

CHAPTER 1

NOvA Test Beam Detector Calibration

The NuMI Off-axis ν_e Appearance (NOvA) Test Beam experiment [1] is a sub-experiment designed to enhance NOvA’s sensitivity to neutrino oscillation measurements by improving the understanding of particle interactions and energy deposition within the NOvA detectors. Initial studies [2] showed that, only by improving the detector calibration, the Test Beam experiment has the potential to reduce the total systematic uncertainty in the measurement of the three flavour oscillation parameters by about 10%.

The NOvA Test Beam experiment consists of a scaled down version of the NOvA detectors placed in a test beam. Using a test beam allows for the study of the response of tagged single particles with known momenta and positions within a NOvA detector. Additionally, this setup enables the determination of the energy resolution and the absolute energy scale without the use of simulation. Furthermore, it permits the comparison of responses between beam muons and cosmic ray muons, study of fibre attenuation, and validation of the NOvA calibration process. The Test Beam detector is equipped with a combination of the Near Detector (ND) and Far Detector (FD) readout electronics and filled with a range of NOvA scintillator oils, enabling a comparison of their respective performance and particle responses [3]. All these advantages require, or benefit from, the calibration of the Test Beam detector, which follows the same calibration procedure as the standard NOvA detectors (Sec. ??).

In this chapter I introduce the NOvA Test Beam experiment in Sec. 1.1, focusing on the Test Beam detector and especially on the aspects that could impact its calibration. Section 1.2 describes the new data-based simulation of cosmic muons that I developed for the Test Beam detector calibration, while Sec. 1.3 discusses the calibration of the Test Beam detector itself.

1.1 The NOvA Test Beam Experiment

The NOvA Test Beam experiment [4] consists of a scaled down version of the NOvA ND and FD, shown in Fig. 1.1, and a series of beamline detectors to measure and identify a range of particles from the MCenter beamline in the Fermilab Test Beam Facility (FTBF) [5].

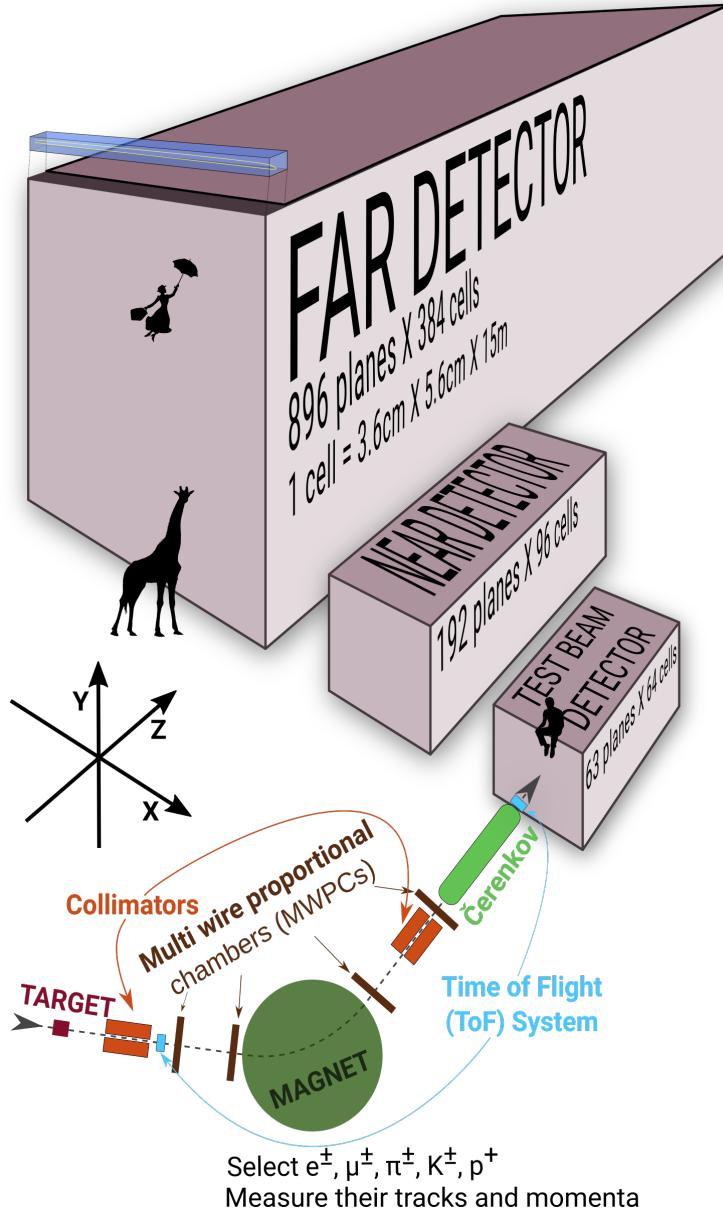


Figure 1.1: Comparison of Test Beam detector scale to the NOvA ND and FD (and a man, giraffe, or Mary Poppins). Also shown are the Test Beam beamline detectors and components (not to scale), with arrows showing the direction of the beam. The three black arrows show the orientation of the detector coordinate system.

The operation of the Test Beam detector started with commissioning runs in June

2019 and ran, with an exception of regular summer shutdowns, until July 2022, after which it was decommissioned. The Test Beam data taking is divided into *periods*, which are defined in Tab. 1.1. Period 1 only lasted for about a month and with only a half-filled detector, as explained below. It was therefore only used for detector commissioning and will not be used in any of the Test Beam physics analysis, or in the calibration.

Table 1.1: Test Beam detector data taking periods.

Period 1	June 3 rd 2019	-	July 6 th 2019
Period 2	December 5 th 2019	-	March 20 th 2020
Period 3	January 12 th 2021	-	June 27 th 2021
Period 4	November 30 th 2021	-	July 10 th 2022

Majority of the Test Beam detector and its instrumentation is identical to the other **NOvA** detectors, with a few exceptions that could have an impact on the calibration. We are going to identify and discuss these differences in this section.

Beamline

The beam for the Test Beam experiment originates from the same 120 GeV Main Injector protons used in Neutrinos from the Main Injector (NuMI), extracted once a minute in a continuous 4.2 s spill [1]. The protons are impinged on a copper target producing mostly protons and pions, which are then directed towards a second target, producing the tertiary beam of particles used in the Test Beam detector. As can be seen in Fig. 1.1, two collimators are used to direct the tertiary beam and a magnet to select the desired momentum. Particle tracking is done using the four Multiwire Proportional Chambers (MWPCs) and particle identification is done with a combination of Time of Flight (ToF) detectors and a Cherenkov detector set for electron detection.

Detector Parameters

The **NOvA** Test Beam detector consists of two 31-plane blocks, each beginning and ending with a vertical plane, with an additional horizontal plane glued in-between them to preserve the alternating pattern [6]. Each plane consists of 2 modules side-by-side, both made up of 32 cells. Each cell is 2.6 m long with an inner (without the

PVC) depth and width of 5.9 cm and 3.8 cm respectively, same as for the other **NOvA** detectors. This brings the final dimensions of the Test Beam detector to 63 planes \times 64 cells, or $2.6 \times 2.6 \times 4.1 \text{ m}^3$.

The 63 planes are numbered from 0 to 62, with even numbers corresponding to vertical planes and odd numbers to horizontal planes. Cells are numbered 0 to 63, going from bottom to top for horizontal planes and left to right, when facing the front of the detector, for vertical planes.

The detector coordinate system is illustrated in Fig. 1.1. It is centred with $(0, 0, 0)$ in the centre of the first plane [7]. The x axis runs left to right when facing the front of the detector, y axis from bottom to top, and z axis goes along the beam direction from front to the back of the detector. Position within each cell (w) is aligned with the x (y) axis for the horizontal (vertical) cells, with $w = 0$ centred in the middle of each cell. The exact geometry of the Test Beam detector was measured in several alignment surveys and is saved in gdml files [8].

In the past we encountered an issue when trying to align the Test Beam detector with the beamline measurements by rotating the detector. This broke several assumptions within the Test Beam geometry [7] and manifested as uncalibrated cells in the back of the detector [9]. This was fixed by realigning both the detector and the beamline separately, based on the last alignment survey, measured during the decommissioning of the detector.

Scintillator

Test Beam used a combination of the leftover **ND** and **FD** production scintillator oils and the oil drained from the **NOvA** Near Detector on the Surface (NDOS) test detector. The used scintillator oils also differ in the way they were stored since the **ND** and **FD** filling, or the **NDOS** draining, which apparently impacted its quality. These factors have a significant effect on the energy deposition within them. The distribution of individual scintillator oils and the relative difference in their energy response can be seen in Fig. 1.2.

There are four distinct samples of **NOvA** scintillator oil used in the Test Beam detector:

1. Mixed **ND** production oil and **NDOS**-drained oil stored in a tanker and four

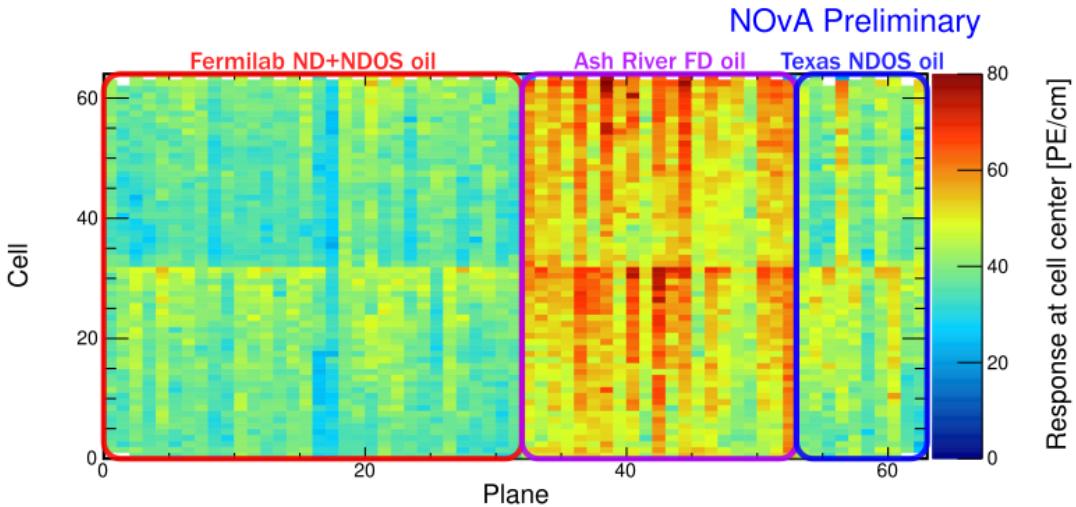


Figure 1.2: Uncorrected energy response in the centre of cells across the Test Beam detector showing a clear distinction between the different scintillator oils, labelled with coloured boxes and descriptions.

tanks outside in Fermi National Accelerator Laboratory (FERMILAB) [10];

2. Separate **ND** production oil and **NDOS**-drained oil stored underground in barrels at the MiniBooNE¹ cavern [2];
3. **FD** production oil stored inside in Ash River in ‘totes’ under several layers of black plastic [12];
4. **NDOS**-drained oil stored mainly inside at Texas A&M University and University of Texas at Austin [13, 14].

The original plan [15] was to only use the tanker/tank scintillator (sample #1). First tests showed acceptable results and the tanker oil was used to fill out almost the entirety of the first block of the detector (first 32 planes) [10]. However, when the oil from tank #2 was loaded into the tanker, it became extremely cloudy and unusable, possibly due to contamination with water accumulated at the bottom of the tanks. The rest of the first block was therefore topped up with high quality scintillator from **NDOS** (sample #2). This is labelled as ‘**FERMILAB ND+NDOS oil**’ in Fig. 1.2.

The first 21 planes of the second block (planes 32 to 52) were filled with the **FD** production scintillator shipped in from Ash River (sample #3) [16]. These planes were again topped up with the **ND+NDOS** scintillator (sample #2).

¹MiniBooNE [11] is a **FERMILAB** experiment located close to the **NOvA ND**

The last 10 planes (planes 53 to 62) [16] were filled with the ‘Texas’ scintillator (sample #4), which has higher light yield than the one from the tanker, but lower than the Ash River one [13].

In total, the Test Beam detector is filled with 5418 gallons of scintillator oil with a weight of approximately 28.6 tons [6].

Readout

The Test Beam detector uses in total 126 Front End Boards (FEBs), each reading out signal from 32 cells [6]. The readout is located on the top and right side (when looking at the front) of the detector. 118 FEBs are version 4.1, same as in the FD, and 8 FEBs, located in pairs on planes 16, 17, 48 and 49, are version 5.2, same as in the ND. As was described in Sec. ??, the ND FEBs are designed to read out data at a faster rate than the FD FEBs and using a mix of FEB types allows us to study the difference in their response and to validate both versions in the same environment [17].

Environment

COMMENT: Not sure what tense to use here. Past as the temperature was monitored, or present as I used in all the rest of this section? The enclosure that houses the Test Beam detector is made up of a concrete platform covered by a metal semi-cylindrical roof. Therefore, unlike the ND and FD, the Test Beam detector does not have any overburden to shield it from cosmic particles, affecting their rate and energies inside the detector. The temperature and humidity are controlled by a humidity, ventilation and air conditioning control system and monitored by a range of sensors. The temperature was kept to around 20 °C and within the range of 18 – 22 °C, except for about three months in the beginning of Period 3 data taking, when it was kept to within 16 – 20 °C. The readout electronics were turned off when the dew point reached 10 °C to limit humidity related noise in the FEBs.

Underfilled Cells Issue

The Test Beam detector is slightly tilted around the z axis by about 0.7° towards the readout. This caused the top cells of both modules of all the horizontal planes (cells 31 and 63) to be underfilled, creating an air bubble on the left side of the detector

and severely affecting the energy response in those cells [17]. This was fixed [18] during the period 3 running by adding extensions to the filling ports and overfilling the horizontal cells with the **ND+NDOS** scintillator (sample #2 from the scintillator description). More details on this issue and its effects and on how it was handled in calibration are detailed in Sec. 1.3.5.

1.2 Data-based Simulation of Cosmic Muons

The standard **NOvA** calibration procedure described in Sec. ?? uses the Cosmic-Ray Shower Generator (**CRY**) Monte Carlo (MC) generator (see Sec. ??) to create the simulated cosmic ray sample for calibration. However, **CRY** proved to be inefficient, generating particles failing to hit the detector, resulting in wasted processing resources and disk space. Moreover, the momentum and angle distributions in **CRY** are not well suited to the **NOvA** sites, potentially impacting the calibration accuracy.

To overcome these challenges, I developed and implemented a data-based simulation that eliminates the need for the **CRY MC** generator. Instead, I use a subset of the cosmic data sample used in calibration, pulling information on the muon vertex position, direction and momentum, to use, after some corrections and smearing, as inputs to the detector simulation to create a new simulated cosmic ray sample.

This approach results in a near-perfect efficiency, ensuring that almost every simulated muon contributes to the final calibration sample, thus saving processing time, file size, and storage. Additionally, the simulated muon distributions are inherently consistent with the distributions from data. Given that the calibration chain itself is a time and computing intensive process, the reduction in the number of simulation files and in their sizes has significant benefits downstream of the file generation.

1.2.1 Reconstruction and Selection of Cosmic Data Events

It is important to choose a data sample that represents the detector in an ideal state, with as few known issues as possible. For Test Beam, we chose the period 4 data sample (see Tab. 1.1), as the other periods had complications such as faulty **FEBs**, or underfilled cells. We only used half of period 4 data by skipping every other sub-run to limit the number of simulated events to that necessary for a successful calibration.

We designed the reconstruction and selection criteria so that the majority of the simulated cosmic muons make it into the final simulation calibration sample. Therefore, we employed a similar process to that used to create the data calibration samples. Additionally, we require all distributions of the selected events to be well-understood and to resemble those of the data calibration samples.

Remove Beam Spills

The first step is to remove beam spill events based on their time relative to the time of the beam spill. For Test Beam the beam spill is 4.2 s long and we remove all events within a 5 s window from the start of the beam spill, as shown in Fig. 1.3. This should leave us with mostly cosmic events.

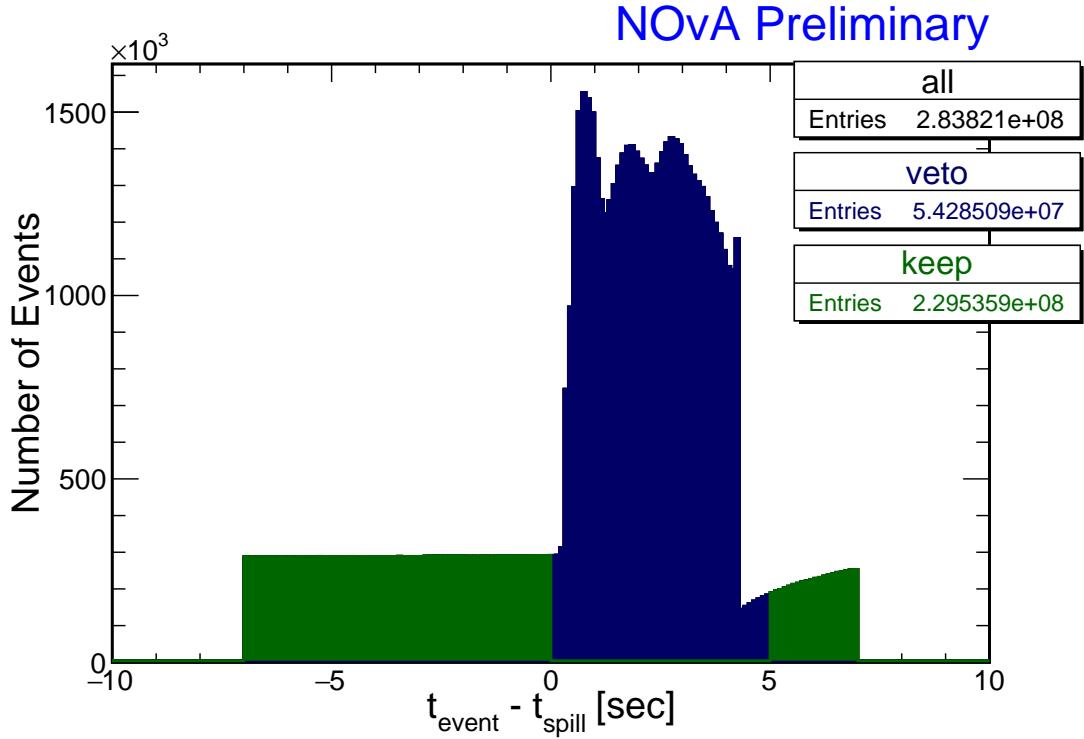


Figure 1.3: Test Beam beam spill events removed (blue) from the calibration samples. The remaining events (green) should mostly consist of cosmic particles. This example and the numbers of entries are for the full period 4 Test Beam sample.

Reconstruction

To use data events in simulation, we need to reconstruct their vertex positions and their initial 4-momenta. We use the standard reconstruction methods from NOvA, described in Sec. ???. First we take the raw hits and group them into slices. Then we

reconstruct cosmic tracks using the window cosmic track algorithm (used for calibration samples). Since we also require the 4-momentum information we have to use the Break Point Fitter (BPF) tracking algorithm to identify muons and assign their momenta. [BPF](#) requires vertex and prong input information, which we get from a cosmic ray vertex and FuzzyK prong algorithms respectively. The first three steps are identical to the full reconstruction applied to both data and simulation to produce the calibration samples. Since we do not need a 4-momentum information for calibration, we do not need to use cosmic ray vertex, FuzzyK vertex, or the [BPF](#) to create calibration samples.

Selection

After the reconstruction process, we proceed to select events based on their slice and [BPF](#) track properties. The overview of all selection criteria and their corresponding cut values are listed in Tab. 1.2. In detail, the following conditions are used to select cosmic muon events for the data-based simulation:

1. We only use successfully reconstructed 3D [BPF](#) tracks with the muon assumption;
2. As we aim to select cosmic events originating outside the detector, we apply a cut based on the distance of each track's start position from the edges of the detector. This cut has a negligible impact on the [BPF](#) tracks, as indicated by the minimal difference between the red and the dotted azure lines in Fig. 1.4;
3. We remove all events whose track is parallel to the beam direction, by requiring the angle from the z axis (parallel to the beam) to be $|\text{Cos}_Z| \leq 0.98$. Figure 1.4 demonstrates the presence of events peaked at track lengths of approximately 410 cm and 200 cm, which correspond to the total and half length of the detector, respectively (or alternatively lengths of both modules and of a single module). These events are strictly parallel to the beam direction and are likely remnants of beam events. Applying a cut on Cos_Z effectively removes these events without affecting the rest of the data. This cut might only be needed for the Test Beam detector and not for the [ND](#) and [FD](#), as it is likely these are particles scattered from the Test Beam beamline, or from the secondary beam.

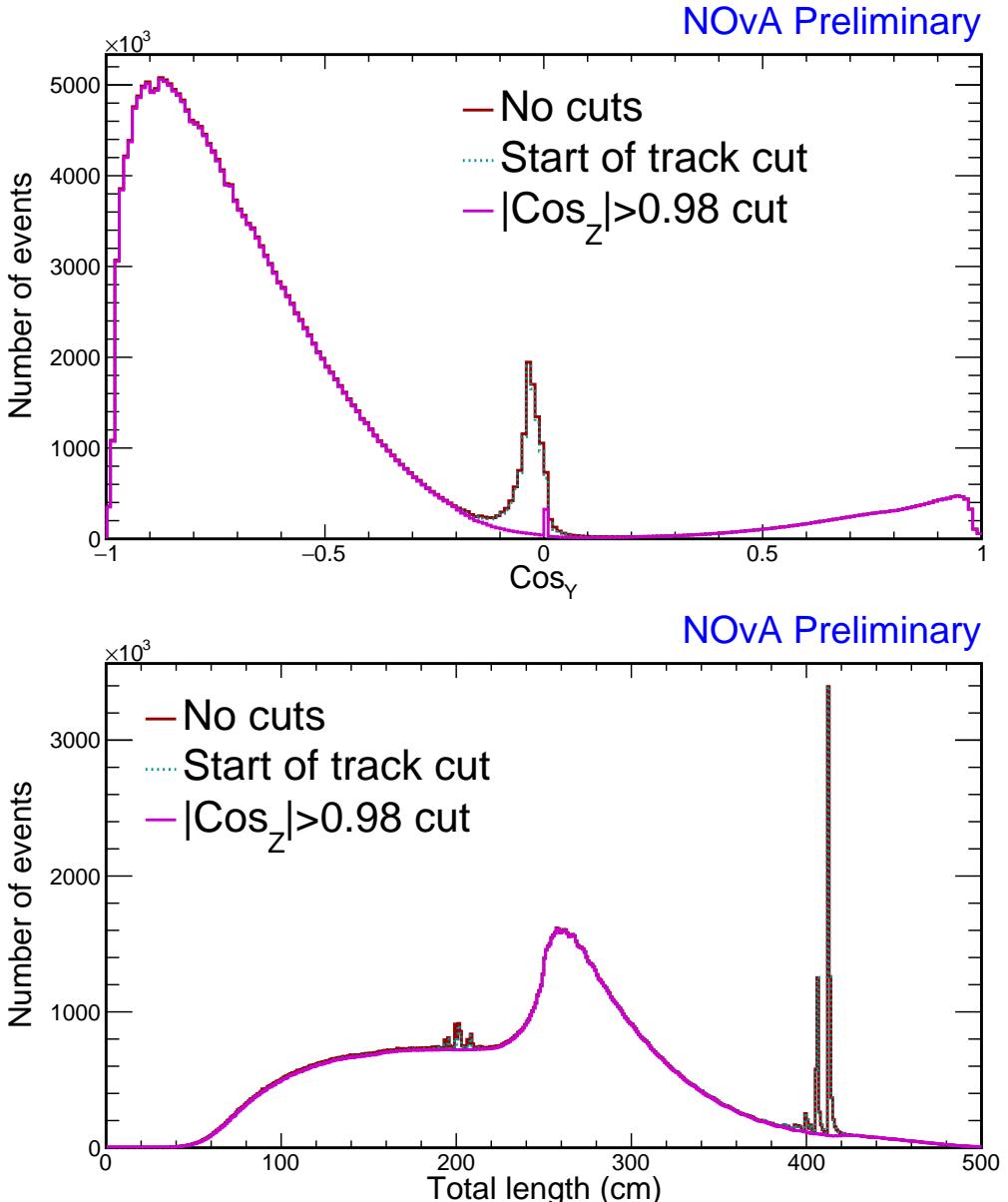


Figure 1.4: Impact of track start and maximum track angle from the z axis (Cos_Z) cuts on the Test Beam data for the data-based simulation of cosmic muons. The track start cut has only negligible effect. The maximum Cos_Z cut effectively removes sharp peaks in the total track length distribution and events perpendicular to the y axis. These events are all parallel with the z axis and are most likely leftover beam events. Top plot shows distribution of the angle from the y axis (Cos_Y) and bottom of the total track length. All of the distributions are made from the period 4 Test Beam data.

