

CHAPTER 1

The NOvA experiment

The NuMI Off-axis ν_e Appearance (NOvA) [1] is a long-baseline neutrino oscillation experiment based at the Fermi National Accelerator Laboratory (Fermilab) [2]. NOvA receives an off-axis ν_μ and $\bar{\nu}_\mu$ beam from Fermilab’s Neutrinos from the Main Injector (NuMI) neutrino source (Sec. 1.1) and measures the ν_e or $\bar{\nu}_e$ appearance and the ν_μ or $\bar{\nu}_\mu$ disappearance between its two highly active and finely segmented detectors (Sec. 1.4) [3].

The capability to measure both the ν_e and the $\bar{\nu}_e$ appearance, coupled with a significant matter effect induced by its long baseline, allows NOvA to address some of the most important questions in neutrino physics to date, such as the neutrino mass ordering, the octant of θ_{23} , and the possible Charge conjugation - Parity (CP) symmetry violation in the neutrino sector [3–7]. NOvA data also enables measurements of θ_{13} , θ_{23} and $|\Delta m_{32}^2|$ [3], measurements of neutrino differential cross sections in the Near Detector (ND) [8–11], constraints on the possible sterile neutrino models [12, 13], monitoring for supernova neutrino activity [14, 15], searches for magnetic monopoles [16], or constraints on the neutrino electromagnetic properties (this thesis). Using two functionally identical detectors mitigates the dominant systematic uncertainties of neutrino oscillation measurements, described in Sec. 1.8.

NOvA started taking data in February 2014 and is expected to run through 2026 [17], or until Fermilab begins redirecting its efforts towards the startup of the upcoming Deep Underground Neutrino Experiment (DUNE) experiment [18].

1.1 The Neutrino Beam

The neutrino beam for NOvA comes from the Fermilab-based NuMI neutrino source [19]. The schematic description of NuMI is shown in Fig. 1.1, starting on the left hand side with 120 GeV protons from the Main Injector (MI), part of the Fermilab accel-

erator complex. The proton beam is divided into $10\ \mu\text{s}$ long pulses, with $\sim 5 \times 10^{13}$ Protons On Target (POT) per spill every $\sim 1.3\ \text{s}$ long cycle time, resulting in a proton beam power of $\sim 800\ \text{kW}$ (current record $959\ \text{kW}$ [20]), with upgrades currently underway to surpass $1\ \text{MW}$ [21].

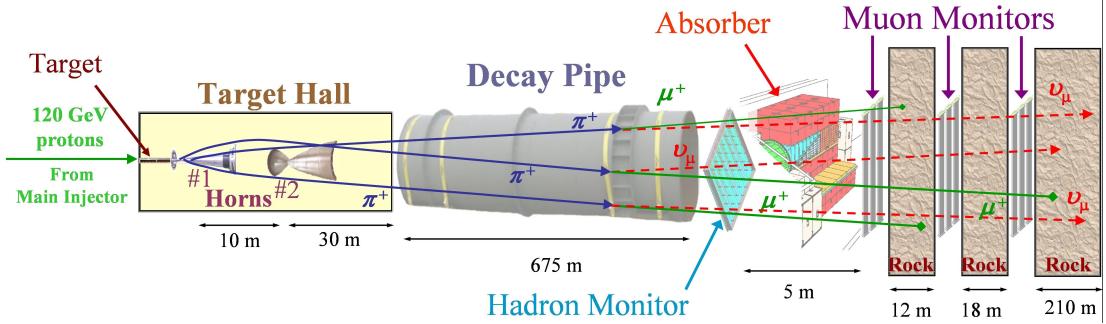


Figure 1.1: The NuMI neutrino beam starts on the left hand side with protons from the MI impinged on a graphite target producing mainly pions and kaons. These are then focused and charge-selected by two focusing horns, after which they decay inside the decay pipe into a high-purity ν_μ or $\bar{\nu}_\mu$ beam. The residual hadrons are stopped and monitored in the hadron absorber and the remaining muons are recorded with muon monitors and absorbed inside the rock. Figure from [19].

The proton beam passes through a collimating baffle before hitting a $\sim 1.2\ \text{m}$ -long (equal to about two interaction lengths) graphite target [22], producing hadrons, predominantly pions and kaons [19]. These are then focused and selected by two parabolic magnetic ‘horns’. The focused hadrons pass through a 675 m-long decay pipe filled with helium to create a low density environment for hadrons to propagate and decay in flight into either neutrinos or antineutrinos. High energy hadrons that do not decay in the decay pipe are absorbed within a massive aluminium, steel, and concrete hadron absorber and monitored with a hadron monitor. The leftover muons are ranged out in dolomite rock after the absorber and monitored using three muon monitors. The hadron and muon monitors are ionization chambers, used to monitor the quality, location and relative intensity of the beam.

Using a positive current inside the horns focuses positively charged particles, which then decay into neutrinos, and removes negatively charged particles. Reversing the horn current focuses negatively charged particles, which decay into antineutrinos, and defocuses positively charged particles. The neutrino mode is therefore called Forward Horn Current (FHC) and the antineutrino mode is called Reverse Horn Current (RHC). The composition of the neutrino beam for both these modes

at the **NOvA ND** is shown in Fig. 1.2, displaying the very high purity of the ν_μ or $\bar{\nu}_\mu$ component in the **FHC** ro **RHC** beam respectively [19].

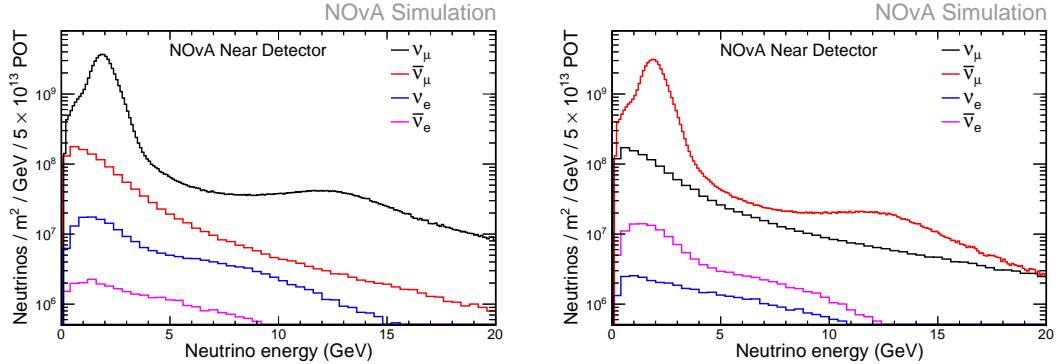


Figure 1.2: The components of the neutrino beam at the **NOvA ND** per one **NuMI** spill in the **FHC** regime shown on the left and the **RHC** regime on the right. The ν_μ ($\bar{\nu}_\mu$) composition in the **FHC** (**RHC**) regime is 93.8% (92.5%), with a wrong sign contribution of 5.3% (6.6%) and only 0.9% (0.9%) contamination by ν_e ($\bar{\nu}_e$), showing the high purity of ν_μ and $\bar{\nu}_\mu$ in the neutrino beam for **NOvA**. Beam composition values calculated for neutrinos with energies between 1 – 5 GeV. Figures are from internal **NOvA** repository [23].

The resulting neutrino beam energy distribution is peaked at ~ 7 GeV with a wide energy band. However, thanks to the kinematics of the dominant pion decay, by placing the **NOvA ND** and Far Detector (FD) 14.6 mrad ($\approx 0.8^\circ$) off the main **NuMI** beam axis, **NOvA** achieves a narrow band neutrino flux peaked at 1.8 GeV [7, 24], as can be seen in Fig. 1.3. Using an off-axis neutrino flux increases the neutrino beam around 2 GeV about 5-fold compared to the on-axis flux and narrow-band peak enhances background rejection for the ν_e appearance analysis [24].

1.2 The **NOvA** Detectors

The two main **NOvA** detectors are the **ND**, located in **Fermilab** ~ 1 km from the **NuMI** target and ~ 100 m under ground, and the **FD**, located ~ 810 km from **Fermilab** at Ash River in north Minnesota, partially underground with a rock overburden [24]. **NOvA** also operated a detector prototype called Near Detector on the Surface (NDOS), which was used for early research and development of detector components and analysis [4]. Additionally, **NOvA** operated a Test Beam detector, described in detail in Sec. ???. The scales of the **ND** and **FD** are shown in Fig. 1.4.

All **NOvA** detectors are highly segmented, highly active, functionally identical

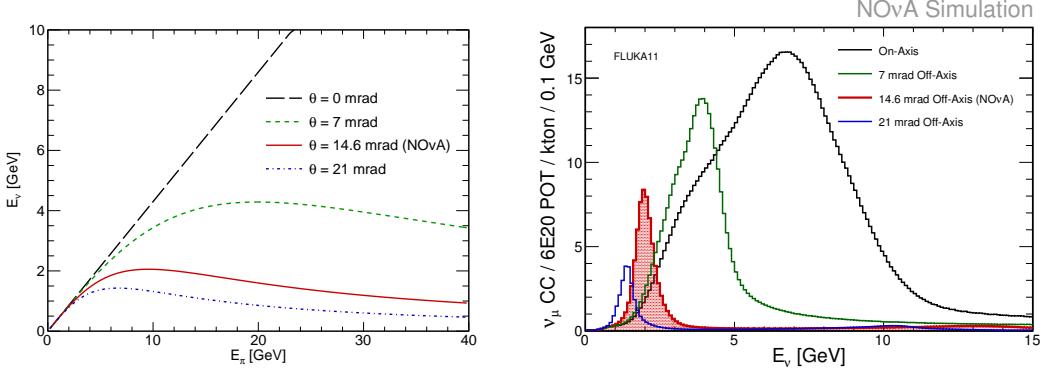


Figure 1.3: (Left) Dependence of the neutrino energy on the parent pion’s energy and (right) neutrino energy distribution for an on-axis beam and three different off-axis beam designs. The case for NOvA is shown in red and results in a narrow neutrino energy distribution around 2 GeV, with limited dependence on the parent pion’s energy. Figure from [24]

tracking calorimeters made up of Polyvinyl chloride (PVC) cells filled with liquid scintillator. Each cell is a long rectangular cuboid with depth of 5.9 cm and width of 3.8 cm (with some variations), with cell length extending to the full width/height of each detector, which is ~ 4.1 m for the ND and ~ 15.6 m for the FD [24]. An example of a FD cell is shown on the right of Fig. 1.4.

Cells are connected side-by-side into a 16 cell-wide extrusions with 3.3 mm-wide walls between cells and 4.9 mm-wide walls on the outsides of the extrusions. The first and last cell of each extrusion are ~ 3 mm narrower than the rest of the cells. Two extrusions are connected side-by-side to form a 32 cell-wide module, with each module having a separate readout (see Sec. 1.3). In the FD, 12 modules are connected side-by-side to form one plane of the detector. In the ND only 3 modules make up a plane. Planes are positioned one after another, alternating between vertical and horizontal orientation, and grouped into diblocks, each containing 64 planes. The FD contains 14 diblocks, totalling 896 planes, whereas the ND contains 3 diblocks totalling 192 planes. The ND also contains a Muon Catcher region, positioned right after the active region, consisting of 22 planes of the normal NOvA detector design, 2 modules high and 3 modules wide, sandwiched with 10 steel plates to help range out muons mainly from the ν_μ charged current interactions [4, 24].

The NOvA coordinate system is centred with $(0, 0, 0)$ in the centre of the first plane, relative to the beam direction. The x axis runs from left to right when facing the detector, y axis from bottom to top and z axis runs perpendicular to the planes

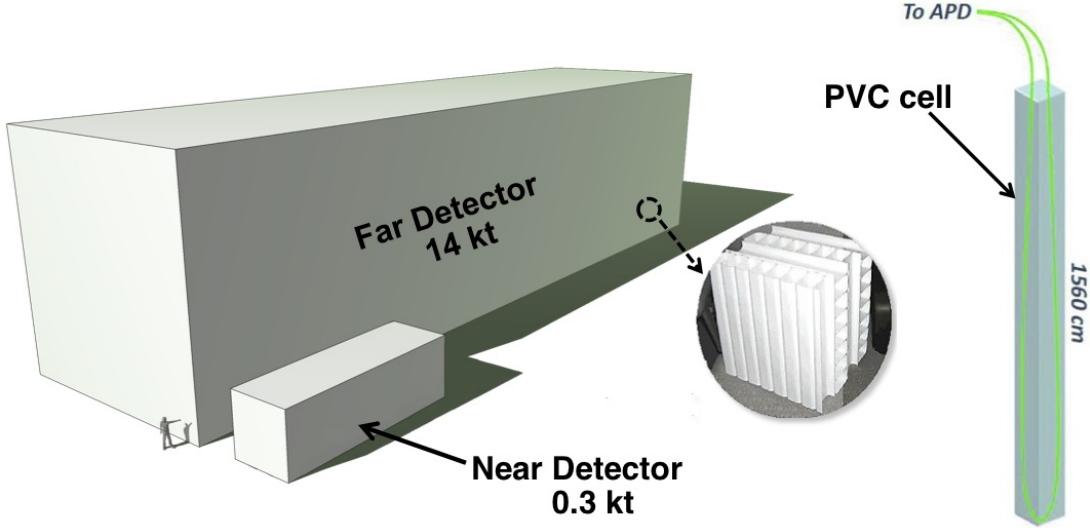


Figure 1.4: Schematic description of the scale and composition of the [NOvA ND](#) and [FD](#). The inset shows a photo of the orthogonal planes made out of [PVC](#) cells. An example of a [FD](#) cell containing liquid scintillator and a looped [WLS](#) fibre attached to an [APD](#) is shown on the right [25].

along the beam direction.

Each cell is filled with a liquid scintillator consisting of mineral oil with 4.1% pseudocumene as the scintillant [26]. Each cell contains a single wavelength shifting fibre with double the length of the cell, looping at one end and connecting to the readout at the other. The [PVC](#) walls of the cells are loaded with highly reflective titanium dioxide, with light typically bouncing off the [PVC](#) walls ~ 8 times before being captured by the Wavelength Shifting (WLS) fibre [24].

The final dimensions of the [FD](#) are $15.6\text{ m} \times 15.6\text{ m} \times 60\text{ m}$ with a total mass of 14 kT and for the [ND](#) the dimensions are $3.8\text{ m} \times 3.8\text{ m} \times 12.8\text{ m}$ with a mass of about 0.3 kT [17]. The active volume, consisting only of the liquid scintillator without the [PVC](#) structure, makes up about 70% of the total detector volume [24].

The [NOvA](#) detectors are specifically designed for electromagnetic shower identification, with a radiation length of 38 cm, which amounts to ~ 7 planes for particles travelling perpendicular to the detector planes [4]. This is particularly useful to distinguish electrons from π^0 s.

We can calculate the minimum energy an electron needs to have to cross one cell (5.9 cm) of the [NOvA](#) detector by using the measured scintillator density 0.86 g/cm^3 [27], which gives us the required range of $\sim 5\text{ g/cm}^2$. Comparing this to measured values for the electron range [28] in the continuous slowing down approximation in

a Polyethylene (approximation of the NOvA scintillator [29]), gives us an estimate of the lowest detectable electron energy as $E_e \gtrsim 10$ MeV.

1.3 Readout and Data Acquisition

The signal from the WLS fibres is read out by an Avalanche Photodiode (APD), converting the scintillation light into electrical signal, with a high quantum efficiency of $\sim 85\%$ and a gain of 100 [24]. An example APD is shown in Fig. 1.5. Both ends of each fibre correspond to a single readout channel and are connected to one of the 32 pixels on the APD, organized in four rows of 8 pixels, with each APD reading out signal from one module. To maximise the signal to noise ratio, the APDs are cooled to -15°C by a thermoelectric cooler, with heat carried away by a water cooling system.

The combination of the APD quantum efficiency and the light yield, determined by the PVC reflectivity and the scintillator and WLS fibre responses, result in a signal requirement of at least 20 Photo Electron (PE) in response to minimum ionizing radiation at the far end of the FD cell.

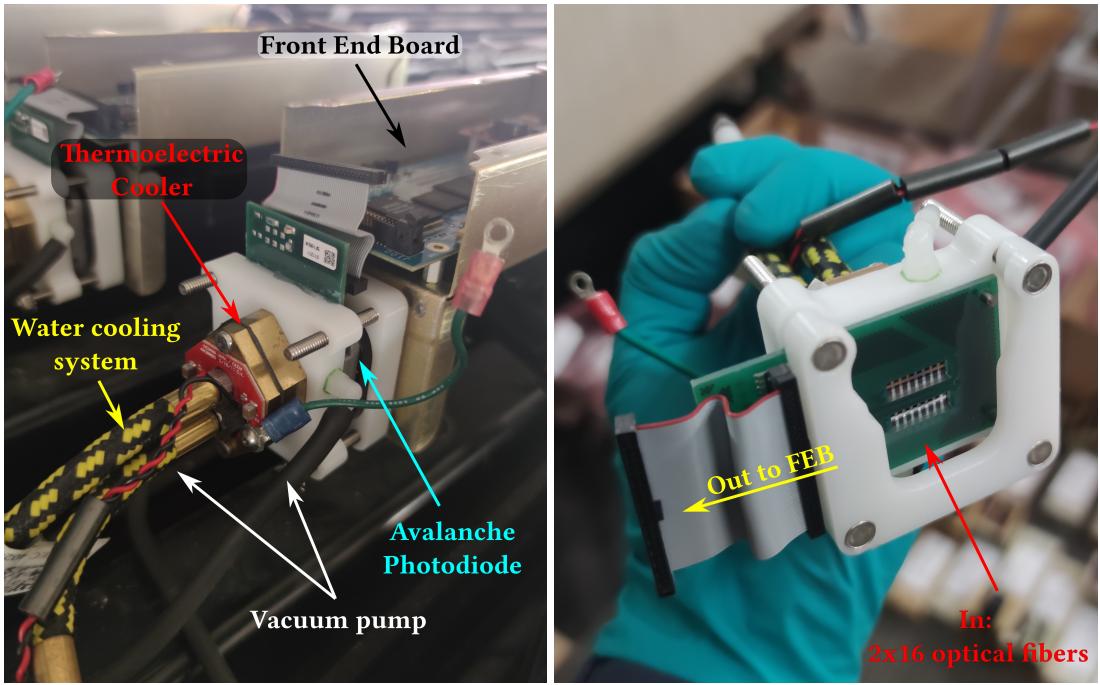


Figure 1.5: The modules with APDs for NOvA mounted on top of the detector on the left picture, and shown from the bottom on the right. The individual components of the module are described. The left picture shows a disconnected ribbon cable and ground cable, which are normally connected to the front end board.

Each APD is connected to a single Front End Board (FEB), shown in Fig. 1.6. The

FEB amplifies and integrates the **APD** signal, determines its amplitude and arrival time, before passing it to the Data Acquisition (DAQ) system. On the **FEB** the **APD** signal is first passed to a custom **NOvA** Application-Specific Integrated Circuit (ASIC), which is designed to maximize the detector sensitivity to small signals. **ASICs** amplify, shape and combine the signal, before sending it to an Analog-to-Digital Converter (ADC). The combined noise from the **APD** and the amplifier is equivalent to about 4 **PEs**, which, compared to an average **PE** yield from the far end of the **FD** cell of 30, results in a good signal and noise separation [24]. The digitized data from an **ADC** is sent to a Field Programmable Gate Array (FPGA), which extracts the time and amplitude of the **ADC** signals, while subtracting noise based on a settable threshold. The **FPGAs** employ multiple correlated sampling methods to reduce noise and improve time resolution of the signal [30].

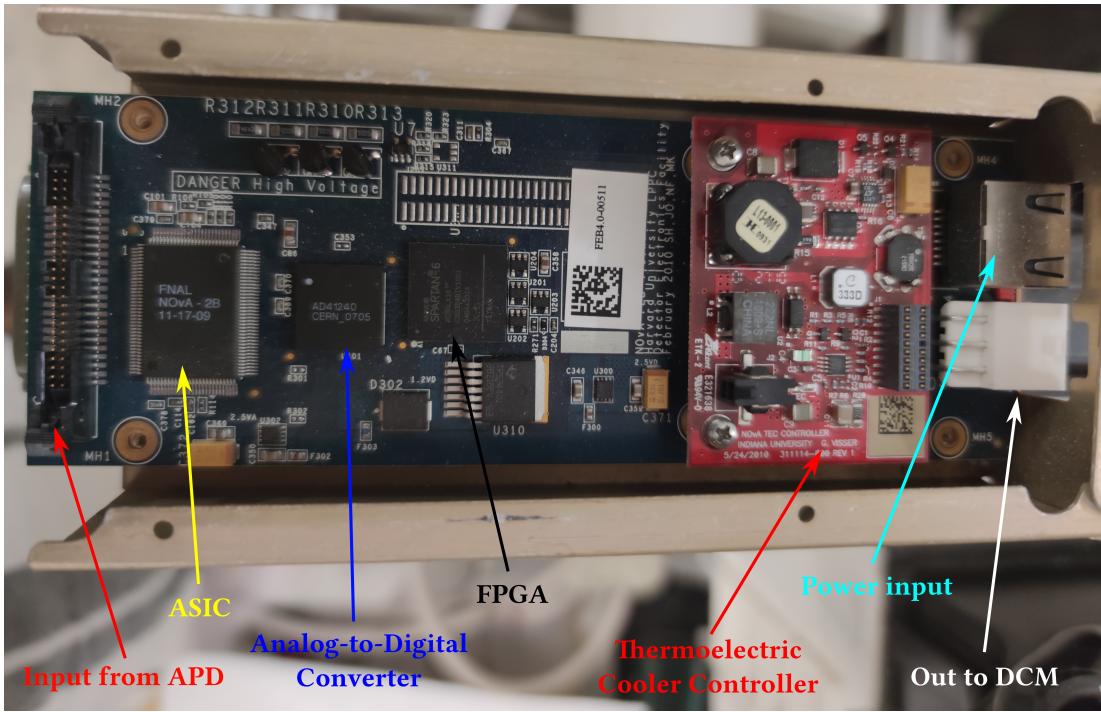


Figure 1.6: An example of a **NOvA** FEB with individual components labelled.

All of the **NOvA** front end electronics (**APDs** and **FEBs**) are operated in a continuous readout mode, without requiring any external triggers [24]. Due to higher detector activity during beam spills, the **ND FEBs** work at a higher frequency of 8 MHz, whereas the **FD FEBs** suffice with 2 MHz sampling frequency [30].

Data from up to 64 **FEBs** are concentrated in a Data Concentration Module (DCM), which concatenates and packages the data into 5 ms time slices, before sending it to

the buffer nodes. [DCMs](#) are also connected to the timing system and pass a single unified timing measurement to the [FEBs](#) to maintain synchronization across the detector [30].

The buffer nodes cache the data for at least 20 seconds while receiving information from the trigger system. Each trigger uses a time window based either on the time of the [NuMI](#) beam spill, on a periodic interval for monitoring and for the readout of cosmic events, or on one of the activity-based data-driven triggers [30]. Data that fall within any of the trigger windows are sent to a data logger system, where they are merged to form events, before being written to files for offline processing and sent to an online monitoring system. Files are organized based on a unique combination of run and subrun numbers, with runs corresponding to data taking periods with constant detector conditions and limited to either 64 subruns or 24 hours. Subruns are delineated either by a 2 GB file size constraint or a 1-hour timeout limit [31].

The detectors are continuously monitored to ensure data stability and quality. Subruns with suboptimal detector conditions or with events failing basic quality criteria are flagged as ‘bad’ and recorded in a ‘bad runs list’ [31]. Additionally, individual readout channel are assessed on a per-subrun basis, with those with too high or too low hit rates marked as ‘bad’ [32]. Both the ‘good runs’ list and the ‘bad channel’ maps are used to inform event processing and during simulation to emulate real detector conditions.

1.4 Simulation

To extract neutrino oscillation parameters, or to test a hypothesis, [NOvA](#) uses a series of simulations to make predictions according to various physical models [33]. The simulation chain can be divided into four parts: simulation of the neutrino beam, simulation of neutrino interactions within the [NOvA](#) detectors, simulation of cosmic particles interacting in the [NOvA](#) detector and simulation of the detector and readout response.

To simulate the neutrino beam, [NOvA](#) uses a simulation based on the GEANT4 v9.2.p03 [34] Monte Carlo (MC) event generator with a detailed model of the [NuMI](#) beamline [35], as described in Sec. 1.1. The simulation starts with the 120 GeV/c

MI protons interacting within the long carbon target and producing hadrons, mainly π , K and secondary protons. This is followed by transport and possible further interaction of hadrons within the focusing system, until finally ending with hadron decays producing the neutrino beam.

To account for the inherently imprecise theoretical models used in GEANT4, **NOvA** uses the Package to Predict the Flux (**PPFX**) to incorporate external measurements of yields and cross sections of hadron interactions inside the target and the other **NuMI** materials into the neutrino beam prediction [36]. The current version of **PPFX** is limited by the results available during its creation and only corrects the most frequent interactions while assigning conservative systematic uncertainties to the rest (see Sec. 1.8). For the most common π , K and p production, **PPFX** uses the NA49 measurements [37–39] of 158 GeV/c protons interacting on a thin (few percent of interaction length) carbon target. To expand the kinematic coverage, **PPFX** uses a few data points from Barton et al [40] for the π production and K/π ratios from the Main Injector Particle Production (MIPP) [41] experiment for the production of K . These results have to be scaled to the 20 – 120 GeV/c incident proton moment seen throughout **NuMI** using the FLUKA [42, 43] **MC** generator.

There are two new experiments that measure the production and interaction of hadrons on various targets and incident energies, specifically designed to improve the prediction of neutrino beams. The most impactful measurements from the NA61 experiment are of the 120 GeV/c protons on a thin carbon target [44–46], of the hadron incident interactions [47], and of the 120 GeV/c protons on a **NuMI** replica target [48]. The Fermilab-based EMPHATIC experiment [49] is currently analysing a broad range of hadron production and secondary and tertiary interaction measurements for neutrino beam prediction with a significant involvement of **NOvA** and **DUNE** collaborators.

The output of the neutrino beam simulation is passed to the simulation of neutrino interactions inside the detectors, which is done with the GENIE v3.0.6 [50] neutrino **MC** generator. GENIE allows users to choose the particular models for different types of neutrino interactions and particle propagation within the nucleus, as well as possible tunes to external measurements. The four main interaction modes in GENIE are the Quasi-Elastic (QE) Charged Current (CC) scattering, the Resonant baryon pro-

duction (Res), the Deep Inelastic Scattering (DIS), and the Coherent π (COH π). The special case of the two particle - two hole (2p2h) interaction via Meson Exchange Current (MEC) and the Final State Interaction (FSI) inside the nucleus are also considered. The initial state of the nucleus is represented by a local Fermi gas in the QE and 2p2h models, while a global relativistic Fermi gas is used for all other processes. All of these are set by the Comprehensive Model Configuration (CMC), which is currently N1810j0000 for NOvA. Additionally, NOvA adds a costume tune to the NOvA ν_μ CC data for a better constraint of the CCMEC interactions. NOvA also uses a set of external π interaction measurements to constrain the FSI model. Table 1.1 shows the list of models and tunes for different interaction modes in NOvA [7].

Table 1.1: Models and tunes used in the NOvA simulation of neutrino interactions.

Interaction	Model	Tune
CCQE	València [51]	External $\nu - D$ data [52]
CCMEC	València [53, 54]	NOvA ν_μ CC data
Res & COH π	Berger-Sehgal [55, 56]	External $\nu - A$ data
DIS	Bodek-Yang [57, 58]	External $\nu - A$ data
FSI	Semi-classical cascade [59]	External $\pi - {}^{12}C$ data

Since the FD is on the surface NOvA also uses a simulation of cosmic rays generated with the MC Cosmic-Ray Shower Generator (CRY) [60]. The simulated cosmic muons are also used to calibrate NOvA detectors [36].

