MINOS-doc-1273
(CC position paper #5)
D. A. Petyt
University of Minnesota
December 31, 2005

Checks on FD data prior to box-opening

Abstract

This note describes the suggested validation procedures for Far Detector beam neutrino data that should be performed prior to box-opening. A summary of the box-opening crtreria and the conclusions of this paper can be found in sections 4.9, 5 and 6.

1 Introduction

This paper is intended to outline a series of checks that should be performed to ensure the integrity and quality of the Far detector spill neutrino dataset prior to box-opening. I interpret my charge as covering a fairly wide range of topics - including low-level checks of detector performance to eliminate "bad" runs from the analysis dataset, checks to ensure that the required timing and beam monitoring information are present for all the spills used in the analysis, verification of the selection procedure, and finally, tests of event reconstruction and data/Monte Carlo comparisions of the selected neutrino candidates. Obviously, these cannot all be covered in depth in this document, and in many cases I refer to more detailed work that is already being performed by others on these topics. Nevertheless, this document is intended to summarise the important tests and checks in order to provide a clean and well-validated dataset for the initial CC measurement.

The tests include:

- basic data quality checks to exclude periods of "bad data"
- checks to ensure that FD neutrino events are selected with high efficiency and negligible background contamination
- data/MC comparisons of basic distributions for the open FD dataset
- basic sanity checks on the unblinded data set

The data taking period for the purposes of this note is taken to be from the start of LE-10 running (May 20th 2005, run 31720) to the closing of the initial dataset, which I have defined as the end of run 33277 on December 6th 2005.

2 Data quality checks

2.1 FD livetime and detector checks

The far detector livetime is one of the important components for calculating the total exposure of the LE-10 dataset. In addition, it is important to know this quantity as a function of time, in order to calculate the integrated protons on target that will be used to normalise the Monte Carlo distributions for data/MC comparisons, and also as input to the oscillation fit.

The far detector livetime is calculated as the fraction of time that the FD is live and taking regular physics data. Inefficiencies include time spent in test and calibration runs, scheduled detector maintenance and unscheduled detector downtime. Scheduled maintenance is wherever possible correlated with beam-off periods in order to maximise beam-on livetime. The most relevant quantity for physics analyses is therefore the pot weighted livetime. This is calculated by integrating the total pot recorded by the beam monitoring devices over the time period when the far detector was live. Figure 1 shows, as a function of time, the fraction of POT lost due to far detector downtime. This quantity is generally much less than 1%, and the pot-weighted FD livetime for this period (May 20th to December 4th) is calculated to be 99.53%. Note that this number does not include inefficiencies due to data quality cuts (discussed below). These will be factored into the livetime calculation in section 2.3.1.

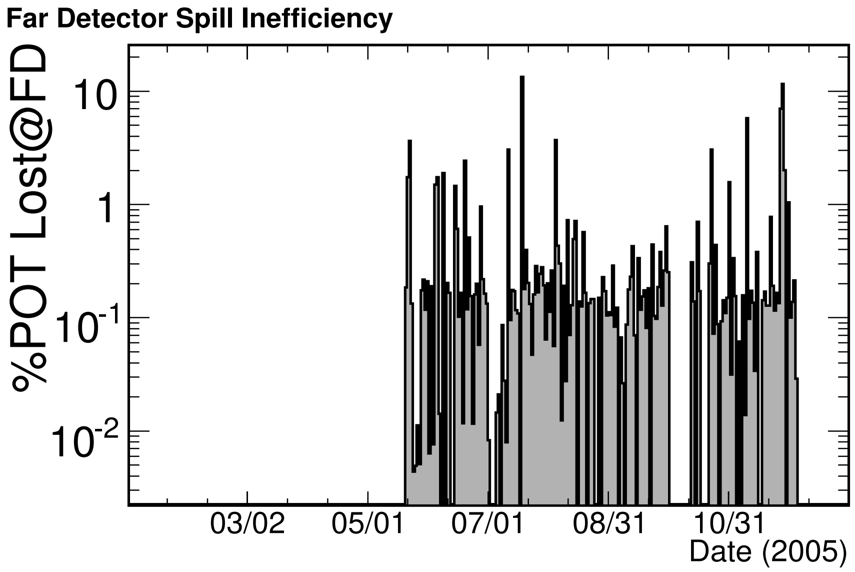


Figure 1: Fraction of pot lost due to FD downtime for the LE-10 datataking period May 20 - December 4th 2005. The average pot-weighted inefficiency for this period is 0.5%.

In addition to selecting time periods when the far detector is live, various quality checks should be performed to ensure that only periods of good quality data are considered in the final analysis. Here "good quality data" is defined by the following criteria:

1. Far detector in "Normal" physics run. Runs with the "modified" bit set are OK (there was a period in September where all Physics runs had the modified bit set, and these were verified to be OK) but those with the "test" bit set (indicating DAQ or hardware tests) should be excluded.

2. SM1 and SM2 magnets on Here "on" is defined as 80 A $\pm$ 0.2 A.

3. HV on, as defined by the singles rates from the raw data.

4. No significant LI leakage in the raw data

5. Guaranteed timing system accuracy (see next section)

For the initial analysis it is proposed that periods of "bad" data should be excluded as follows:

- The start and end times of each instance are identified and an exclusion list is built up.

- The timestamp (in seconds) of each neutrino candidate is compared to this list and the snarl is rejected if the snarl time coincides with any of these periods of bad data quality.

This procedure is documented in more detail in [1], which also describes the C function that has been written to perform this check and is available for general use.

In order to illustrate the kind of detector problems I am referring to, I show two examples below. The first case, which is illustrated in figure 2 shows a trip that occured in both supermodule 1 and supermodule 2 magnet power supplies. The data is derived from raw DCS data and shows that the coil current dropped to zero for a period of $\sim$ 1 hour before the problem was fixed. For this particular problem (which would obviously result in a bias in the energy measurement of any CC events that occurred during this time) the data from at most 2 subruns would be excluded. Note that the second drop in magnet current was part of a controlled test which took place during beam downtime. A more detailed investigation of magnet downtime is given in Appendix A.

The second example, shown in figure 3 shows a short duration drop in the singles rate in crate 1 (and also in crates 3,5 and 7) that was caused by trips in high voltage Mainframe 4 that occurred in late August and early September. This plot was created from the RawTPSinglesSummaryBlocks that are written out with the raw data. These trips were infrequent (24 in total) and were typically of 2-4 seconds duration. One side-effect of a HV trip in crate 1 is that the trigger PMT (which is located in crate 1) will not be read out during this period and therefore there is a possibility of LI leakage in the data. However, given that these time periods are excluded from the data, this problem should not occur.

The general problem of LI leakage has been investigated from scanning the spill data during the LE-10 data-taking period to date. The vast majority of LI events are accompanied by a trigger PMT hit and are thus excluded. For LI events that do not have a trigger PMT hit associated with them (5 events were observed in a scan of all spill triggers from May 20 to Nov 15), toplogical cuts that are very efficient ($>$ 99.9%) at LI removal

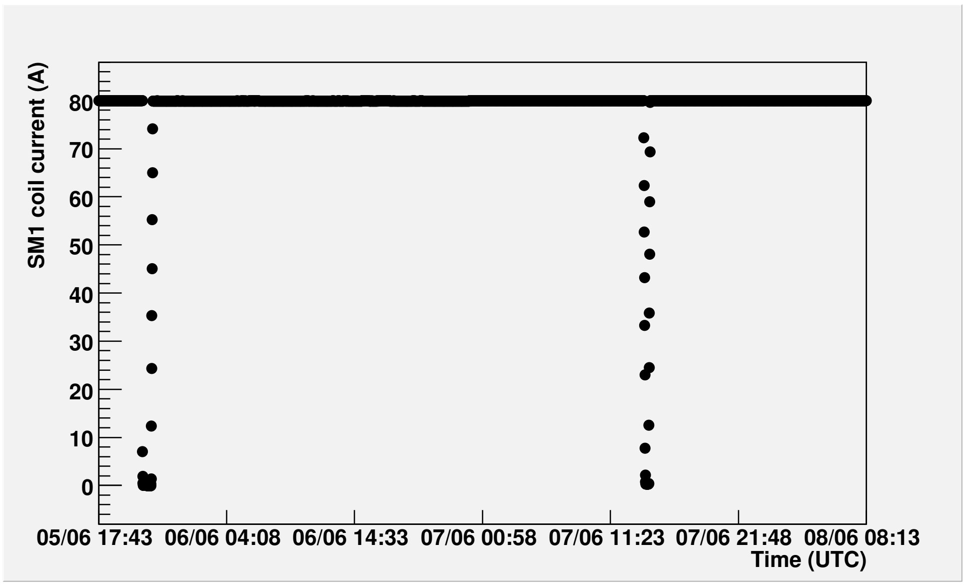


Figure 2: Magnet current (in amperes) for FD supermodule 1 as reported by the DCS system for early June 2005. The drop observed on June 5th was caused by a magnet trip. The drop on June 7th was a controlled test taken during beam downtime..

are applied [1] and so the expected contamination of LI events in the data is therefore negligible.

2.2 Timing and triggering checks

The important check to make here is to ensure that all the beam spills are present in the far detector spill data files. The SpillServer process should ensure that there is one spill-tagged event read out in the FD for each beam spill. Even when the SpillServer is down, it has been pointed out [2] that this has a very small effect on the neutrino selection effieiency, since only the lowest energy NC events that fail the E4 trigger will not be read out in this case.

The first number to check is the spill server efficiency, which is defined as the number of spill trigger events in the FD (trigger source 131072) divided by the total number of beam spills (extracted from the SpillTime database) for the time periods that the FD was live. This number has been calculated for the period May 20th-December 6th, and efficiencies computed on a month-by month basis are shown in table 1. The SpillServer efficiency for this period is calculated to be 95.6%, where the major source of inefficiency is the period June 10 - June 20 when the SpillServer was down. The efficiency for this period exculding June is 99.4%. Figure 4 shows the SpillServer inefficiency as a function of time.

Given this inefficiency, it follows that the total number of "lost spills" can be at most 4.4%. If the inefficiency is weighted by the beam intensity, this number will be slightly smaller ($\sim 3\%$) since the beam intensity was lower in early June than it is now. Of course, we expect a negligible loss of events, since beam-coincident events in the FD without the spill trigger bit set will be tagged by the Spill time database, as explained below.

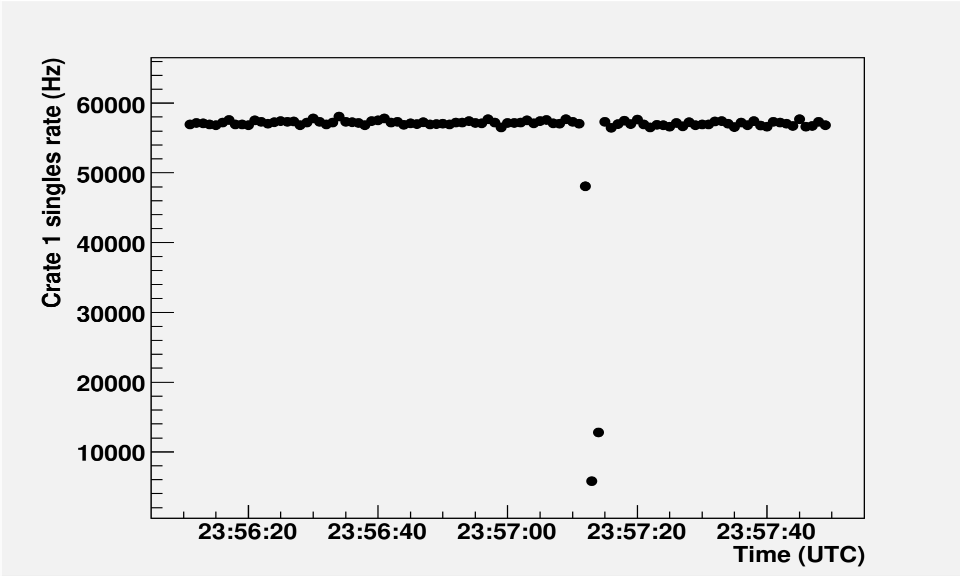


Figure 3: Singles rates (in Hz) for FD crate 1 for run 32263 (September 1, 2005). This shows a drop in the singles rate for a period of 2-3 seconds caused by a trip of high voltage Mainframe 4.

Month	ND spills	FD spills served	SpillServer Efficiency (%)
May	284436	283879	99.8
June	694576	498450	71.8
July	752710	741857	98.6
August	894977	886001	99.0
September	925201	920380	99.5
October	511131	509994	99.8
November	876000	875160	99,9
December	200174	199728	99.8

Table 1: ND and FD spill events during FD physics runs, May 20 - December 6 2005

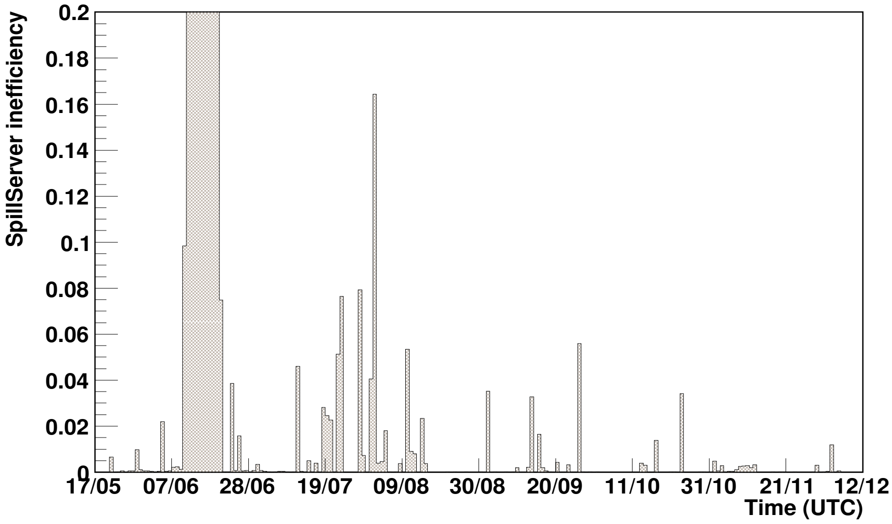


Figure 4: SpillServer inefficiency as a function of time. Only data from FD physics runs are considered. The inefficiency from June 10 - 20 is due to a known period of SpillServer downtime.

When the spillserver is down, the ND spill times in the database are used in the batch processing to tag FD events that are coincident in time with beam spills and write them to the spill stream. This process of course relies on the database containing a complete record of the beam spills. This can be checked by comparing the spill times in the database to the timestamps of SGATE snarls in the near detector. For the months of June and July, where the spillserver efficiency was below 99% the following results were obtained:

- June: 2/689476 SGATE spills had no corresponding entry in the spill database
- July: 1/737636 SGATE spill had no corresponding entry in the database

From these numbers it is clear that the integrity of the spill time database is high.

The proof that the process of tagging FD events that are coincident with the beam using the database must be made using the data itself, by checking that the number of events are observed as a function of time is a constant and does not depend on whether the spill server is up or down. Figure 9 in the next section is intended to demonstrate this, and shows that the rate of neutrino candidates observed during the period June 10 - June 20 is consistent with the total integrated pot taken during this period..

Another important check is to verify the accuracy of the timestamps that are applied to the spill data. One way of checking this is to monitor the timing accuracy reported by the GPS timing system (either by using monitoring data from the TimeGoblin process directly or by using the RawSpillServerMonitorBlocks that are written to the raw data output stream). A set of plots using the latter method are shown in figures 5 and 6. Figure 5 shows, for the total LE-10 dataset, the total number of FD spillserver monitoring blocks written out as function of time (these are reported once per second). Also shown are those monitoring blocks where the GPS system reports a worst case timing uncertainty of greater than 1 μs (the nominal uncertainty is 200 ns).

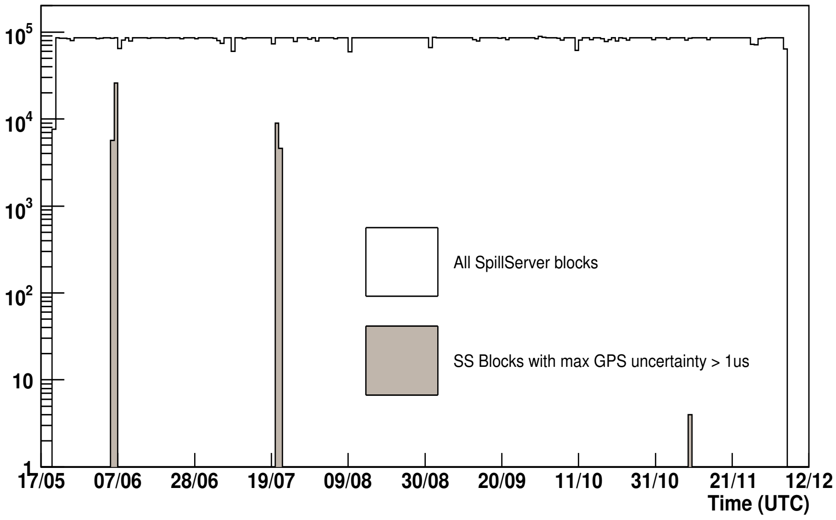


Figure 5: SpillServer monitoring blocks recorded during May 20 - December 6 2005. Those blocks that have a GPS worst case timing uncertainty of $> 1\mu$s are shown in grey .

