

# CHAPTER 1

## NOvA Test Beam Detector Calibration

The NuMI Off-axis  $\nu_e$  Appearance (NOvA) Test Beam experiment [1] aims to improve NOvA’s sensitivity to the neutrino oscillation parameters by improving our understanding of particle interactions and energy deposition in the NOvA detectors, with the hope of reducing the total systematic uncertainty on the three flavour oscillation parameters by about 10% [2].

Specifically, using a test beam allows us to study the response of tagged single particles as a function of their measured energies and compare it to the simulated prediction. It also enables us to determine the energy resolution and the absolute energy scale of these particles. Additionally, we are able to compare the response of beam and cosmic ray muons, to study fibre attenuation, or to validate the entire NOvA calibration process. The Test Beam detector was also equipped with a combination of Near Detector (ND) and Far Detector (FD) readout electronics and filled with a variety of NOvA scintillator oils, which makes it possible to make a comparison of their responses [3].

All the aforementioned benefits of running the NOvA Test Beam experiment require, or benefit from, the Test Beam detector calibration, which follows the same calibration procedure as the standard NOvA detectors (outlined in Sec. ??). This allows for a direct comparison of the deposited energy in the Test Beam detector with the standard NOvA detectors. On top of that, the unique qualities of Test Beam allow us to use the Test Beam calibration to validate the calibration process and possibly to provide a simulation-independent absolute energy scale.

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

as part of my PhD and Sec. 1 discusses the Test Beam detector calibration itself. Here we present the differences from the calibration of the **NOvA ND** and **FD**, as well as the results and their discussion and validation.

## 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 on 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].

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.

Table 1.1: Test Beam detector data taking periods.

Period 1	June 3 <sup>rd</sup> 2019	-	July 6 <sup>th</sup> 2019
Period 2	December 5 <sup>th</sup> 2019	-	March 20 <sup>th</sup> 2020
Period 3	January 12 <sup>th</sup> 2021	-	June 27 <sup>th</sup> 2021
Period 4	November 30 <sup>th</sup> 2021	-	July 10 <sup>th</sup> 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 on Fig. 1.1, we use two collimators to direct the tertiary beam and a magnet to select the desired momentum. Particle tracking is done using the four Multiwire

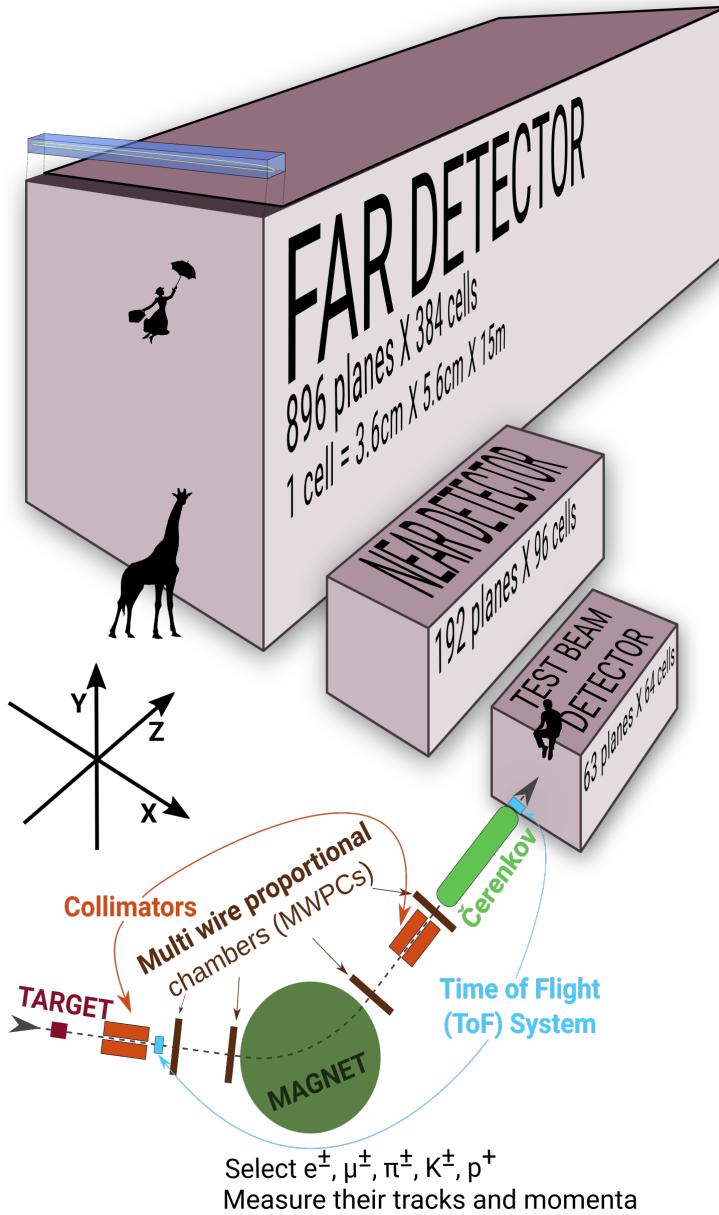


Figure 1.1: Comparison of Test Beam detector scale to the NOvA ND and FD (and a man, giraffe, or Mary Poppins). Also shown are the Test Beam beamline detectors and components (not to scale), with arrows showing the direction of the beam. The three black arrows show the orientation of the detector coordinate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

Proportional Chamber (MWPC)s and particle identification is done with a combination of the Time of Flight (ToF) detectors and the Cherenkov detector, which is set for electron detection. There was an additional ToF detector installed ... during period ...

**COMMENT:** *Where do I find the information on this?*

## 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 on 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. *COMMENT: Should I keep this paragraph in or no?*

## Scintillator

Test Beam used a combination of the leftover **ND** and **FD** production scintillator oils and the oil drained from the 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 on Fig. 1.2.

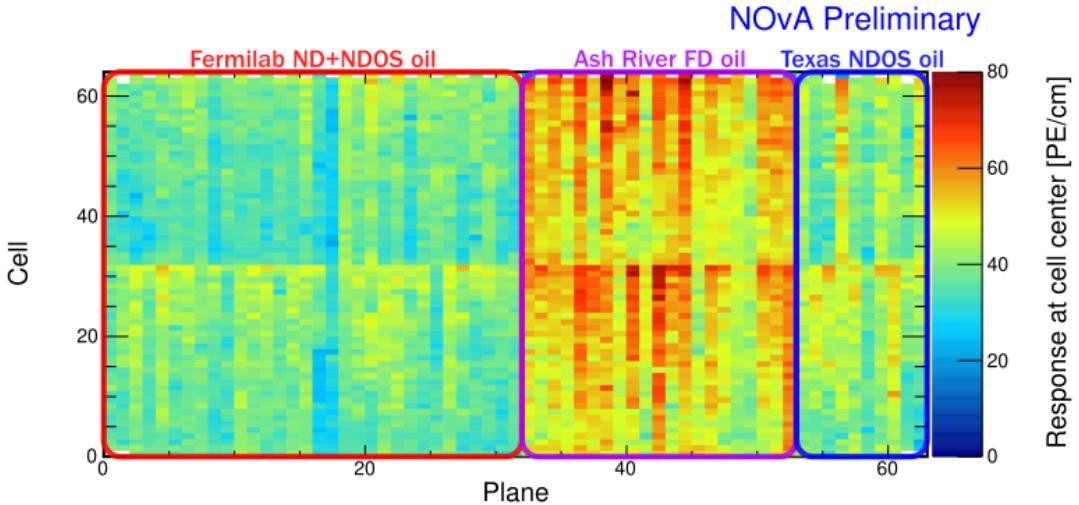


Figure 1.2: Uncorrected energy response in the centre of cells across the Test Beam detector showing a clear distinction between the different scintillator oils, labelled with coloured boxes and descriptions.

We can distinguish four 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 tanks outside in Fermi National Accelerator Laboratory (FERMILAB) [10];
2. Separate ND production oil and NDOS-drained oil stored underground in barrels<sup>1</sup> at the MiniBooNE<sup>1</sup> 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 we loaded oil from tank #2 into the tanker, it became extremely cloudy and unusable, possibly due to contamination with water accumulated at the bottom of the tanks.

---

<sup>1</sup>MiniBooNE [11] is a FERMILAB experiment located close to the NOvA ND

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" on 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]. We again topped up these planes with the **ND+NDOS** scintillator (sample #2).

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 Board (FEB)s, 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 in a fester rate and we used a mix of **FEB** types to study the difference in their response and to validate both versions in the same environment [17].

## Environment

Unlike the **ND** and the **FD**, the Test Beam detector does not have any overburden to shield it from cosmic particles, which affects their rate and energies inside the detector. There is also less precise control of temperature and humidity than in the other detectors [source?], which can potentially impact the scintillator and readout performance. *TO DO: Finish this description of environmental control, maybe mention HVAC? - depends how much I want to talk about this effect...*

## Underfilled cells issue

The Test Beam detector is slightly tilted around the z axis by about  $0.7^\circ$  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 has been 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, the **CRY** simulation proved to be highly inefficient, with only a small fraction of the simulated cosmic-ray activity resulting in selected calibration hits, and the majority of particles failing to even hit the detector. This inefficiency consumed significant CPU resources, disk space, and file usage. Moreover, the momentum and angle distributions in **CRY** were not well suited to the **NOvA** sites, potentially impacting the calibration accuracy.

To overcome these challenges, we have implemented a data-based simulation method that eliminates the need for the **CRY MC** generator. Instead, we use a subset of the cosmic data sample used in calibration and pass it through the beam removal filter, reconstruction chain and a selection of high-quality cosmic muons. The selected cosmic muon events are then being used in a custom simulation to create a new simulated cosmic ray sample.

This approach results in 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 data. Given that the calibration chain itself is a time and CPU intensive process, the reduction in simulation files and their sizes has significant benefits downstream of the file generation. On the other hand, using real data to seed the simulation could introduce undesirable bias, if we do not carefully consider possible misreconstruction or selection bias.**COMMENT: Can I re-word the last sentence?**

### 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 used half of period 4 data sample (see Tab. 1.1), as the other periods had complications such as faulty [FEBs](#), or underfilled cells. We sampled through the entire period 4 by using every other sub-run for the simulation.

We design the reconstruction and selection criteria so that the majority of the simulated cosmic muons will make it into the final simulation calibration sample. Therefore, we employ 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 the event 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 on Fig. 1.3. This should leave us with mostly cosmic events.

#### Reconstruction

The [MC](#) generator requires the vertex position and the initial 4-momentum for each simulated event. For that 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) algorithm to identify muons and assign their momenta. [BPF](#) required a vertex and prong input information, which we get from a cosmic ray vertex and FuzzyK algorithms respectively. The first three steps are identical to the full reconstruction applied to get 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.

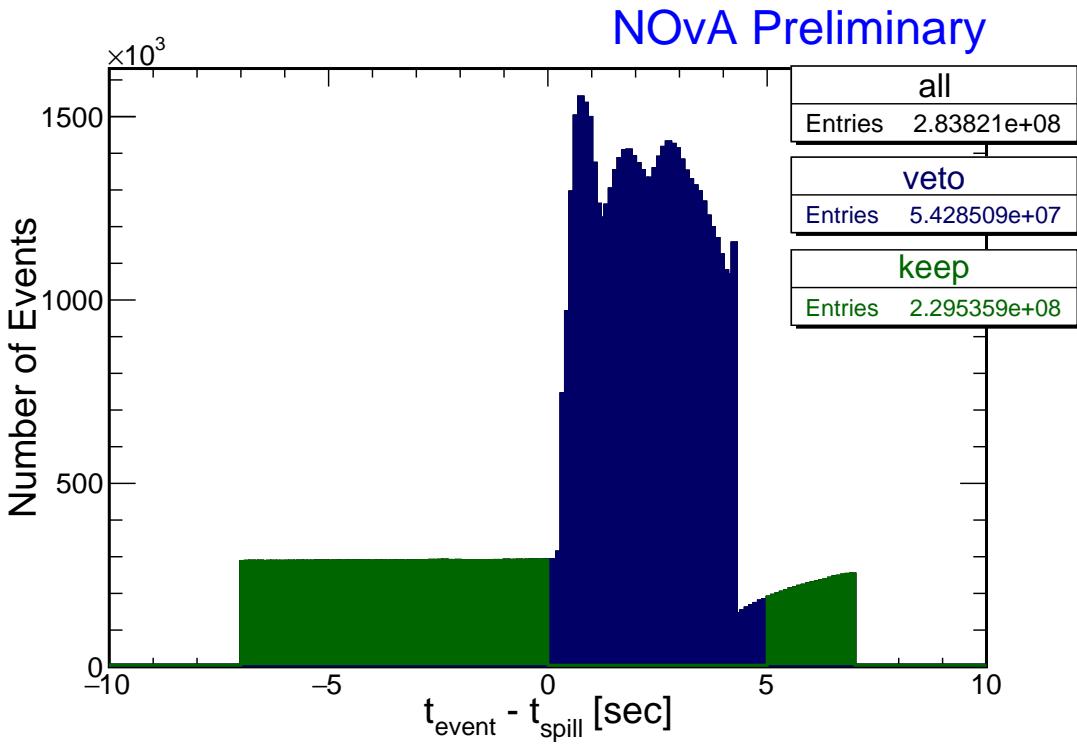


Figure 1.3: Test Beam beam spill events removed (blue) from the calibration samples. The remaining events (green) should mostly consist of cosmic particles. This example and the numbers of entries are for the full period 4 Test Beam sample.

## Selection

After the reconstruction process, we proceed to select events based on their slice and **BreakPointFitter (BPF) track** properties.

To select an event, the following conditions must be met (Tab. 1.2 provides an overview of all cuts and their corresponding values):

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 from the Top/Front/Back/Sides 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 (beam) axis to be  $|\cos_z| \leq 0.98$ . Figure 1.4 demonstrates the presence of events peaked at track lengths of approximately 410 cm and

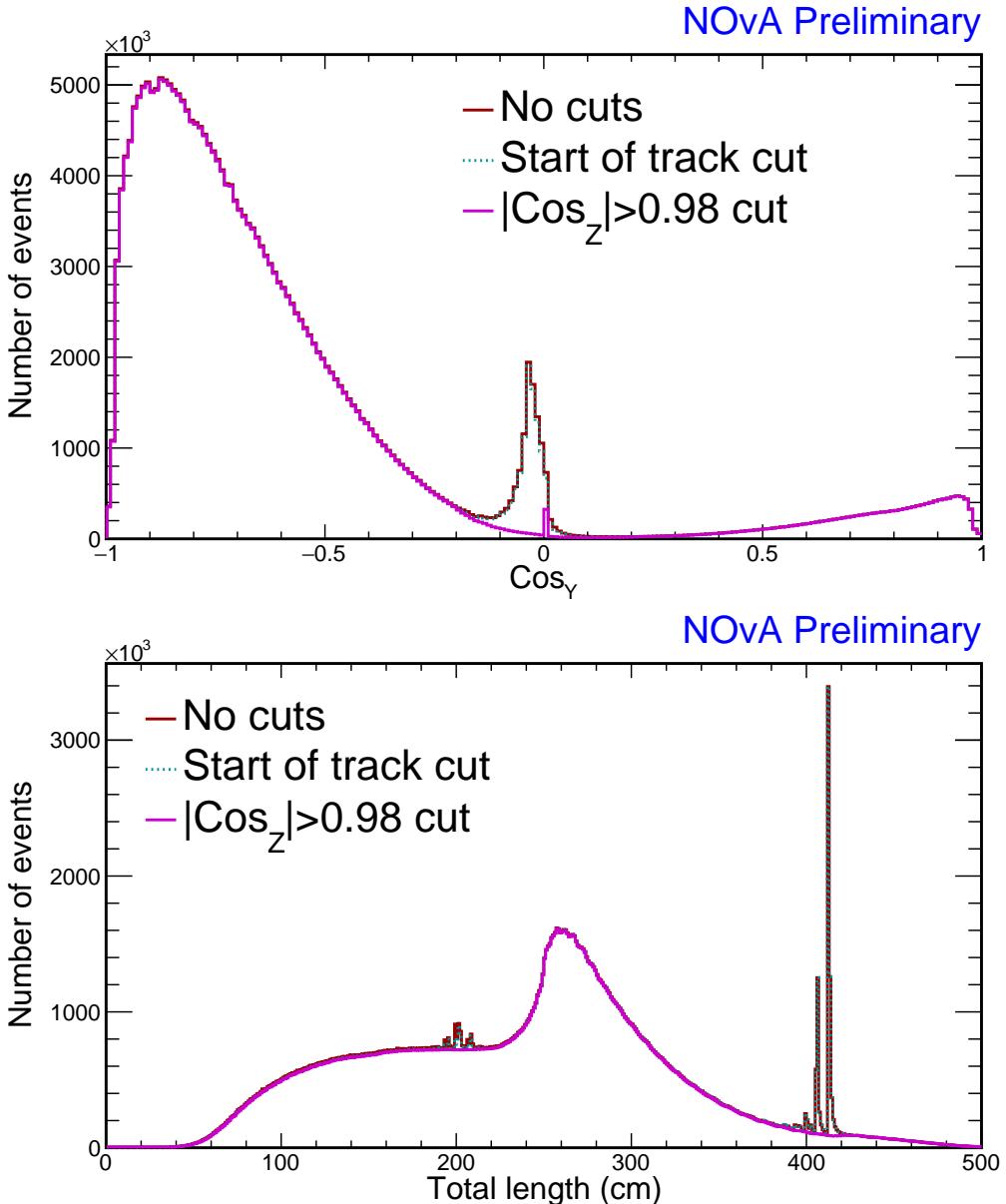


Figure 1.4: Impact of track start and maximum track angle from the z axis ( $\text{Cos}_Z$ ) cuts on the Test Beam data for the data-based simulation of cosmic muons. The track start cut has only negligible effect. The maximum  $\text{Cos}_Z$  cut effectively removes sharp peaks in the total track length distribution and events perpendicular to the Y axis. These events are all parallel with the Z axes and are most likely leftover beam events. All of the distributions are made from the period 4 Test Beam data.

200 cm, which correspond to the total and half length of the detector, respectively (or alternatively lengths of both modules and a single module). These events are strictly parallel to the beam direction and are likely remnants of beam events. Applying a cut on  $\text{Cos}_Z$  effectively removes these events without affecting the rest of the data. This cut might only be needed for Test Beam and not for the near and far detectors.

4. To ensure that only events contributing to the final calibration sample are simulated, we use a selection based on cuts from the Calibration/PCHitsList module. This module is used for both data and simulation to create the calibration samples. Let's call these cuts the **calibration cuts**. However, there are two caveats we need to consider:
  - (a) First, to create calibration samples, we apply selection to tracks from the **Window cosmic track** algorithm instead of the Break Point Fitter algorithm, which yield different distributions as depicted on 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  $\text{Cos}_Z$ , which is not present for Window cosmic tracks. This is likely caused by the detector structure, as shown on 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 have loosened the full calibration cuts to create a "buffer" around the selected events, allowing for fluctuations of the reconstruction algorithms while maintaining track quality. 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 on 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<sup>2</sup>. This

---

<sup>2</sup>For Test Beam we say it is a stopping muon if its track ends at least 20 cm from any edge of the

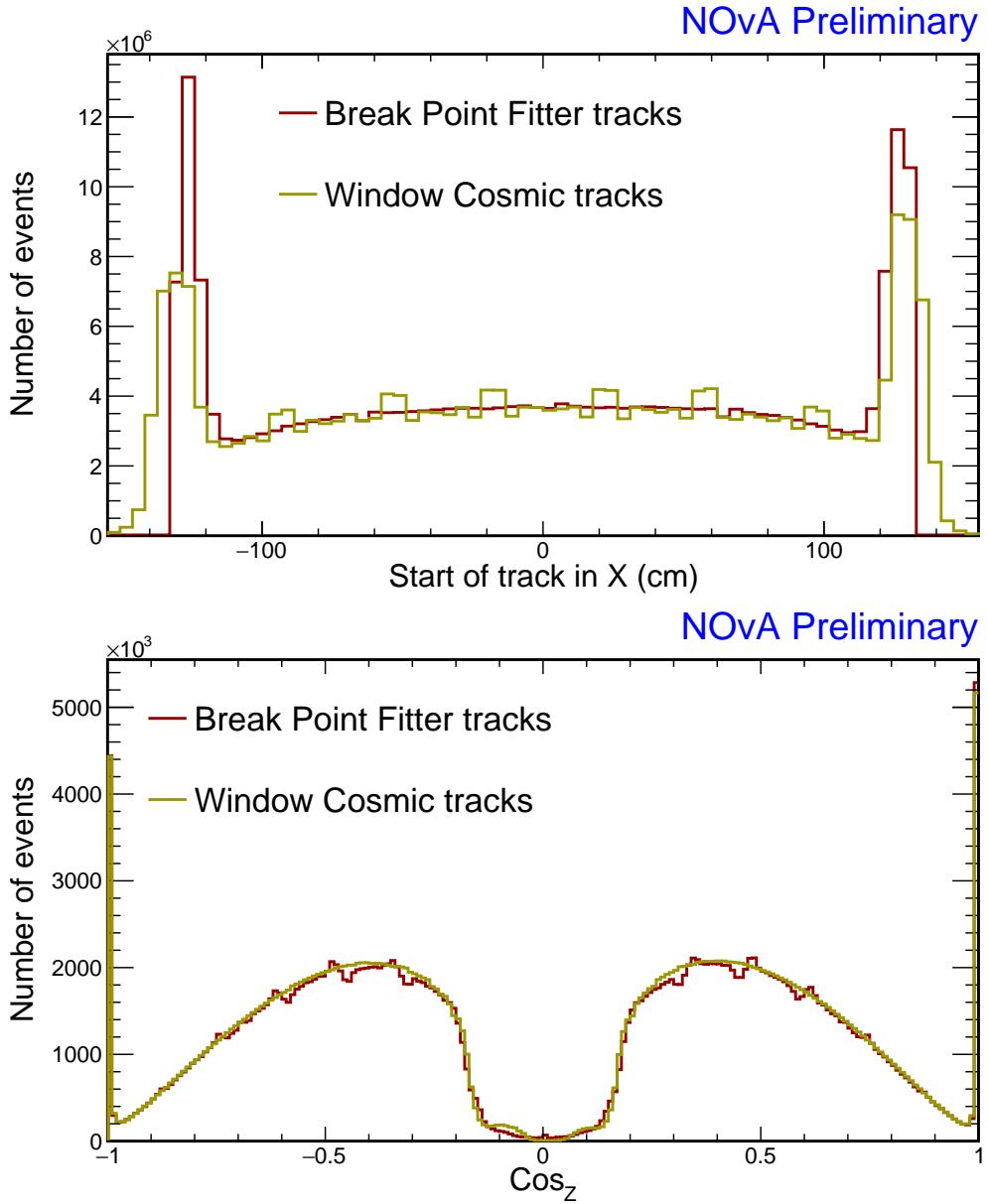


Figure 1.5: Difference between the tracks reconstructed with the break point fitter and with the Window cosmic track algorithms. Both distributions are for the period 4 Test Beam data (with removed beam spill) without applying any selection.

information assists in correcting the energy of through-going muons, as outlined in Sec. 1.2.2.

### 1.2.2 Energy correction, charge assignment and smearing

Once we have the kinematic information for the selected events performs several tasks, including correcting energies of the through-going muons, assigning a charge

---

detector. For the far and near detector this is 50 cm. This value was chosen by Kevin Mulder [?] as 50 cm removed too many cosmic events from the Test Beam detector.

