A Comparison of Monte-Carlo Simulations and Data from MicroBooNE - Public Note - MICROBOONE-NOTE-1014-PUB

The MicroBooNE Collaboration

July 5, 2016

Abstract

This note presents a first comparison of data and Monte-Carlo (MC) simulation at the MicroBooNE experiment, a surface-level liquid argon time projection chamber (LArTPC) located in the Booster Neutrino Beam (BNB) at Fermi National Accelerator Laboratory. Before any analysis can be validated, it is important to ensure understanding of both the performance of the detector and the reconstruction algorithms. Comparing data with simulation can contribute to this effort.

1 Introduction

MicroBooNE is a liquid argon time projection chamber (LArTPC) located in the Booster Neutrino Beam (BNB) at the Fermi National Accelerator Laboratory. Its flagship analysis will be the investigation of the low energy excess observed by the MiniBooNE experiment [1, 2], however it will also carry out important work on neutrino cross section measurements on argon. The first of these will be the muon neutrino charged current (ν_μ CC) inclusive cross section, for which the reconstruction chain and current selection procedures are outlined in reference [3]. This note is intended to act in support of the ν_μ CC inclusive effort, and for this reason many of the distributions found here correspond to variables used in that analysis.

This is a first comparison of data and Monte-Carlo (MC) simulation. There is generally good agreement between the two, however there are several areas with large disagreements. Many of these discrepancies are thought to be due to known issues such as un-modeled noisy wires and space charge effects, and it is also possible that the modeling of the angular distribution of cosmogenic tracks in the simulation is incorrect. Further detail is supplied on each of these topics

in the body of the note.

All plots in this note are area normalized such that the two data sets (see Table 1) can be properly compared. Error bars, when shown, are statistical only.

The reconstruction chain, outlined in reference [3], can roughly be separated into two parts. The cosmic pass is the first part of the reconstruction and is used to remove cosmogenic tracks, which is done by a "geometrical tagging" of tracks that are reconstructed as through-going. During the cosmic pass, all tracks are reconstructed using the pandoraCosmic algorithm (the details of which can be found in references [4] and [5]) and the trackkalmanhit algorithm, which employs a Kalman filter for track fitting. All hits that are associated with through-going tracks in either (or both) algorithms are tagged as cosmogenic tracks and are removed. The neutrino pass can then be run on the remaining hits. Further cosmic removal is performed downstream of the reconstruction using the optical system (see, for example, reference [3]).

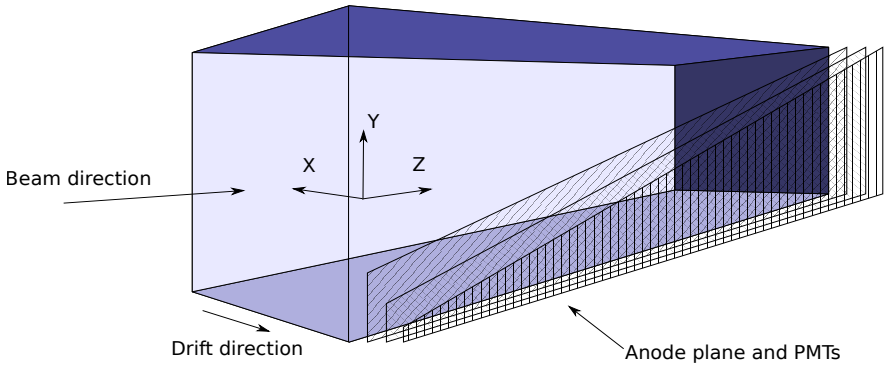


Figure 1: The MicroBooNE co-ordinate system. The three wire planes are vertical (collection plane) and at ± 60 degrees to the vertical (induction planes). The dimensions of the TPC are 256.35 cm × 233 cm × 1036.8 cm (x × y × z). The fiducial volume of the detector is 236.35 cm × 203 cm × 1026.8 cm.

The orientation of the axes in the following plots is standard in LArSoft, and makes a right handed co-ordinate system: the x coordinate (256.35 cm) points along the negative drift direction with 0 placed at the anode plane, y (233 cm) points vertically upward with 0 at the center of the detector, and z (1036.8 cm) points along the direction of the beam, with 0 at the upstream edge of the detector. It is worth noting that the readout window is longer than the time taken for electrons to drift the distance from the cathode plane to the anode plane, and so the plots pertaining to the x-direction cover several drift

windows. The effect of this is that the x-direction plots run from x $\simeq -45$ cm to x $\simeq 305$ cm, where values less than 0 cm and greater than 256.35 cm correspond to activity in the frames previous to and following the trigger, respectively. A visual representation of the co-ordinate system is shown in Figure 1.

1.1 Data and MC files

The CORSIKA [6] generator was chosen to simulate cosmic rays due to the many improvements it makes over the CRY generator, as outlined in reference [7]. It reproduces the cosmogenic background observed at MicroBooNE by simulating cosmic rays with multiple primary particle types and allows simulation at the MicroBooNE altitude. CORSIKA is paired with FLUKA to model the cosmic flux below 50 MeV. The generated events are then propagated using GEANT4 [8], and passed through the detector simulation stage. The detector simulation attempts to simulate the detector response as precisely as possible, meaning that the current state of the detector is reflected, including known dead wires. Other known effects, such as the space charge effect [9] and correlated noise [10] are not currently modeled in the simulation, but work is currently being undertaken to include them in future versions of the MC simulation.

The following plots have been made using beam-off data (meaning that this is a cosmogenic sample) and an MC simulation of an in-time cosmic sample. The data comprises of only "good runs" where the detector was considered to be stable, as outlined in reference [11]. The CORSIKA in-time sample has been chosen such that it only admits events in which at least one cosmic ray passes the software trigger[1], meaning that the cosmogenic tracks may mimic a neutrino interaction. Here, an event is defined to be the whole readout window. This is comparable to the off-beam sample which is being used which was also passed through the software trigger. The number of events in each sample is listed in Table 1.

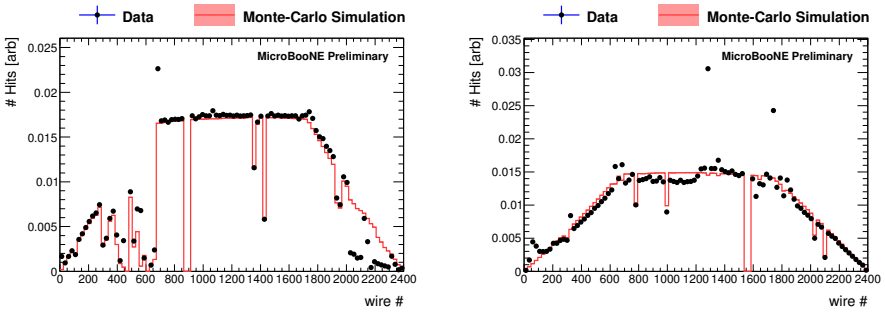
Type	# Events
off-beam data	388,471
CORSIKA in-time	29,300

Table 1: Number of events in the two samples used in this note.