4. To ensure that only events contributing to the final calibration sample are simulated, we use a selection based on the cuts used to select events for the data calibration samples (see Sec. ??). We call these cuts the **calibration cuts**. However, there are two caveats we need to consider when applying the calibration cuts:

- (a) First, to create calibration samples, we apply the selection on tracks from the **window cosmic track** algorithm instead of the **BPF** algorithm, which yield different distributions as depicted in Fig. 1.5. Notably, the **BPF** tracks have a hard cut-off at the detector edges, whereas the window cosmic tracks are allowed to start beyond these limits. Also, the **BPF** tracks have a rugged distribution in Cos_Z , which is not present for window cosmic tracks. This is likely caused by the detector structure, as shown in Fig. 1.6, but it is not clear how. We concluded that the rugged shape does not have any impact on the resulting simulation. Given these differences between the tracking algorithms, applying the calibration cuts on **BPF** tracks could mistakenly remove events that would pass the same selection when applied to the window cosmic tracks.
- (b) Second, each reconstruction algorithm has intrinsic deficiencies that can lead to misreconstructions. Applying the full calibration cuts may remove misreconstructed events that should have been included in the simulation, introducing a bias.

To address these concerns, we loosened the full calibration cuts to create a ‘buffer’ around the selected events, allowing for fluctuations of the reconstruction algorithms while maintaining track quality. This way, events that would have been removed based on the calibration cuts applied to their reconstructed **BPF** tracks, but kept based on the calibration cuts applied to their Window cosmic tracks, now have a chance to make it into the final selection and therefore calibration sample. The differences between the full calibration cuts and the employed loosened calibration cuts applied to the **BPF** tracks are listed in Tab. 1.2 and shown in Fig. 1.5. There we also show the data calibration sample, which was created by applying the full calibration cuts on window cosmic tracks from the same artdaq data sample.

During the selection process, we determine whether the muon is stopping inside the detector or passing through, based on the reconstructed track’s end position. For Test Beam we say it is a stopping muon if its track ends at least 20 cm from any edge of the detector. For the far and near detector this is 50 cm. This information assists in correcting the energy of through-going muons, as outlined in the following Sec. 1.2.2.

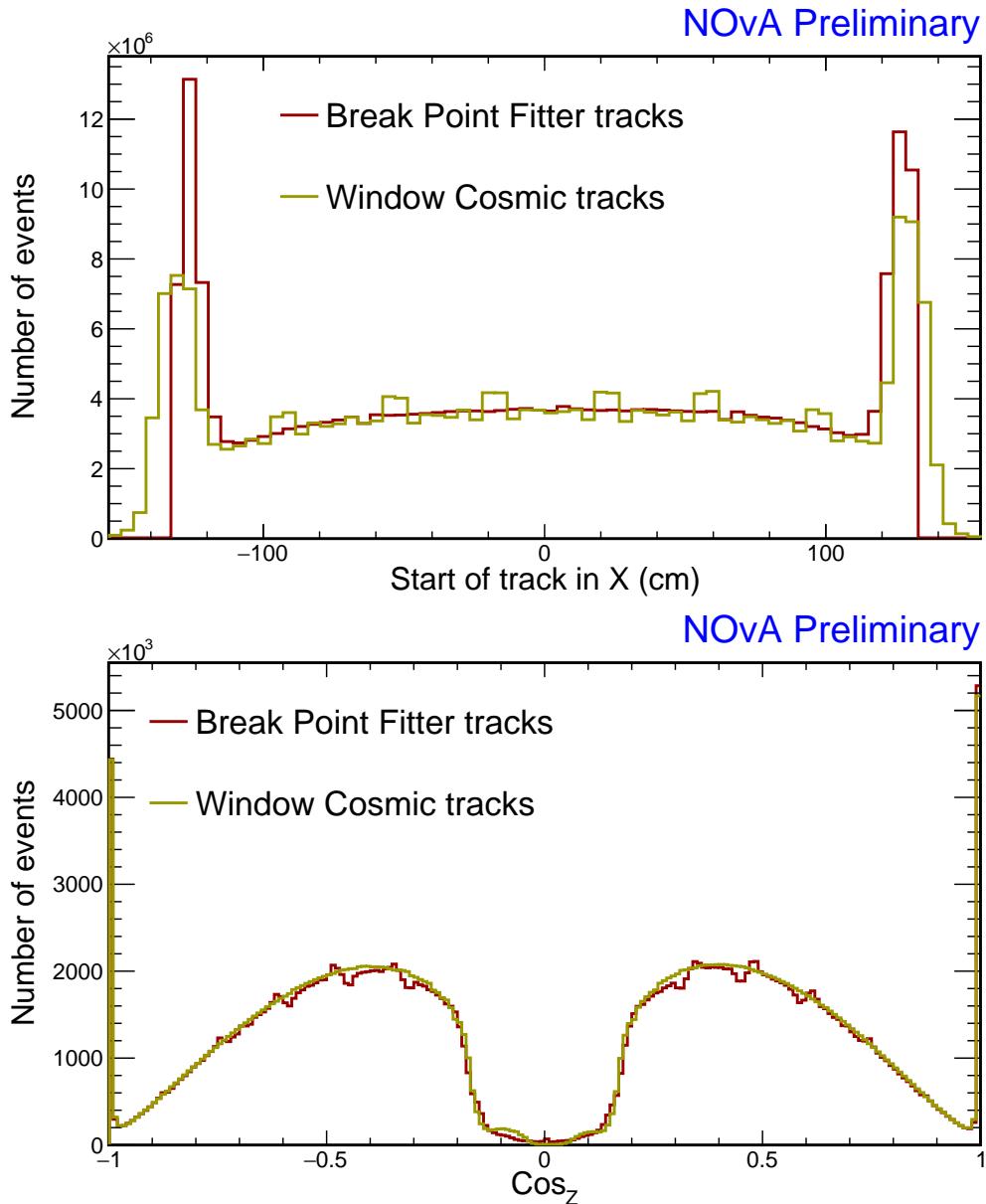


Figure 1.5: Difference between the tracks reconstructed with the BPF and with the window cosmic track algorithms. Both distributions are for the period 4 Test Beam data (with removed beam spill) without applying any selection. The bottom plot shows a distribution of the angle from the z axis.

1.2.2 Energy Correction, Charge Assignment and Smearing

Once we have the kinematic information for the selected events, we perform several tasks to get the final sample of cosmic muon events for the data-based simulation. This includes correcting energies of the through-going muons, assigning a charge to each muon event, and smearing and converting the information into the correct format required by the generator.

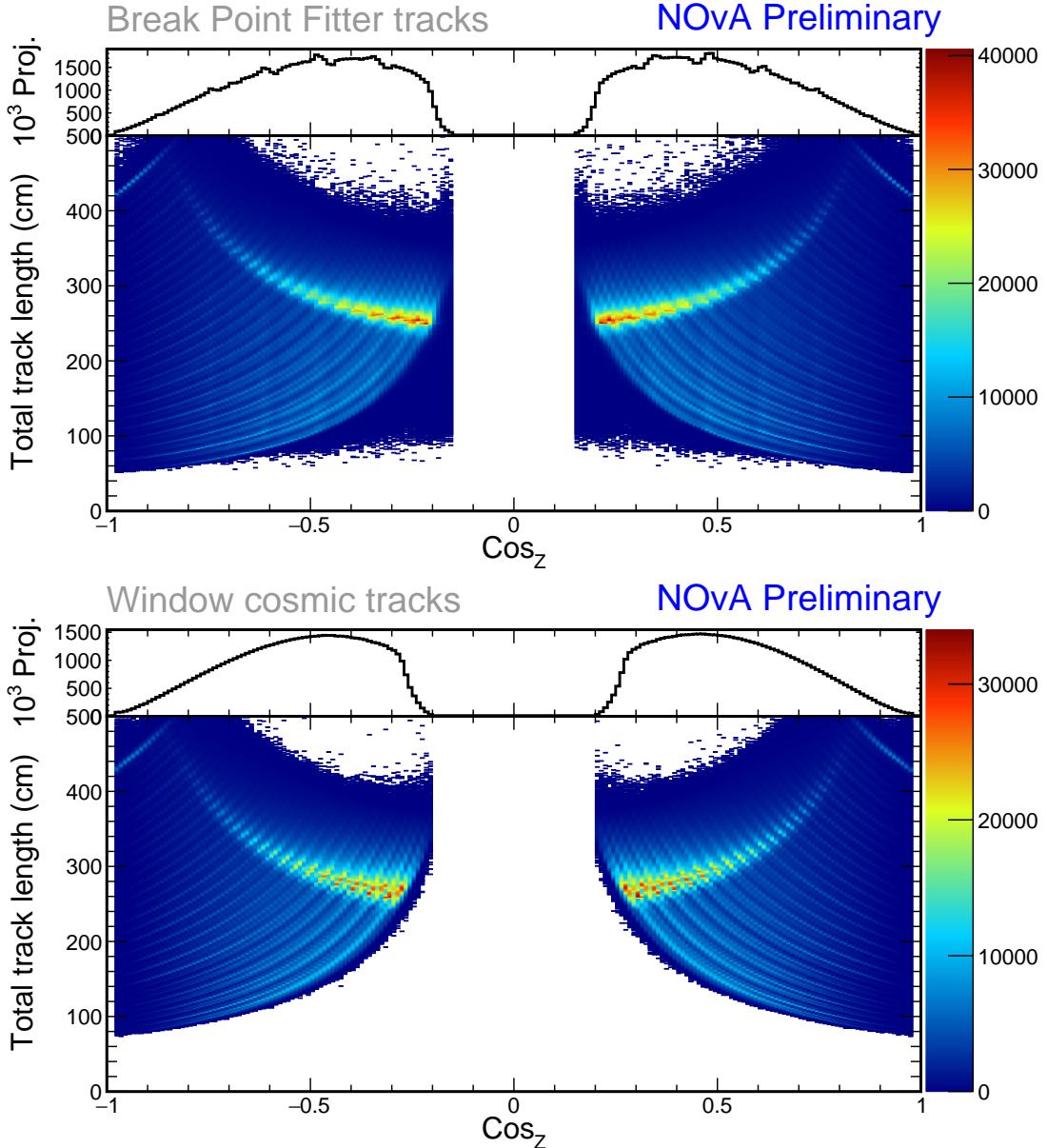


Figure 1.6: Investigating the origin of the rugged shape in the Cos_Z distribution of BPF tracks. The top plot is created with the Loose calibration cuts and the bottom plot with the Full calibration cuts, as described on Tab. 1.2. However, this difference in selection shouldn't matter. The long lines on the 2D plots are likely the effects of the detector structure. We can see that for the BPF tracks, each Cos_Z angle corresponds to a specific track length, whereas for the Window cosmic tracks there is multiple track length for each angle. This could cause the resulting shape in the Cos_Z distribution of BPF tracks.

Energy Correction

Through-going muons do not deposit all of their energy inside the detector. From the reconstructed information we cannot reliably calculate their initial energies, but we can estimate an energy that could leave the same track. In general, the energy

Table 1.2: Event selection of cosmic muons used for the data-based simulation (in green under Loose selection) and comparison to the Full selection cuts used to create the calibration samples in blue. The last two rows are not used for Test Beam, but are employed for the Near and Far detectors and should be examined before creating another data-based simulation for them.

Cut	Selection	
	Full	Loose
Muon assumption and 3D track from BPF		
Max. track start distance from edge	50 cm	
Max. Cos_Z	0.98	
Calibration sample selection	Min. number of hits in X or Y	2
	Min. difference between $Stop_Z$ and $Start_Z$	70 cm
	Min. Cos_Z	0.2
	Min. frac. of slice hits in track in each view	0.8
	Max. number of cells per plane in each view	6
	Max. difference in X-Y for first (last) plane	3
	Max. plane asymmetry	0.1
	Max. step size to median step size ratio	3
	Max. vertex distance from edge	10 cm
	Max. track end distance from edge	10 cm

spectrum of cosmic muons can be approximately described by a power law $E^{-\alpha}$, with $\alpha \approx 2.7$ [19, 20]. The expectation value for the ‘true’ initial energy of through-going muons can be therefore calculated as

$$\langle E \rangle = \frac{\int_{E_R}^{E_C} E \cdot E^{-\alpha}}{\int_{E_R}^{E_C} E^{-\alpha}} = \left(\frac{\alpha - 1}{\alpha - 2} \right) \left(\frac{E_C^{2-\alpha} - E_R^{2-\alpha}}{E_C^{1-\alpha} - E_R^{1-\alpha}} \right), \quad (1.1)$$

where E_R is the reconstructed energy we got from the BPF. E_C is the critical energy chosen to be 300 GeV, as we do not expect muons with higher energies to be selected due to large showers along their paths.

We use this corrected initial energy for all muons that do not stop inside the detector, as identified during selection described in Sec. 1.2.1. Figure 1.8 shows the corrected energy distribution of our selected events and demonstrates that the choice of the critical energy does not significantly change the correction.

This corrected energy is **not** a good representation of the true energy spectrum of cosmic muons on surface level and getting a correct energy distribution from data would require a much more dedicated effort. The corrected energy would also be different for different NOvA detectors, since the reconstructed energy is calculated from the track length. For example, the corrected energy of cosmic muons when

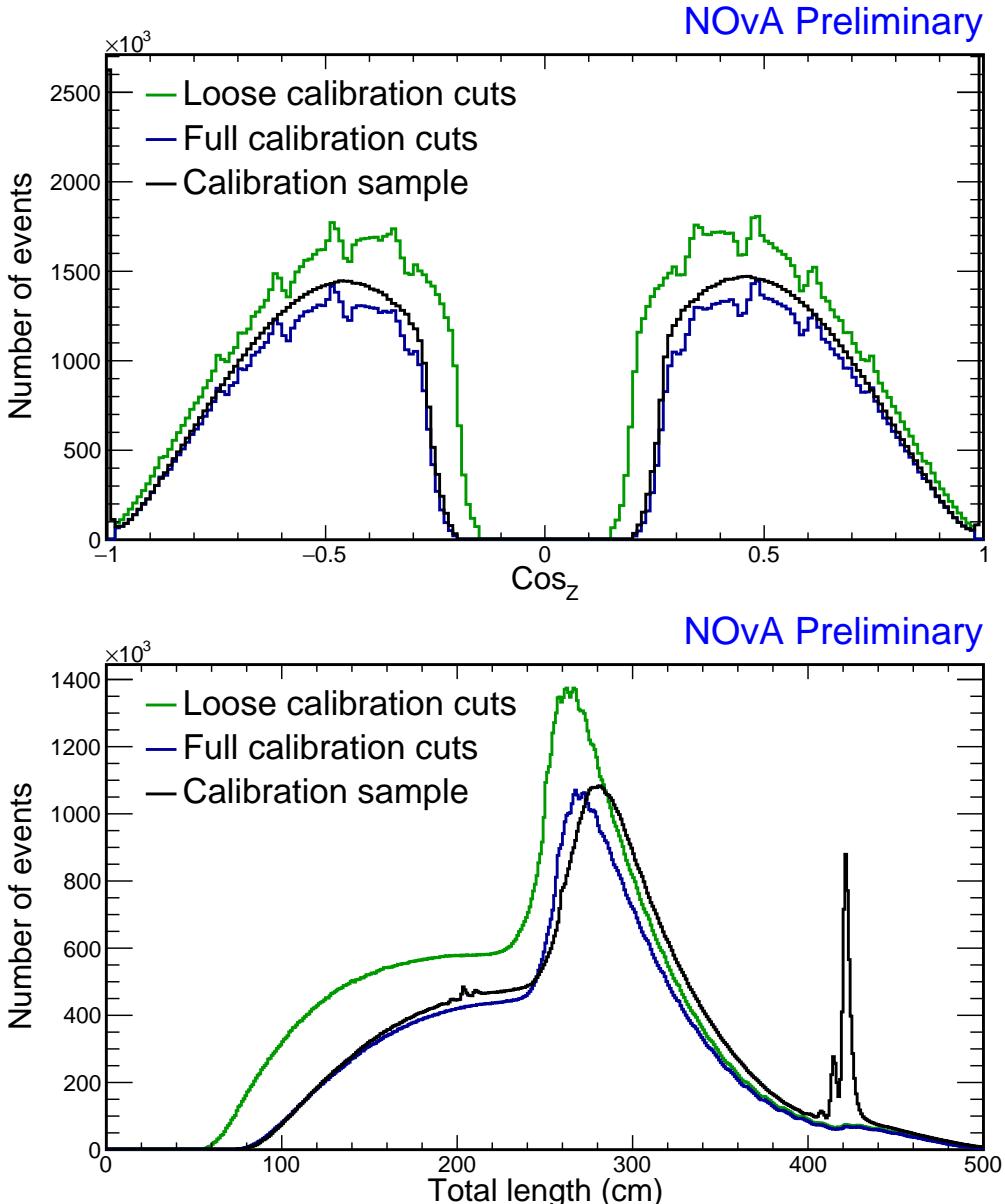


Figure 1.7: Comparison of event selections for the data-based simulation and of the corresponding data calibration sample in black. The green line represents the final selection used for the simulation, using the loosened calibration cuts, as described in text and in Tab. 1.2. The blue line shows the distributions with full calibration cuts applied to the same sample and using the same tracks. The ‘calibration sample’ shown in black was made with the same full calibration cuts as the blue line (without the track start cut and maximum Cos_Z cuts), but applied to the window cosmic tracks instead of the BPF tracks. All of the distributions are made from the period 4 Test Beam data.

entering the detector would be larger for the bigger ND than for Test Beam, even though the ND is underground.

However, since this simulation is intended to be used for calibration, where we use through-going muons only for relative calibration, we do not need a perfect repre-

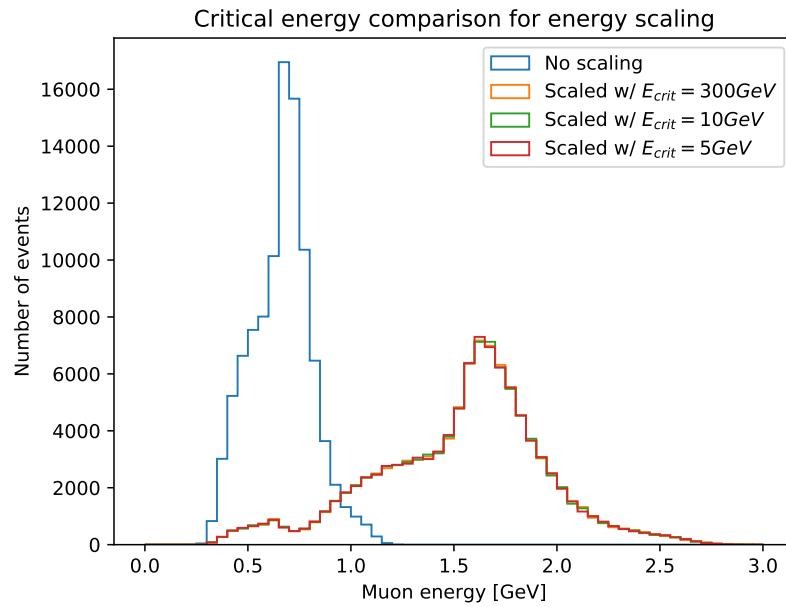


Figure 1.8: The effect of energy correction for through-going muons with various critical energies. No significant difference can be seen when using different critical energies.

sentation of the cosmic muon energy spectrum. Not including more energetic cosmic muons into the simulation does bias the energy deposition towards lower values, but this is corrected for during absolute calibration which only uses stopping muons, for which we assume we reconstruct their energy well from BPF.

If someone were to use this simulation for something other than calibration, it would be necessary to rethink the energy correction, either by changing the energy estimation from track based algorithms to energy deposition, or by including information from external sources. It would also be necessary to include angular dependence for the energy correction as described in the PDG [20].

Smeearing

The reconstructed data is influenced by the detector structure, reconstruction efficiencies and other effects that can bias the simulation. To avoid this influence, we smear the reconstructed values by randomly changing

- the total momentum within 2%,
- the azimuthal angle uniformly,
- the polar angle within 4 mrad,

- and the X/Y and Z vertex positions within the width or depth of the cell respectively.

The size of the smearing has been decided as the best estimate of variations of these variables for cosmic muons.

Charge Assignment

We need to tell the detector simulation whether to simulate a muon or an anti-muon. However we do not reconstruct the charge of the muons, so we have to randomly assign it based on a statistical distribution from external measurements [19]:

$$P_+ \simeq 0.539 + \frac{x}{34.5} - \left(\frac{x}{9.48} \right)^2 + \left(\frac{x}{8.27} \right)^3, \quad (1.2)$$

where x is the logarithm of the total momentum in GeV.

Running the Simulation

We save the vertex positions, the four momenta and the assigned charge into a text file, that is then fed into the same detector and readout simulation chain, as was described for the ND and FD in Sec. ???. We use the fibre brightness map that is used in calibration (see Sec. ??) to inform the simulation about the real detector conditions. Since we want the simulated detectors to be functional copies of the ideal versions of the real detectors, it is important to provide a correct brightness file without any defects. For this simulation we use the fibre brightness map described in Sec. 1.3.1.

1.2.3 Validation

To validate whether the newly created simulation works as expected, we compare the new simulation with the original data it was created from. Additionally, we use the new simulation as ‘fake data’ and pass the simulated events through the same reconstruction, selection and simulation processes as were used to create the first simulation, creating a ‘re-simulation’ sample. This is used to validate the stability of the simulation process.

For the data-simulation comparisons we use the events from the actual calibration samples as data. This is equivalent to looking at the Window cosmic tracks with

full calibration cuts described in the selection Sec. 1.2.1. We are expecting the new simulation to be similar to the data calibration sample, without a bias from the original data used for the simulation.