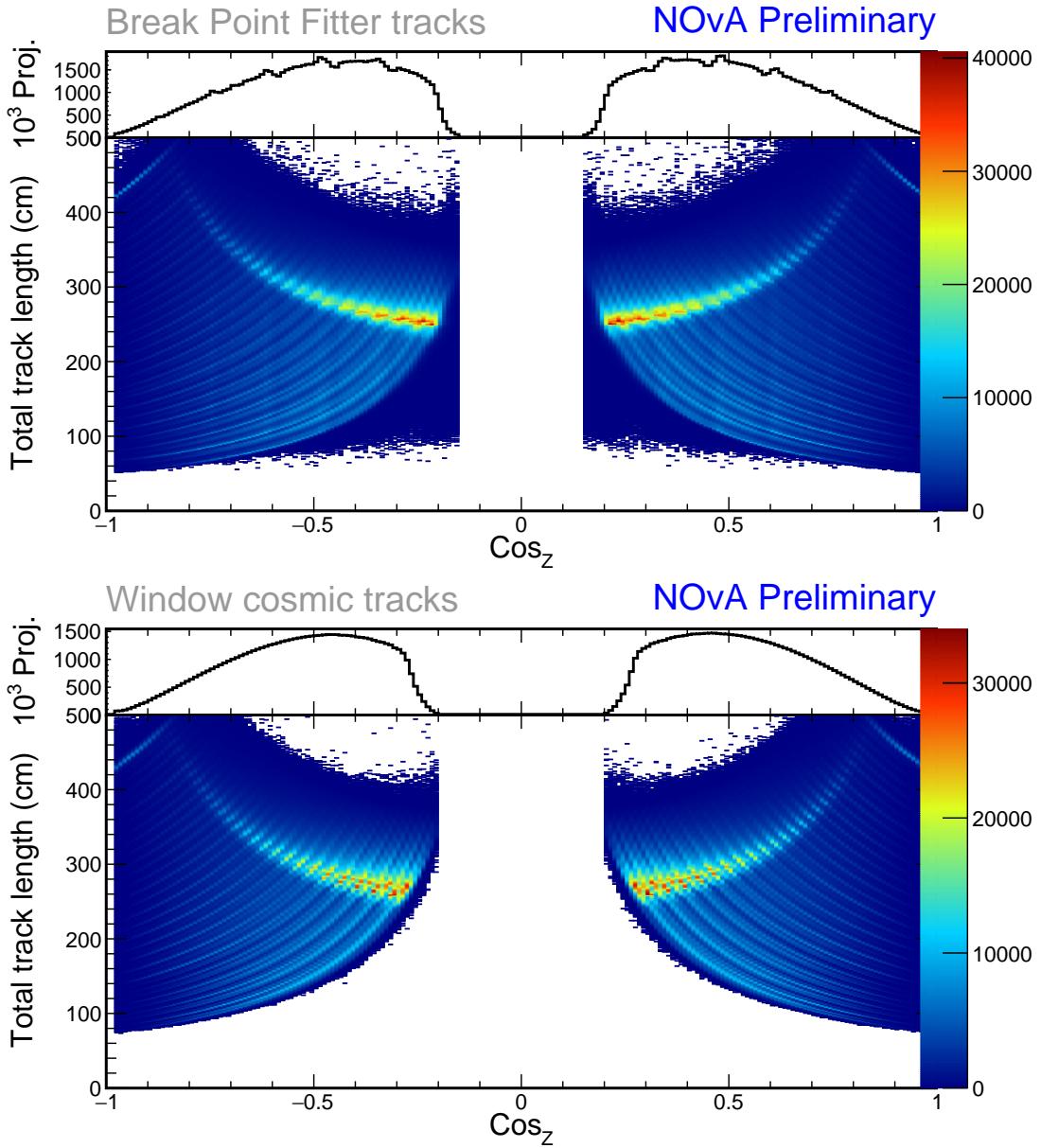


Figure 1.6: Investigating the origin of the rugged shape in the  $\text{Cos}_Z$  distribution of Break Point Fitter tracks. The top plot is created with the loose calibration cuts and the bottom plot with full calibration cuts. This difference in selection shouldn't matter. The long lines on the 2D plots are likely the effects of the detector structure. We can see that for the Break Point Fitter tracks, each  $\text{Cos}_Z$  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  $\text{Cos}_Z$  distribution of Break Point Fitter tracks.

to each muon, and smearing and converting the information into the correct format required by the generator.

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

Cut	Selection	
	Full	Loose
Calibration sample selection	Muon assumption and 3D track from BPF	
	Max. track start distance from edge	50 cm
	Max. $Cos_Z$	0.98
	Max. number of hits in X or Y	2
	Min. difference between $Stop_Z$ and $Start_Z$	70 cm
	Min. $Cos_Z$	0.15
	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

## Energy correction

Through-going muons do not deposit all of their energy inside the detector. From data we cannot reliably calculate their initial energies, but we can estimate an energy that would leave the same track. In general, the energy spectrum of cosmic muons can be approximately described by a power law  $E^{-\alpha}$ , with  $\alpha \approx 2.7$  [? ? ]. 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 from the Break Point Fitter.  $E_C$  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), as described in Sec. 1.2.1. Plot 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

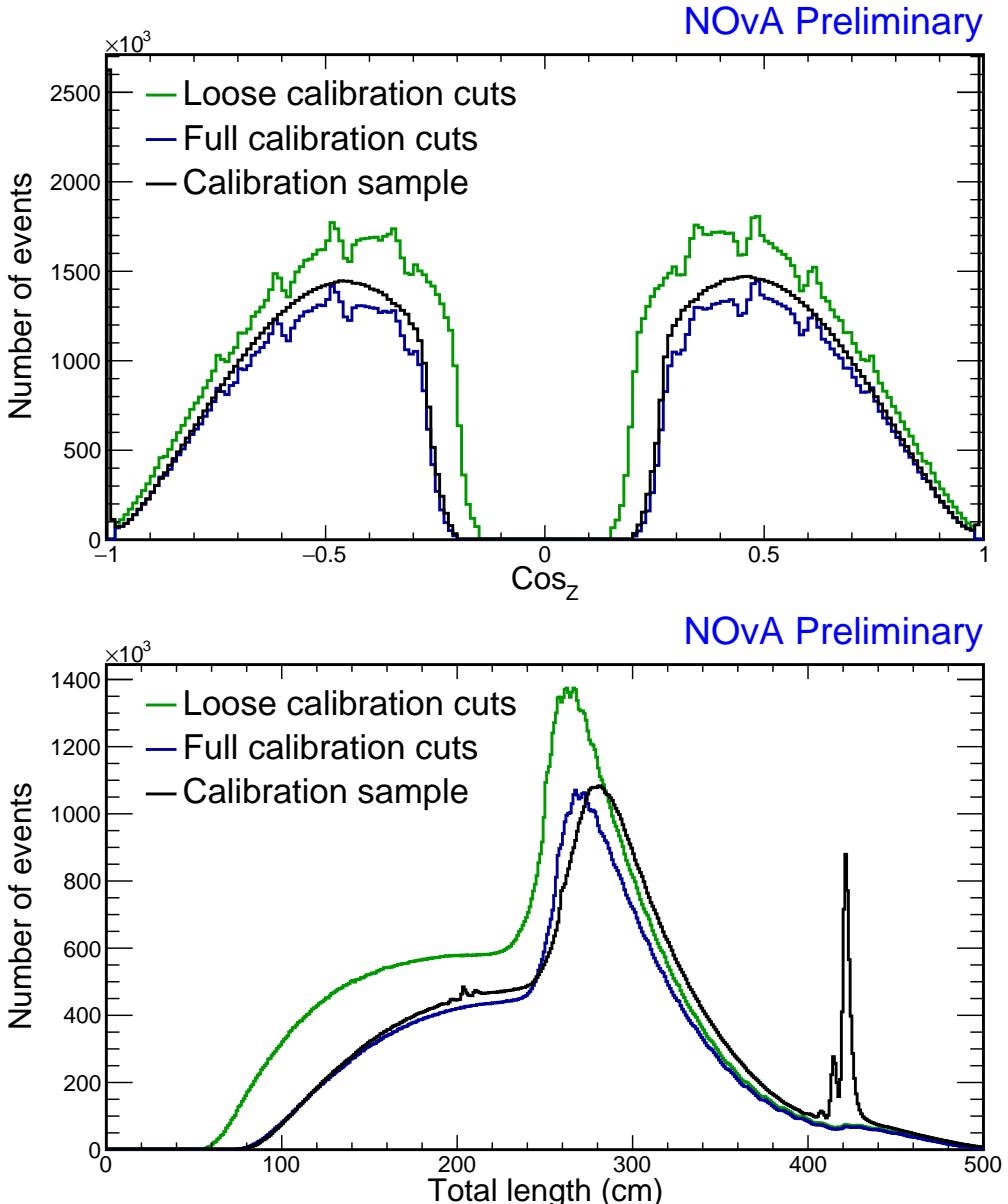


Figure 1.7: Comparison of event selections for the data-based simulation and of the corresponding data calibration sample in black. The green line represents the final selection used for the simulation, using the loosened calibration cuts, while the blue line shows the distributions with full calibration cuts. The data calibration sample was made with the same full calibration cuts as the blue line (without the track start cut and maximum  $\text{Cos}_Z$  cuts), but applied to the window cosmic tracks instead of the Break Point Fitter tracks. All of the distributions are made from the period 4 Test Beam 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 entering the detector would be larger for the bigger near detector than for Test Beam,

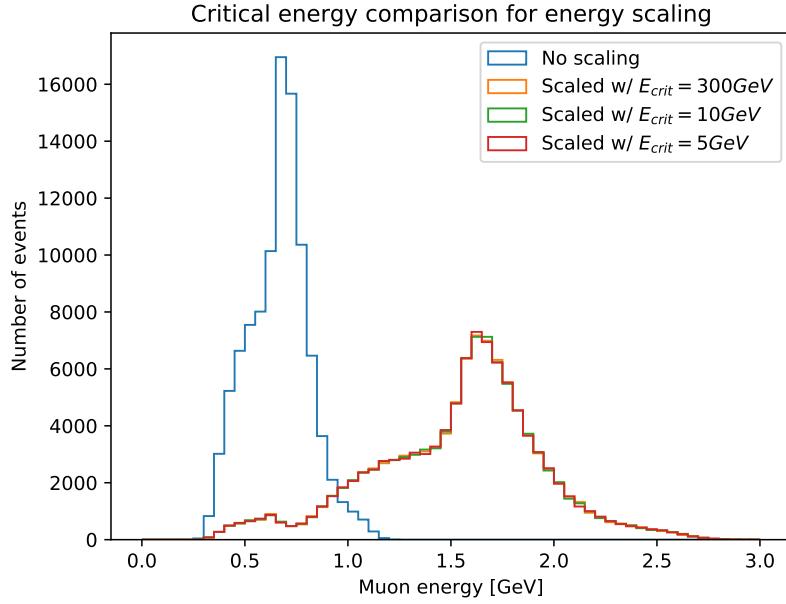


Figure 1.8: The effect of energy correction for through-going muons with various critical energies. No significant difference can be seen when using different critical energies.

even though the near detector 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 representation of cosmic muon energy spectrum. Not including more energetic cosmic muons into the simulation does bias the energy deposition towards lower average, but this is corrected for during absolute calibration which only uses stopping muons. See Test Beam calibration technical note for further information [? ].

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 [? ].

## Smearing

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

- total momentum within 2%,
- azimuthal angle uniformly,
- 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 MC generator 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 [? ]:

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

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

### Submitting the simulation jobs

The brightness file describes the relative differences in energy response across the different cells and planes. These differences mainly arise from the variability of each fibre's brightness, and specifically for Test Beam, also from the different scintillators used. 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.

In the first iteration of the data-based simulation, Teresa used a test beam brightness file created from period 2 data, which contains faulty FEBs and underfilled cells, resulting in a simulation also containing these defects. In the second iteration Robert created a new brightness file from period 4 test beam data, which are free from any irregularities and supplied that to the simulation.

### 1.2.3 Validation

To validate whether our simulation is working as expected, we

1. compare the new simulation with the original data and
2. use the new simulation (specifically the artdaq-stage sample) as "fake data" and feed it into the exact same simulation process to create a "re-simulation" sample.

For the data-simulation comparisons we use the events from the calibration (plist)<sup>1</sup> samples, which are equivalent to the Window cosmic tracks with full calibration cuts from the CosmicGenAna module, as 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 distributions of Break Point Fitter 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 Break Point Fitter tracks and the Window cosmic tracks. This can also be seen on the total track length distribution on Fig. 1.10. This is disappointing, but the distributions of the new simulation look reasonable and are close enough to the data, that we've decided to proceed with this simulation and use it in the Test Beam calibration. It is unlikely that we could mitigate these differences by changing the selection even more.

The start of track comparison between data and simulation on Fig. 1.11 and 1.12 show that there are fewer events that start at the edge of the detector. This is likely the result of the event smearing.

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

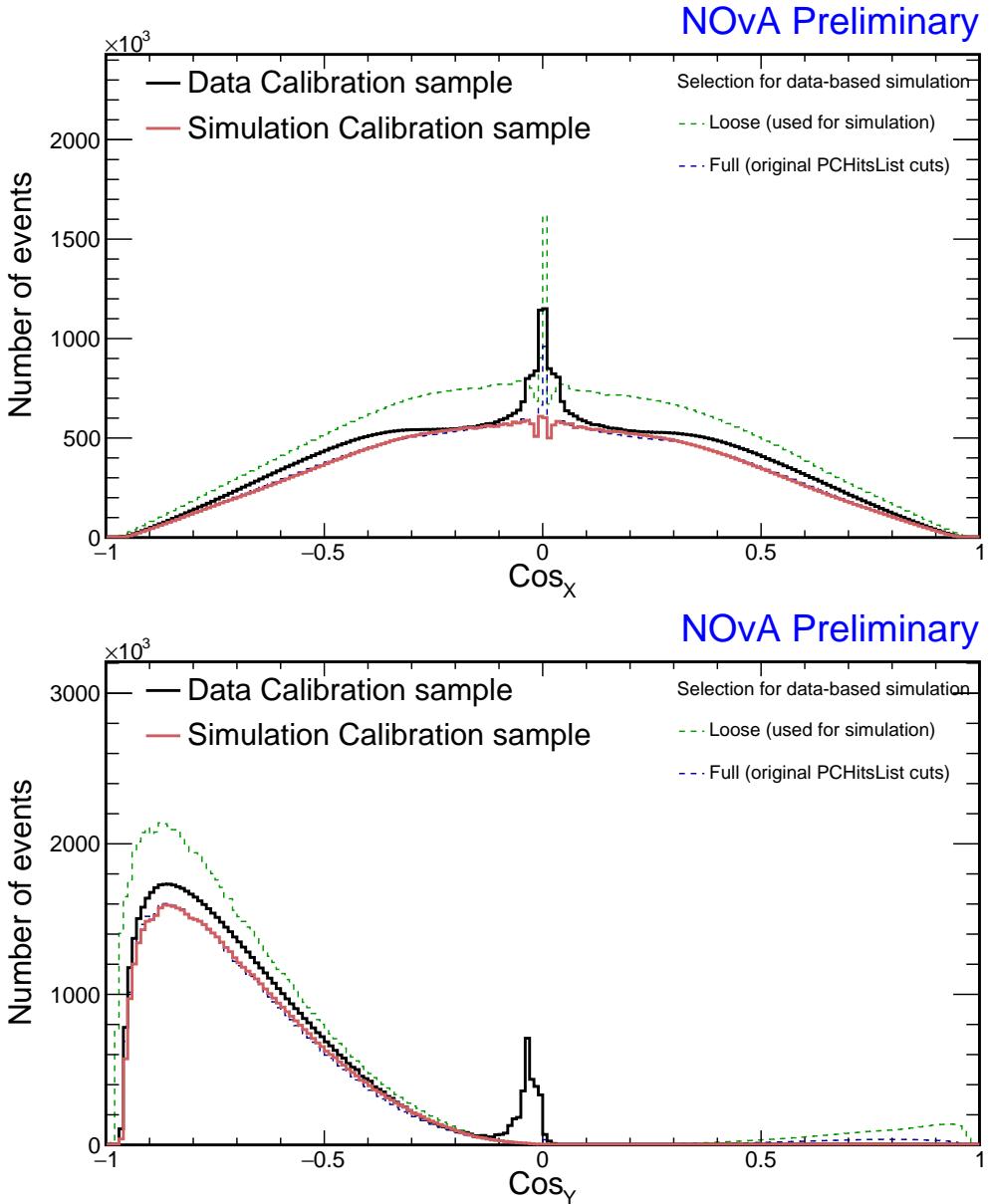


Figure 1.9: Angular distribution comparison of the newly created simulation calibration sample and the corresponding data calibration sample. We are also showing the selection of data used to create the new simulation in green and a "full calibration cuts" selection (same cuts as used for the simulation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

#### 1.2.4 Conclusions

We created a new version of data-based simulation of cosmic muons for the Test Beam calibration. The new simulation is performing better than the old version of the data-based simulation and also than the CRY based simulations used in the earlier stages. The results of the new simulation inside the Test Beam calibration process are described in the Test Beam calibration technical note [? ].

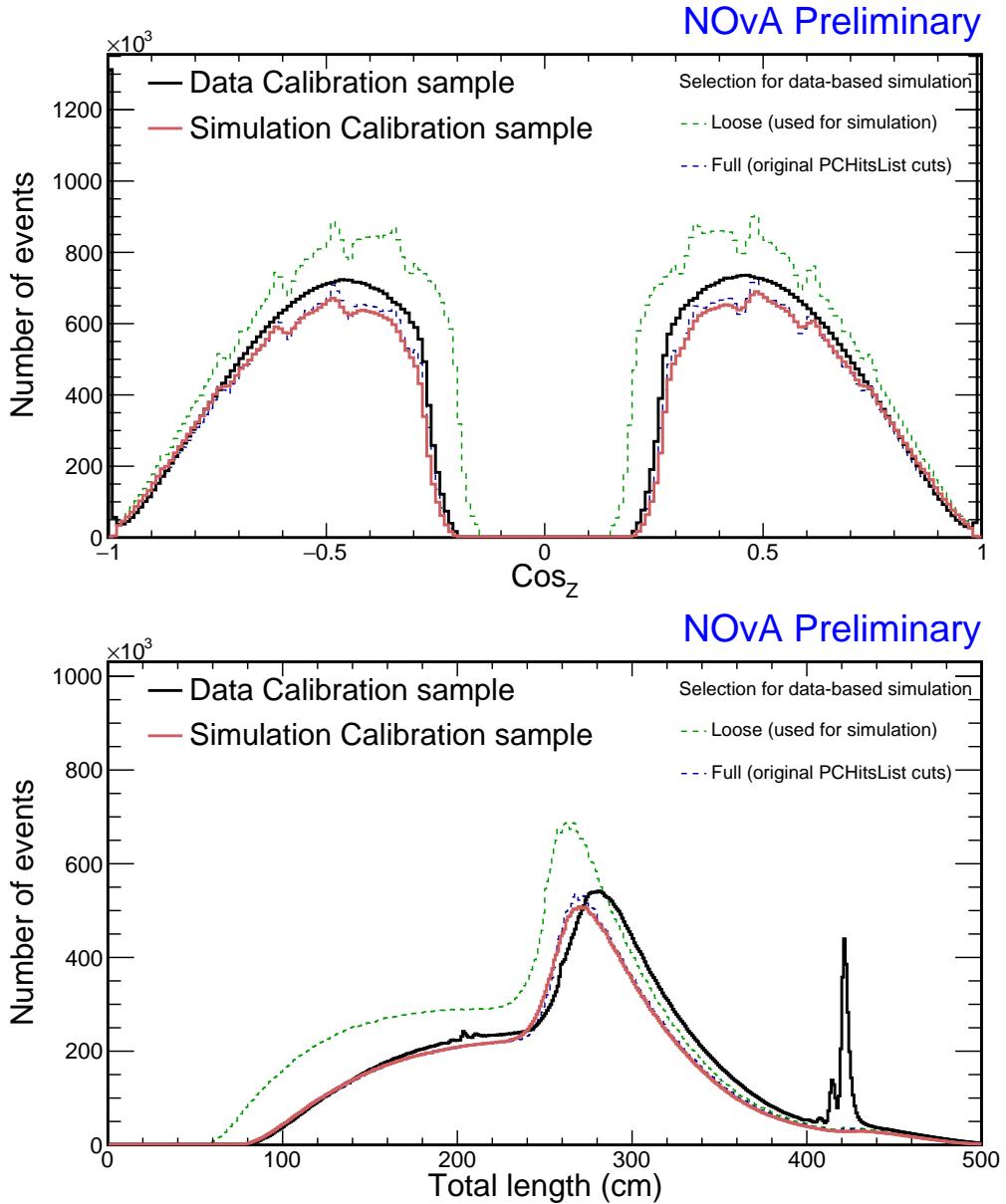


Figure 1.10: Angular and total track length distribution comparison of the newly created simulation calibration sample and the corresponding data calibration sample. We are also showing the selection of data used to create the new simulation in green and a "full calibration cuts" selection (same cuts as used for the simulation calibration sample but applied to Break Point Fitter 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 measured Test Beam data, with the exception of period 1 data. Period 1 only includes one month of data, with half-filled detector and several issues including the beam

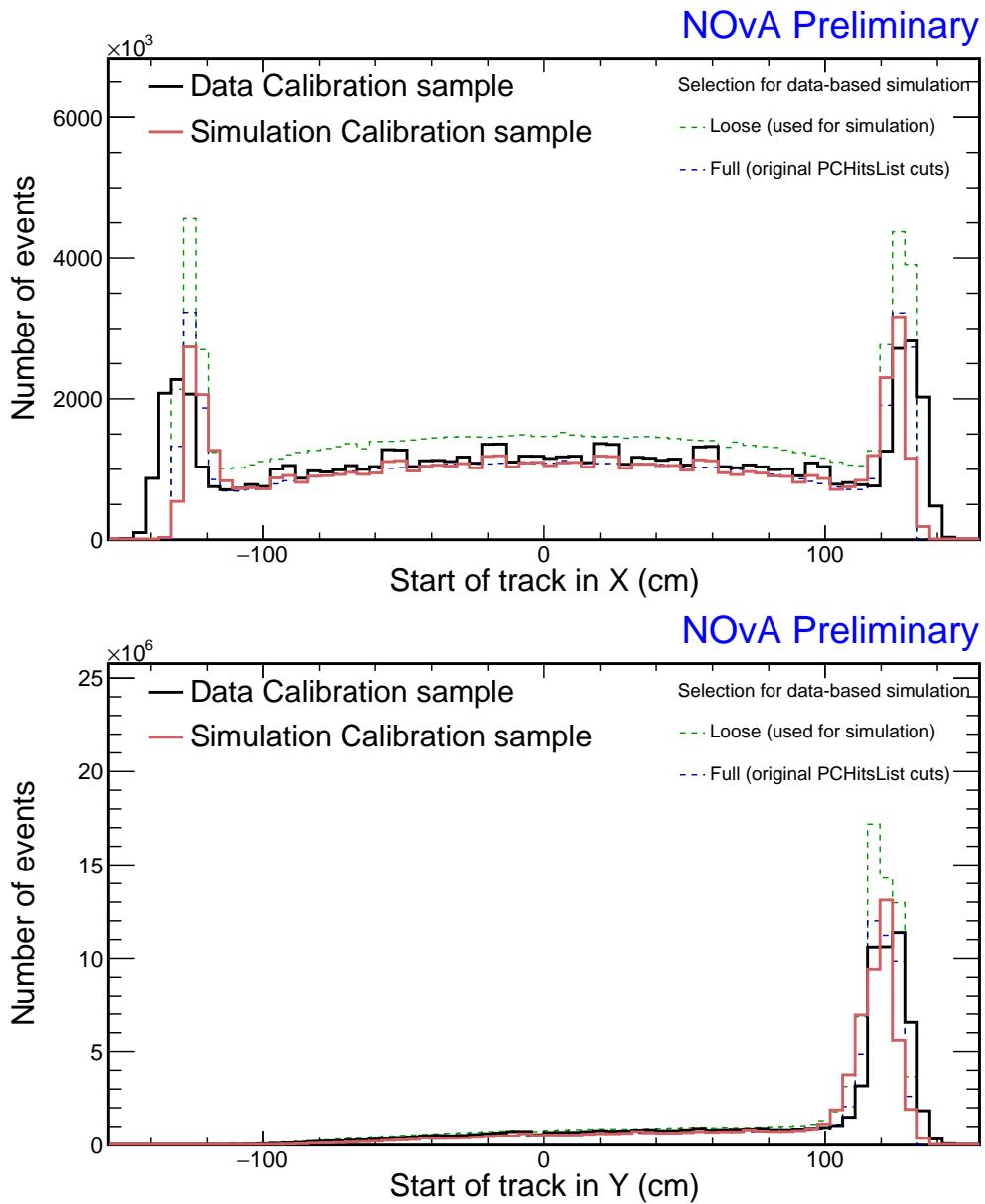


Figure 1.11: Start of tracks comparison of the newly created simulation calibration sample and the corresponding (period 4) data calibration sample. We are also showing the selection of data used to create the new simulation in green and a "full calibration cuts" selection (same cuts as used for the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

halo [? ] and was only used for commissioning and is not used in any Test Beam analysis.