Particles that are created from neutrino interactions and cosmic rays are propagated through the NOvA detectors using the GEANT4 v10.4.p02 [34], which outputs the energy deposited in the scintillator. This is then passed to a custom NOvA software of the light model [36], which calculates the amount of scintillation light produced for the deposited energy based on a Poisson distribution. The scintillation light production is parametrized using the Birks-Chou model [61], which corrects for the recombination in organic scintillators at high deposited energies. The scintillator light yield and the inherent production of the Cherenkov light, which can affect the light readout, are tuned to NOvA data [9]. The light collection by the WLS fibres, its transport to the APDs, and the APD response use a parametrized simulation, as the NOvA cells and their readout are generally the same across the detectors [36]. The simulation of the readout electronics is done by another custom NOvA parametrized model, which accounts for random noise in the readout electronics and outputs true

events in the same format as the real data.

Due to the high neutrino rate in the **ND**, there are neutrinos interacting in the surrounding rock creating particles, mainly muons, that make it to the detector and act as background. However, since only a few ‘rock muons’ make it into the detector, it would be very time consuming to run a simulation which includes the rock around the **ND** for every neutrino. Instead, **NOvA** creates a separate simulation that includes the surrounding rock and then overlays these results into the nominal **NOvA** simulation chain to match the **NuMI** neutrino rate [36].

1.5 Data Processing and Event Reconstruction

Both data and simulation events for all **NOvA** detectors are passed through the same event reconstruction and particle identification algorithms. The reconstruction was specifically developed with the ν_e appearance search in mind, focusing on identifying the ν_e **CC** signal against the ν_μ **CC** and Neutral Current (NC) backgrounds. Each **NOvA** detector has to deal with different challenges, with multiple neutrinos interacting during one beam spill in the **ND**, and a large cosmic background in the **FD** [62].

The output from the **DAQ** system for each channel is called a *raw hit*. Hits are grouped into 550 μs -long windows and passed to an offline reconstruction chain [62]. Reconstruction starts by grouping hits into *slices* based on their proximity to other hits in both time and space [63]. Slices are designed to ideally contain only a single neutrino interaction event.

For events that produce hadronic and electromagnetic showers, reconstruction first identifies straight lines through major features using a modified Hough transform [64], representing particle directions. These lines are passed to the Elastic Arms algorithm [65] to identify *vertex* candidates from their intersection points. Hits are then clustered into *prongs*, which are collections of hits with a start point, based on the vertex, and a direction, using a k-means algorithm called FuzzyK [66, 67]. Here ‘fuzzy’ means that each hit can belong to multiple prongs. Prongs are first created separately for each view (also called 2D prongs) and then, if possible, view-matched into 3D prongs (from here on referred to as prongs) [62]. Figure 1.7 shows an example of a simulated electron shower, where the reconstructed vertex is shown as a red

cross and the prong as a red shaded area. The prong groups together all the hits that are part of the shower, while removing the background hits, shown in grey.

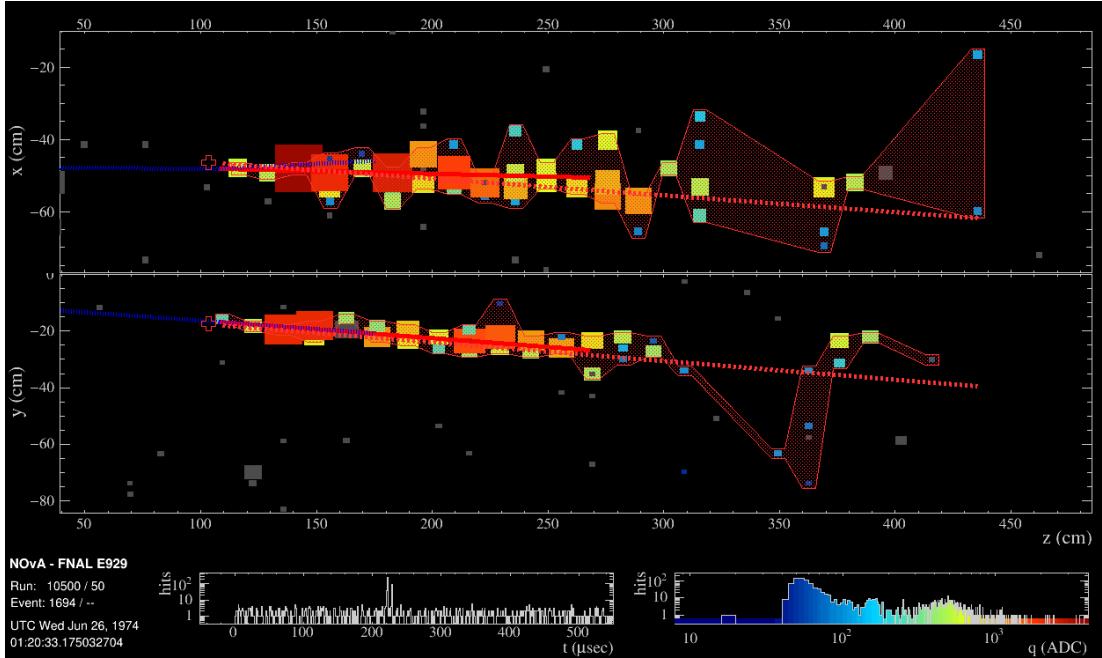


Figure 1.7: Reconstruction of a simulated single electron event in the NOvA ND. The red cross is the reconstructed vertex, the shaded area shows the cluster of hits into a prong and the dotted red line shows the estimated shower direction. The blue dotted line shows the true direct of the scattering neutrino and the solid red line the true momentum of the scattered electron. Figure from internal NOvA database [68].

For particles that are represented by tracks rather than showers (especially muons), the reconstruction takes the slice hits and forms ‘Kalman tracks’ based on a Kalman filter [69]. In addition to the start point and the direction, which exist also for prongs, tracks also contain information on the vector of trajectory points that make up the track and on the end point - and therefore on the track length. A parallel tracking algorithm takes in the Elastic Arms vertex and the Fuzzy-K prongs and forms Break Point Fitter (BPF) tracks [70, 71], using a model of Coulomb scattering and energy loss. BPF tracks also contain an information on the particle 4-momenta based on various particle assumptions, most notably the muon assumption. For cosmic particles, mostly muons, NOvA uses another track reconstruction algorithm, called ‘window cosmic track’ [72]. It uses a sliding 5 plane-long window, in which it fits a straight line to the recorded hits. The window starts from the end of the detector and then slides forward and repeats the fitting process until all hits are processed. This way it accounts for possible Coulomb scattering of cosmic muons. The intersection of each

cosmic track with the edge of the detector (or extrapolation of the track to the edge of the detector) is reconstructed as the ‘cosmic ray vertex’.

To identify individual particles and remove backgrounds, NOvA uses several Machine Learning (ML) algorithms, outputs of which are used in combination with the information from classical reconstruction algorithms for Particle Identification (PID). The most common topologies for particles interacting in NOvA detectors are shown in Fig. 1.8. Muons are easily identifiable as single long tracks which decay into an electron (or positron) if stopping inside of the detector. Both electrons and π^0 's produce electromagnetic showers, but thanks to the low-Z composition and high granularity of the detector, there is a gap between the interaction vertex and the electromagnetic shower for the π^0 .

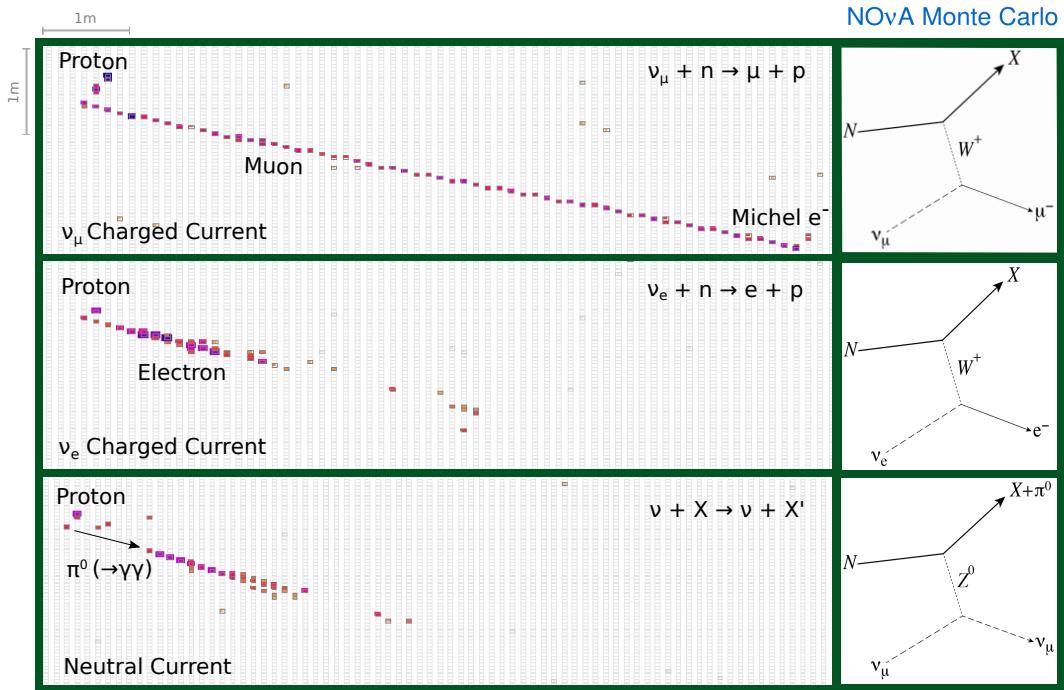


Figure 1.8: Different event topologies as seen in the NOvA detectors with corresponding Feynman diagrams [62]. Each event is a simulated 2.15 GeV neutrino interacting in a NOvA detector producing a 0.78 GeV proton and a second 1.86 GeV particle depending on the interactions type. The figure shows only one view and the colouring represents the deposited energy.

One of the ML algorithms that NOvA employs is a Convolutional Neural Network (CNN) based on the GoogLeNet [73] architecture named Convolutional Visual Network (CVN) [74]. When it is applied to identify entire events it is called *EventCVN* and uses slice hits to classify interactions into one of the five categories: ν_e , ν_μ , ν_τ ,

NC, or cosmic. The same architecture, but applied to the Fuzzy-K prongs, is called *ProngCVN* [75], and is used to identify what particles the prongs most likely correspond to. This assignment is useful in calculation of prong energy, as described in Sec. 1.7. Another **ML** algorithm is specifically designed for identifying muons and is based on a Boosted Decision Tree (BDT). It is called Reconstructed Muon Identifier (ReMId) [69] and uses the reconstructed Kalman tracks as inputs.

1.6 Detector Calibration

The energy deposited within **NOvA** detectors is represented by the peak **ADC** values for each cell the particle passed through, obtained from the readout electronics, as described in Sec. 1.3. The conversion of the peak **ADC** values into physical units of energy requires calibrating the **NOvA** detectors [76], while accounting for the attenuation of light along the **WLS** fibres, or for differences between individual cells. The purpose of calibration is to calculate a conversion factor from $\text{ADC} \rightarrow \text{MeV}$ for every part of the detector, so that the same energy deposited anywhere and at any time, is recorded as the same value of the reconstructed energy.

NOvA uses cosmic ray muons for calibration due to their abundance in the **NOvA** detectors and their consistent energy deposition. To calculate the absolute energy scale, **NOvA** selects a subsample of muons stopping inside of the detectors when they are almost exactly Minimum Ionising Particle (MIP) and therefore have a well understood energy deposition. The cosmic muons are collected using a periodic trigger with the same length as the beam trigger, whilst removing events with timestamps overlapping with the beam spill window. The simulation of cosmic muons is created using the **CRY** [60] **MC** generator, as outlined in Sec. 1.4.

Cosmic muon tracks are reconstructed using the window cosmic track algorithm described in Sec. 1.5. The selection of well reconstructed cosmic tracks requires that at least 80% of all hits from the reconstructed slice contribute to the track [29]. Each track must have at least 2 hits in both the x and y views and the difference in the number of planes the track crossed between the views must be at most 10% of the total number of planes. Also, the plane where each track starts or stops in one view must be within 3 planes of the start or stop plane in the other view. Additionally,

since tracks that do not cross many planes tend to not be reconstructed very well, the extent of each track in the z direction must be at least 70 cm and tracks must have at least 20% of their total track direction in the z axis. Tracks with on average more than 6 cells per plane and with path lengths through the cell larger than 10 cm are removed for the same reason. Furthermore, all the reconstructed tracks must start at most 10 cm from the edge inside of the detector and stop at most 10 cm outside of the detector. Lastly, tracks with trajectory points far away from each other are also removed. The selection of stopping muons for the absolute energy scale relies on identifying Michel electrons, which are produced by decaying muons at the end of their tracks, as can be seen on the top panel of Fig. 1.8.

Since the energy deposited in a cell is proportional to the distance the particle travels through the cell, the input variable for calibration is the deposited energy divided by the path length through the cell PE/cm . To ensure the path length is well calculated, all hits used in calibration must satisfy the so-called ‘tricell’ condition, shown in Fig. 1.9. This means that for each calibration hit, there must be a corresponding hit in both of the surrounding cells in the same plane for the same track. The path length can then be calculated simply from the height of the cell and the angle of the reconstructed track. In case there is a bad channel in a neighbouring cell (right side of Fig. 1.9), this channel is ignored and the tricell condition looks one cell further [76]. If the tricell condition fails, the hit can still pass the ‘z tricell’ condition, which is a longitudinal equivalent of the tricell condition and requires a hit in both the neighbouring planes in the same view and with the same cell number. The ‘z tricell’ hits are saved separately and may be used if there are no hits satisfying the original tricell condition. This is especially useful for the cells on the edge of the detector, which fail the tricell condition due to only having one neighbouring cell.

The calibration conversion factor from the signal recorded by the detector readout to the deposited energy can be expressed by as

$$E_{dep} \text{ [MeV]} = \text{Signal [ADC]} \times S_d \times TS_{d,i}^{\text{CALIB}} \times R_{d,i}(t) \times A_d(t). \quad (1.1)$$

The calibration scale therefore consists of four separate and complementary factors: the Scale (S_d), the Threshold and Shielding correction ($TS_{d,i}$), the Relative calibration ($R_{d,i}(t)$) and the Absolute calibration ($A_d(t)$), all described below. Each part is

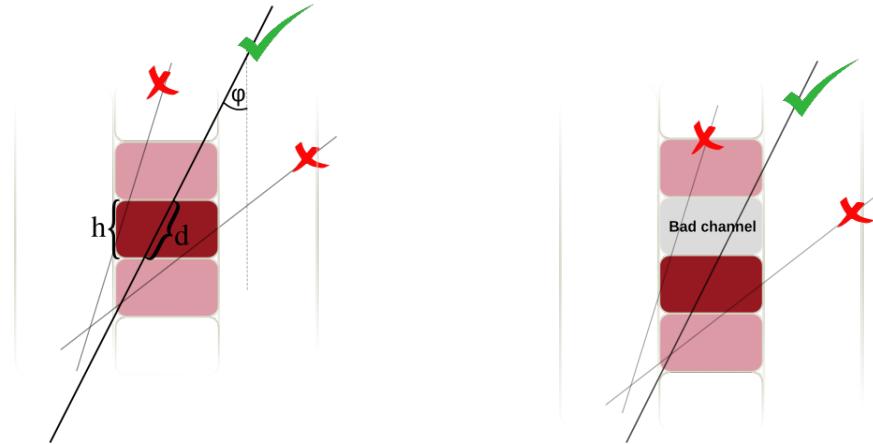


Figure 1.9: Illustration of the tricell condition. Only the hits with two surrounding hits in the same plane are used in the NOvA calibration, as shown on the left plot. This is to ensure a good quality of the path length (d) reconstruction, which is calculated from the known cell height (h) and the reconstructed track angle (φ). In case the hit is next to a bad channel, as shown on the right plot, the bad channel is ignored and the tricell condition requires a hit in the next cell over.

calculated for each detector separately, as indicated by the subscript d . The threshold and shielding correction is only used during calibration and is omitted when applying the calibration results. The relative and absolute calibrations are calculated for each time period separately to account for possible changes in the energy deposition throughout the time, possibly caused by the ageing of the scintillator oil, or of the readout electronics. The time periods are either determined by a fixed time interval, or by running conditions separated by significant changes to the readout or the DAQ systems, including the summer shutdown.

The threshold and shielding correction and the relative calibration calculate a calibration factor for each position within the detector to account for variations caused by the attenuation of light as it travels through the WLS fibres, or by differences between individual cells. This is expressed with a subscript i in Eq. 1.1. For data, the position of a hit in the detector is described by the plane number, cell number and the position within the cell (w). w is calculated as the projection of the cosmic track to the central cell axis and its value is equivalent to the x axis (y axis) coordinate of the projection for the horizontal (vertical) cells, with the 0 value at the centre of the cell [76].

For simulation, the calibration does not use the plane number to determine the position within a detector, as by construction all detector planes should have the same

readout. This significantly reduces the requirements for the number of events that need to be simulated, reconstructed, and calibrated, especially for the [FD](#) with 896 planes. However, in reality there are some variations in the detector response between individual planes, caused by different *brightness* qualities of the fibres, zipped or twisted fibres, different qualities of the scintillator, possible air bubbles, and potentially other factors. To include these differences in simulation without having to simulate every cell individually, all the cells are divided into 12 equally populated Fibre Brightness (FB) bins based on the uncorrected average response in the center of that cell, as shown in Fig. 1.10. These FB bins describe the relative differences in the detector response between individual cells [77].

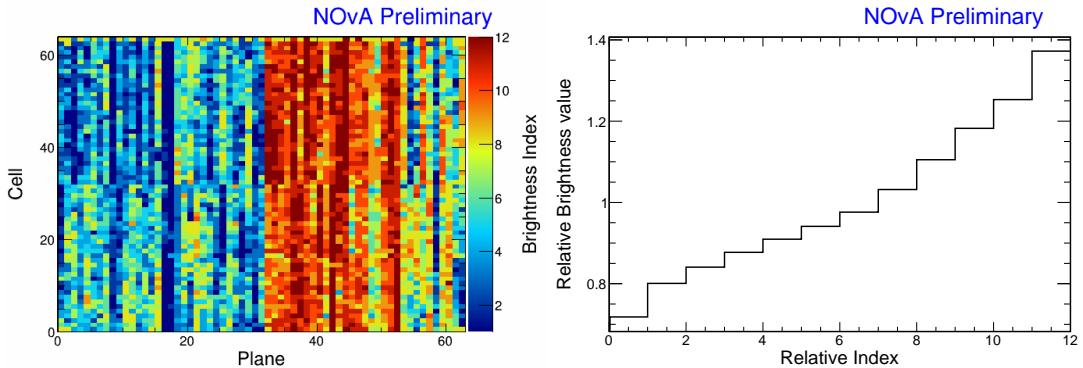


Figure 1.10: Distribution of the [NOvA](#) detector cells into 12 brightness bins (left plot), each representing a relative difference in energy response (right plot) due to different brightnesses of the fibres, scintillators, or readout. This is an example from the [NOvA](#) Test Beam detector, described in Sec. 2.1, where the left side of the detector (planes 1–32) has clearly lower response relative to the right side of the detector (planes 33–64).

1.6.1 Scale

The scale calibration factor from Eq. 1.1 is a simple conversion from the peak [ADC](#) value into the number of [PEs](#). This factor only depends on the [APD](#) gain (which was different in the beginning of [NOvA](#) data taking) and on the [FEB](#) type (different between detectors, as described in Sec. 1.3).

1.6.2 Threshold and Shielding Correction

The threshold and shielding correction accounts for two assumptions, which hold true in most cases in [NOvA](#), but fall short for some hits at the bottom of the detector, or far away from the readout, especially for the [FD](#) [76].

The first assumption is that the **ADC** response to the photon signal is linear, which is mostly true except close to the **APD** threshold. Energy deposited far away from the readout may produce photons that get attenuated enough to be shifted below the threshold. However, due to natural fluctuations of the number of photons created by the energy deposition, the same deposited energy may also produce photons that would make it over the threshold, therefore making it appear that the actual deposited energy was higher than in reality, introducing a bias to the calibration. The threshold correction is calculated using simulation, as the ratio between the mean of the Poisson distribution of the true number of the created **PE** ($\text{PE}_{\text{Poisson}\lambda}$) and the number of the ‘reconstructed’ **PE** seen by the **APD** (PE_{Reco}).

The second assumption is that the spectrum of cosmic muons is uniform within each detector. Again, this is generally true, but breaks down in the **FD**, which is big enough for the top of the detector to shield the bottom of the detector and therefore affect the energy distribution. The shielding correction is calculated from simulation as a ratio between the expected deposited energy if the particle was a **MIP** (E_{MIP}), which is estimated from simulation for the **NOvA** scintillator as $E_{\text{MIP}} = 1.78 \text{ MeV/cm}$ and the true deposited energy (E_{true}).

The total threshold and shielding correction is calculated for simulated events in each cell, **FB** bin and w as

$$TS_i = \frac{\text{PE}_{\text{Poisson}\lambda}}{\text{PE}_{\text{Reco}}} \frac{E_{\text{MIP}}}{E_{\text{true}}}. \quad (1.2)$$

To ensure that the correction changes smoothly across each cell position, the final correction is calculated as a fit to the mean correction value along w in each cell and **FB** bin.

1.6.3 Relative Calibration

The main goal of the relative calibration is to correct for the attenuation of the scintillator light as it travels through the **WLS** fibre to the readout. The attenuation in each cell is estimated by performing an ‘attenuation fit’ to the mean response in **PE/cm**, as shown in Fig. 1.11. The relative calibration scale is then calculated as the ratio between the average response in **PE/cm** across the entire detector (can differ between

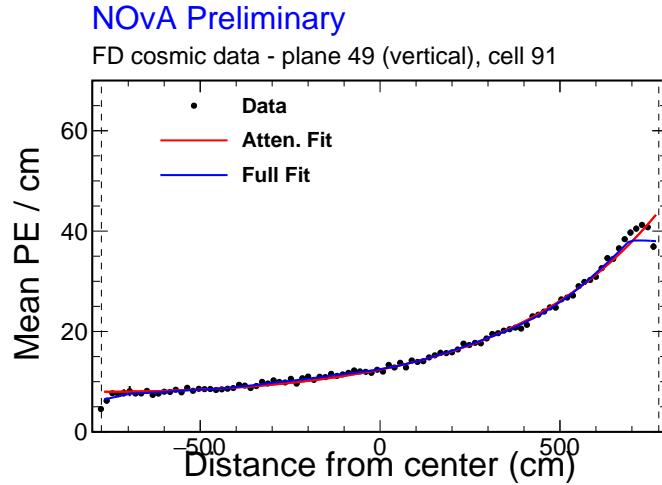


Figure 1.11: Example attenuation fit for a single cell in the NOvA FD across its full length, as shown by dashed vertical lines. The red line shows the initial exponential fit and the blue line shows the full fit after the LOWESS correction, both described in text. Figure from [78].

detectors) and the result of the attenuation fit in each particular position within the detector. The response after applying the relative calibration scale is expressed as Corrected Photo Electronss (PECorrs). Since the relative calibration scale is calculated for each cell independently, it effectively corrects for the relative differences between detector cells as well as for the attenuation. Therefore, the resulting distribution of PECorr/cm should be uniform across the detector, especially along the plane, cell and w [76].

The first step to do the attenuation fit is to create ‘attenuation profiles’ for each cell. Attenuation profiles are profile histograms of mean detector response over the path length through the cell, in the units of PE/cm , along the position within the cell. An example attenuation profile is shown in Fig. 1.11 as black dots. The threshold and shielding correction described in Sec. 1.6.2 is applied to the attenuation profiles before doing the attenuation fit, which consists of two steps.

1. The first step is a three-parameter exponential fit according to

$$y = C + A \left(\exp \left(\frac{w}{X} \right) + \exp \left(-\frac{L+w}{X} \right) \right), \quad (1.3)$$

where y is the fitted response, L is the length of the cell and C , A and X are the fitted parameters representing the background, attenuation scale and atten-

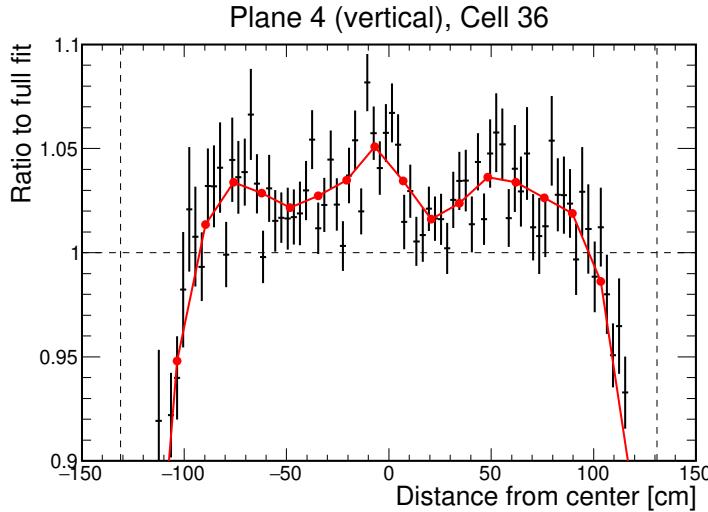


Figure 1.12: Example LOWESS correction for the residual differences after the exponential part of the attenuation fit of the NOvA relative calibration. This is an example for a single cell in the NOvA Test Beam detector with black points showing the residual differences and red line the LOWESS correction, both described in text.

uation length respectively. An example of the exponential fit is shown as a red curve in Fig. 1.11.

2. The second step is the smoothing out of residual differences between the exponential fit and the original distribution with the Locally Weighted Scatter plot Smoothing (LOWESS) method, shown in Fig. 1.12. The residual differences get evened out by creating a smooth distribution of 20 locally weighted points across the length of each cell. The result of the LOWESS correction is then combined with the exponential fit into the full attenuation fit, shown as a blue line in Fig. 1.11.

Even after applying the LOWESS correction, there are sometimes large differences between the attenuation fit and the fitted response. This is usually caused by a small number of events in that cell, common for cells at the edge of the detector. To ensure a good quality of the attenuation fit, the total χ^2 between the attenuation fit and the fitted response is calculated and only cells with the final $\chi^2 \leq 0.2$ are counted as *calibrated*. Cells with $\chi^2 > 0.2$ are ignored in further processing and marked as *uncalibrated*.

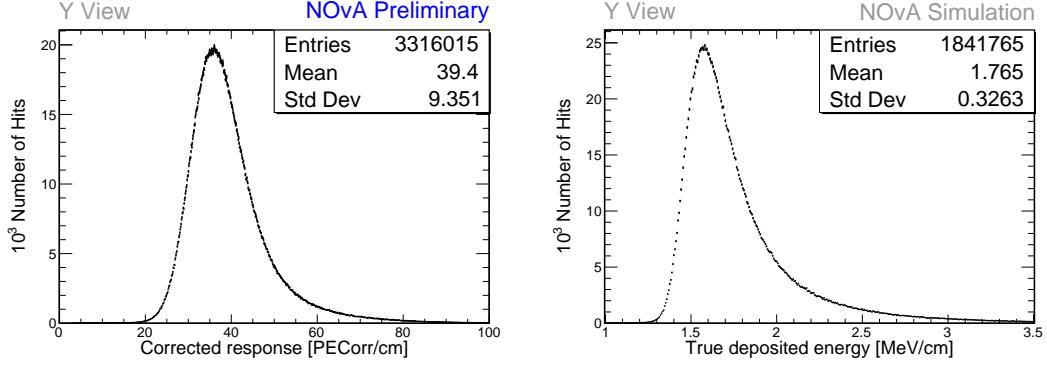


Figure 1.13: The absolute energy scale is calculated as the ratio between the simulated mean true deposited energy (right) and the mean reconstructed energy response (left) for selected stopping muons in each view and each data period or simulation.