Figures 1.9 and 1.10 show that the angular distributions of the new simulation (pink lines) are almost the same as the distributions of the BPF tracks with full calibration cuts (blue dashed lines). This means that loosening up the calibration cuts (green dashed lines) did not help as expected with compensating for the underlying differences between the BPF tracks and the Window cosmic tracks. This can also be seen on the total track length distribution in Fig. 1.10. During the development of the event selection we created three more versions of the simulation with various event selections, including with full calibration cuts and with different loosened calibration cut values. It became clear that it is unlikely we could mitigate these track algorithm differences by changing the selection even more. Since the entire simulation process is fairly time and resource consuming and the distributions of the new simulation look reasonable and are close enough to the data, we've decided to proceed with this version of the simulation and use it in the Test Beam calibration.

The start of track comparison between data and simulation in Fig. 1.11 and 1.12 show that there are fewer events that start at the edge of the detector and the vertex positions are moved slightly towards the inside of the detector. This is likely the result of the smearing of the vertex positions and we do not expect this to have an effect on the calibration. *COMMENT: Should this have been mitigated? Should I mention here that the smearing is still a good idea?*

After adding the distributions for the re-simulation calibration sample, shown in Fig. 1.13, we can see that the tracks' starts are shifted even further towards the inside of the detector. This would support the hypothesis that this effect is caused by the smearing of the reconstructed variables. This is also likely directly related to the loss of events with longer track lengths as shown in Fig. 1.13. If tracks start a few centimetres later in the detector their tracks would get shorter by the same amount.

COMMENT: Should I describe every single aspect of the validation plots in here? For example the black peaks in Fig. 1.9 and 1.10? These have been explained during selection (leftover beam events) but might be good to explain them here as well and reference the selection

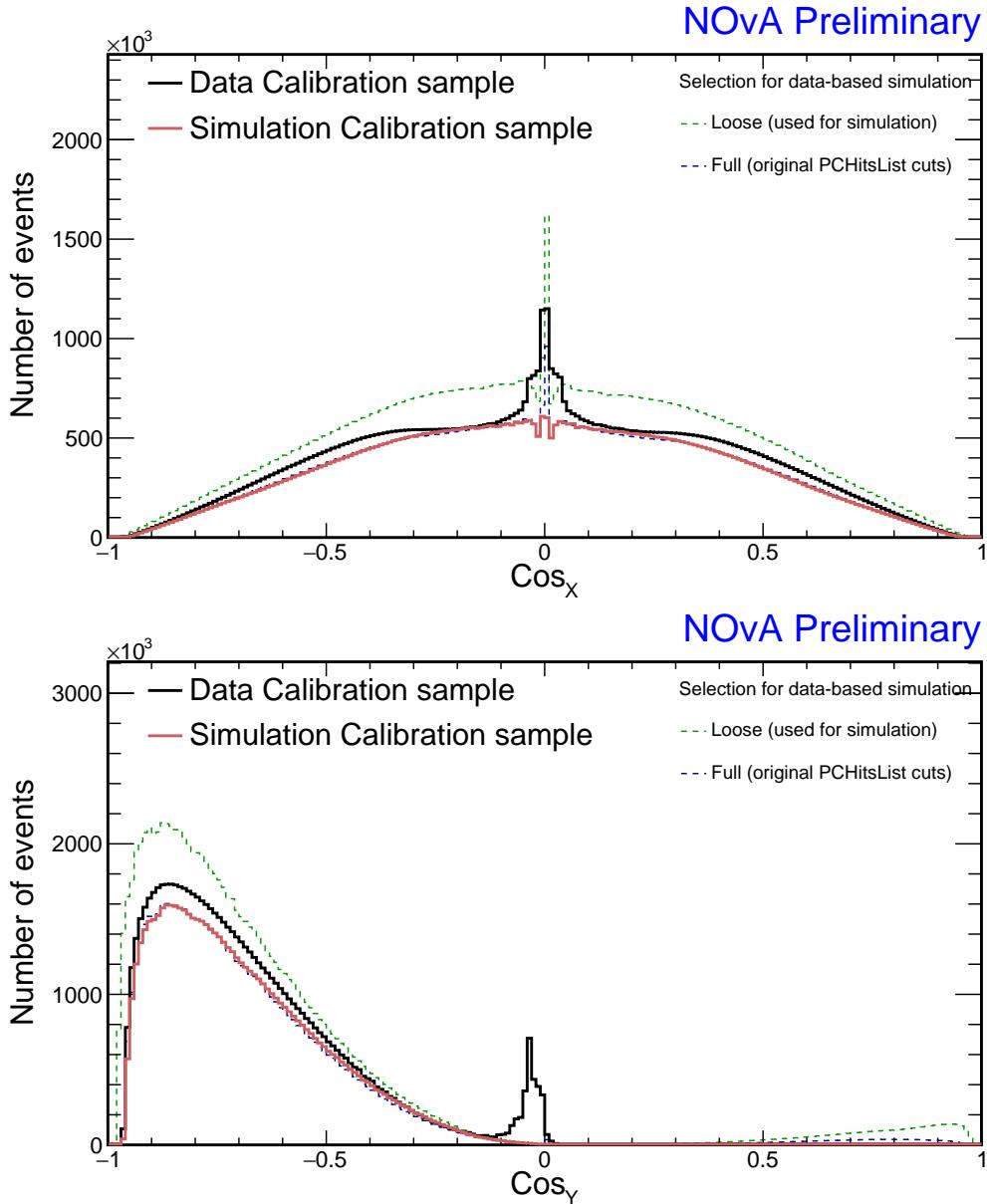


Figure 1.9: Angular distribution comparison of the newly created simulation calibration sample and the corresponding data calibration sample. We are also showing the selection of data used to create the new simulation in green and a ‘full calibration cuts’ selection (same cuts as used for the simulation calibration sample but applied to BPF tracks instead of Window cosmic tracks) in blue.

1.3 NOvA Test Beam Detector Calibration

In this section we describe the details of the Test Beam detector calibration as it was finalized in June 2023. This version includes a new purpose-made simulation and all the measured Test Beam data, with the exception of the period 1 data.

The data calibration samples for Test Beam were creating using the same procedures as the ND and FD calibration samples, described in Sec. ???. However, there are

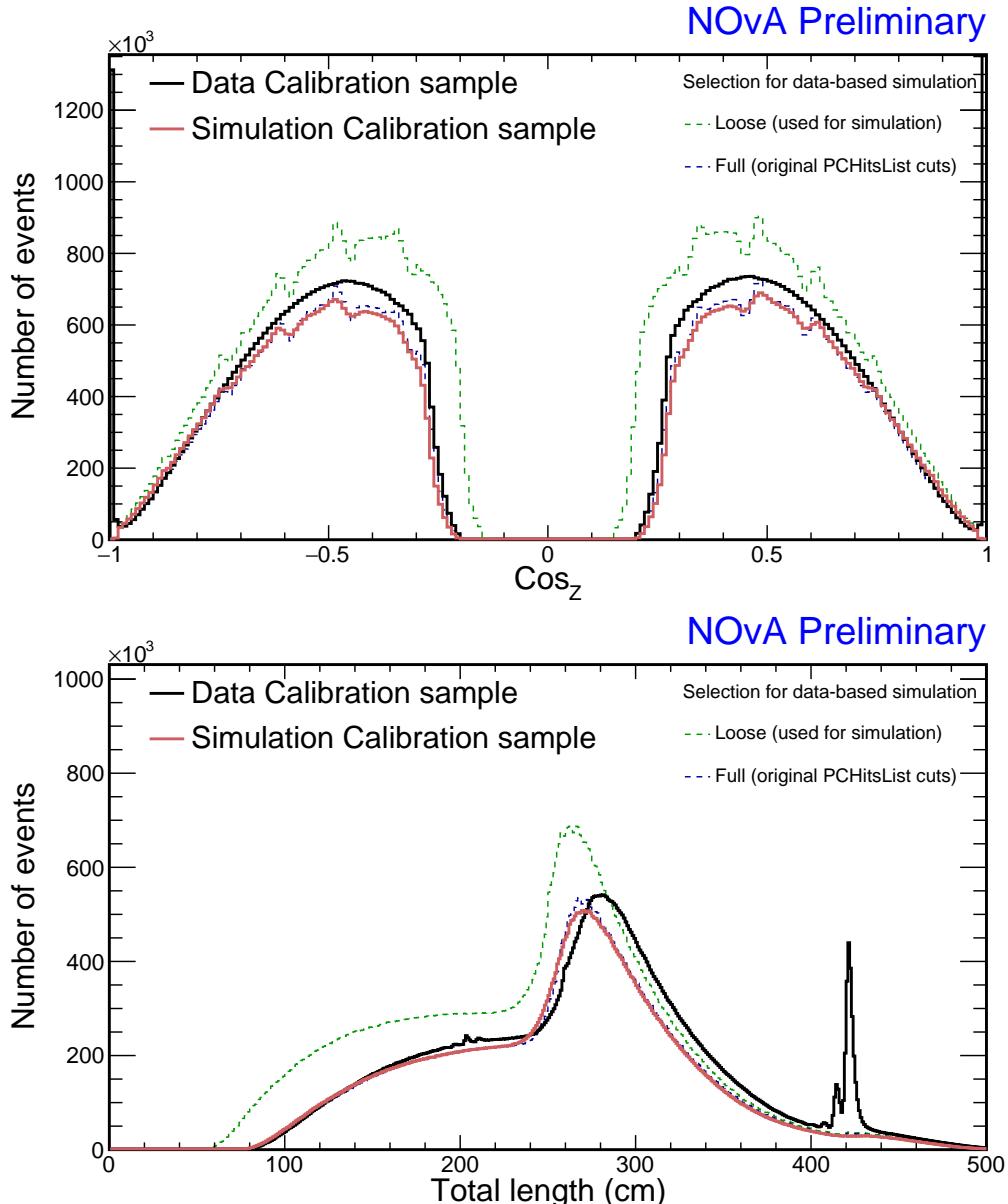


Figure 1.10: Angular and total track length distribution comparison of the newly created simulation calibration sample and the corresponding data calibration sample. We are also showing the selection of data used to create the new simulation in green and a ‘full calibration cuts’ selection (same cuts as used for the simulation calibration sample but applied to BPF tracks instead of Window cosmic tracks) in blue.

two cuts from the event election, that were by accident not included for Test Beam during the processing of the data samples. This can be seen on Tab. 1.2, where the two bottom rows show the two excluded cuts. One cut contains the vertex close to the edge of the detector ensuring we only use cosmic events, the other contains the end of track close to the edge, ensuring we only use through-going muons for the relative calibration. Given that we remove beam events and that all the other cuts

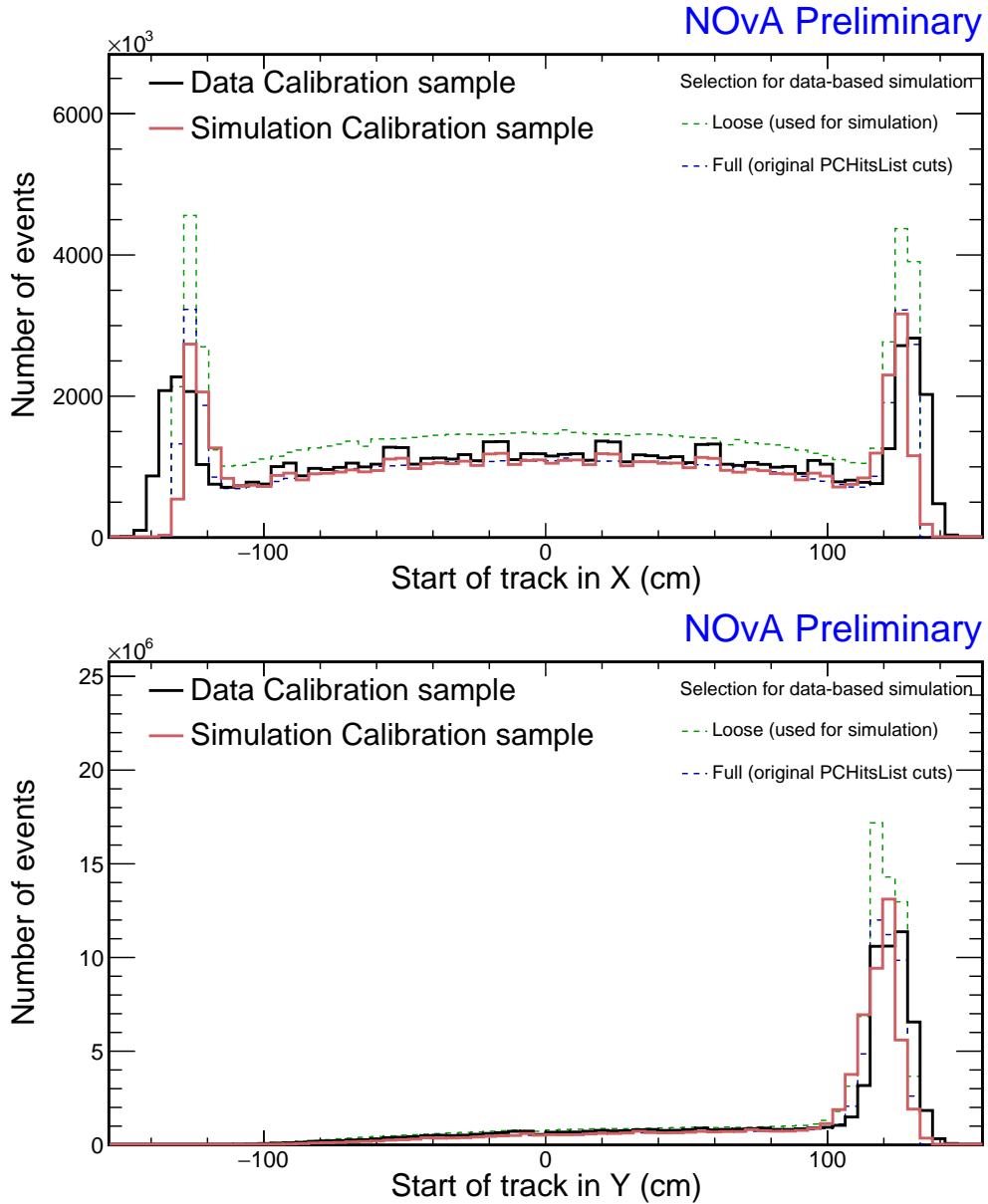


Figure 1.11: Start of tracks comparison of the newly created simulation calibration sample and the corresponding (period 4) data calibration sample. We are also showing the selection of data used to create the new simulation in green and a ‘full calibration cuts’ selection (same cuts as used for the simulation calibration sample but applied to BPF tracks instead of Window cosmic tracks) in blue.

are designed to select cosmic events, the first cut has only a negligible effect on the final selection. Additionally, the stopping muons only make up a small fraction of the total cosmic muon events, rendering the second cut also with only limited effect. Therefore, we concluded that the lack of the two event selection cuts for the Test Beam calibration samples does not have a substantial impact on the result and does not necessitate re-processing of the calibration samples, which would cost time and

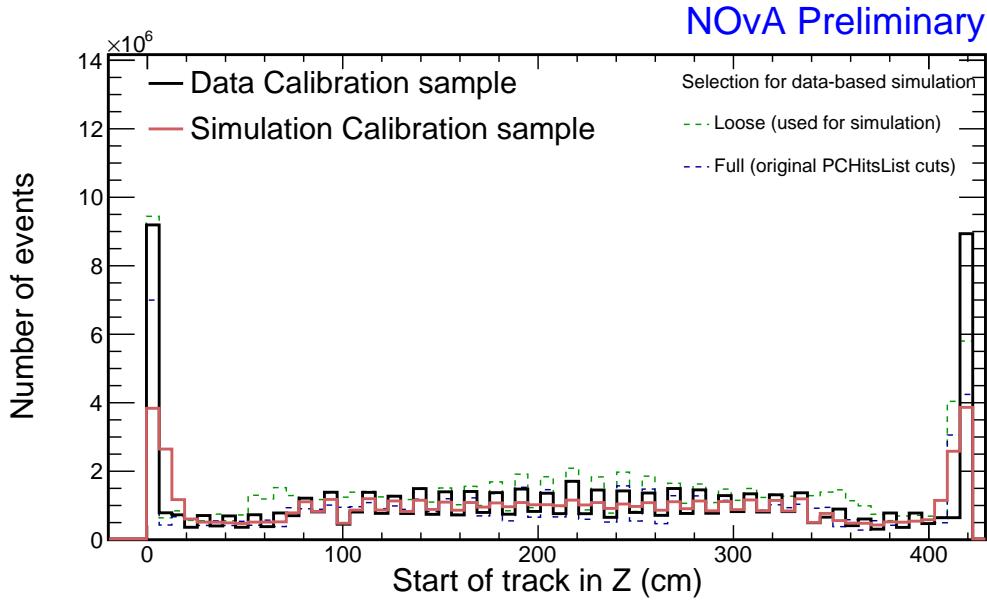


Figure 1.12: Start of tracks comparison of the newly created simulation calibration sample and the corresponding (period 4) data calibration sample. We are also showing the selection of data used to create the new simulation in green and a ‘full calibration cuts’ selection (same cuts as used for the simulation calibration sample but applied to BPF tracks instead of Window cosmic tracks) in blue.

resources.

This section is organized as follows. We first describe the Test Beam versions of the fibre brightness map and the threshold and shielding correction introduced in Sec. ???. We then go over the simulation and the three data samples and for each one we introduce their respective detector conditions and how they may affect the energy deposition. We show the state before the calibration process and discuss the effects that should be corrected for during calibration. We then show a selection of attenuation fit results and an overview of the relative calibration effects. Afterwards, we discuss the absolute calibration for all the samples combined, as well as the validation and conclusion of the Test Beam calibration.

1.3.1 Fibre Brightness

To divide the Test Beam detector into fibre brightness bins we used the attenuation fit results for period 4 Test Beam data (described in Sec. 1.3.6), as that is the best detector conditions data we have. Since we need the fibre brightness map in order to run the attenuation fits and we need the attenuation fit results to create the brightness file, we proceeded iteratively and first ran the attenuation fit with an older version of the

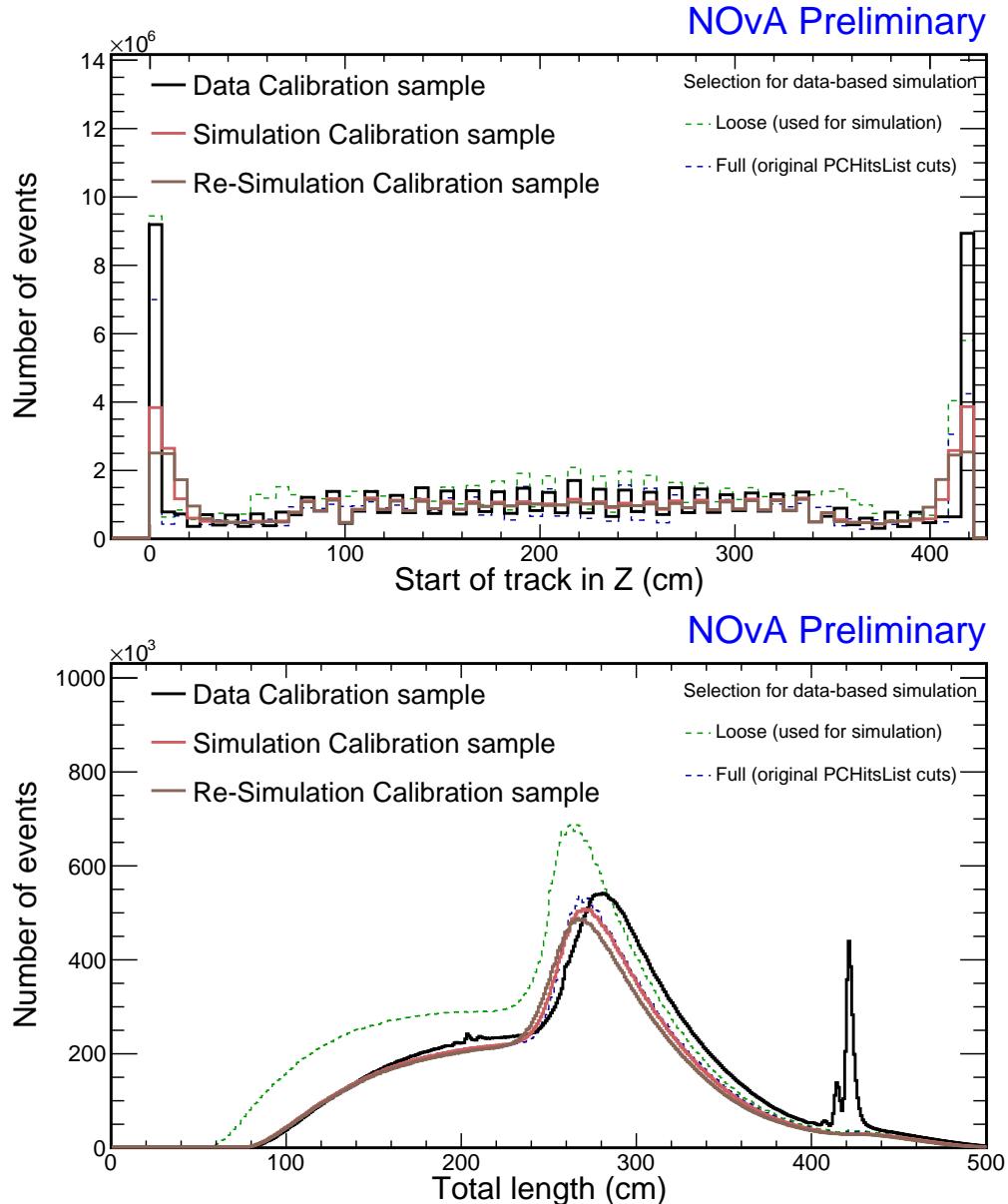


Figure 1.13: Distribution of the re-simulated events, when the new simulation is used as ‘fake data’ for a new iteration of the simulation process discussed in this document. This is compared to the newly created simulation calibration sample and the corresponding data calibration sample. We are also showing the selection of data used to create the new simulation in green and a ‘full calibration cuts’ selection (same cuts as used for the simulation calibration sample but applied to BPF tracks instead of Window cosmic tracks) in blue.

brightness file and then used the newer fit results to create a new brightness file to be used in a new attenuation fit.

As we are only using the attenuation fit results in the centre of each cell to create the fibre brightness map, we’ve decided to allow some cells that initially failed the calibration condition ($\chi^2 > 0.2$), to be still used for the creation of the brightness file.

Otherwise, all the officially uncalibrated cells would be assigned an average response and we would lose the information on their relative brightness. As can be seen in Fig. 1.14, some attenuation fits have $\chi^2 > 0.2$, even though they correctly represent the energy deposition in the centre of that cell. By carefully investigating all Test Beam cells with $\chi^2 > 0.2$ (which is doable for Test Beam, due to its small number of cells), we concluded it is safe to use all the attenuation fit results with $\chi^2 < 0.7$. We use this loosened calibration condition only to create the fibre brightness file and we keep the original condition for the actual calibration results.