The calibration samples used as inputs for the Test Beam detector calibration are listed in Tab. ???. We are using data from one of the two Test Beam data-driven activity-based triggers (DDActivity1).

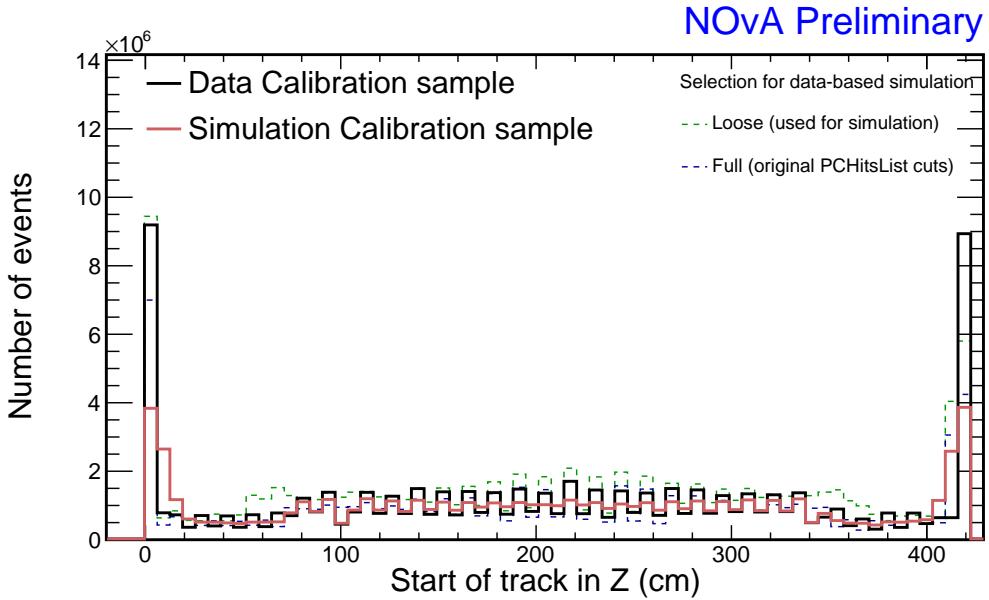


Figure 1.12: Start of tracks comparison of the newly created simulation calibration sample and the corresponding (period 4) data calibration sample. We are also showing the selection of data used to create the new simulation in green and a "full calibration cuts" selection (same cuts as used for the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

### 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 file 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 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, 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 on 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 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

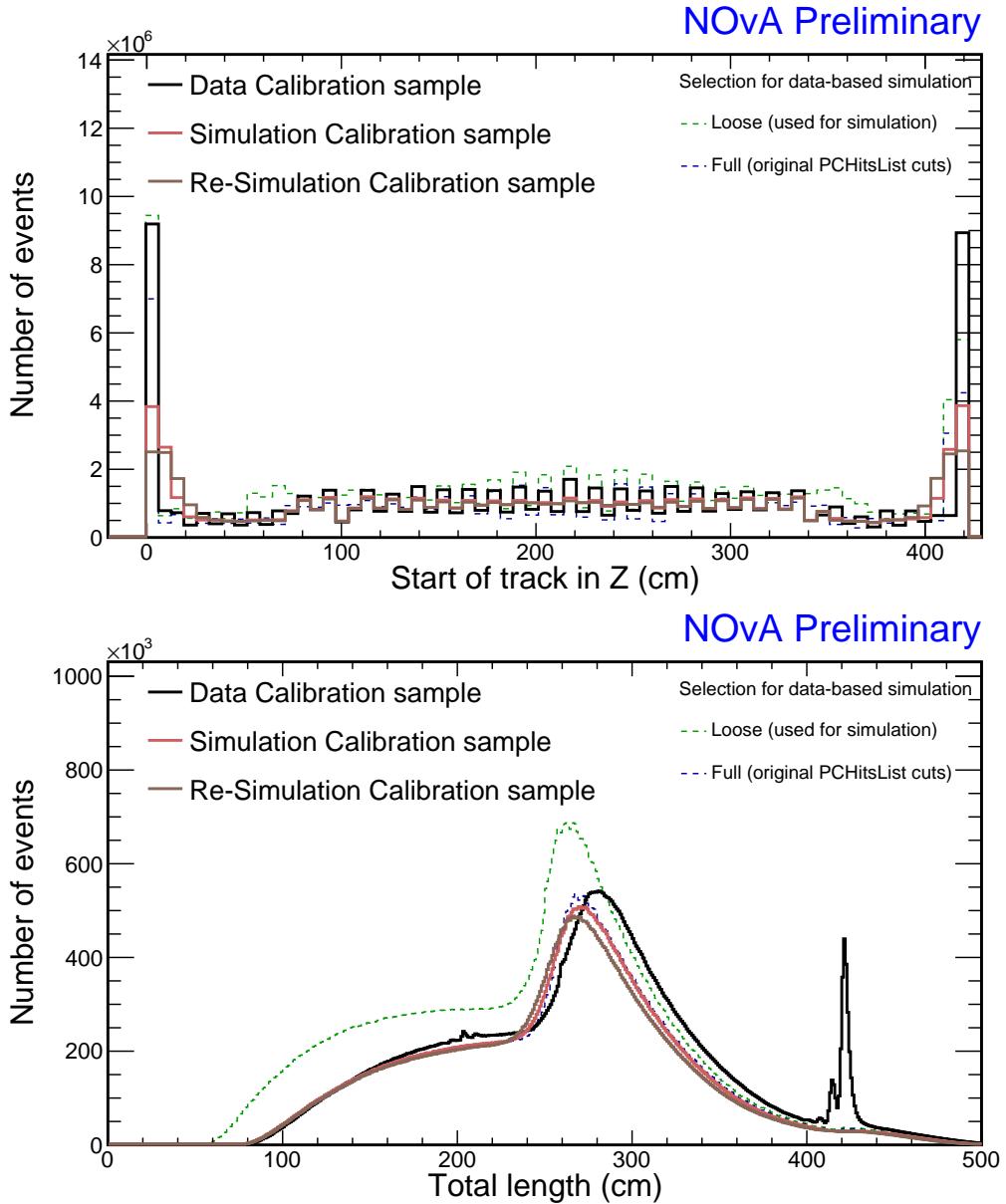


Figure 1.13: Distribution of the re-simulated events, when the new simulation’s artdaq sample is used as “fake data” for a new iteration of the simulation process discussed in this document. This is compared to the newly created simulation calibration sample and the corresponding data calibration sample. 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 Break Point Fitter tracks instead of Window cosmic tracks) in blue.

to create the brightness file and keep the original condition for the actual calibration results.

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

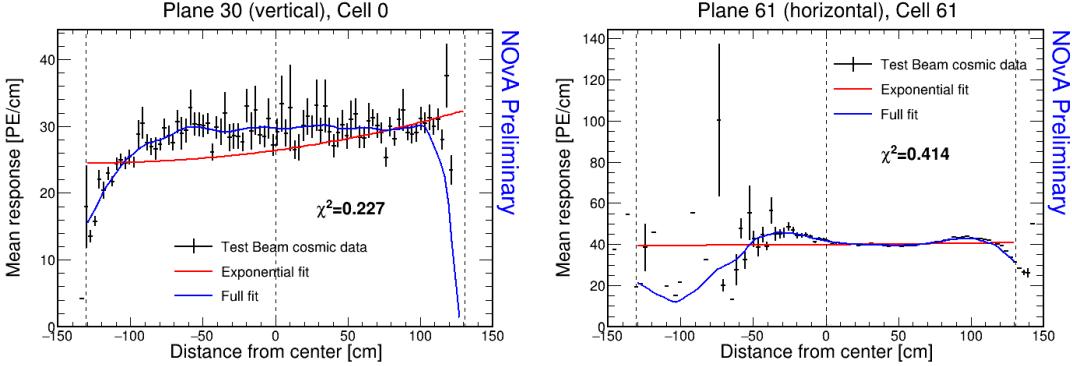


Figure 1.14: Attenuation fits for two cells that fail the calibration condition, but the fit (blue line) correctly represents the energy deposition in the centre of that cell (yellow dashed line).

### 1.3.2 Threshold and shielding corrections

We created the threshold and shielding correction for Test Beam from the new simulation described in the next Sec. 1.3.3. As can be seen on Fig. 1.15, the correction is almost uniform as a function of both  $w$  and cell number. This is the case for all fibre brightness bins and for both views.

The uniform distribution is expected, as the Test Beam detector is much smaller than the far detector and the corrected effects don't have a significant effect. The cell length of 2.6 m has only a negligible effect 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 the relative calibration only cares about relative differences across the detector, a normalization factor does not have any impact on its results.

### 1.3.3 Simulation

We used a custom made data-based simulation of cosmic muons for the Test Beam detector calibration. The details on the simulation and how it was created are described in the "*Data-based simulation of cosmic muons (not only) for calibration technical note*" [? ]. We used half of period 4 data (used every second event as saved in the root file, therefore sampled from the entire period 4) as the inputs. We also used the newly created fibre brightness file (Sec. 1.3.1) to inform the simulation on the realistic detector conditions.

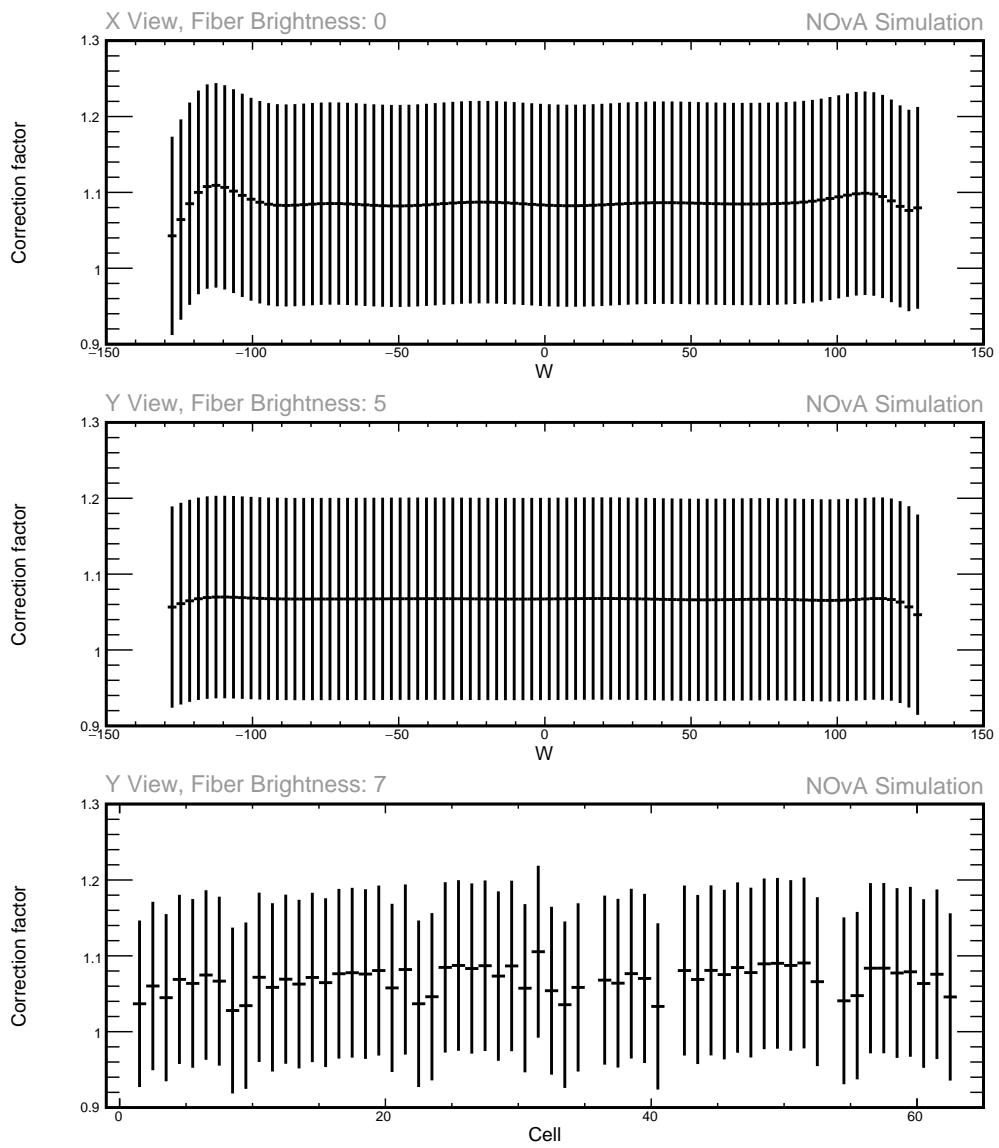


Figure 1.15: Examples of threshold and shielding corrections for the Test Beam detector

The distribution of cosmic muon events from the new simulation across the detector is shown on Fig. 1.16.

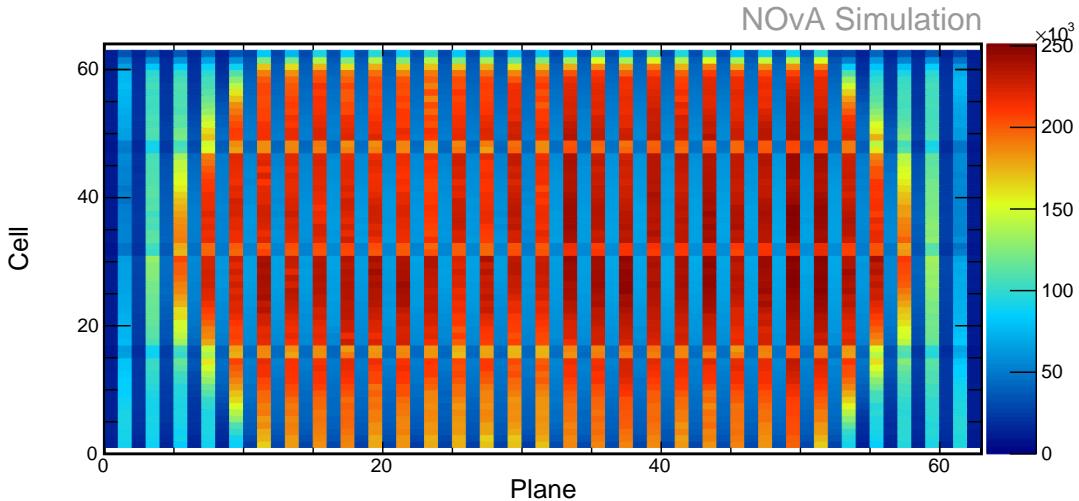


Figure 1.16: Distribution of events in the Test Beam simulation calibration sample.

An overview of the results of the attenuation fit are shown on Fig. 1.17 as a map of the response in the centre of each cell. The blank cells show the uncalibrated cells which failed the calibration condition ( $\chi^2 > 0.2$ ). 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.

Examples of a standard detector response and of the response for cells on the edge of the detector are shown of Fig. 1.18. Here the red line shows the initial exponential fit and the blue line the final attenuation fit, after the LOWESS correction, as described in Sec. ???. Most cells have an expected response with a slow rise towards the readout (right side of the plots), with drops on the edges, as shown on the top left plot.

There is only one cell in the middle of the detector that is left uncalibrated, which is the cell 32 in a vertical plane in the brightness bin 5, shown on the top right of Fig.1.18. The corresponding  $\chi^2 = 0.227$ . It seems the reason the  $\chi^2$  is  $> 0.2$  is the high response with a large uncertainty in the very last fitted bin.

Overall, this is a much better result of the relative calibration (attenuation fit) for a simulation than any of the previous versions of the cosmic muon simulations in the Test Beam detector.

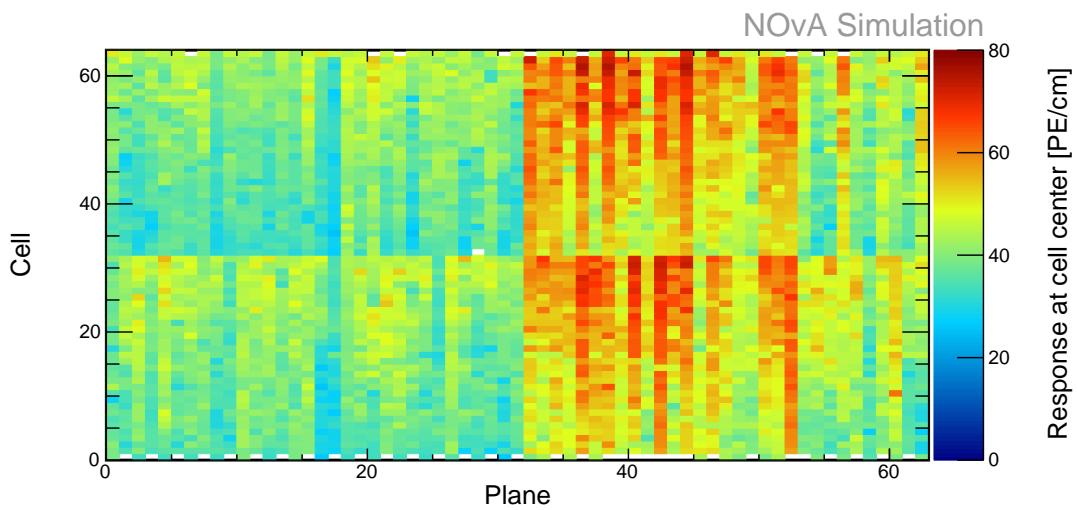


Figure 1.17: Overview of the attenuation fit results for the Test Beam detector calibration simulation. Each cell represents the result of the attenuation fit to the energy response in the centre of that cell. The blank cells are uncalibrated.

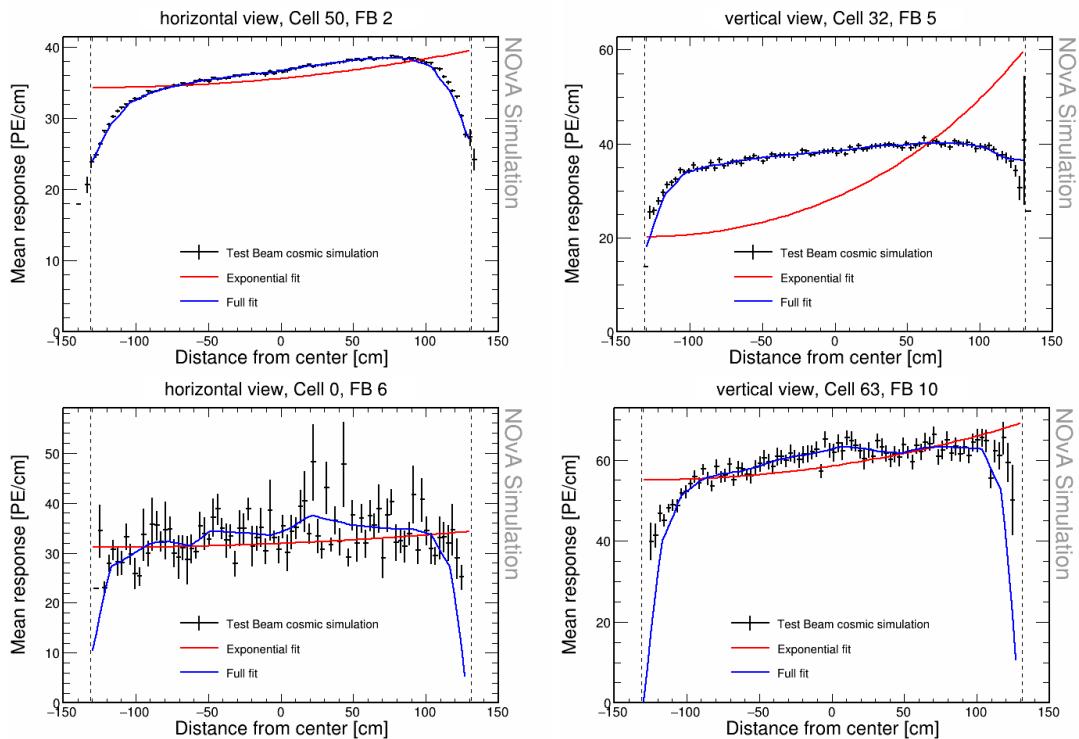


Figure 1.18: Attenuation fits for a selection of cells in the Test Beam calibration simulation. Top left is an example of a successful attenuation fit, top right is a failed fit due to statistical fluctuation in the last bin and the bottom plots show failed fits for cells on the edges of the detector.

### 1.3.4 Period 2 data

The issue with underfilled cells described in Sec. ?? was present throughout the period 2 data taking. This can be clearly seen on Fig. 1.19, 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. ??).

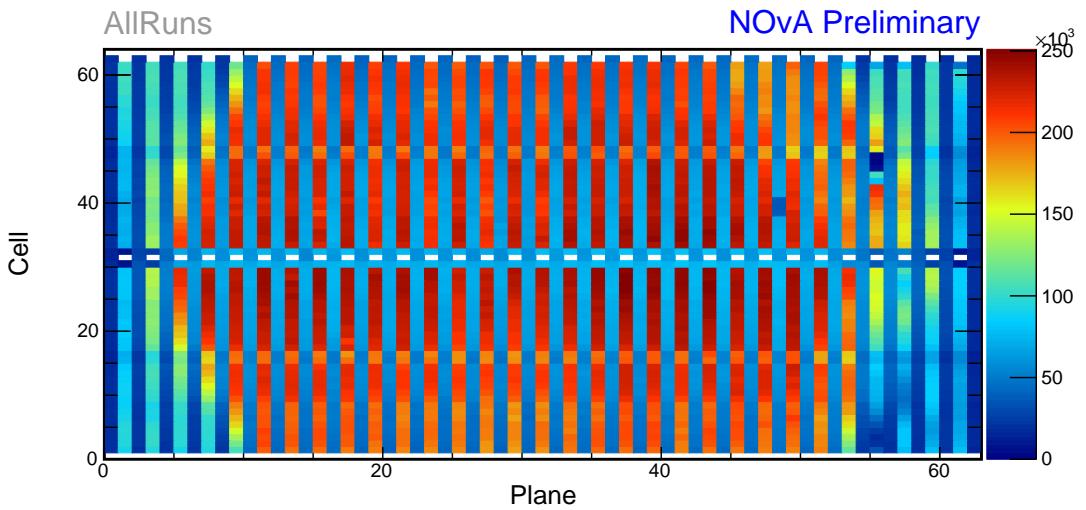


Figure 1.19: Distribution of events in the period 2 Test Beam data calibration sample.

