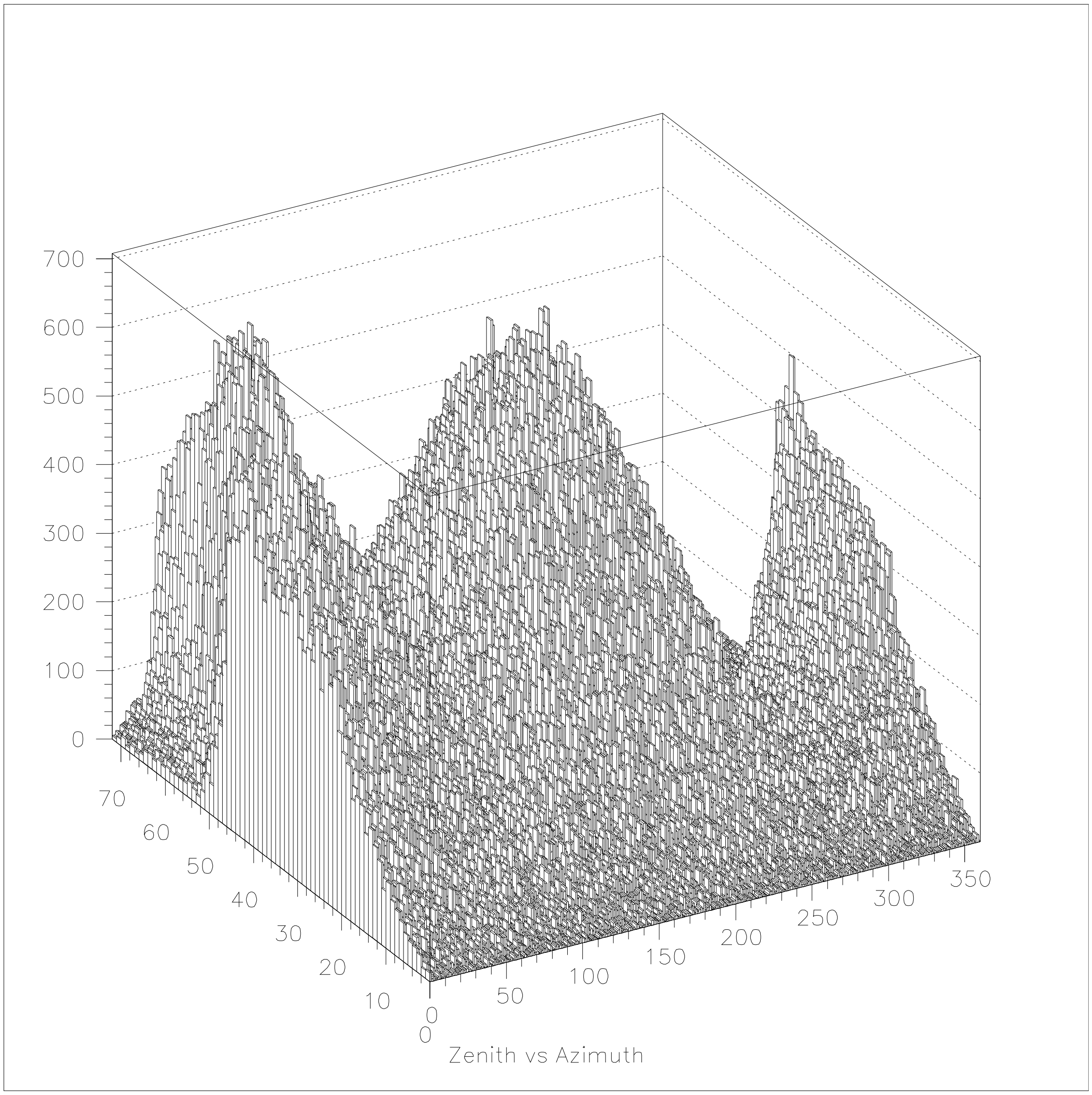

Figure 1: The Zenith and Azimuthal angle distribution for muons in the “95_1" dataset.

Since the arrival direction and arrival time distributions of muons are well-known, the question of "What would the data look like if the muons arrived randomly (and thus, if the sky was sourceless) can be answered. For each data run, the start time of the run is known. The mean muon rate for that run is also known. For each muon in the run, an elapsed time from the last muon can be drawn from a Poissonian distribution[3], and a simulated arrival time calculated. This produces a time sequence for the simulated events that explicitly accounts for down time and overall efficiency. For each such event, an arrival direction for the muon is drawn from the two dimensional distribution of event arrival directions appropriate for that data run using CERN's HRNDM2[4] routine. Given an arrival direction and time, the position of the simulated muon in celestial coordinates is directly computable.

Each run in the real data is simulated in this manner. The simulated muons are then subjected to the same analysis as the real muons. However, the statistical fluctuations in the simulated data can be suppressed by averaging over many simulated data sets. The muon count for each point in the sky, for any given period, has been computed under the assumption of an sourceless muon sky using this technique. When the real data are compared to this background, any differences in the real data from the background will be due either to a statistical fluctuation in the data or an astronomical source. Example slices in right ascension and sin(declination) comparing the data and background are shown in Figure 2.