2 Detector status

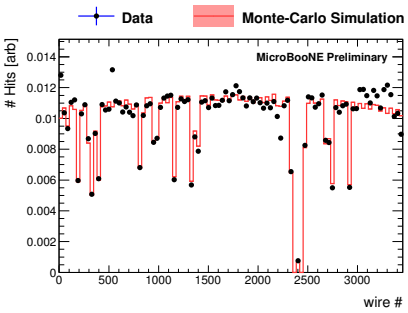
The anode plane in MicroBooNE is instrumented with three planes of wires: two induction planes (U and V, each with 2400 wires), and one collection plane

[1]The software demands that there is a flash in coincidence with the 1.6 μs beam window. More information regarding this can be found in Section 5 of this note.



(a) U plane

(b) V plane



(c) Y plane

Figure 2: Plots of the number of hits on each wire in the two induction planes (a and b) and the collection plane (c).

(Y, with 3456 wires). The induction planes are held at a bias voltage such that ionization electrons pass the wires inducing a signal on the wires. The collection plane is also held at a bias voltage such that electrons are collected on this plane, producing a signal on the wires.

The plots in Figure 2 show the state of the modeling of the dead, chirping, noisy and mis-configured wires – which are known to affect the data distributions – in the MC simulation. It is currently known that the "chirping", due to saturation of the application specific integrated circuits (ASICs) and noisy wires are not modeled in the simulation, and are also not masked in the data, meaning differences are to be expected. The sloped structure of the U and V planes is a geometrical effect due to the fact that these are the planes of wires at

± 60 degrees to the vertical, meaning that wires close to the end of the detector are shorter, and so register fewer hits.

The increase in the data over the MC simulation at low wire number in the U plane (Figure 2(a)) is due to noisy wires. For these wires more hits will be seen than are expected from the simulation, leading to an excess in data over MC. There is also a problem area at high wire number in the U plane caused by a mis-configuration of some of the ASICs, leading to a lower number of hits in the data than expected from the simulation. The jagged structure seen around wire number 500 in this plane is known to be due to dead wires.

There is also a clear discrepancy between data and MC simulation in the V plane. Two bins around wire numbers ∼1300 and ∼1800 show a large increase in the data over the MC. The cause of this is known to be due to particularly noisy wires. The dip in the data towards the center of the TPC is caused by ASIC saturation, meaning that there is significant dead time on the wires in this region, leading to a reduction of hits in data. Finally, the surplus in the data near to the noisy channel at wire number ∼1300 is due to a range of channels having anomalous signals, thought to be due to ionization electrons in the large dead region in the Y plane being collected on the V plane. These signals get deconvolved into spiky structures which have extra hits, leading to the peak in this region.

Finally, the Y plane in Figure 2(c) shows where the large (known) dead region is, at around wire ∼2400. There are also a number of spikes at low wire number in the Y plane and this is thought to be due to noise.

There are two features of these plots which are not yet accounted for, and that require further investigation: the surplus in data at very low wire number in the V plane, and the surplus in data at high wire number in the Y plane.

A more in-depth discussion of some of the issues presented in this section can be found in references [10] and [12].

3 Tracking Algorithms

3.1 The Cosmic Pass

The two algorithms addressed in this section are pandoraCosmic [4, 5] and trackkalmanhit, as these are the algorithms which feature in the cosmic pass part of the reconstruction, as previously discussed. The pandoraCosmic algorithm assumes that tracks are down-going (as most tracks are of cosmic origin) and so it defines the track start position, and its associated vertex, to be the highest end of the reconstructed track. The trackkalmanhit algorithm does not assume a specific direction, and so the tracks are roughly evenly divided between up-

and down-going. In practice, the algorithm is given a collection of tracks sorted by drift time, and the track start positions are chosen to be the end with the lowest value of x. This determines that the resulting tracks travel away from the anode plane.

3.1.1 Track start and track end x-position

The plots in Figures 3 and 4 show the starting and ending positions of tracks in the x-direction, respectively. As the x-direction acts as an effective time coordinate, the plots here have a range corresponding to one readout window, and not to the detector width as might be expected. This aids in cosmic rejection as tracks which are outside of the beam window can immediately be rejected.

We effectively see the sum of two distributions here. The peaks at ~ 0 cm and ~ 250 cm (along with the slightly raised flat distribution between these points) are caused by tracks which have been triggered on, meaning that enough light has been produced in the detector to allow determination of the interaction time with respect to the beam trigger (t_0). This allows the precise determination of the track position within the detector. The peaks at ~ -50 cm and ~ 300 cm are caused by non-triggering tracks. These tracks do not produce enough light in the detector to get a precise t_0, and so the distribution of non-triggered tracks is effectively random in the x direction. Because there is generally only one triggering track per readout window, and many non-triggering tracks, the triggering track distribution is a much smaller contribution to the overall distribution. The data and the MC simulations show very good agreement for these plots. There are discrepancies between the data and the MC at the edges of these distributions which are currently unexplained.

The differences between the algorithms in the start and end positions here are due to the different assumed directionality in the two algorithms. The pandoraCosmic algorithm does not see any preferred direction in x, while the trackkalmanhit algorithm sees most tracks entering the detector through the anode plane and exiting through the cathode plane.

3.1.2 Track start and track end y-position

The plots in Figures 5 (6) are heavily biased towards the top (bottom) of the detector for the pandoraCosmic algorithm, as is expected when considering a cosmogenic sample. The trackkalmanhit algorithm assigns most tracks to be traveling away from the anode plane, as previously discussed. For both algorithms the shoulders that appear in the data but not the MC are thought to be due to the effect of space charge, which is not currently modeled in the simulation. The space charge effect is due to the accumulation of slow-moving argon ions which can be modeled as a region of positive charge in the center of the detector, shifted towards the cathode plane in x. This has the effect of modifying

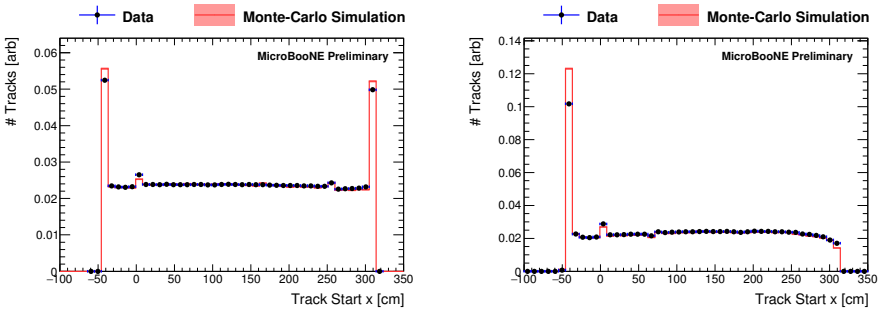


Figure 3: Plots showing the starting position of tracks in the x direction using the pandoraCosmic algorithm (left) and the trackkalmanhit algorithm (right). The region between 0 cm and 256.35 cm is the triggered drift frame, with x $<$ 0 cm and x $>$ 256.35 cm being for the earlier and subsequent drift frames, respectively.