There is a period around June 5-6th where the GPS uncertainty was larger that this value, and this is correlated with a known GPS hardware problem (see CRL entry 41298). The 21 physics runs and subruns that are affected by this problem are shown in grey in figure 6[1]

There is a second period where the GPS worst case uncertainty exceeds 1μs. This occured at Wednesday 20th July at 18:29:42 (UTC) and lasted for 9 hours. This particular problem is believed to be due to a failure of the ND GPS timing unit during an electrical storm (see CRL entry 42990). This affects the data from 10 subruns (run 32212 subrun 7, run 32215 subruns 0-7 and run 32216 subrun 0).

Finally, there is a third short period in run 33119 subrun 9 where the GPS uncertainty was large. In this instance, the TimeGoblin process went into error and stopped the run. The period in question starts on Thursday November 10th at 03:06:20 (UTC) and lasts for 9 seconds.

So far, these are the only known periods where timing inaccuracies of this level have been observed. In total 32 subruns are affected, and the time periods within these runs where the GPS timing system uncertainty exceeds 1μs should be excluded from the final dataset.

[1] For this particular case, given that the maximum uncertainty recorded (1.2 ms) was large with respect to the size of the spill trigger window (meaning that many of the beam events recorded during this time period will not be present in the spill files), it is recommended that data taken during these subruns not be used for the initial analysis [3]. This data could be rescued at a later date by searching through the "all" (or cosmic) reconstructed data files for this time period.

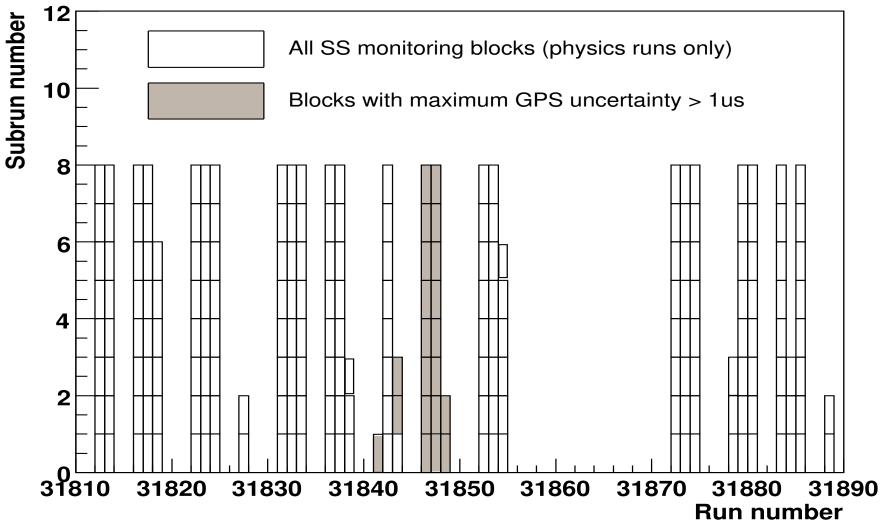


Figure 6: Run and subrun numbers of physics runs taken during early June 2005. Runs that contain spills with GPS timing uncertainties $> 1\mu$s are shown in grey.

2.3 Beam monitoring information

It is necessary to use the beam monitoring information to select data that is taken under stable and well-understood beam conditions, and also to provide the pot normalisation for the data/MC comparison plots, and ultimately the oscillation fit itself. Note that a thorough description of beam monitoring related efficiencies and accuracies will be provided in a separate note for the 1e20 measurement [4], and that this topic was also covered in detail at the October collaboration meeting [5] and so I will only briefly summarise things here.

The beam monitoring information in both the standard ntuples and the offline database is provided by the Beam Data Process (BDP), which is calculated (from comparison to raw beam monitoring data) to have an efficiency of $> 99\%$ for the time period May 1 - Sept 26th [5]. One simple way of monitoring the efficiency of the BDP is by comparing the timestamps of SGATE snarls in the near detector with the nearest beam monitoring spill provided by the database (which is populated by data taken by the BDP). The database will always return the closest spill in time, so it is necessary to check whether the beam information is coincident (within 0.5 secs) of the SGATE snarl. If this is not the case, then there is no beam information for that particular spill. Figure 7 shows this comparison for June through October 2005 data. It can be seen that the rate of "missing spills" is generally low $< 1\%$, except for certain short periods when the BDP is down. The BDP inefficiency for this period (in terms of spills, not pot) is $\sim 2\%$, the major components are one particular period in mid-June when the BDP was down, a similar period in early-September and a period in mid-October where there were problems with duplicate spill information appearing in the raw beam monitoring files. Other months show much lower inefficiency rates [5].

In order to select data taken under stable and well-understood beam conditions, the

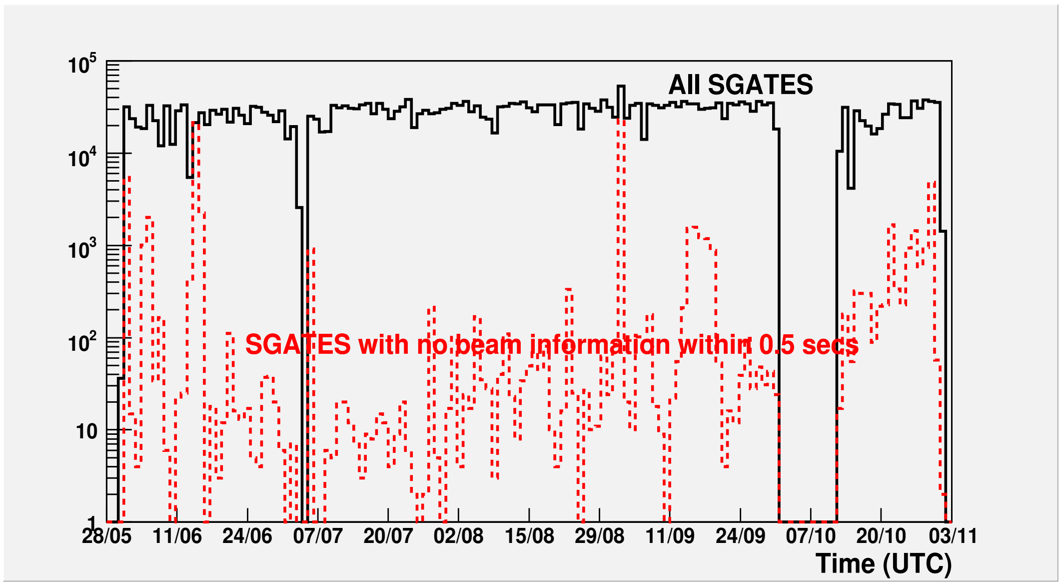


Figure 7: Timestamps of near detector SGATE snarls for June 1 - Oct 31 2005. The timestamps of the snarls that do not have time-coincident beam monitoring information are shown by the red-dotted line.

standard beam monitoring cuts will be applied as follows:

1. maximum time difference between spill and nearest beam monitoring data < 1 second
2. $0.5 \times 10^{12} <$ tortgt $< 50 \times 10^{12}$ protons on target
3. -200 kA $<$ horn current < 155 kA
4. X-position of beam at target: $-2.0 < x < 0.0$ mm
5. Y-position of beam at target: $0 < y < 2.0$ mm
6. Horizontal width of beam at last profile monitor before target: $0.1 < x_{RMS} < 1.5$ mm
7. Vertical width of beam at last profile monitor before target: $0.1 < y_{RMS} < 1.5$ mm

Note that the lower limit on the horn current has been lowered from -163 kA to -155 kA to accommodate the special horn current runs that were taken in late-July and early August. The inefficiency of each of these cuts is detailed in [5].

Applying these quailty cuts to the 165 neutrino candidates that have been isolated in the "open" sample results in the loss of five events (3%). All events are judged to be good neutrino candidates from scanning. Of these, two are due to BDP inefficiency, one appears to be due to tortgt readout errors and the cause of the other two is under investigation. Details of these events can be found in Appendix B.

2.3.1 Total POT normalisation

Once all the above inefficiencies (FD livetime, FD data quality and timing cuts and beam quality cuts) are known, the total calibrated POT for the LE-10 dataset can be calculated and the results for the LE-10 dataset are shown in table 2 and figure 8. The total POT for this period is calculated to be 9.303×10^{19}.

	Total POT	Percentage of total
Total	9.408×10^{19}	100
FD live (physics run)	9.303×10^{19}	98.8897
Fails data quality cuts	6.257×10^{17}	0.6650
During no run	4.554×10^{17}	0.4841
In calinject runs	8.46×10^{16}	0.0899
In pedestal runs	3.10×10^{16}	0.0330
In test runs	5.27×10^{15}	0.0056
In detector checkout runs	4.67×10^{14}	0.0005

Table 2: Breakdown of FD livetime and pot-weighted inefficiencies. Courtesy A. Habig.

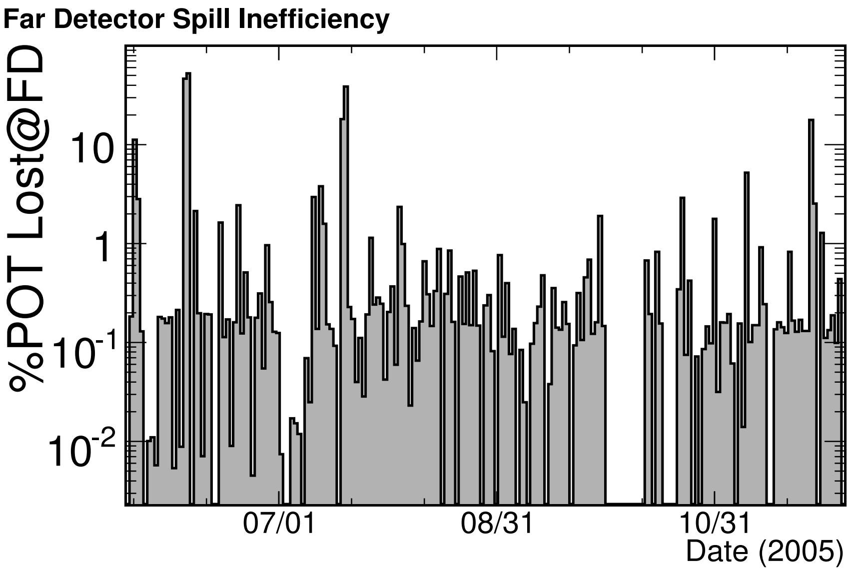


Figure 8: Fraction of pot lost due to FD downtime, with beam selection cuts and data quality cuts applied, for the LE-10 datataking period May 20 - December 6th 2005. The average pot-weighted inefficiency for this period is 1.1%.

2.4 Basic data integrity

Simple sanity checks should be performed to ensure the completeness and integrity of the FD reconstructed ntuple files that will be used for the analysis: It is expected that most (if not all) of these checks will be performed as a matter of course in the standard validation checks that will be applied by the Batch processing and Data validation groups.

- Verify that the file lists for the spill sample are complete and that there are the same number of files in both the open and blinded set. **Result of test: PASSED** - the same number of files are observed in both samples and there is no evidence of missing data (see table 3 for a breakdown of the number of existing files as a function of calendar month for both datasets).

Month	Hidden files	Open files	Open files (R1_18)
May	267	267	267
June	715	715	714
July	734	734	734
August	742	742	741
September	722	722	719
October	739	739	739
November	720	720	720
December	140	140	N/A
Total	4769	4769	4634

Table 3: Number of spill files (subruns) in the hidden and open datasets for the LE-10 data taking period (runs 31720-33277) processed with release R1_18_2. There are an indentical number of subruns in the two samples. Also shown for comparison are the number of subruns in the open dataset for the previous round of reconstruction (which has a small number of missing files and does not include data from December 2005).

- Verify that the ntuple files are readable, both in a bare ROOT session and looping through ntuple entries within a loon job. This will of course be tested thoroughly for the open data set - the same basic checks should also be performed for the blind set. **Result of test: PASSED** - all files are readable.

3 Data selection checks

In order to demonstrate that far detector beam events are selected with high efficiency and negligible background, the results of three independent selection methods are compared for the open dataset:

1. A visual scan of all spill trigger events that are coincident in time with beam spills and contain sufficient activity in the detector to form at least one reconstructed event.
2. An automated method to select neutrino candidates based on topology and timing. This method uses the same event sample as the visual scan, applies fiducial cuts on the event vertices and requires that each event contain at least one reconstructed track. The efficiency of this method for selecting neutrino events is calculated to be $\sim 63\%$ [1, 6].

3. An automated method to select neutrino candidates based on topology and timing which uses events with tracks and/or showers and does not apply fiducial cuts on the event vertices. The efficiency of this method for selecting neutrino events is calculated to be $\sim 90\%$ [7]. In order to provide a direct comparison with method 2, only fiducial events with a reconstructed track are considered here.

3.1 Comparison of scan and automated methods

A comparison between the scan and automated selections has been carried out using offline reconstruction version R1.18. The dataset considered is from run 31720 (Friday May 20th 2005) to run 33141 (Tuesday Nov 15th 2005). During this period a total of 1698157 "true" spills were recorded in the Far detector, of which 2080 produced sufficient activity to form a reconstructed event in the detector. The number of neutrino events found in this period by the various methods is as follows (note that beam and data quality cuts have not been applied to these events):

1. 227 scanned neutrino candidates
2. 137 selected neutrino candidates (method 2)
3. 142 selected neutrino candidates (method 3)

The relative efficiency of method 2 to the scan is 60% which is consistent with the 63% efficiency quoted above. The level of agreement between the two automated selections is very good. All 137 of the events tagged as neutrino candidates in method 2 are also present in the 142 events selected by method 3. This gives us confidence that the methods used to isolate neutrino candidates that contain a reconstructed track are robust.

Figure 9 shows the number of events per 10^{17} pot as a function of time for the two selections and the scan. This plot shows that there is no discernible time dependence of these quantities, which indicates that there are no obvious selection problems due to missing or incomplete data. As previously noted, there are no selection issues relating to known periods of spillserver downtime evident in this plot. Of the 2080 spill events considered, 29 (1.5%) were not tagged with the spill trigger bit, and 8 (3.5%) of the 227 neutrino candidates were similarly tagged as E4 or 4/5 triggers. This is consitent with the pot-weighted SpillServer inefficiency of $\sim 3-4\%$ which was estimated in section 2.2.

3.2 Comparison of automated methods for the full dataset

A comparison of the two automated selection algorithms has been made for the full LE-10 dataset with data reconstructed with software release R1.18.2. A check of these results against a visual scan of the events is in progress and will be presented at the Oxford meeting. The numbers of events selected by methods #2 and #3 are 159 and 163 events respectively. Table 4 shows the breakdown of the selection cuts for method #2. Most of the events in the initial sample of 2035 snarls are random noise in the Spill Trigger window and these are efficiently rejected by the topological cuts. The five events lost due to beam quality cuts have previously been discussed, and the single event lost due to data quality cuts corresponds to the aforementioned period in July 2005 where the GPS timing uncertainty was large.

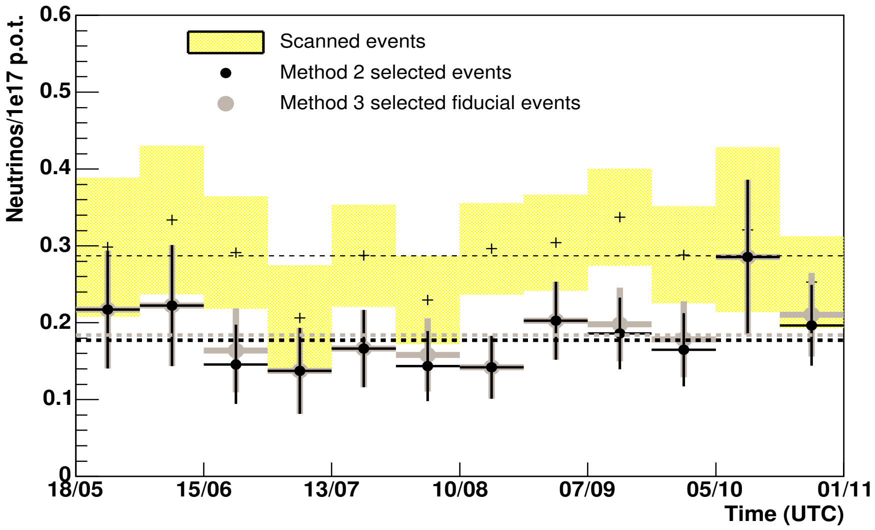
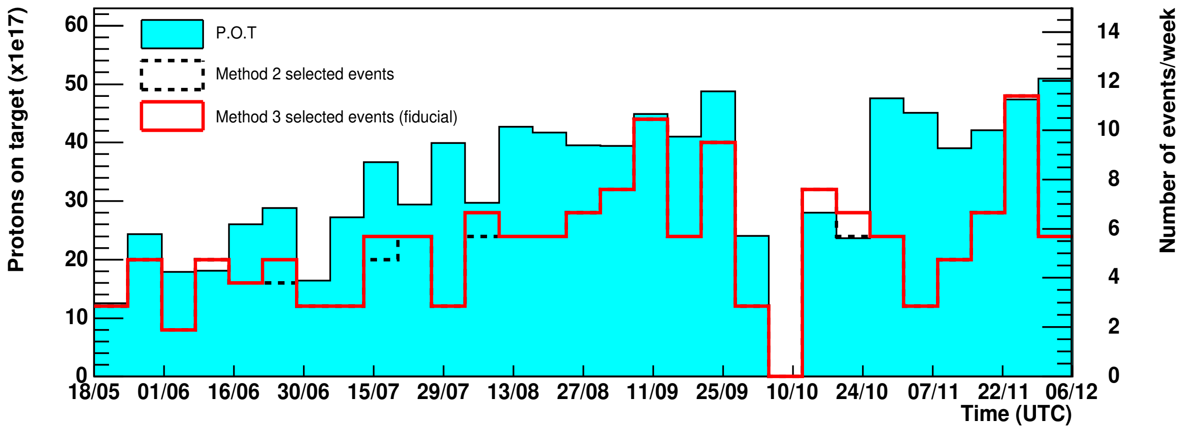


Figure 9: Number of neutrinos per 1e17 protons on target as a function of time for the data taking period May 20 - November 1. The yellow boxes show this quantity for scanned neutrino candidates, the data points are for events that pass automated selection method #2 and the grey points are for events that pass automated selection method #3 that contain a reconstructed track and have vertices within the fiducial region of the detector. The dashed lines show, for the three selections, the average number of neutrinos per 1e17 pot for the entire time period. Note that there is no significant dependence of this quantity on time.