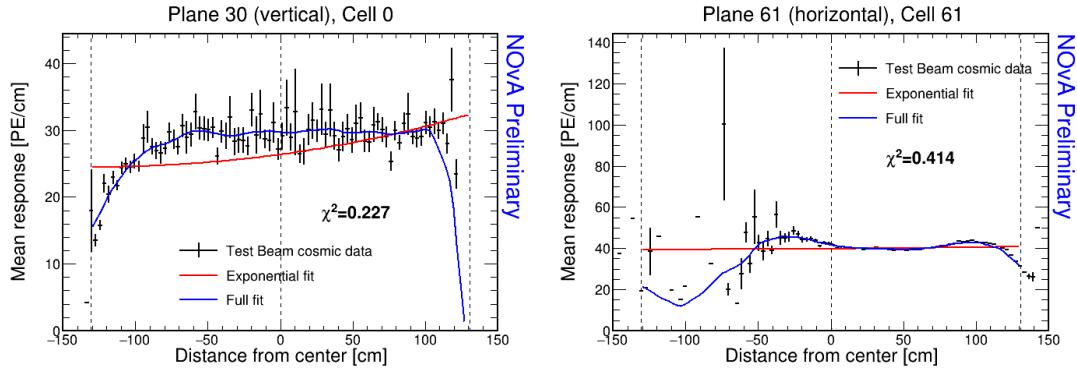


Figure 1.14: Examples of attenuation fits for two cells that fail the calibration condition, but the fit (blue line) still correctly represents the energy deposition in the centre of that cell (dashed vertical line in the middle).

The final distribution of fibre brightness bins and their corresponding relative brightnesses for the Test Beam detector is shown in Fig. ??.

1.3.2 Threshold and Shielding Corrections

We created the threshold and shielding correction for Test Beam from the new simulation described in Sec. 1.2. As can be seen in Fig. 1.15, the correction is almost uniform as a function of both the position within each cell and the cell number. This is the case for all the Test Beam fibre brightness bins and for both views.

The uniformity of the distributions is expected, as the Test Beam detector is much smaller than the FD, results of which originated the study of the threshold and the shielding effects. The cell length of 2.6 m has only a negligible impact on the energy distribution of cosmic muons or on the threshold saturation. Therefore the threshold and shielding correction for Test Beam is only a normalization factor, except for the cell edges, where there is a large variation in the energy response there anyway due

to low number of events. Since we apply the correction prior to the relative calibration, which only cares about the relative differences across the detector, a correction consisting of only a shape-less normalization factor does not have any impact on the calibration results.

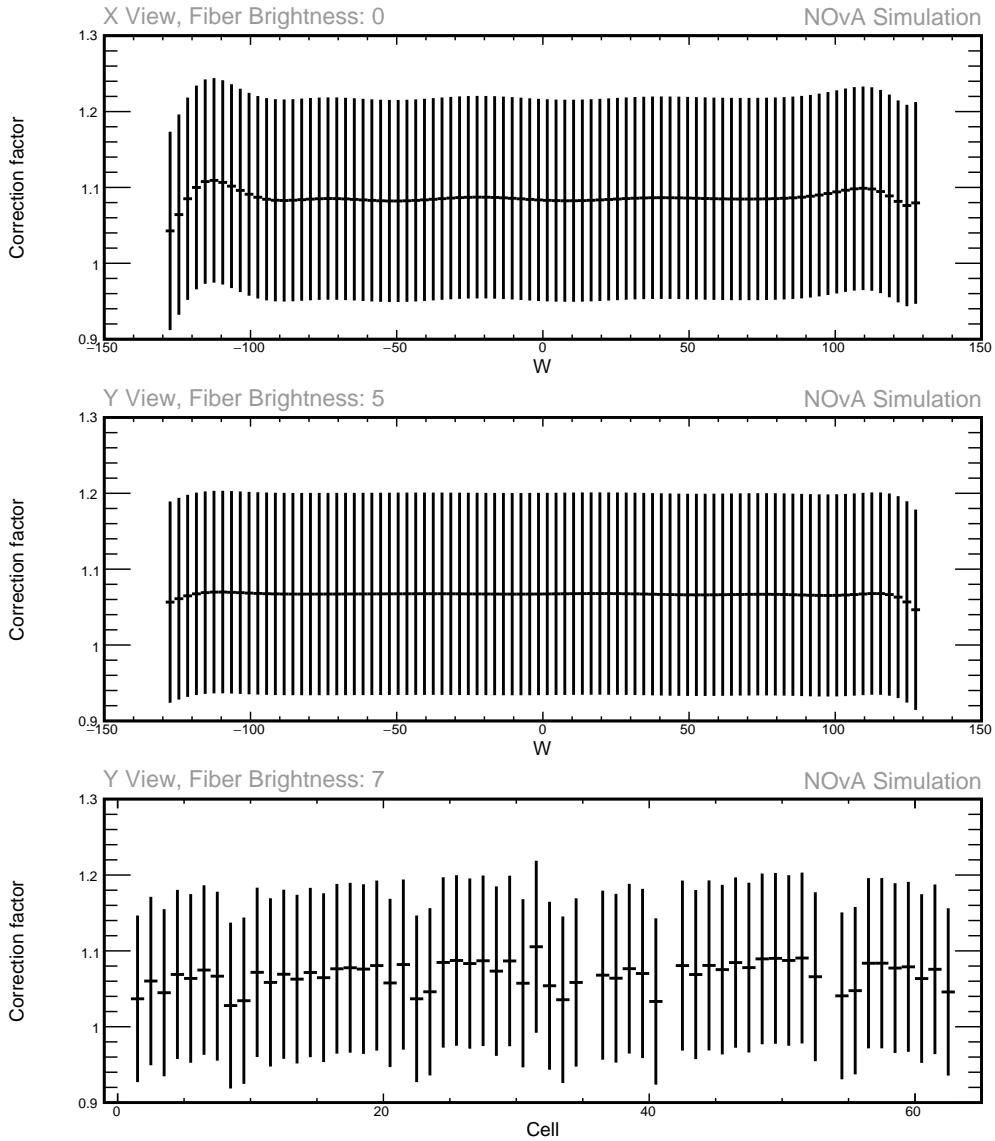


Figure 1.15: Examples of threshold and shielding corrections for the Test Beam detector.

1.3.3 Simulation

We use the data-based simulation described in Sec. 1.2. Figure 1.16 shows the distribution of the tricell hits from the simulated cosmic muon events selected for calibration. The features on this plot illustrate the distribution of tricell hits in the Test Beam de-

tector in ideal conditions and are present in all the data samples as well. We can clearly see the difference between the number of events in the vertical (even) and the horizontal (odd) planes. This is expected as cosmic muons are generally vertical and a single cosmic track often passes more horizontal planes than vertical planes. We can also see that due to the tricell condition there are no hits in cells 0 and 63, which are on the edge of the detector. The clear horizontal lines are made up by cells (0), 15, 16, 31, 32, 47, 48, (63), or in other words by the first and last cell of each 16 cell-wide extrusion that makes up half of a module (which makes up half of the Test Beam plane). As was mentioned in Sec. ??, these cells are 3 mm narrower than the rest, which results in fewer hits and lower deposited energy, but consistent deposited energy per path length. Overall, Fig. 1.16 shows that the tricell hits are distributed fairly uniformly in the centre of the detector, with the number of hits dropping off towards the front, back and corners of the detector. This is due to the event selection applied to the cosmic tracks to be used in calibration.

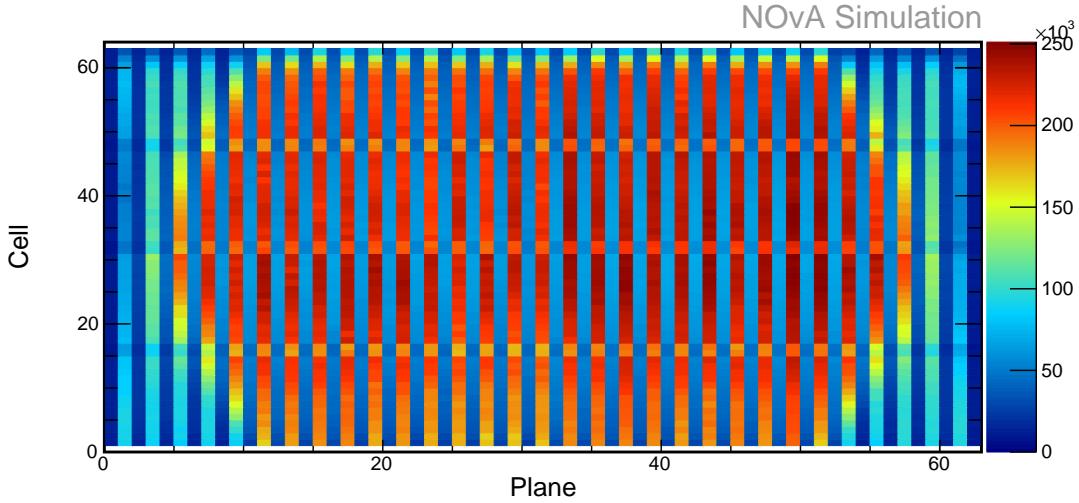


Figure 1.16: Distribution of hits used in the calibration for events in the Test Beam simulation calibration sample.

The distributions of deposited energy before calibration in units of PE/cm are shown in Fig. 1.17. We are showing the dependence on the positions within a cell w , Cell number and Plane number. These are the distributions that are supposed to be uniform after applying the results of the calibration. We can identify the main features that will need to be corrected for during calibration.

The rise of the energy response along w shown in Fig. 1.17 is caused by the attenuation of light along the Wavelength Shifting (WLS) fibres. The drop of the response at

the edges of the cell is caused by the fibres looping and connecting to the Avalanche Photodiode (APD) and the larger uncertainties at the edges of the cell are caused by lower number of hits passing the event selection and the tricell condition.

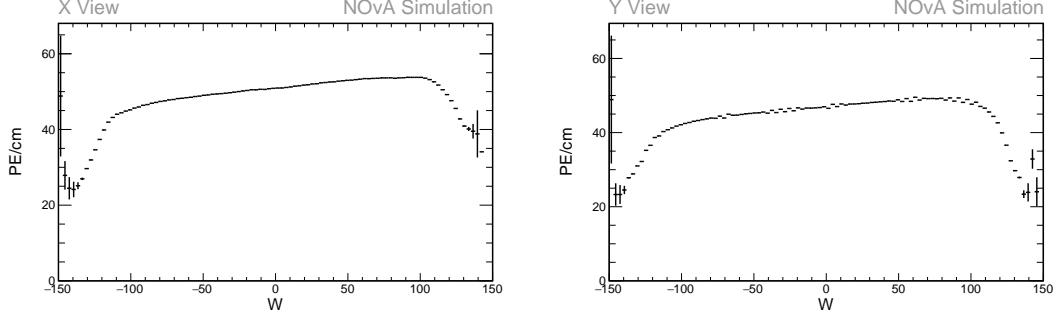


Figure 1.17: Uncorrected average energy response along the position within a cell (w) for simulation. Left side shows distributions for the X view planes and right side for the Y view planes.

The rise of the response along the cell number shown on the middle plots of Fig. 1.18 is due to the varying distance of the cells to the readout. Since the APDs are located on a side of each module, the light from the cells on the opposite side has to travel along the WLS fibre for the additional width of the module, compared to the cells close to the readout. Light undergoes additional attenuation along these so-called pig tails, causing the difference of the energy response.

We can also see additional drops in the uncorrected energy response for cells 0, 1, 9, 10, 23, 24, 31 and 32 (and the corresponding cells in the second module). These are most likely related to the organization of the WLS fibres' connections to the APDs. As can be seen in Fig. ??, the fibres are connected to the total 32 APD pixels in four rows of 8 pixels. Therefore, if one side of each APD has a lower response than the rest, this could explain the drop in the energy response for the aforementioned mentioned cells. However, this has not been confirmed yet and the reason for one side of the APD having a lower response than the rest is unclear. **COMMENT:** *Also I need to make sure that the pixel organization on the APD is the way I explain it here*

The distribution of the uncorrected response along the planes is shown in Fig. 1.19. Here we can see large fluctuations in both views, but we can clearly identify the three distinctly different responses corresponding to the various scintillators used, as we described in Sec.1.1. Additionally, planes 16, 17, 48 and 49 have a lower response relative to their neighbouring planes due to the different readout electronics used in

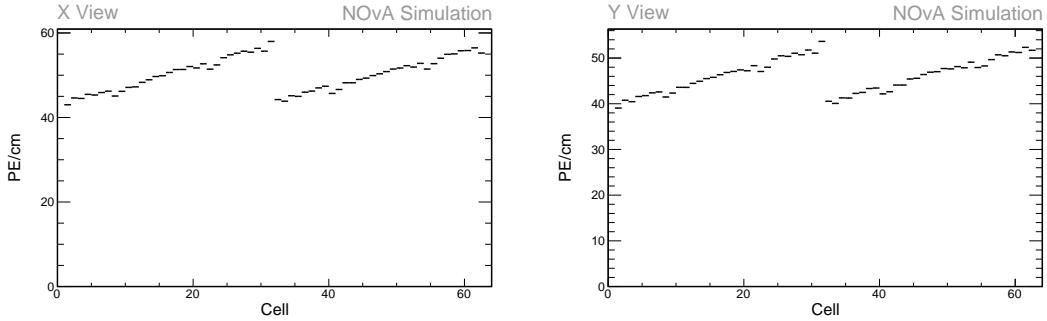


Figure 1.18: Uncorrected average energy response along cells for simulation.

these planes.

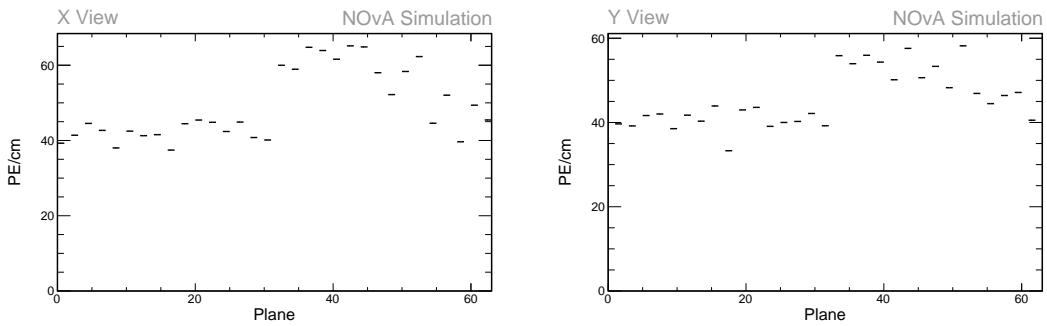


Figure 1.19: Uncorrected average energy response along planes for simulation.

Simulation Relative Calibration Results

An overview of the attenuation fit results for simulation are shown in Fig. 1.20 as a map of the average fitted response in the centre of each cell. The blank cells show the uncalibrated cells which failed the calibration condition (attenuation fit $\chi^2 > 0.2$). Most of the uncalibrated cells are on the edges of the detector, which is expected as those have much fewer events that pass the calibration sample selection than the rest. There are a total of 43 uncalibrated cells out of the total 4032 cells, resulting in 1.07% of the simulated detector uncalibrated.

Examples of the standard detector response and of the response for cells on the edge of the detector are shown on the top left plot and on the two bottom plots of Fig. 1.21 respectively. Here the red line shows the initial exponential fit and the blue line the final attenuation fit after the Locally Weighted Scatter plot Smoothing (LOWESS) correction, as described in Sec. ???. The cells on the edge of the detector failed the calibration conditions due to the low number of entries causing large fluctuation in the mean response.

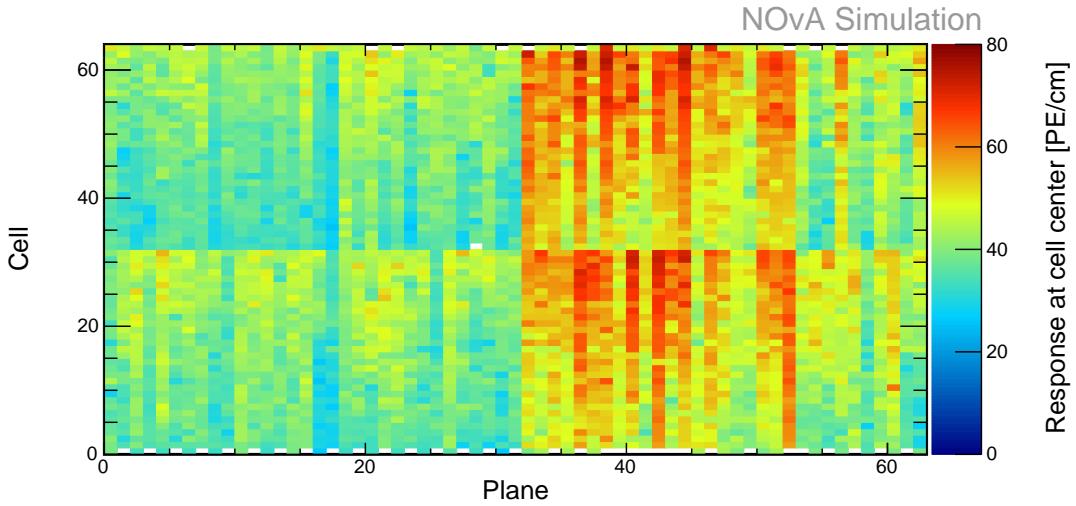


Figure 1.20: Overview of the attenuation fit results for the calibration of the simulated Test Beam detector. Each cell represents the result of the attenuation fit of the energy response in the centre of that cell. The blank cells are uncalibrated as the attenuation fit did not satisfy the calibration condition.

There is only one cell in the middle of the detector that is left uncalibrated. This is the cell 32 in a vertical plane in the brightness bin 5, shown on the top right of Fig. 1.21, with $\chi^2 = 0.227$. It seems the reason this cell has a $\chi^2 > 0.2$ and therefore failed the calibration condition is the unusually high response with a large uncertainty in the right-most bin. It is unclear why this bin has such an elevated mean response, but since this only causes an issue for a single cell, we decided to ignore it.

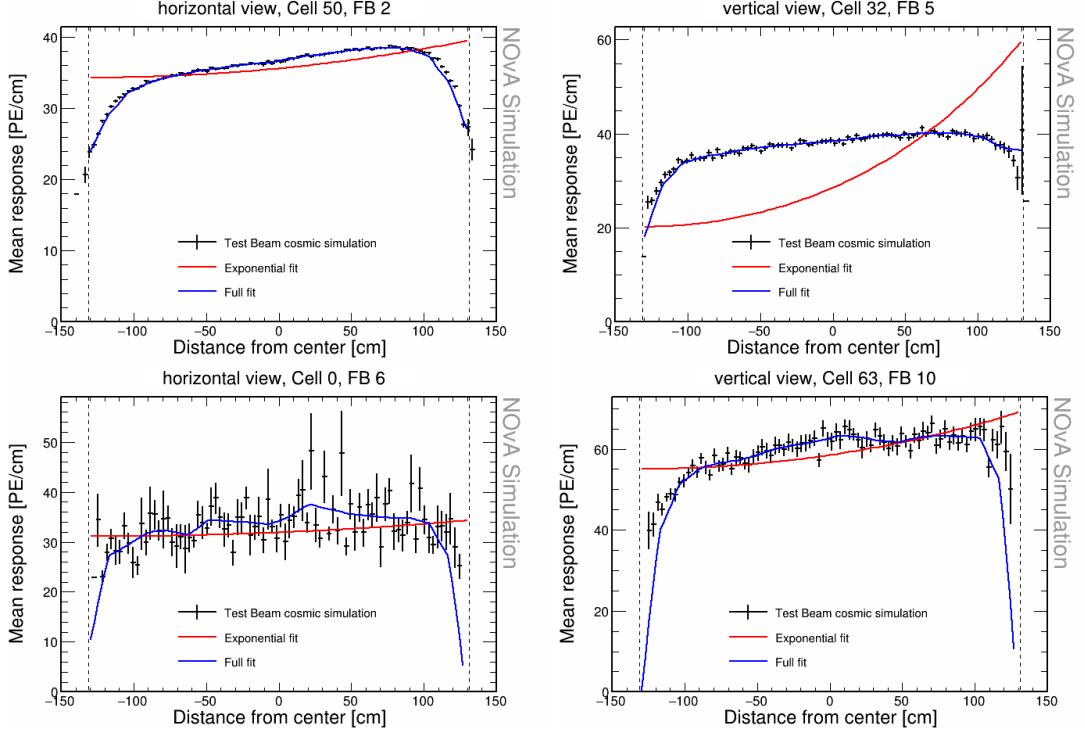


Figure 1.21: Attenuation fits for a selection of cells in the Test Beam calibration simulation. Top left is an example of a successful attenuation fit, top right is a failed fit due to statistical fluctuation in the last bin and the bottom plots show failed fits for cells on the edges of the detector.

1.3.4 Period 2 Data

The issue with underfilled cells described in Sec. 1.1 was present throughout the period 2 data taking. This can be clearly seen in Fig. 1.22, represented by the empty cells 31 and 63 in the horizontal planes, which were marked as bad channels and therefore ignored during production of calibration samples. This also affects the neighbouring cells to the underfilled cells, which have fewer events due to the tricell condition (see Sec. ??).

We can also see three noticeably darker spots than their neighbours. Specifically in plane 48 cells 38-40 and in plane 55 cells 2-4 and 45-47. These are all three cell wide, so it is likely that the issue is only in the middle cells and their immediate neighbours are affected due to the tricell condition. The three affected cells had most likely dead channels for some portion of period 2, with cell 46 in plane 55 being dead the longest. One possible explanation proposed for dead channels in plane 55 is that it is due to switched cables from the readout to the Data Acquisition (DAQ) [21], which is manifested as fewer total number of events in those cells. However, this wouldn't

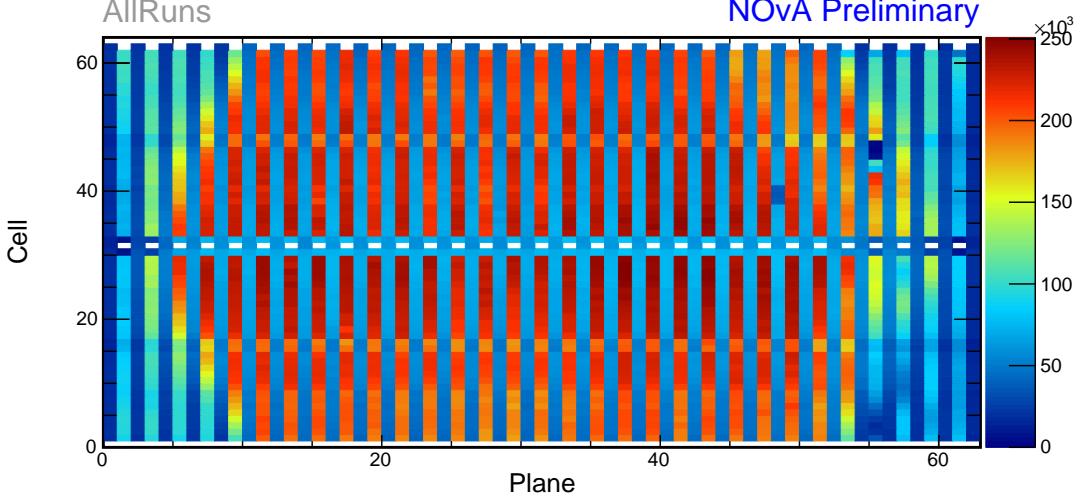


Figure 1.22: Distribution of events in the period 2 Test Beam data calibration sample.

explain the dead channel in cell 39 of plane 48, which was caused by a different issue.