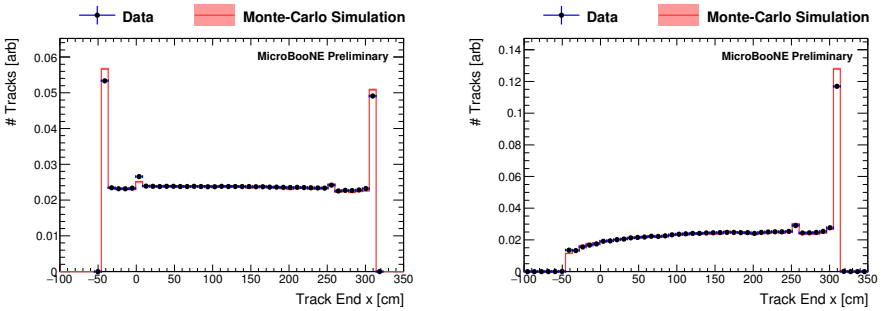


Figure 4: Plots showing the end point of tracks in the x direction using the pandoraCosmic algorithm (left) and the trackkalmanhit algorithm (right). The region between 0 cm and 256.35 cm is the triggered drift frame, with x $<$ 0 cm and x $>$ 256.35 cm being for the earlier and subsequent drift frames, respectively.

the electric field such that tracks appear to be shortened, rotated, and bowed towards the cathode plane. Further information about the space charge effect can be found in reference [9].

The preference for the top of the detector for both the start and end point distributions for the trackkalmanhit algorithm can be explained by the existence of a large number of short tracks near the top of the detector. As this is a cosmic sample, it is expected that a majority of tracks will enter the TPC at the top of the detector, and many of these stop shortly afterwards. It seems probable that this is a significant contribution to this short track excess, but whether this completely explains the effect requires further study.

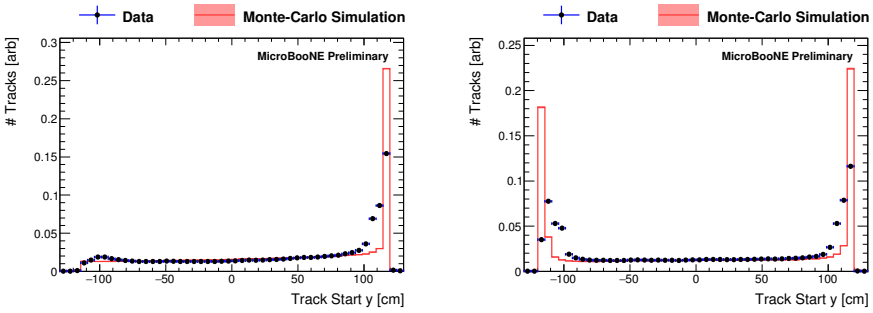


Figure 5: Plots showing the starting position of tracks in the y-direction using the pandoraCosmic algorithm (left) and the trackkalmanhit algorithm (right).

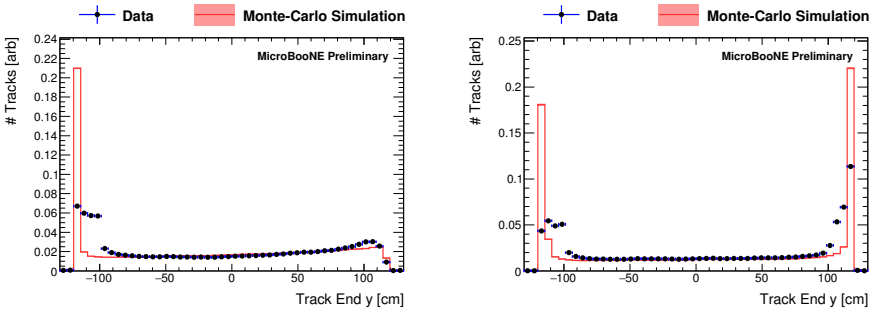


Figure 6: Plots showing the end point of tracks in the y-direction using the pandoraCosmic algorithm (left) and the trackkalmanhit algorithm (right).

3.1.3 Track start and track end z-direction

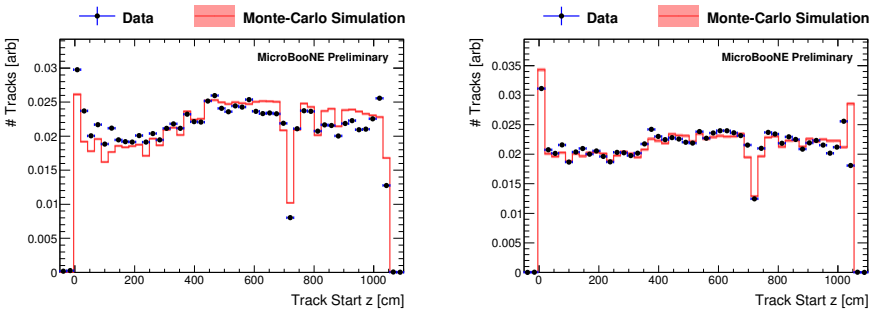


Figure 7: Plots showing the starting position of tracks in the z-direction using the pandoraCosmic algorithm (left) and the trackkalmanhit algorithm (right).

The track start and end position plots in Figures 7 and 8 show a dip in number of tracks at between $z \simeq 700$ cm and $z \simeq 750$ cm. This is due to a

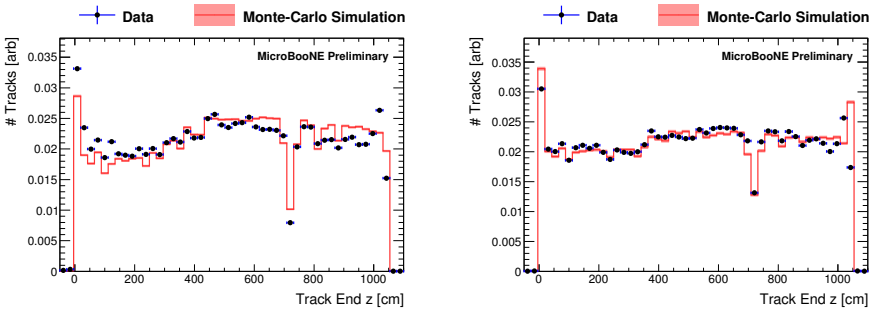


Figure 8: Plots showing the end point of tracks in the z-direction using the pandoraCosmic algorithm (left) and the trackkalmanhit algorithm (right).

significant region of dead wires in the detector (and modeled in the simulation) in the Y plane located at this point, as shown in Section 2. There is also a slight reduction in signal toward the $z = 0$ cm part of the plot, and this is again due to dead wires. The dead regions in this area are typically due to a small number of wires and so the algorithms can be trained to traverse the gap. This is why the dip in signal is not as severe as in the $z \simeq 700$ cm case. There are significant differences at the start and end of the detector in both algorithms, and the cause of this is currently not known, although this could be related to the effect seen at the edges of the detector in Figure 3. The space charge effect could go some way to explain these differences, but it is not expected to mitigate the effect completely. Considering the known issues surrounding dead and chirping wires (see Section 2), these plots show relatively good agreement.