All events	2035
Pass topological cuts	166
Pass LI filter	165
Pass timing cuts	165
Pass beam quality cuts	160
Pass data quality cuts	159

Table 4: Breakdown of cuts for automated selection method 2.

The differential and cumulative distributions of the numbers selected neutrino events using the two methods as a function of time is shown in figure 10, which also shows - for the time period May 20 to December 6 - the number of pot recorded per week and the cumulative pot. The total number of neutrinos (unblinded, no oscillations) is indicated by the right-hand axis on the bottom plot. This plot shows that both the numbers of selected events for the two methods follow the total pot closely. Figure 11 shows the number of events per 10^{17} pot as a function of time for the two selections and again shows no significant time dependence of this quantity.



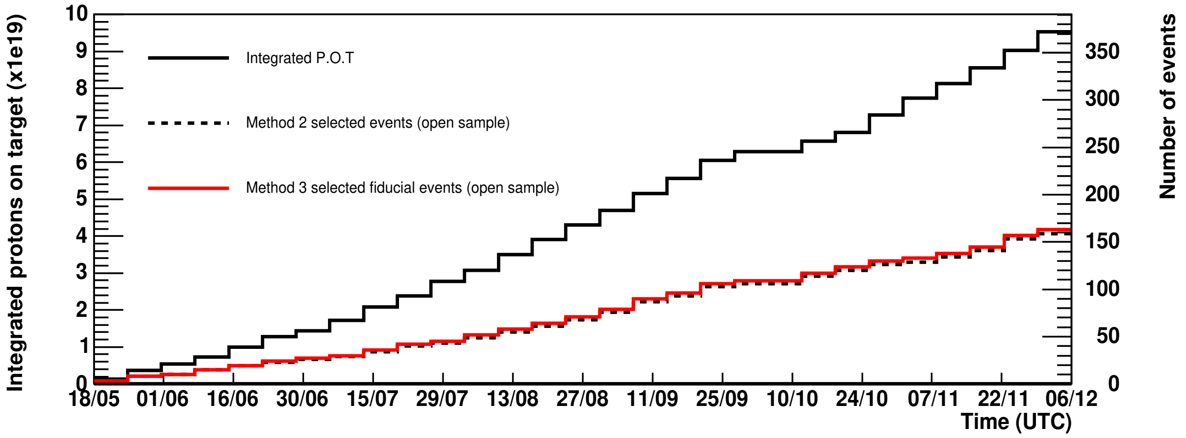


Figure 10: Neutrino selection as a function of time. Top plot - number of neutrinos recorded per week for the period May 20 to December 2005. The dashed black line shows the number of neutrino candidates identified by automated selection method #2, which has an efficiency of $\sim 63\%$. The number of fiducial events selected by method #3 is shown by the red solid line. The number of events per week is indicated by the scale on the right-hand y-axis. Also shown is the number of pot per week (tortgt calibrations applied) recorded over the same period, in units (as shown on the left-hand y-axis) of 10^{17} protons on target. The bottom plot shows the cumulative distributions of the above. The number of neutrino events expected for no oscillations and in the absence of blinding, can be determined by projecting the solid black line to the right-hand y-axis

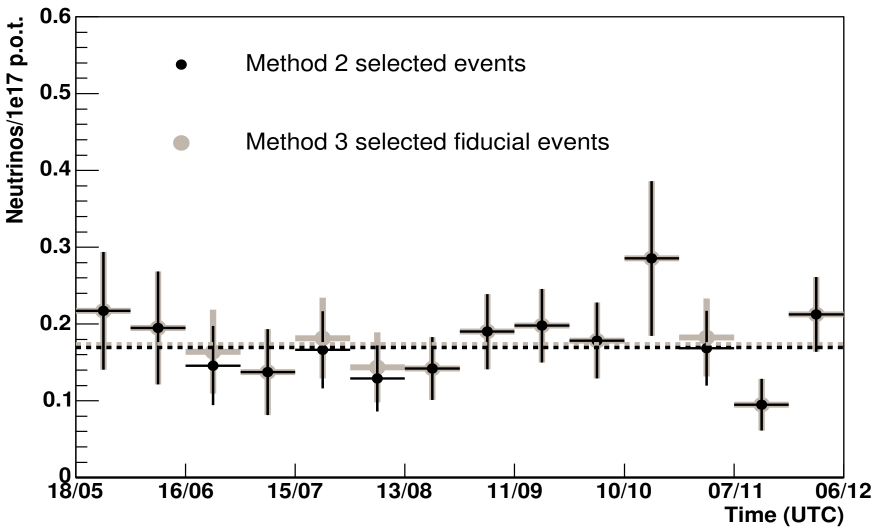


Figure 11: Number of neutrinos per 1e17 protons on target as a function of time for the open dataset. The data points are for events that pass automated selection method #2 and the grey points are for events that pass automated selection method #3 that contain a reconstructed track and have vertices within the fiducial region of the detector. The dashed lines show, for the two selections, the average number of neutrinos per 1e17 pot for the entire time period. Note that there is no significant dependence of this quantity on time.

3.3 Backgrounds

A comparison of the scan and automated analysis method 2 shows that, with the various timing, topological and fiducial cuts, the number of background events surviving the cuts is a negligible fraction of the total sample. This can be further checked by applying the same selection criteria to the "fake" spill events (spill triggers taken in anti-coincidence with beam spills). Here, fake spills from the hidden dataset were used (to increase statistics) and 2209 of these produced at least one reconstructed event (to be compared to 2035 in the open dataset). Of these events, zero passed the selection and timing cuts of both methods 2 and 3. The total number of fake spills used was slightly larger than the number of true spills in the open dataset (2650757 versus 1965980) so, given that no events pass the selection cuts in the fake sample, the background in the open sample cannot be larger than 1.7 events (at 90% C.L.) - and can be shown to be zero by a visual scan - for a signal of 159 neutrino candidates.

3.4 Event selection for the hidden dataset

The same selection methods have been applied to the hidden dataset in order to ensure that the selections perform as expected and that are no periods of missing data. Figure 12 shows the number of events per $X\times10^{17}$ pot as a function of time, where an arbitrary scale factor has been applied to hide the overall normalisation. There is no significant variation of this quantity with time.

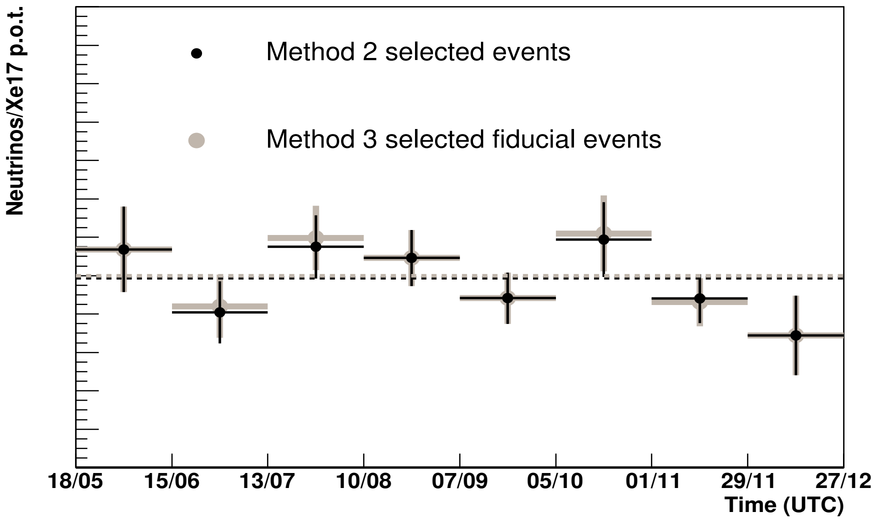


Figure 12: Number of neutrinos per Xe17 protons on target (arbitrary scale) as a function of time for the hidden dataset. Note that there is no significant dependence of this quantity on time.

4 Checks on unblinded data sample

These checks are intended to provide a verification of the data quality of the selected neutrino candidates and to compare the various reconstructed quantities with expectations from Monte Carlo events. Given that the data sample is subject to blinding and presumably oscillations, and that it is statistically limited, these checks cannot therefore be precise, but they should serve as basic sanity checks to ensure that there are no significant reconstruction and/or detector effects that would affect the data distributions. Note that the MC prediction for all plots in this section is scaled to the same number of events as the data.

- *Role of hidden dataset:* The "hidden" data sample is expected to remain closed until the box is formally opened. However, it is allowable to use this sample for basic sanity checks, and for tests of the selected neutrino events if the statistics of the open sample proves insufficient, if there are significant discrepanices observed in the open dataset that cannot be explained by blinding and oscillations, or if there are basic (non-physics) validation tests that cannot be performed because the distributions are distorted by the blinding function. I do not consider that tests involving physics quantities (such as track momentum and visible energy) should be performed on this sample, even if you confine the check to regions in phase space where oscillations should be minimised. In that case you will end up with relatively few events, which will not allow a very precise test of data/MC agreement - this kind of test is much better performed in the near detector where there are large numbers of events

4.1 Check 1: Visual scan

The first check of the selected events is to visually scan them. This will provide a detailed check of the events in terms of overall data quality, reconstruction performance and background conatamination that may not be possible by comparing distributions of small numbers of events. Both detector and reconstruction pathologies should be clearly visible here. This test should be performed with the final reconstruction pass of the entire LE-10 dataset (currently in progress), but from scanning the neutrino candidates from LE-10 running up to the end of November with $R1.18$ processing, I, and several others, have not observed any serious detector or reconstruction problems so far.

4.2 Check 2: Event timing

Figure 13 shows the distribution of the event time of far detector neutrino candidates relative to the SpillServer prediction (where the trigger time has been corrected for the time of the reconstructed event within the snarl). The purpose of this check is to verify that the selected events are in coincidence with NuMI spills and that the width of the distribution is consistent with the expected 8.9μs spill length. In this particular case, the distribution is shifted to slighly negative times, but the shift is very small in relation to the Spill trigger window of 100μs. The fact that there are no other events within the plotted 100μs time window that are not precisely correlated with beam spills indicates that the level of background in this sample is very small.

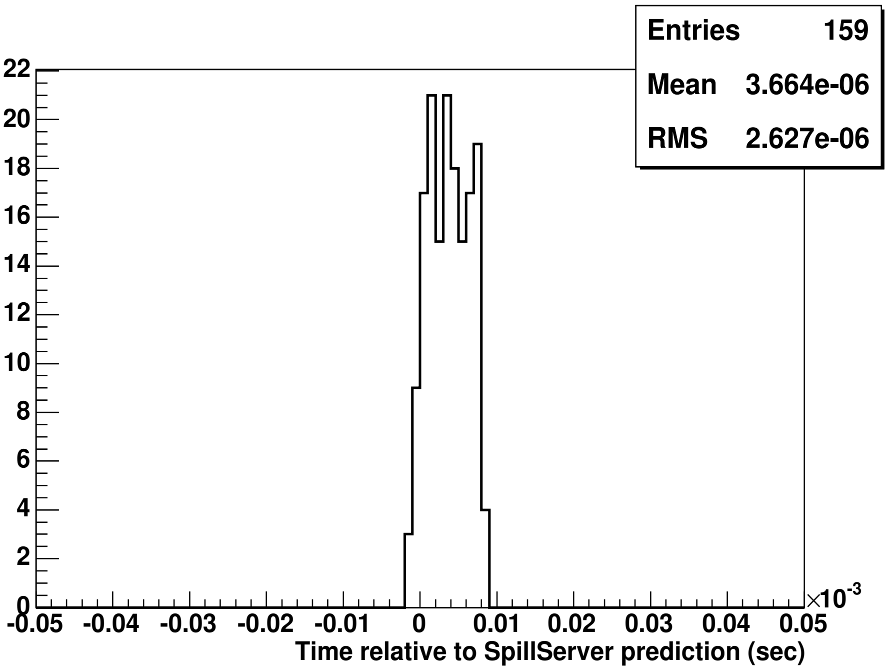


Figure 13: The time difference between the earliest hit of selected neutrino candidates relative to the spill server prediction for these events. The width of the x-axis is 100μs and the neutrino candidates are bunched in a $\sim$10 μs wide region, as expected from the time duration of NuMI beam spills

4.3 Check 3: Basic event quantities - vertices and endpoints

Comparing data and MC track vertex positions (shown in figure 14) is a very basic sanity check that does not depend at all on blindness or oscillations. The plot, and all subsequent plots in this section, shows the selected FD events as data points with error bars and the Monte Carlo distribution as the solid histogram. The MC distribution is shown for no oscillations and is scaled to the same number of events as the data. The vertex positions should be uniform in space and non-uniformities might point to either background contamination (most likely from cosmic muons) or gross reconstruction/detector pathologies.

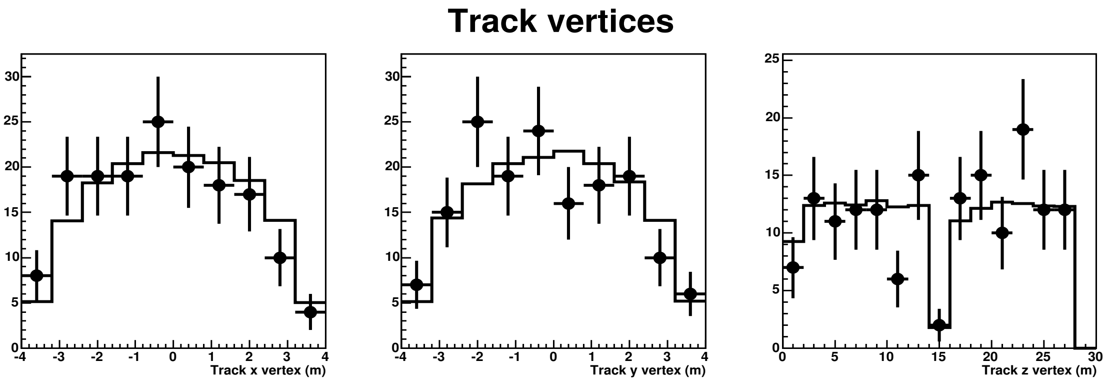


Figure 14: Distribution of the reconstructed track vertices of selected neutrino candidates (black points), compared with scaled (to the same area) distributions from LE-10 MC events.

The distributions of data and MC track end positions (shown in figure 15) test tracking efficiencies and also the modelling of the magnetic field in the MC. Negatively charged tracks ($\sim 90\%$ of the total sample) should be focussed towards the coil hole and the remaining anti-neutrino events should be defocussed. This test does not depend in particular on blindness or oscillations(there will be a weak second-order energy-dependent effect).

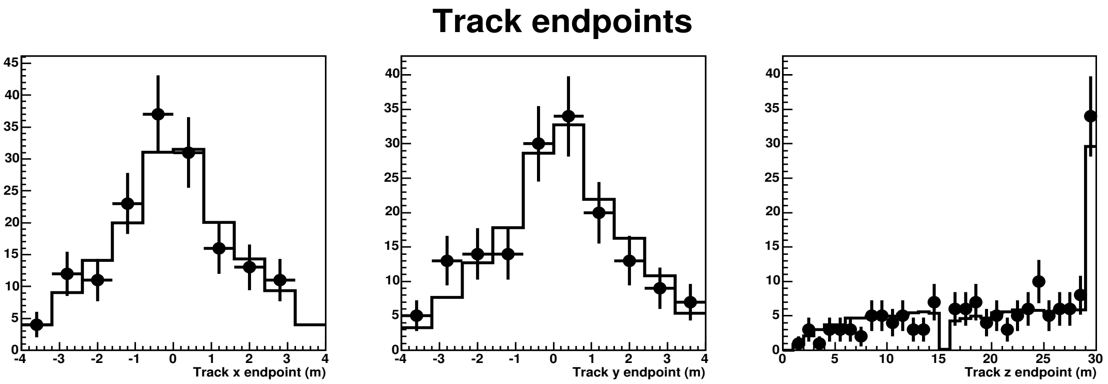


Figure 15: Distribution of the reconstructed track endpoints of selected neutrino candidates (black points), compared with scaled (to the same area) distributions from LE-10 MC events.