There was also an issue of likely switched cables from the readout in plane 55 between cells 3 and 46 [?], which can also be seen on Fig. 1.19 as dark spots. This is manifested as fewer total number of events in those cells and in their neighbours, again due to the tricell condition.

Officially, period 2 is divided into 6 epochs 2a - 2f, compared on figures 1.20, 1.21 and 1.22. The epochs mostly differ in the use of various FEB firmwares, with epoch 2c being a trigger study with paddles. As can be seen on the plots, the individual epochs vary only slightly, and only in a small normalization difference. We decided to calibrate the entire period 2 together, without splitting it into any smaller samples.

### Period 2 relative calibration results

The results of the attenuation fit for period 2 are summarised on Fig. ??, showing the fitted response at the centre of each cell, or a blank cell if it failed calibration.

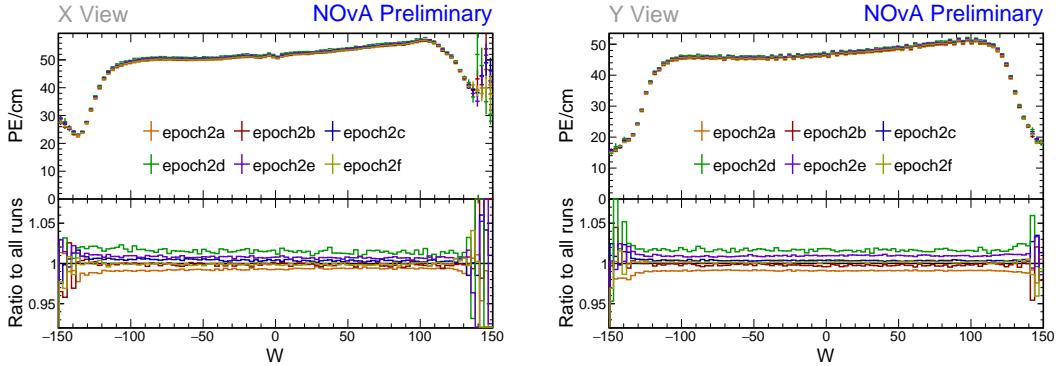


Figure 1.20: Uncorrected average energy response as a function of the position within a cell ( $w$ ) for epochs in period 2. It is clear that there is no significant difference between the various epochs.

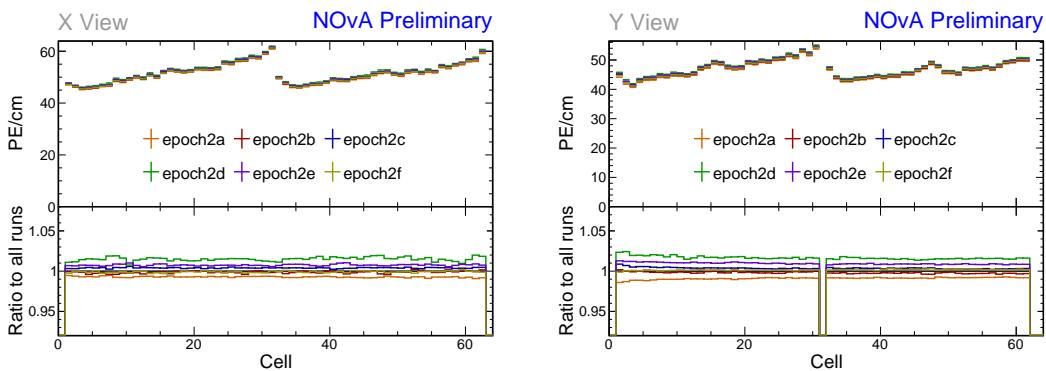


Figure 1.21: Uncorrected average energy response as a function of cells for epochs in period 2.

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. 1.24.

Some cells have a non-flat response across the cell, with one or more regions with a drop in the energy response, as shown on the right plot of Fig. 1.24. These low regions are (almost certainly) a real physical effect caused by zipped, or possibly even twisted fibres [?], present in all of NOvA’s detectors. Relative calibration corrects for this effect in data, but zipped fibres are not included in simulation, for any of the detectors. This could potentially cause issues with the ADC threshold in 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 on Fig. 1.25. The neighbouring cells in plane 1 don’t pass calibration due to low statistics and therefore large fluctuations, as shown on Fig. 1.26. 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

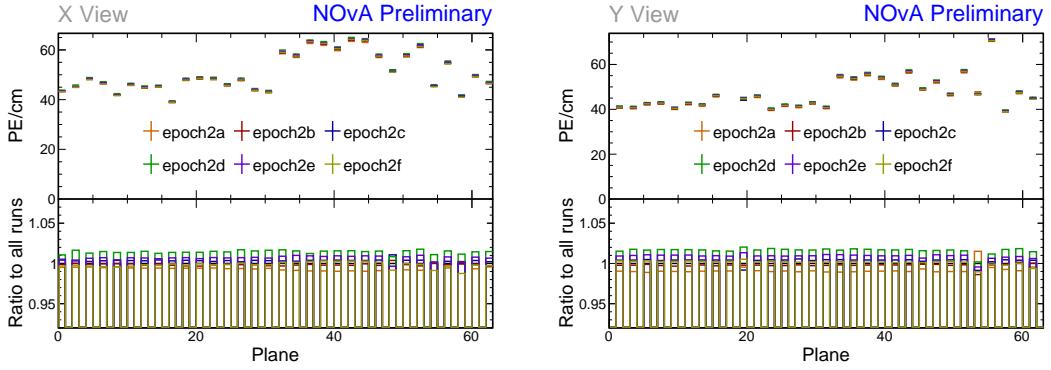


Figure 1.22: Uncorrected average energy response as a function of planes for epochs in period 2.

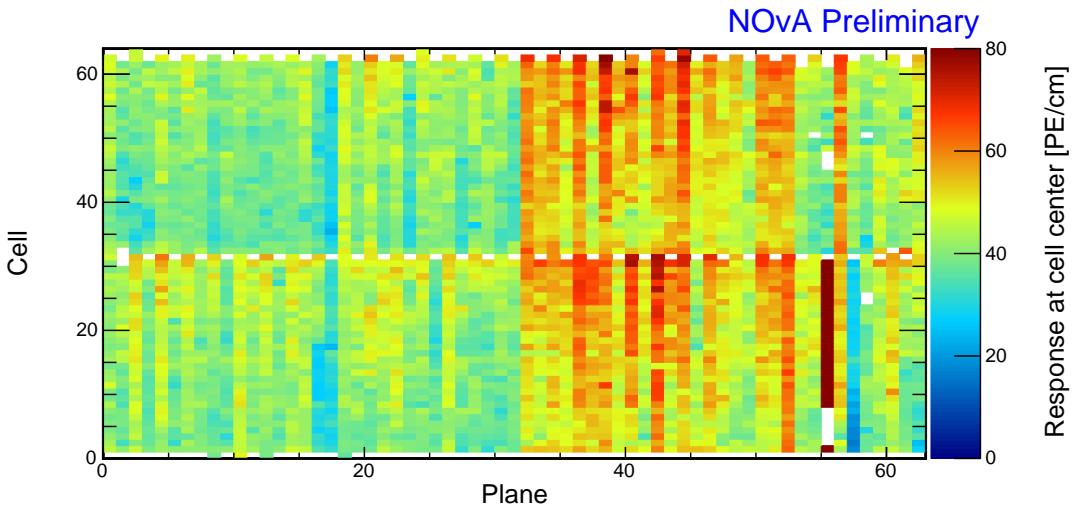


Figure 1.23: Overview of the relative calibration results for the Test Beam detector period 2 data. Each cell represents the result of the attenuation fit to the energy response in the centre of that cell. The blank cells are uncalibrated.

(accepted) hits than the center, as shown on Fig. 1.19.

The left half of plane 55 has more than  $3\times$  larger response than it's surrounding planes, as shown on the left plot of Fig. 1.27. Similarly, the left half of plane 57 has slightly lower response than the surrounding planes, as shown on the right plot of Fig. 1.27. This is due to the corresponding APDs/FEBs incorrectly recording different energy response than the real energy deposited in the detector. Since this effect is present for all data, not only for the cosmic muons used for calibration, it is important to correctly calibrate it out. The issue can arise if these FEBs have been "faulty" only for a limited time 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,

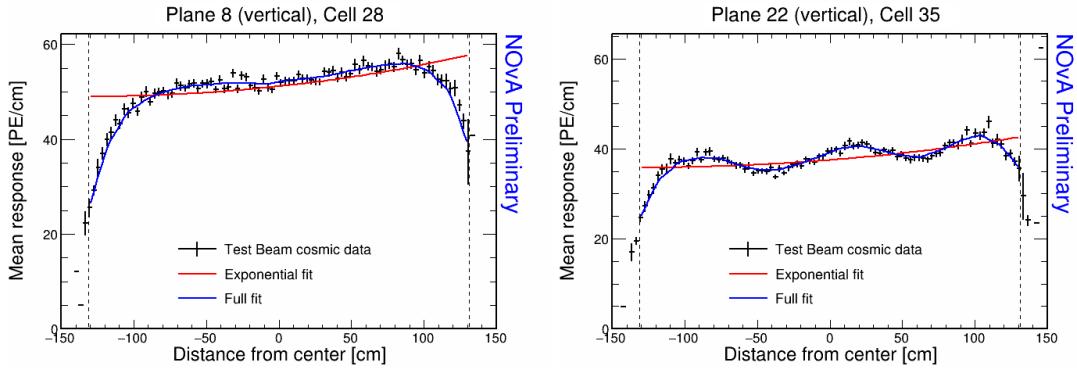


Figure 1.24: Attenuation fits for a selection of cells in period 2. Left plot shows an example of the standard energy deposition in the Test Beam. Right plot shows the effect of zipped fibres.

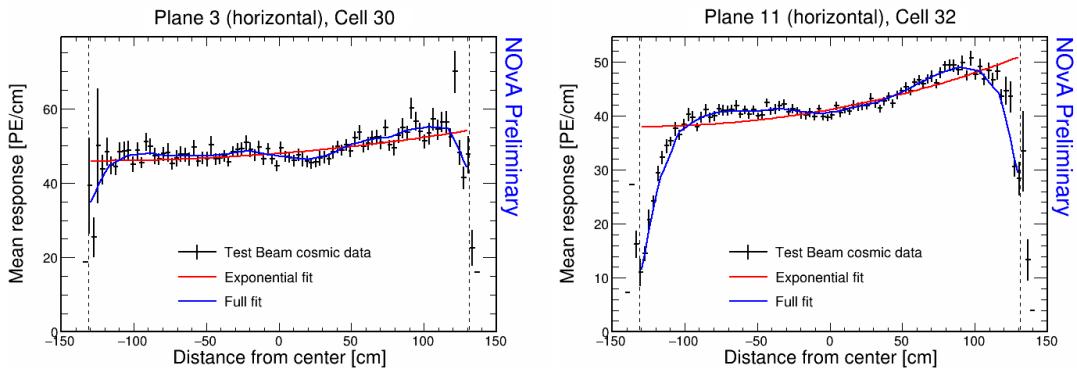


Figure 1.25: Fit to the energy response in period 2. The cells neighbouring the under-filled cells have fewer events and therefore larger fluctuations than the "usual" Test Beam cell.

calibration is going to assume the response was  $4\times$  larger the entire time, which is incorrect. Since both of these planes are in the back of the detector, we decided to ignore this effect for period 2.

The cells with (probably) swapped cables in plane 55 have almost no events, which affects both them and their neighbours as shown of Fig. 1.28.

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 on Fig. 1.29. It is unknown if this is a real physical effect, possibly related to the fibres, or if it is unfiltered noise hits, or something else entirely. Since these cells are in the end of the detector, it is fairly safe to ignore them.

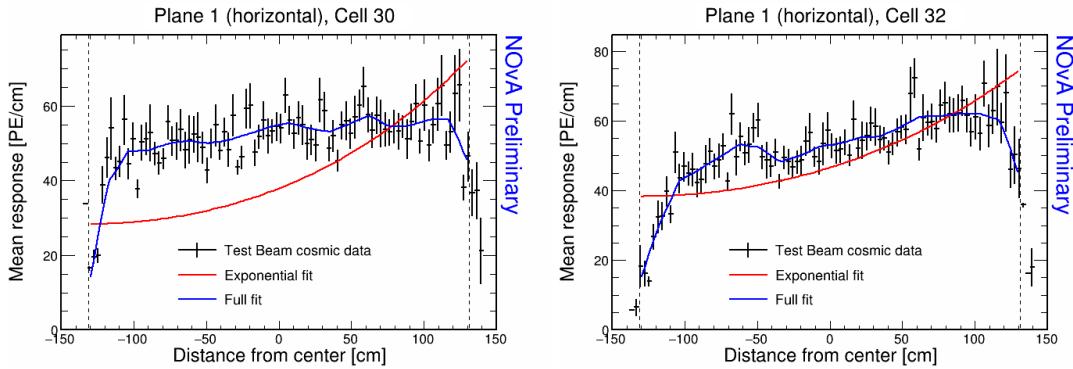


Figure 1.26: Fit to the energy response in period 2. The neighbouring cells to the underfilled cells in plane 1 are uncalibrated due to low statistics.

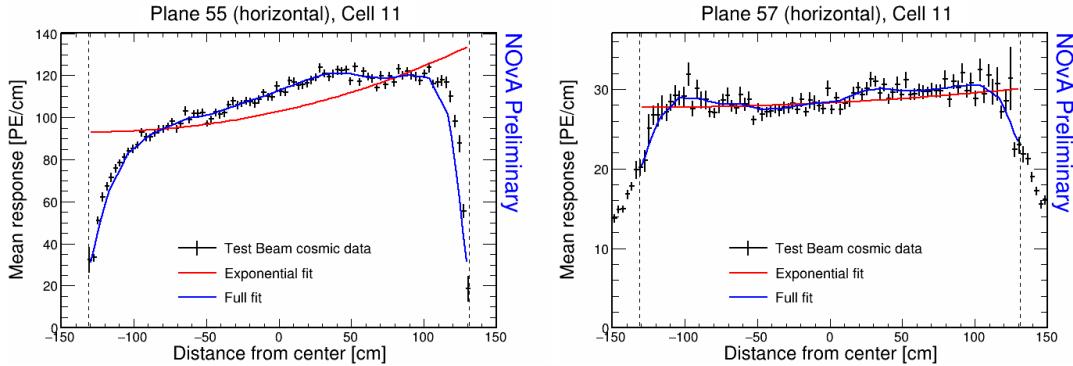


Figure 1.27: Fit to the energy response in period 2. Lower halves of planes 55 and 57 have a different scale of energy response than the surrounding planes.

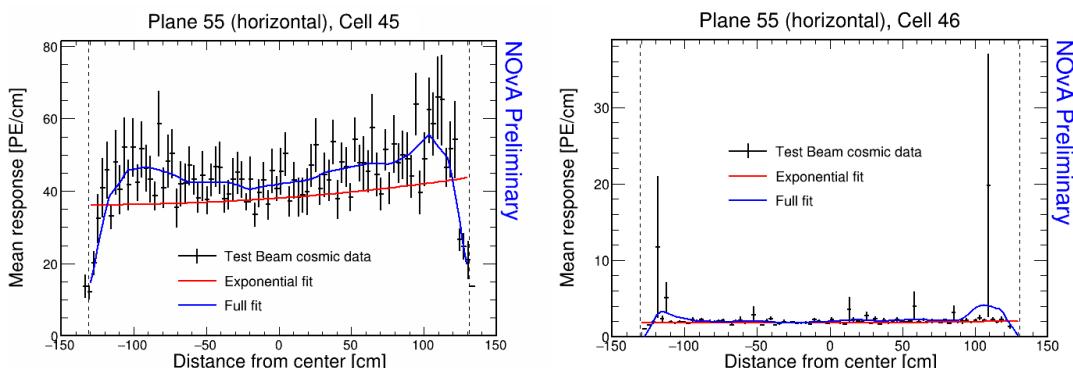


Figure 1.28: Fit to the energy response in period 2. Cells with swapped readout cables have almost no recorded events as shown on the right plot. This also affect their neighbouring cells due to the tricell condition as shown on the left plot.

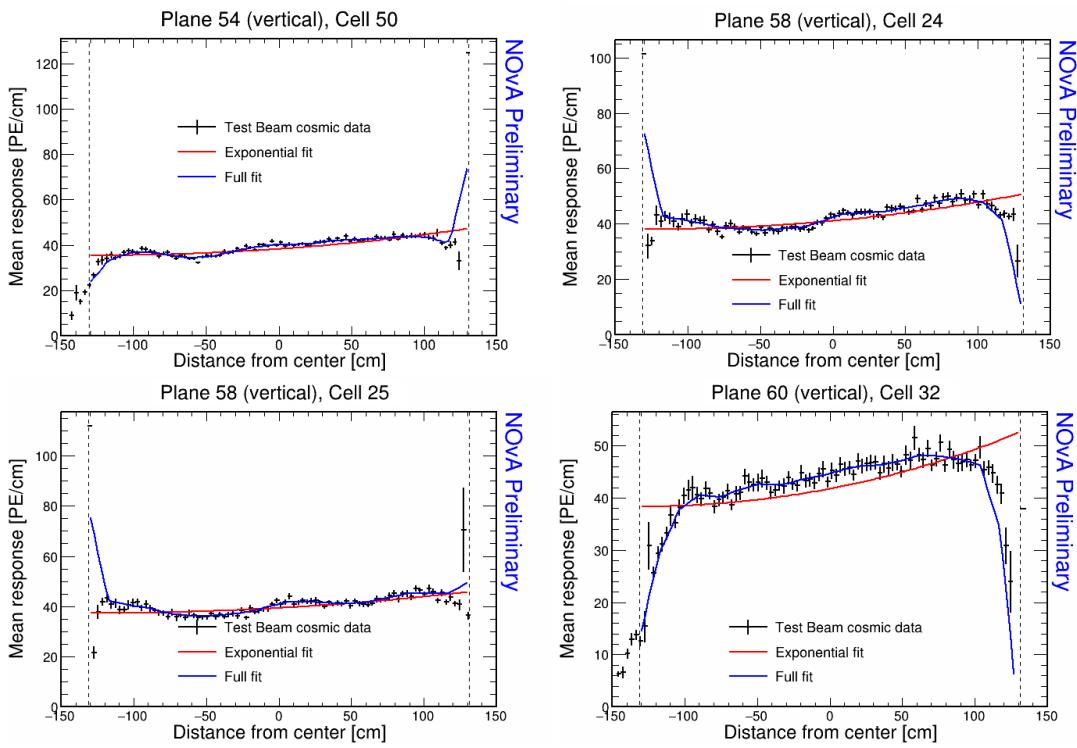


Figure 1.29: Fit to the energy response in period 2. Examples of cells that have an unusually high or low energy response at the edge of the cell. These cells are not calibrated.

### 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. One more major event that could impact the Test Beam data is the replacement of several faulty FEBs, which motivated the creation of epoch 3e.

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

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

The refilling of the underfilled cells can be clearly seen on the cell hits distribution on Fig. 1.30 and on the distribution of energy deposition across horizontal cells (Y view) on Fig. 1.32.

From the cell hits distributions 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.

The energy distributions across cells and planes in the X view (vertical) on Fig. 1.32 and 1.33 shows, that the top half of plane 58 has a very distinctly different energy deposition compared to the rest of the cells. However Fig. 1.30 shows that this part has the same number of events. This is the FEB, that has the largest impact on the calibration (and overall) out of the faulty FEBs replaced before the start of epoch 3e.

From these considerations, we decided to calibrate epochs 3a, 3b and 3c together (all epochs containing any underfilled cells) and to separately calibrate epochs 3d and 3e. 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

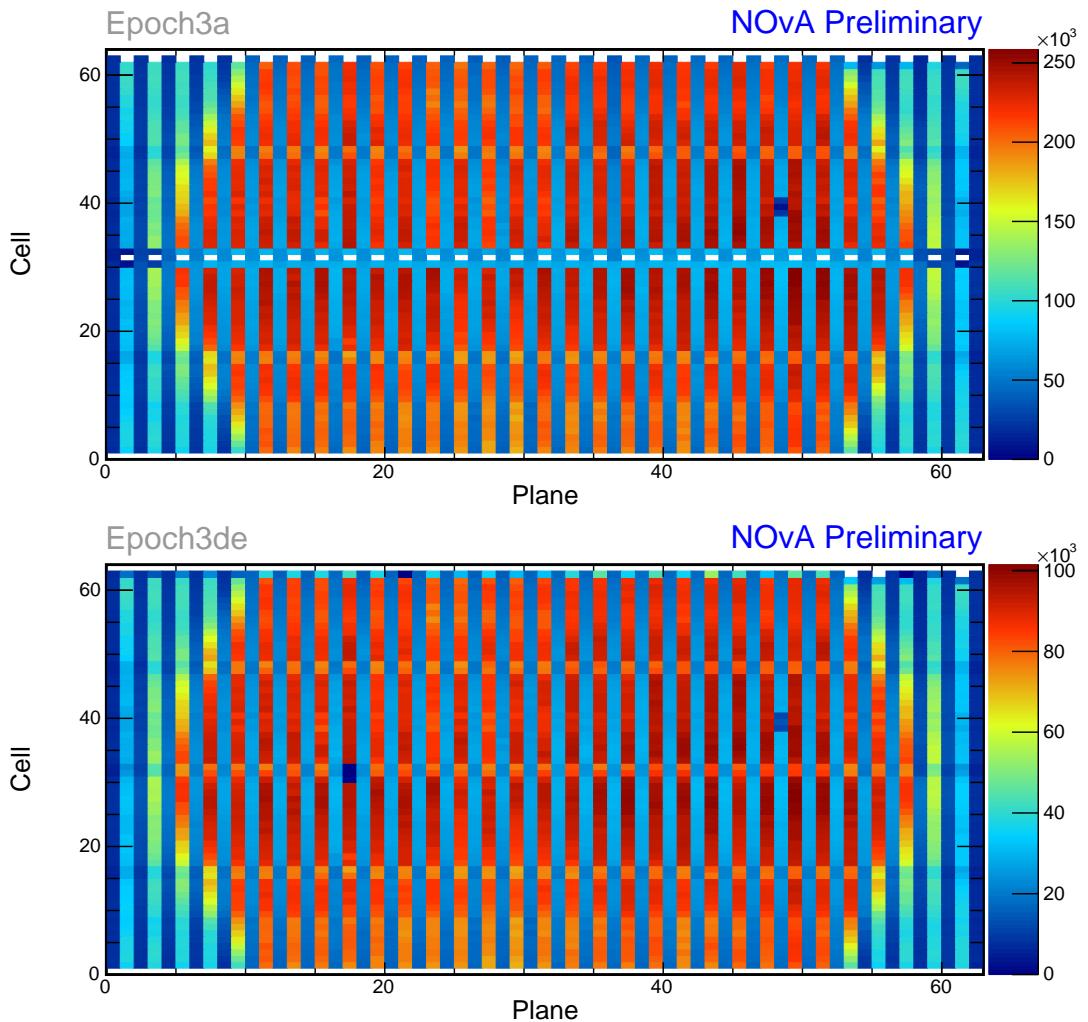


Figure 1.30: Distribution of events in the period 3 Test Beam data calibration sample. Comparison of Epoch 3a data before the refilling of the underfilled cells and Epoch 3de (combination of Epochs 3d and 3e) after the full refilling.