3.1.4 Theta

The theta angle is defined such that 0 is the beam direction, and tracks at $\theta = \pi/2$ correspond to tracks traveling in only the x-y plane (not in the z direction). Correspondingly, the cosine of this angle runs between 1 (the beam direction) and -1 (negative beam direction). There is a 30 cm cut on the track length in these plots in order to mitigate the contribution of delta rays. The peaks at $\cos(\theta) \simeq \pm 0.85$ in the pandoraCosmic plot (Figure 9) correspond to the angles of the wires on two of the planes of readout wires (the U and V planes). It seems likely that these may be caused by noise in the detector which is not modeled in the MC simulation, although whether this effect is localized to specific wire numbers, or is constant across the detector, is not known. The hit requirement for pandora is set to be low, and any noisy wire in one plane may produce a track if there is any correlated activity in the other two planes. This does not happen when the trackkalmanhit algorithm is used, as more hits are required to start a track during the track fitting process[2]. For both algorithms

[2]pandoraCosmic requires far fewer hits to make a track than trackkalmanhit. The pandoraCosmic algorithm can fit a track from as little as 1 hit per cluster per view. After looking at the data, it is suspected that this threshold is set to be too low. The trackkalmanhit algorithm

there is an increase in the data above the simulation for forwards going and backwards going tracks. One possible reason for this may be noise in the beam direction which is known to exist in MicroBooNE. The slight increase in the MC above the data in several areas over both plots is currently under study.

There appears to be a reduction in tracks traveling at $\cos(\theta) = 0$, however this effect is not physical and is due to several factors. Tracks parallel to wire planes are known to be difficult to reconstruct (as one view of the track is essentially lost). In addition, these bins include tracks which are traveling only in the x direction, which adds further difficulty to the task. Pandora seems to do a much better job at reconstructing tracks in this region, although care should be taken when drawing conclusions from this, as this apparent improvement over trackkalmanhit may just be contributions from noise. Until the effects of noise on the reconstruction are understood, no definite statement can be made.

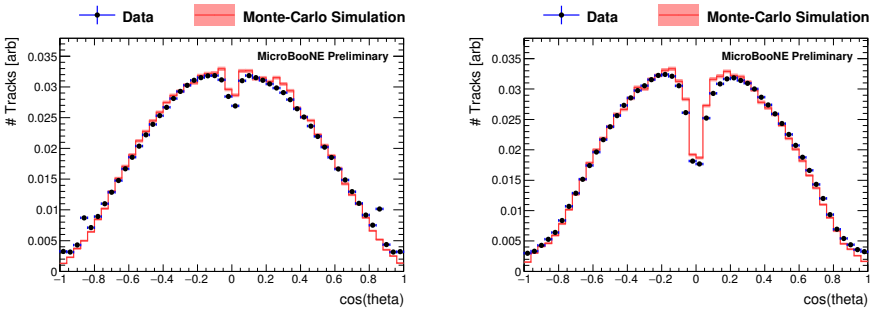


Figure 9: Plots showing the θ angular distribution of the track with respect to the beam direction using the pandoraCosmic algorithm (left) and the trackkalmanhit algorithm (right). There is a cut on the track length for both of these plots at 30 cm in an effort to remove delta rays.

3.1.5 Phi

Phi is defined as the angle around the beam direction, where 0 and $\pm\pi$ correspond to the x axis, and ϕ = -$\pi/2$ corresponds to downwards going tracks. The plots in Figure 10 very clearly show the different assumptions that go into the pandoraCosmic and trackkalmanhit algorithms. The pandora plot shows the tracks mostly going downwards due to the assumption that everything in the cosmic pass is downwards going, while the trackkalmanhit algorithm shows around half of tracks having their direction reversed. There does seem to be a slight enhancement of the data over MC at each of the peaks in both algorithms, which correspond to tracks traveling close to vertically. This underestimation of MC of the downwards going flux, along with MC-data differences seen at the edges of the detector in Figure 3 and similar distributions, may be a subtle

requires a larger number of hits in order to form a good "seed" for starting a track.

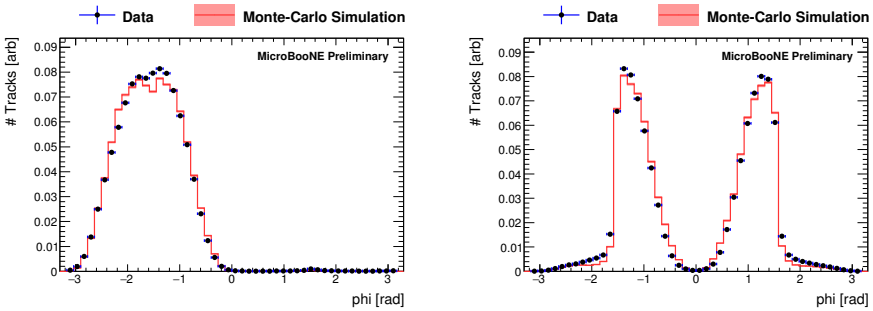


Figure 10: Plots showing the ϕ angular distribution of tracks with respect to the beam direction using the pandoraCosmic algorithm (left) and the trackkalmanhit algorithm (right). There is a cut on the track length for both of these plots at 30 cm in an effort to remove delta rays.

indication that the angular distribution in the MC simulation is not modeled correctly. Equally, this could be caused by differences in the tracking efficiencies between data and simulation.

3.1.6 Track length

Pandora reconstructs a large number of very short tracks (such as delta rays and other stubs, along with noise), most of which is cut out by track length cuts in the ν_μ CC inclusive analysis [3]. The plots here contain only tracks which have a length of 30 cm or greater in order to easily compare the two algorithms. The peaks at between $\sim$ 230 cm and $\sim$ 250 cm in both plots correspond to through-going cosmic rays traveling the whole height (or width) of the detector. The data shows a less severe peak at $\sim$ 230 cm, but this can be explained by the space charge effect which is discussed in relation to Figures 5 and 6.

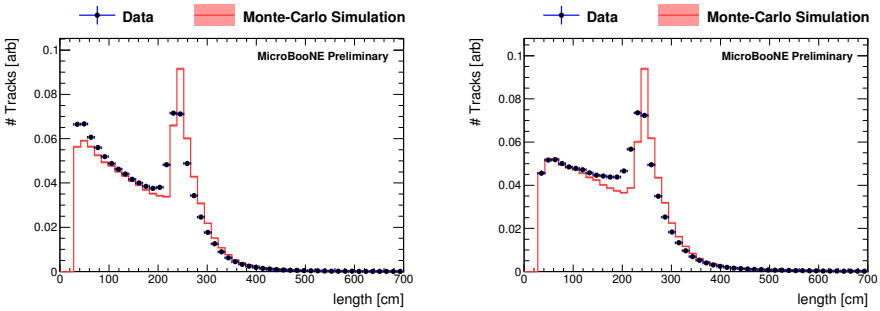


Figure 11: Plots showing the length of tracks using the pandoraCosmic algorithm (left) and the trackkalmanhit algorithm (right). This plot is cut at 30 cm so as to allow an easy comparison of the two plots.