Officially, period 2 is divided into 6 epochs labelled by letters, 2a - 2f, based on the specific running conditions. The epochs mostly differ in the use of various FEB firmwares or trigger studies. We compare the energy deposition in various epochs in Fig. 1.23, 1.24 and 1.25. As can be seen, the difference between the energy response across the individual epochs is fairly small and only in normalization. There's also no clear trend of energy response falling or raising with time. The largest outliers seem to be epochs 2a and 2d. Since each individual epoch would not have enough statistics for a successful attenuation fit, we decided to calibrate the entire period 2 together, without splitting it into any smaller samples.

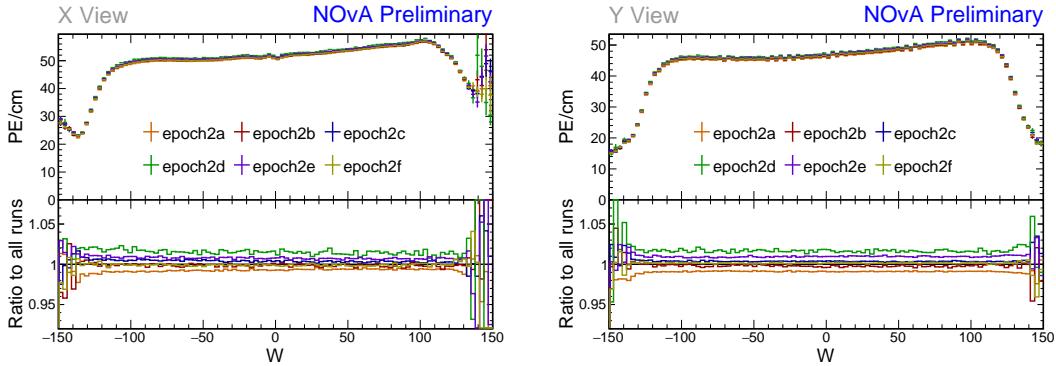


Figure 1.23: Uncorrected average energy response along the position within a cell (w) for epochs in period 2. It is clear that there is no significant difference between the various epochs.

The only variation of energy response in shape can be seen on the distributions along planes in Fig. 1.25. Here, in the top panel of the right plot, we can see that the

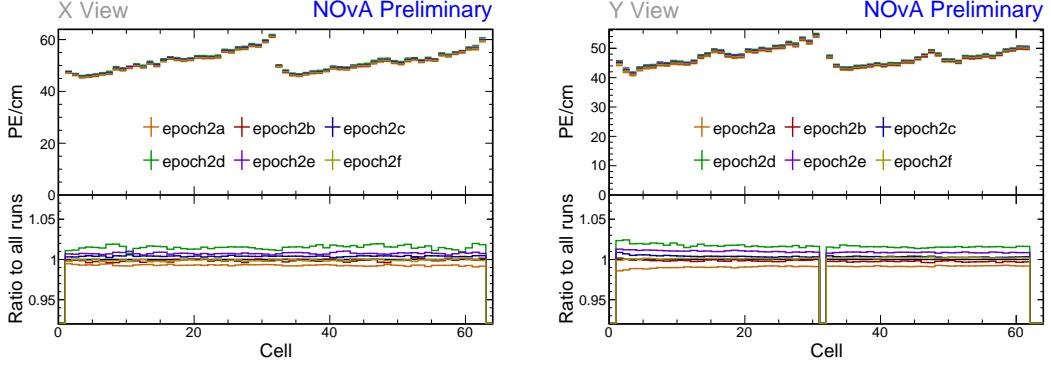


Figure 1.24: Uncorrected energy response along cells for epochs in period 2 data.

uncorrected response in plane 55 is noticeably higher than the rest of the detector. The exact reason for this is unknown, but it is likely caused by a fault in one of the two FEBs that make up the plane readout.

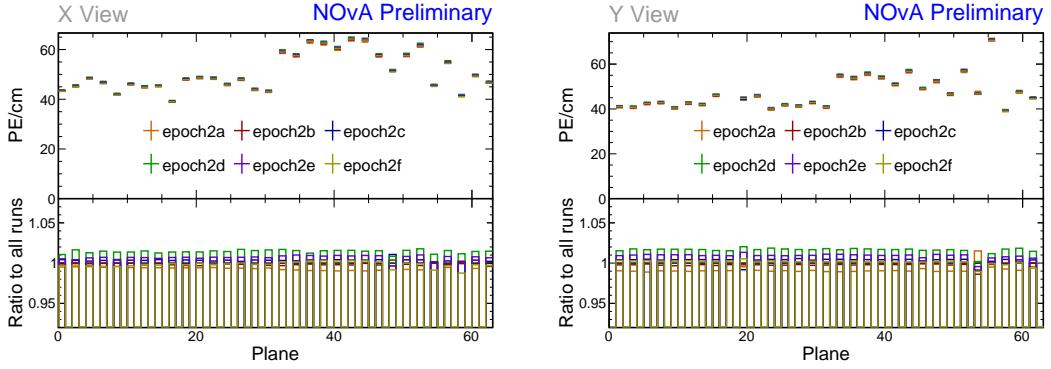


Figure 1.25: Uncorrected average energy response along planes for epochs in period 2 data.

Period 2 Relative Calibration Results

The results of the attenuation fit for period 2 are summarised in Fig. 1.26, showing the map of the fitted response at the centre of each cell. Same as for simulation, the blank cells failed the calibration condition for the attenuation fit. There are 199 cells that failed the calibration condition out of the total 4032 cells, constituting 4.94% of the detector left uncalibrated for period 2.

Most of the cells have an expected response, with steady rise towards the readout and a drop on the edges, as shown on the left plot of Fig. 1.27. This is the same as was shown for simulation.

Some cells have a non-regular response across the cell, with one or more regions with a drop in the energy response, as shown on the right plot of Fig. 1.27. These low

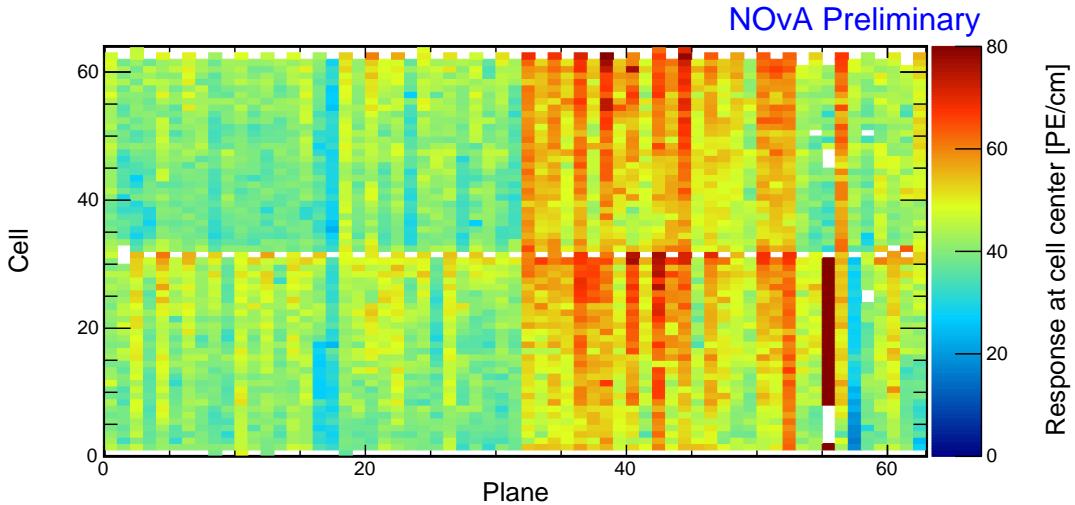


Figure 1.26: Overview of the relative calibration results for the Test Beam detector period 2 data. Each cell represents the result of the attenuation fit to the energy response in the centre of that cell. The blank cells are uncalibrated.

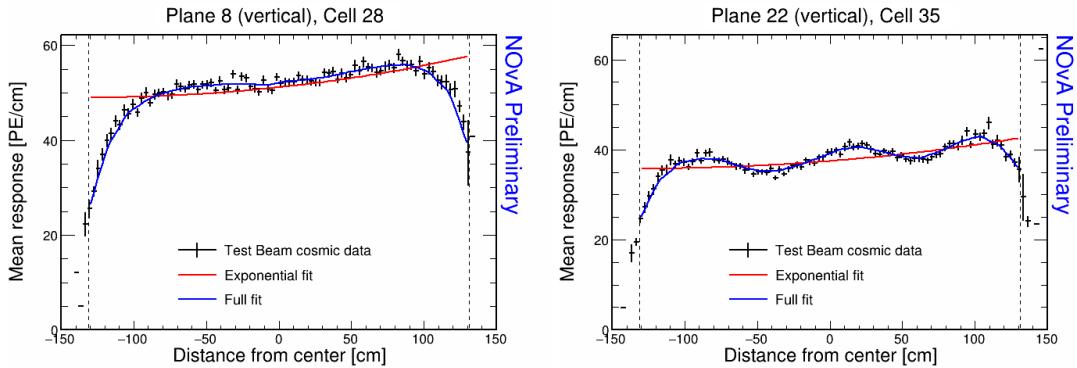


Figure 1.27: Attenuation fits for a selection of cells in period 2. Left plot shows an example of the standard energy deposition in the Test Beam. Right plot shows the effect of zipped fibres.

regions are (almost certainly) a real physical effect caused by zipped, or possibly even twisted, WLS fibres [22]. This effect is present in all the NOvA detectors. As can be seen, the attenuation fit is capable of fitting this response and therefore the relative calibration corrects for this effect in data. However, zipped fibres are not included in simulation for any of the detectors, which could potentially cause issues with the Analog-to-Digital Converter (ADC) threshold in simulation. It was decided that this is not does not have a significant impact and it would not be worth the amount of work required to include all the zipped fibres into the simulation.

Since the underfilled cells were marked as bad channels, we didn't attempt to calibrate them. Their neighbours have fewer events due to the tricell condition, but majority of them pass the calibration condition, as shown in Fig. 1.28. The decision

to mark the underfilled cells as bad channel was motivated by the fact that bad channels get skipped by the tricell condition and the neighbouring cells to the underfilled cells can therefore be included in calibration. The fact that majority of the neighbouring cells to the underfilled cells do get calibrated clearly proves that this was a good decision.

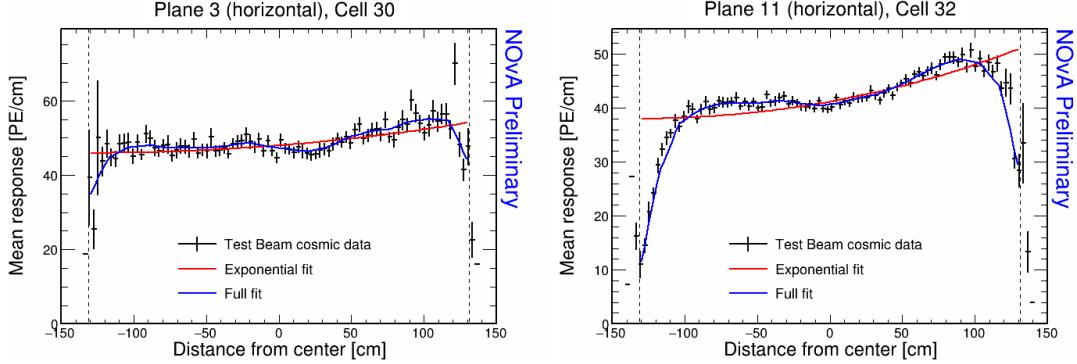


Figure 1.28: Fit to the energy response in period 2. The cells neighbouring the underfilled cells have fewer events and therefore larger fluctuations than the ‘usual’ Test Beam cell.

The neighbouring cells in plane 1 don’t pass the calibration condition due to low statistics and therefore large fluctuations, as shown in Fig. 1.29. This is likely due to a combination of the tricell condition and plane 1 being on the edge of the detector, which typically has fewer (accepted) hits than the center, as shown in Fig. 1.22.

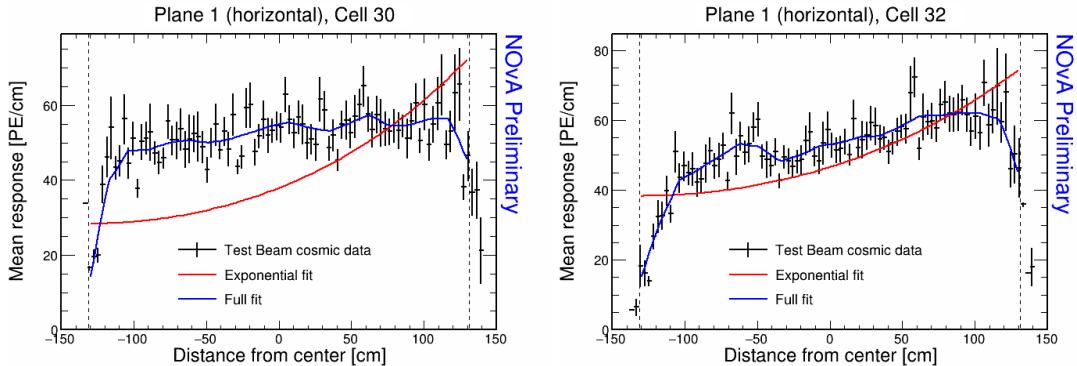


Figure 1.29: Fit to the energy response in period 2. The neighbouring cells to the underfilled cells in plane 1 are uncalibrated due to low statistics.

The left half of plane 55 has more than $3\times$ larger response than the surrounding planes, as shown on the left plot of Fig. 1.30. Similarly, the left half of plane 57 has slightly lower response than the surrounding planes, as shown on the right plot of Fig. 1.30. This is due to the corresponding APDs/FEBs incorrectly recording a scaled

up/down energy response than the real energy deposited in the detector. The cause of this scaled response is not known. Since this effect is present for all data, not only for the cosmic muons used for the calibration, it is important to correctly calibrate it out. A reason for concern is that this issue can arise if these FEBs were only affected for a limited time out of the entire calibrated period. Since we are doing the attenuation fit on the average response across the whole calibrated period, if an FEB records a standard response for half of the time and $7\times$ larger response for the second half, calibration is going to assume the response was $4\times$ larger the entire time, which would be incorrect. However, since both of the affected planes are in the back of the detector, we decided to ignore this effect for period 2.

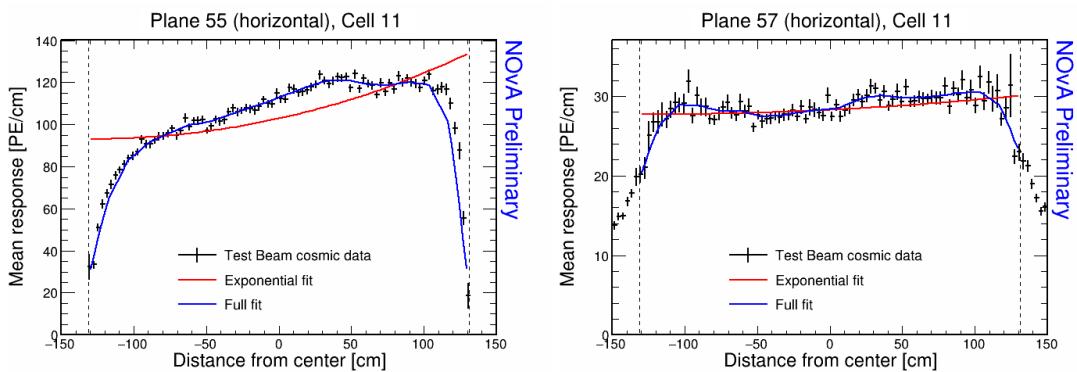


Figure 1.30: Fit to the energy response in period 2. Lower halves of planes 55 and 57 have a different scale of energy response than the surrounding planes.

As was discussed above, there are three dead channels with lower number of hits, possibly caused by swapped cables. These are located in planes 48 and 55. An example of one of the dead channels in plane 55 is shown on the left of Fig. 1.31 together with one of the neighbouring cells on the right. As can be seen in Fig. 1.26, the dead channel in plane 48 and its neighbours were successfully calibrated despite the lower number of hits.

Several cells in the end of the Test Beam detector are uncalibrated due to the histogram bins on the edges of the cell having an unusually high response, or no events at all, as shown in Fig. 1.32. It is unknown if this is a real physical effect, possibly related to the fibres, or if it is unfiltered noise hits, issue with the binning, or something else entirely. Since these cells are in the end of the detector, we decided to ignore them.

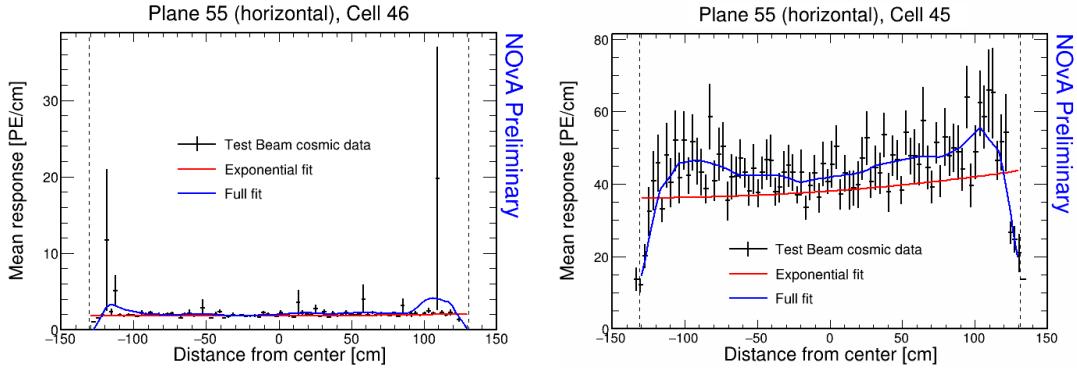


Figure 1.31: Fit to the energy response in period 2. Cells with swapped readout cables have almost no recorded events as shown on the left. This also affect their neighbouring cells due to the tricell condition as shown on the right.

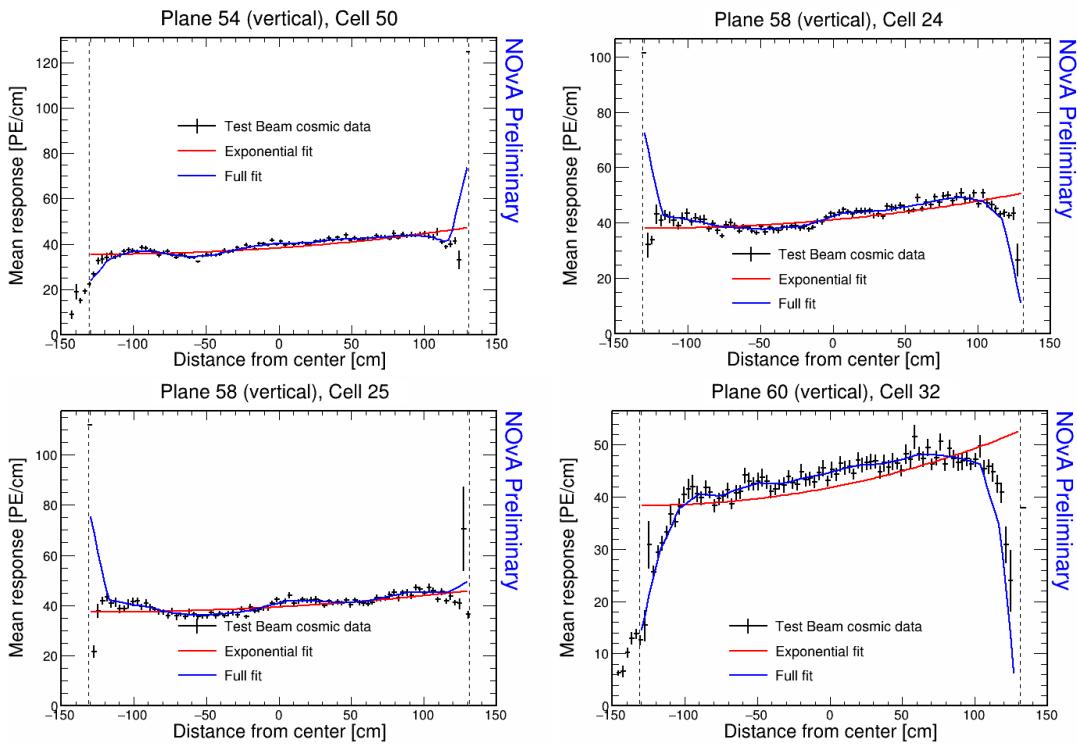


Figure 1.32: Fit to the energy response in period 2. Examples of cells that have an unusually high or low energy response at the edge of the cell. These cells are not calibrated.

1.3.5 Period 3 Data

The underfilled cells were refilled (or overfilled) during the period 3 data taking. This was the main motivation for dividing period 3 into individual epochs as shown on Tab. 1.3. Another major event that could impact the Test Beam data is the replacement of several faulty FEBs, which motivated the creation of epoch 3e.

Table 1.3: Test Beam period 3 epochs, their start dates and the reason for their separation.

Name	Start date	Reason for creating the epoch
Epoch 3a	January 12 th 2021	Underfilled cells
Epoch 3b	April 21 st 2021	Overfilling the back 9 horizontal planes and the 7th horizontal plane from the front
Epoch 3c	April 27 th 2021	Overfilling of the 15 front horizontal planes (except the 7th, which was already done) and the 14th horizontal plane
Epoch 3d	April 30 th 2021	Overfilling of the remaining 8 horizontal planes
Epoch 3e	May 12 th 2021	FEB swaps

The refilling of the underfilled cells can be clearly seen on the cell hits distribution in Fig. 1.33 and on the distribution of energy deposition across horizontal cells (Y view) in Fig. 1.35.

From the cell hits distributions in Fig. 1.33 we can also see there are a few channels (cells) that were likely dead for a certain time and weren't recording the same number of events as the surrounding cells. This is specifically plane 48, cell 39 in all of period 3 and plane 18, cell 31 in epochs 3d and 3e. Cell 39 in plane 48 was also affected in period 2.