1.6.4 Absolute Calibration

The absolute calibration only uses hits from muons stopping inside of the detector, in a track window 1 – 2 m from the end of their tracks. This is when they are approximately MIP and their energy deposition is well understood. Additionally, hits at the edges of each cell are removed to mitigate the effects at the end of the WLS fibres and the lower number of events at the edge of the detector [29].

First, the relative calibration results are applied to the selected stopping muon hits to get a distribution of the corrected detector response in *PECorr*/cm, as shown on the left of Fig. 1.13. The mean of this distribution is called the *reconstructed Muon Energy Unit* (MEU) and is calculated separately in each of the two views, and in each time period or version of simulation. Analogously, the mean of the true deposited energy in MeV/cm from simulation, shown on the right of Fig. 1.13, is called the *true MEU*. The absolute energy scale (the absolute calibration scale) is then the ratio between the true and the reconstructed MEU value, where both the MEU values are taken as a simple average over the two views

$$\text{Absolute Energy Scale} = \frac{\text{MEU}_{\text{True}} \text{ [MeV/cm]}}{\text{MEU}_{\text{Reco}} \text{ [PECorr/cm]}}. \quad (1.4)$$

The values of the absolute energy scales for each data period and simulation, as well as the results of the attenuation fit, are saved in a set of lookup tables, which are then used any time a hit is recorded in the NOvA detector and processed and reconstructed with the NOvA algorithms described above.

1.7 Energy Estimation

The deposited energy from detector calibration (Sec. 1.6) is only the first step in estimating the neutrino energy (E_ν) required for the main **NOvA** analyses.

For the ν_μ disappearance analysis, the ν_μ energy is measured as the sum of the muon energy and the energy of the hadronic shower [7]. The muon energy is identified from the length of its track, without the need of the calibration results. The energy of the hadronic shower is estimated from simulation as a fit to the 2D distribution of the true ν_μ energy minus the reconstructed muon energy, versus the visible (not corrected for the dead material) deposited energy of the hadronic system [75].

For the ν_e appearance analysis, the ν_e energy is calculated using a quartic fit to the 2D distribution of the electromagnetic versus the hadronic calorimetric energies, both corrected for the energy deposition in the dead material (PVC cells) [75]. The dead material correction is currently just a simple scaling of the deposited energy from calibration for all particles and is calculated from the measurement of the π^0 mass peak in the **NOvA ND**. This correction is correct only for electromagnetic showers and is not directly applicable to hadronic showers. The fit to determine the ν_e energy keeps the normalization of both the electromagnetic and the hadronic energies free, so the exact value of the dead material correction is not important. It is however used in other, non-neutrino oscillation analyses.

1.8 Systematic Uncertainties at **NOvA**

Systematic uncertainties in **NOvA** analyses arise from the imperfect knowledge on the individual components of the **NOvA** experiment, or from the known shortcomings of the prediction used to extract the measured parameters. Even though different analyses in **NOvA** need to consider different systematic uncertainties and their effect on the results varies, there are a few commonalities across all **NOvA** analyses that are explained below.

Both the 3-flavour [7] and the sterile neutrino [13] oscillation analyses in **NOvA** use the **ND** to constrain the **FD** prediction, which significantly reduces the effect of the neutrino beam and interaction prediction systematic uncertainties. On the other hand, these are the leading sources of systematic uncertainties for the **ND**-only anal-

yses, such as the cross section analyses [8–11]. The leading systematic uncertainty for the neutrino oscillation measurements comes from the detector calibration uncertainty. Significant uncertainties for all NOvA measurements also come from the neutron modelling, the detector simulation and the muon energy estimation. There are other sources of systematic uncertainties that are not mentioned here as they are sub-dominant, or specific to a certain analysis.

The systematic uncertainty on the prediction of the neutrino beam consists of two parts: the hadron production and the beam focusing uncertainties [36]. The uncertainty for hadron production is estimated by the **PPFX** (describe in Sec. 1.4) using the multi-universe technique. Here we create 100 **PPFX** universes in which the inputs from the external measurements used to constrain the hadron production are randomly floated around their central values within their respective systematic uncertainties. Parts of the hadron production that are not constrained by external measurements are given a conservatively large systematic uncertainty. The beam focusing systematic uncertainties account for the uncertainties on the horn and target positions, the horn current, the beam position on the target, the beam spot size, and the effect of Earth’s magnetic field in the beam pipe. **TO DO: *Describe the beam PCA and how can we limit the beam uncertainty***

TO DO: *Describe the neutrino interaction modelling systematic uncertainties*

TO DO: *Describe the detector modelling systematic uncertainties*

TO DO: *Describe the neutron uncertainty - only if I decide I need it*

TO DO: *Describe the calibration uncertainties*

CHAPTER 2

NOvA Test Beam Detector Calibration

The NOvA Test Beam experiment [79] 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 [80] 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 ND and FD readout electronics and filled with a range of NOvA scintillator oils, enabling a comparison of their respective performance and particle responses [81]. 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. 1.6).

In this chapter I introduce the NOvA Test Beam experiment in Sec. 2.1, focusing on the Test Beam detector and especially on the aspects that could impact its calibration. Section 2.2 describes the new data-based simulation of cosmic muons that I developed for the Test Beam detector calibration, while Sec. 2.3 discusses the calibration of the Test Beam detector itself.

2.1 The NOvA Test Beam Experiment

The NOvA Test Beam experiment [82] consists of a scaled down version of the NOvA ND and FD, shown in Fig. 2.1, 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) [83].

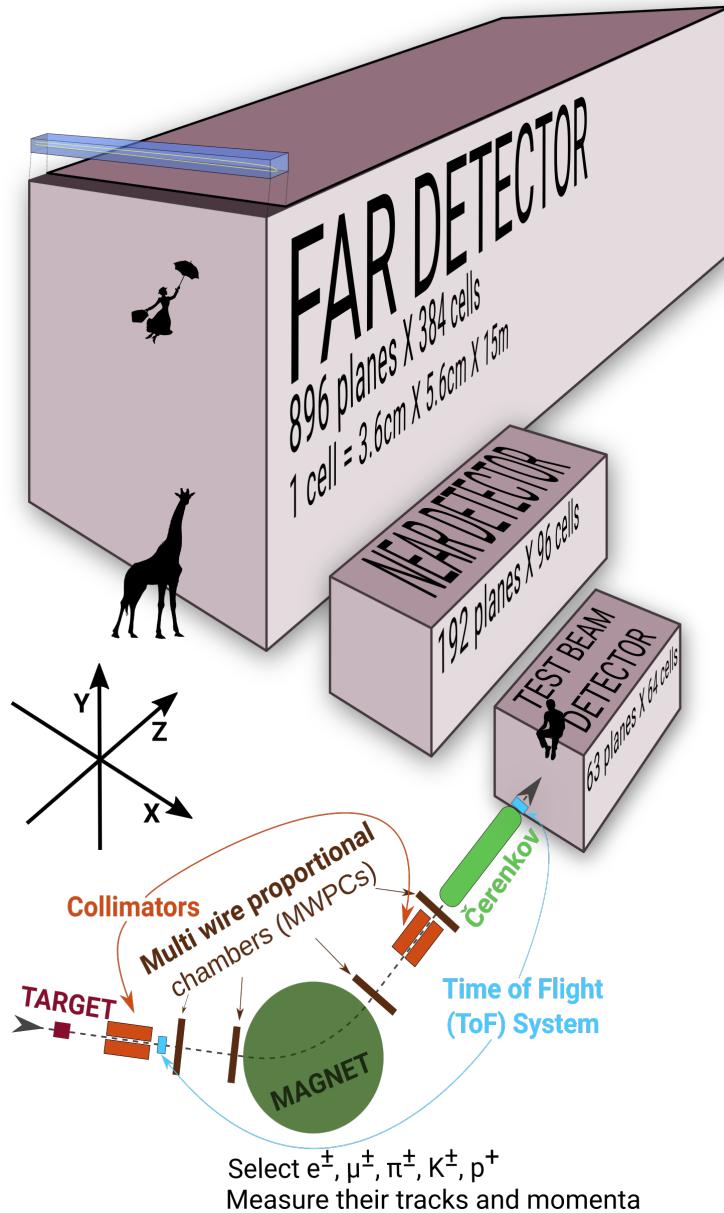


Figure 2.1: Comparison of Test Beam detector scale to the NOvA ND and FD (and a man, giraffe, or Mary Poppins). Also shown are the Test Beam beamline detectors and components (not to scale), with arrows showing the direction of the beam. The three black arrows show the orientation of the detector coordinate 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. 2.1. 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 2.1: 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 [NuMI](#), extracted once a minute in a continuous 4.2 s spill [79]. 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. 2.1, 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 [84]. 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. 2.1. It is centred with $(0, 0, 0)$ in the centre of the first plane [85]. 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 [86].

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 [85] and manifested as uncalibrated cells in the back of the detector [87]. 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 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. 2.2.

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 tanks outside in Fermilab [88];

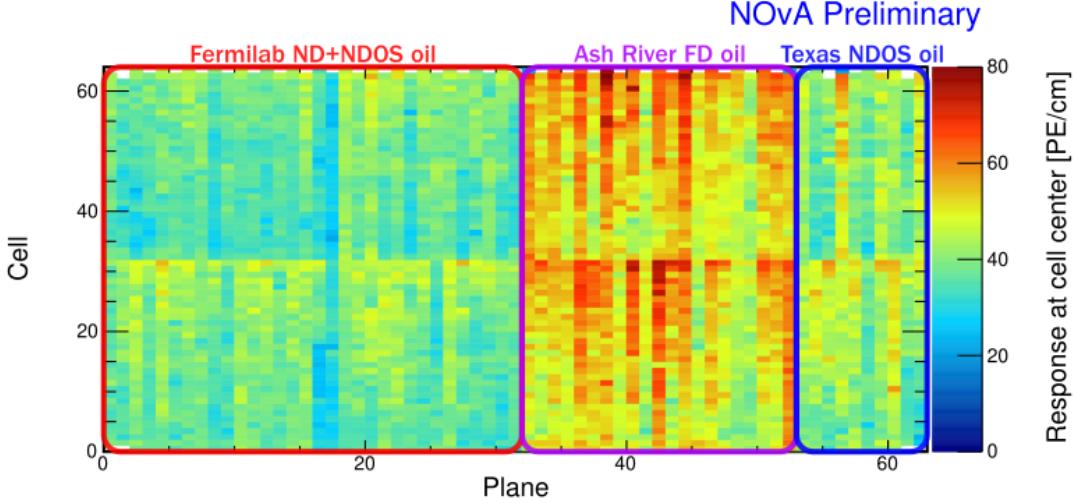


Figure 2.2: Uncorrected energy response in the centre of cells across the Test Beam detector showing a clear distinction between different scintillator oils, labelled with coloured boxes and descriptions.

2. Separate **ND** production oil and **NDOS**-drained oil stored underground in barrels at the MiniBooNE¹ cavern [80];
3. **FD** production oil stored inside in Ash River in ‘totes’ under several layers of black plastic [90];
4. **NDOS**-drained oil stored mainly inside at Texas A&M University and University of Texas at Austin [91, 92].

The original plan [93] 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) [88]. 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. 2.2.

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) [94]. These planes were again topped up with the **ND+NDOS** scintillator (sample #2).

The last 10 planes (planes 53 to 62) [94] were filled with the ‘Texas’ scintillator (sample #4), which has higher light yield than the one from the tanker, but lower

¹MiniBooNE [89] is a **Fermilab** experiment located close to the NOvA ND

than the Ash River one [91].

In total, the Test Beam detector is filled with 5418 gallons of scintillator oil with a weight of approximately 28.6 tons [84].

Readout

The Test Beam detector uses in total 126 [FEBs](#), each reading out signal from 32 cells [84]. 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. 1.3, 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 [95].

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 [95]. This was fixed [96] 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. 2.3.5.

2.2 Data-based Simulation of Cosmic Muons

The standard **NOvA** calibration procedure described in Sec. 1.6 uses the **CRY MC** generator (see Sec. 1.4) 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.

2.2.1 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. 2.1), 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. There-

fore, 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. 2.3. This should leave us with mostly cosmic events.

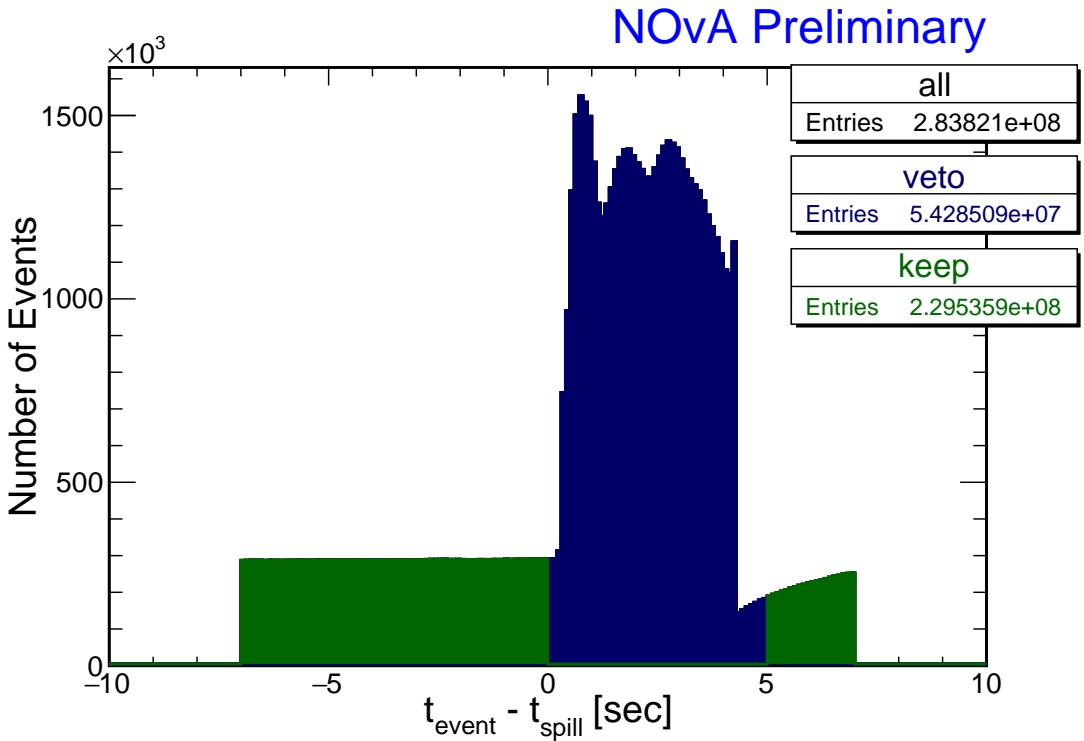


Figure 2.3: Test Beam beam spill events removed (blue) from the calibration samples. The remaining events (green) should mostly consist of cosmic particles. This example and the numbers of entries are for the full period 4 Test Beam sample.

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. 1.5. 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 **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. 2.2. 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. 2.4;
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 2.4 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 blocks and of a single block). 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. **COMMENT:**
Do we actually know why would there be a peak, even if small, after a single block? Is it due to reconstruction? Or more glue?

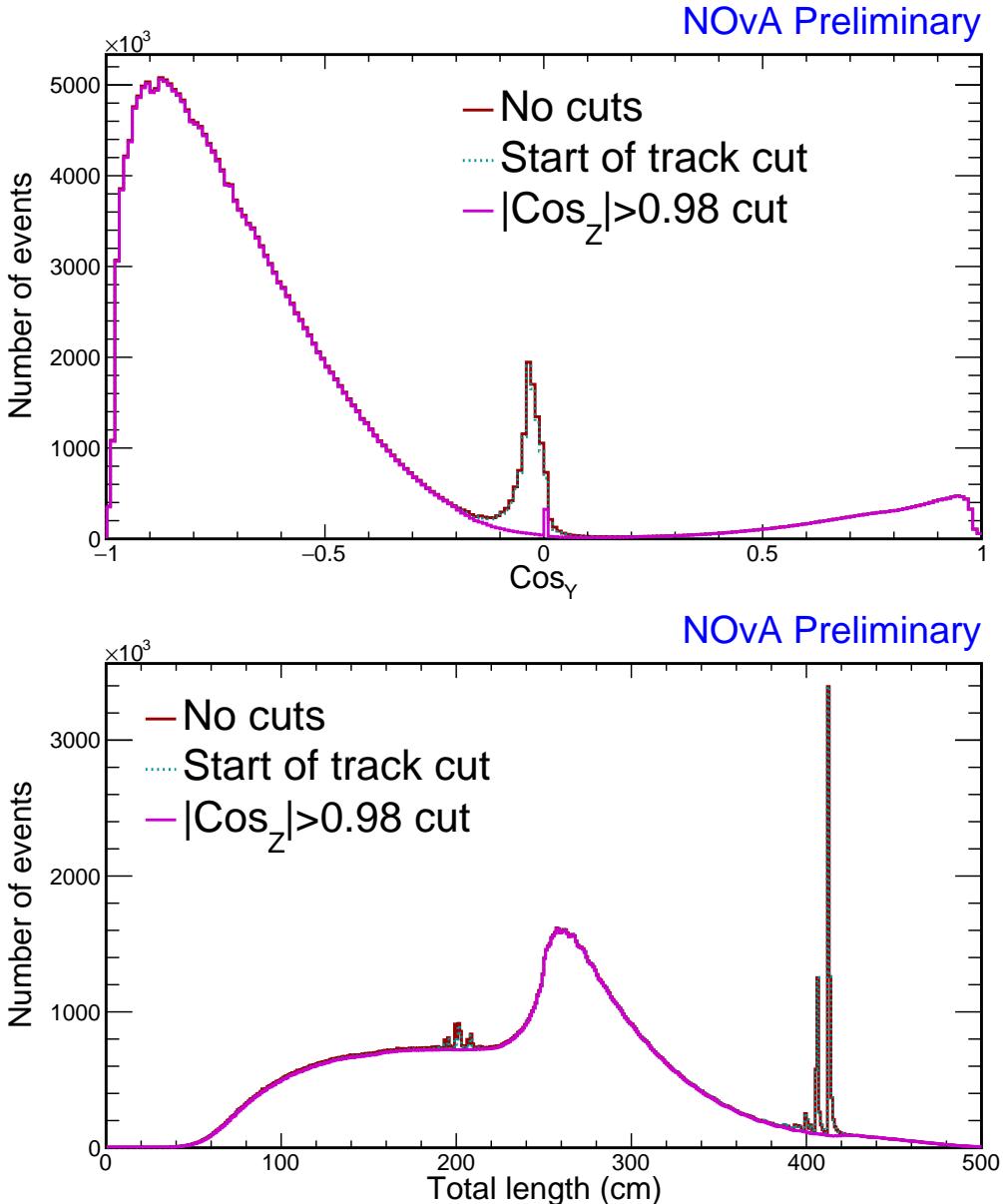


Figure 2.4: Impact of the cuts on the track start and on the maximum track angle from the z axis (Cos_Z) on the Test Beam data for the data-based simulation of cosmic muons. Top plot shows distribution of the angle from the y axis and bottom of the total track length, both made from the period 4 Test 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 (described in Sec. 1.6 and listed as Calibration sample selection in Tab. 2.2). 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. 2.5. 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, as can be seen in the bottom plot of Fig. 2.6, the **BPF** tracks have a rugged distribution in Cos_Z , which is not present for window cosmic tracks. The origin of this shape is not exactly known, but it is likely caused by the detector structure, as shown in Fig. 2.6. We concluded that the rugged shape should not have any impact on the resulting simulation. However given these differences between the tracking algorithms, applying the calibration cuts on the **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. Therefore, applying the full calibration cuts on misreconstructed events may remove them, even though they would have passed if they were reconstructed correctly and hence they should have been included in the simulation. These events would then be missing from the resulting simulation sample, introducing a bias.
COMMENT: Is this clear enough as to what I mean? Also, is the tense use correct?

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. 2.2 and shown in Fig. 2.7. There we also show the data calibration sample, which was created by applying the full calibration cuts on window cosmic tracks.

During the selection process, we determine whether the muon is stopping inside the detector or passing through, based on the end position of the reconstructed track. For Test Beam we say it is a stopping muon if its track ends at least 20 cm from

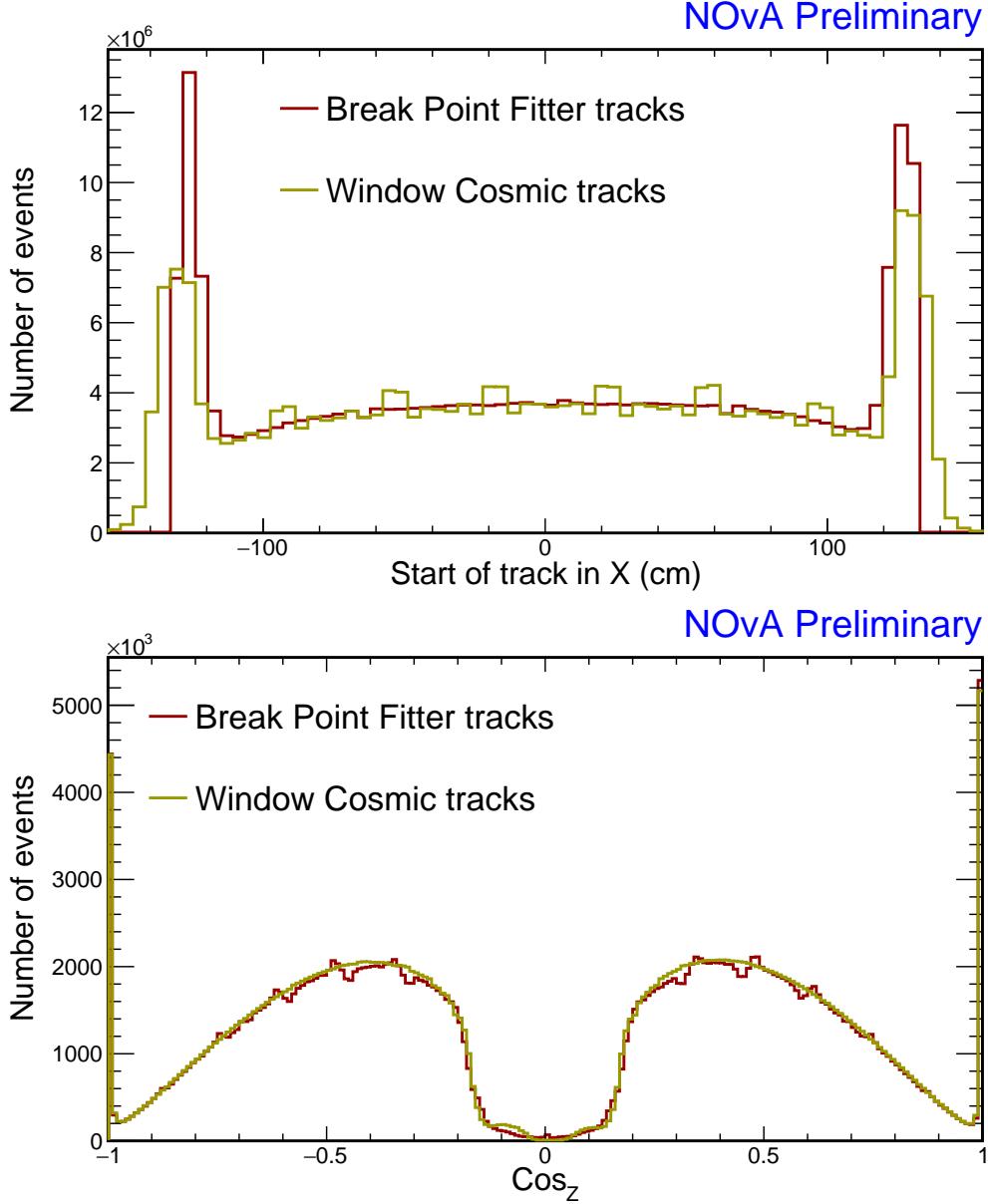


Figure 2.5: Difference between the tracks reconstructed with the [BPF](#) and with the window cosmic track algorithms. Top plot shows the distribution of the start of track along the x axis and bottom plot shows the distribution of the angle from the z axis, both for the period 4 Test Beam data (with removed beam spill) without applying any selection.

any edge of the detector. For the [ND](#) and [FD](#) this is 50 cm. This information assists in correcting the energy of through-going muons, as outlined in the following Sec. 2.2.2.

2.2.2 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 muons for the data-based simulation. This

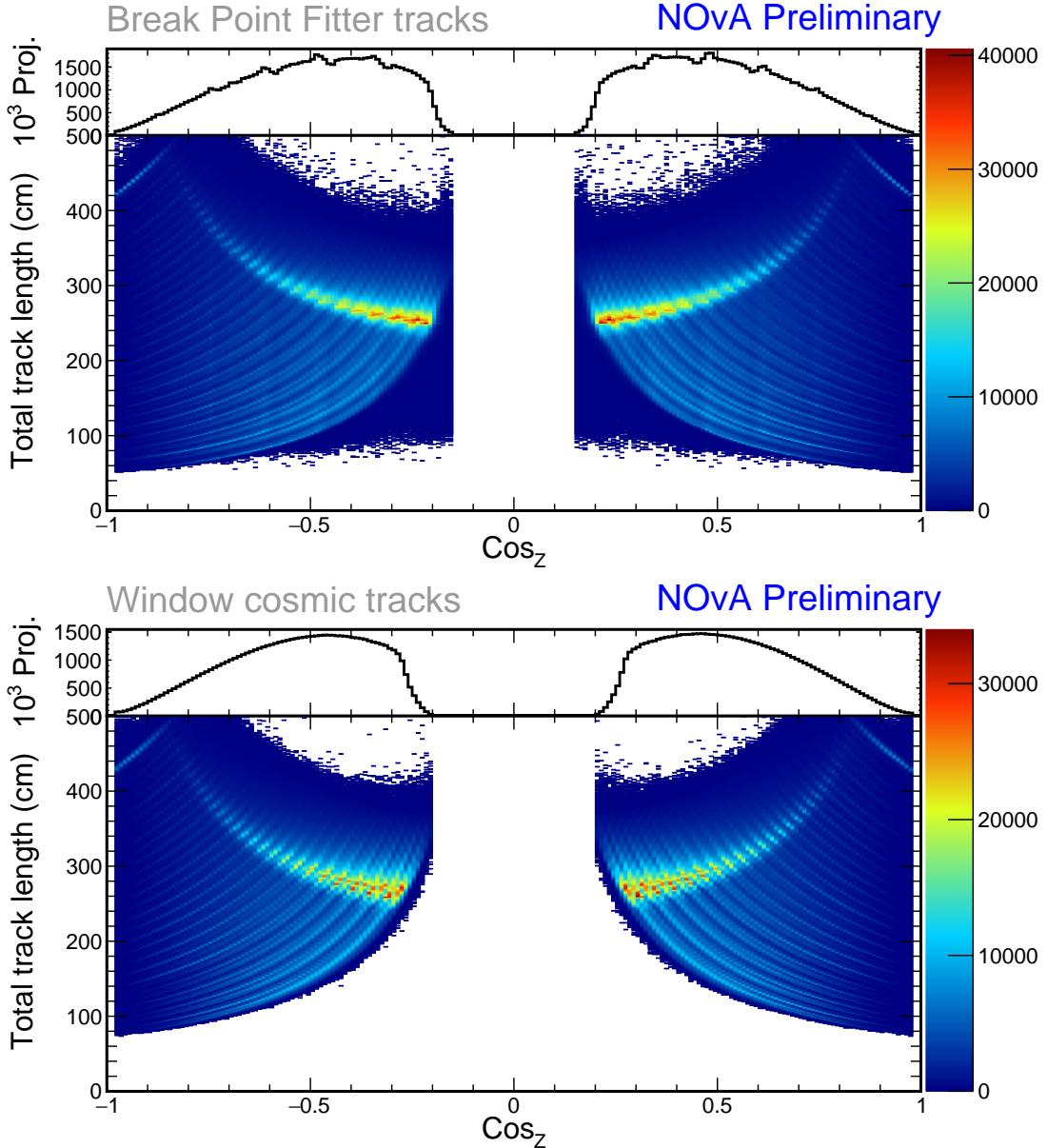


Figure 2.6: Comparison of the angle from the z axis (Cos_Z) and the total track length distributions between the [BPF](#) tracks (top) and window cosmic tracks (bottom). Top parts of both plots show the 1D Cos_Z distributions, scaled by $1/10^3$. The top plot is created with loose calibration cuts and the bottom plot with full calibration cuts per Tab. 2.2, although this difference in selection shouldn't matter. These plots are investigating the origin of the rugged shape in the Cos_Z distribution of [BPF](#) tracks, as can be seen in the top part of the top plot. The long curved light blue/green lines in the 2D plots correspond to constant values of $|\text{Cos}_Z| \times \text{Tot. length} = \text{Tot. length}_Z$, equal to the extent of the track length in the z direction, and are distinct from each other due to the structure of the detector (segmentation into planes). It is clear that the [BPF](#) tracks are peaked more sharply in Tot. length_Z than the window cosmic tracks, which are more spread out. This discrepancy could cause the resulting shape in the Cos_Z distribution of [BPF](#) tracks. **COMMENT:** *To be honest I don't fully understand where do the BPF peaks come from, I but I have an idea and I'm hoping I'm explaining it enough here. If not, happy to discuss how to improve, cause I'm struggling to explain it well myself*

Table 2.2: 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 (described in Sec. 1.6) in blue. The last two rows are not used for Test Beam, but are employed for the ND and FD.