A third check is to examine the track vertices and endpoints in U,V space. These are shown in figure 16. Although the level of agreement between data and MC is generally

OK, there are some noticable discrepancies in the track endpoint distribution in the U co-ordinate. In particular there is an excess at -3m and deficits at ± 1 m. The value of χ^2 for this distribution is quite poor ($\chi^2 = 19.7/10, P = 0.032$). In order to test if this was due to any reconstruction or detector pathology, I scanned the candidate neutrino events with track endpoints between -3 and -2.4 m in U. I did not find any obvious problems and the events were generally well-reconstructed. In this case, I think it is reasonable to examine these distributions for the hidden dataset (see figure 17) to see if the discrepancies persist with increased statistics. The figure shows that the agreement is generally better and for the distribution in question, the value of χ^2 is improved to $13.7/10, P = 0.187$.

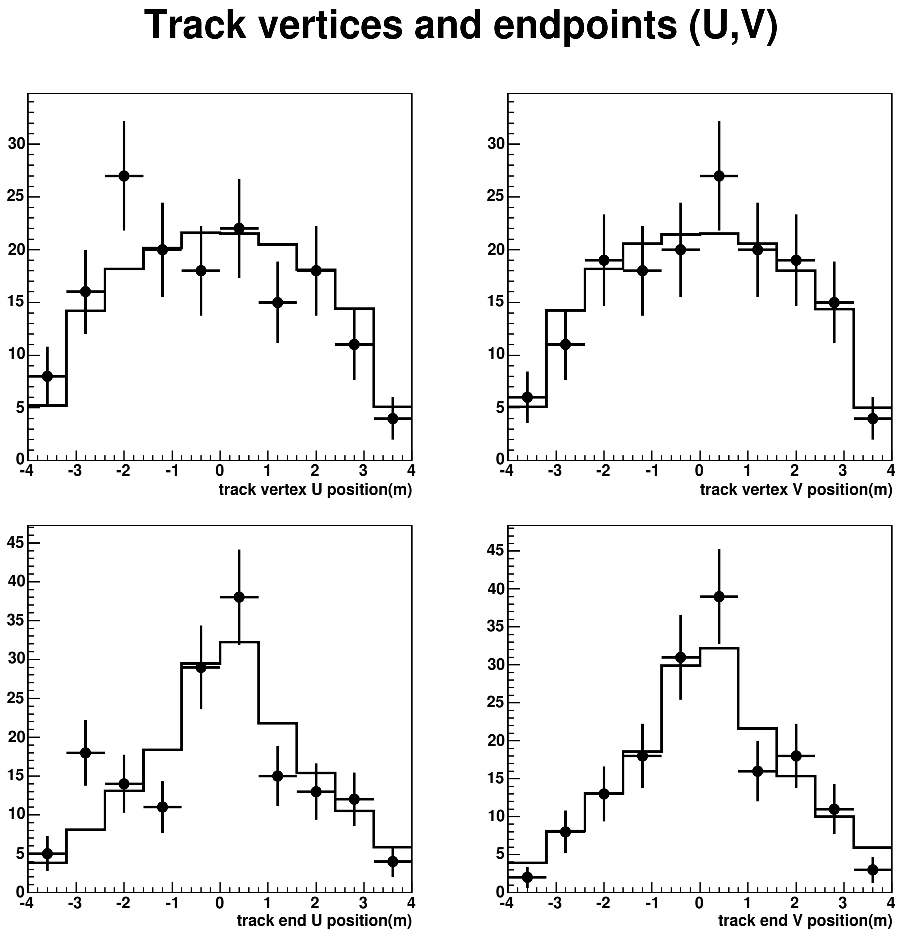


Figure 16: Distribution of the reconstructed track vertices and endpoints in U and V co-ordinates of selected neutrino candidates (black points), compared with scaled (to the same area) distributions from LE-10 MC events.

4.4 Check 4: Basic Event Distributions

Figures 18 to 22 show basic distributions of global event quantities. These plots are provided for direct comparison with figures of the same quantities in the Near Detector that are presented in [8]. The plots are: event pulse height (sigcor ADC) per plane - this tests, amongst other things, pulse height calibration; fraction of event pulse height in planes 1-3.

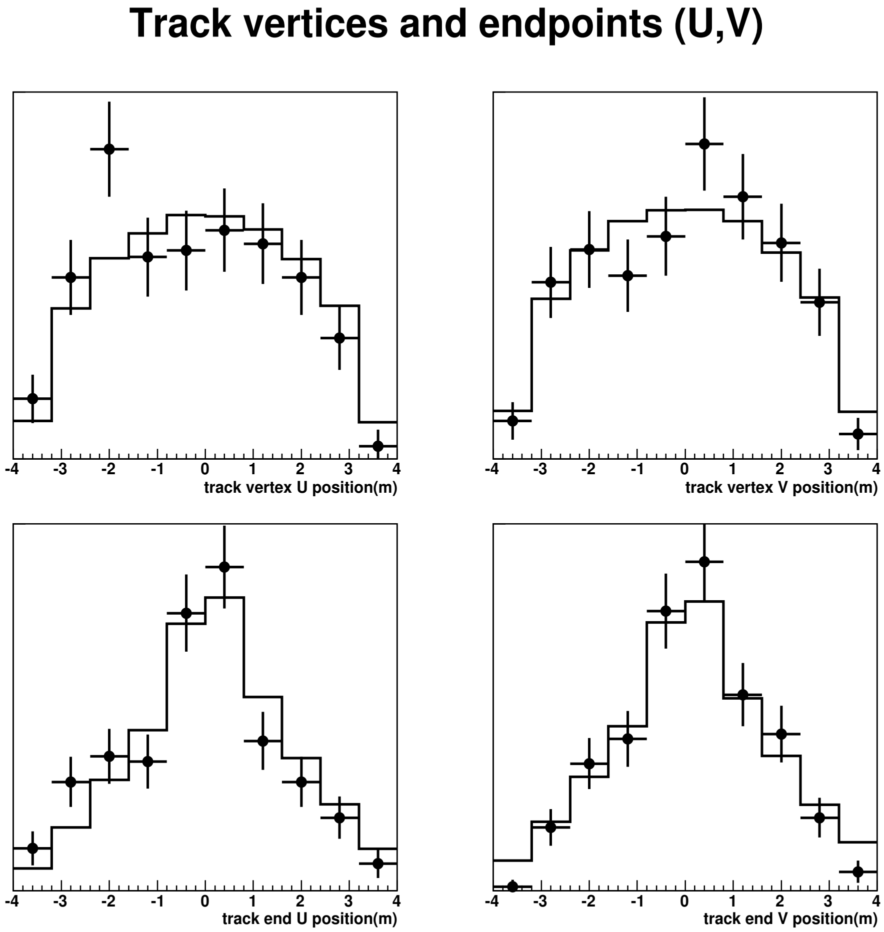


Figure 17: Distribution of the reconstructed track vertices and endpoints in U and V coordinates of selected neutrino candidates (black points) for the hidden dataset. Note that the irregularities observed in the open dataset are reduced with the increased statistics of the hidden dataset

3-6 and > 6 - these distributions describe the longitudinal deposition of energy in the event and test shower modelling in the MC.

Figure 18 shows these distributions for all events in the open sample and figure 19 shows the same plots for a clean sample of CC-like events (selected using a cut on the PID variable - see later - of -0.2). This cut preferentially selects low-y events, and this is reflected in changes to the pulse height fraction plots. The agreement between data and MC is fair in these plots, although there are some discrepancies - most notably in the bottom-right plot of each figure (event pulse height fraction beyond plane 6).

In order to remove the effect of blinding, the hidden sample was also used for this comparison and the results are shown in figures 20 and 21. The effect of oscillations with representative values of Δm^2 on the shape of these distributions is considered in figure 22 and is shown to be small. The agreement between data and MC for the hidden sample is generally good although there is still a minor discrepancy in the "event pulse height fraction beyond plane 6" plot, where the data tends to lower values of this ratio than the MC. This effect was also observed (with much higher significance) in the ND distributions shown in [8], and is interpreted as being due to a higher electromagnetic shower fraction in data showers than in the MC. The other effect observed in the ND distributions - that the average event pulse height per plane is slightly lower (by $\sim 2\%$) in data than MC - cannot be observed with the statistics shown here.

In summary, I would characterise the data/MC agreement for these quantities to be good, with a slight difference in shower energy profiles that is also observed in ND data.

4.5 Check 5: Basic track distributions

Figure 23 shows some very basic track distributions. The distribution of track length (left plot) is of course not expected to agree with the MC since a) it is affected by oscillations and b) it is one of the principal variables used in the blinding algorithm. In this distribution, the sinusoidal dependence of the blinding function is clearly evident. Although the data/MC comparison is meaningless here, it can at least be shown that there are no pathological spikes in the distribution that would be evidence of reconstruction or detector problems. The plot on the right shows the average number of track digits per plane - this is a rough measure of tracking efficiency and this distribution shows that the efficiency is slightly lower for data than MC. It is unlikely though that a small difference such as this could cause a significant difference in data/MC track reconstruction.

Figure 24 shows distributions relating to the quality of the track curvature fit. These distributions are intended to verify that the track fitter is performing in a consistent way for both data and MC events. The top-left plot shows that the rate of fitting failures is comparable for data and MC (actually almost identical). The top-right plot shows the χ^2 per degree of freedom, which is also very similar. The bottom-left plot shows the (signed) muon momentum returned by the fit. Of course, this will be affected by blindness and oscillations, but it can be verified that the fit returns reasonable values and that the ratio of μ^+/μ^- events indicated by the fit is also consistent with expectations. The plot at bottom-right shows the fractional error on the momentum measurement, which also shows good agreement between data and MC.

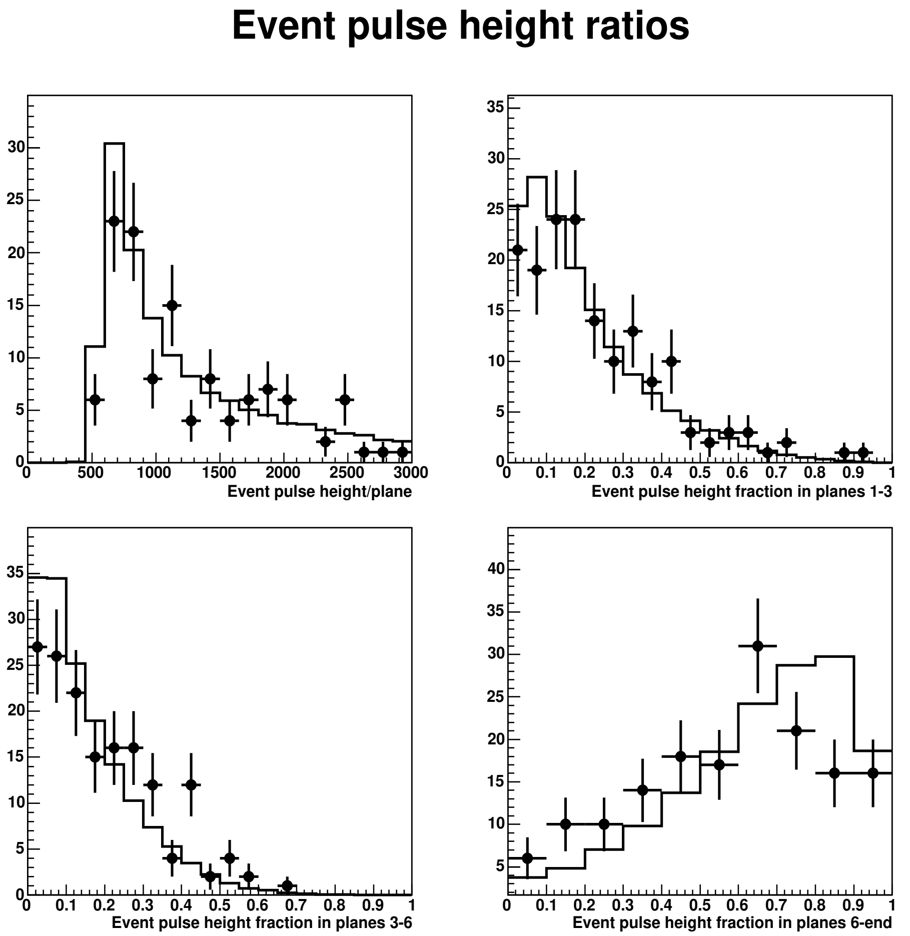


Figure 18: Event-based quantities for the open dataset. These are (from left to right and top to bottom): event pulse height per plane; fraction of total pulse height in planes 1 to 3; fraction of total pulse height in planes 3-6; fraction of total pulse height beyond plane 6.

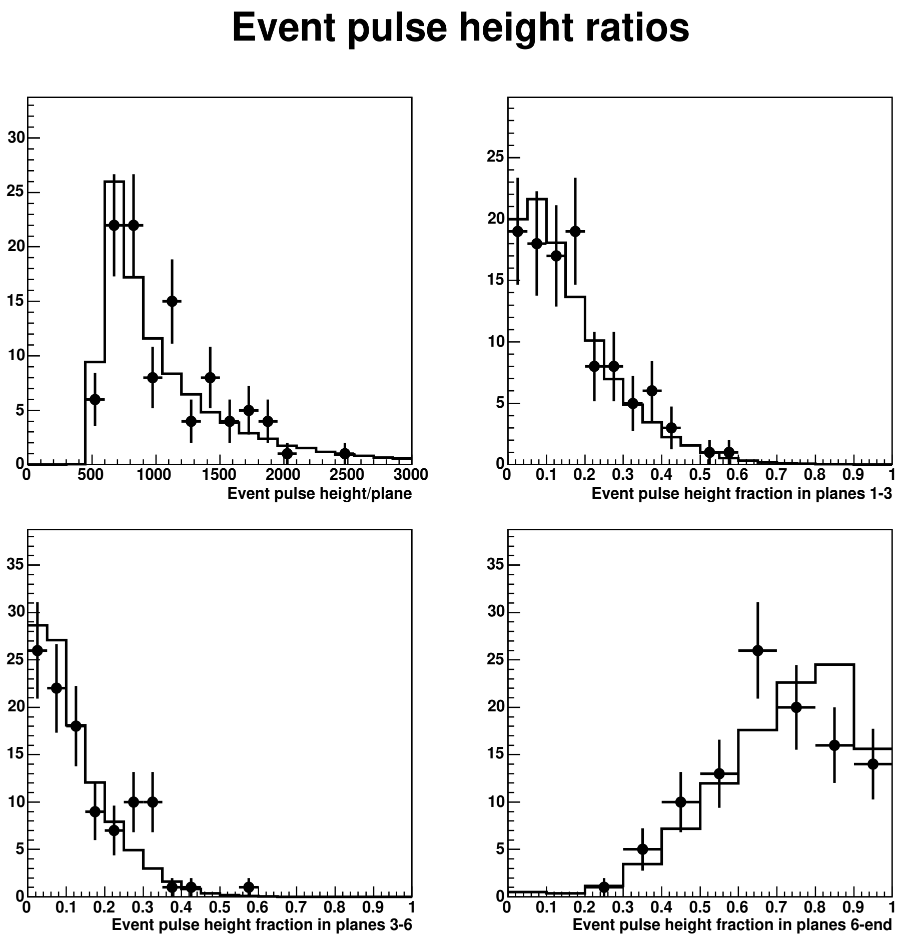


Figure 19: As figure 18 but with a PID cut of -0.2 applied to select CC-like events

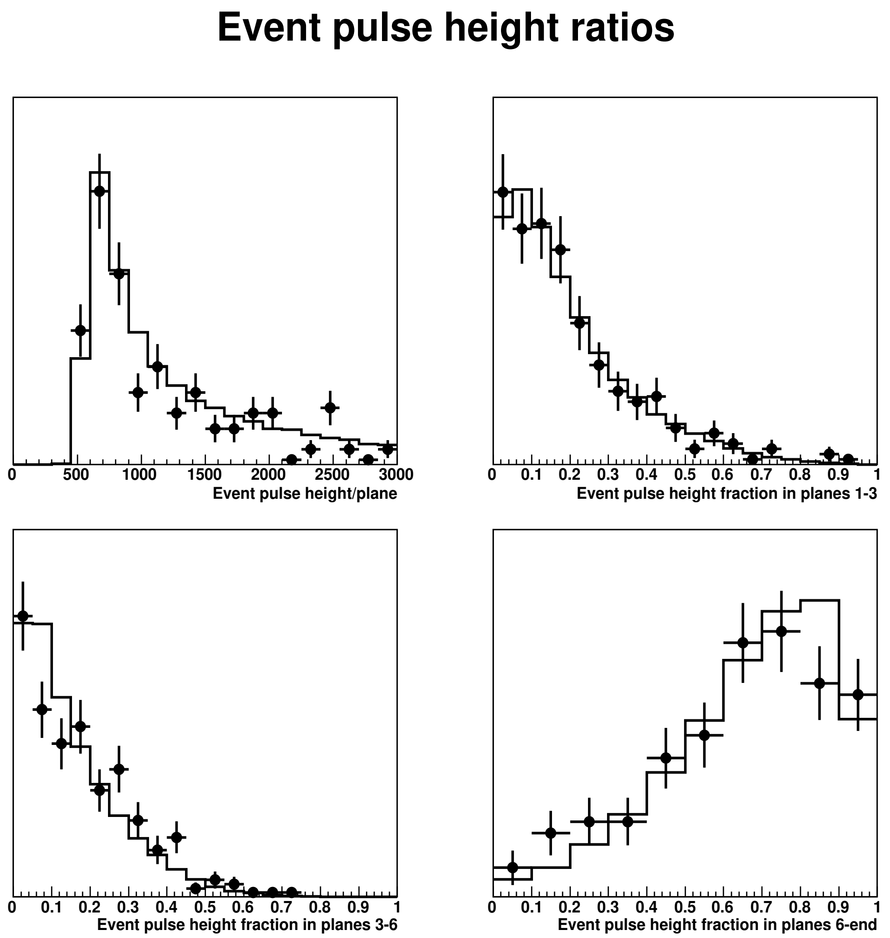


Figure 20: As figure 18 but for the hidden dataset.