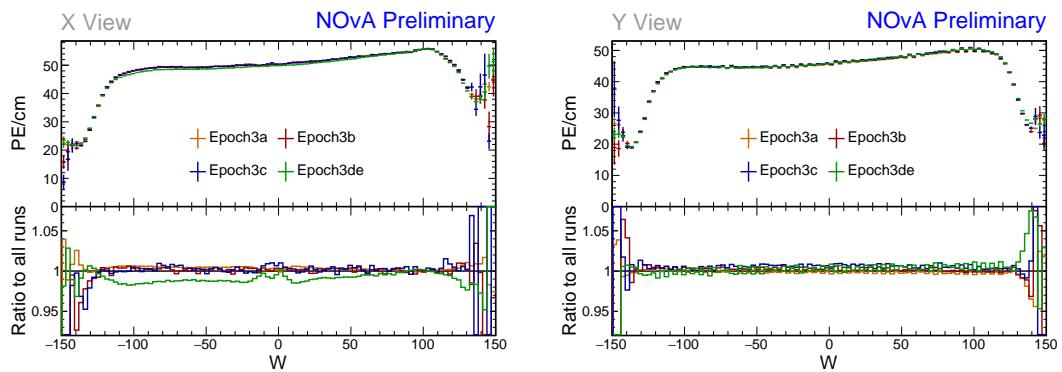


Figure 1.31: Uncorrected average energy response as a function of the position within a cell ( $w$ ) for epochs in period 3.

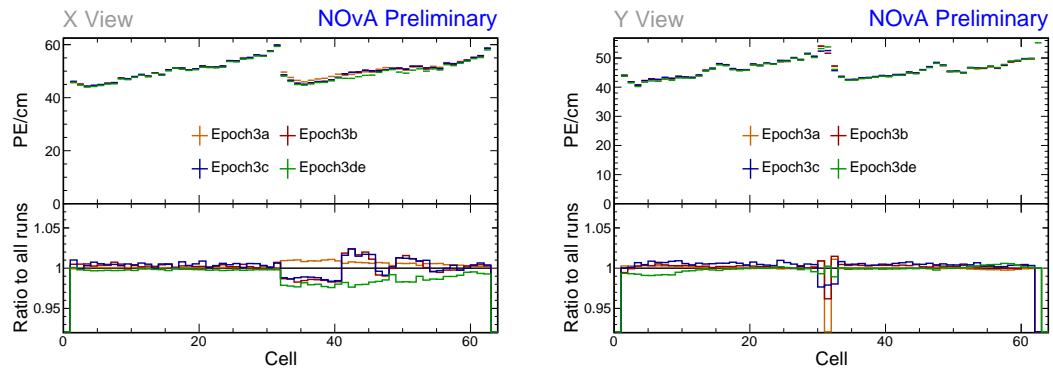


Figure 1.32: Uncorrected average energy response as a function of cells for epochs in period 3.

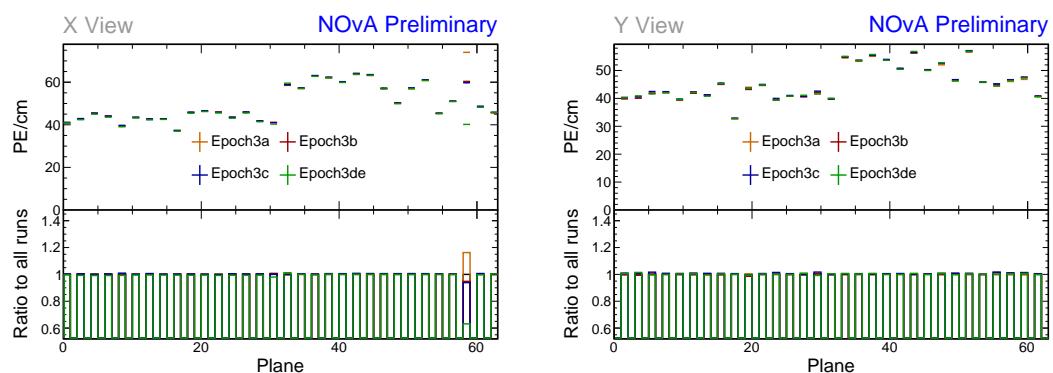


Figure 1.33: Uncorrected average energy response as a function of planes for epochs in period 3.

3b and 3c only contain a few days worth of data, which wouldn't be enough for a successful attenuation fit.

### Combined epochs 3a, 3b and 3c relative calibration results

The results of the attenuation fit are summarised on Fig. 1.34 showing cell  $\times$  plane distribution of the fitted response at the centre of each cell.

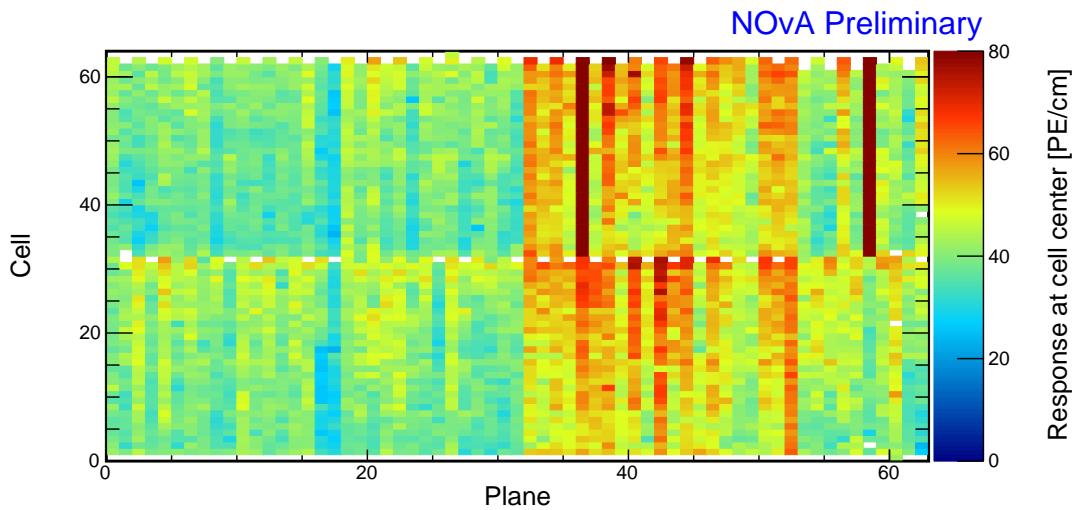


Figure 1.34: Overview of the relative calibration results for the Test Beam detector period 3, combined epochs 3a, 3b and 3c data. Each cell represents the result of the attenuation fit to the energy response in the centre of that cell. The blank cells are uncalibrated.

We can see that some of the underfilled cells that have been refilled for epochs 3b or 3c 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 plot of Fig. 1.35.

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. 1.35. This is due to the low statistics at the edges of the detector.

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, it's 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. The energy deposition for these cells is shown on Fig. 1.36.

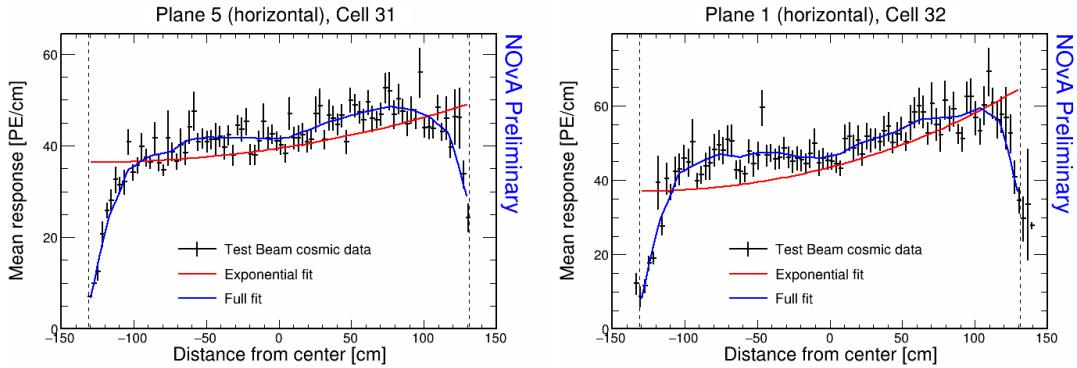


Figure 1.35: Fit to the energy response in epochs 3 a, b and c. Some underfilled cells that have been refilled in epochs 3b and 3c are now calibrated as shown on the left plot. Cell 32 in plane 1 is the only neighbouring cell to the underfilled cell that didn't manage to get calibrated due to low number of events.

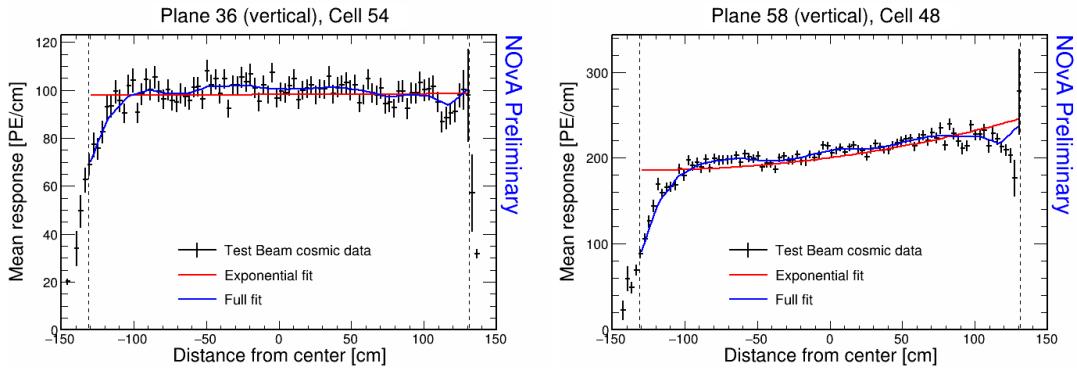


Figure 1.36: Fit to the energy response in epochs 3a, 3b and 3c. The most obvious faulty FEBs that have a significantly larger energy response than their neighbours.

Similarly as in 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 them to be uncalibrated. This can be seen on Fig. 1.37.

### Combined epochs 3d and 3e relative calibration results

The results of the attenuation fits for epochs 3d and 3e are shown on Fig. 1.38. There we can see the expected uncalibrated cells in plane 17 related to the dead channel (or possibly still underfilled cell). The energy deposition for this cell and one of its neighbours is shown on Fig. 1.39.

Epochs 3d and 3e should have all the previously underfilled cells now refilled, but as can be seen on Fig. 1.38, there's several of these cells that are still (officially) uncalibrated. The energy deposition in these cells is shown on Fig. 1.40. Here we can see that these cells have a fairly large discrepancy between the left and right side

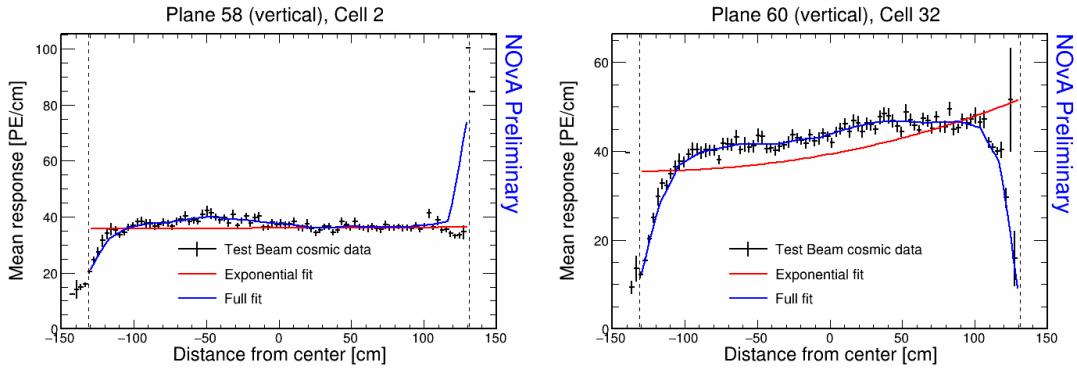


Figure 1.37: Fit to the energy response in epochs 3a, 3b and 3c. Some cells are not calibrated due to large fluctuations at one edge of the cells.

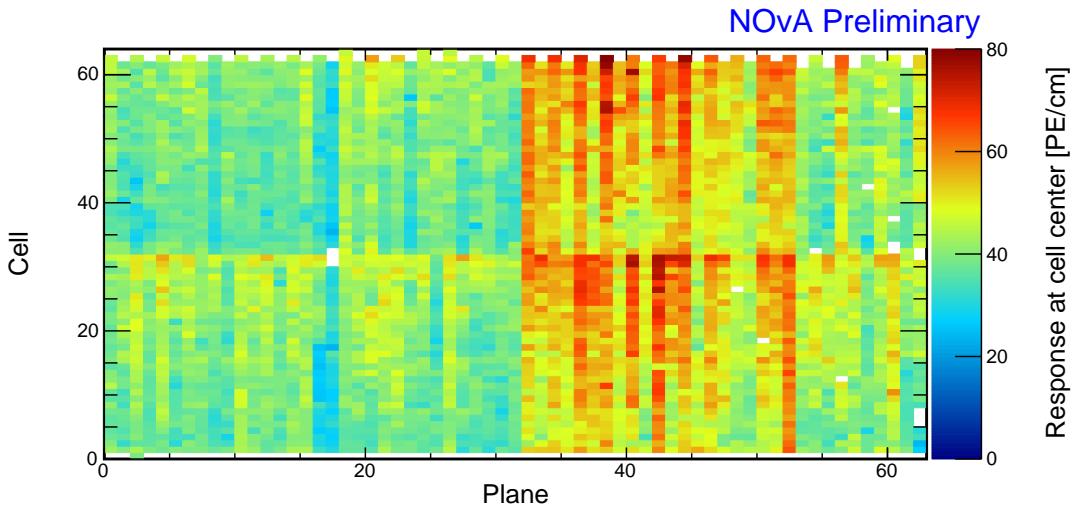


Figure 1.38: Overview of the relative calibration results for the Test Beam detector period 3, combined epochs 3d and 3e data. Each cell represents the result of the attenuation fit to the energy response in the centre of that cell. The blank cells are uncalibrated.

of the cells. This is caused by using different scintillator oils for the initial filling of the cells and for the refilling. Specifically, 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 causes a different 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. The large  $\chi^2$  value is most likely caused only by the unusual shape of the distribution, which the fit is not built for. **We have therefore decided to manually change the  $\chi^2$  inside the cvs tables (results of the attenuation fits), so that the  $\chi^2 < 0.2$  and these cells are officially considered calibrated.**

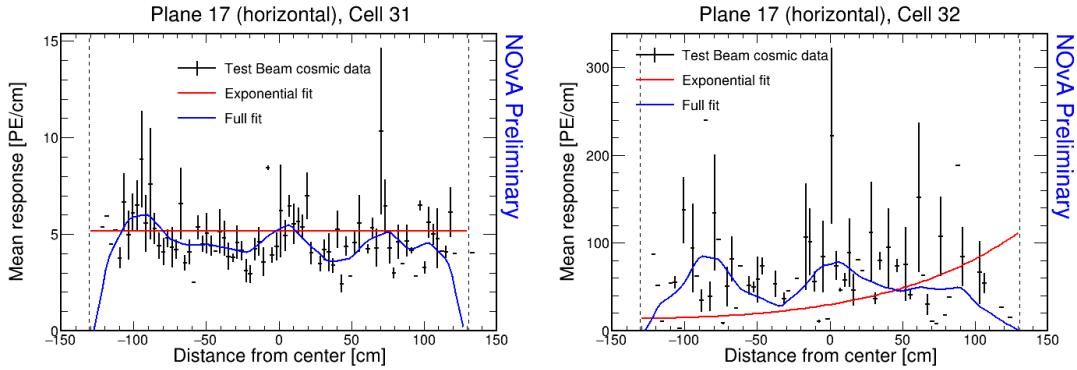


Figure 1.39: Fit to the energy response in epochs 3d and 3e. Possibly dead channel or still underfilled cell.

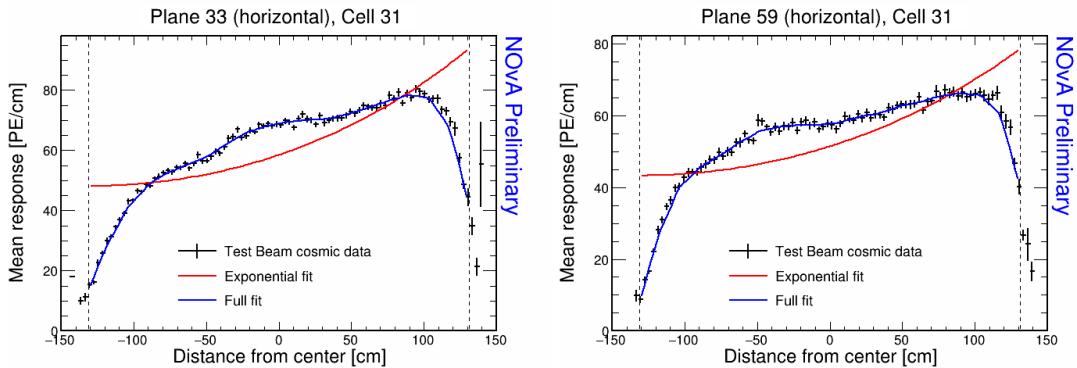


Figure 1.40: Fit to the energy response in epochs 3 d and e. The scintillator oil used for refilling of the underfilled cells has lower energy response than the oil used for the initial filling. These oils didn't mix properly causing a different energy response in the left and right side of the cell.

Some of the cells in the back of the detector have a rise, or drop in energy deposition at the edge of the cell, as can be seen on Fig. 1.41. 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.

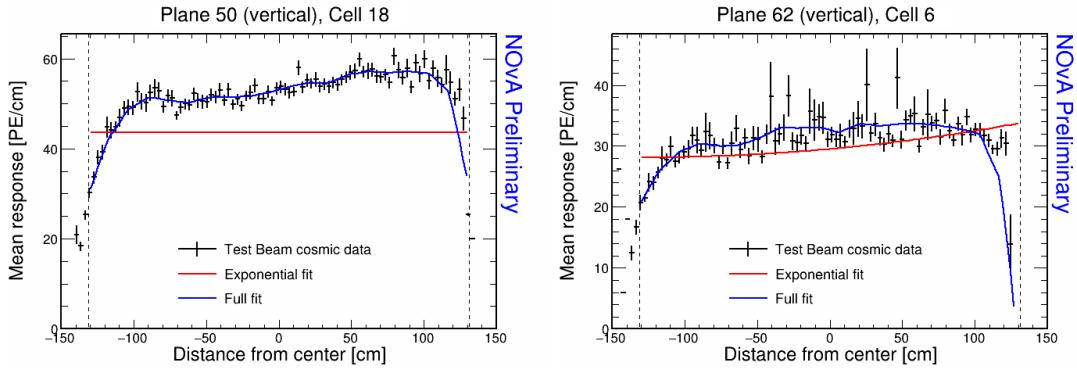


Figure 1.41: Fit to the energy response in epochs 3d and 3e. Some cells have a drop, or a rise of energy response at the edge of the cell. This can be caused by low statistics.

### 1.3.6 Period 4 data

The period 4 Test Beam data taking period is the best data we managed to collect with almost an ideal detector conditions. There's only 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 make up epoch 4a, shown on the top plot of Fig. 1.42. There is also a few runs during which we masked parts of the detector to help with FEB saturation [?], which can clearly be seen on the middle plot of Fig. 1.42.

Figures 1.43, 1.44 and 1.45 show that the epoch 4a and the cell masking study did have a noticeable impact on the energy deposition across the detector. We have therefore decided to ignore these runs and only calibrate the rest of the period 4 data.

#### Period 4 relative calibration results

Results of the attenuation fits for period 4 are summarised on Fig. 1.46. We can see that majority of the detector is calibrated, besides some cells on the edge of the detector, a few formerly underfilled cells (left plot on Fig. 1.47), and one cell with an unusually high response at the edge of the cell (right plot on Fig. 1.47).

We treated the formerly underfilled cells the same way as in epochs 3d and 3e, by manually changing their  $\chi^2$  inside the csv files to be  $< 0.2$  and therefore making them officially calibrated.

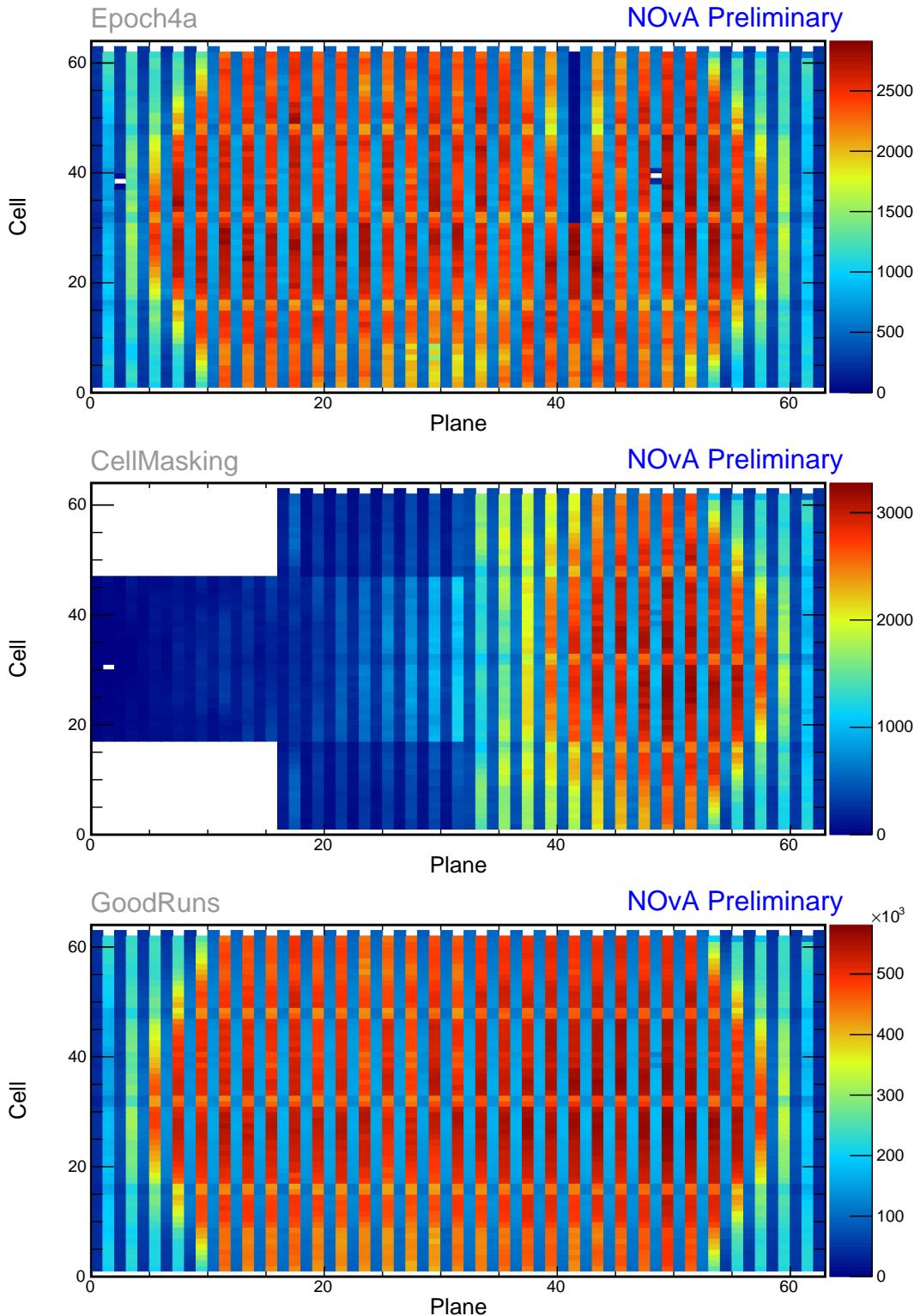


Figure 1.42: Distribution of events in the Test Beam period 4 data calibration sample. The top plot shows the first three commissioning runs, the middle plot the status of the detector during the Cell Masking studies and the bottom plot shows the rest.