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
	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

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.

Energy Correction

Through-going muons do not deposit all of their energy inside the detector. Therefore we cannot reliably calculate their initial energies from the reconstructed information, but we can estimate an energy that could leave the same track. In general, the energy spectrum of cosmic muons can be approximately described by a power law $E^{-\alpha}$, with $\alpha \approx 2.7$ [97, 98]. The expectation value for the ‘true’ initial energy of through-going muons can be therefore calculated as

$$\langle E \rangle = \frac{\int_{E_R}^{E_C} E \cdot E^{-\alpha}}{\int_{E_R}^{E_C} E^{-\alpha}} = \left(\frac{\alpha - 1}{\alpha - 2} \right) \left(\frac{E_C^{2-\alpha} - E_R^{2-\alpha}}{E_C^{1-\alpha} - E_R^{1-\alpha}} \right), \quad (2.1)$$

where E_R is the reconstructed energy we got from the BPF. E_C is the critical energy chosen conservatively 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

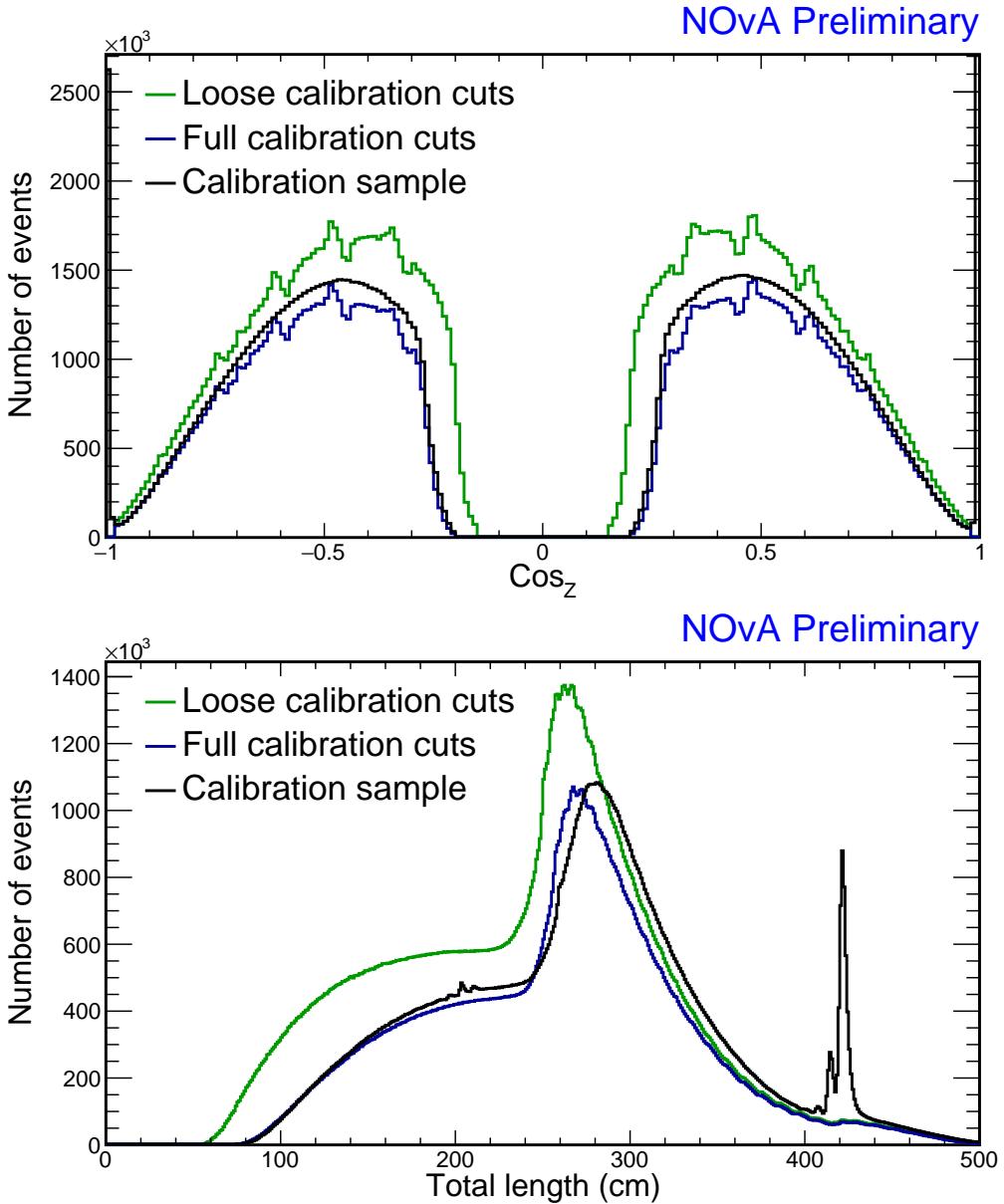


Figure 2.7: Comparison of full and loose event selections for the data-based simulation as per Tab. 2.2 and of the corresponding data calibration sample in black. As described in text, the loose calibration cuts applied to the BPF tracks (green) were used for simulation to mitigate the discrepancy between applying the full calibration cuts to the BPF tracks (blue) and the window cosmic tracks (black). All of the distributions are made from the period 4 Test Beam data.

selection described in Sec. 2.2.1. Figure 2.8 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 however **not** a good representation of the true energy spectrum of cosmic muons at surface level and getting a correct energy distribution from data would require a much more dedicated effort. The corrected energy would

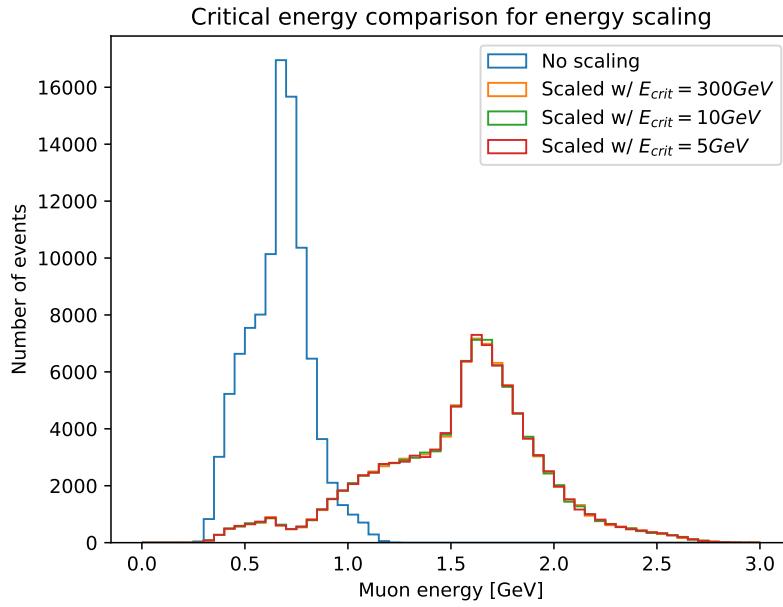


Figure 2.8: The effect of energy correction for through-going muons with various critical energies. No significant difference can be seen when using different critical energies.

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 **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 representation 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 [98].

Smearing

The reconstructed distributions in data are 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 varying

- 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.

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. The probability that a muon has a positive charge can be expressed as [97]

$$P_+ \simeq 0.539 + \frac{x}{34.5} - \left(\frac{x}{9.48}\right)^2 + \left(\frac{x}{8.27}\right)^3, \quad (2.2)$$

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. 1.4. We use the fibre brightness map that is used in calibration (see Sec. 1.6) 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. 2.3.1.

2.2.3 Validation

To validate whether the newly created simulation works as expected, we create the simulation calibration sample using the reconstruction and selection of cosmic muons for calibrations described in Sec. 1.6. We then compare this to the equivalent data calibration sample, created from the same data that was used to seed the simulation. Additionally, we use the newly simulated events as ‘fake data’ and pass them through the same reconstruction, selection and simulation processes as were used to create the first iteration, hence creating a ‘re-simulation’ sample. This is used to validate the stability of the simulation process.

The data-simulation comparisons are shown in Fig. 2.9-2.12. The data and simulation calibration samples are shown in black and pink solid lines respectively. Both are equivalent to applying the full calibration cuts to the window cosmic tracks, as described in Sec. 2.2.1. For comparison, we are also showing distributions of the **BPF** tracks with full and loose calibration cuts in blue and green dashed lines respectively. We are expecting that the distributions of the simulation calibration sample (pink) match the distributions of the data calibration sample (black), without being affected by the different shape of the **BPF** tracks (dashed lines).

It can be seen that the distributions of the new simulation calibration sample resemble those of the data **BPF** tracks with full calibration cuts applied (blue dashed lines) more closely than those of the data calibration sample. This indicates that the simulated window cosmic tracks have similar properties to the data **BPF** tracks, meaning that the differences between the tracking algorithms have an impact on the resulting simulation. Therefore, loosening the calibration cuts did not help with mitigating these differences as expected. However, there were in total four versions of simulation created, with varying event selections and smearing applied, including using full calibration cuts and with different loosened calibration cut values. It is clear that it is unlikely we can mitigate the differences between the window cosmic tracks and the **BPF** tracks simply by changing the event selection or smearing and it would require adapting (or developing) a different reconstruction algorithm for the simulation. Due to the time and effort required for such a task, and due to the relative similarity between the simulation and data calibration samples, we decided to proceed with the reconstruction, selection, and correction as described above.

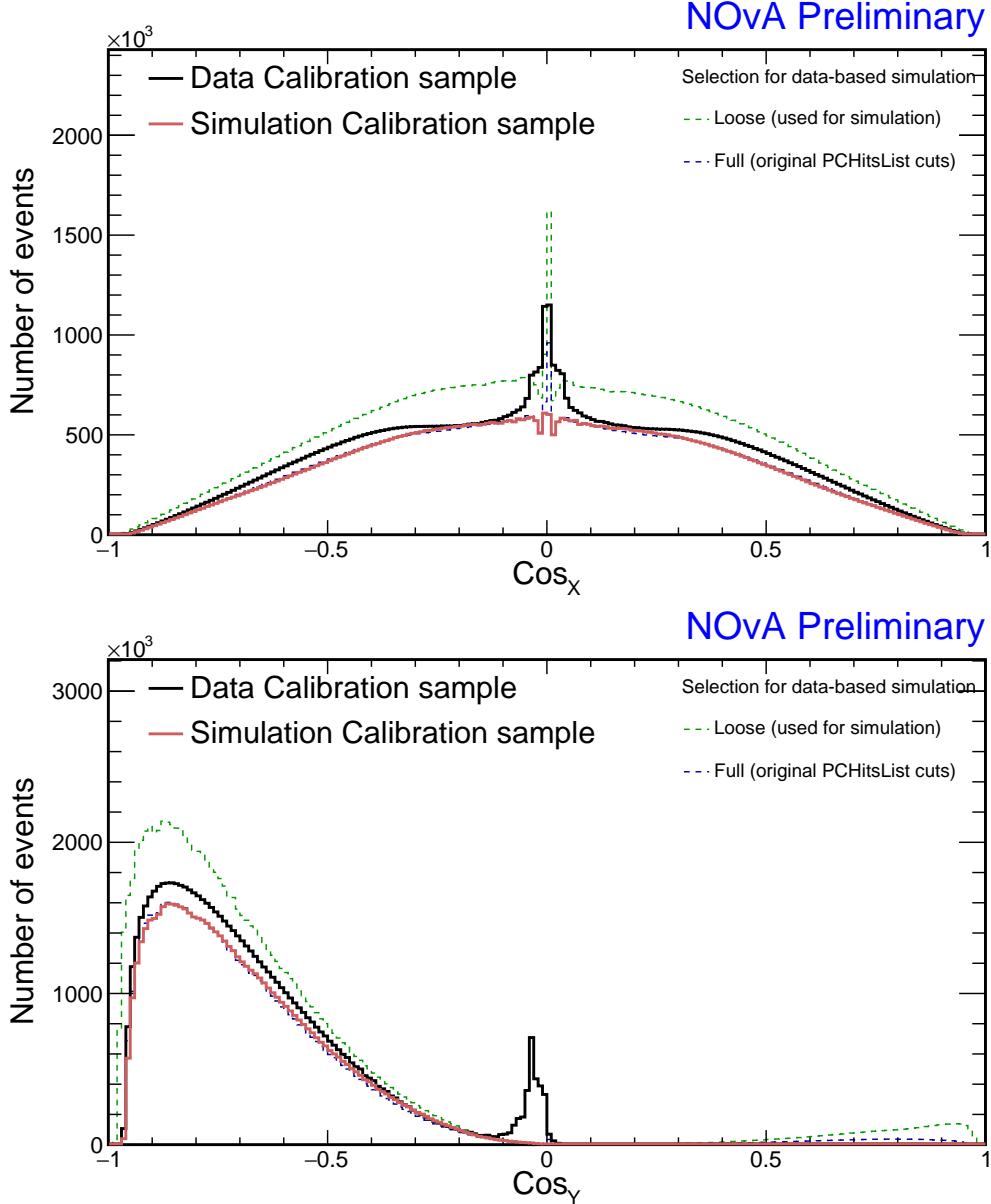


Figure 2.9: Comparison of the angle from the x axis (top) and y axis (bottom) between the data (black) and simulation (pink) calibration samples, as detailed in the text. Additionally, the distributions of data with full (blue) and loose (green) selections applied to the BPF tracks are shown, where the loose selection was used to create the simulation.

Figures 2.9 and 2.10 show the effect of removing the beam-like events with a $|\text{Cos}_Z| < 0.98$ cut as described in Sec. 2.2.1. While we applied this cut to the data events used to create the simulation (green dashed lines), it is not included in the selection process for the calibration samples. This means that we were successful in removing these undesirable events from the new simulation. This can be seen by the wide peaks at $\text{Cos}_{X/Y} = 0$ in Fig. 2.9 which are present for the data calibration sam-

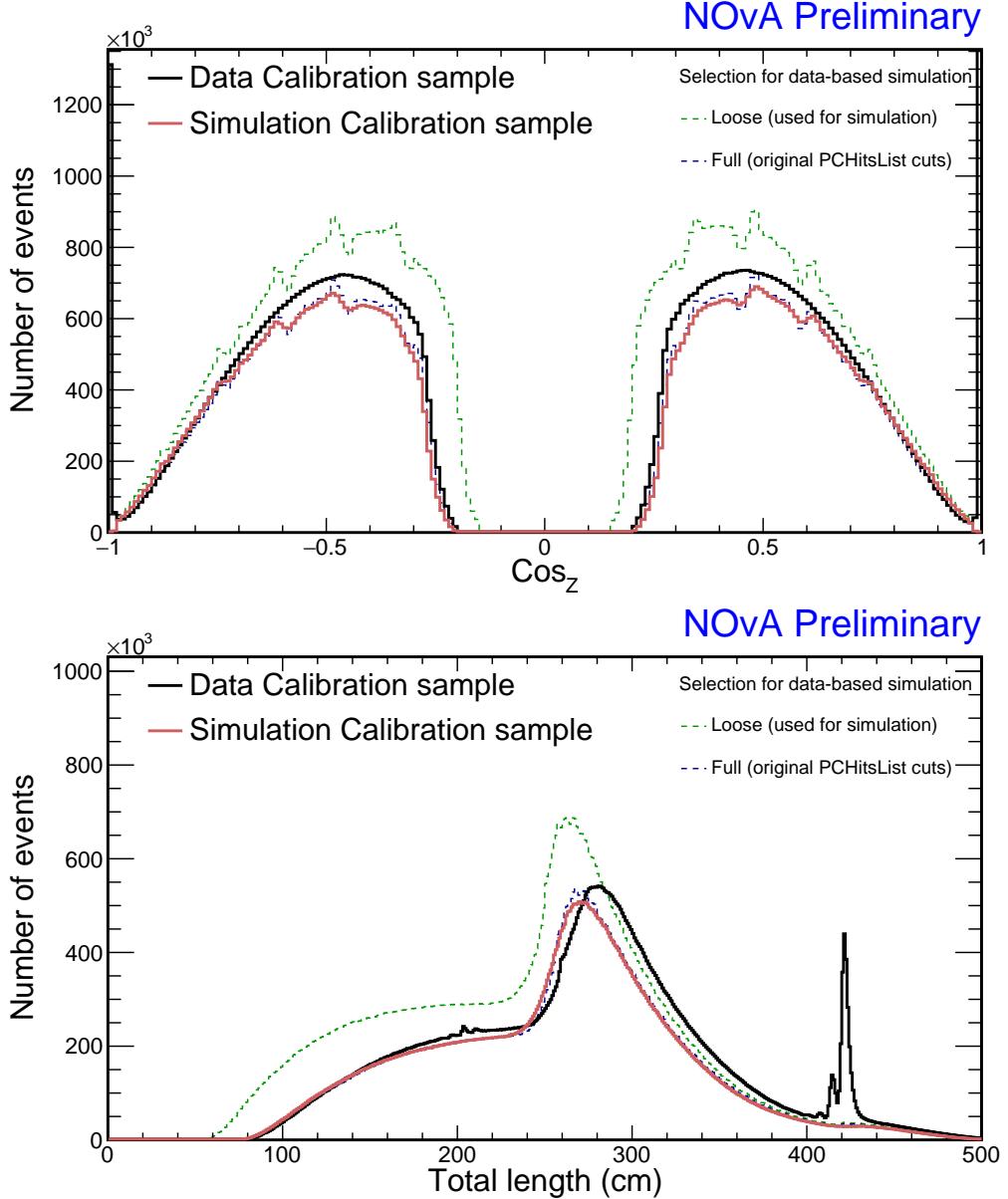


Figure 2.10: Comparison of the angle from the z axis and total track length between the data (black) and simulation (pink) calibration samples, as detailed in the text. Additionally, the distributions of data with full (blue) and loose (green) selections applied to the BPF tracks are shown, where the loose selection was used to create the simulation.

ple, by not for the simulation. Additionally, sharp peaks in the total track length and in the edges of the angle from the z axis distributions in Fig. 2.10 further illustrate this distinction.

The effect of smearing of vertex positions and four momenta is visible in the top plot of Fig. 2.10 and in Fig. 2.12. Here, the rugged distributions of the data used for simulation (green dashed lines) appears smoother for the simulation calibration sam-

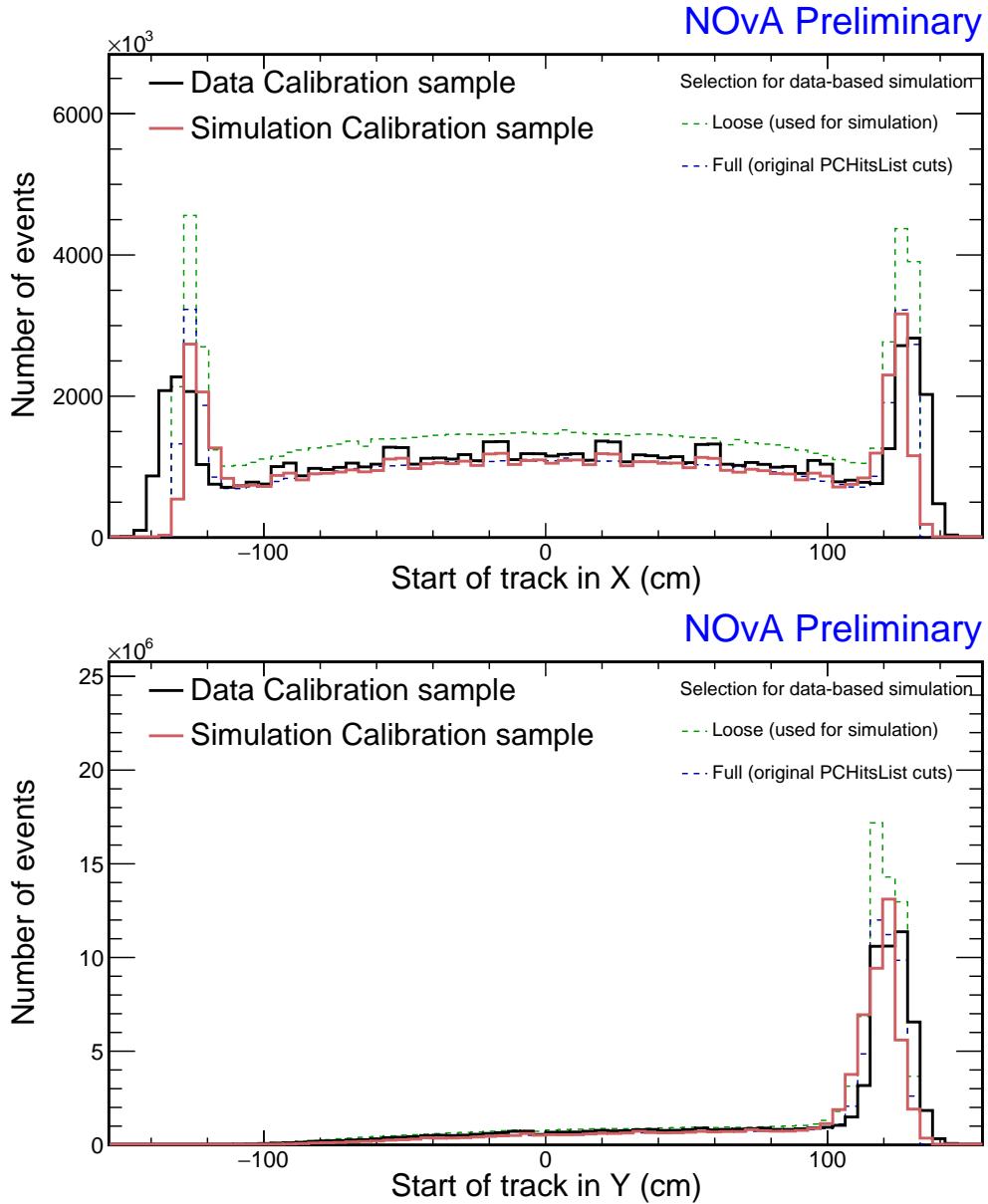


Figure 2.11: Comparison of the x (top) and y (bottom) track start position between the data (black) and simulation (pink) calibration samples, as detailed in the text. Additionally, the distributions of data with full (blue) and loose (green) selections applied to the BPF tracks are shown, where the loose selection was used to create the simulation.

ple (pink solid lines). However, comparisons of track start positions in Fig. 2.11 and 2.12 show that the simulated events start further away from the detector edge compared to data. While this is likely also a result of smearing, we anticipate it will not impact the calibration of the simulated detector.

After adding the distributions for the re-simulation calibration sample, shown as solid brown lines Fig. 2.13, we can see that the track start positions are shifted even

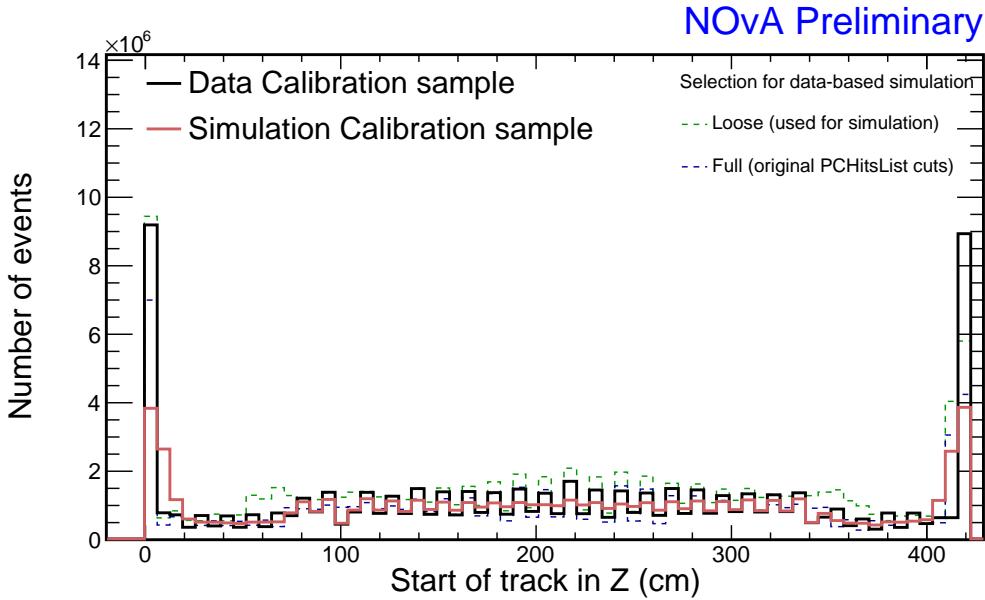


Figure 2.12: Comparison of the z track start position between the data (black) and simulation (pink) calibration samples, as detailed in the text. Additionally, the distributions of data with full (blue) and loose (green) selections applied to the BPF tracks are shown, where the loose selection was used to create the simulation. Each bin corresponds to a single detector plane and the shape of the distribution is caused by the fact that cosmic rays are naturally more vertical and therefore more likely to be detected by horizontal than vertical planes.

further towards the inside of the detector. This would support the hypothesis that this effect is caused by smearing and is also likely related to the loss of events with longer track lengths, as shown in Fig. 2.13, since if tracks start a few centimetres later in the detector their tracks would get shorter by the same amount. Since there are only minimal discrepancies between the simulation and the re-simulation, we conclude that the reconstruction, selection and correction processes used to create the simulation do not significantly bias the simulation, which is self-consistent.

2.2.4 Conclusion

COMMENT: *Should I include this conclusion section here?*

I have successfully developed a new simulation of cosmic muons intended for Test Beam calibration. The reconstruction, selection, and correction processes were designed to minimize the number of simulated events, resulting in an efficient and concise simulation while avoiding a significant bias, particularly in the applicability to the calibration procedures.