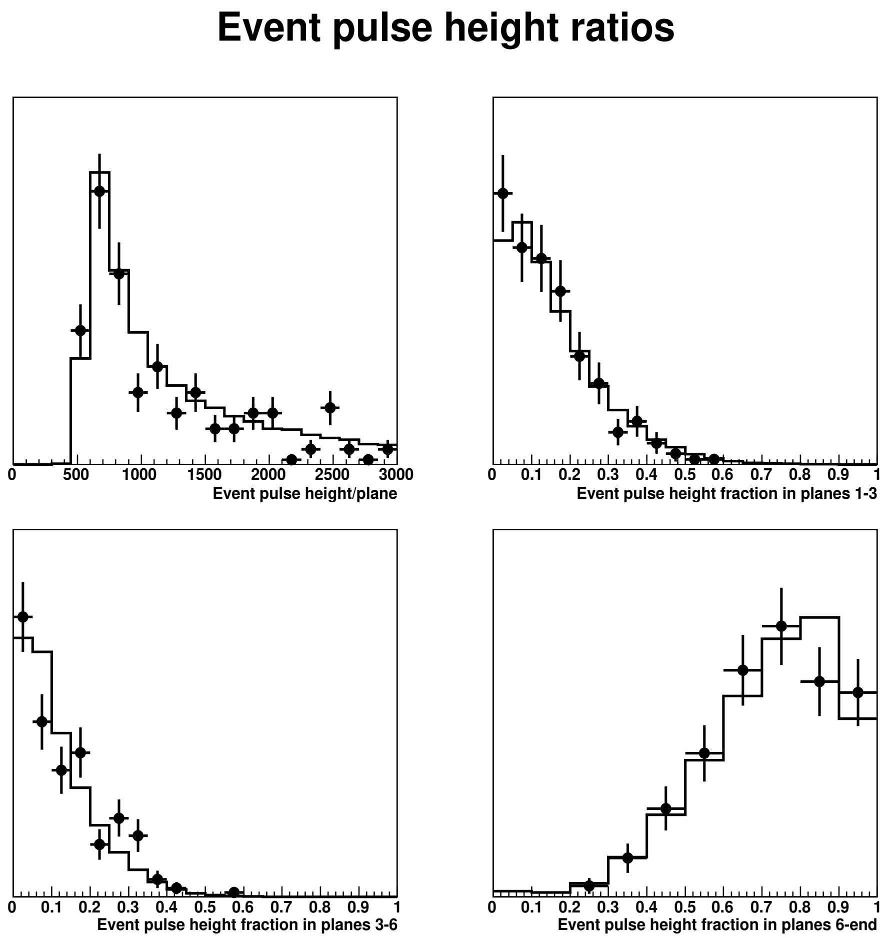


Figure 21: As figure 19 but for the hidden dataset.

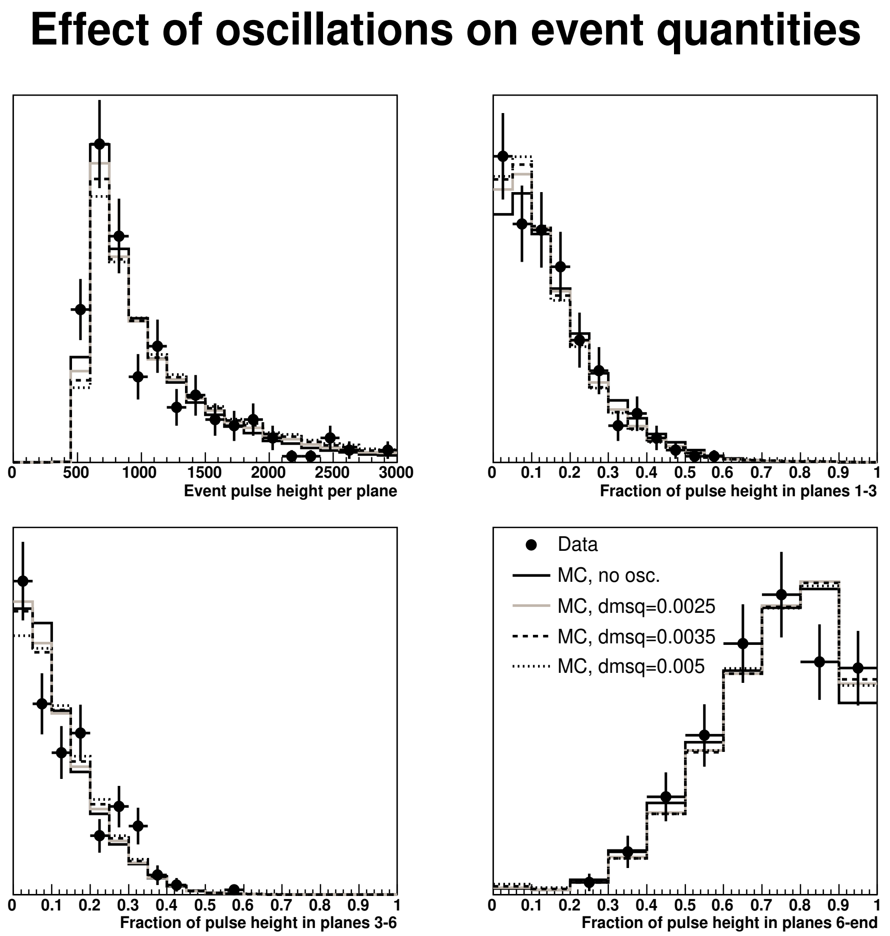


Figure 22: Effect of oscillations on the shape of the event-based distributions. Oscillations with $\Delta m^2 = 0.0025, 0.0035, 0.005$ eV^2 and $\sin^2 2\theta = 1$ are considered, and the hidden dataset is used.

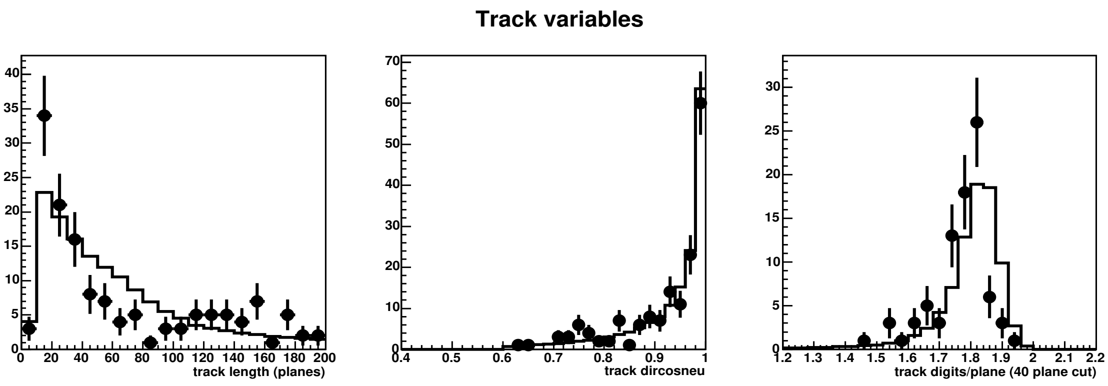


Figure 23: Basic track-related variables: left - track length in planes (the distortion introduced by the blinding algorithm is obvious here); centre - cosine of the angle of the reconstucted track relative to the beam direction; right average number of track digits per plane (a measure of tracking efficiency), for tracks that are greater than 40 planes in length.

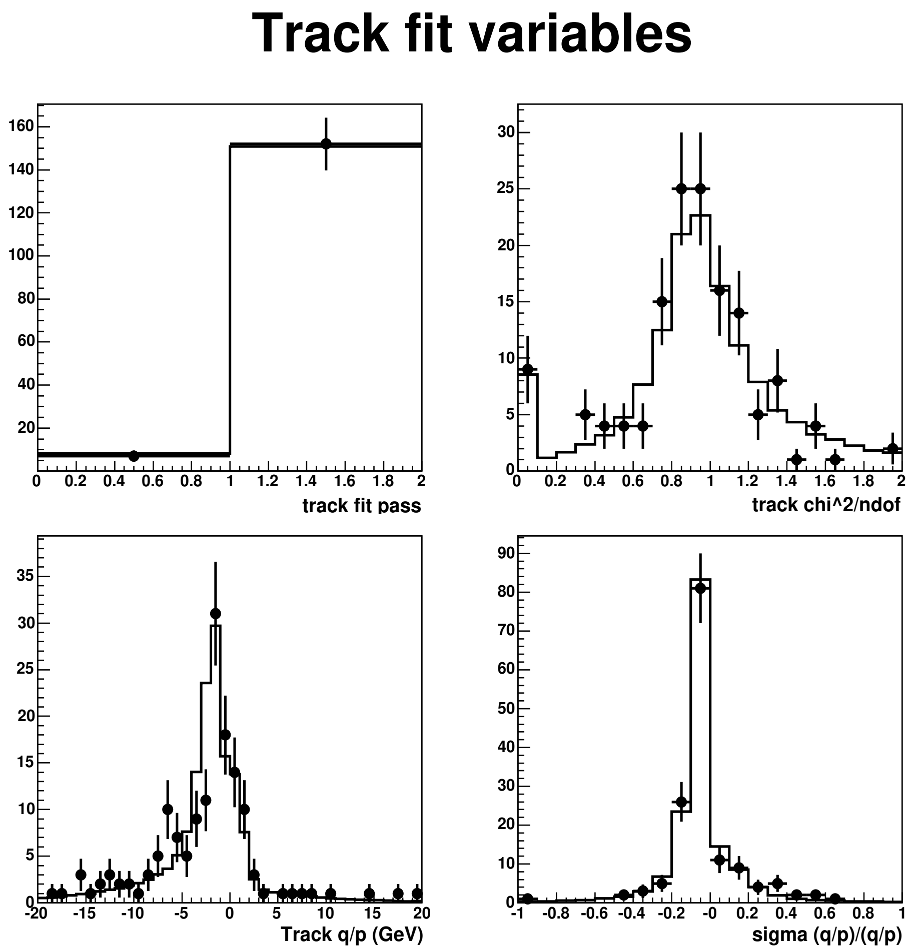


Figure 24: Variables related to track fit quality. Top left - track fit quality flag returned by the track fitter (0 denotes track fitting failure, 1 denotes an acceptable fit). Top-right - χ^2 per degree of freedom for the fitted track. Bottom-left - (signed) track momentum from curvature (this distribution is of course influenced by blindness/oscillations). Bottom-right: fractional uncertainty returned on the fitted momentum from curvature.

Figures 25 and 26 show plots of the track emission angle with respect to the beam direction, and the track angle as a function of muon momentum for the open and hidden datasets respectively. These are shown for comparison with the corresponding plots in [8]. The agreement is quite good although these distributions are obviously affected by blinding and oscillations, which will tend to deplete events at moderate (1-5 GeV) values of muon momentum. This is clearly evident in the right-hand plots of the two figures, which show higher mean muon momenta for small emission angles - as would be expected from these two effects.

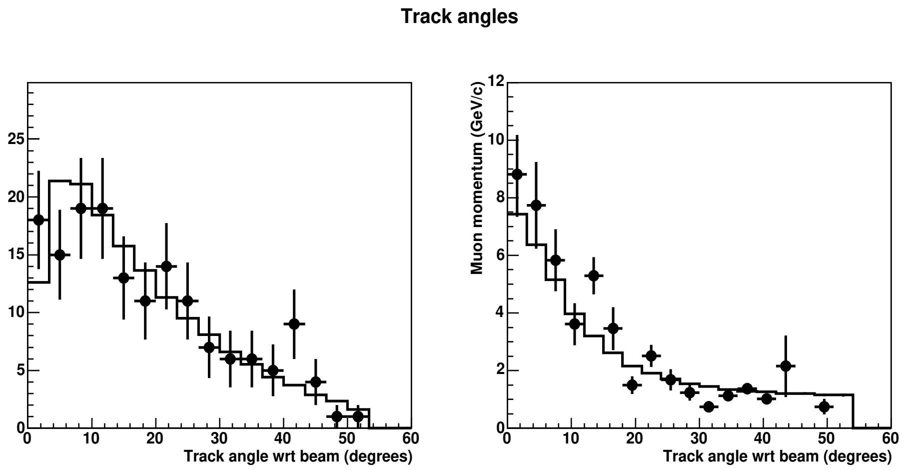


Figure 25: Track emission angle in degrees, and track emission angle as a function of muon momentum for the open dataset.

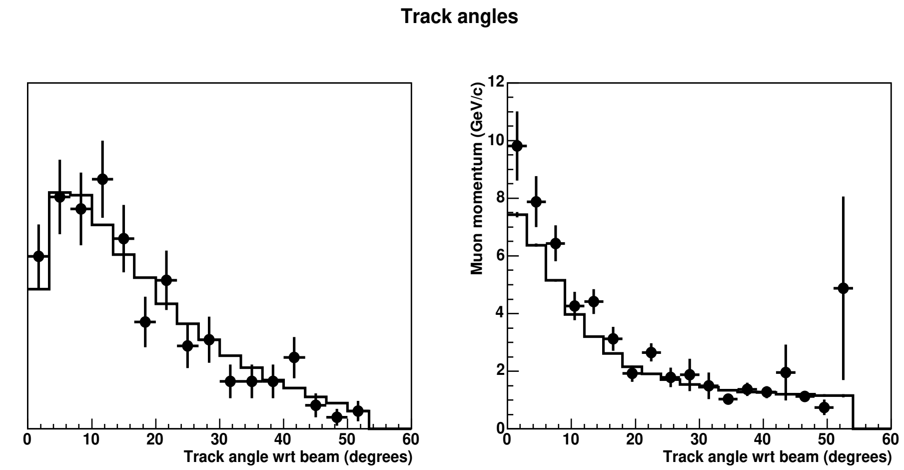


Figure 26: As figure 25 but for the hidden dataset.

One useful consistency check of the track reconstruction is to compare the track momentum derived from curvature and range for events that are fully contained in the detector. Figure 27 shows this comparison for 65 contained events that track lengths longer than 20 planes and $(q/p)/\sigma_{q/p} > 3$. In this event sample, the agreement between the two measurements is generally good.

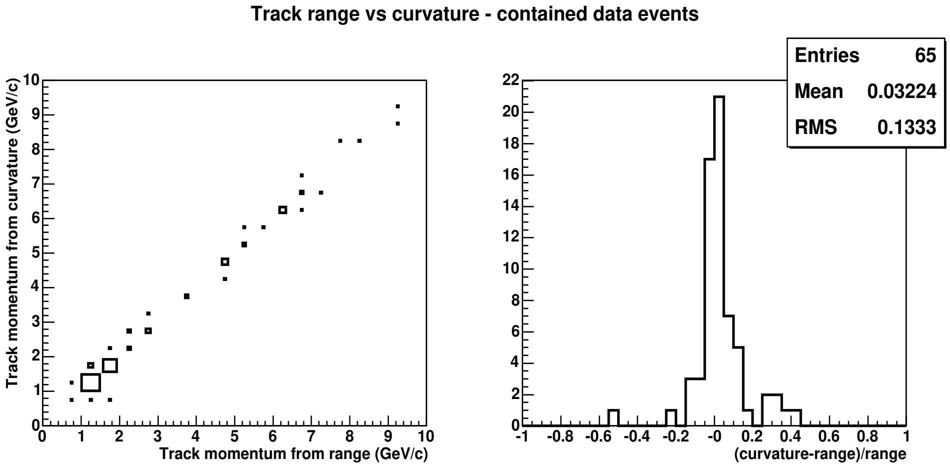


Figure 27: Comparison of track momentum from range versus curvature for fully contained events with the following quality cuts applied: track quality flag set ($trk.fit.pass = 1$); ratio of track momentum from curvature to quoted error on the fit > 3; track length greater than 20 planes. The left-hand plot shows a 2D distribution of track momentum from range vs curvature. The right-hand plot shows the distribution of the difference in the two momentum measurements divided by the momentum from range for these events.

4.6 Check 6: Shower distributions

The basic characteristics and energies of reconstructed showers should be compared to those from MC events. Figure 28 shows such a comparison. The number of reconstructed showers agrees fairly well between data and MC, although the total reconstructed pulse height in the primary shower seems to be somewhat higher in data than in Monte Carlo. The bottom-left plot shows the average shower pulse height per digit for data and Monte Carlo events. There is a scale shift of approximately +10% between data and MC (the dashed histogram shows the MC distribution scaled up by 10% and the agreement is obviously better). The right-hand plot of figure 29 shows profile histograms of this quantity as a function of shower pulse height, and it can be seen that the scale factor remains roughly constant over this range.

In order to understand these differences, I have considered the effect of blinding and oscillations on these distributions. Figure 30 shows how these quantities are affected by oscillations with three representative values of Δm^2. The effect of oscillations is to deplete the number of low (1-5 GeV) events, and therefore to preferentially suppress lower energy showers. The surviving events tend to have higher average shower energy, more shower digits and larger average pulse height per plane/digit (since higher energy showers have higher strip occupancies and therefore per-strip energy deposition than lower energy showers). These features can be clearly seen in the oscillated histograms that are shown in figure 30.