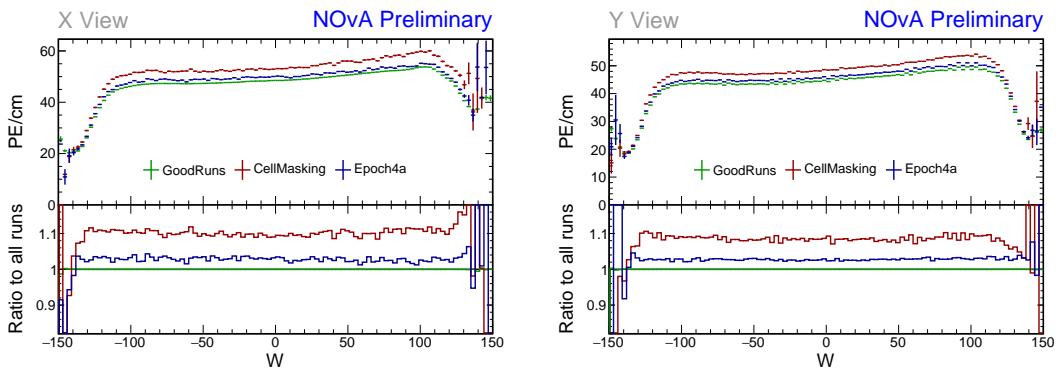


Figure 1.43: Uncorrected average energy response as a function of the position within a cell ( $w$ ) for epochs in period 4.

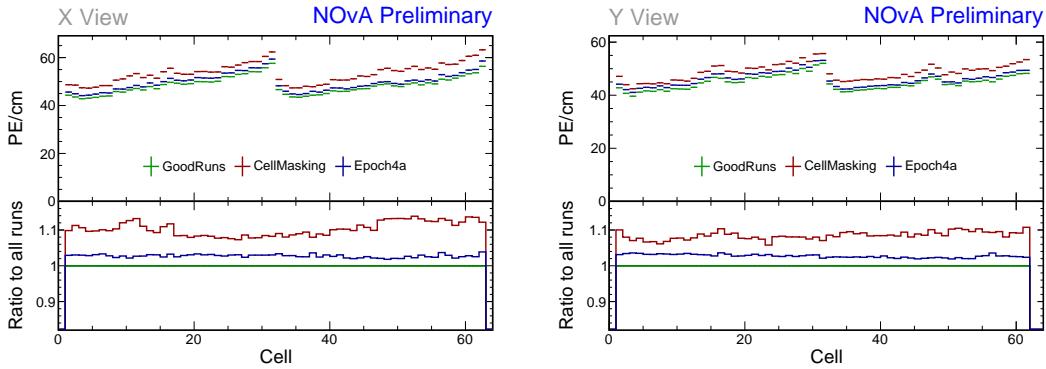


Figure 1.44: Uncorrected average energy response as a function of cells for epochs in period 4.

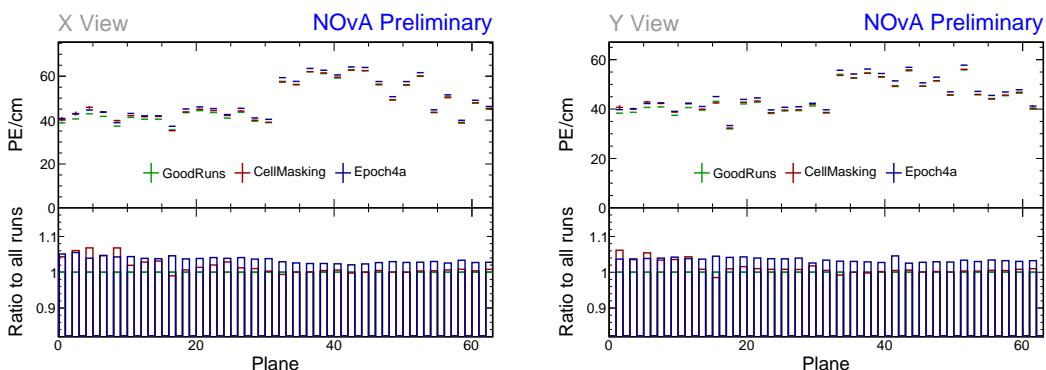


Figure 1.45: Uncorrected average energy response as a function of planes for epochs in period 4.

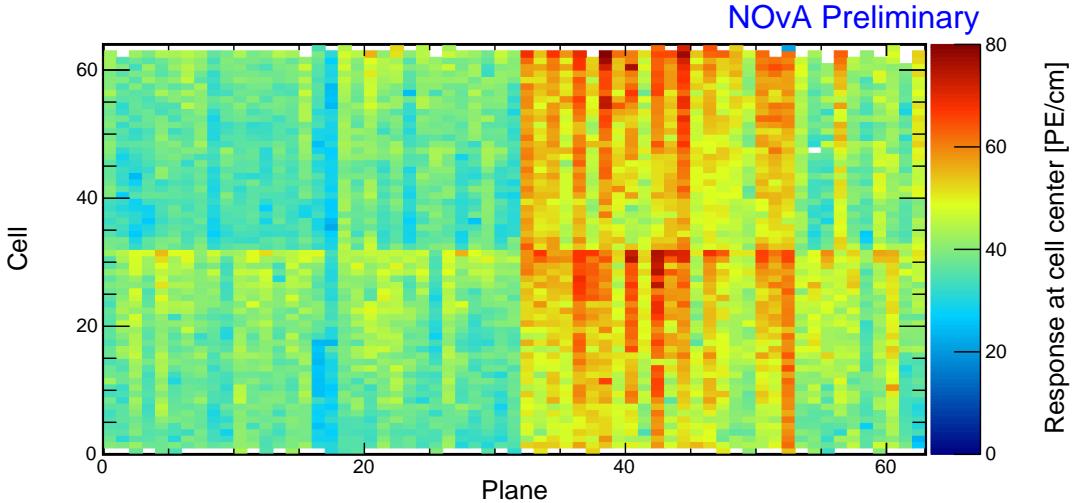


Figure 1.46: Overview of the relative calibration results for the Test Beam detector period 4 data. Each cell represents the result of the attenuation fit to the energy response in the centre of that cell. The blank cells are uncalibrated.

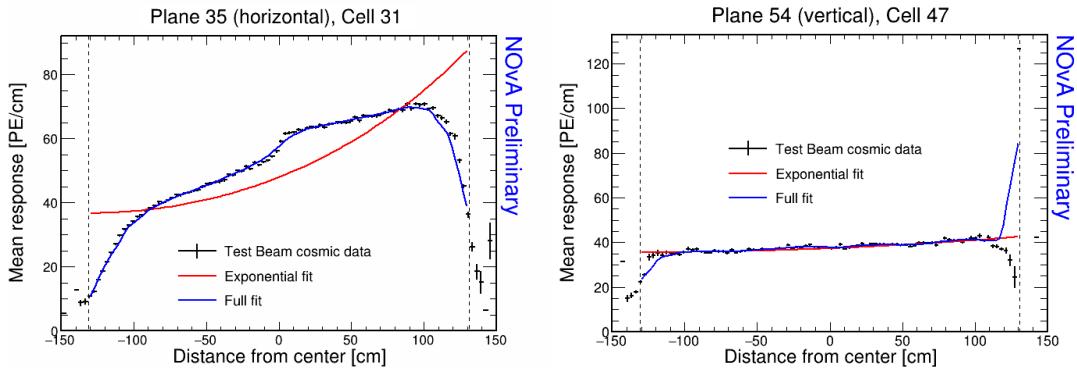


Figure 1.47: Fit to the energy response in period 4. Previously underfilled cells refilled with a scintillator of a different quality causing an unusual distribution of energy deposition (left). Unusually high energy response at the edge of the cell 47 (right).

### 1.3.7 Absolute calibration results

To get the absolute energy scale we look at the stopping muon sample, apply the relative calibration results and the absolute calibration cuts to select only well-understood minimum ionising muons. The absolute calibration cuts are mostly the same as for the other detectors (hits 1-2 m from the end of track, pathlength  $> 0$ , PE  $> 0$ , PECorr  $> 0$ , PECorr/cm  $< 100$ ), but with a smaller cell window  $-80 < w < 80$  cm to remove hits at the cell edges.

We then look at the distributions of the reconstructed energy response in units of PECorr/cm (for all data and simulation samples), and true energy response in units of MeV/cm (only for simulation) in each view, as shown on Fig. 1.48. The mean

of these distributions is the  $\text{MEU}_{\text{Reco}/\text{True}}$  value for each view, with an uncertainty calculated as  $\text{StdDev}/\sqrt{N_{\text{Entries}}}$  from the distribution. The MEU for each sample and view is shown on Tab. 1.4.

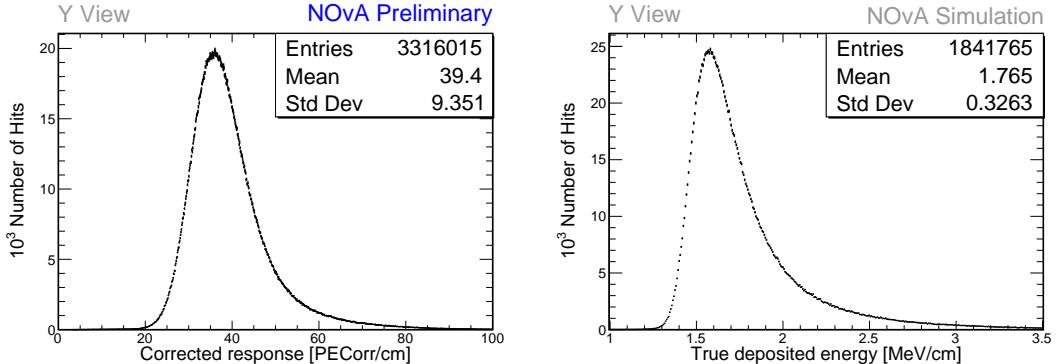


Figure 1.48: Example distributions of reconstructed (left) and true (right) energy response of stopping muons 1-2 m from the end of their tracks. The mean of the reconstructed (true) response is the reconstructed (true) MEU unit.

Sample		X view		Y view		Combined	
		NHits	MEU	NHits	MEU	$\text{MEU}_{\text{Reco}}$	$\sigma_{\text{MEU}_{\text{Reco}}}$
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

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

Table 1.4: Table of the absolute calibration results.  $\text{MEU}_{\text{Reco}}$  values (top table) are in units of PECorr/cm and  $\text{MEU}_{\text{True}}$  values (bottom table) are in units of MeV/cm

We don't apply the absolute energy scale separately for each view, instead we combine the two views into a single  $\text{MEU}_{\text{Reco}}$ , or  $\text{MEU}_{\text{True}}$  value. To combine the MEU value we take a simple average  $(\text{MEU}_X + \text{MEU}_Y)/2$ , without accounting for the different statistics of the two views. To get the uncertainty on the final value, we calculate it as  $\sigma_{\text{Combined}}^2 = \sigma_X^2 + \sigma_Y^2$ . In the past (and for the other NOvA detectors), this uncertainty was calculated as  $1/\sigma_{\text{Combined}}^2 = 1/\sigma_X^2 + 1/\sigma_Y^2$ . This uncertainty is however **not** the uncertainty used in NOvA for the absolute energy scale. Instead, we use a data-simulation comparison of special samples to derive an uncertainty on the absolute energy scale [? ]. The final combined values are highlighted in Tab. 1.4. Here  $\text{MEU}_{\text{Reco}}$  values are all in units of PECorr/cm.

For each calibrated sample we write the combined  $MEU_{Reco}$  value into an `calib-abs_consts.csv` file, together with its uncertainty and with the  $MEU_{True}$  value, which is common for all samples.

### 1.3.8 Results

The results of the relative and the absolute calibration, in form of the csv files, are stored in `/grid/fermiapp/products/nova/externals/calibcsvs/abs_consts.csv` and are applied within NOvAsoft in the calibration tag v15.09 and higher.

The csv files follow the official NOvA calibration naming convention, which is `calib_{abs/atten}_{consts/points}.{nd/fd/tb}.{data/mc}.{version}.{period}`. Here `version` is the calibration tag (i.e. v15) and `period` is the range of runs for that sample (i.e. `r100857-r101356` for the combined epochs 3a, 3b, and 3c, or `r-r` for simulation, since it is not divided into different periods).

To create the calibration tag we've asked Lisa Koerner for help. Lisa is a NOvA calibration expert who wrote the instructions for calibration tagging [? ]. It is possible to do it ourselves following these instructions, but it is advised to consult the detector systematics group before hand.

We have also stored the final calibration results in a special location created for safekeeping of Test Beam calibration files: `/nova/ana/testbeam/calibration.abs_consts.csv`. Here we have also copied all of the attenuation profiles used in the relative calibration. These can be very useful in case someone wants to re-do the calibration, as it allows to skip the prestaging of the calibration plist samples (the pliststop samples are much smaller and therefore easier to prestage). If there has been no change to the calibration samples, it is possible to skip the creation of attenuation profiles and reuse the existing files.

### 1.3.9 Validation

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 on Fig. 1.49-1.62 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  $PECorr/cm$  distributions are mostly flat across all studied variables.

The distribution of  $PECorr/cm$  across cells in X view on Fig. 1.53 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  $PECorr/cm$  across planes in the X view (Fig. 1.53) 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 on Fig. 1.36), 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 underestimated for the period when the FEB behaved "normally". The  $PECorr/cm$  over Plane plot shows the average over these responses.

The corrected response across planes in Y view (Fig. 1.53) 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. 1.61 and 1.62) 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  $PECorr/cm$  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  $PECorr/cm$ , 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.

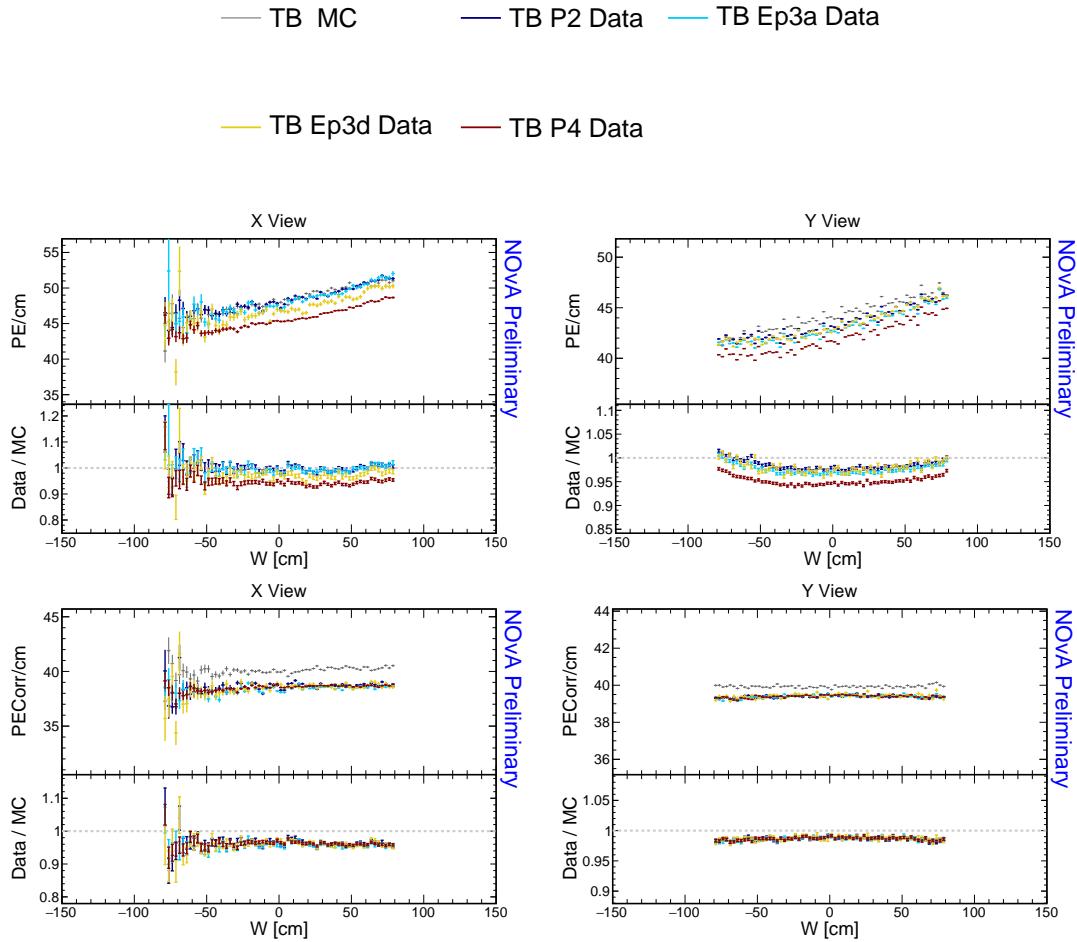


Figure 1.49: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the position within a cell.

## 1.4 Conclusion

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

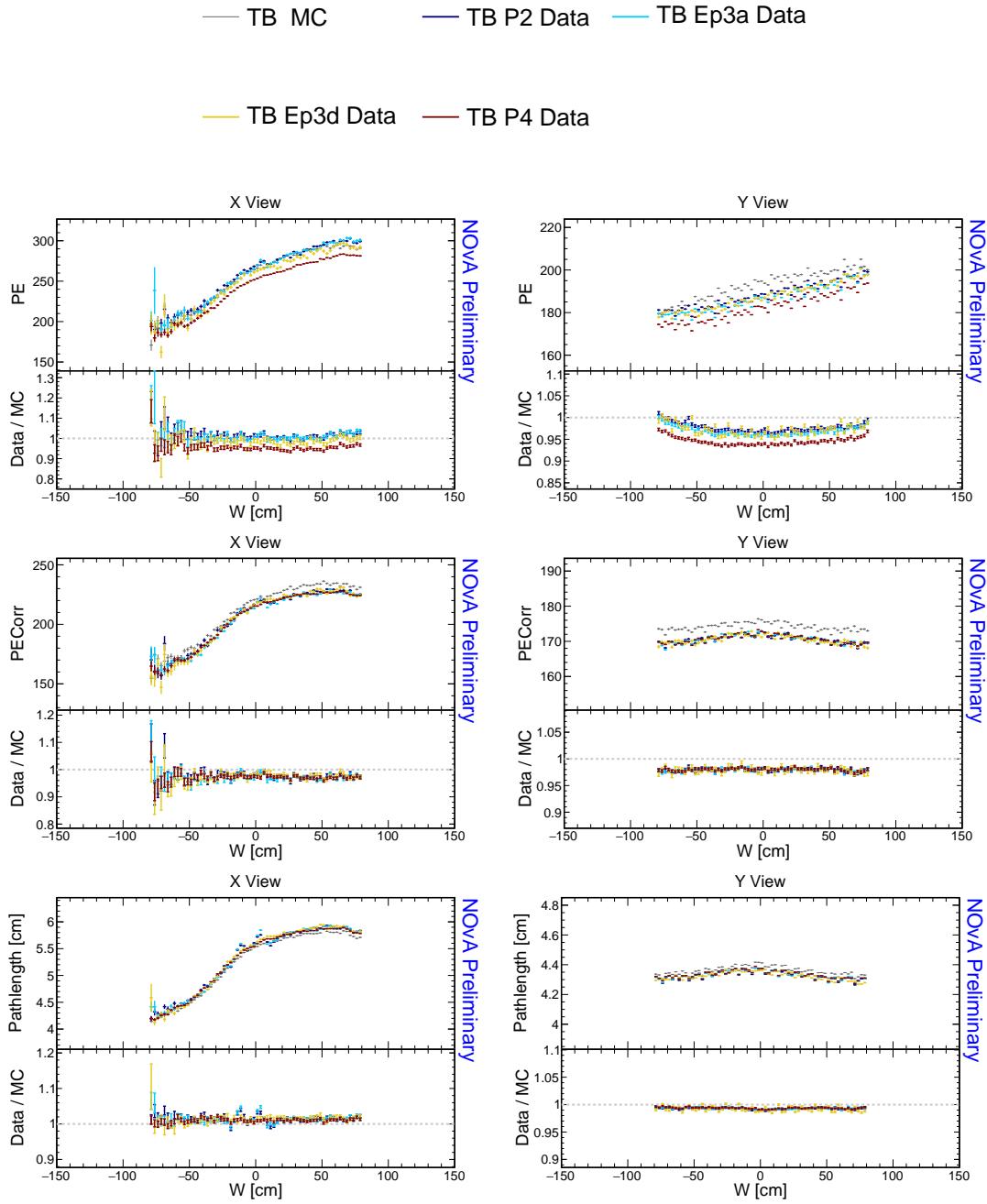


Figure 1.50: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the position within a cell.

technical note.

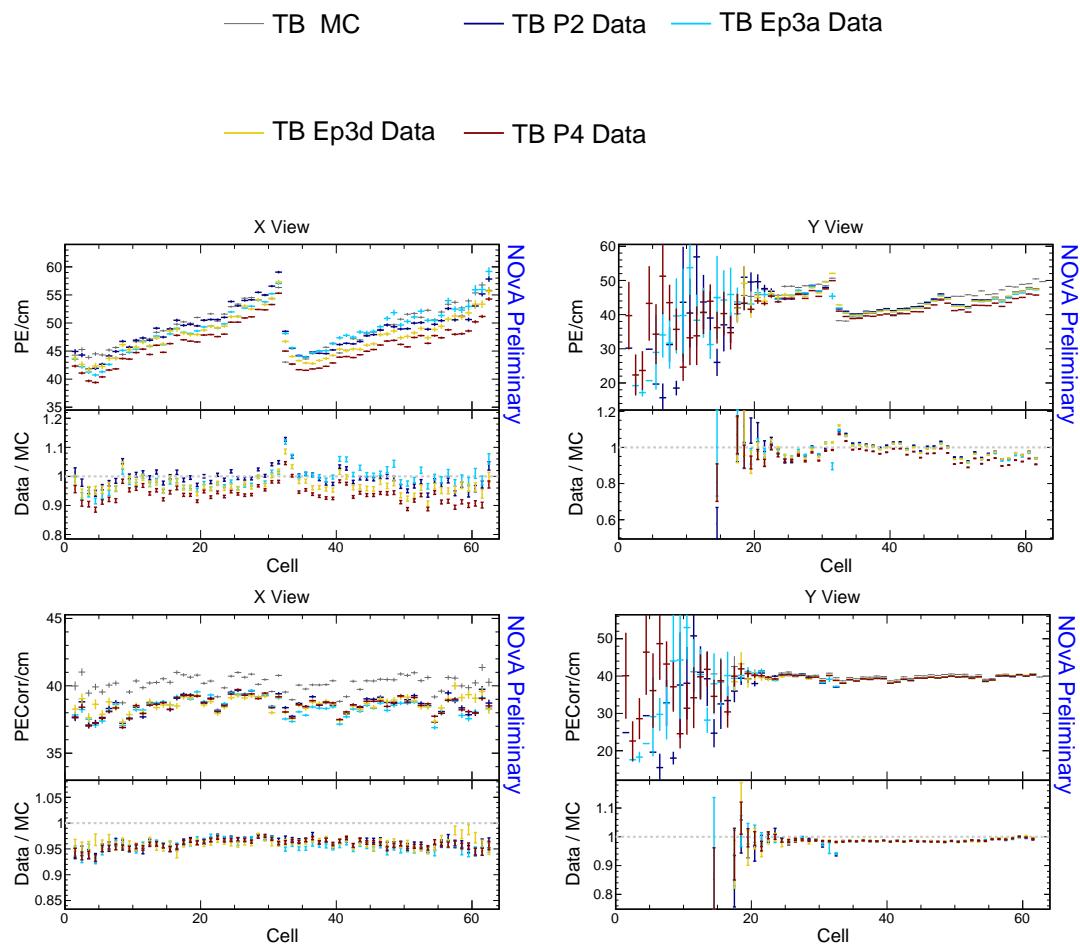


Figure 1.51: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the cells of the detector.