3.2 The Neutrino Pass

The two algorithms that have been used to produce the plots in this section, pandoraNu and pandoraNuPMA, are run after the cosmic pass, leading to an expected reduction in statistics. The pandoraNu algorithm is run on the hits that remain after the cosmic pass, producing 2D clusters. From here, 3D tracks and vertices can be found internally, producing the pandoraNu output. Alternatively, these clusters can be passed to the projection matching algorithm (PMA) [13], which will run track and vertex finding on the pandoraNu clusters (producing pandoraNuPMA). The PMA algorithm can also be run on the hits which have been passed through the linecluster clustering algorithm, again producing 3D tracks and vertices (this is called pmtrack). This is done as part of the ν_μ CC inclusive analysis [3], but only the vertices are kept in this case, and so the tracking part of the pmtrack algorithm is not considered here. A schematic of these reconstruction channels can be found in Figure 2 in reference [3].

These algorithms are designed to deal with beam-generated interactions, which are generally forward going, and so the reconstructed vertex is usually placed at the end of the track that is closest to the upstream end of the detector.

Note that, as previously mentioned, there are no neutrino events in these data sets. These distributions represent the tracks that have not been tagged as cosmics by the cosmic pass, and so constitute a potential background to the real neutrino interactions. As previously stated, the cosmic tagging done in the cosmic pass is purely geometrical, and does not make use of the optical information from the event, meaning that further cosmic rejection can be performed downstream of the reconstruction.

3.2.1 Track start and track end x-position

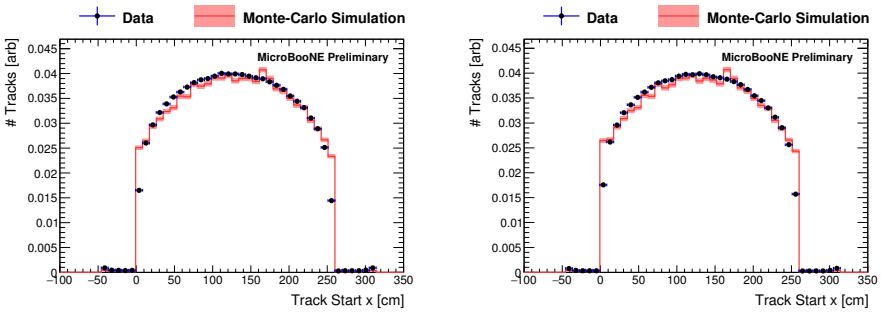


Figure 12: Plots showing the starting position of tracks in the x direction using pandoraNu algorithm (left) and the pandoraNuPMA algorithm (right).

The plots in Figures 12 and 13 show the starting and ending positions of tracks in the x-direction. Unlike the previous plots (Figures 3 and 4), almost all

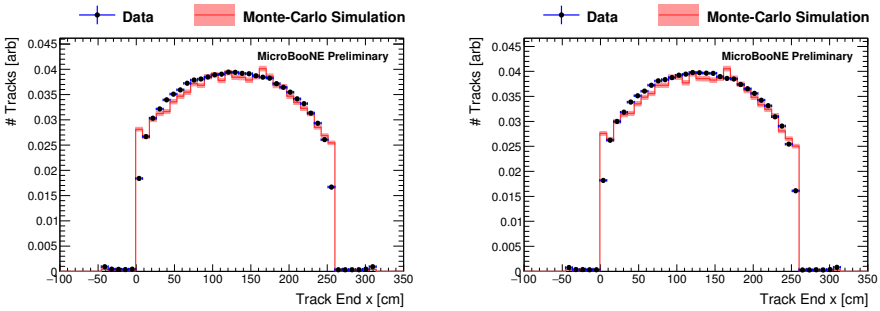


Figure 13: Plots showing the end point of tracks in the x direction using the pandoraNu algorithm (left) and the pandoraNuPMA algorithm (right).

of the start and end points here are located within the dimensions of the TPC. This is because the cosmic rejection has already been performed and through-going tracks have been removed. One feature of interest is that there seems to be an enhancement in the MC over the data at the edges of the detector. This is currently under study, however it is not expected to impact the distributions in the ν_μ CC inclusive note due to the fiducial cut of a 20 cm boundary in the x direction imposed during both selections of that analysis. The small number of tracks outside of the TPC boundary seems to be due to short tracks created by hits left over after the cosmic pass. Because these tracks are not through-going, they are not removed during the cosmic pass. The small peak in MC at $\sim$ 180 cm is not currently understood.

3.2.2 Track start and track end y-position

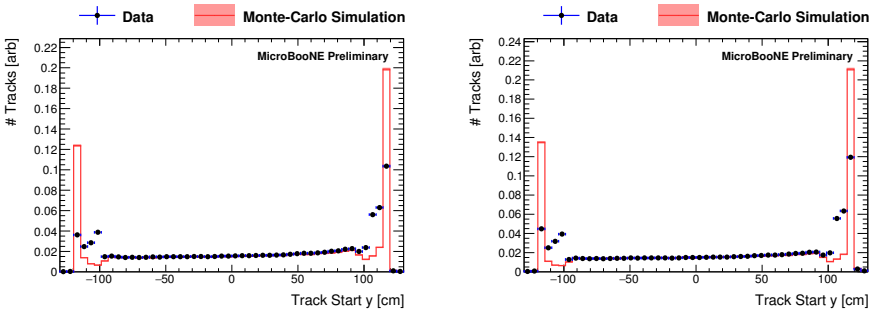


Figure 14: Plots showing the starting position of tracks in the y-direction using pandoraNu algorithm (left) and the pandoraNuPMA algorithm (right).

The plots in Figures 14 and 15 suffer from the same space charge issue discussed in relation to Figures 5 and 6. The dip in the MC distribution towards the edges of these distributions is understood to be an acceptance effect from

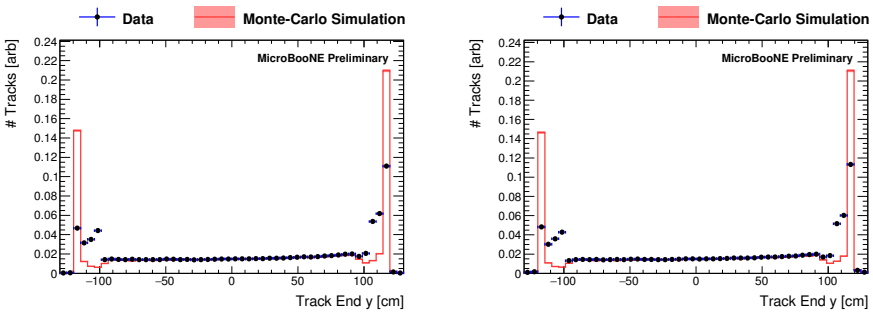


Figure 15: Plots showing the end point of tracks in the y-direction using the pandoraNu algorithm (left) and the pandoraNuPMA algorithm (right).

removal of through-going cosmics. Although there are significant differences between the data and the MC simulation here, this is not expected to affect the distributions in the ν_μ CC inclusive note [3] because of the fiducial volume cut of a 20 cm boundary imposed in the y-direction during both selections used in the analysis.

The preference for the top of the detector for both the start and end point distributions for the plots in Figures 14 and 15 can be explained by a large number of short tracks near the top of the detector. As this is a cosmic sample, it is expected than a majority of tracks will enter the TPC at the top of the detector, and many of these should stop shortly afterwards. It seems probable that this is a significant contribution to this short track excess, but whether this completely explains the effect requires further study.