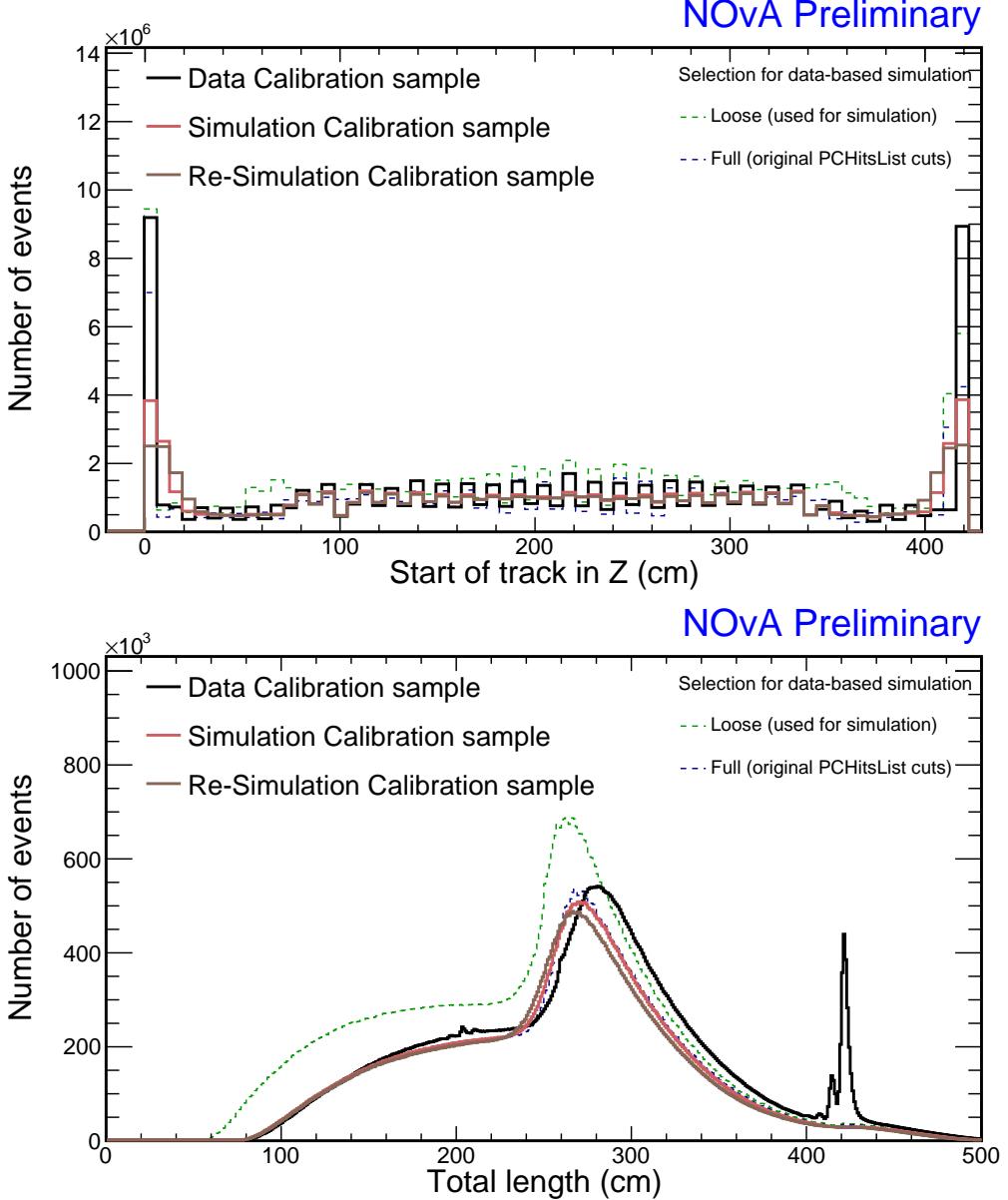


Figure 2.13: Comparison of the z track start position and total track length between the re-simulated events (brown), which use simulation as ‘fake data’ for the second iteration of the same simulation process, and the data (black) and simulation (pink) events. Additionally, the distributions of data with full (blue) and loose (green) selections applied to the [BPF](#) tracks are shown, where the loose selection was used to create the simulation.

While the simulation achieved its objectives, there are possible improvements that could be introduced to refine it. Using a different track reconstruction algorithm than the [BPF](#) more consistent with the window cosmic track algorithm could make the simulation calibration sample more alike the data calibration sample. Additionally, developing a more sophisticated energy correction procedure, potentially involving external measurements of energy distribution of cosmic muons, would enable simu-

lation of through-going muons with accurate incident energies.

Looking ahead, there are plans to use the data-based simulation approach for cosmic muons in the **NOvA ND** and **FD**. While the primary motivation is again detector calibration, there is potential for broader applications in cosmic ray studies. However, using the simulation for a different detector would require re-validating the event selection process, especially with the inclusion of the calibration cuts missing for the Test Beam detector, as outline in Sec. 2.2.1 and showed shaded out in two bottom rows of Tab. 2.2.

2.3 NOvA Test Beam Detector Calibration

In this section I 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 created using the same procedures as the **ND** and **FD** calibration samples, described in Sec. 1.6. However, there are two cuts from the event election, that were not included for Test Beam during the processing of the data samples. This can be seen on Tab. 2.2, 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 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 negligible.

This section is organized as follows. I first describe the Test Beam versions of the fibre brightness map in Sec. 2.3.1 and the threshold and shielding correction in Sec. 2.3.2, as they were introduced in Sec. 1.6. I then go over the simulation sample and the three data samples (for periods 2, 3, and 4), showing distributions of hits selected for calibration and of the uncorrected energy deposition before calibration. I discuss considerations going into calibration, including splitting the individual periods into smaller samples, or describing issues that could affect the calibration results. After-

wards, I am showing a selection of attenuation fit results for each sample together with an overview of the relative calibration effects. Lastly, I discuss the absolute calibration results for all the samples combined, as well as the validation and conclusion of the Test Beam calibration.

2.3.1 Fibre Brightness

To divide the Test Beam detector into **FB** bins we use the attenuation fit results for Test Beam period 4 data (described in Sec. 2.3.6), as that is the best detector conditions data we have. Since we need the **FB** map in order to run the attenuation fits and we need the attenuation fit results to create the **FB** map, we proceeded iteratively. We first run the attenuation fit with an older version of the **FB** map and use the results to create a new **FB** map, discussed here, which is then used in a new attenuation fit.

We are only using the attenuation fit results in the centre of each cell to create the **FB** map, therefore, we decided to allow some cells that failed the calibration condition ($\chi^2 > 0.2$), to be still used for the creation of the **FB** map. Otherwise, all the officially uncalibrated cells are assigned an average response across the entire detector, resulting in a loss of information on their relative brightness. As can be seen in Fig. 2.14, 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 the uncalibrated Test Beam cells (doable for Test Beam, due to its small number of cells), we concluded that all the cells with $\chi^2 < 0.7$ can be used to create the **FB** map, since the response in their centre is described reasonably well by their attenuation fits. We use this loosened calibration condition only to create the **FB** map and we keep the original condition for the actual calibration results.

The final distribution of relative **FB** values that are used to populate the **FB** bins for the Test Beam detector is shown in Fig. 2.15. The resulting map of **FB** bins and their corresponding relative brightnesses was shown in the previous chapter in Fig. 1.10.

2.3.2 Threshold and Shielding Corrections

The threshold and shielding correction is intended to mitigate biases arising from differences between cosmic events used for calibration and beam events. It is only used prior to the attenuation fits and is omitted when applying the results of the relative

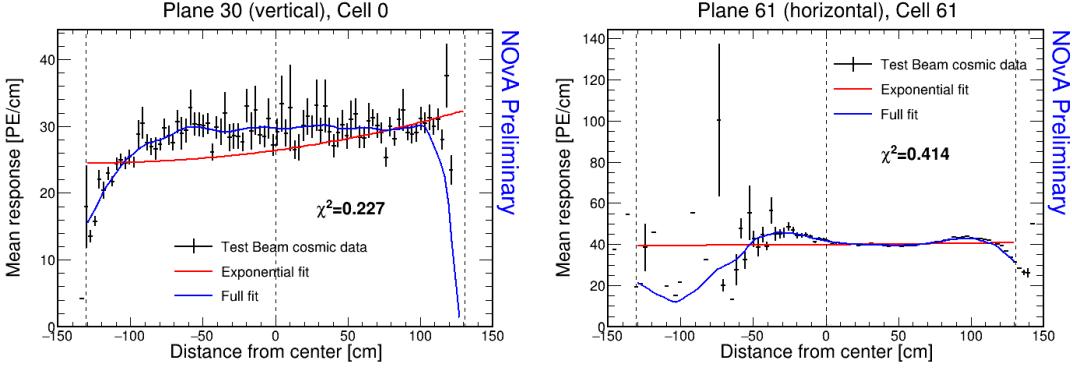


Figure 2.14: Examples of attenuation fits for two cells that fail the calibration condition, but the fit (blue line) still correctly represents the energy deposition in the centre of that cell (dashed vertical line in the middle). The total χ^2 between the data (black) and the attenuation fit for both plots are included.

calibration, whether during the absolute calibration or for beam events. Additionally, it is derived exclusively from simulation.

We created a new threshold and shielding correction for the Test Beam detector using the new simulation described in Sec. 2.2. The correction is calculated for both views, across 12 FB bins, 64 cells, and 100 w bins, where $w \in (-130 \text{ cm}, 130 \text{ cm})$. Two examples of the correction as a function of w are shown in Fig. 2.16, demonstrating an almost uniform behaviour along a cell. Relative variations of the correction in the X view range from 1 – 2 %, primarily concentrated at the edges of the cell. In the Y view, the correction exhibits sub-1 % variations. These trends are consistent across all the FB bins and views. Given that the threshold and shielding correction precedes relative calibration, the absolute value of the correction is irrelevant and only the relative variations along w and between cells matter.

This uniformity of the distributions is expected, considering the relatively small size of the Test Beam detector compared to the FD, which prompted the investigation into threshold and shielding effects. The Test Beam detector’s cell length of 2.6 m has only a negligible impact on the threshold saturation or on the energy distribution of cosmic muons, resulting in the uniformity of the threshold and shielding correction for Test Beam detector. The larger correction at cell edges is likely caused by lower event counts in those areas. However, since this relative sparsity of events also influences relative calibration due to large variation in the energy response, the relatively larger threshold and shielding correction at cell edges is not detrimental.

The distribution of the threshold and shielding correction across Test Beam detec-

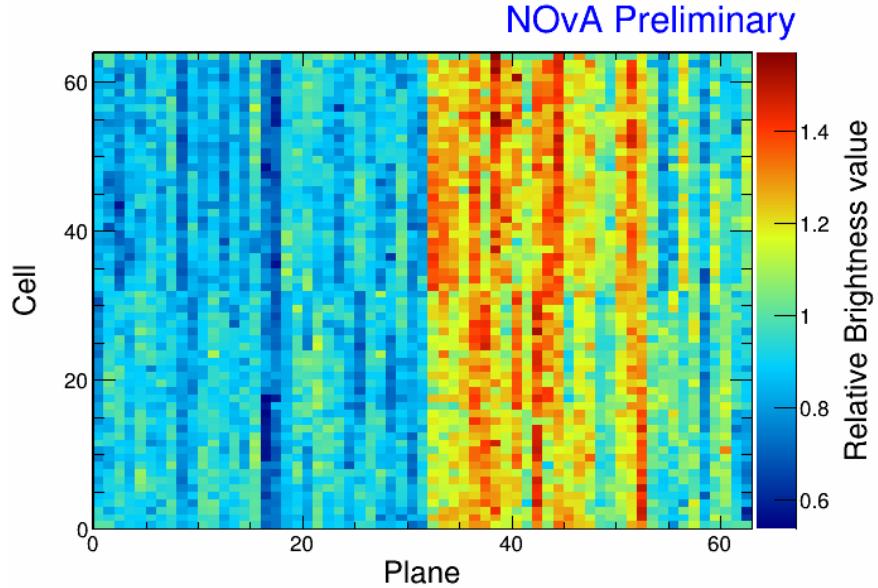


Figure 2.15: FB map representing relative differences in energy response due to different brightnesses of the fibres, scintillators, or readout. Create from the attenuation fit results of the NOvA Test Beam detector with a shifted calibration condition from $\chi^2 < 0.2 \rightarrow 0.7$ to enable using the attenuation fits that are officially uncalibrated, but correctly represent energy deposition in cell centre. Otherwise, all the uncalibrated cells get assigned a mean detector response, represented by number 1 on this map.

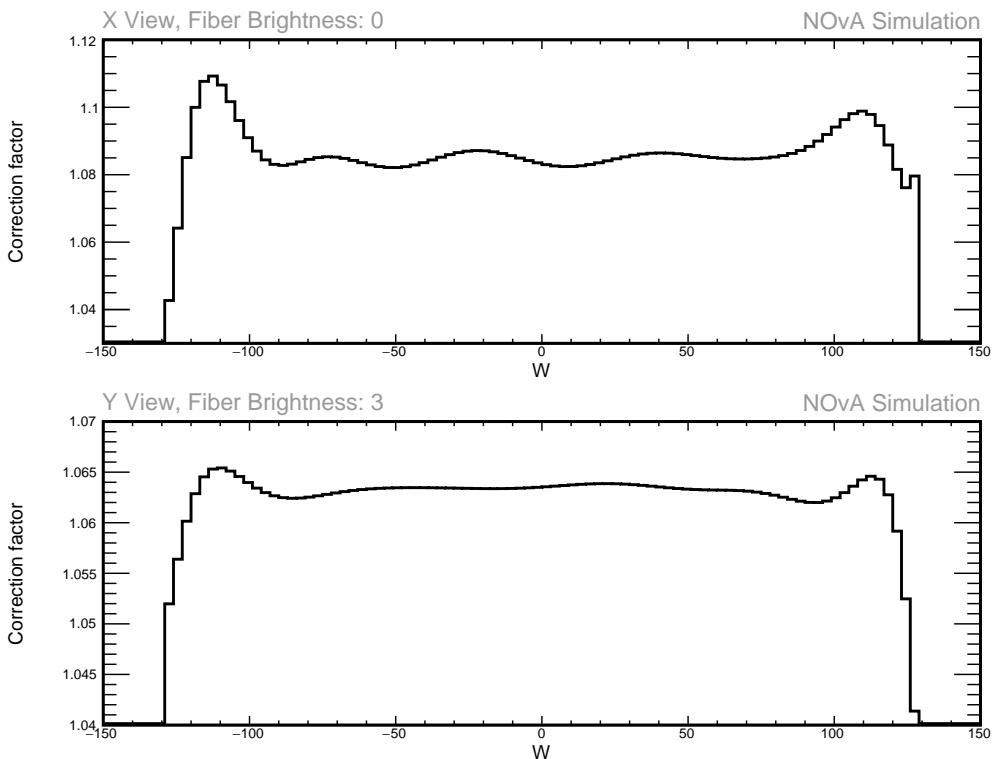


Figure 2.16: Examples of threshold and shielding corrections as a function of the position within a cell in X view (top) and Y view (bottom) for the Test Beam detector.

tor's cells and planes, shown in the top of Fig. 2.17, demonstrates, that while the correction is expected to be generally uniform across the detector, there are notable variations between cells and planes forming a discernible pattern. These variations and their shape primarily stem from the threshold component of the correction, shown in the bottom of Fig. 2.17.

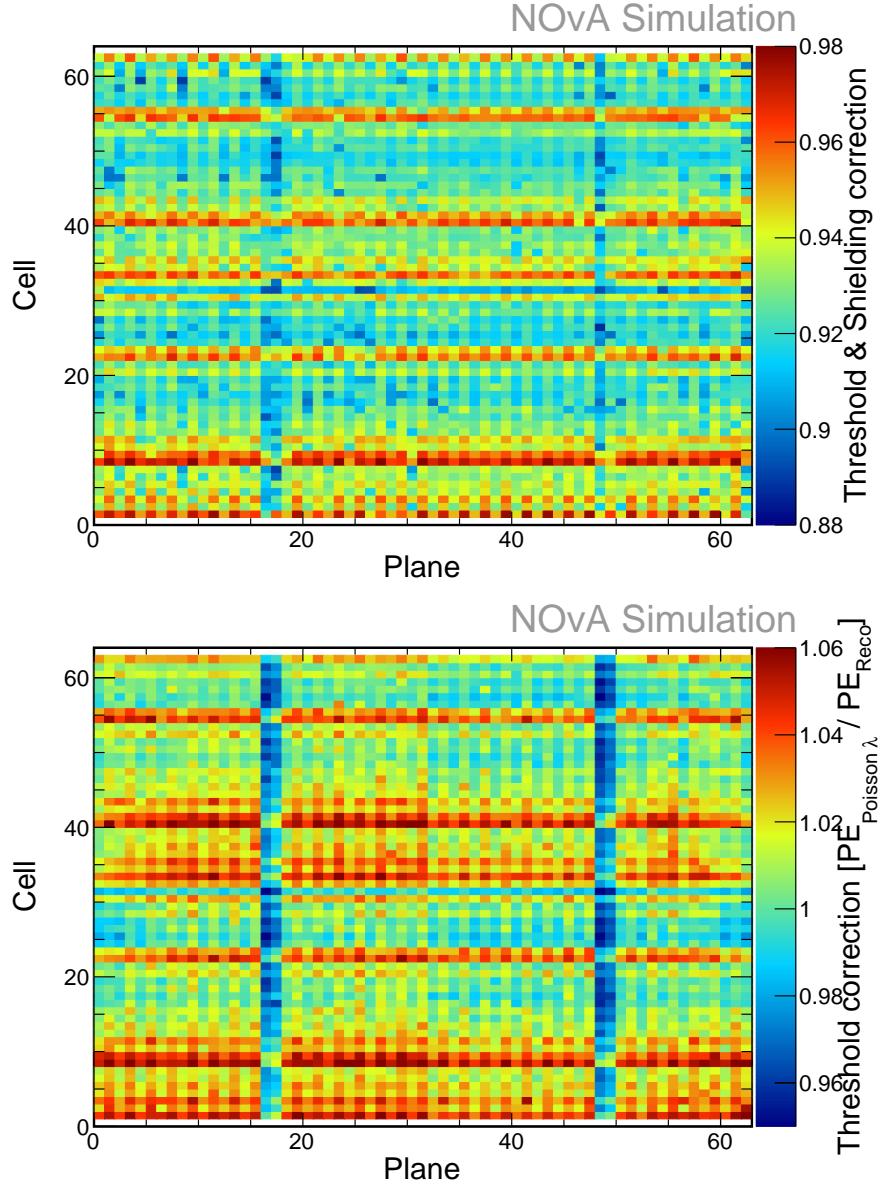


Figure 2.17: Map of the threshold and shielding correction (top) and only of the threshold part of the correction (bottom) as a function of the Test Beam detector's cell and plane number. Each bin shows the mean correction for all the simulated events in that cell.

The threshold part of the correction can be expressed as

$$\text{Threshold correction} = \frac{PE_{Poisson\lambda}}{PE_{Reco}}, \quad (2.3)$$

where $PE_{Poisson\lambda}$ represents the mean of the Poisson distribution of the true deposited energy (in terms of PE_{True}), and PE_{Reco} is the reconstructed number of PE from simulation. Both $PE_{Poisson\lambda}$ and PE_{True} are direct outputs of the light model simulation, as detailed in Sec. 1.4. After the light model simulation, PE_{True} is passed through the readout simulation, which includes a PE-to-ADC function for calculating the peak ADC value. This value is then converted into PE_{Reco} using the ADC-to-PE scale described in Sec. 1.6. The observed shape in the threshold correction can thus be attributed to differences between PE_{True} and $PE_{Poisson\lambda}$, as well as to various effects introduced by the readout simulation. However, the differences between PE_{True} and $PE_{Poisson\lambda}$ are marginal (below 1 %) and contribute minimally to the threshold correction. Therefore, the predominant influence on the observed pattern comes from the effects introduced by the readout simulation.

There are two prominent features in the threshold correction variations in Fig. 2.17. Firstly, the two blue vertical lines in planes 16-17 and 48-49. These planes are using the FEB version 5.2, used in the ND, instead of the FEB version 4.1, used in the FD and in all the other Test Beam planes, as explained in Sec. 2.1. Both the readout simulation and the ADC-to-PE scale do account for the expected disparity in the ADC/PE ratio between the two FEB versions. However, it is expected that FEBv5 would exhibit a lower response to the same energy compared to FEBv4. Therefore, for the same $PE_{Poisson\lambda}$ values, the PE_{Reco} for FEBv5 should be smaller than that for FEBv4. Consequently, the FEBv5 planes should have a larger threshold correction compared to the FEBv4. However, as was shown in Fig. 2.17, the observed correction is contrary to this expectation, suggesting a potential error in the readout simulation regarding the handling of different FEB versions.

The second notable feature in Fig. 2.17 is the variation of the threshold correction across cells, which appears to be consistent across all planes, depicted by the presence of red horizontal lines. The origin of this dependency is in the APD structure, where each APD collects signal from 32 cells arranged in 4 rows of 8 APD pixels, as explained in Sec. 1.3. Manufacturing discrepancies [99] lead to relative gain variations among

the [APD](#) pixels, typically exhibiting either increasing or decreasing trend along each of the four rows. To incorporate these variations into the readout simulation, the mean relative gain across cells of every module (comprising 32 cells) is used in the [PE-to-ADC](#) function. Consequently, these variations are consistent across all modules in the simulated detector, despite their inherent randomness in actual data.

The distribution of the relative gain for each ‘pixel number’ is shown on the left of Fig. 2.18. However, it is important to note that ‘pixel number’ is a [NOvA](#) jargon and does not correspond directly to the [APD](#) pixel position or cell number; instead it denotes the purely technical routing of [APD](#) pixels to the [FEB](#) [100]. Therefore, the depicted distribution of gain variation on the left of Fig. 2.18 is incorrect and should instead describe the distribution with respect to the cell number rather than the ‘pixel number’. Simply translating ‘pixel numbers’ to cell numbers yields the distribution shown on the right of Fig. 2.18. Comparing this to the positions of the red horizontal lines in Fig. 2.17 demonstrates that this (incorrect) relative gain variation is responsible for the observed pattern in the threshold correction.

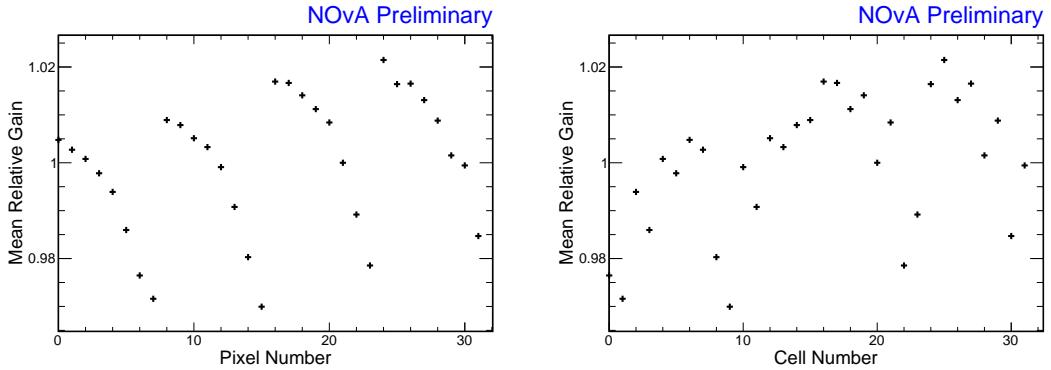


Figure 2.18: The relative gain variation as a function of the ‘pixel number’ (left) and cell number (right).

In summary, the threshold and shielding correction exhibits significant variations concentrated within specific planes and cells, arising from various effects in the readout simulation. However, it is evident that these effects are not limited to cosmic events and therefore should not be incorporated into the threshold and shielding correction. Given that these effects are corrected out for simulation before the attenuation fits, they are not accounted for in the relative calibration and therefore remain present for the absolute calibration and for beam events. Moreover, the two main effects outline above are not implemented into the simulation correctly, result-

ing in discrepancies between actual data and simulation. This means, that in data these variations are either not present or present in a different way than in simulation. Therefore, applying the simulation-based threshold and shielding corrections to data introduces new variations that would otherwise not exist for data. As a result, these new variations are incorporated into the attenuation fits for data, resulting in incorrect relative calibration results applied to both absolute calibration and beam events.

Several approaches can address these issues. For simulation-related discrepancies, the only viable solution is to rectify the identified faults and to remake the simulation, albeit this would be computationally very intensive. However, for data-related concerns, efforts are underway to devise a new data-driven threshold and shielding correction [101], eliminating any influence of simulation on the relative calibration of data. If a purely data-driven correction is not viable, there is another possible improvement to the threshold correction while still using simulation, which is to not use the $PE_{Poisson\lambda}$ directly, but to pass it through the readout simulation in the same way as PE_{True} and create an alternative $PE_{Poisson\lambda Reco}$.

2.3.3 Simulation

The distribution of tricell hits from the simulated cosmic muon events selected for calibration, mapped across the Test Beam detector’s planes and cells, is shown in Fig. 2.19. As this is a simulated detector, we will use this ‘ideal conditions’ distribution of tricell hits to illustrate the main features, which are also present in all the data samples discussed below. We can clearly see the difference in the number of events between 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. These cells can still be calibrated by including hits from the ‘z tricell’ condition, which is not shown in the plot. The three clear horizontal lines of relatively lower response going across the detector correspond to pairs of cells 15 + 16, 31 + 32, and 47 + 48. Together with cells 0 and 63, they represent the first and the last cells of each 16 cell-wide extrusion, which makes up half of a module, which in turn makes up half of a Test Beam plane. As

was mentioned in Sec. 1.2, these cells are 3 mm narrower than the rest, resulting in fewer hits and a lower deposited energy. However, using the deposited energy divided by path length for calibration should compensate for this effect. Overall, Fig. 2.19 demonstrates 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 a result of the event selection applied to the cosmic tracks for calibration.

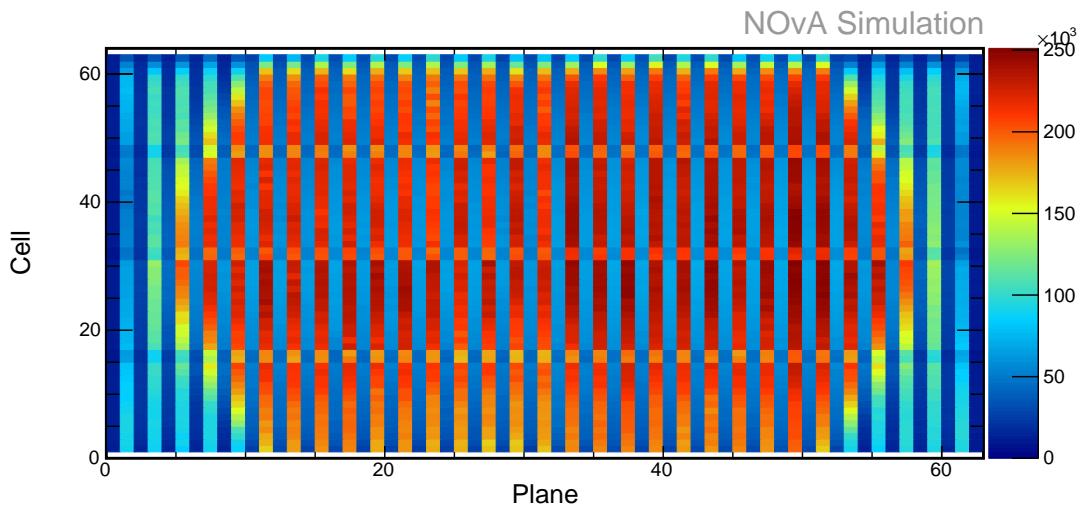


Figure 2.19: Distribution of tricell hits used for the calibration of the simulated Test Beam detector. Features are described in text.

The distributions of deposited energy per path length through the cell before the calibration in units of PE/cm as a function of w , cell and plane number, are shown in Fig. 2.20. These distributions should be uniform after applying the results of the calibration and can be used to identify the main features that will need to be corrected for during the calibration. The shallow rise of the energy response along w is caused by the attenuation of light along the [WLS](#) fibres. The drop in the response at the edges of the cell is caused by the fibres looping and connecting to the [APDs](#), while the larger statistical uncertainties at the edges of the cell reflect the lower number of hits passing the event selection including the tricell condition.