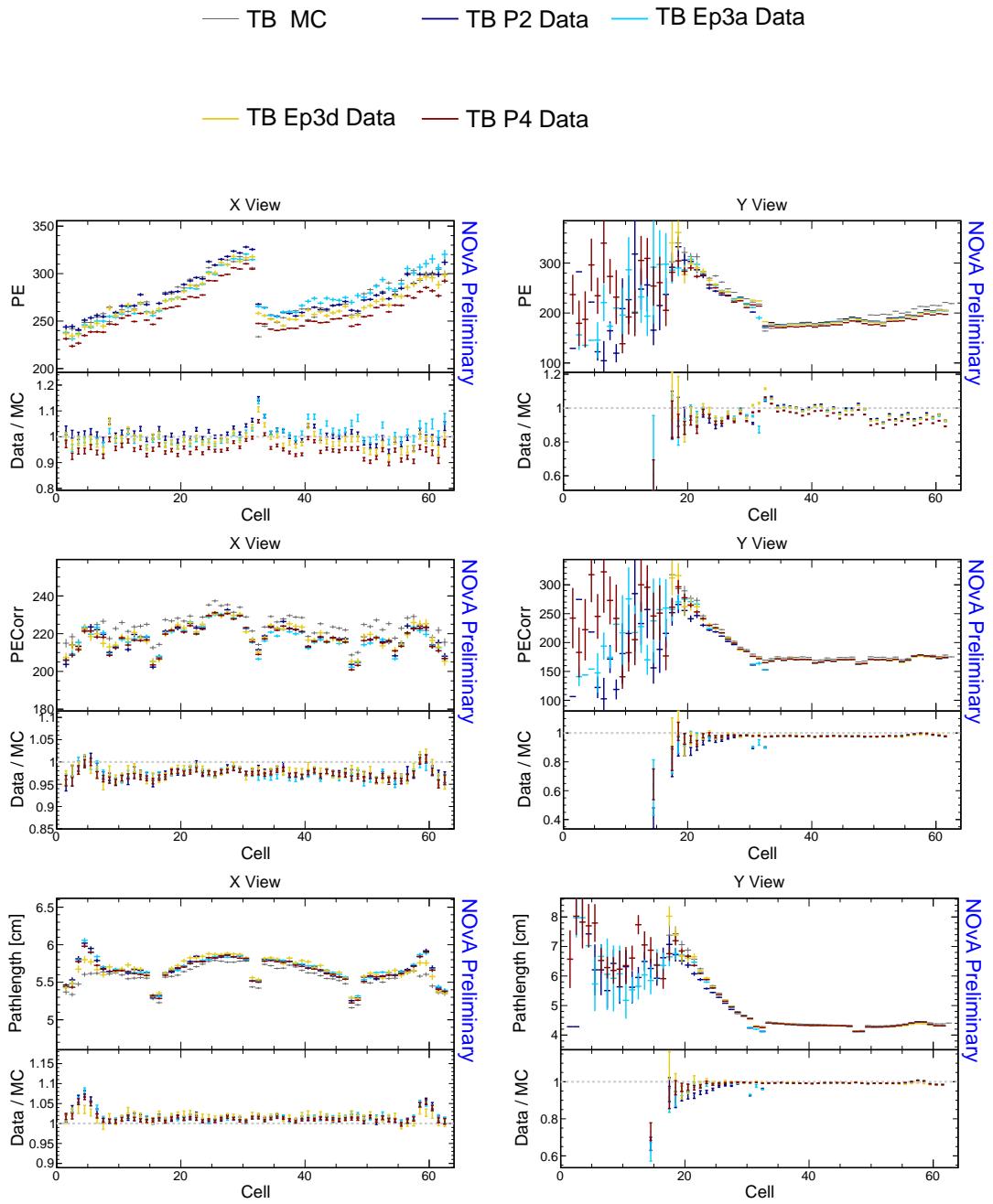


Figure 1.52: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the cells of the detector.

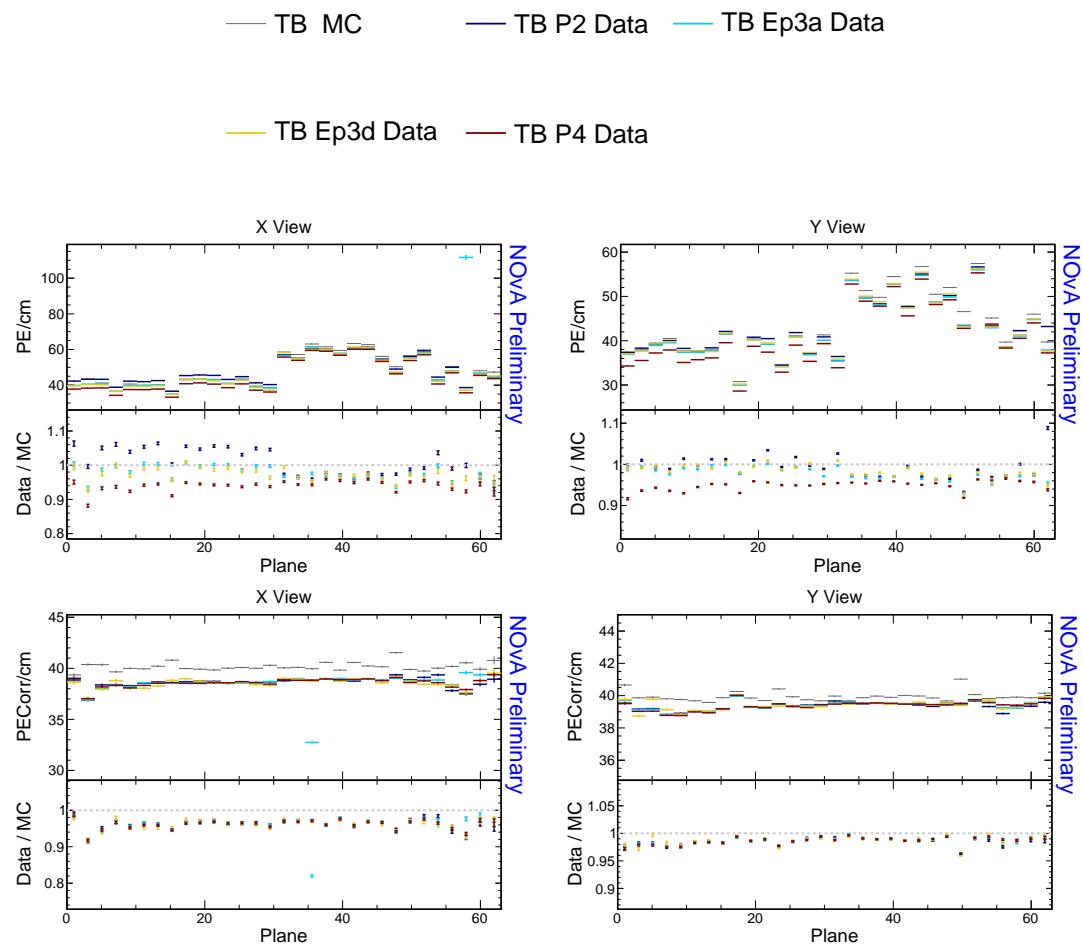


Figure 1.53: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the planes of the detector.

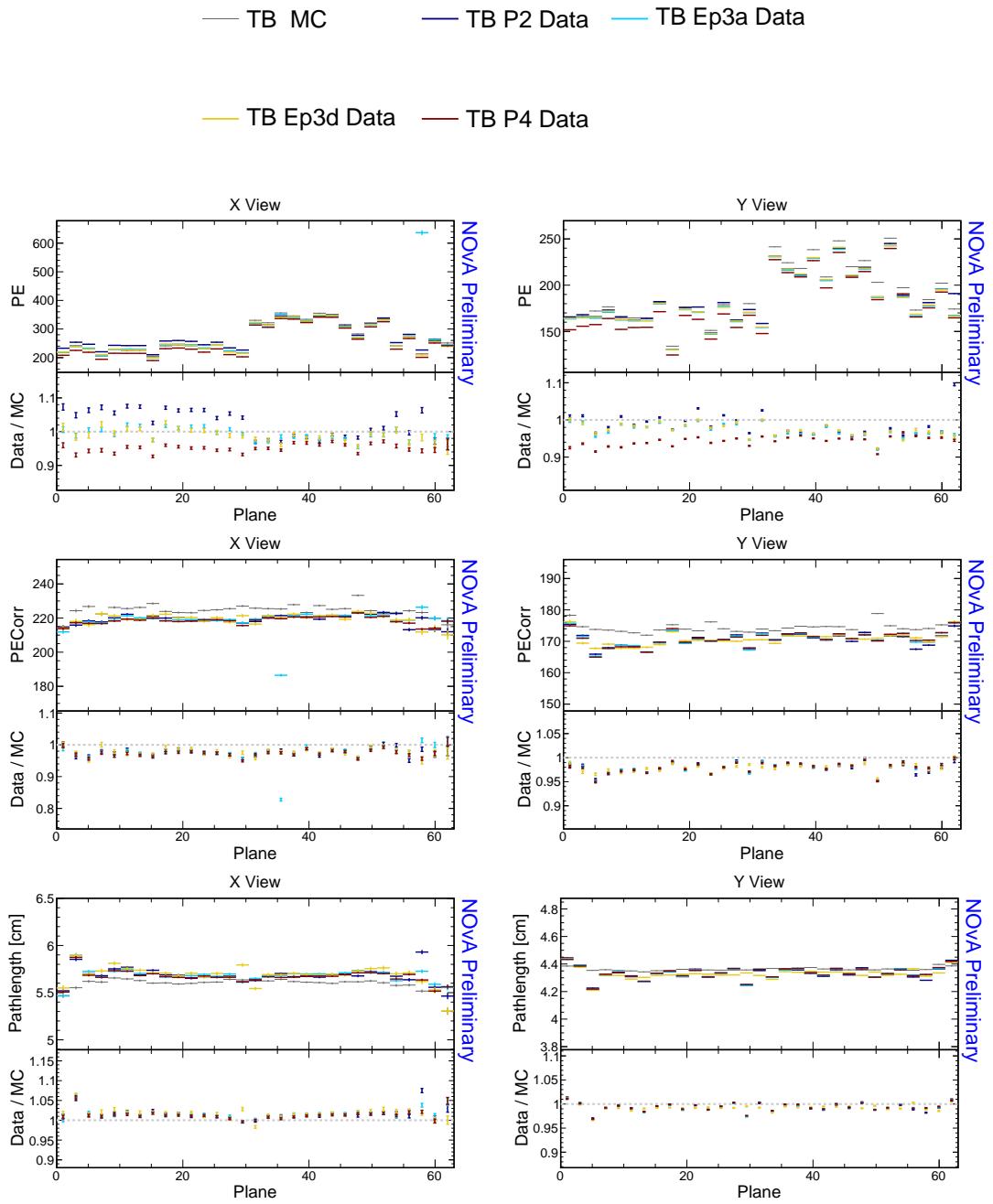


Figure 1.54: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the planes of the detector.

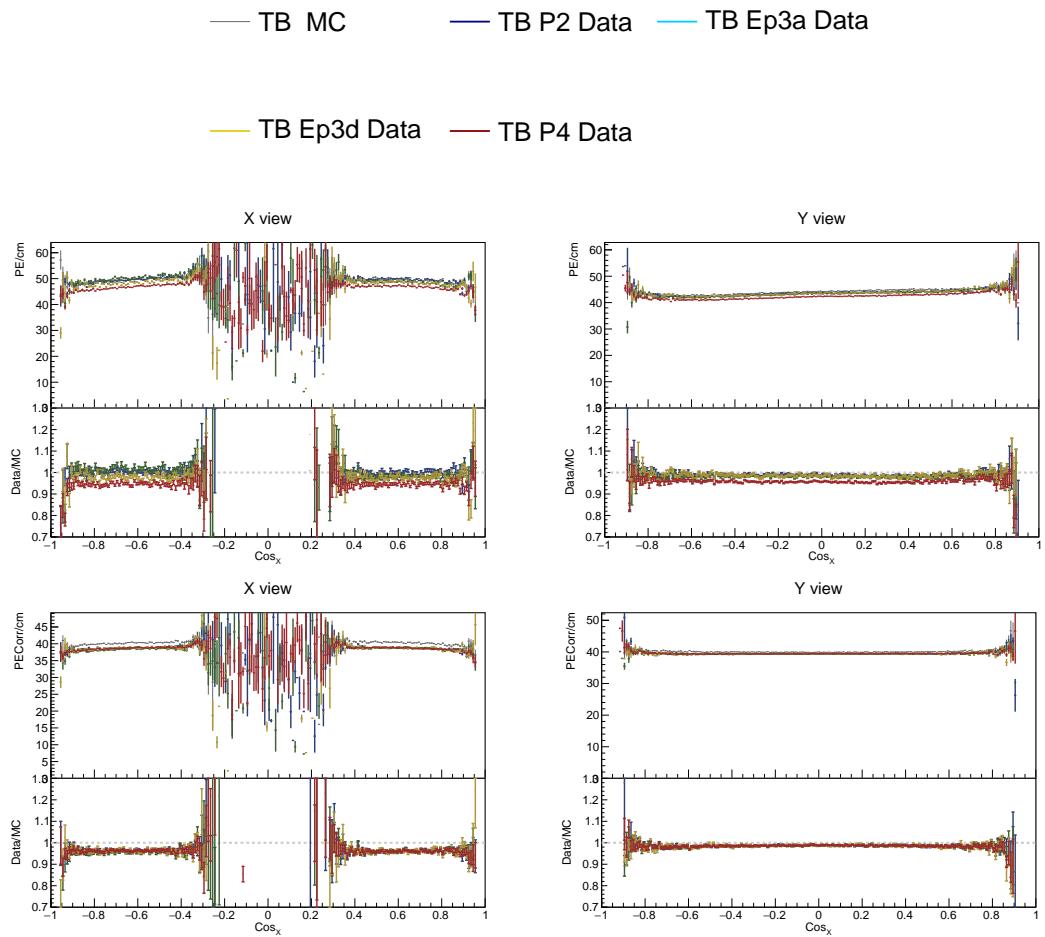


Figure 1.55: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the cosine of the track angle from the X (horizontal) axis.

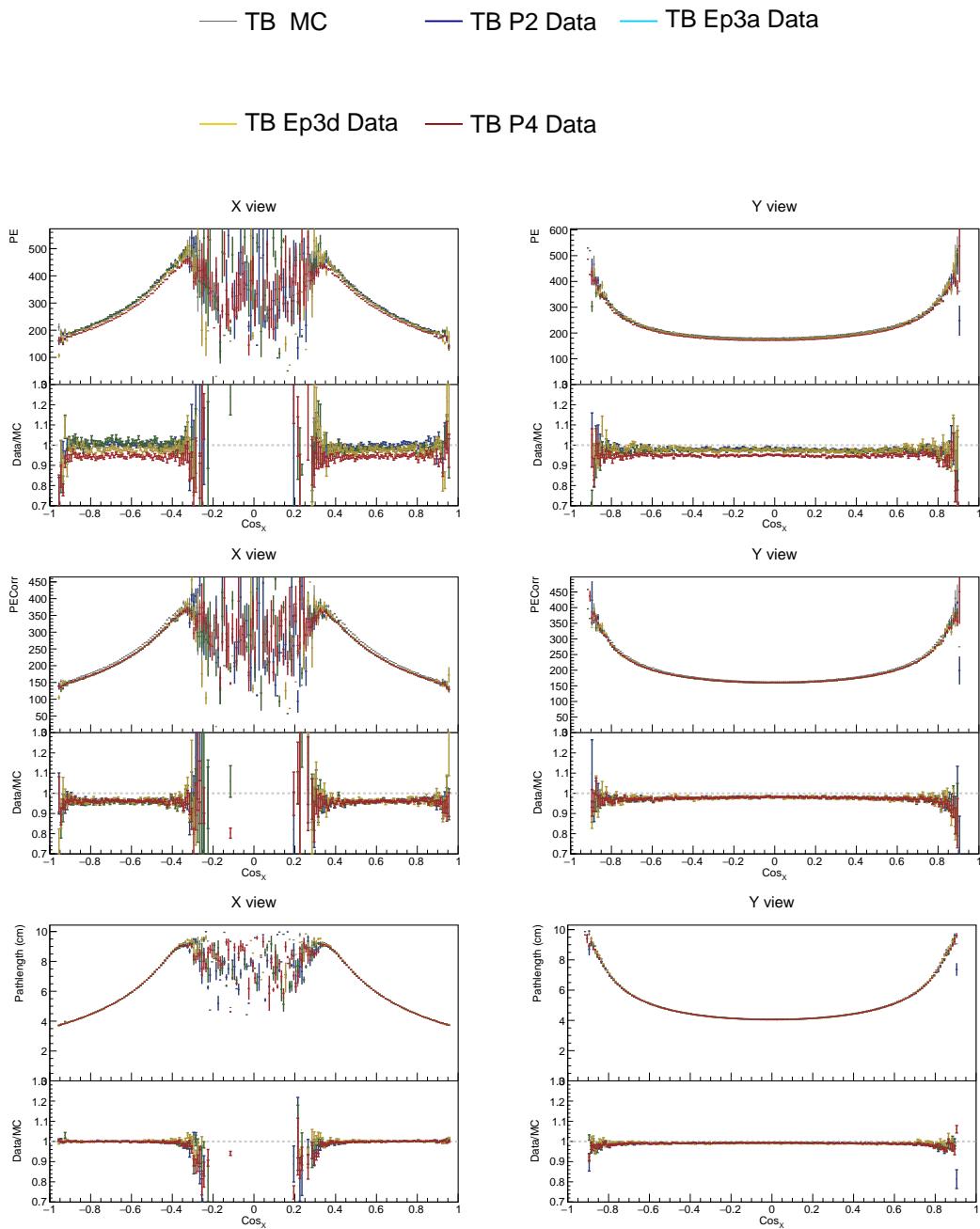


Figure 1.56: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the cosine of the track angle from the X (horizontal) axis.

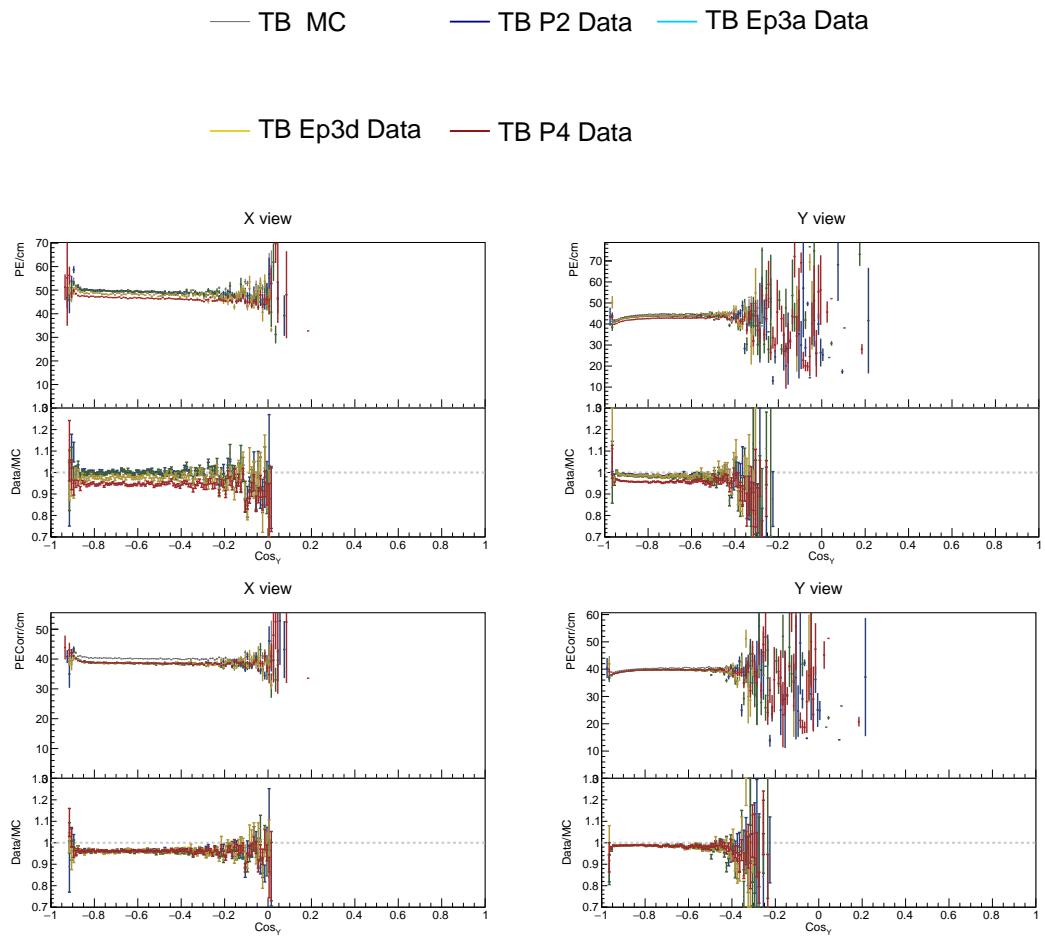


Figure 1.57: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the cosine of the track angle from the Y (vertical) axis.

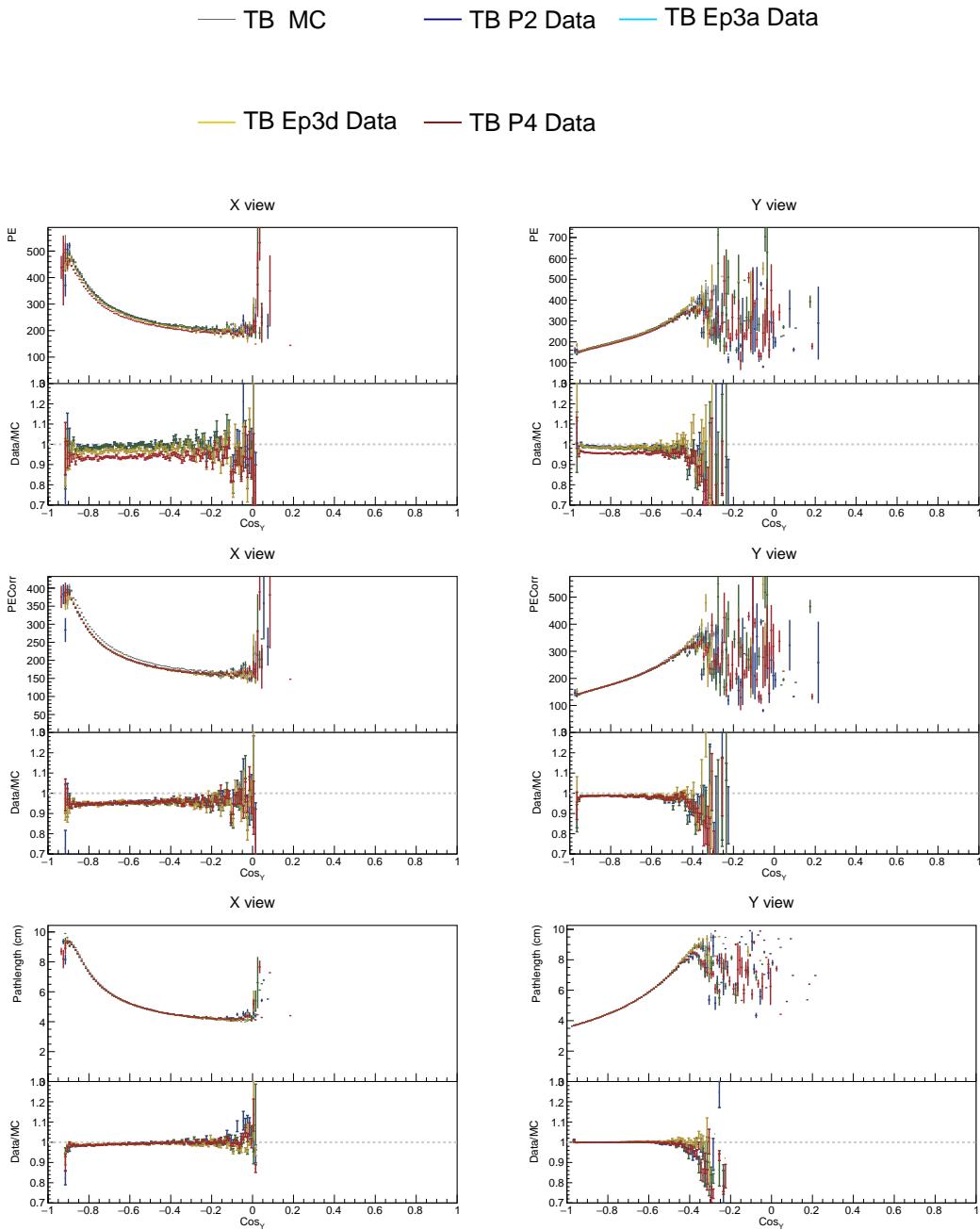


Figure 1.58: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the cosine of the track angle from the Y (vertical) axis.