3.2.3 Track start and track end z-position

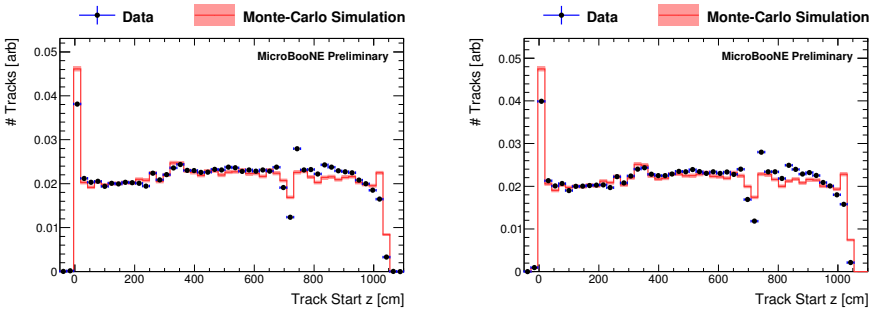


Figure 16: Plots showing the starting position of tracks in the z-direction using pandoraNu algorithm (left) and the pandoraNuPMA algorithm (right).

The track start and end position plots in Figures 16 and 17 show the same

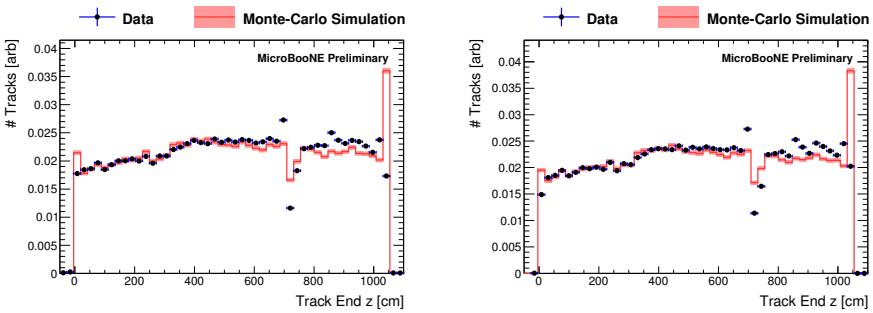


Figure 17: Plots showing the end point of tracks in the z-direction using the pandoraNu algorithm (left) and the pandoraNuPMA algorithm (right).

dip in number of tracks between $z \simeq 700$ cm and $z \simeq 750$ cm as in Figure 7 corresponding to the large dead region in the Y plane which is modeled in both the MC and the data. It seems that both algorithms deal with this drop in signal better in the MC than in the data. There are significant differences between the data and the MC simulation at the start and end of the detector, as seen during the cosmic pass (Figure 8).

It might be expected that the end peaks in the distribution would disappear here, for the same reason that we see no peaks in Figures 12 and 13, however this is not what is observed in these histograms. This can possibly be explained by tracks that would usually be through-going being reconstructed as two tracks due to the large dead region at z $\simeq$ 700 cm, and other dead regions. This would mean that the tracks are not removed in the cosmic pass. The end of the detector has mis-configured ASICs meaning that the track start and end positions may be reconstructed as being upstream of their true positions. It is thought that this is the reason for the enhancement of the simulation over the data at the end of the detector. As previously stated, these discrepancies are not expected to change the distributions in reference [3] due to the fiducial volume cut of a 10 cm boundary in the z direction present in that analysis.

3.2.4 Theta

Figure 18 shows the θ distribution where, again, θ is defined such that $\theta = 0$ is the beam direction, and $\theta = \pi/2$ corresponds to tracks traveling in only the x-y plane. Correspondingly, plotting the cosine of theta means that 1 is defined to be forward going and -1 is defined to be backwards going. There is a 30 cm cut on the track length here in order to mitigate the contribution of delta rays. These plots indicate that most of the tracks are downwards going (which is expected, as they are cosmic in origin), but are also reconstructed as forwards going due to the direction assumption mentioned in the preamble to this section. The difference between the two algorithms is notable: data and MC are found to be in much better agreement when using the pandoraNuPMA algorithm than

when using the pandoraNu algorithm. This is not currently understood. The peaks which align with the U and V planes in the pandoraCosmic algorithm (which are prevalent in Figure 9) make another appearance here although the effect is much less prominent. The increase in the data above the MC simulation for $\cos(\theta) = 1$ may be due to noise, or due to broken through-going tracks.

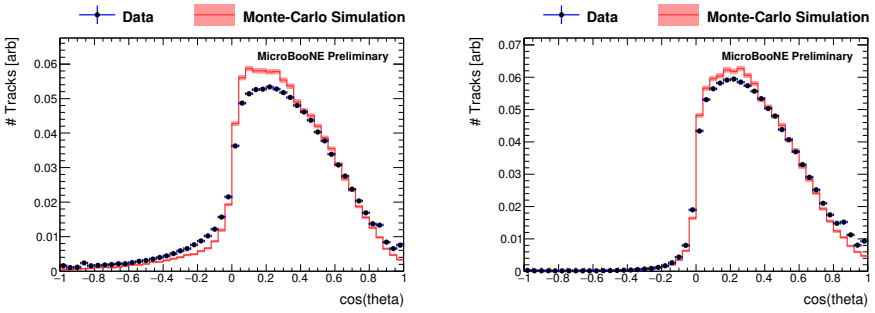


Figure 18: Plots showing the θ angle with respect to the track start position using pandoraNu algorithm (left) and the pandoraNuPMA algorithm (right). There is a cut on the track length for both of these plots at 30 cm in an effort to remove delta rays.

3.2.5 Phi

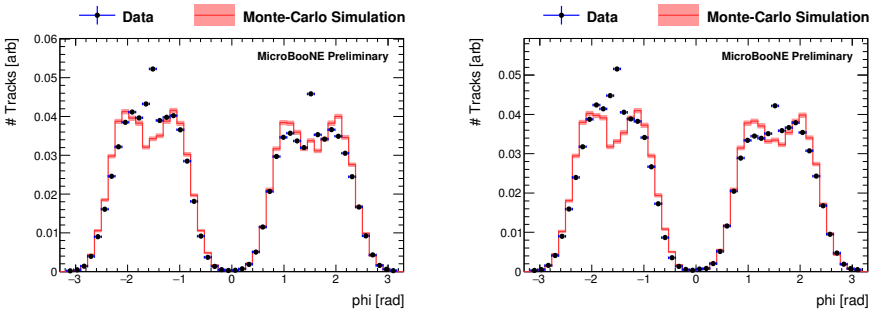


Figure 19: Plots showing the ϕ angular distribution with respect to the track start position using pandoraNu algorithm (left) and the pandoraNuPMA algorithm (right). There is a cut on the track length for both of these plots at 30 cm in an effort to remove delta rays.