The rise of the response with the cell number, visible in the middle plots in Fig. 2.20, is due to the varying distance of cells to the readout. Since the [APDs](#) are located on one side of each module, light from cells on the opposite side has to travel along the [WLS](#) fibre for an additional module width, compared to the cells closer to the readout. Light undergoes additional attenuation along these so-called ‘pig tails’, causing

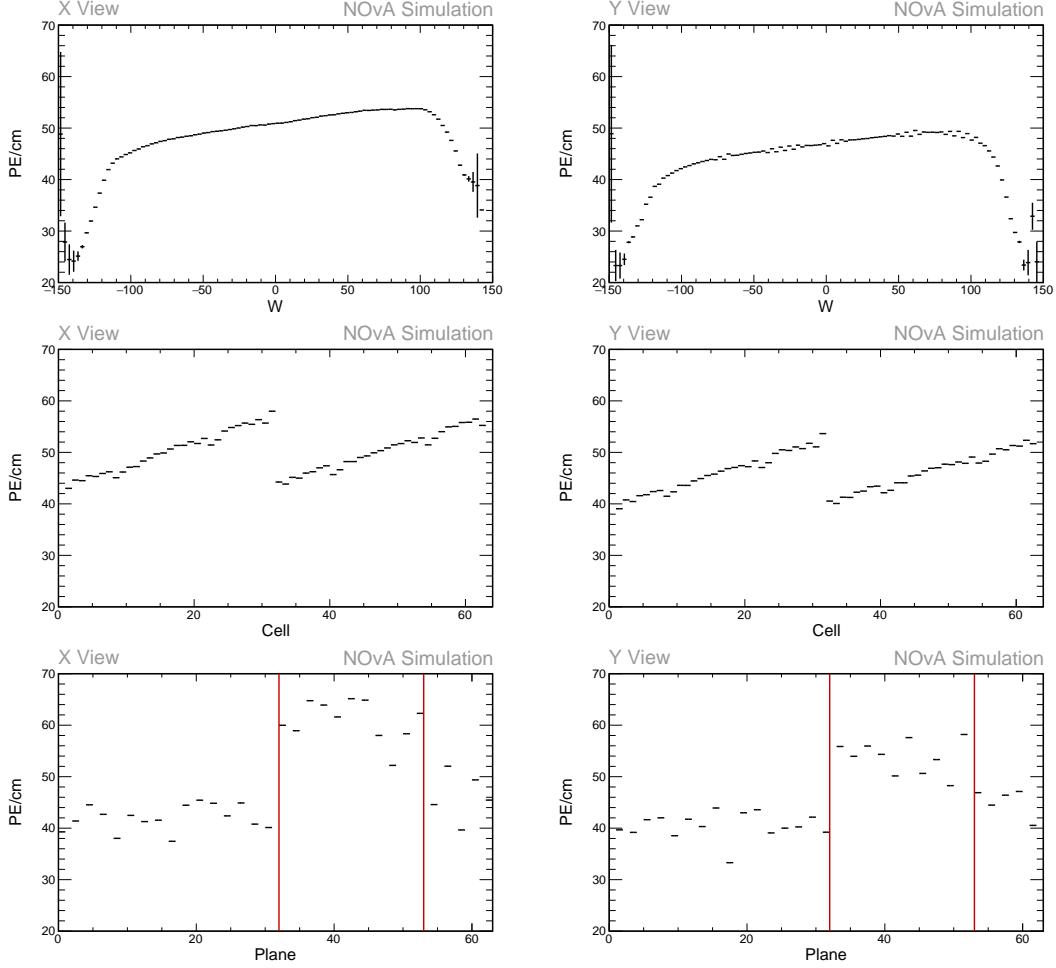


Figure 2.20: Uncorrected average energy response as a function of the position within a cell (w - top), cell number (middle), or plane number (bottom) for the Test Beam detector simulation of cosmic muon hits selected for calibration. Left side shows distributions for the X view (vertical) planes and right side for the Y view (horizontal) planes. Each plot is a profile histogram, with uncertainties representing statistical variations. Red lines on the bottom two plots depict the boundaries between different scintillators. Features explained in text.

the difference of the energy response. The additional variations across cells within a module, notably the relatively lower response in cells 0, 1, 9, 10, 23, 24, 31 and 32, is caused by including the relative gain differences into the simulation, as explained above in Sec. 2.3.2.

The uncorrected energy response as a function of plane number is shown in the bottom row of Fig. 2.20, illustrating large fluctuations between planes in both views. We can clearly identify the three distinctly different responses delineated by red lines, corresponding to the three scintillator variations used, as described in Sec. 2.1. Additionally, planes 16, 17, 48 and 49 use the FEBv5.2 instead of FEBv4.1, resulting in a

relatively lower response. All these variations between planes originate from the FB map (Sec. 2.3.1), which is used for simulation to emulate real detector conditions. The rest of the variations are caused by differences between readout electronics and individual cells, but are exacerbated by the FB binning, which groups otherwise smooth variations across planes into 12 discrete bins, thus amplifying them. **COMMENT: Not sure if the last point is clear and if it makes sense**

Simulation Relative Calibration Results

An overview of the attenuation fit results for simulation is shown in Fig. 2.21 as a map of average fitted response in the centre of each cell. Blank cells mark the uncalibrated cells which failed the calibration condition (attenuation fit $\chi^2 > 0.2$). All the uncalibrated cells but one are on the edges of the detector, which is expected, as they have much fewer events that pass the calibration sample selection. There are 43 uncalibrated cells out of the total 4032 cells in the Test Beam detector, resulting in 1.07% of the simulated detector remaining uncalibrated.

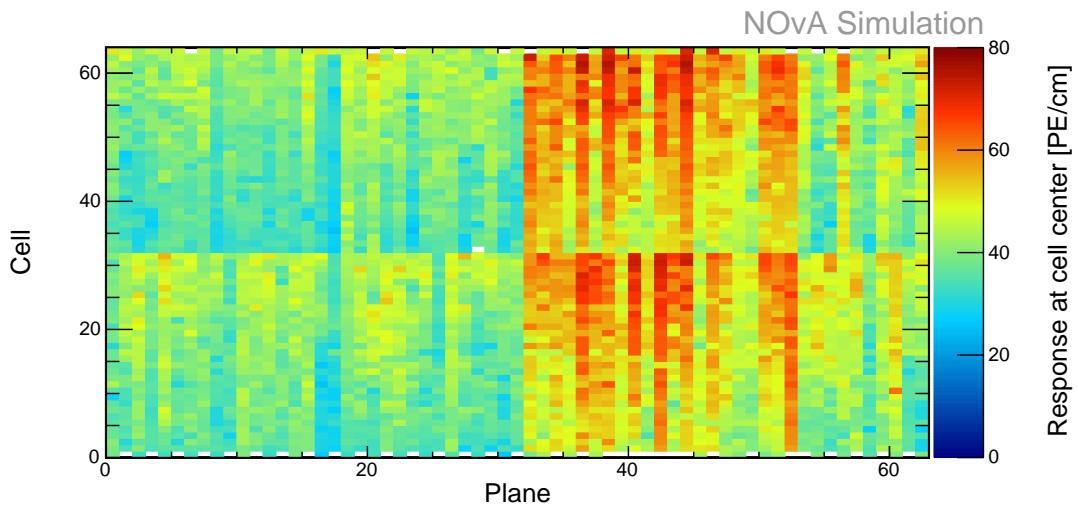


Figure 2.21: Overview of the attenuation fit results for the simulated Test Beam detector. Each cell represents the result of the attenuation fit to the energy response in the centre of that cell. The blank cells are uncalibrated as the attenuation fit did not satisfy the calibration condition.

For simulation, the attenuation fit is done for each FB bin and each cell separately. Examples of detector response for different cells in various FB bins are shown in Fig. 2.22. Here the red line shows the initial exponential fit and the blue line depicts the final attenuation fit after the LOWESS correction, as described in Sec. 1.6. 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.

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 FB bin 5, shown on the top right of Fig. 2.22, with $\chi^2 = 0.227$. Apparently, the reason the attenuation fit for this cell 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 and leave it uncalibrated **COMMENT: You are right that I should've marked this as calibrated manually as I'm doing with some other relative calibration results, but I didn't. Should I discuss it here, or rather in the discussion at the end of the chapter?**.

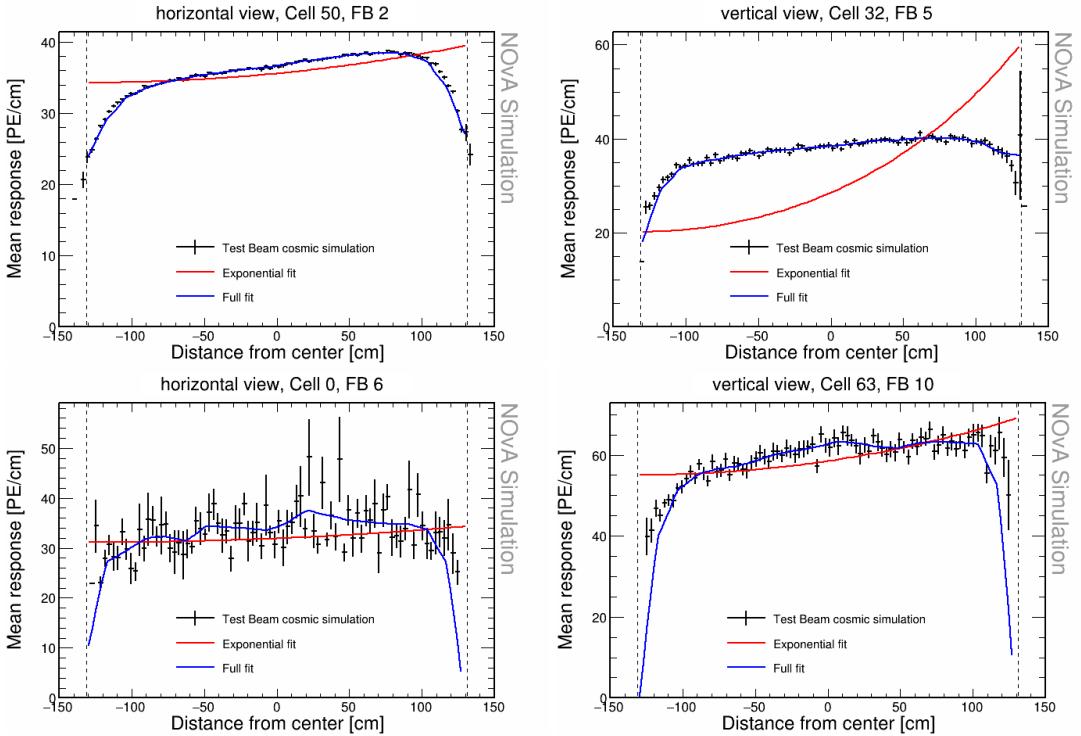


Figure 2.22: Attenuation fits for a selection of cells in various FB bins in the calibration of the Test Beam simulation. Top left is an example of a successful attenuation fit, top right is a failed fit due to statistical fluctuation in the last bin and the bottom plots show failed fits for cells on the edges of the detector.

2.3.4 Period 2 Data

The distribution of cosmic muon tricell hits selected for calibration in Test Beam period 2 data is shown in Fig. 2.23. The issue with underfilled cells described in Sec. 2.1

was present throughout period 2. The underfilled cells were marked as bad channels and therefore ignored during production of calibration samples. This also visibly affects the event count in the neighbouring cells to the underfilled cells, which have fewer calibration hits due to the tricell condition (see Sec. 1.6). However, since the underfilled cells 63 are also on the edge of the detector, labelling them as bad channels can't mitigate the effect on the neighbouring cells 62.

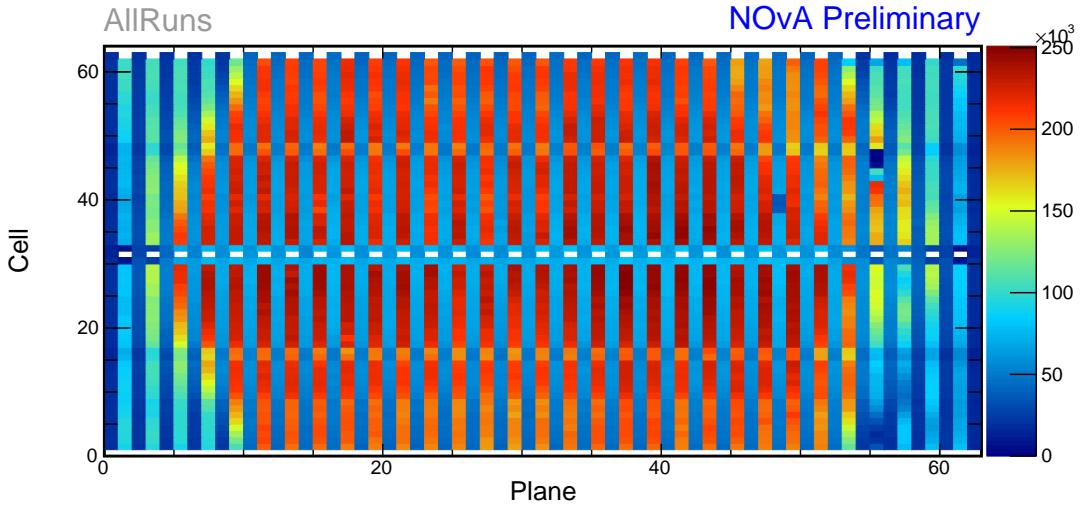


Figure 2.23: Distribution of tricell hits as a function of Test Beam detector cells and planes in the entire period 2 data calibration sample. The rows of empty cells 31 and 62 across all the horizontal planes are caused by the underfilled cells (and tricell condition), as explained in text. There are several areas with relatively fewer hits. Notably cells 38-40 in plane 48 and cells 45-47 in plane 55. Both of these spots comprise of three cells, pointing towards the middle cell being a dead channel (for a limited time) and the two surrounding cells being affected by the tricell condition. Additionally, the bottom half of planes 55 and 57 have noticeably lower number of hits than their top halves (one half corresponds to a single readout).

We can also observe areas with relatively fewer hits, likely due to channels that were dead for some time. This also affects their immediate neighbours due to the tricell condition. Additionally, there are planes that have noticeably fewer hits in one half than in the other, and since half of a plane corresponds to a single readout (one FEB and APD), which means an entire readout was faulty for a certain time.

Officially, period 2 is divided into six epochs labelled 2a - 2f, based on specific Test Beam detector running conditions. Generally, smaller calibration samples reduce time-dependent effects on calibration, such as detector ageing or temperature and humidity variations. However, smaller samples also increase the number of cells with issues in the attenuation fit (examples shown below). Therefore, it is important to

choose an optimal calibration sample size to balance both concerns. Since individual epochs in period 2 do not contain enough events for a successful attenuation fit, and variations between epochs are minimal, we decided to calibrate the entire period 2 together, without splitting it into any smaller calibration samples.

The epochs in period 2 mostly differ in the use of various [FEB](#) firmwares or in the presence of trigger studies. We compare the energy deposition during the individual epochs in Fig. 2.24. As can be seen, the difference between the energy response across the individual epochs is fairly small (within 2 %) and only in normalization, with the largest outliers seemingly epochs 2a and 2d. There is also no clear trend of energy response falling or raising with time (epoch labels are organized in time alphabetically).

The only noticeable variation of energy response across epochs in both normalization and shape can be seen on the distributions of the energy response as a function of planes, where the uncorrected response in plane 55 is noticeably higher than the rest of the period. The exact reason for this is unknown, although it is likely caused by a fault in one of the two [FEBs](#) that make up the plane readout.

Period 2 Relative Calibration Results

The results of the attenuation fit for period 2 are summarised in Fig. 2.25, showing the map of the fitted response at the centre of each cell, with blank bins representing cells that failed the calibration condition and are left uncalibrated. Summary of the relative calibration results is shown in Tab. 2.3. 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. The largest contribution to the uncalibrated cells are the peripheral cells on the edge of the detector, which contain too few events due to the tricell condition.

Most cells have the standard response, as discussed for simulation. However, some cells have one or more regions with a drop in the energy response, as shown in Fig. 2.26. These low regions are a real physical effect caused by zipped, or possibly even twisted, [WLS](#) fibres [102]. This effect is present in all the [NOvA](#) detectors. As can be seen, the attenuation fit is capable of fitting this irregular response and therefore the relative calibration corrects for this effect in data. However, zipped fibres

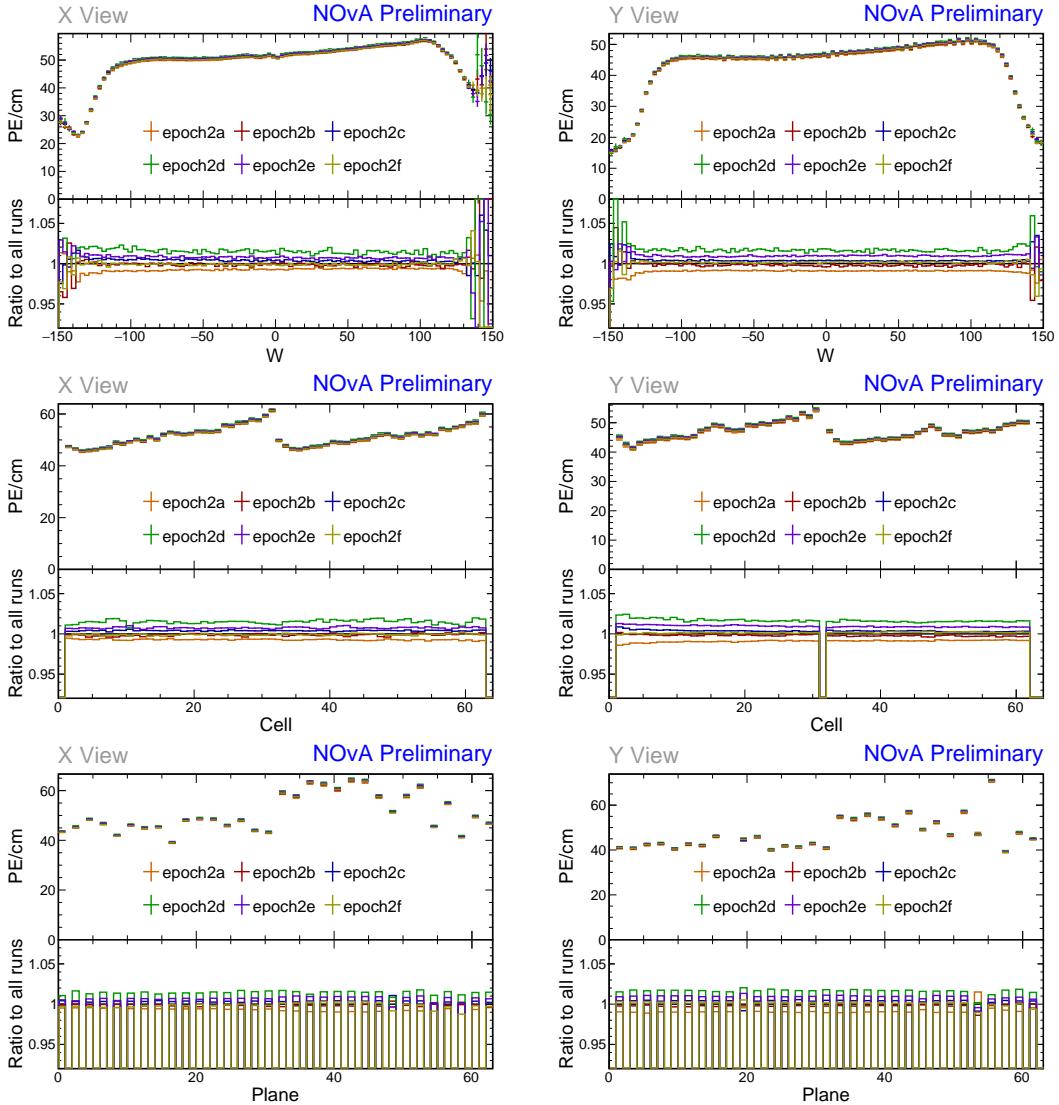


Figure 2.24: Uncorrected average energy response as a function of the position within a cell (w - top), cell number (middle), or plane number (bottom) for various epochs in the Test Beam detector period 2 data of cosmic muons hits selected for calibration. Left side shows distributions for the X view (vertical) planes and right side for the Y view (horizontal) planes. Each plot is a profile histogram, with uncertainties representing statistical variations. It is clear that there is no significant difference in shape between the various epochs. The one exception is plane 55, which has a visibly higher energy response than the rest of the planes, especially in epoch 2a, as can be seen in the bottom right plot.

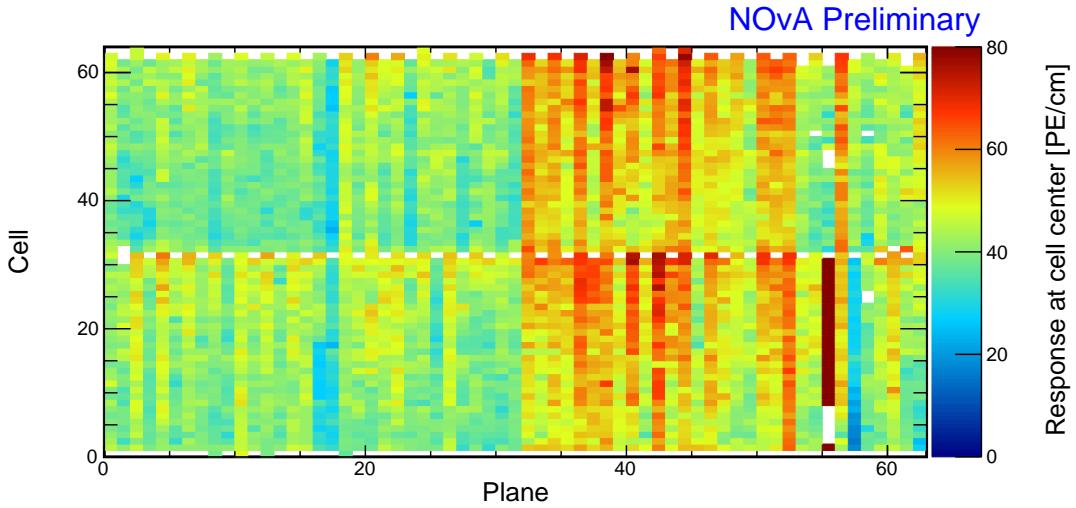


Figure 2.25: Overview of the attenuation fit results for the Test Beam detector period 2 data. Each cell represents the result of the attenuation fit to the energy response in the centre of that cell, with blank cells failing the calibration condition ($\chi^2 > 0.2$). Cells 0 and 63, which are on the edges of the detector are mostly uncalibrated due to low statistics of calibration hits. Cell 31 and 63 in horizontal planes are underfilled, showing as rows of blank cells across the detector. This affects some of their neighbouring cells, such as cells 30 and 32 in plane 1, or cells 62 in all of the horizontal planes. Cells 0-31 for planes 55 and 57 have a visibly higher (plane 55) and lower (plane 57) energy response, caused by faulty FEBs, which for some time wrongly recorded scaled response. Cells 2-4 and 45-47 in plane 55 were dead for some time during period 2, resulting in failing the calibration condition. There are a few other uncalibrated cells, which are concentrated at the end of the detector (right hand side), which failed the calibration condition due to large fluctuations at cell edges.

Table 2.3: Summary of relative calibration results for period 2 with the uncalibrated cells divided into four categories based on the main reason of failure, all described in text. **COMMENT:** *To be honest, some cells have a combination of reasons to be uncalibrated - usually both being on the edge of the detector and maybe binning, or readout, or something. Should I keep this then?*

Calibration status	Number of cells	Detector proportion
Uncalibrated	Calibrated	95.06 %
	Peripheral cells	3.00 %
	Underfilled cells	1.59 %
	Readout	0.22 %
	Binning	0.12 %

are not included in simulation for any of the detectors, which could potentially cause discrepancy from data due to the [ADC](#) threshold. It was decided that this 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.

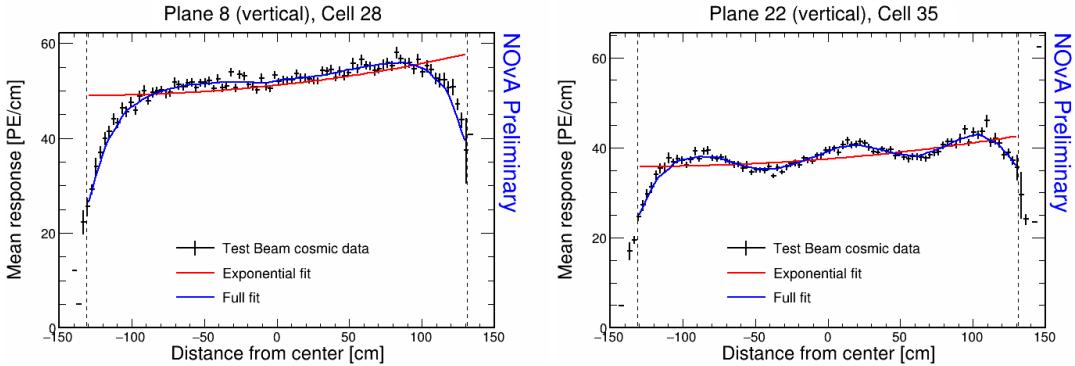


Figure 2.26: Attenuation fits for a selection of cells in period 2. Left plot shows an example of the standard energy deposition in the Test Beam and right plot shows the effect of zipped fibres.

The attenuation fits for the underfilled cells fail the calibration condition as expected. On the other hand, most of their neighbouring cells in the middle of the detector (cells 30 and 32) successfully pass the calibration condition despite having fewer events. This is thanks to the decision to label the underfilled cells as bad channels, as shown in Fig. 2.27. However, it appears that some cells neighbouring the underfilled cells near the edge of the detector have too few events to have satisfactory attenuation fits.

The effects of the issues with dead channels and with faulty readout electronics occurring during period 2, which were discussed above, can be clearly seen on the map of the attenuation fit results in Fig. 2.25 and on the attenuation fits themselves in Fig. 2.28. The (temporarily) dead channel in plane 55 contains too few events to pass the calibration condition. However, the channel in plane 48 was likely dead for a shorter duration, resulting in a successful attenuation fit, despite the lower number of hits compared to a standard cell. Cells corresponding to the entire readout affected have lower number of hits, resulting in some of them having attenuation fits failing the calibration condition. Furthermore, these cells have a strikingly different energy response, even $3\times$ larger than the average in the case of plane 55. This is due to the corresponding [APDs](#) or [FEBs](#) incorrectly recording a scaled-up or scaled-down energy response than the real energy deposited in the detector. The cause for this

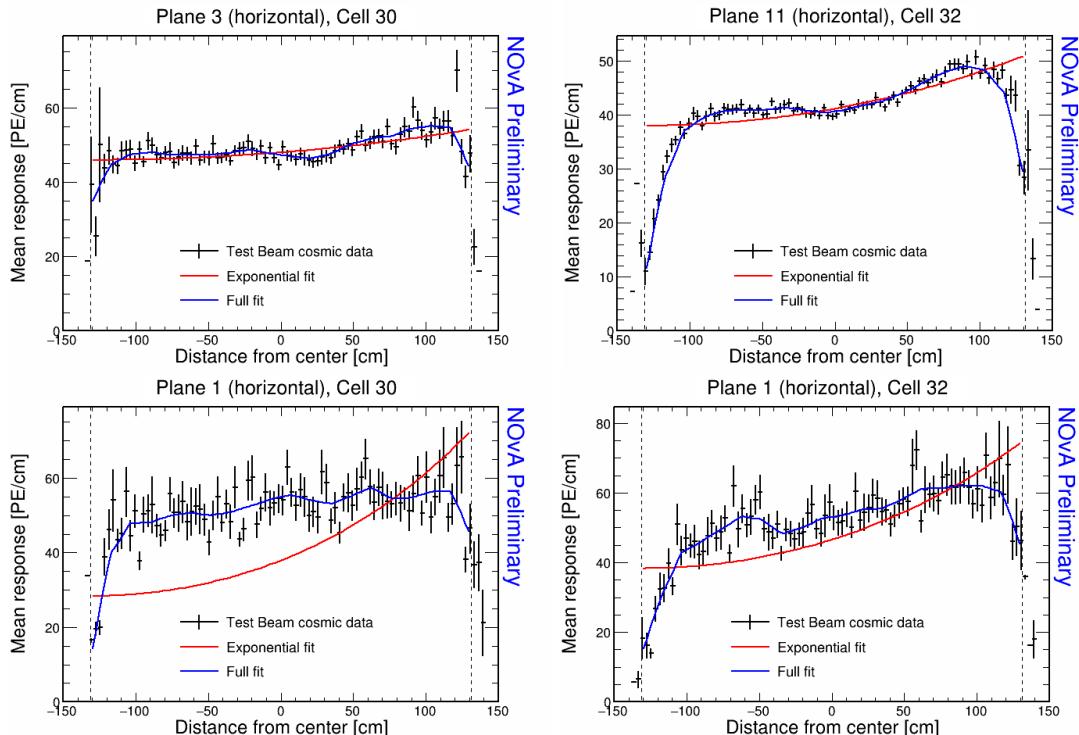


Figure 2.27: Fit to the energy response in period 2. Showing examples of cells neighbouring the underfilled cells which have fewer events and therefore larger fluctuations than the ‘usual’ Test Beam cell. Bottom two plots show examples of neighbouring cells to the underfilled cells, specifically in plane 1, which failed the calibration condition due to low statistics. This is a result of the combined effect of being a neighbour to the underfilled cell and on the edge of the detector.

scaled recorded response is not known. Since this effect is present for all data, not only for the cosmic muons used for calibration, it is important to correctly account for it in calibration. However, there is a reason for concern, as this issue can arise even if these FEBs (or possibly APDs) were only affected for a limited time out of the entire calibrated period. Since we are performing the attenuation fits on the average response across the entire 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.