Figure 31 shows the same plots for the hidden data sample where the effect of the blinding function is taken out. Here it is clear to see that the agreement between data and MC is significantly better when oscillations are taken into account.

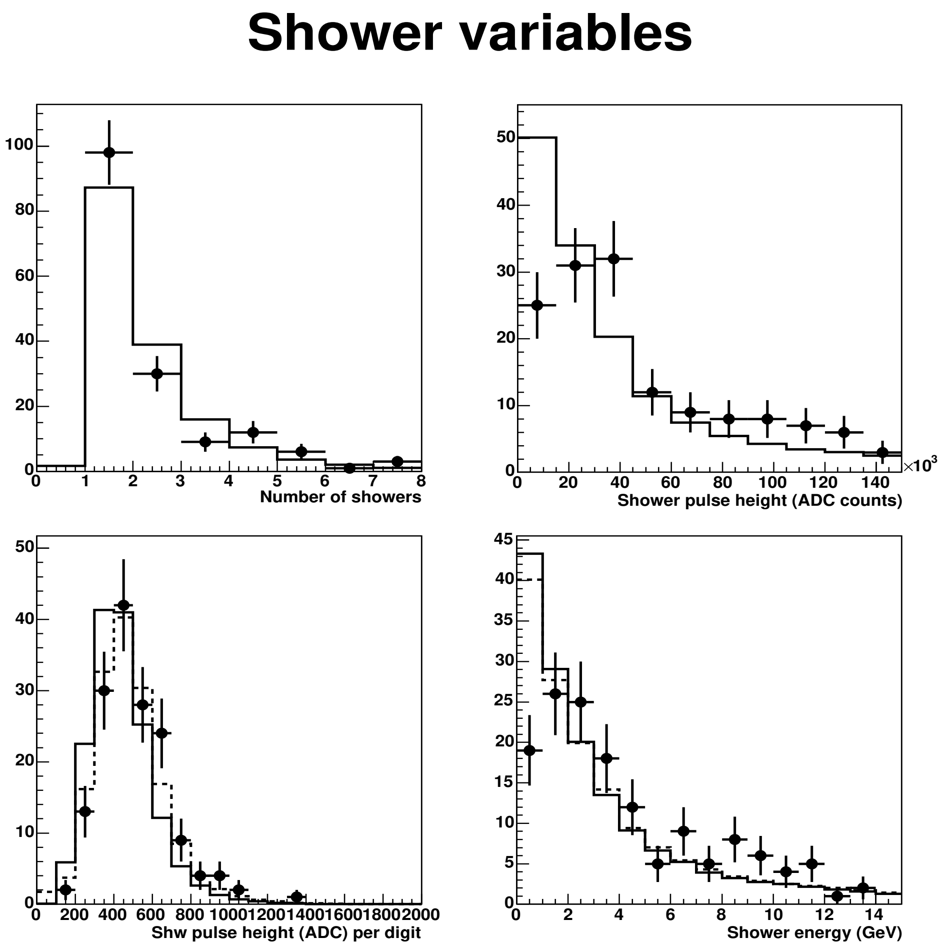


Figure 28: Variables related to shower finding and shower energy estimation. Top left - Number of reconstructed showers per event. Top-right - total pulse height in the shower. Bottom-left: shower pulse height per digit. Bottom-right - shower energy in GeV. The dashed lines in these two bottom plots shows the effect on the MC distributions of a +10% scale shift in the shower pulse height and reconstructed shower energy respectively. The agreement is better for the former, but barely changed for the latter.

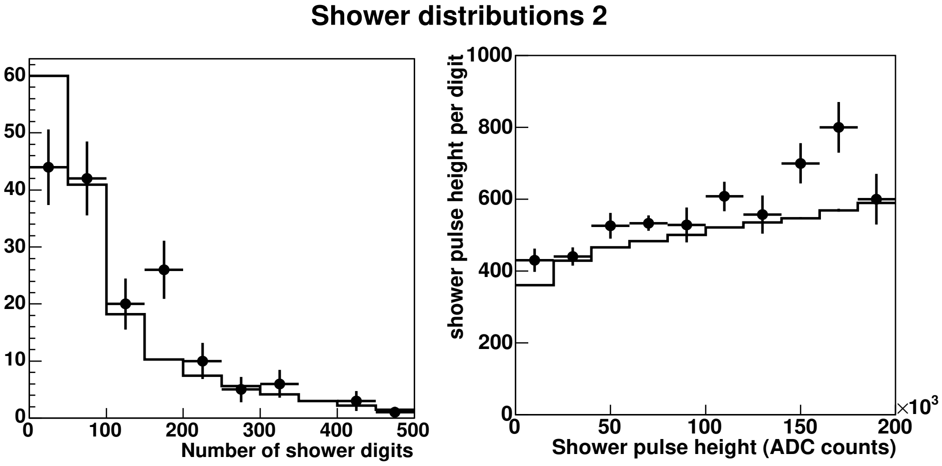


Figure 29: More variables related to shower finding and shower energy estimation. Top left - Total number of digits in the primary shower. Data showers tend to contain more digits than MC showers. Top-right - profile plot of shower pulse height per digit versus total shower pulse height. Note the scale shift between data and MC.

The effect of blinding is shown in figures 32 and 33. Figure 32 shows the shower energy distribution that results from the blinding function that is applied to the data, with five randomly generated sets of blinding parameters (the true parameters used are of course unknown). This plot gives a sense of the distortion that is induced by blinding. Figure 33 shows, for the open sample, data and MC distributions for four shower quantities, where the MC sample has been weighted with oscillation parameters $\Delta m^2 = 0.0025, \sin^2 2\theta = 1$. The variation of these distributions due to blinding has been computed by randomly generating 100 sets of blinding parameters, weighting the MC events by the resulting blinding functions, and calculating, for each bin, the mean deviation between the blinded and unblinded distribtions. These variations are represented by the grey areas in 33. The agreement between data and MC with blinding and oscillations applied - while not perfect - is significantly better than the comparison between data and unoscillated MC.

My conclusion from the comparison of shower quantities is that, once oscillations and blinding are accounted for, the agreement between data and MC is generally good, and figure 31 - which factors out oscillations and blinding - is a good demonstration of this. In the Near Detector, the average shower pulse height per plane was shown to be 5% lower in data than MC [8]. This is not clearly observed in the FD distributions, although the statistics are much poorer, and the precise shape of this distribution does depend somewhat on assumed oiscllation parameters. In any case, a difference between ND and FD shower energy scales of 5% is within the quoted relative ND-FD calibration uncertainty [10].

4.7 Check 7: Event separation

Figure 34 shows the variables that are input to the likelihood-based event separation method that is used to distinguish between ν_μ CC and NC events. Some of the variables are of course distorted by blinding and oscillations, in particular the event length variable (which is related to muon momentum) and the track pulse height fraction (which

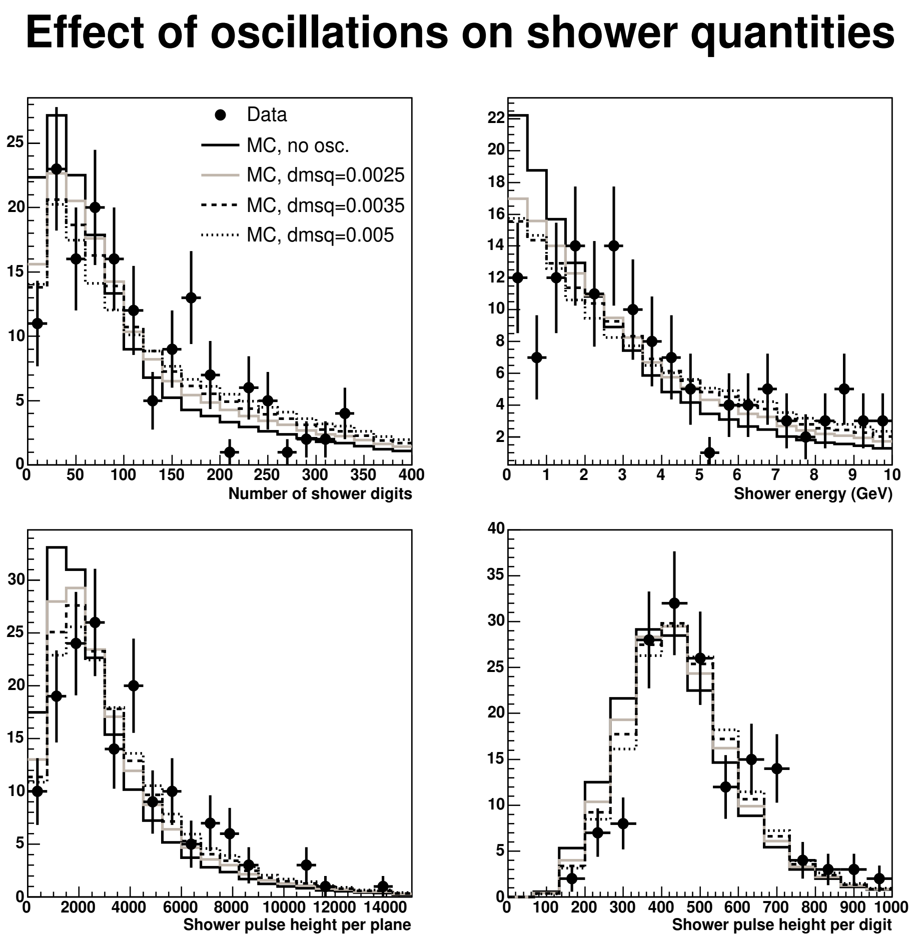


Figure 30: Effect of oscillations on shower quantities for the open dataset. Top-left: number of shower digits; top-right: reconstructed shower energy; bottom-left: average shower pulse height (sigcor ADC) per plane; bottom-right: average shower pulse height per digit. It can be seen that the agreement between data and MC is improved when oscillations are taken into account.

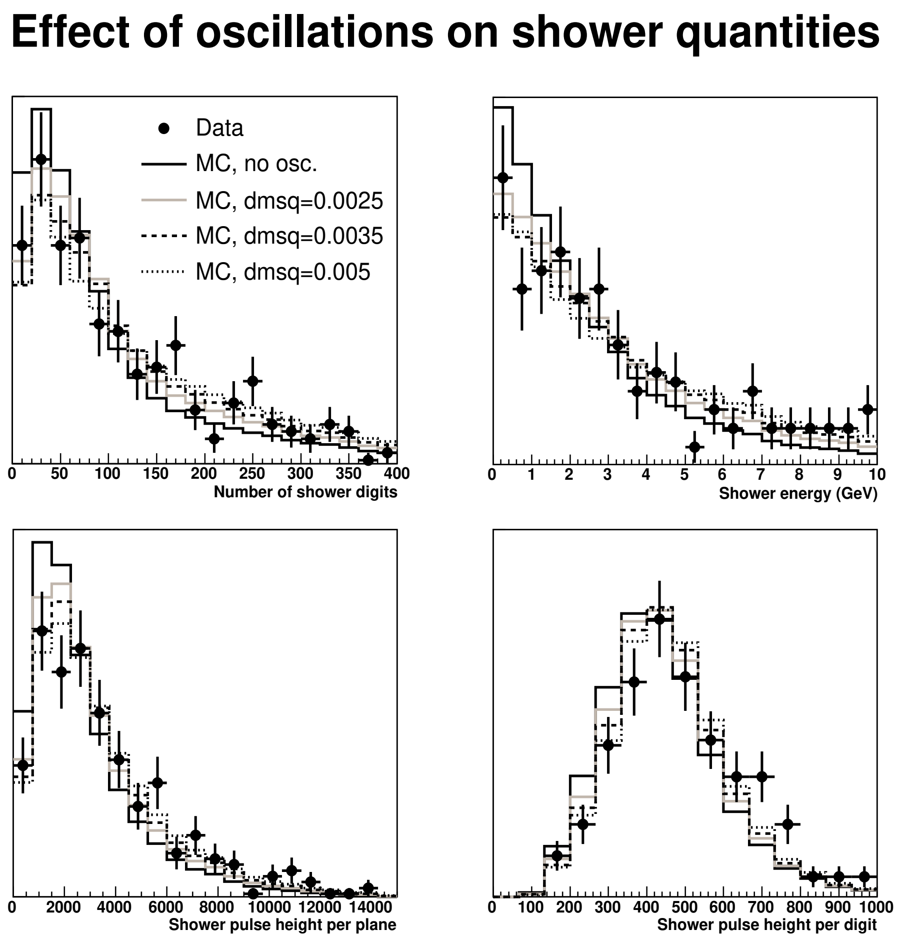


Figure 31: As figure 30 but for the hidden sample.

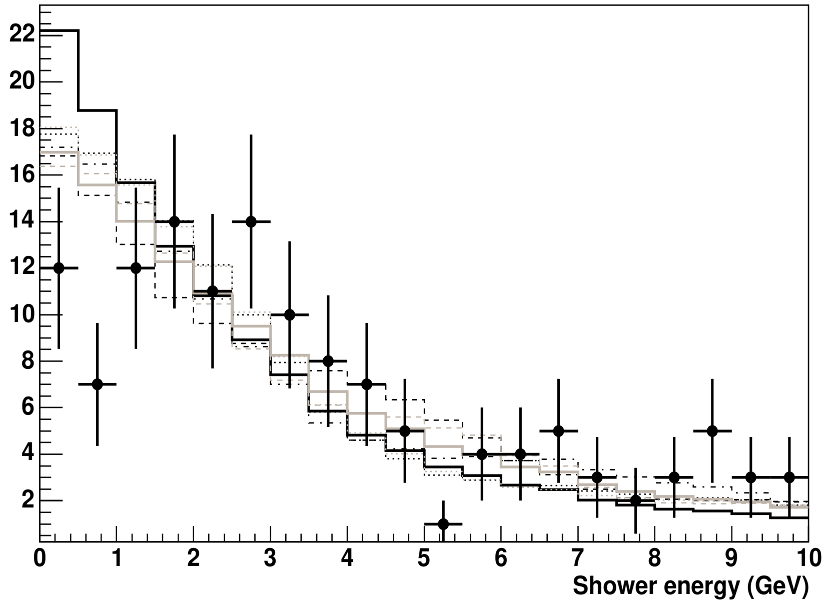


Figure 32: Effect of blinding on the reconstructed shower energy distribution. The thick black line shows the MC expectation for no oscillations and the thick grey line shows the expected distribution assuming oscillations with $\Delta m^2 = 0.0025, \sin^2 2\theta = 1$. The thin lines show the effect of five randomly generated blinding functions on the oscillated distribution. The data points are for the open dataset.

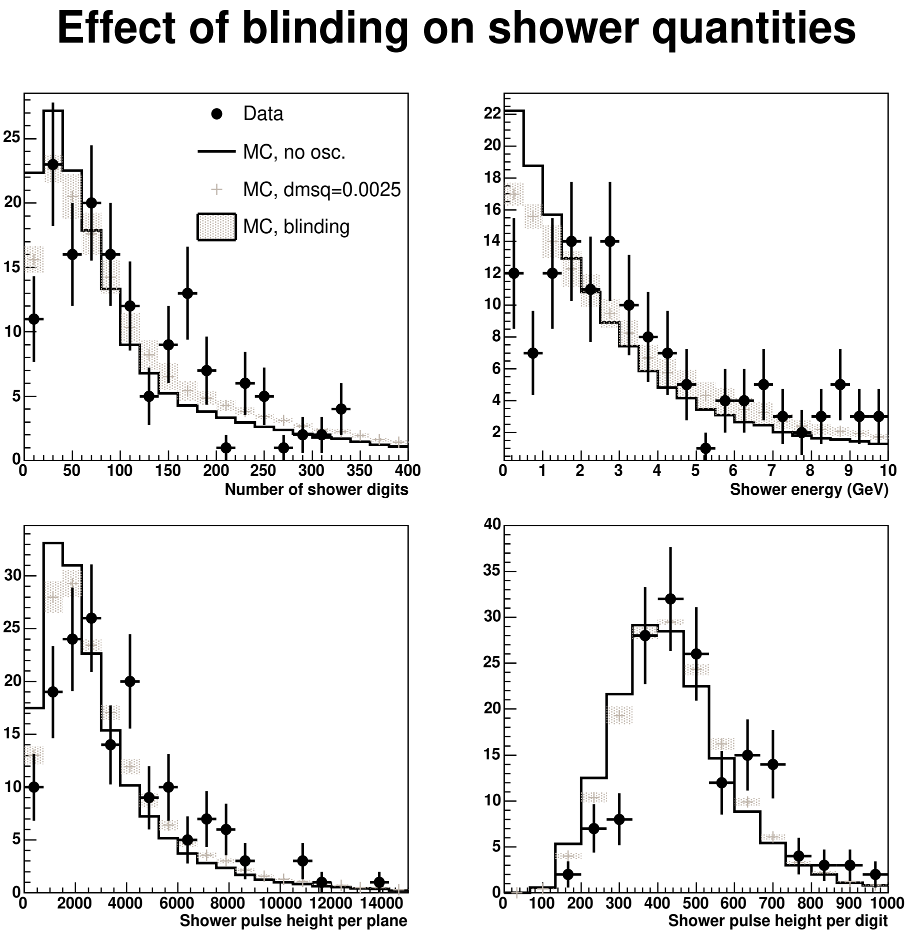


Figure 33: Effect of blinding on the four shower variables plotted in figure 30. The grey bars show the mean deviation from the nominal oscillated distribution ($\Delta m^2 = 0.0025, \sin^2 2\theta = 1$) introduced by the blinding function that is applied to the FD data (with randomly generated blinding parameters, since the true parameters are not known).

is somwehat related to the inelasticity of the event). The third variable (track pulse height per plane) is, as expected, invariant to blinding and oscillations. The PID parameter is defined as:

$$PID = -[\sqrt{(-\log P_{\mu})} - \sqrt{(-\log P_{NC})}] \tag{1}$$

where:

$$P_{\mu,NC} = \prod_{i=1}^{3} f_i^{\mu,NC}(\epsilon_i) \tag{2}$$

here, ϵ is the value of each of the three PDFs for CC and NC events respectively.