The energy distributions across cells and planes in the X view (vertical) in Fig. 1.35 and 1.36 shows, that the top half of plane 58 has a very distinctly different energy deposition compared to the rest of the cells. Specifically, that the energy response in this module was larger in epoch 3a, then got lower in epochs 3b and 3c, until getting significantly lower for epochs 3d and 3e. However, Fig. 1.33 shows that module has the same number of events as the surrounding modules. This is one of the FEB that got replaced between epochs 3d and 3e. and as will be shown below this is the FEB with the largest impact on the calibration out of the faulty FEBs replaced before the

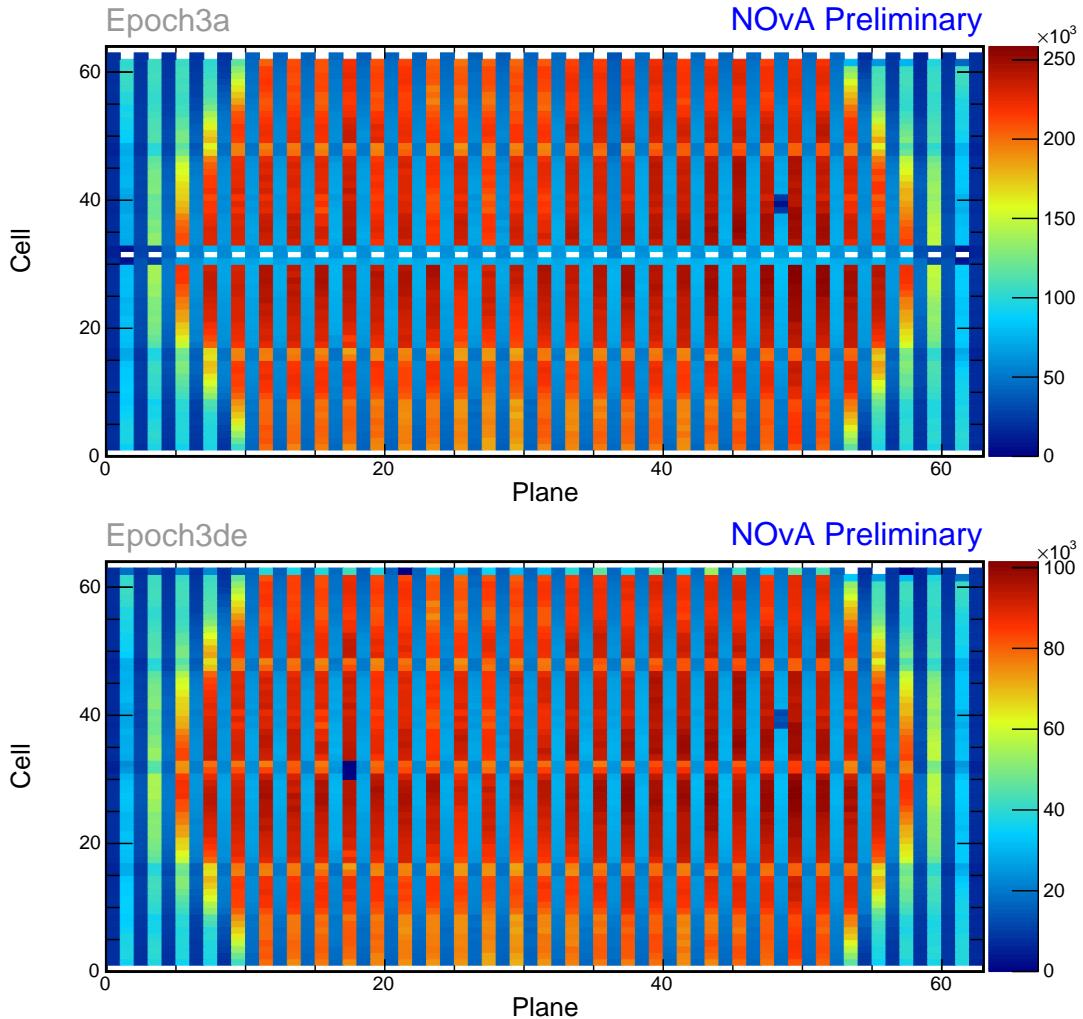


Figure 1.33: Distribution of events in the period 3 Test Beam data calibration sample. Comparison of Epoch 3a data before the refilling of the underfilled cells and Epoch 3de (combination of Epochs 3d and 3e) after the full refilling.

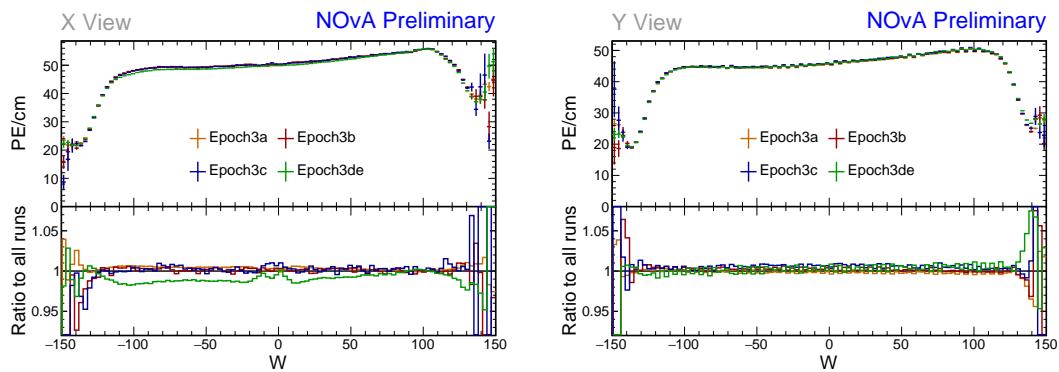


Figure 1.34: Uncorrected average energy response along the position within a cell for epochs in period 3 data.

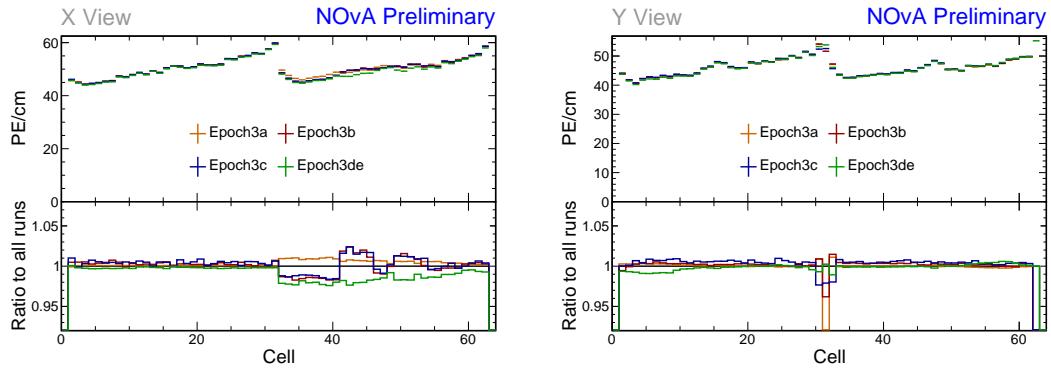


Figure 1.35: Uncorrected average energy response along cells for epochs in period 3 data.

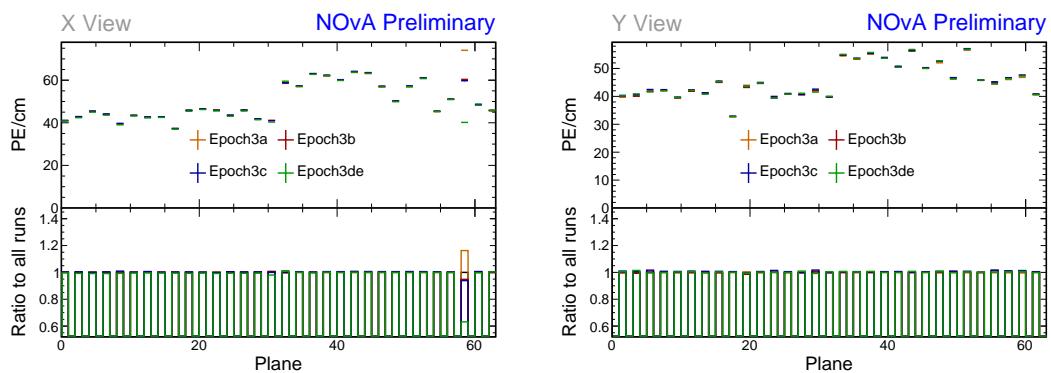


Figure 1.36: Uncorrected average energy response along planes for epochs in period 3 data.

start of epoch 3e.

From the aforementioned considerations, we decided to calibrate epochs 3a, 3b and 3c together, which are all the epochs containing any underfilled cells, and to separately calibrate epochs 3d and 3e together. The faulty **FEB** in the top of plane 58 is far enough in the back of the detector, that we didn't find it necessary to calibrate epochs 3d and 3e separately. Also epochs 3b and 3c only contain a few days worth of data, therefore they wouldn't have enough statistics for a successful attenuation fit.

Combined Epochs 3a, 3b and 3c Relative Calibration Results

The results of the attenuation fit for the combined epochs 3a, 3b and 3c are summarised in Fig. 1.37, showing the map of the fitted response at the centre of each cell. There are 182 uncalibrated cells out of 4032, constituting 4.51% of the detector.

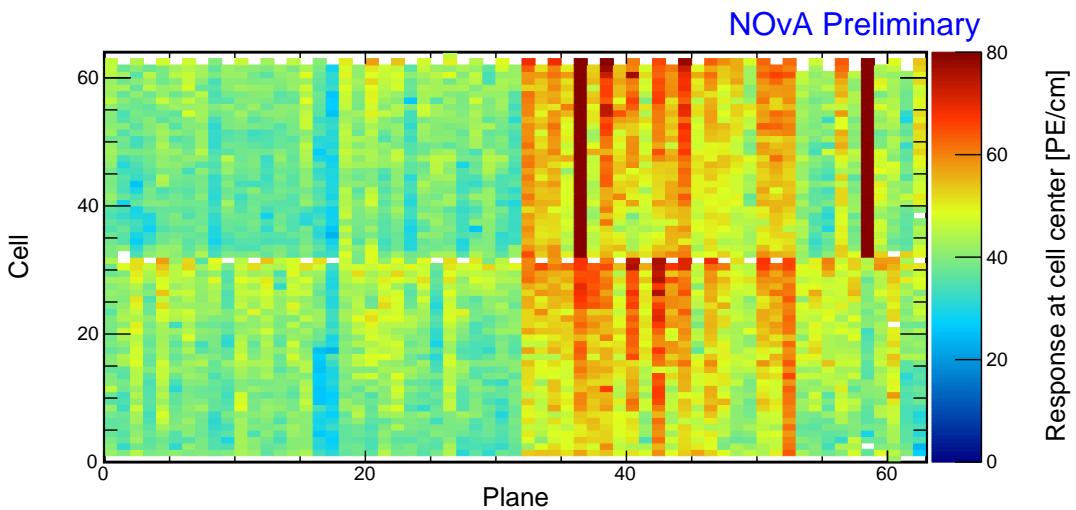


Figure 1.37: Overview of the relative calibration results for the Test Beam detector period 3, combined epochs 3a, 3b and 3c data. Each cell represents the result of the attenuation fit to the energy response in the centre of that cell. The blank cells are uncalibrated.

We can see that some of the underfilled cells that have been refilled for epochs 3b or 3c, but were underfilled for epoch 3a which makes up the majority of this calibrated data, are now calibrated thanks to including these two short epochs into the same attenuation fit. An example of energy deposition in such a cell is on the left side of Fig. 1.38.

Same as in period 2, most of the neighbouring cells to the underfilled cells are calibrated, except for cell 32 in plane 1, shown on the right of Fig. 1.38. This is due to

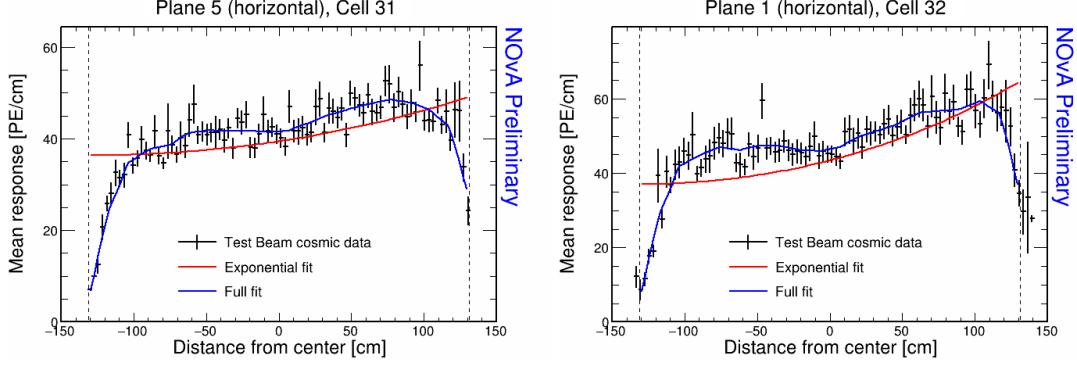


Figure 1.38: Fit to the energy response in epochs 3a, 3b and 3c. Some underfilled cells that have been refilled in epochs 3b and 3c are now calibrated as shown on the left plot. Cell 32 in plane 1 is the only neighbouring cell to the underfilled cell that didn't manage to get calibrated due to low number of events.

the low statistics at the edges of the detector, same as in period 2.

There is a couple of notably faulty FEBs with a different energy response than their neighbours. Besides the expected top half of plane 58, which has about $5\times$ larger response than the usual, there is also the top half of plane 36, which has about $2.5\times$ larger response as its neighbours. This could mean that the FEB in plane 36 was faulty only for a limited time compared to the FEB in plane 58. This is a reason for concern, as this could mean that using the results of this attenuation fit for hits in this module when the FEB wasn't faulty would give an incorrectly large correction (and therefore small 'corrected response'), whereas hits during the period when the FEB was faulty would have smaller than required correction (and therefore larger corrected response). Given that plane 36 is in the middle of the detector, this might affect some Test Beam analysis results. It is possible this might have to be mitigated in the future, whether with an additional uncertainty, or by improving the calibration.

COMMENT: *Should I talk about this further? We decided not to correct this since it would take more time and it might not be a huge problem. But this is probably the most problematic part of the calibration.* The energy deposition for these cells is shown in Fig. 1.39. As plane 58 is in the end of the detector and its readout was likely faulty for the majority of the calibrated period, we decided to ignore this FEB.

Similarly to period 2, there are a few cell in the back of the detector that have a sharp rise in the energy response at the edge of the cell, which causes the attenuation fit to fail the calibration condition. This can be seen in Fig. 1.40 with significantly different mean responses at the edge bins pulling the attenuation fit to incorrect val-

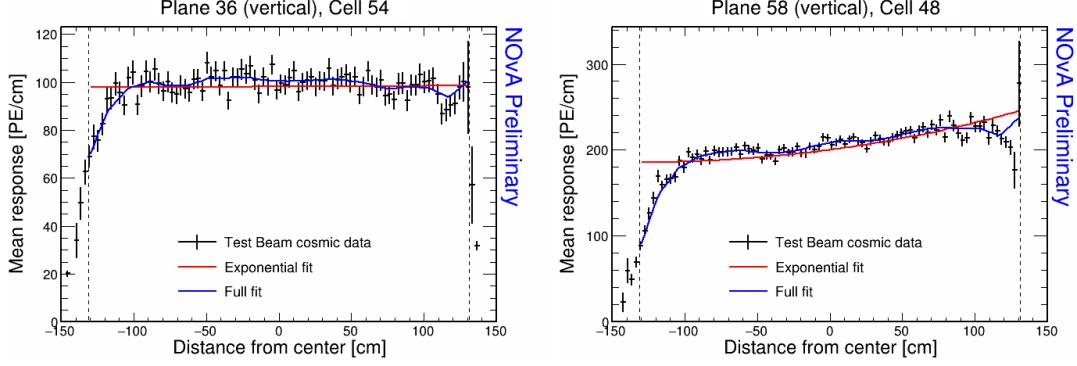


Figure 1.39: Fit to the energy response in epochs 3a, 3b and 3c. The most obvious faulty FEBs that have a significantly larger energy response than their neighbours.

ues. Given this is happening only in the end of the detector, we decided it should be safe to ignore this effect.

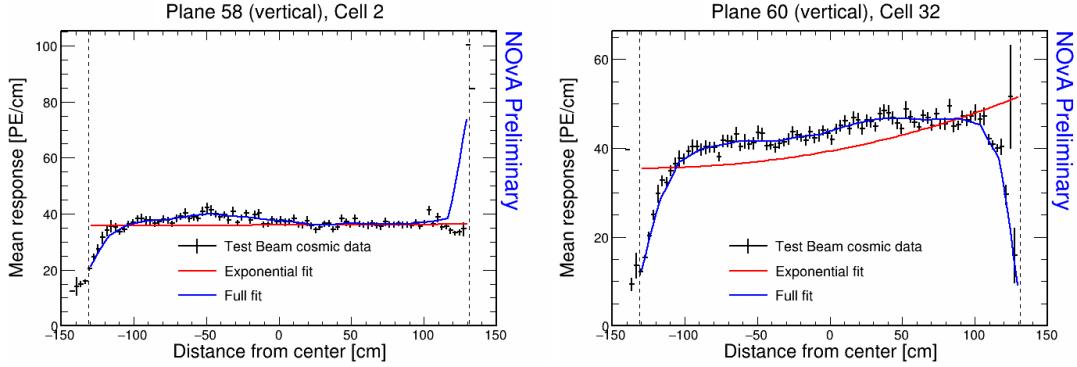


Figure 1.40: Fit to the energy response in epochs 3a, 3b and 3c. Some cells are not calibrated due to large fluctuations at one edge of the cells.

Combined Epochs 3d and 3e Relative Calibration Results

The results of the attenuation fits for epochs 3d and 3e are shown in Fig. 1.41. There are 182 uncalibrated cells out of 4032 total cells, making up 4.51% of the detector. The uncalibrated are now however almost entirely concentrated at the edges and at the end of the detector.

Figure 1.41 shows the expected uncalibrated cells in plane 17 surrounding the dead channel discussed above (or possibly still an underfilled cell). The energy deposition for this cell and one of its neighbours is shown in Fig. 1.42.

Epochs 3d and 3e should have all the previously underfilled cells now refilled, but as can be seen in Fig. 1.41, there's several of these cells that are still uncalibrated. The energy deposition in these cells is shown in Fig. 1.43. Here we can see, that these

NOvA Preliminary

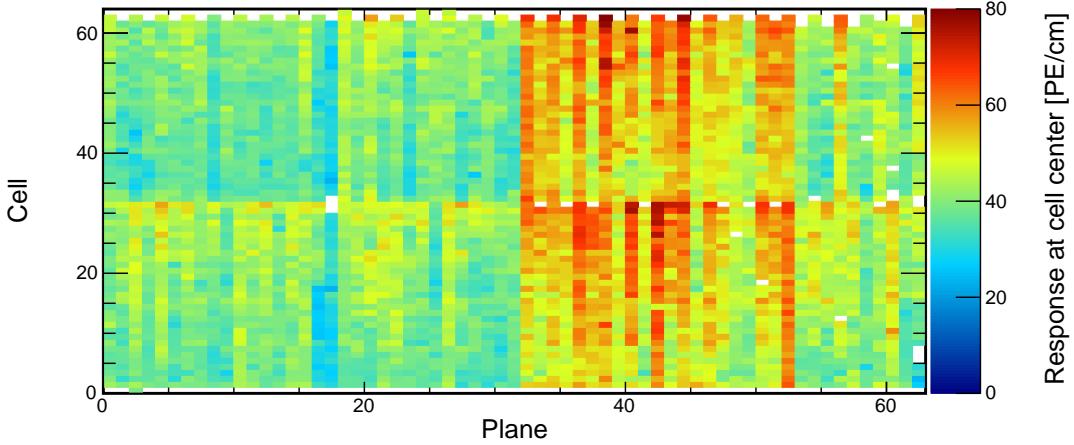


Figure 1.41: Overview of the relative calibration results for the Test Beam detector period 3, combined epochs 3d and 3e data. Each cell represents the result of the attenuation fit to the energy response in the centre of that cell. The blank cells are uncalibrated.

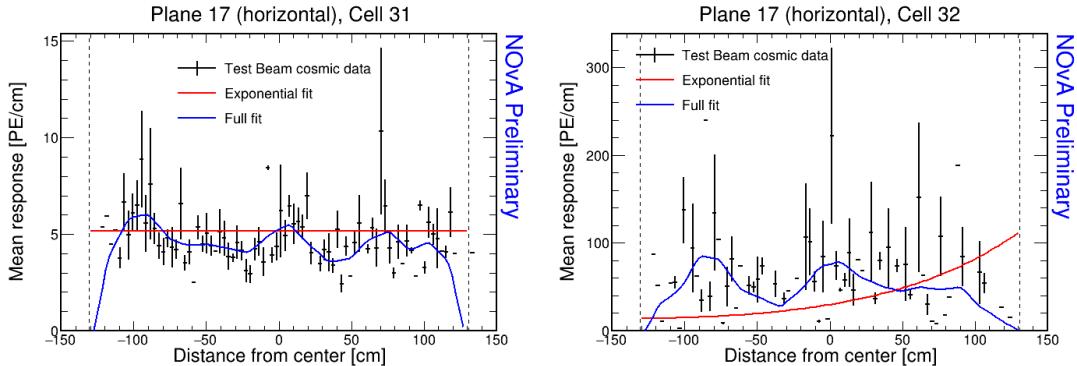


Figure 1.42: Fit to the energy response in epochs 3d and 3e. Possibly dead channel or still underfilled cell.

cells have a fairly large discrepancy between the left and right side of the cells. This is caused by using different scintillator oils for the initial filling of the cells and for the refilling. Specifically, as was described in Sec. 1.1, these cells have been initially filled with the Ash River and the Texas oils, which have higher energy depositions compared to the NDOS oil that was used for the refilling during period 3. These oils clearly didn't mix properly, which caused a difference in energy deposition in different parts of the cells.

This is a physical effect that should be accounted for in the calibration and as we can see, the attenuation fits are actually performing reasonably well. Additionally, these cells are in the middle of the detector and leaving them uncalibrated would almost certainly have an impact on the Test Beam analyses. The large χ^2 value of the

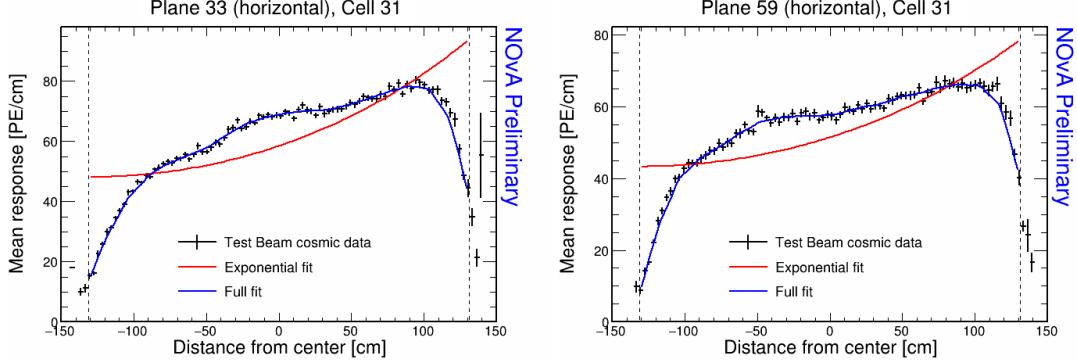


Figure 1.43: Fit to the energy response in epochs 3d and 3e. The scintillator oil used for refilling of the underfilled cells has lower energy response than the oil used for the initial filling. These oils didn't mix properly causing a different energy response in the left and right side of the cell.

attenuation fit is most likely caused only by the unusual shape of the distribution, which the fit is not designed for. We have therefore decided to manually change the χ^2 values for these cells inside the cvs tables (which hold the results of the attenuation fits), so that their $\chi^2 < 0.2$ and these cells are officially considered calibrated, even if they originally weren't. The ‘corrected’ distribution of the attenuation fit results for epochs 3d and 3e is shown in Fig. 1.44. The number of uncalibrated cells was reduced by 8 to 174, or 4.32% of the detector.