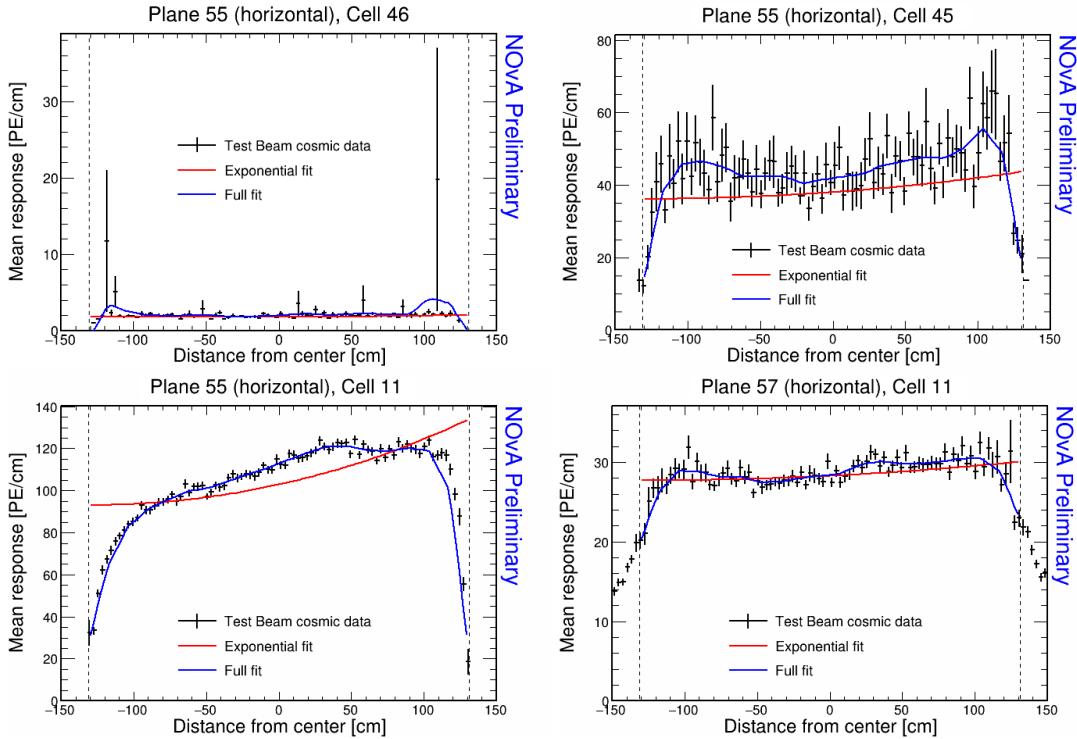


Figure 2.28: Fit to the energy response in period 2. Some channels were likely dead for some time, resulting with significantly less recorded events as shown on the top left plot. This also affect their neighbouring cells due to the tricell condition as shown on the top right. Planes 55 and 57, shown on the bottom left and bottom right plots respectively, correspond to one of the ‘faulty’ FEBs affected for some time. This results in a significantly different scale of energy response, which is much higher than the rest of the detector for plane 55, and smaller for plane 57.

An unexpected issue appeared for several cells located near the end of the Test Beam detector (relative to the beam). These cells have attenuation fits failing the calibration condition due to the unusually high response or a lack of events in histogram

bins at the edges of the cell, as shown in Fig. 2.29. This is a combination of a real physical effect - caused by fewer hits at the edge of the detector, possibly also due to the fibre loops and fibre ends - and of the choice of binning for the attenuation profiles. All attenuation profiles for all the NOvA detectors are created with 100 bins, extending beyond the physical dimensions of the detector. For example, in the Test Beam detector, the attenuation profiles range from -150 cm to 150 cm , while the actual half-length of a Test Beam cell is 131.07 cm . This means that the attenuation profile bins near the physical edges of the cell contain fewer hits from inside of the detector, resulting in larger fluctuations. Since the attenuation fits are limited to the physical cell boundaries, these bins with larger variations can skew their results. This effect can be addressed either by changing the binning of the attenuation profiles to better match the physical dimension of the cell, by loosening the calibration condition for hits on the edges of the cell, or using larger samples for the attenuation fits to reduce variations. However, since the affected uncalibrated cells are in the end of the detector, we decided to ignore them.

2.3.5 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 in Tab. 2.4. Another major event that could impact calibration is the replacement of several faulty FEBs, which motivated the creation of epoch 3e.

Table 2.4: 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

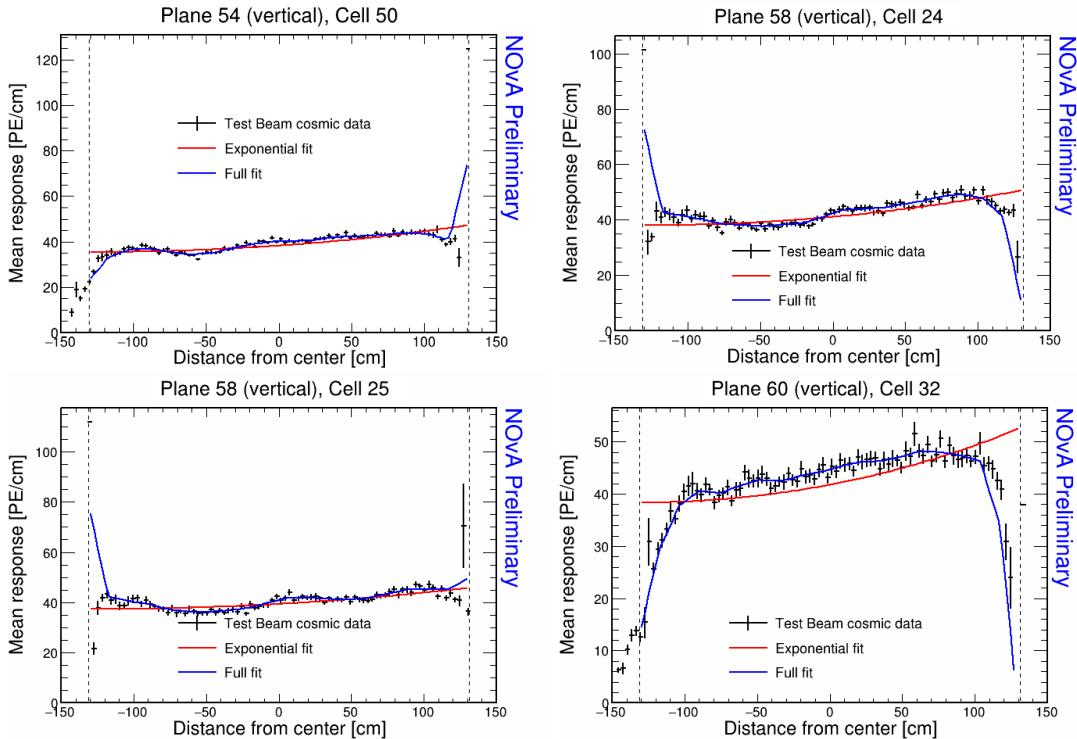


Figure 2.29: Fit to the energy response in period 2. Examples of cells that have an unusually high or low energy response at the edge of the cell, skewing the attenuation fits and resulting in them getting labelled as not calibrated. Cells shown on the two top plots and on the bottom left plot have a single bin on the edge of the fitted region (marked by dotted vertical lines) with noticeably higher average energy response. These anomalous bins typically only have a single entry that skews that attenuation fits and their χ^2 calculations. Cell 32 in plane 60, shown in the bottom right plot, has bins on the edge of the cell with no entries, resulting in the same effect as the other cells mentioned above.

The refilling of the underfilled cells can be clearly seen on the cell and plane distribution of hits in Fig. 2.30 and on the distribution of energy deposition across horizontal (Y view) cells in Fig. 2.31. The distributions of hits also shows a few channels that were dead for a certain time. Additionally, the energy deposition distributions show, that one of the FEBs was recording a scaled up/down energy response, similarly to the faulty FEBs in period 2. However, a can be seen in the distribution of hits, this particular faulty FEB recorded the same number of events as were recorded in 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 start of epoch 3e.

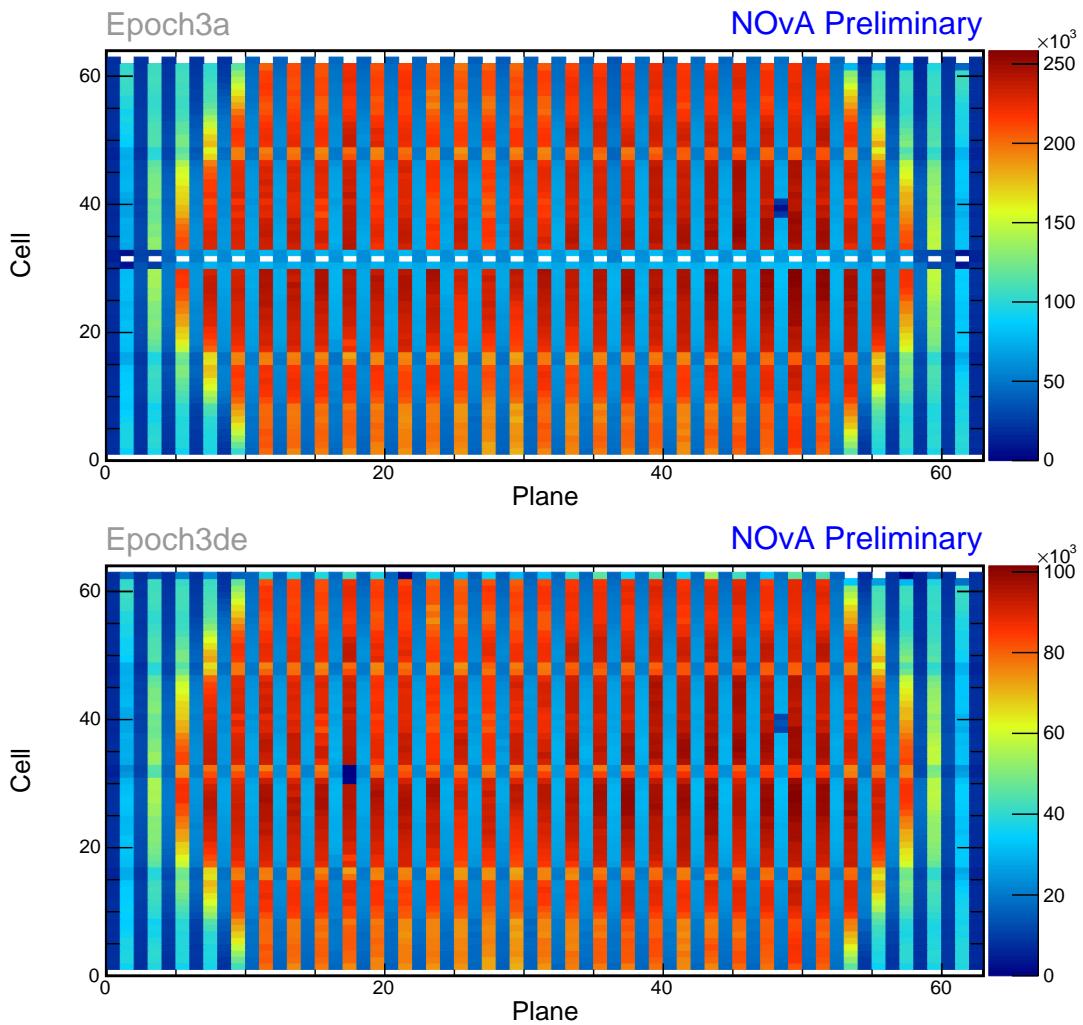


Figure 2.30: Distribution of events in the period 3 Test Beam data calibration sample. Comparison of the epoch 3a data before the refilling of the underfilled cells 31 and 63, clearly visible by a row of empty bins, and the combination of epochs 3d and 3e after the full refilling. There are also several cells that experienced readout issues, specifically cell 39 in plane 48 and cell 31 in plane 18.

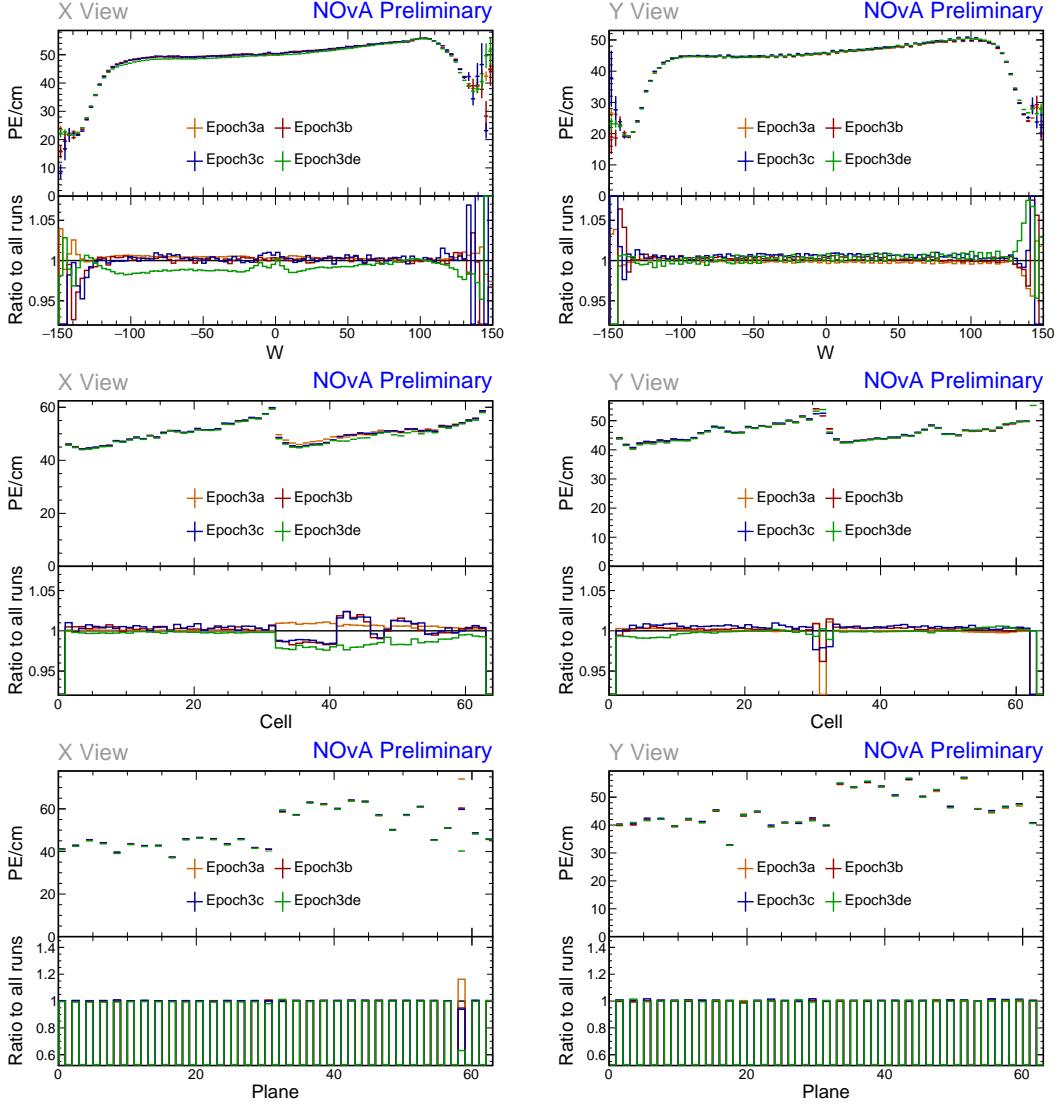


Figure 2.31: Uncorrected average energy response as a function of the position within a cell (w - top), cell number (middle), or plane number (bottom) for various epochs in the Test Beam detector period 3 data of cosmic muons hits selected for calibration. Left side shows distributions for the X view (vertical) planes and right side for the Y view (horizontal) planes. Each plot is a profile histogram, with uncertainties representing statistical variations. The effect of staged refilling of the underfilled cells between the epochs can be seen in the middle right plot, where epoch 3a (orange) has all no underfilled cells refilled, and epochs 3d and 3e (green) have all the cells filled to the top. Comparing the distributions of energy deposition in X view between the cell and plane plots, it can be seen that the top FEB/APD in plane 58, which correspond cells 32-63, was faulty throughout period 3. 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.

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. Additionally, epochs 3b and 3c contain only few days worth of data, therefore they wouldn't have enough events for successful independent attenuation fits.

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

The results of attenuation fits for the combined epochs 3a, 3b and 3c are summarised in Fig. 2.32, 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, as shown in Tab. 2.5.

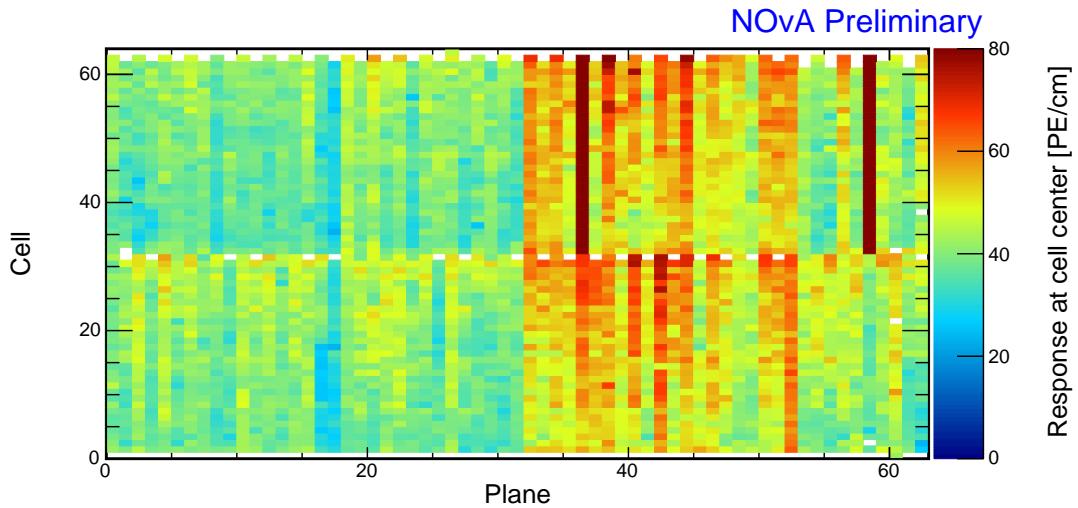


Figure 2.32: Overview of the relative calibration results for the Test Beam detector period 3, combined epochs 3a, 3b and 3c data. Each cell represents the result of the attenuation fit to the energy response in the centre of that cell. The blank bins represent uncalibrated cells. The rows of uncalibrated cells 31 and 62 are caused by the underfilled cells together with the tricell condition. The same effect affects cell 32 in plane 1. The two dark-red stripes correspond to two faulty FEBs in planes 36 and 58. There are five additional uncalibrated cells, specifically cell 2 in plane 58, cells 21 and 32 in plane 60, and cells 31 and 38 in plane 63, which are uncalibrated due to large fluctuations at cell edges.

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

Table 2.5: Summary of relative calibration results for the combined epochs 3a, 3b and 3c with the uncalibrated cells divided into four categories based on the main reason of failure, all described in text.

Calibration status	Number of cells	Detector proportion
Calibrated	3850	95.49 %
Uncalibrated	128	3.17 %
	49	1.22 %
	0	0.00 %
	5	0.12 %

the same attenuation fit. Example of energy deposition in such a cell is shown on the left side of Fig. 2.33. Same as in period 2, most of the neighbouring cells to the underfilled cells are calibrated, except for cells on the edge of the detector due to lower statistics.

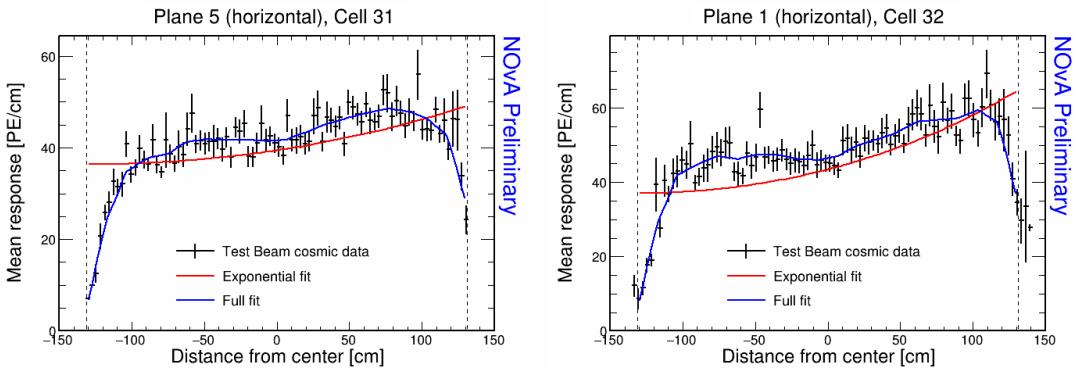


Figure 2.33: Fit to the energy response in epochs 3a, 3b and 3c. Some underfilled cells that have been refilled in epochs 3b and 3c are now calibrated as shown on the left plot. Cell 32 in plane 1 is the only neighbouring cell to the underfilled cell that didn't manage to get calibrated due to low number of events.

There is a couple of noticeably faulty FEBs with a scaled energy response, shown in Fig. 2.34. Besides the expected FEB in plane 58, which has about $5\times$ larger response, there is also the FEB in plane 36, which has about $2.5\times$ larger response compared to the average. 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 the relative calibration correction for hits in this module, during the time when the FEB wasn't faulty, would be too large (and therefore the 'corrected response' would be too small). On the other hand, during the time when the FEB was faulty, the correction would be too small and hence the corrected response would be too large. Given that plane 36 is

in the middle of the detector, there is a chance this might noticeably affect some Test Beam analysis results. Therefore, it is possible these issues might have to be mitigated in the future, whether with an additional uncertainty, or by improving the calibration. It is currently difficult to address issues such as this in the NOvA calibration. However, there is currently an effort underway to split the inputs for calibration by cells, rather than by time, which would make solving these issues much simpler. For the time being, we decided to ignore these faulty FEBs.

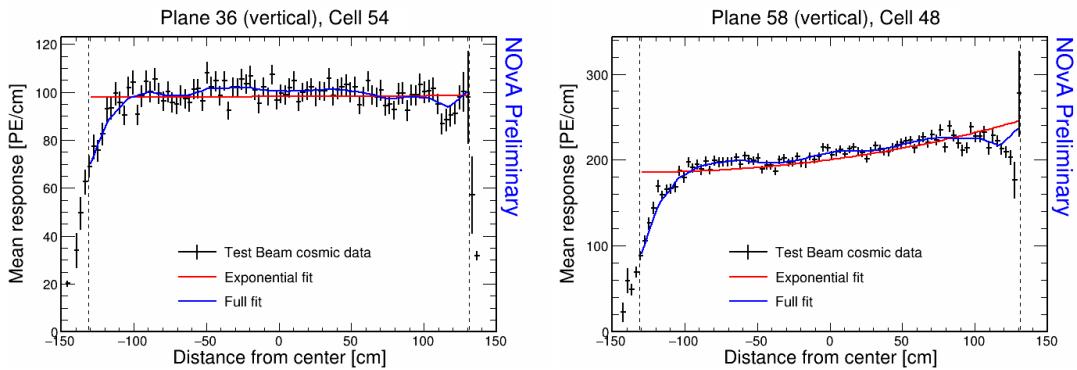


Figure 2.34: Fit to the energy response in epochs 3a, 3b and 3c. The most obvious faulty FEBs that have a significantly larger energy response than their neighbours.

Similarly to period 2, there are a few cells in the back of the detector that have a sharp rise in energy response at their edge, which causes their attenuation fit to fail the calibration condition. This can be seen in Fig. 2.35, where the significantly different mean responses at the edge bins is pulling the attenuation fit to incorrect values. Given this is concentrated in cells in the end of the detector, we decided to ignore this effect and leave these cells uncalibrated.

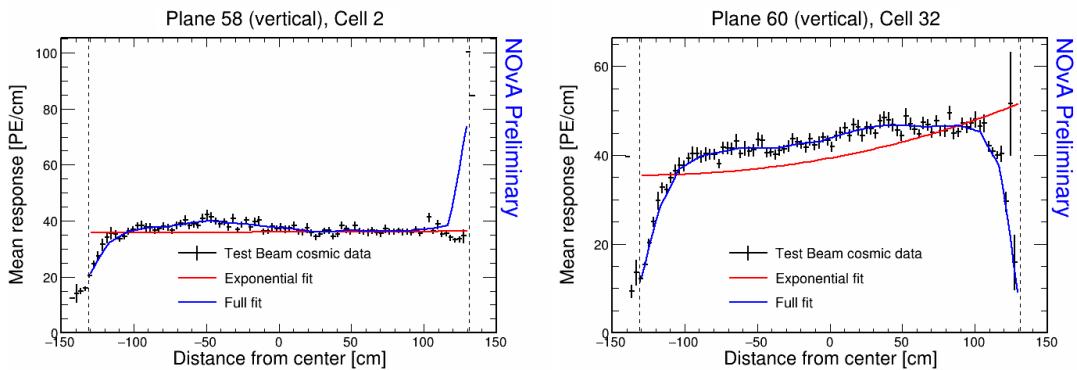


Figure 2.35: Fit to the energy response in epochs 3a, 3b and 3c. Some cells are not calibrated due to large fluctuations at one edge of the cells.

Combined Epochs 3d and 3e Relative Calibration Results

The attenuation fits results for epochs 3d and 3e are shown in Fig. 2.36. There are 182 uncalibrated cells out of 4032 total cells, making up 4.51% of the detector. The uncalibrated cells are now however almost entirely concentrated at the edges of the detector. Summary of the relative calibration results is shown in Tab. 2.6.