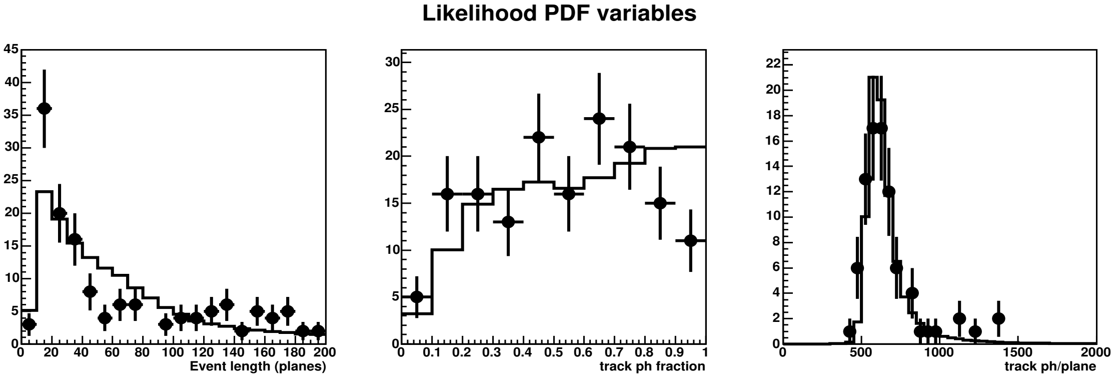


Figure 34: Distribution of the three variables that are used as PDFs in the likelihood-based ν_μ CC selection algorithm. Left - event length (in planes); centre - fraction of total event pulse height contained in the reconstructed track; right: average track pulse height per plane (measured in units of sigcor ADC).

It is a little difficult to conclude much from these distributions as the first two are heavily distorted by blinding (and oscillations in the case of the event length distribution). Given that the track pulse height fraction depends only to second order on oscillations and it is important to verify that the PID selection algorithm is performing as expected, I believe it is legitimate in this case to use the hidden sample to check that the MC adequately describes the data for this variable. For the open dataset, the χ^2 of this distribution is quite poor ($\chi^2 = 15.5/10, P = 0.115$), but for the hidden dataset (shown in figure 35) the agreement is much better (reflected in an improved $\chi^2 = 5.9/10, P = 0.82$). This level of agreement gives additional confidence in the robustness of the PID selection algorithm. Note that the mean track pulse height per plane is slightly lower in data than MC. This feature was also observed for ND data [8] and could be related to a known issue in the current MC generation where the muon dE/dX is too high by $\sim 2\%$.

The distribution of the PID parameter for data and MC events is shown in figure 36. The shapes are fairly similar, although there is an excess of NC-like events ($PID < -0.2$) in the data which is expected both from blinding and oscillations. Figure 37 shows a comparison of the PID variable for events that are shorter than or greater than 40 planes in length. When the events are divided in this manner, which factors out some of the effects of oscillations/blinding, the distributions become more similar.

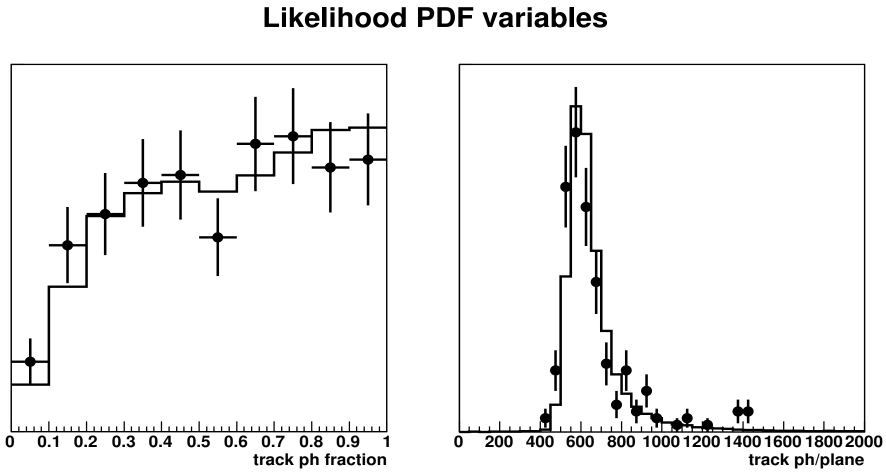


Figure 35: Distribution of the two oscillation invariant variables that are used as PDFs in the likelihood-based ν_μ CC selection algorithm. Here the hidden dataset is used. Left - fraction of total event pulse height contained in the reconstructed track; right: average track pulse height per plane (measured in units of sigcor ADC). Notice the much improved agreement between data and MC for the track ph fraction variable over the same distribution for the open dataset

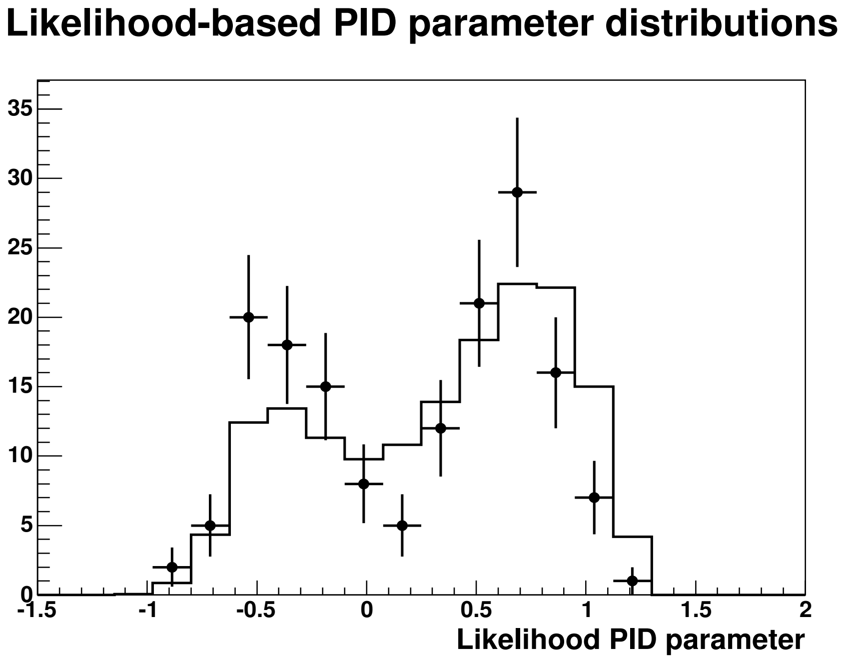


Figure 36: The PID variable derived from the three individual PDFs. Events with negative values of this quantity are more likely to be NC and events with positive values are more likely to be CC. A cut of PID> -0.2 is used to select a clean sample of CC events.

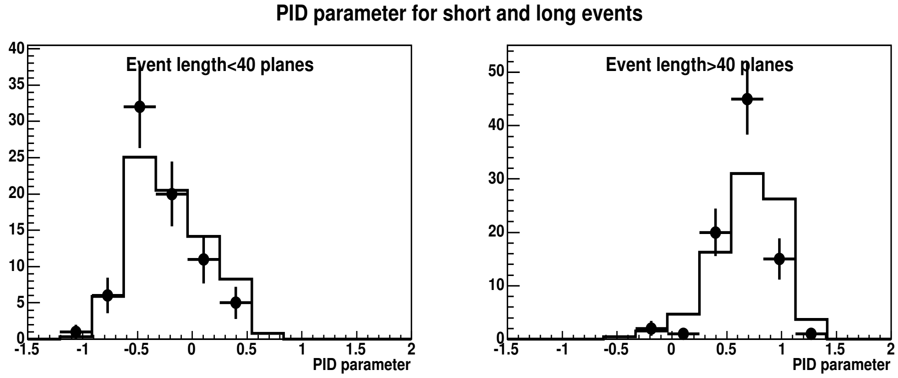


Figure 37: The PID variable plotted for events that are less than 40 planes in length (left-hand plot) and for those that are greater than 40 planes (right hand plot). Short events are a mixture of CC and NC events, long events are more predominantly CC-like.

These checks should also be carried out for other well-tested selection algorithms, such as the ANN PID method [9] utilised in [7]. Comparison of the two PID variables for the same events will provide a useful consistency check of the selection methods. Figure 38 shows the distribution (for both data and Monte Carlo) of the ANN PID variable that is used to discriminate between CC and NC events. Here, events with low values of the ANNPID parameter are considered CC-like (ANNPID < 0.2), and those with large values are classified as NC-like (ANNPID ≥ 0.2). The left-hand plot shows the distributions obtained for all events and the right-hand plot shows only events that are shorter than 40 planes in length. The overall shape of the distributions look reasonable, although the same trend that was seen in the PDF PID variable is evident - that there are relatively more NC-like events than CC-like events in the sample, as compared to the unoscillated MC. This difference could be caused by both oscillations and blinding.

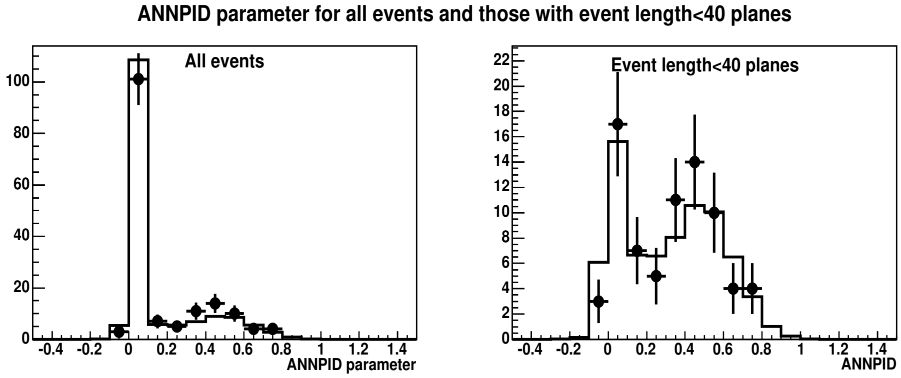


Figure 38: The ANNPID variable plotted for all events (left-hand plot) and that are less than 40 planes in length (right-hand plot). Events that have low values of the ANNPID parameter (here the cut is defined to be ANNPID<0.2) are considered to be CC-like.

Figure 39 shows distributions of one PID quantity for those events that are classified as CC-like or NC-like by the other selection algorithm. This shows that CC-like events selected by one algorithm, tend to be classified as CC-like by the other. This is more clearly illustrated by figure 40 which shows the correlation between the two PID variables for all events (left-hand plot) and events that are shorter than 40 planes (right-hand plot). It can be seen that the agreement between the two selections is very good. Because of oscillations and blinding, it is only possible to obtain a general sense of the agreement between data and MC for a particular selection algorithm. However, for this test - which uses data only - to show such good agreement between two quite different selection algorithms is very encouraging, and gives us confidence that the selection methods are robust and working as expected.

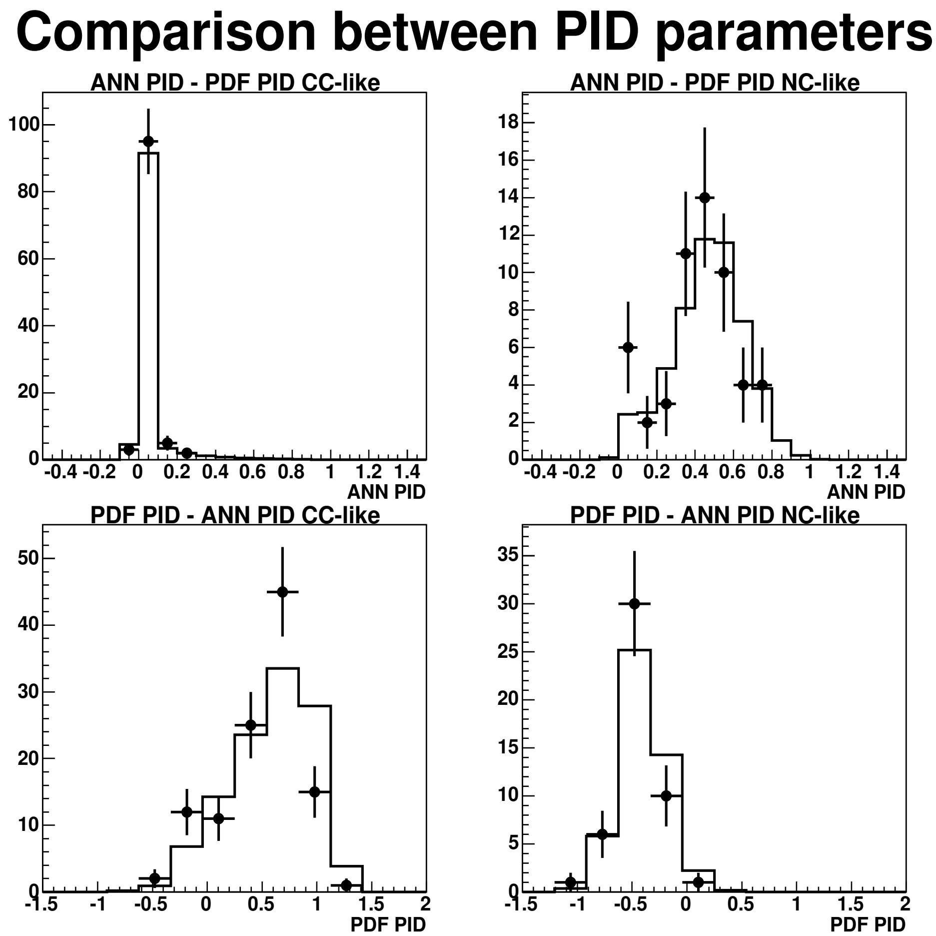


Figure 39: Comparison of ANN and PDF PID parameter distributions. The top two plots show the PDF PID variable distributions for events that are classified as CC-like by the ANN selection method (ANNPID<0.2) and those that are classified as NC-like (ANNPID> 0.2). The bottom two plots show the ANN PID variable distributions for events that are classified as CC-like by the PDF selection method (PDFPID> -0.2) and those that are classified as NC-like (PDFPID< -0.2).

4.8 Check 8: Physics distributions

Figure 41 shows the distributions of physics quantities (reconstructed muon, shower and neutrino energies and reconstucted y) for both data and MC events. Of course these distri-

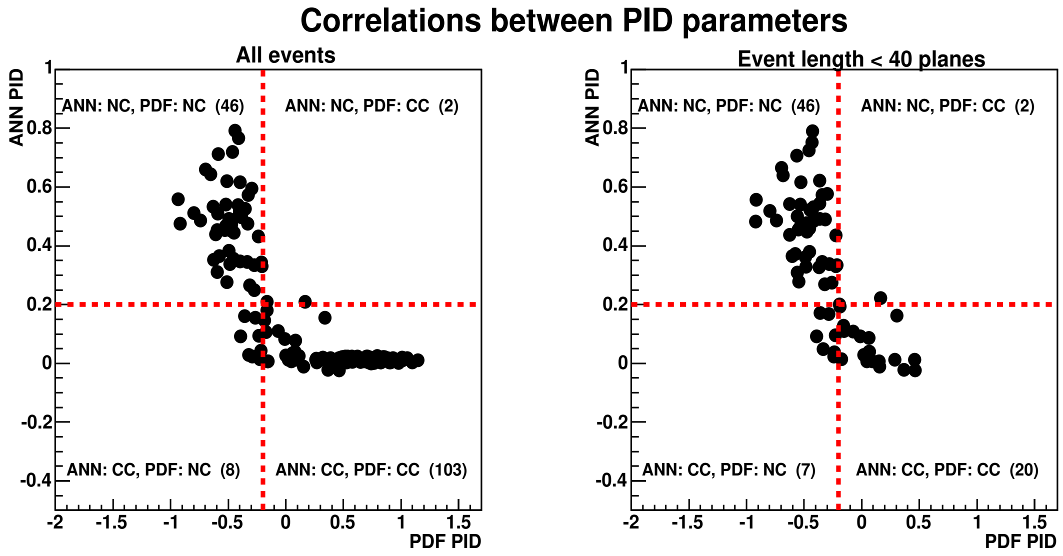


Figure 40: Correlation between the ANNPID and PDFPIF variables for neutrino candidate events selected by method #2. The left-hand plot shows all selected events, and the right-hand plot shows events that are shorter than 40 planes in length. The red dashed lines denote the cuts that are applied to select CC-like events for the two selections. These are ANNPID<0.2 and PDFPID>-0.2. It can be seen that the agreement between the two methods on the classification of these events as CC-like or NC-like is very good.

butions are not expected to agree, due to both blinding and oscillations, but the purpose of examining them is to search for pathologies in the distributions and inconsistencies between them. We know what the gross effects of blinding and oscillations are, and these distributions should exhibit these features. The muon momentum distribution clearly shows the effect of blinding; there is also a suppression of low-y events, which has components of blinding, neutrino oscillations (low-y events are preferentially suppressed for particular values of Δm^2), and the shower energy discrepancy that is seen in the first bin in the top-right plot (which was shown earlier to be consistent with the effect of oscillations).