Phi is again defined as the angle around the beam direction, where 0 and $\pm\pi$ correspond to the x axis, and ϕ = -$\pi/2$ corresponds to downwards going tracks. The plots in Figure 19 show that the tracks are reconstructed as upward- or downward-going, as is expected from a cosmic data sample. A deficit for com-

pletely upwards or downwards going tracks would be expected due to difficulties in reconstructing such a track, but this is clearly not the case in the data. This is thought to be due to an inability to properly capture charge deposition for tracks that are parallel to one of the wire planes.

3.2.6 Track length

As both of these algorithms are using pandoraNu, they both reconstruct a large amount of short tracks which correspond to delta rays and other stubs. To be able to compare the algorithms in a reasonable way, a 30 cm cut has been introduced, as done previously. The peak at $\sim$ 230 cm which corresponded to tracks traveling through the whole height (or width) of the detector in Figure 11 has been reduced to a kink in the distribution, as through-going cosmic muons have been removed by the cosmic pass. These distributions show remarkable agreement. The small discrepancies between the data and MC are expected to be removed when the space charge effect is properly accounted for.

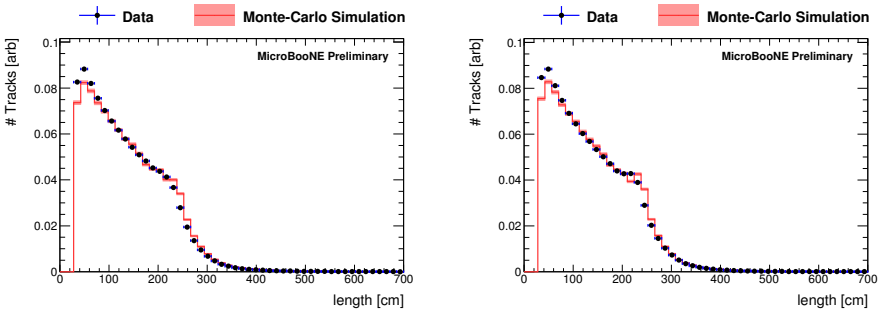


Figure 20: Plots showing the length of tracks using pandoraNu algorithm (left) and the pandoraNuPMA algorithm (right).

4 Vertexing Algorithms

The plots in this section show the positions of vertices in the neutrino pass. The two algorithms used for this in the ν_μ CC inclusive analysis are pandoraNu and pmtrack. PandoraNu assumes that all of the tracks it reconstructs are due to neutrino interactions, and it thus assigns a vertex to every reconstructed particle, including the neutrino. Pmtrack, on the other hand, is not a dedicated vertexing algorithm and does not reconstruct neutrino vertices. In general, the vertices for both algorithms are placed at the most upstream end of the track. Further details can be found at references [4, 5] for pandoraNu, and at reference [13] for pmtrack.

Note, as discussed in Section 3.2, that there are no actual neutrino events in the samples used here. These distributions represent the events that have not

been tagged as cosmics by the cosmic pass.

4.1 Vertex x-position

Figure 21 shows plots of the vertex x-position. They look very similar to the corresponding track start point plots in Figure 12, as expected. Data and MC generally agree quite well. There is some discrepancy at the edges, which may be related to the edge effect observed in several tracking plots presented earlier in this note. It may also have some contribution from the space charge effect. These discrepancies should not affect the ν_μ CC inclusive analysis because of the 20 cm fiducial cut in the x direction present in that analysis.

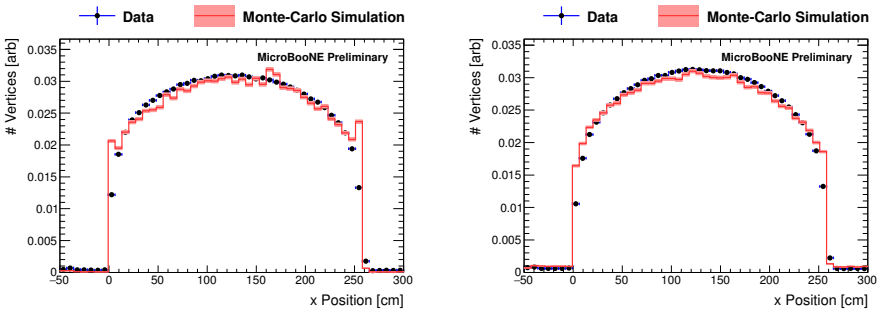


Figure 21: Vertex x-position for the pandoraNu algorithm (left) and the pmtrack algorithm (right).

4.2 Vertex y-position

Figure 22 shows plots of the vertex y-position. Again, they look very similar to the corresponding track start point plots in Figure 14 and agreement between data and MC is good. The edge effects are again due to space charge, however as previously mentioned, these are not expected to be important for distributions presented in the ν_μ CC inclusive note [3] due to the 20 cm fiducial volume cut in the y direction used in the selections of that analysis.

4.3 Vertex z-position

Figure 23 shows plots of the vertex z-position. They once again look similar to the corresponding track start point plots in Figure 16, and much of the discussion around those plots is relevant here. Data and MC agree well, including near the dead region at z $\simeq$ 700 cm. Also of note is that pandoraNu is able to mostly skip over this region, showing only a small decrease in the number of reconstructed vertices here.

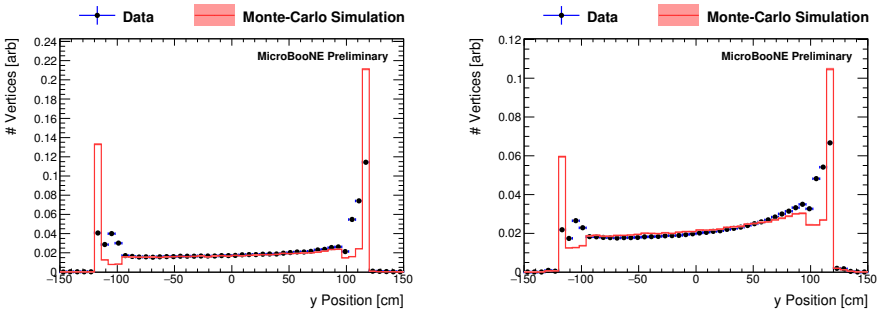


Figure 22: Vertex y-position for the pandoraNu algorithm (left) and the pmtrack algorithm (right).

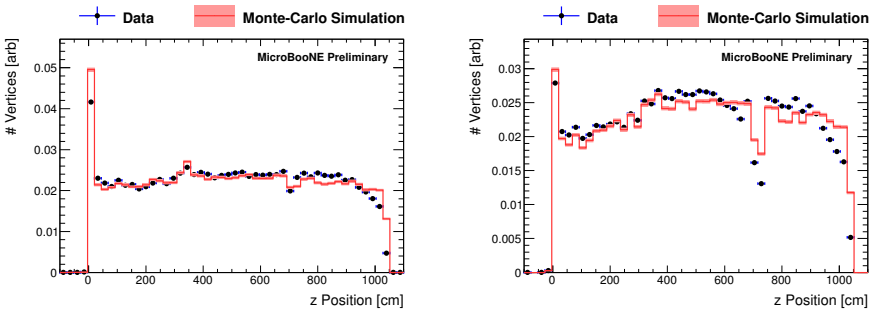


Figure 23: Vertex z-position for the pandoraNu algorithm (left) and the pmtrack algorithm (right).