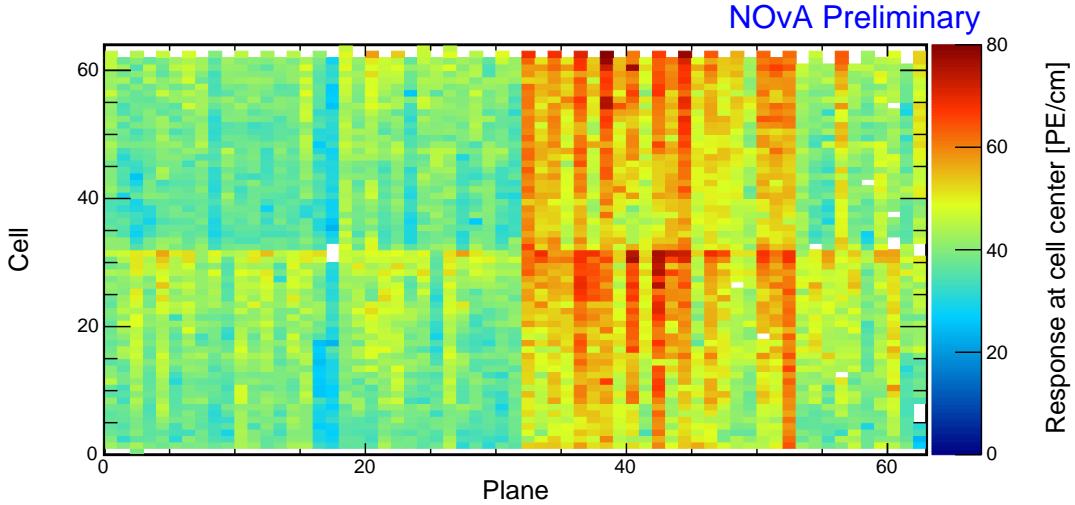


Figure 1.44: Overview of the relative calibration results for the Test Beam detector period 3, combined epochs 3d and 3e data. Each cell represents the result of the attenuation fit to the energy response in the centre of that cell. The blank cells are uncalibrated. The previously uncalibrated cells in the middle of the detector were artificially marked as calibrated after careful considerations.

Some of the cells in the back of the detector have a rise, or drop in energy depo-

sition at the edge of the cell, as can be seen in Fig. 1.45. This is similar to the effect seen in period 2 and epochs 3a+3b+3c and since it's again concentrated in the end of the detector, we ignored these cells and left them uncalibrated.

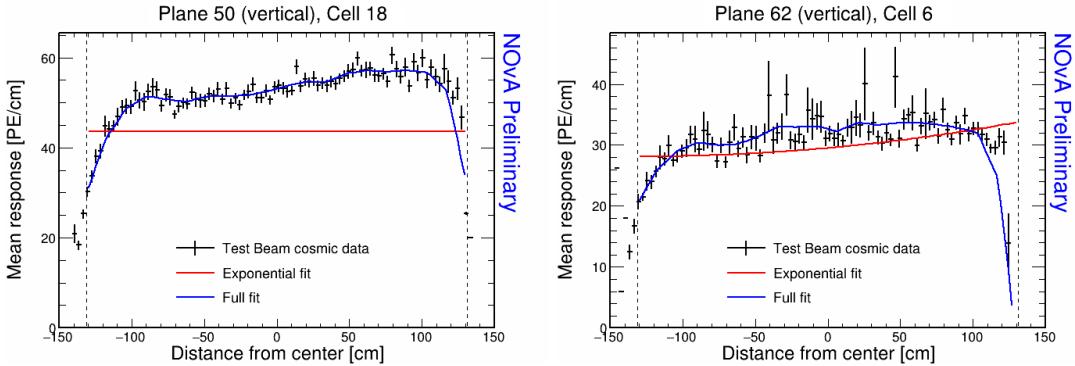


Figure 1.45: Fit to the energy response in epochs 3d and 3e. Some cells have a drop, or a rise of energy response at the edge of the cell. This can be cause by low statistics.

1.3.6 Period 4 Data

The period 4 Test Beam data taking period is the best data we managed to collect with almost ideal detector conditions. There are a few commissioning runs in the very beginning of period 4, which uncovered some dead channels or faulty FEBs that were immediately fixed. These runs constitute epoch 4a, shown on the top of Fig. 1.46. There is also a few runs during which we performed studies in which we masked parts of the detector to help with the FEB saturation [23], which can clearly be seen in the middle of Fig. 1.46. Bottom part of Fig. 1.46 shows that the rest of the period 4 data does not have any noticeable faults in the distribution of hits across the detector.

Figures 1.47, 1.48 and 1.49 show that the epoch 4a and the cell masking study did have a noticeable impact on the energy deposition across the detector. Both of these special periods only span a short time and contain a very limited number of hits. We therefore decided to ignore these runs and only calibrate the rest of the period 4 data, using these results for all runs in period 4. *COMMENT: I assume here that the runs from the cell masking studies will not be used in the TB analyses. Is that correct?*

Period 4 Relative Calibration Results

Results of the attenuation fits for period 4 are summarised in Fig. 1.50. We can see that almost the entire detector is now calibrated. The only exceptions consists of cells

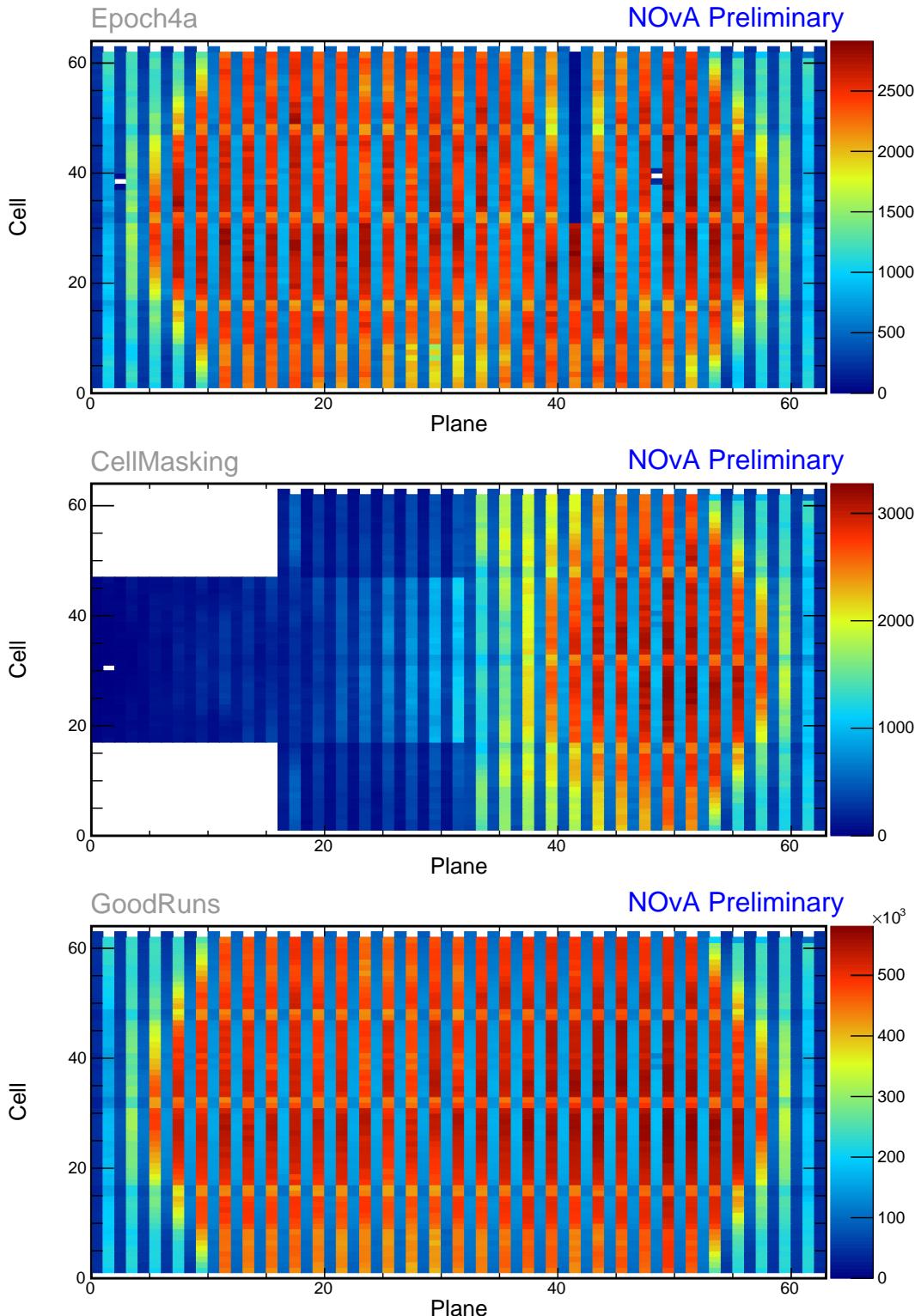


Figure 1.46: Distribution of events in the Test Beam period 4 data calibration sample. The top plot shows the first three commissioning runs, the middle plot the status of the detector during the Cell Masking studies and the bottom plot shows the rest.

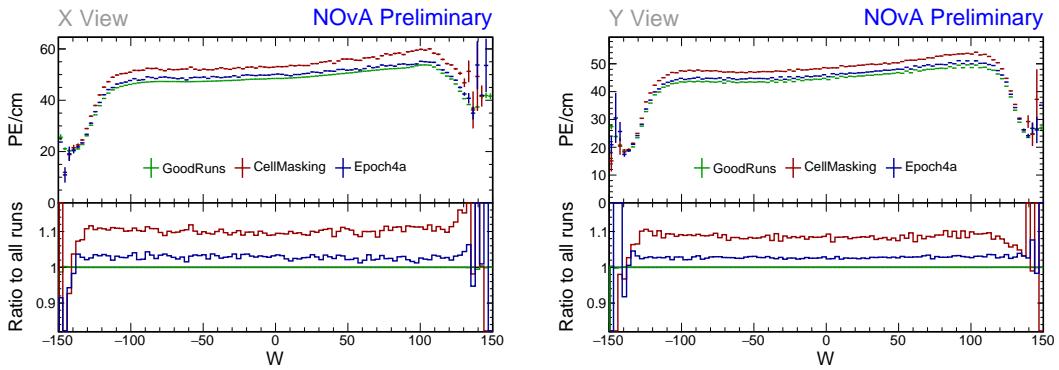


Figure 1.47: Uncorrected average energy response along the position within a cell (w) for epochs in period 4 data.

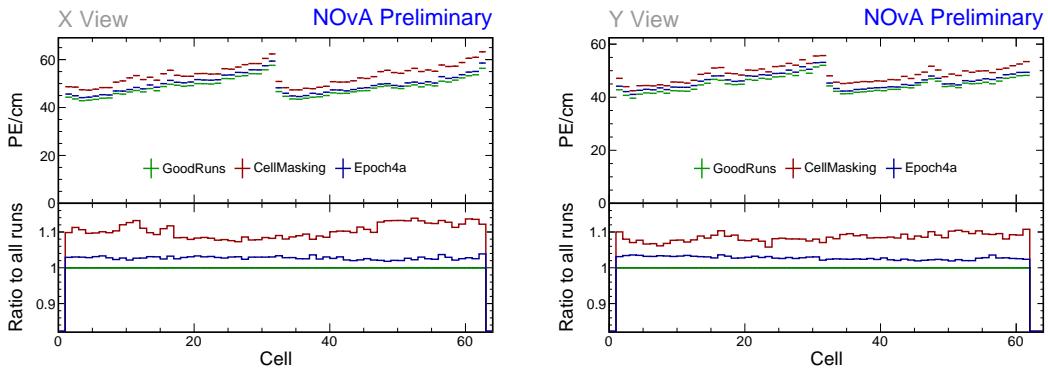


Figure 1.48: Uncorrected average energy response along cells for epochs in period 4 data.

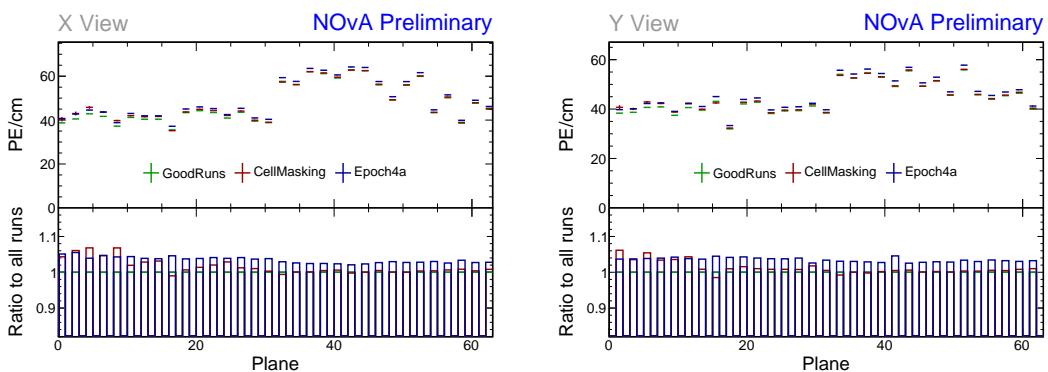


Figure 1.49: Uncorrected average energy response along planes for epochs in period 4 data.

on the edge of the detector, 7 formerly underfilled cells (left plot in Fig. 1.51), and one cell 47 in plane 54 with an unusually high response at the edge of the cell (right plot in Fig. 1.51).

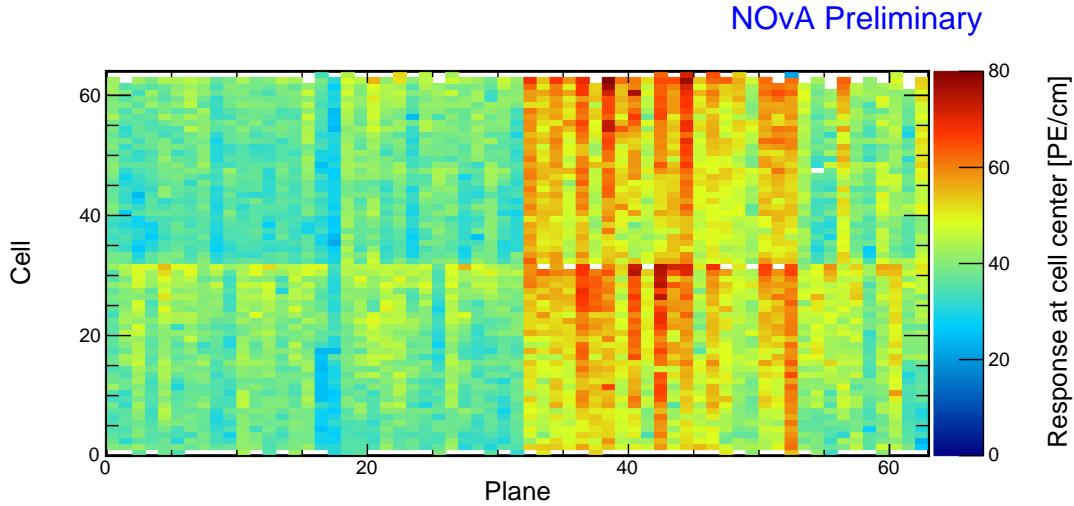


Figure 1.50: Overview of the relative calibration results for the Test Beam detector period 4 data. Each cell represents the result of the attenuation fit to the energy response in the centre of that cell. The blank cells are uncalibrated.

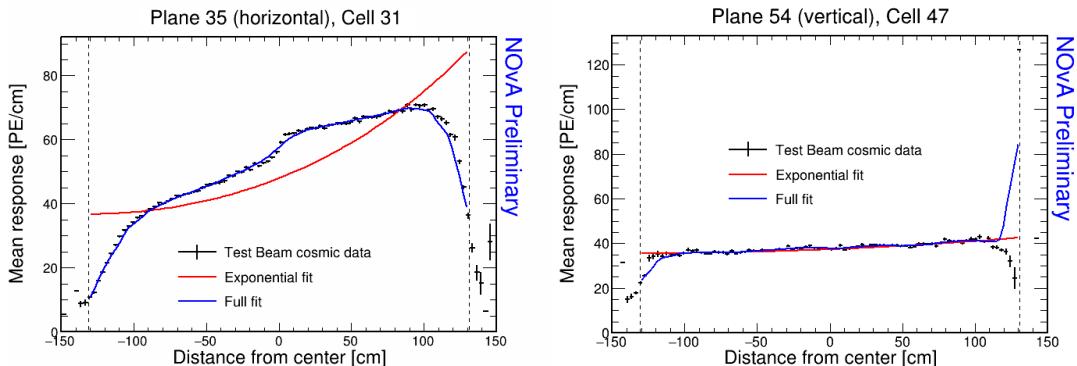


Figure 1.51: Fit to the energy response in period 4. Previously underfilled cells refilled with a scintillator of a different quality causing an unusual distribution of energy deposition (left). Unusually high energy response at the edge of the cell 47 (right).

We treated the formerly underfilled cells the same way as in epochs 3d and 3e, by manually changing the χ^2 of their attenuation fits inside the csv files to be < 0.2 and therefore making them officially calibrated. The ‘corrected’ distribution of the relative calibration results is shown in Fig. 1.52. There are 108 uncalibrated cells out of 4032, totalling 2.68% of the detector, almost exclusively on the edges.

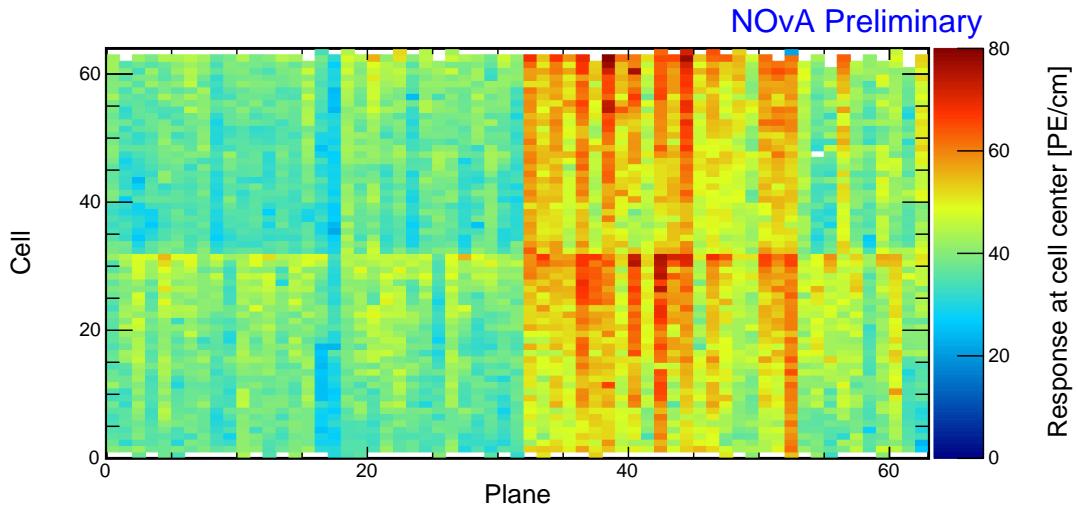


Figure 1.52: Overview of the relative calibration results for the Test Beam detector period 4 data. Each cell represents the result of the attenuation fit to the energy response in the centre of that cell. The blank cells are uncalibrated. The previously uncalibrated cells in the middle of the detector were artificially marked as calibrated after careful considerations.

End of the current version of the thesis!

Anything after this point was simply copied from the technote and hasn't been review yet

1.3.7 Absolute Calibration Results

As was described in Sec. ??, we apply the relative calibration results on the stopping muon sample and use the absolute calibration cuts to select only well-understood minimum ionising muons. These events represent a very well understood source of energy deposition we use as a standard candle to calculate the absolute energy scale. The absolute calibration cuts are mostly the same as for the other NOvA detectors, selecting hits $1 - 2$ m from the end of their tracks and removing uncalibrated and wrongly reconstructed hits by requiring non-zero pathlength, $\text{PhotoElectron(PE)} > 0$, and $\text{CorrectedPhotoElectrons(PECORR)} > 0$, as well as $\text{PECORR}/\text{cm} < 100$. To remove hits at the edge of the cells, we require that $-80 < w < 80$ cm, therefore removing hits approximately 0.5 m from each side of the detector.

The distributions of the reconstructed and true energy responses, for both views, and for each data and simulation sample, are shown in Fig. 1.53. The means of these distributions are the $\text{MuonEnergyUnit(MEU)}_{\text{Reco}}$ and MEU_{True} values for each view. The uncertainty on the means are calculated as the standard deviation of the distributions divided by the square root of the number of entries. We average the two MEU values from both views and combine the uncertainties in the sum of squares. The total number of entries, the MEU values for each sample and view, and the combined MEU values with corresponding uncertainties are shown in Tab. 1.4. Given the large number of entries in the energy response distributions, the total uncertainties on the MEU values are negligible (around 0.05%). This are however not the final uncertainties of the absolute energy scale used in NOvA. Instead, we use comparison to other standard candles, as was explained in Sec. ??.

Table 1.4: Summary of absolute calibration results. MEU_{Reco} values (top table) are in units of PECORR/cm and MEU_{True} values (bottom table) are in units of MeV/cm

Sample		X view		Y view		Combined	
		NHits	MEU	NHits	MEU	MEU_{Reco}	$\sigma_{\text{MEU}_{\text{Reco}}}$
Data	Period 2	2.322e+05	38.70	1.413e+06	39.40	39.05	0.02
	Epochs 3abc	2.638e+05	38.49	1.621e+06	39.40	38.94	0.02
	Epochs 3de	1.049e+05	38.63	6.725e+05	39.42	39.02	0.03
	Period 4	5.268e+05	38.63	3.316e+06	39.40	39.01	0.01
Simulation		2.829e+05	40.17	1.842e+06	39.93	40.05	0.02

$$\text{MEU}_{\text{True}} = 1.7722 \text{ MeV/cm} \quad \sigma_{\text{MEU}_{\text{True}}} = 0.0003 \text{ MeV/cm}$$

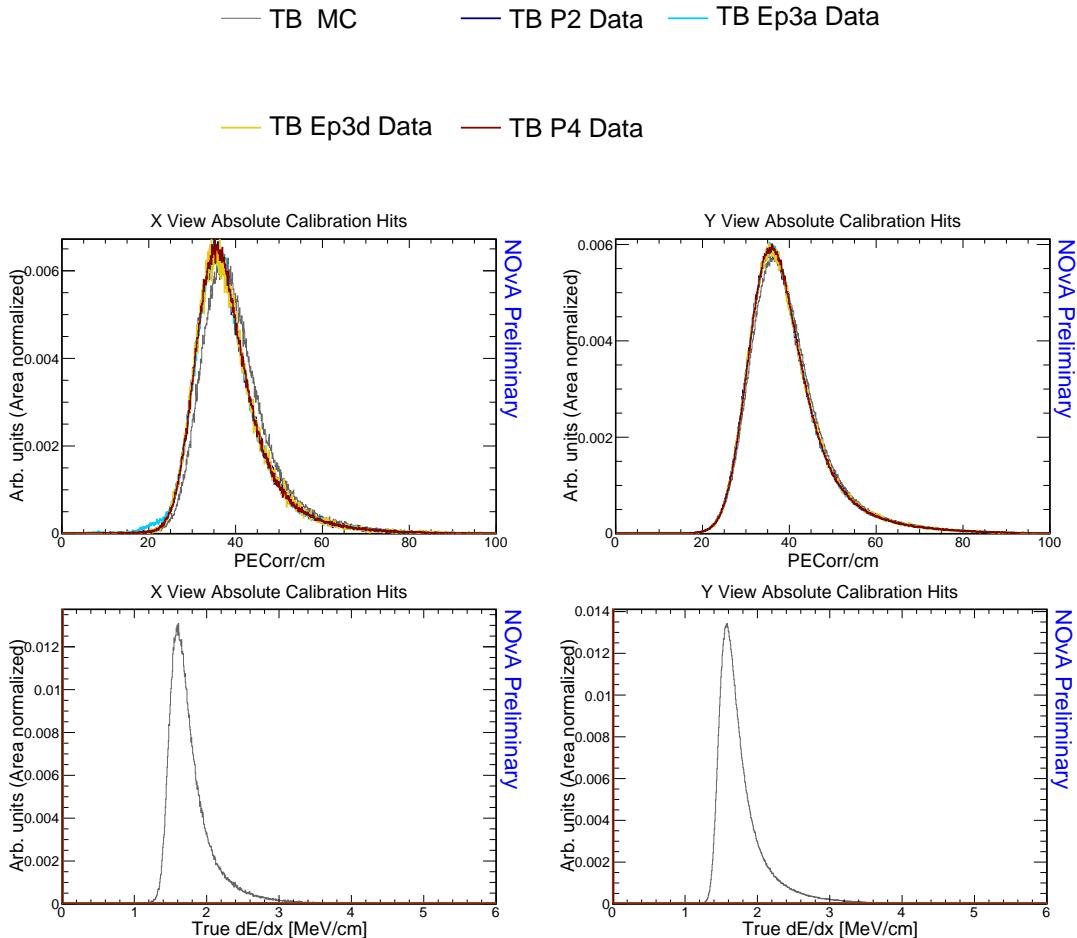


Figure 1.53: Distributions of the reconstructed (top) and true (bottom) energy response of stopping muons in the X (left) and Y (right) view within a 1 – 2 m track window from the end of their tracks. The mean of the reconstructed and true distributions of the response are the reconstructed and true MEU values respectively for the corresponding views.