Figure 42 shows the same distributions for events with $PID > -0.2$, which should select a pure sample of CC events. The features of the plots are similar to before, except that the selection biases the events to lower values of y. Again, there are no obvious pathologies evident here.

4.9 Summary on blinded checks

The checks on the blinded sample can be summarised as follows:

- Visual scan to ensure that selected events are neutrino interactions, that there is negligible background and that the events are well-reconstructed.
  - **Result:** *Preliminary OK - A complete scan of R1.18 reconstructed events showed that the selected events were good neutrino candidates and that the events were generallly well-reconstructed. This check is underway for the R1.18.2 dataset*

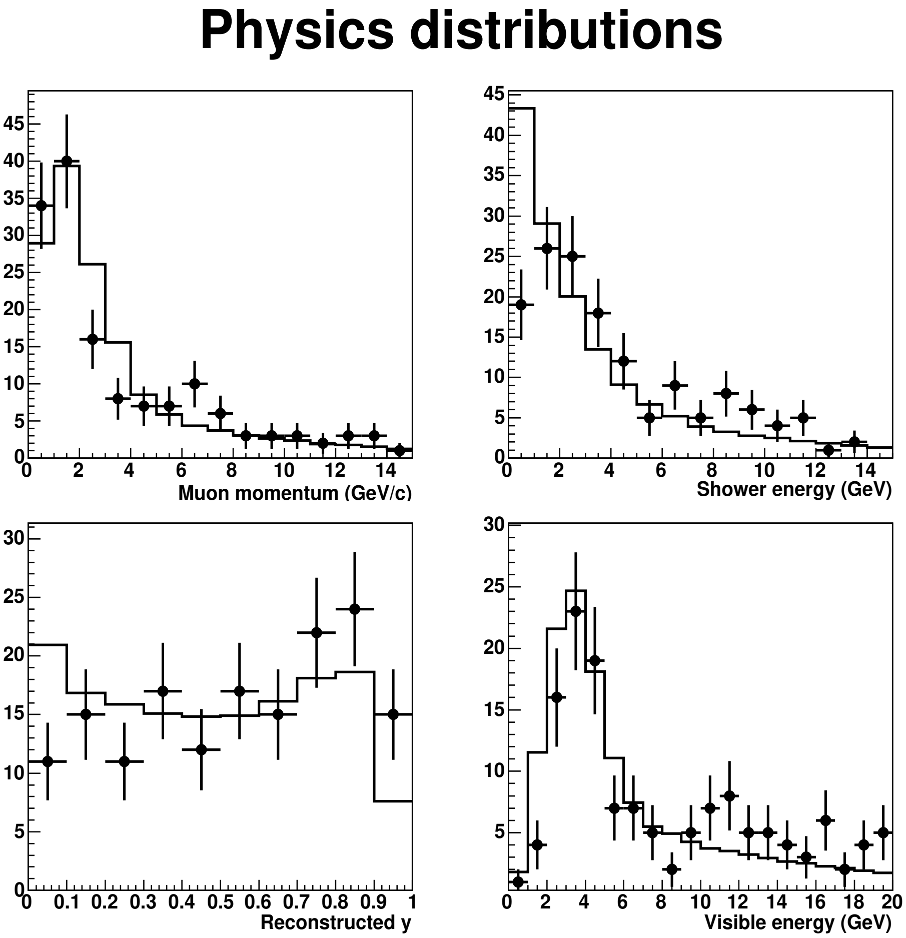


Figure 41: Distributions of “physics” variables for data and LE-10 Monte Carlo events. Top-left - reconstructed muon monentum (from range or curvature as appropriate); Top-right - reconstructed shower energy; Bottom-left - reconstructed y; Bottom-right - visible energy (or reconstructed neutrino energy). Of course, these distributions are expected to be significantly distorted by both oscillations and the blinding algorithm applied to the open dataset. The purpose of examining these plots is to ensure that there are no gross shifts or distortions that cannot be accounted for by either of these two effects.

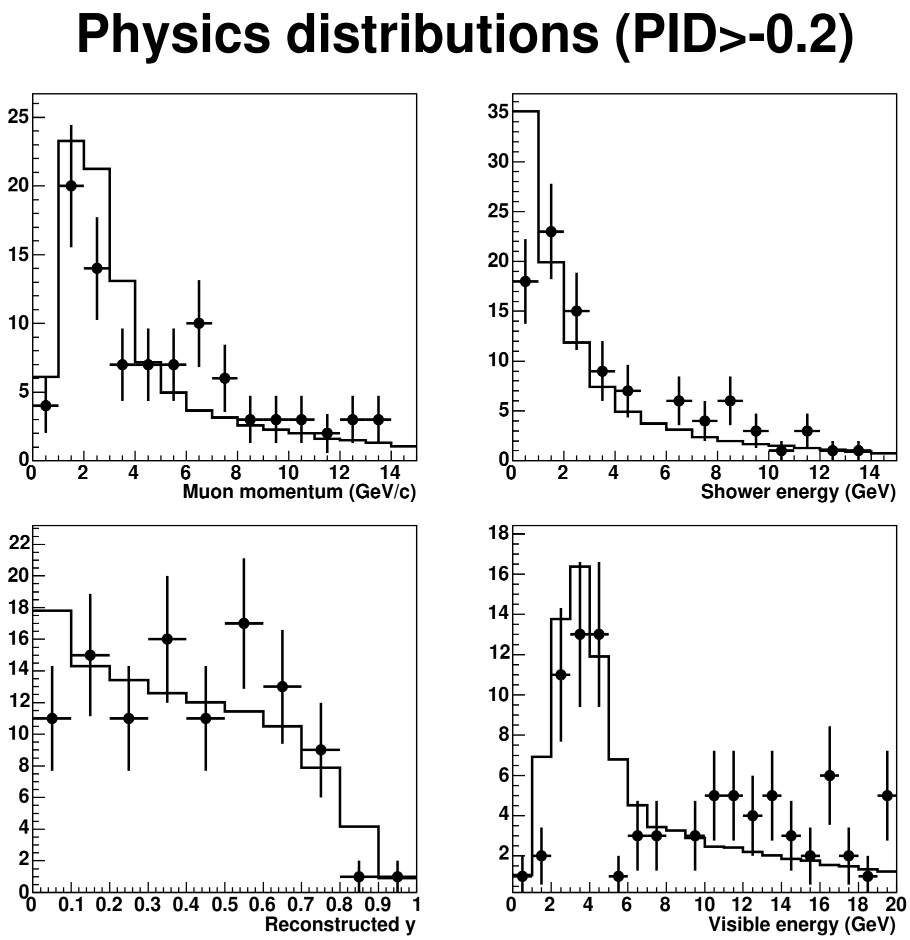


Figure 42: Distributions of “physics” variables for data and LE-10 Monte Carlo events with a cut on the likelihood PID parameter of $PID > 0.2$. This cut should remove the vast majority of neutral current interactions from the sample. The same four quantities are shown as those in figure 41.

- Timing check to verify that the selected events are coincident with beam spills and that the background contamination is again small.
  - **Result:** *PASSED - All neutrino candidates are localised in a 10μs window around the time of beam spills and there is no evidence of background in a $\pm 50\mu s$ window around this time (see figure 13).*
- Checks on basic quantities that are largely invariant under oscillations and blinding (such as event vertices and pulse height distributions). The data and MC distributions should agree.
  - **Result:** *PASSED - The invariant variables show reasonable agreement between data and MC. There are some discrepancies in the track endpoint distributions for the open set, which were reduced by comparing the same quantities in the hidden sample. Pulse height distributions for tracks look fine and the discrepancy between shower pulse height distributions for data and MC can be accommodated when second-order effects of oscillations and blinding are taken into account.*
- Checks on track quality and track fitting. Consistency checks between range and curvature measurements.
  - **Result:** *PASSED - Good agreement between data and MC for invariant track quantities (see figure 18. Agreement between range and curvature measurement of contained data events is very good (see figure 19).*
- Check of event separation techniques - do PID distributions make sense given the input distributions? Consistency between independent selection methods.
  - **Result:** *PASSED - Good agreement between data and MC for invariant PDF quantities for the likelihood-based analysis (see figures 34 and 35). Use was made of the hidden sample since one PDF (which is largely invariant to oscillations) was badly distorted by blinding. Very good agreement between the likelihood and ANN-based selection algorithms for the same data events (see figure 40).*
- Sanity check of physics distributions. Search for pathologies and inconsistencies.
  - **Result:** *PASSED - No obvious pathologies observed. Discrepancies in the track and shower-related quantities are consistent with the effects of oscillations and blinding.*

To summarise, I think that the agreement between data and MC distributions is satisfactory. The distributions that should agree do so, and those that do not, show features that are consistent with the effects of oscillations and blinding. There are some minor discrepancies (in certain shower shape variables for example) but these should not cause any significant biases in event selection and energy measurement. In addition, several distributions can be directly (or indirectly, once oscillations are taken into account) compared with their Near Detector equivalents that are presented in [8] and the results obtained are very consistent with each other.

5 Criteria for opening the box

I think the criteria for opening the box are as follows:

- Demonstrate that the analysis data sample only contains “good” runs - that is data periods with detector problems that would affect event selection and reconstruction are eliminated. **STATUS: OK**
- Validate that the required beam monitoring and timing information to select spill events is present and correct. **STATUS: OK**
- Show that we are not missing any of the data in both the open and hidden datasets. **STATUS: OK**
- Validate the selection of neutrino candidates by comparing independent selection methods. Show that the background contamination is negligible. **STATUS: OK**
- Show that there are no obvious reconstruction or detector effects in the selected event sample that would produce a significant bias in energy measurement. **STATUS: OK for tracks and showers**
- Show that distributions that do not depend on oscillations or blinding are in agreement between data and MC. **STATUS: OK, after hidden file checks**
- Show that distributions that do depend on oscillations or blinding do not show pathological effects, and possess features that are consistent with the predicted effects of oscillations/blinding. **STATUS: OK for tracks and showers**

6 Summary and conclusions

In this paper, I have outlined the following procedures:

- Data quality checks: Defined procedures to define bad runs and eliminate runs with significant problems on a per-timeframe basis. The loss of data is very small ($\sim 0.6\%$), as there are rather few runs where serious problems occured.
- Timing checks: Defined a procedure to validate that the combination of SpillServer uptime plus the use of the spill time database when the spill server is down ensures that no spills are lost in the spill data stream (except the very lowest energy NC events). Also defined a simple method to eliminate data taking periods with large timing errors (these are quite rare - only 3 occurrances during the LE-10 dataset).
- Event selection: this is validated by comparing independent selection algorithms and evaluating their selection efficiencies and background rejection factors using Monte Carlo and “fake” spill data respectively. The automated methods are additionally verified by comparing to a scan-based analysis.

- Event quality checks and Data/MC comparison: Defined a set of distributions to be checked, which test event quality and background rejection, reconstruction quality and event selection. Some of the distributions are oscillation/blindness invariant, and can be directly compared, others are dependent on these factors and are therefore used for basic sanity checks

6.1 Recommendation

I believe the decision to open the box based on the contents of this paper is contingent on two factors: that we demonstrate the ability to reliably select neutrino interactions with negligible background contamination, and that the event reconstruction and derived quantities for these events show satisfactory agreement with Monte Carlo expectations.

I believe that the first requirement is satified by the procedures detailed in this note. The results of three independent selection algorithms are presented and are shown to produce consistent results. A series of data quality cuts are applied to eliminate periods of "bad" data and the number of selected events recorded as a function of time is very consistent with the number of pot recorded, which demonstrates that we are not missing significant amounts of data. In addition, the background contamination in this is demonstrated to be small.

A number of comparisons have been made between selected data events and Monte Carlo. These tests are a little limited given the low statistics of the data and the presence of blinding and oscillations, but agreement (where it is expected) is generally good. In particular, event reconstruction and event selection algorithms appear to be performing well. There are no obvious pathologies evident in the FD sample that would cause significant biases in energy measurement.

I therefore conclude that, from the results of the tests that are described in this note, no glaring issues have been uncovered that would preclude box-opening at this time.

Acknowledgments

I would like to thank Niki Saoulidou and Alec Habig for their valuable contributions to this paper.

References

[1] D. Petyt, Event selection and data quality position paper, MINOS-doc-1351.

[2] N. Saoulidou, talk presented at 21st September 2005 ND phone meeting; D. Petyt, talk presented at October 2005 MINOS Collaboration Meeting, MINOS-doc-1178

[3] N. Tagg, private communication

[4] M. Bishai, NuMI note in preparation

[5] M. Dierckxsens, talk presented at October 2005 MINOS Collaboration Meeting, MINOS-doc-1140

[6] D. Petyt, talk presented at October 2005 MINOS Collaboration Meeting, MINOS-doc-1179

[7] N. Saoulidou, talk presented at October 2005 MINOS Collaboration Meeting, MINOS-doc-1226

[8] N. Saoulidou, CC Position Paper: "ND Data/MC Comparisons", MINOS-doc-1278

[9] N. Saoulidou, talk presented at March 2004 MINOS Collaboration Meeting, NuMI-Note 1012

[10] N. Tagg et. al., Calibration Position Paper, MINOS-doc-1289

A FD Magnet downtime

This appendix lists the time periods during LE-10 running where the FD magnet was down (here "down" is defined as current<79.8 A for either of the two supermodules). Several periods of downtime are found - these are listed in table 5 along with the affected FD run ranges and whether the beam was on or not. In total, 4 subruns taken when the beam was on are affected by magnet downtime and the time periods where these problems occurred should be excluded from the analysis dataset.

Date	Start time (UTC)	End Time (UTC)	Start/End run	Beam on	Subruns affected
22/05/05	05:26:41	06:54:12	31727/31729	YES	31729_0000 31729_0001
05/06/05	21:16:56	22:11:47	31843/31846	YES	31843_0000 31843_0001
07/06/05	14:06:19	14:38:05	31855/31872	NO	
08/07/05	14:01:34	20:16:25	32121/32126	NO	
10/10/05	17:52:39	16:48:09 10/12/05	32876/32892	NO	

Table 5: Magnet downtime periods for May 18- Oct 31 2005.

B Events lost due to beam and data quality cuts

The following is a list of those neutrino candidates that fail either beam quality or data quality cuts. All these events have been verified to be good neutrino candidates from a visual scan.

1. *run 31812, snarl 55343, June 1 2005, t=1117599756*
   - **Problem:** $\Delta t \sim 4000$ sec
   - Comments: BDP inefficiency
2. *run 33006, snarl 140889, October 24 2005, t=1130146649*
   - **Problem:** tortgt=0
   - Comments: tortgt readout error?
3. *run 33029, snarl 146819, October 30 2005, t=1130631794*
   - **Problem:** pot~ 0
   - Comments: Difference in time between spill and beam information ~ 0.67 s. Incorrect information?
4. *run 33141, snarl 220622, November 15 2005, t=1132045304*

- **Problem:** $\Delta t \sim 3700$ sec
- Comments: BDP inefficiency

5. *run 33159, snarl 258273, November 18 2005, t=1132331398*
   - **Problem:** $\Delta t \sim 2$ sec
   - Comments: Missing spill in database?

6. *run 32215, snarl 25436, July 20 2005, t=1121891287*
   - **Problem:** GPS max uncertainty $> 1\mu$ sec
   - Comments: Timing system drift caused by GPS failure at ND (during electrical storm). Good event within spill trigger window, but at edge of neutrino candidate timing distribution (see figure 43).

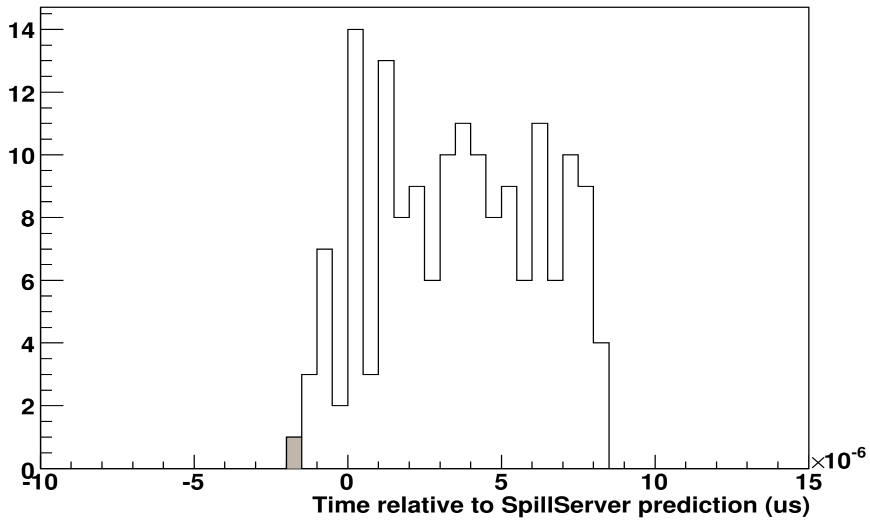


Figure 43: Snarl time relative to SpillServer prediction for the 160 neutrino candidates in the open sample. The single event that occurred during a period where the GPS worst case uncertainty was greater than 1μs is shown in grey.