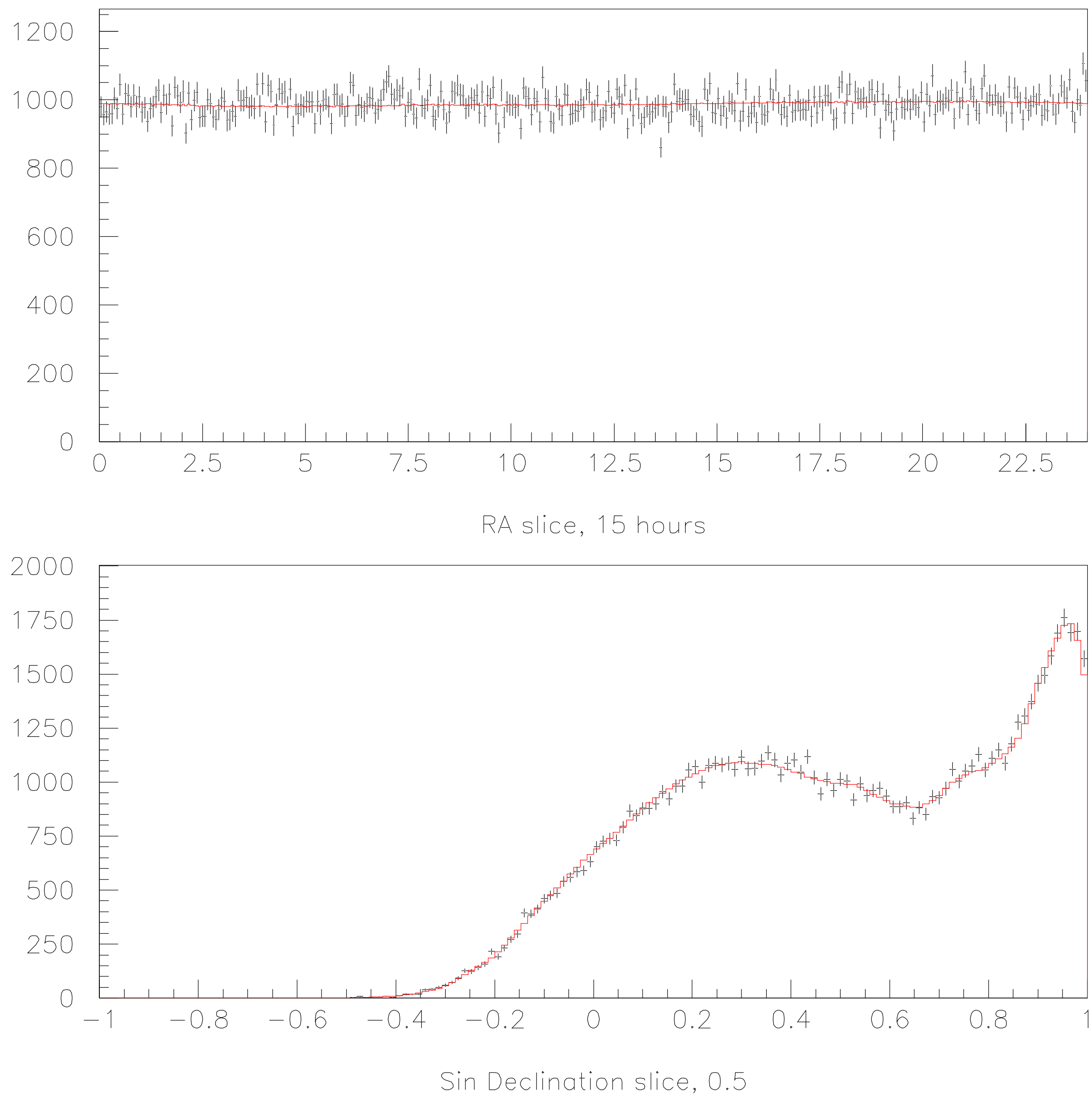

Figure 2: Slices of the data in RA and Declination, with statistical error bars, plotted over the simulated background (lines).

References

1. Ambrosio, M. *et al.*, The MACRO Collaboration, 1995, “Vertical Muon Intensity Measured with MACRO at the Gran Sasso Laboratory”, *Physical Review D*, **52**, 3793.

2. Ahlen, S. *et al.*, The MACRO Collaboration, 1992, “Arrival Time Distributions of Very High Energy Cosmic Ray Muons in MACRO”, *Nucl. Phys. B*, **370**, 432.

3. Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T., **Numerical Recipes: The Art of Scientific Computing**. Cambridge University Press, Cambridge, 1986. pp 206-208.

4. Brun, R. *et al.*, 1987, **HBOOK - User Guide**. CERN program library, Geneva. p.54.