5 Flash Reconstruction

In order to determine if a BNB spill possibly produced a neutrino interaction in the detector, a system of 32 PMTs is used to measure scintillation light. A "flash" consists of all of the photoelectrons (PE) detected in a narrow window of time (on the order of 100s of ns). The ν_μ CC inclusive analysis requires that flashes are produced coincident with the spill, and that they contain at least 50 PE. For more details, see Section 3.1 of the ν_μ CC inclusive note [3].

The plots in Figure 24 show the number of PE per flash and the number of flashes per event that have > 50 PE. There are slightly more events with high numbers of flashes for MC than data, but the distributions generally agree quite well.

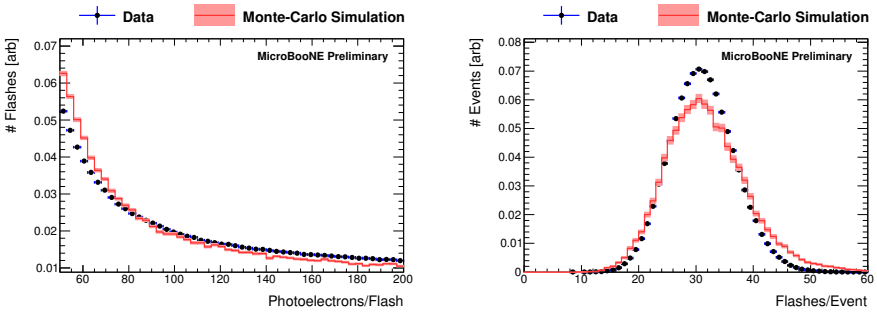


Figure 24: Number of photoelectrons per flash (left) and number of flashes per event, with a > 50 PE cut (right).

6 Calorimetry

Figure 25 shows the charge loss per unit path length (dQ/ds) in the collection plane, reconstructed using the pandoraNuPMA algorithm. The method of extracting charge distributions from the signals on the wire planes is described in detail in reference [14]. In short, the signals are first deconvolved, removing the effect of field and electronics responses in order to better estimate the true signal. Corrections are then applied for differences due to track angle and for channel-to-channel variations in signal response.

Calorimetry is an ongoing subject of study, and the agreement between data and MC in this plot shows promise. Further details of both our current understanding and plans for future developments can be found in references [10, 14].

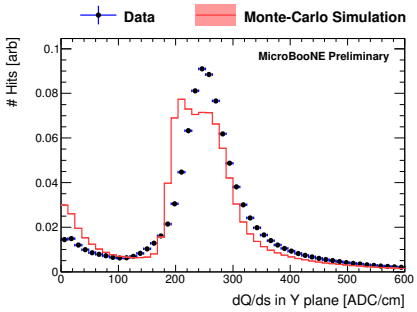


Figure 25: dQ/ds in the collection plane, using the pandoraNuPMA algorithm.

7 Conclusions

The distributions which have been presented in this note generally show very good agreement between data and MC simulation, especially considering that this is a first comparison. Discrepancies can generally be attributed to one of two effects:

- Noise/ASIC saturation and mis-configured channels currently un-modeled in the MC simulation, as outlined in Section 2.
- The space charge effect, by which slow-moving ions act to modify the electric field in the TPC, meaning that tracks appear shortened, rotated, and bowed towards the cathode plane [9].

These issues are known to the collaboration and work is being undertaken to address them. There is also the possibility that the angular distribution of cosmogenic tracks may be slightly mis-modeled in the simulation, although further study is necessary to confirm this. The large discrepancies located near the edges of the detector in the track start (and vertex) distributions are not expected to affect the ν_μ CC inclusive analysis due to the fiducial volume cut applied by the two selection schemes.

Further study of the agreement between data and MC simulation will be necessary when the known issues have been addressed, however the outlook from this first comparison is very positive.

References

[1] A. A. Aguilar-Arevalo et al. Search for electron neutrino appearance at the $\Delta m^2 \sim 1\text{eV}^2$ scale. *Phys. Rev. Lett.*, 98:231801, Jun 2007. doi: 10.1103/PhysRevLett.98.231801. URL http://link.aps.org/doi/10.1103/PhysRevLett.98.231801.

[2] A. A. Aguilar-Arevalo et al. Search for electron antineutrino appearance at the $\Delta m^2 \sim 1\text{eV}^2$ scale. *Phys. Rev. Lett.*, 103:111801, Sep 2009. doi: 10.1103/PhysRevLett.103.111801. URL http://link.aps.org/doi/10.1103/PhysRevLett.103.111801.

[3] The MicroBooNE Collaboration. Selection of charged-current ν_μ inclusive events at MicroBooNE. *MICROBOONE-NOTE-1010-PUB*, 2016. URL http://www-microboone.fnal.gov/publications/publicnotes/index.html.

[4] J. S. Marshall and M. A. Thomson. The Pandora Software Development Kit for Pattern Recognition. *Eur. Phys. J.*, C75(9):439, 2015. doi: 10.1140/epjc/s10052-015-3659-3.

[5] The MicroBooNE Collaboration. The Pandora multi-algorithm approach to automated pattern recognition in LAr TPC detectors. *MICROBOONE-NOTE-1015-PUB*, 2016. URL http://www-microboone.fnal.gov/publications/publicnotes/index.html.

[6] D. Heck, G. Schatz, T. Thouw, J. Knapp, and J. N. Capdevielle. CORSIKA: A Monte Carlo code to simulate extensive air showers. 1998.

[7] The MicroBooNE Collaboration. Cosmic Shielding Studies at MicroBooNE. *MICROBOONE-NOTE-1005-PUB*, 2016. URL http://www-microboone.fnal.gov/publications/publicnotes/index.html.

[8] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003. doi: 10.1016/S0168-9002(03)01368-8.

[9] The MicroBooNE Collaboration. Space Charge Effect Measurements and Corrections. *MICROBOONE-NOTE-1018-PUB*, 2016. URL http://www-microboone.fnal.gov/publications/publicnotes/index.html.

[10] The MicroBooNE Collaboration. Noise Characterization and Filtering in the MicroBooNE TPC. *MICROBOONE-NOTE-1016-PUB*, 2016. URL http://www-microboone.fnal.gov/publications/publicnotes/index.html.

[11] The MicroBooNE Collaboration. MicroBooNE Detector Stability. *MICROBOONE-NOTE-1013-PUB*, 2016. URL http://www-microboone.fnal.gov/publications/publicnotes/index.html.

[12] The MicroBooNE Collaboration. MicroBooNE Detector Stability. *MICROBOONE-NOTE-1013-PUB*, 2016. URL `http://www-microboone.fnal.gov/publications/publicnotes/index.html`.

[13] M. Antonello et al. Precise 3D track reconstruction algorithm for the ICARUS T600 liquid argon time projection chamber detector. *Adv. High Energy Phys.*, 2013:260820, 2013. doi: 10.1155/2013/260820.

[14] The MicroBooNE Collaboration. A Method to Extract the Charge Distribution Arriving at the TPC Wire Planes in MicroBooNE. *MICROBOONE-NOTE-1017-PUB*, 2016. URL `http://www-microboone.fnal.gov/publications/publicnotes/index.html`.