1.3.8 Discussion

To validate the results of the Test Beam calibration we look at the stopping muon sample used for the absolute calibration, since these events have the most consistent and reliable energy deposition.

In plots in Fig. 1.54-?? we look at distributions of variables used during the calibration, namely PE , $PECorr$, $Pathlength$, PE/cm and $PECorr/cm$. Their distributions are over a range of variables we tried to correct the energy deposition in, namely position within a cell w , cell number, plane number, track angles and time.

The most important validation plots are the distributions of $PECorr/cm$, which should be completely flat. This would mean that all the deposited energy results in

an equivalent recorded energy wherever and whenever in the Test Beam detector it occurred. As can be seen on the validation plots, this was successfully achieved and the PEC_{corr}/cm distributions are mostly flat across all studied variables.

The distribution of PEC_{corr}/cm across cells in X view in Fig. 1.56 seems fairly scattered, however this is mostly due to the better resolution of this plot and the dispersion of the energy deposition across cells isn't large enough to constitute further investigation.

The distributions of PEC_{corr}/cm across planes in the X view (Fig. 1.56) shows a noticeable smaller corrected energy response of stopping muons in plane 36. This means that the relative calibration over-corrected the energy response due to the through-going muons having unusually high energy response (as shown in Fig. 1.39), but not the selected stopping muons. The most likely cause is that the impacted FEB was ‘faulty’ only for a certain period of time. In that case the corrected energy response would be correct for the period when the FEB was faulty, but would be under-estimated for the period when the FEB behaved ‘normally’. The PEC_{corr}/cm over Plane plot shows the average over these responses.

The corrected response across planes in Y view (Fig. 1.56) shows a slight incline in the first half of the detector. We do not know where does this slope come from, but it is not big enough to be of concern and we decided to ignore it.

The distributions of energy deposition in time (Fig. 1.57 and ??) show a non-trivial dependency. The detector response could be influenced by environmental factors (temperature and humidity) and by scintillator or readout ageing. Neither of these factors are well understood within NOvA and Test Beam detector could be potentially used to shine more light on this issue. However this is a topic for a separate study and is out of scope of this technical note.

Technically, we would expect the distributions of PEC_{corr}/cm to also have the same **scale** for all data samples and for simulation. As can be seen on all the validation plots, the data samples have a reasonably similar scale of PEC_{corr}/cm , but this is noticeably different for simulation. This is caused due to the data-based simulation we are using does not have a correct energy estimation for through-going muons, which have generally underestimated energies [?]. This results in an over-estimated correction from the relative calibration. However, this is not an issue, since we only

use stopping muons to calculate the absolute energy scale and stopping muons have correct energies in the new simulation.

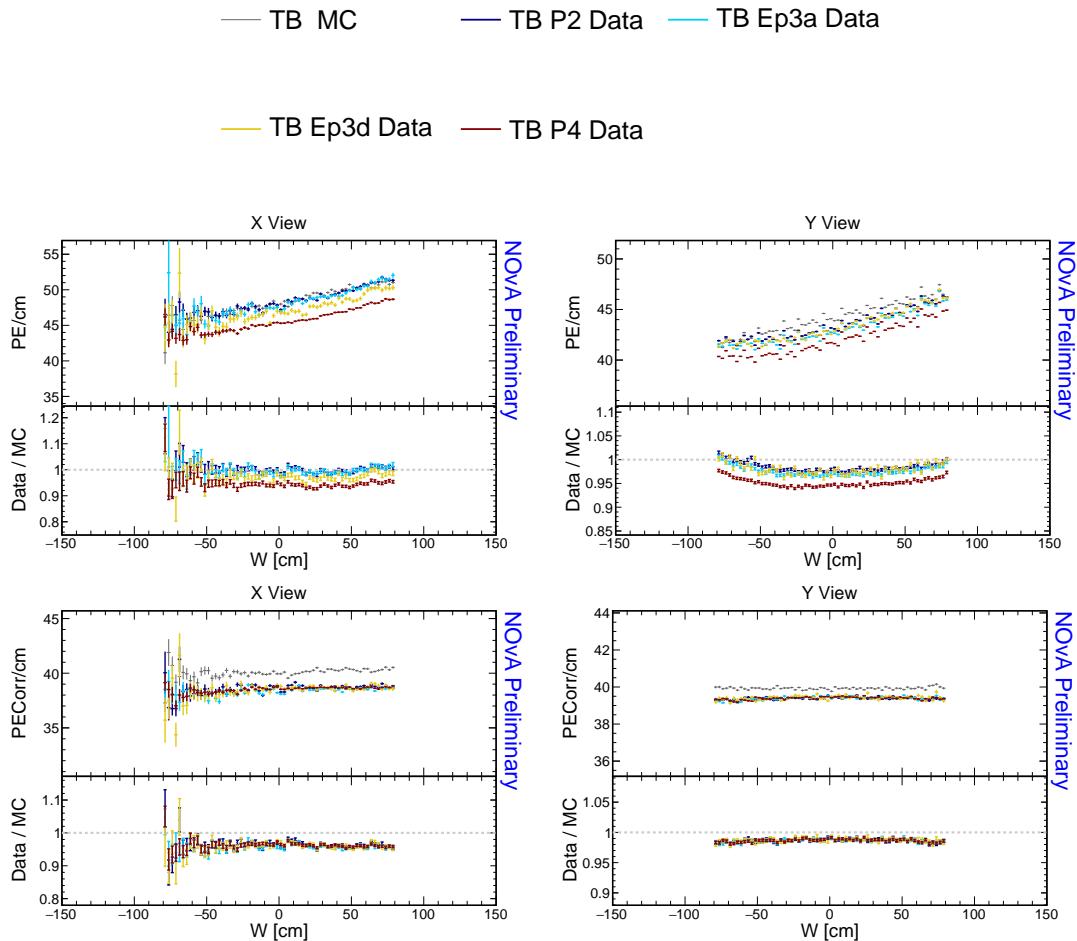


Figure 1.54: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the position within a cell.

1.4 Conclusion

The results of the relative and the absolute calibration, in form of the csv files, are stored in and are applied within NOvAsoft in the calibration tag v15.09 and higher.

TO DO: Include a table with the percentages of calibrated cells

We have successfully calibrated the NOvA Test Beam detector for all the Test Beam run periods in both data and simulation. The calibration results are implemented in the v15.09 version of the NOvAsoft calibration tag. We haven't attempted to estimate the uncertainty of the calibration, which is a separate task out of scope of this technical note.

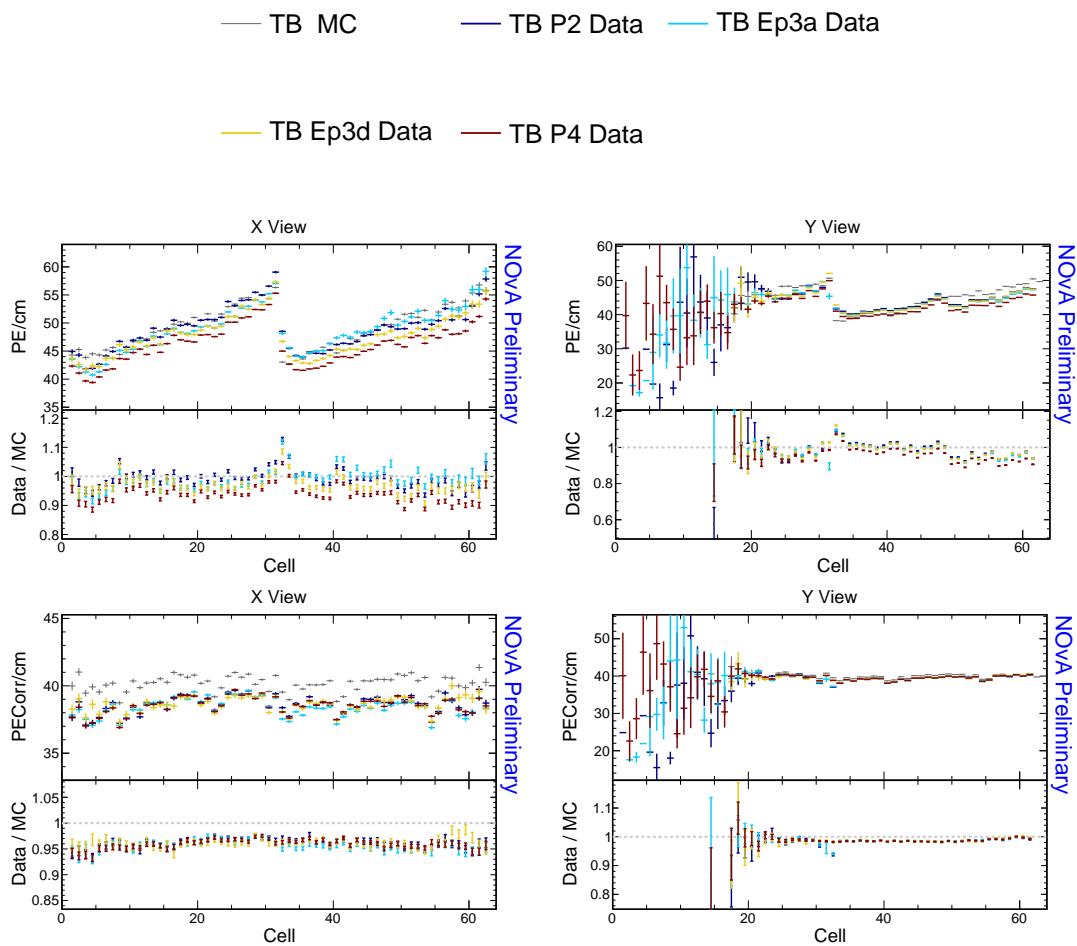


Figure 1.55: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the cells of the detector.

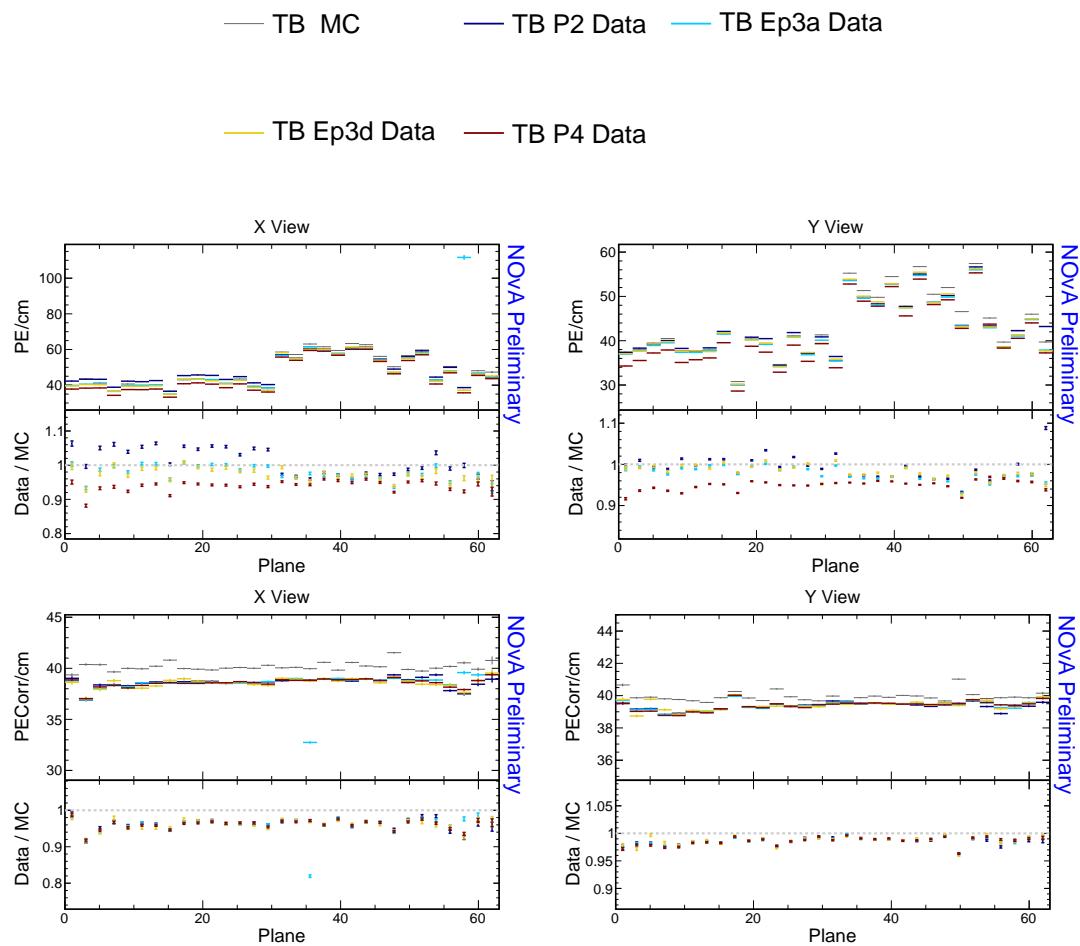


Figure 1.56: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the planes of the detector.

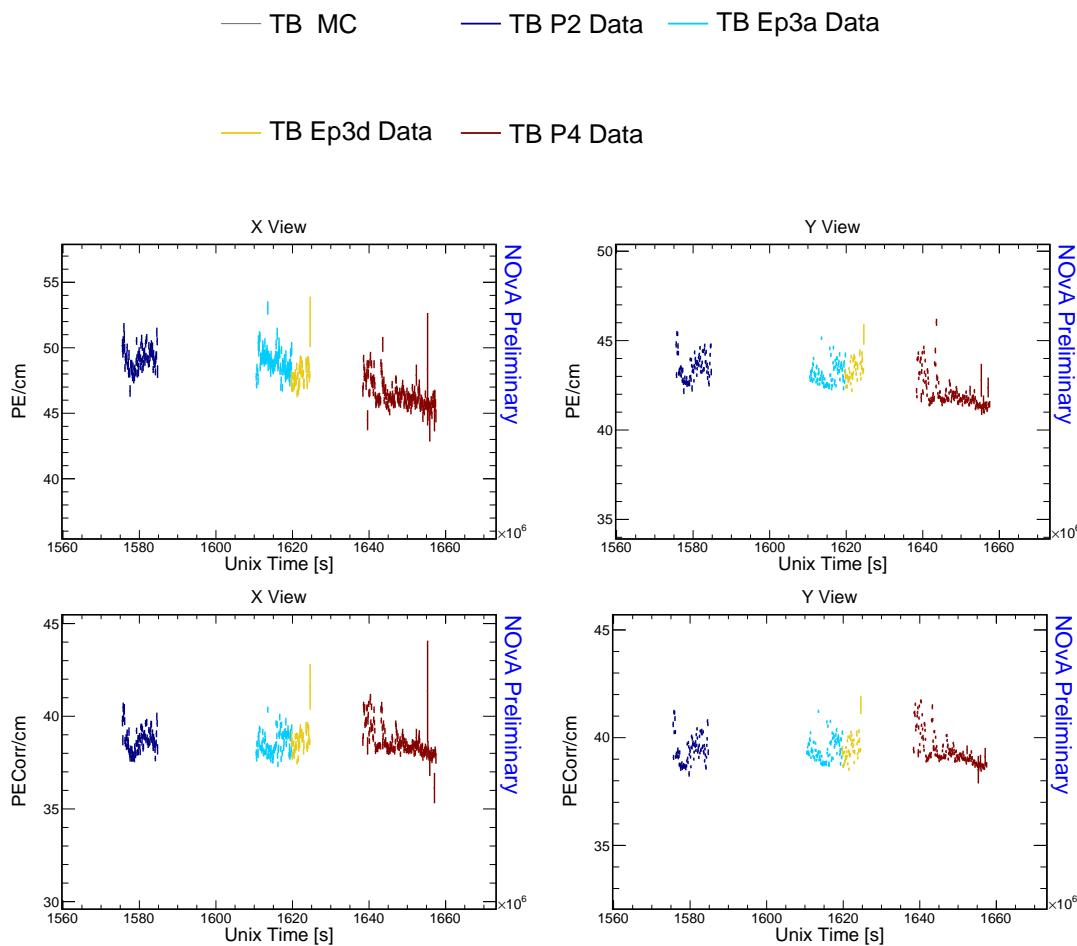


Figure 1.57: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the event UNIX time.

Acronyms

ADC Analog-to-Digital Converter. [33](#)

APD Avalanche Photodiode. [27](#), [34](#)

BPF Break Point Fitter. [9](#), [11–16](#), [18–23](#)

CRY Cosmic-Ray Shower Generator. [7](#)

DAQ Data Acquisition. [30](#)

FD Far Detector. [1](#), [2](#), [4–6](#), [9](#), [17](#), [19](#), [24](#)

FEB Front End Board. [6](#), [7](#), [32](#), [34](#), [35](#), [37](#), [40](#), [41](#), [45](#)

FERMILAB Fermi National Accelerator Laboratory. [5](#)

FTBF Fermilab Test Beam Facility. [2](#)

LOWESS Locally Weighted Scatter plot Smoothing. [28](#)

MC Monte Carlo. [7](#)

MEU Muon Energy Unit. [50](#)

MWPC Multiwire Proportional Chamber. [3](#)

ND Near Detector. [1](#), [2](#), [4–7](#), [9](#), [15](#), [17](#), [19](#)

NDOS Near Detector on the Surface. [4](#), [5](#), [7](#), [43](#)

NOvA NuMI Off-axis ν_e Appearance (experiment). [1–5](#), [7](#), [8](#), [14](#), [33](#), [50](#)

NuMI Neutrinos from the Main Injector. [3](#)

PE Photo Electron. [50](#)

PECORR Corrected Photo Electrons. [50](#)

ToF Time of Flight. [3](#)

WLS Wavelength Shifting (fibre). [26](#), [27](#), [33](#)

Bibliography

- [1] Michael Wallbank. The NOvA Test Beam Program. ICHEP2020:188, . doi:[10.22323/1.390.0188](https://doi.org/10.22323/1.390.0188).
- [2] Alex Sousa. Test Beam Plenary Update - FNAL September 2018. NOVA Document 33012, . URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=33012>. NOvA internal document.
- [3] Alex Sousa, Ryan Nichol, Karol Lang, and Jeff Nelson. NOvA Test Beam Task Force Report. NOVA Document 15750. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=15750>. NOvA technical note.
- [4] Alex Sousa. NOvA Test Beam Status and Plans - Support Documentation. NOVA Document 22172-v2, . URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=22172>. NOvA technical note.
- [5] Fermilab test beam facility official website. URL <https://ftbf.fnal.gov/>. Cited March 2024.
- [6] Alex Sousa. NOvA Test Beam Plenary @ IU Collaboration Meeting. NOVA Document 29543, . URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=29543>. NOvA internal document.
- [7] Michael Wallbank. Final Test Beam Updates (Geometry and Other!). NOVA Document 58388, . URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=58388>. NOvA internal document.
- [8] Michael Wallbank. Understanding, Improving, Validating the Test Beam Geometry. NOVA Document 57955, February 2023. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=57955>. NOvA internal document.
- [9] Robert Kralik. Test beam calibration update. NOVA Document 57516-v2. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=57516>. NOvA internal document.

- [10] Alex Sousa. Test Beam Plenary Update - Jun. 6, 2019. NOVA Document 38349, .
 - . URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=38349>. NOvA internal document.

- [11] Miniboone experiments official website. URL <https://www-boone.fnal.gov/>.
 - Cited March 2024.

- [12] Alex Sousa. Filling System and Scintillator Status. NOVA Document 34067, .
 - . URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=34067>. NOvA internal document.

- [13] Junting Huang, Will Flanagan, and Beatriz Tapia Oregui. Test Beam: Light Yield of the Liquid Scintillator Drained from the NDOS Detector. NOVA Document 38740. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=38740>. NOvA internal document.

- [14] Dung Phan. Test Beam: Tintometer Measurement of Texas A&M oil. NOVA Document 39088. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=39088>. NOvA internal document.

- [15] Alex Sousa. Test Beam Scintillator Fill Plan. NOVA Document 34196, .
 - . URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=34196>. NOvA internal document.

- [16] Alex Sousa. 2nd Block Filling Status - Nov. 18, 2019. NOVA Document 41961, .
 - . URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=41961>. NOvA internal document.

- [17] Teresa Megan Lackey. *Proton Scattering in NOvA Test Beam*. PhD thesis.

- [18] David Northacker, Alex Sousa, and Yagmur Torun. Test Beam - Overfilling Horizontal Planes. NOVA Document 49439. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=49439>. NOvA internal document.

- [19] Mark Messier and Teresa Lackey. Data driven cosmic generation. NOVA Document 51327-v3. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=51327&version=3>. NOvA internal document.

- [20] R. L. Workman and Others. Review of Particle Physics. 2022:083C01.
doi:[10.1093/ptep/ptac097](https://doi.org/10.1093/ptep/ptac097).
- [21] Anna Maureen Hall. TB P2 Calib. Summary. NOVA Document 49674. URL
<https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=49674>. NOnA internal document.
- [22] Matthew Strait. Update on light level tuning. NOVA Document 43249. URL
<https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=43249>. NOnA internal document.
- [23] Artur Sztuc. DCM/FEB Shut-Off Studies. NOVA Document 53658. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=53658>. NOnA internal document.