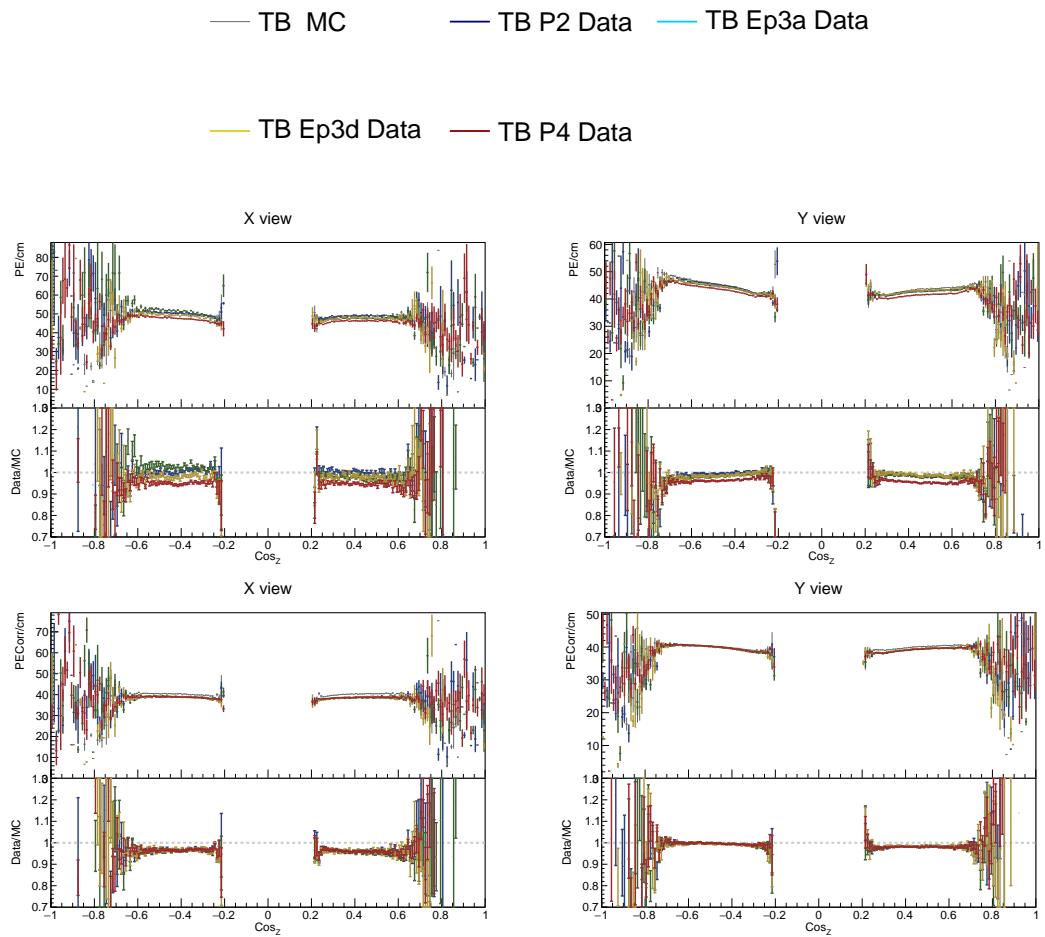


Figure 1.59: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the cosine of the track angle from the Z (beam) axis.

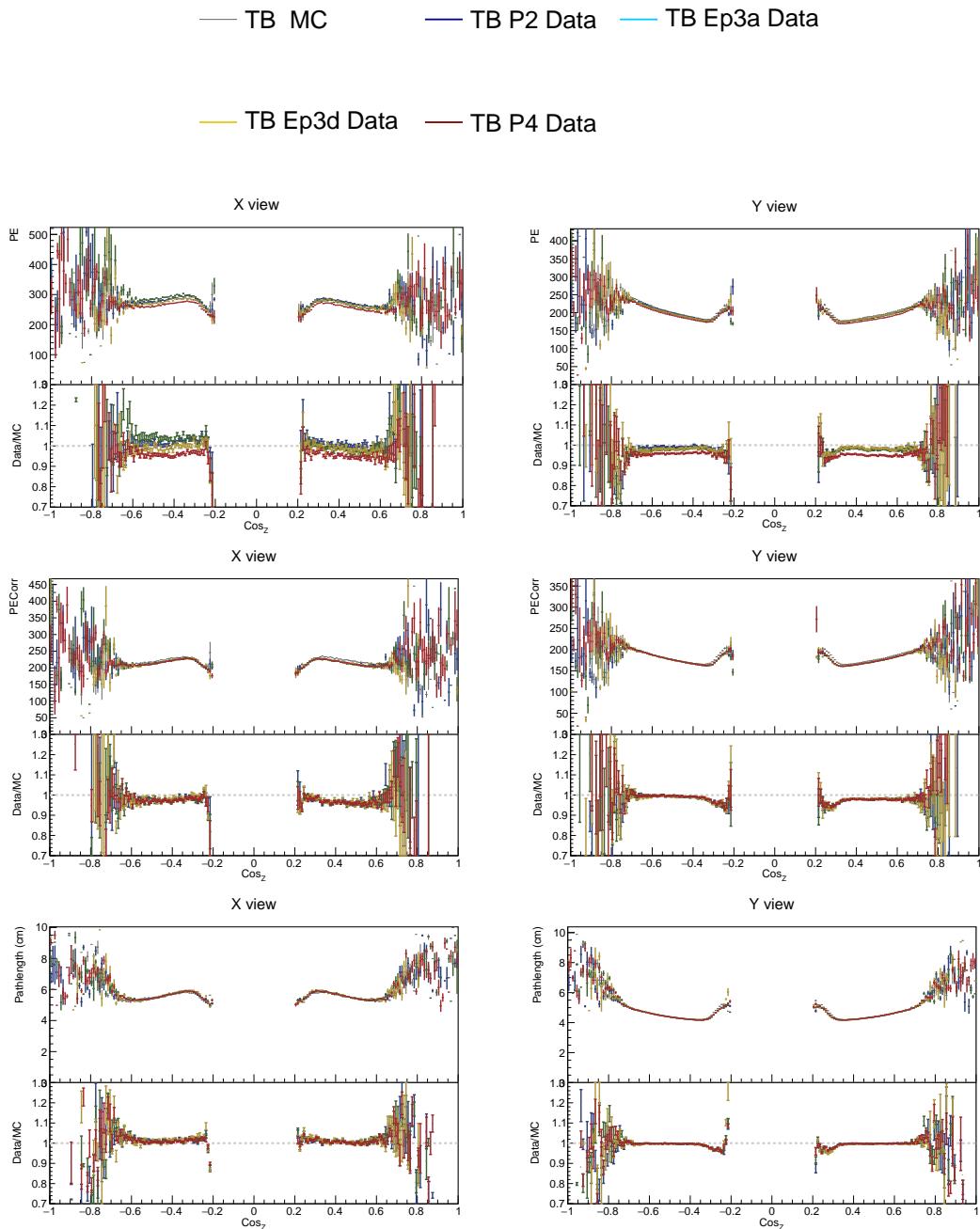


Figure 1.60: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the cosine of the track angle from the Z (beam) axis.

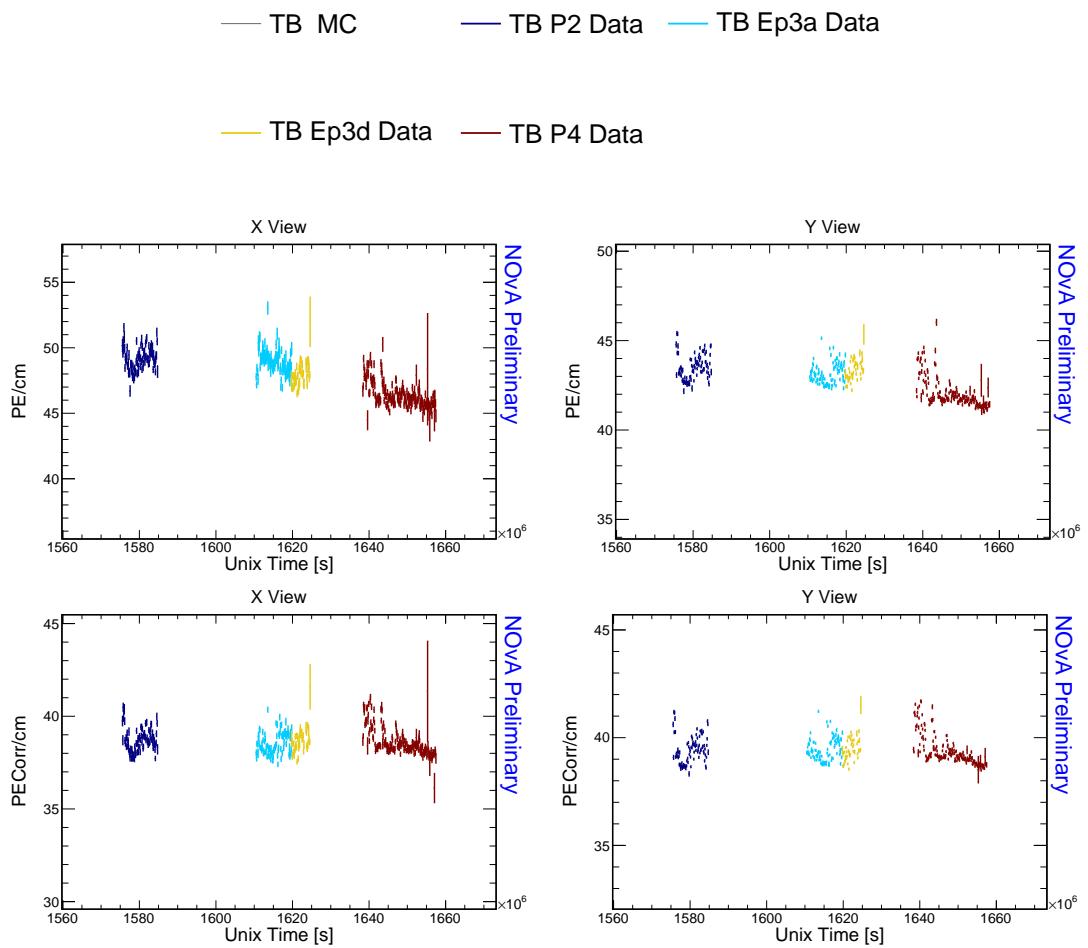


Figure 1.61: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the event UNIX time.

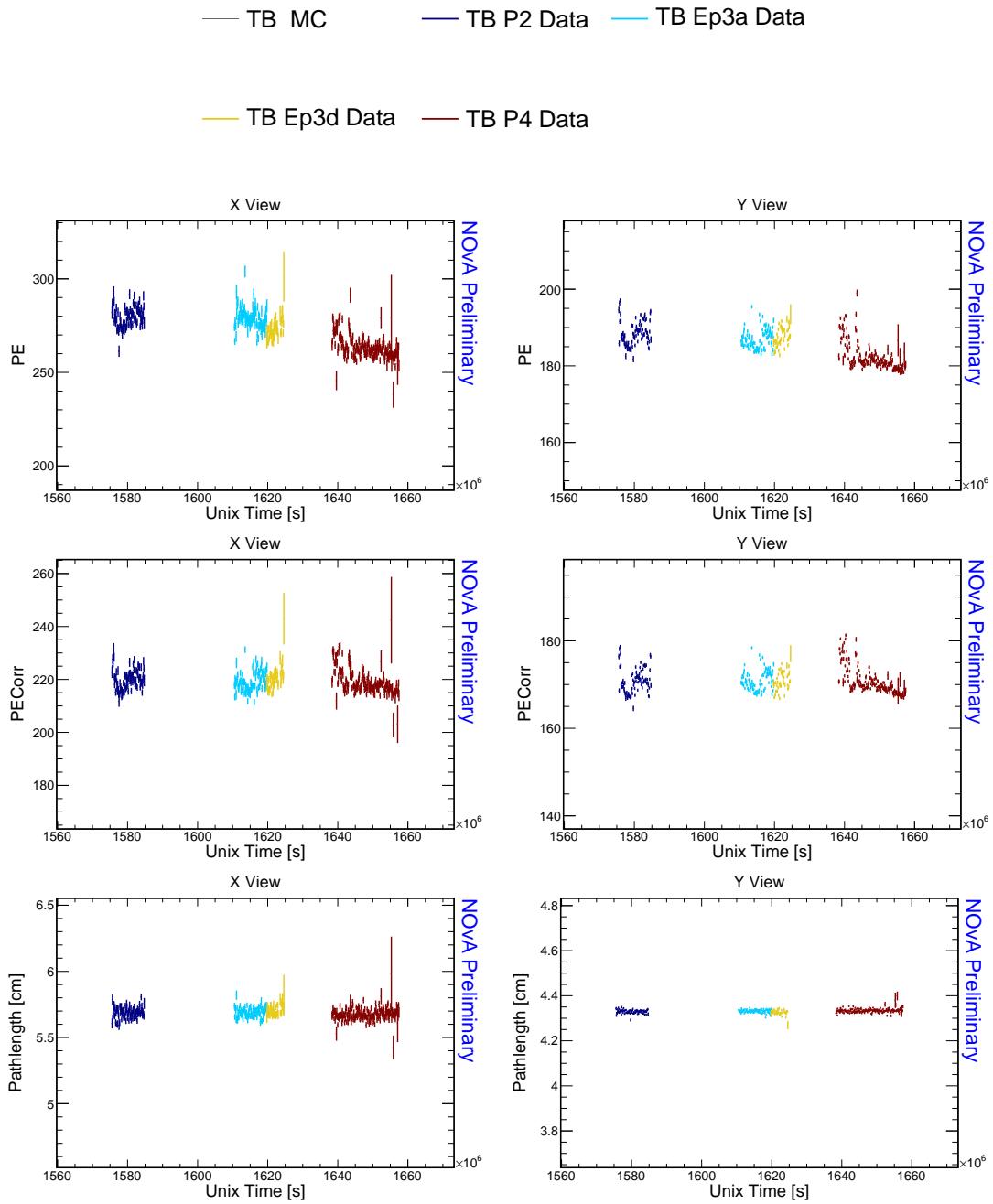


Figure 1.62: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the event UNIX time.

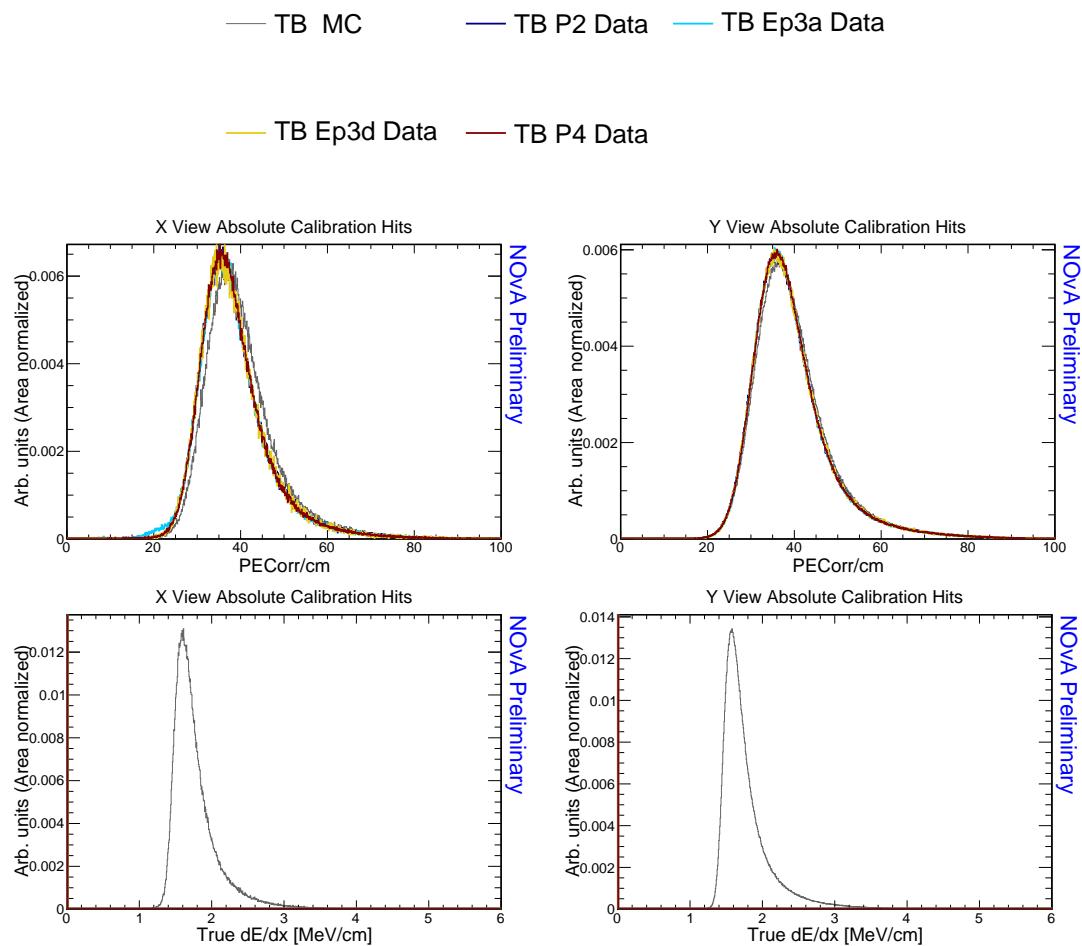


Figure 1.63: Distributions of stopping muons within a 1-2 m track window from the end of their tracks.

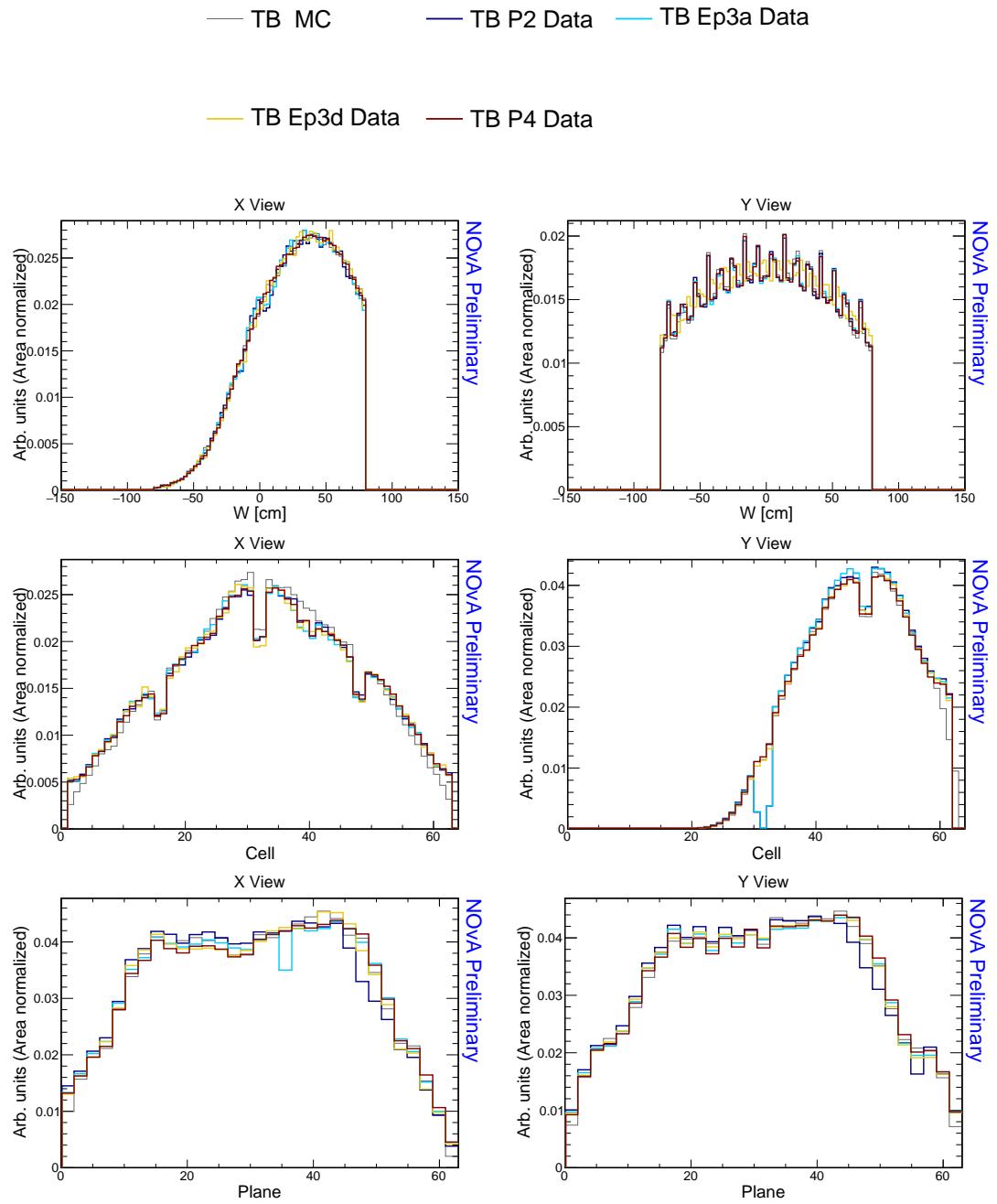


Figure 1.64: Distributions of stopping muons within a 1-2 m track window from the end of their tracks.



# Acronyms

**BPF** Break Point Fitter. [8](#)

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

**FD** Far Detector. [1–6](#)

**FEB** Front End Board. [6, 8](#)

**FERMILAB** Fermi National Accelerator Laboratory. [5, 6](#)

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

**MC** Monte Carlo. [7, 8, 17](#)

**MWPC** Multiwire Proportional Chamber. [2](#)

**ND** Near Detector. [1–7](#)

**NDOS** Near Detector on the Surface. [4–7](#)

**NOvA** NuMI Off-axis  $\nu_e$  Appearance (experiment). [1–5, 7, 8](#)

**NuMI** Neutrinos from the Main Injector. [2](#)

**ToF** Time of Flight. [3](#)

# 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.