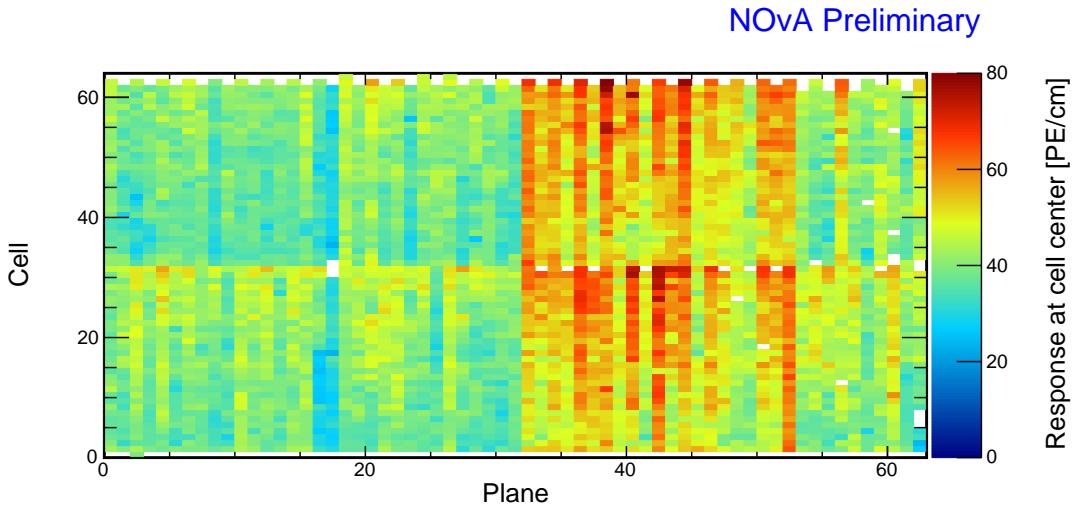


Figure 2.36: Overview of the relative calibration results for the Test Beam detector period 3, combined epochs 3d and 3e data. Each cell represents the result of the attenuation fit to the energy response in the centre of that cell. The blank cells are uncalibrated. The uncalibrated cells 30-32 in plane 17 and cells 5-7 in plane 63 are caused by a dead channel coupled with the effect of the tricell condition. The 8 previously underfilled cells 31 in planes 33, 35, 37, 41, 47, 49, 51 and 59 are uncalibrated due to the difference in the scintillator used for refilling, as described in text. There are 11 cells that are uncalibrated due to low number of events combined with the attenuation profile binning.

The expected effect of one of the two dead channels is shown in Fig. 2.37 together with some of the cells in the back of the detector, which have a rise or drop in energy deposition at their edge. This is similar to the effects seen in period 2 and epochs 3a+3b+3c and since it's again concentrated in the end of the detector, we ignore these cells and leave them uncalibrated.

Epochs 3d and 3e should have all the previously underfilled cells now refilled, but as can be seen in Fig. 2.36, there are several of these previously underfilled cells that are still uncalibrated. The energy deposition in these cells is shown in Fig. 2.38. Here we can see that these cells have a fairly large discrepancy between the left and right sides of the cell. This is caused by using different scintillator oils for the initial filling and for the refilling (or overfilling). Specifically, as was described in Sec. 2.1, these

Table 2.6: Summary of relative calibration results for the combined epochs 3d and 3e with the uncalibrated cells divided into four categories based on the main reason of failure, all described in text. Brackets show the number of cells that were originally calibrated (or uncalibrated, depending on the row) before the manual alteration of their χ^2 values, as described in text. Proportions are calculated from the final cell counts.

Calibration status	Number of cells	Detector proportion
Calibrated	3858 (3850)	95.68 %
Uncalibrated	Peripheral cells	3.13 %
	Underfilled cells	0.77 %
	Readout	0.15 %
	Binning	0.27 %

cells have been initially filled with the Ash River oil, or with the Texas oils, depending on the cell, which have a higher energy response compared to the NDOS oil that was used for their overfilling. These scintillator oils clearly did not mix properly, which caused a discrepancy in the energy deposition in different parts of the cells. This is a physical effect that should be accounted for in 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 Test Beam analyses. The large χ^2 value of the attenuation fit is most likely caused only by the unusual shape of the distribution, which the fit is not designed for. Therefore, we decided to manually change the χ^2 values for these cells inside the csv tables (which hold the results of the attenuation fits), so that their $\chi^2 < 0.2$ and these cells are officially considered calibrated when applying the calibration results, even if they originally weren't. The map of the ‘corrected’ distribution of the attenuation fit results for epochs 3d and 3e is shown in Fig. 2.39.

2.3.6 Period 4 Data

The data collected during period 4 of the Test Beam run represent our best dataset, with nearly ideal detector conditions. There were a few commissioning runs in the very beginning of period 4, which uncovered some dead channels or faulty FEBs that were immediately fixed. These initial runs constitute epoch 4a, shown on the top of Fig. 2.40. Additionally, a few runs included studies where parts of the detector

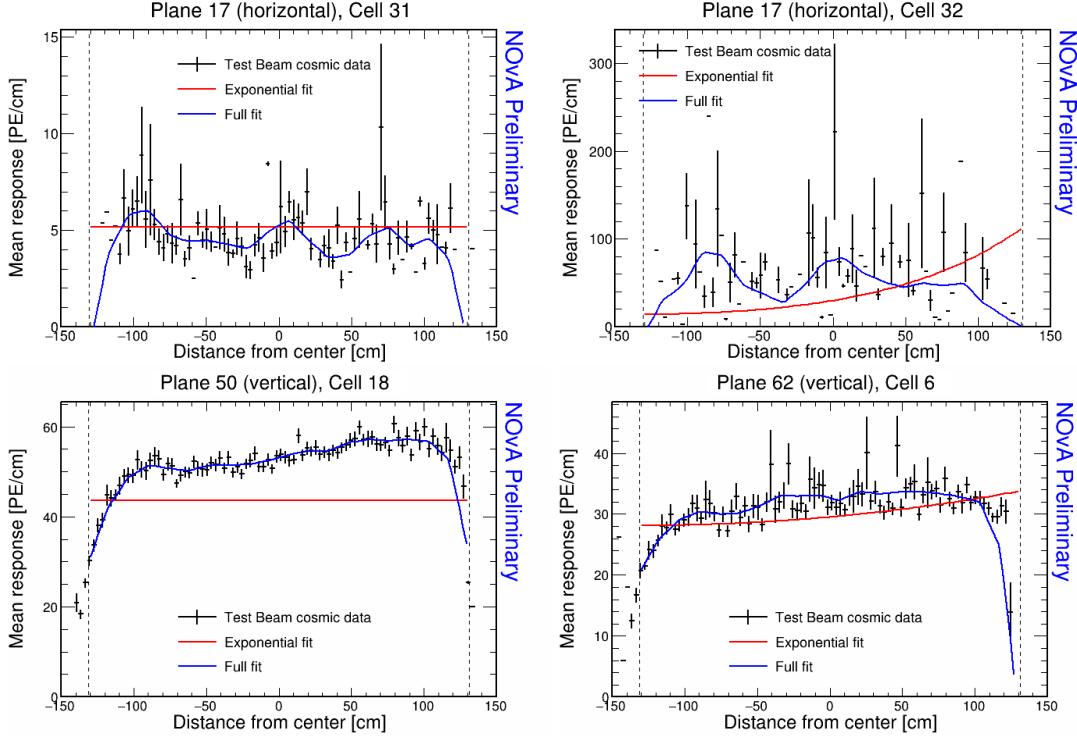


Figure 2.37: Fit to the energy response in epochs 3d and 3e. Top plots show the dead channel (left) and its immediate neighbour (right) affected by the tricell condition. Bottom plots show examples of cells with large fluctuations on their edges likely caused by low number of events combined with binning of attenuation profiles.

were masked to address FEB saturation issues [103], clearly visible in the middle of Fig. 2.40. The bottom part of Fig. 2.40 shows the remainder of period 4 data, which do not have any noticeable faults in their hit distribution across the detector.

Figure 2.41 shows, that the epoch 4a and the cell masking study had noticeable impacts on the energy deposition across the detector. Both of these special periods only span a short time and therefore contain very limited number of hits. We decided to ignore these runs and only calibrate the rest of period 4 data, using their 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. 2.42 and Tab. 2.7. We can see that almost the entire detector is now calibrated, with only few exceptions on the edges of the detector and a single cell with an unusually high response at the edge (right plot of Fig. 2.43). We treated the formerly underfilled cells the same way as in

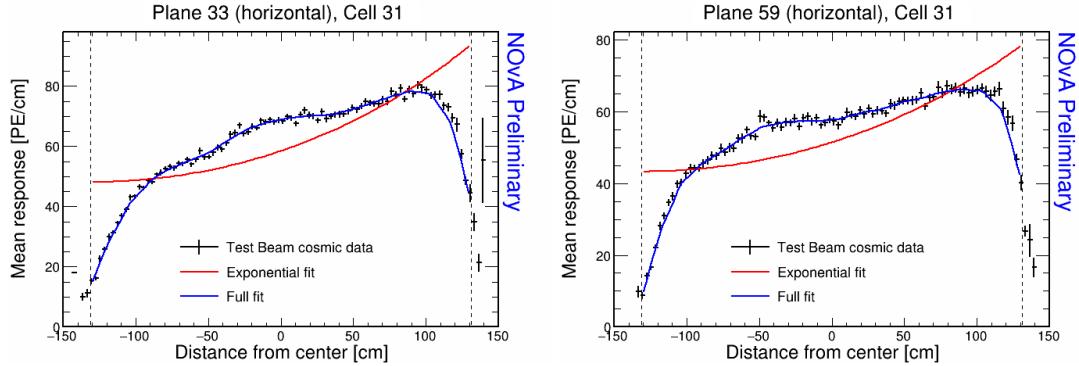


Figure 2.38: Fit to the energy response in epochs 3d and 3e. The scintillator oil used for refilling of the underfilled cells has lower energy response than the oil used for the initial filling. These oils didn't mix properly causing a different energy response in the left and right side of the cell.

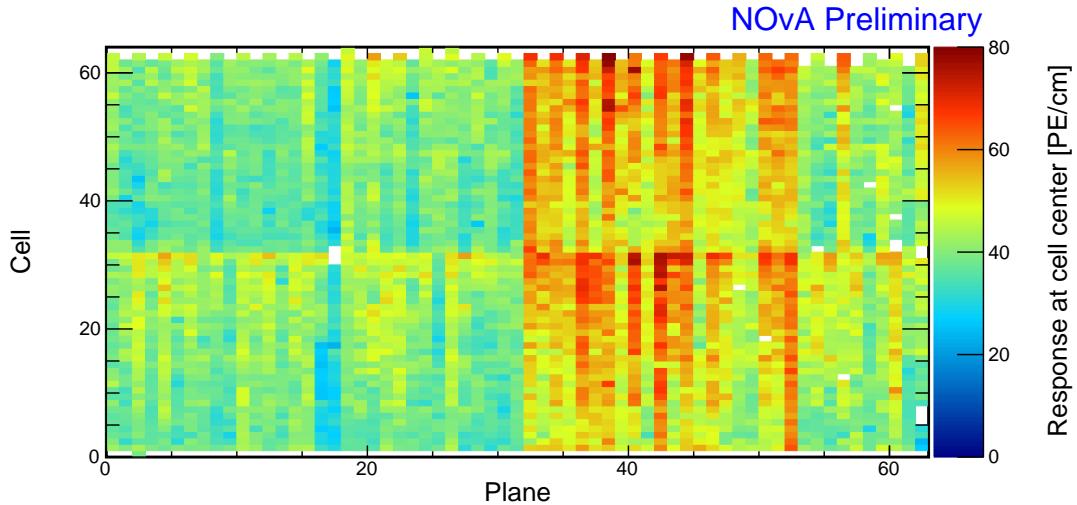


Figure 2.39: Overview of the final relative calibration results for the combined epochs 3d and 3e data after manually labelling the originally uncalibrated refilled cells as calibrated. Each cell represents the result of the attenuation fit to the energy response in the centre of that cell. The blank cells are uncalibrated and described in text.

epochs 3d and 3e, manually changing the χ^2 of their attenuation fits inside the csv files to < 0.2 , therefore making them officially calibrated. There are 108 uncalibrated cells out of 4032, totalling 2.68% of the detector.

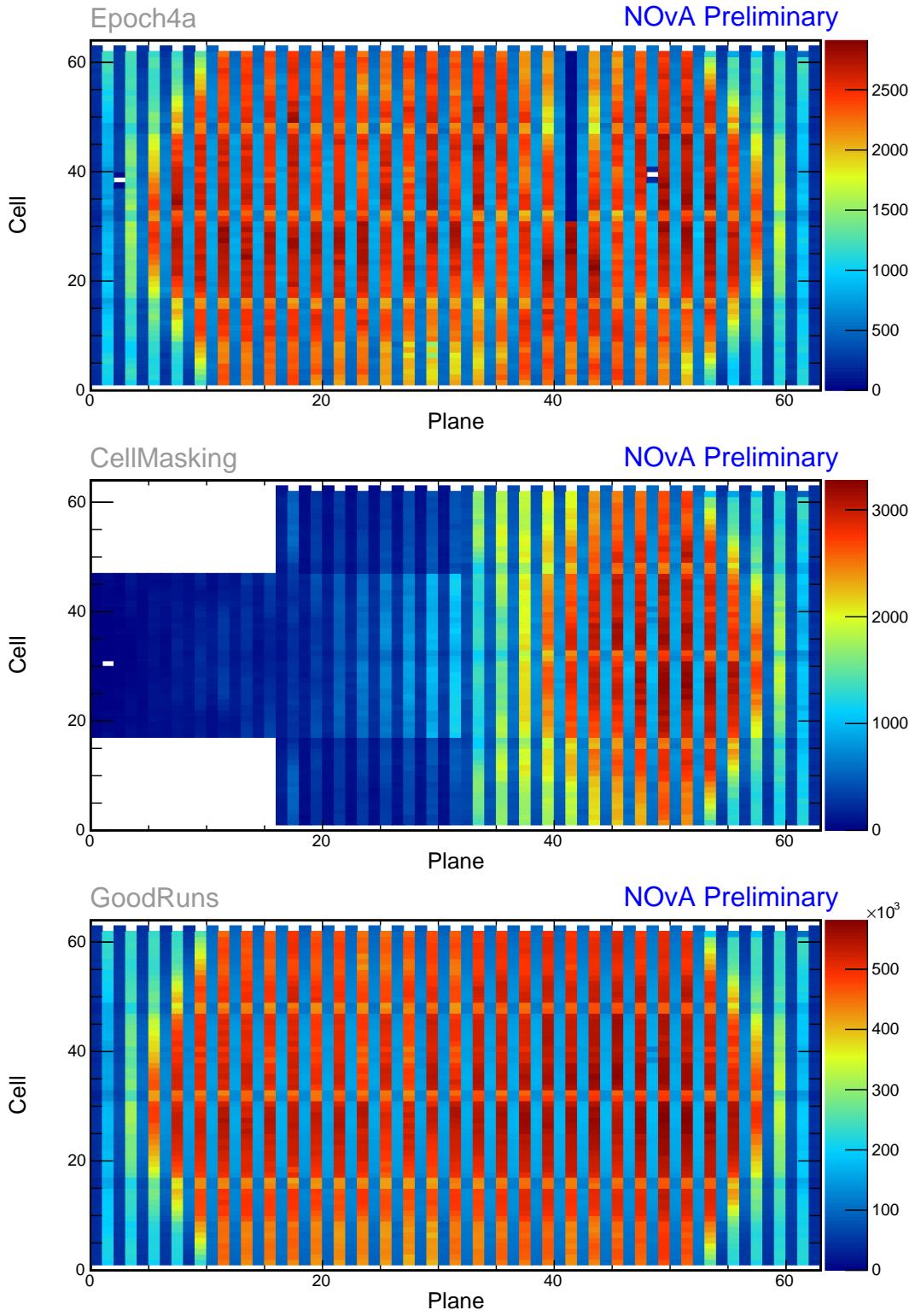


Figure 2.40: Distribution of events in the Test Beam period 4 data calibration sample. The top plot shows the first three commissioning runs with readout issues, the middle plot shows the status of the detector during the cell masking studies and the bottom plot shows the rest of the runs. Only the runs from the bottom plot (marked GoodRuns) are used for calibration.

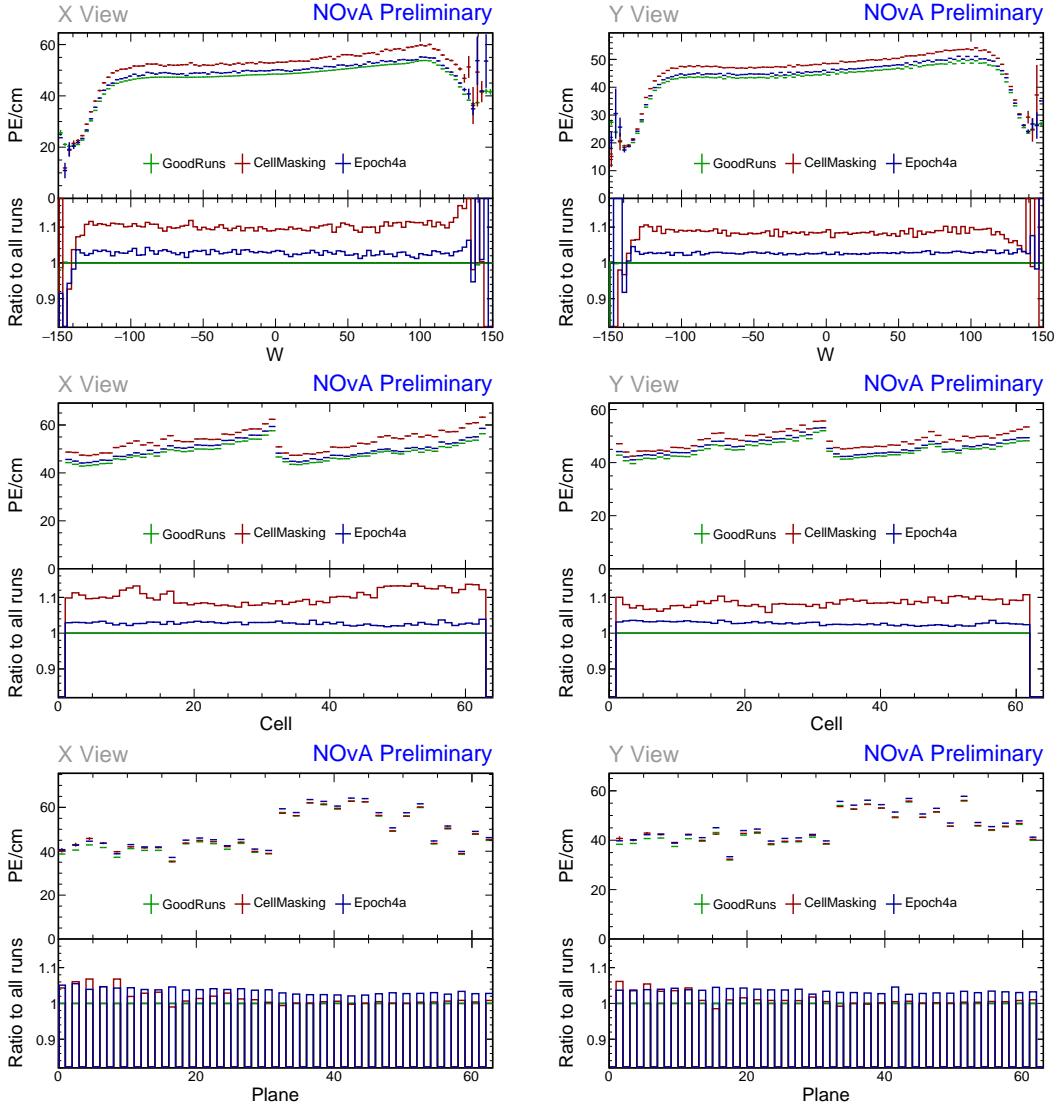


Figure 2.41: Uncorrected average energy response as a function of the position within a cell (w - top), cell number (middle), or plane number (bottom) for the Test Beam detector period 4 data of cosmic muons hits selected for calibration. Left side shows distributions for the X view (vertical) planes and right side for the Y view (horizontal) planes. Each plot is a profile histogram, with uncertainties representing statistical variations. The commissioning runs in epoch 4a and the runs during the cell masking studies have a visibly different energy deposition across all the shown variables compared to the rest of the period 4 runs.

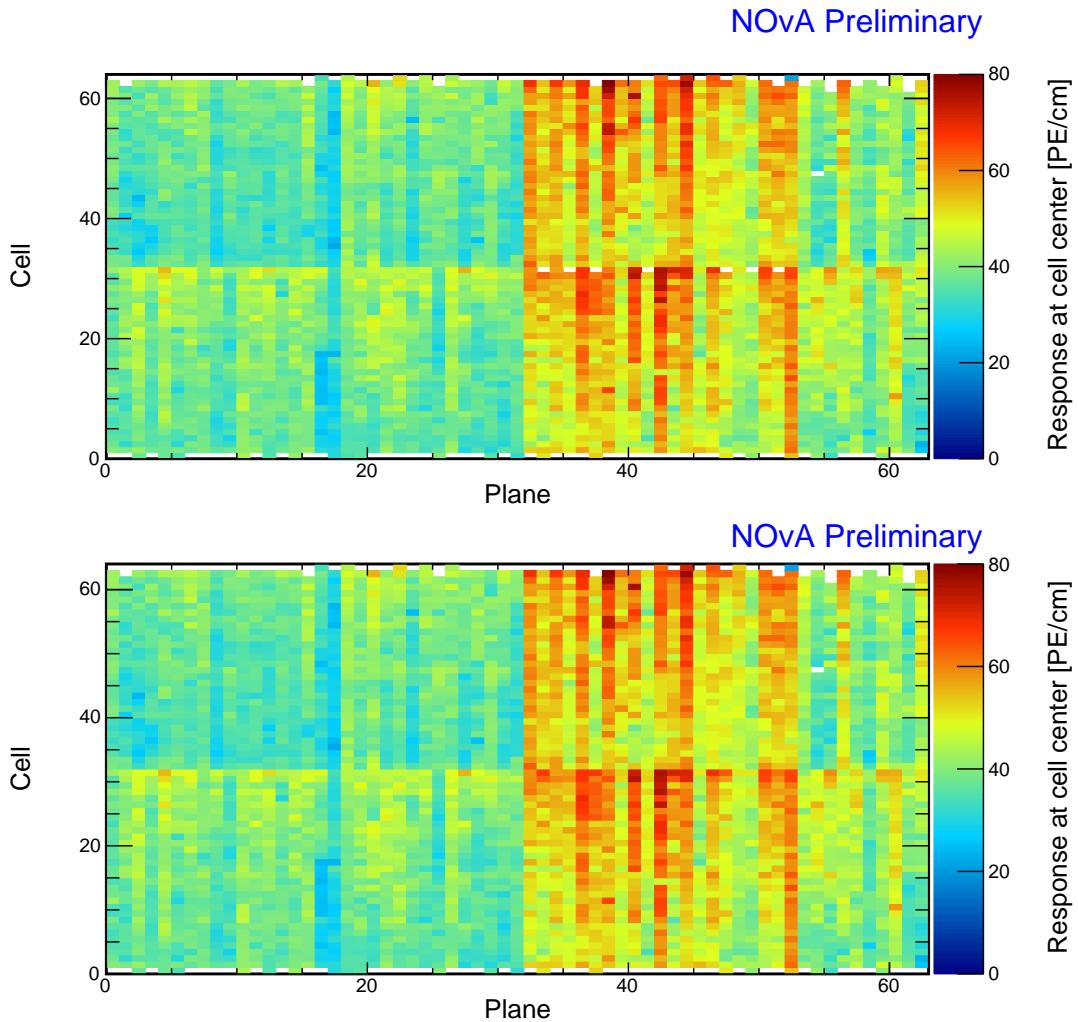


Figure 2.42: Overview of the relative calibration results for the Test Beam detector period 4 data. Top plot shows the results of the attenuation fit and bottom plot shows the final result for period 4 after manually labelling the originally uncalibrated refilled cells as calibrated. Each cell represents the result of the attenuation fit to the energy response in the centre of that cell. The blank cells are uncalibrated. The uncalibrated cells are concentrated on the edge of the detector, with a single cell 47 in plane 54 with an unusually high response at the edge of the cell. The 7 previously uncalibrated cells in the middle of the detector were artificially marked as calibrated after careful considerations.

Table 2.7: Summary of relative calibration results for period 4 with the uncalibrated cells divided into four categories based on the main reason of failure, all described in text. Brackets show the number of cells that were originally (un)calibrated before the manual alteration of their χ^2 values, as described in text. Proportions are calculated from the final cell counts.

Calibration status	Number of cells	Detector proportion
Calibrated	3924(3917)	97.32 %
Uncalibrated	97	2.41 %
	10(17)	0.25 %
	0	0.00 %
	1	0.02 %

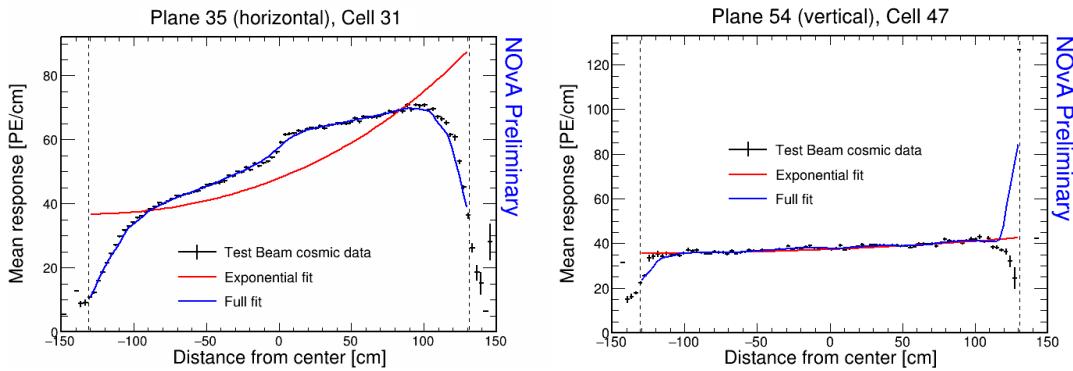


Figure 2.43: Fit to the energy response in period 4. Previously underfilled cells refilled with a scintillator of a different quality causing an unusual distribution of energy deposition (left). Unusually high energy response at the edge of the cell 47 (right).

2.3.7 Absolute Calibration Results

The results of the relative calibration (without the threshold and shielding correction) are applied to the stopping muon sample to calculate the absolute energy scale, which translates the energy response from PECorr to GeV, as described in Sec. 1.6. We apply the absolute calibration cuts to select minimum ionising muons, which represent a very well understood source of energy deposition. 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 path lengths, $\text{PE} > 0$, $\text{PECorr} > 0$, as well as $\text{PECorr}/\text{cm} < 100$. Additionally, we constrain w to a smaller allowed range: $-80 < w < 80$ cm, reflecting the smaller Test Beam cell length, removing hits approximately 0.5 m from each side of the detector.

Distributions of reconstructed and true energy responses, for both views, and for each data and simulation sample, are shown in Fig. 2.44. The mean of each of these distributions constitute the MEU_{Reco} or MEU_{True} values for both views. We calculate the statistical uncertainty on the MEU values as the standard deviation of the corresponding distributions divided by the square root of the number of entries. To combine the result from the two views, we take the average over the view-dependent MEU values to obtain the final MEU value for each sample. This is the first time in the calibration chain where the two views, which were treated completely independently so far, are combined together. The uncertainties are added in the sum of squares. The total number of entries, the MEU values for each sample and view, as well as the combined MEU values with corresponding statistical uncertainties are shown in Tab. 2.8. Given the large number of entries in the energy response distributions, the statistical 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. 1.8.

Table 2.8: Summary of absolute calibration results. MEU_{Reco} values (top table), including the statistical uncertainty $\sigma_{\text{MEU}_{\text{Reco}}}$, 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

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

As expected, the comparison of the absolute calibration results in Fig. 2.44 and Tab. 2.8 demonstrates that the MEU values across the four data samples are consistent, particularly in the Y view, which has larger statistics. However, the MEU_{Reco} values are noticeably higher for simulation than for data, especially in the X view (vertical planes). This discrepancy is anticipated, as through-going muons in the new data-based simulation (Sec. 2.2) have incorrect (smaller) incident energies, leading to smaller mean deposited energies used in the attenuation fits and consequently larger relative calibration corrections. These larger corrections are then applied to the cor-

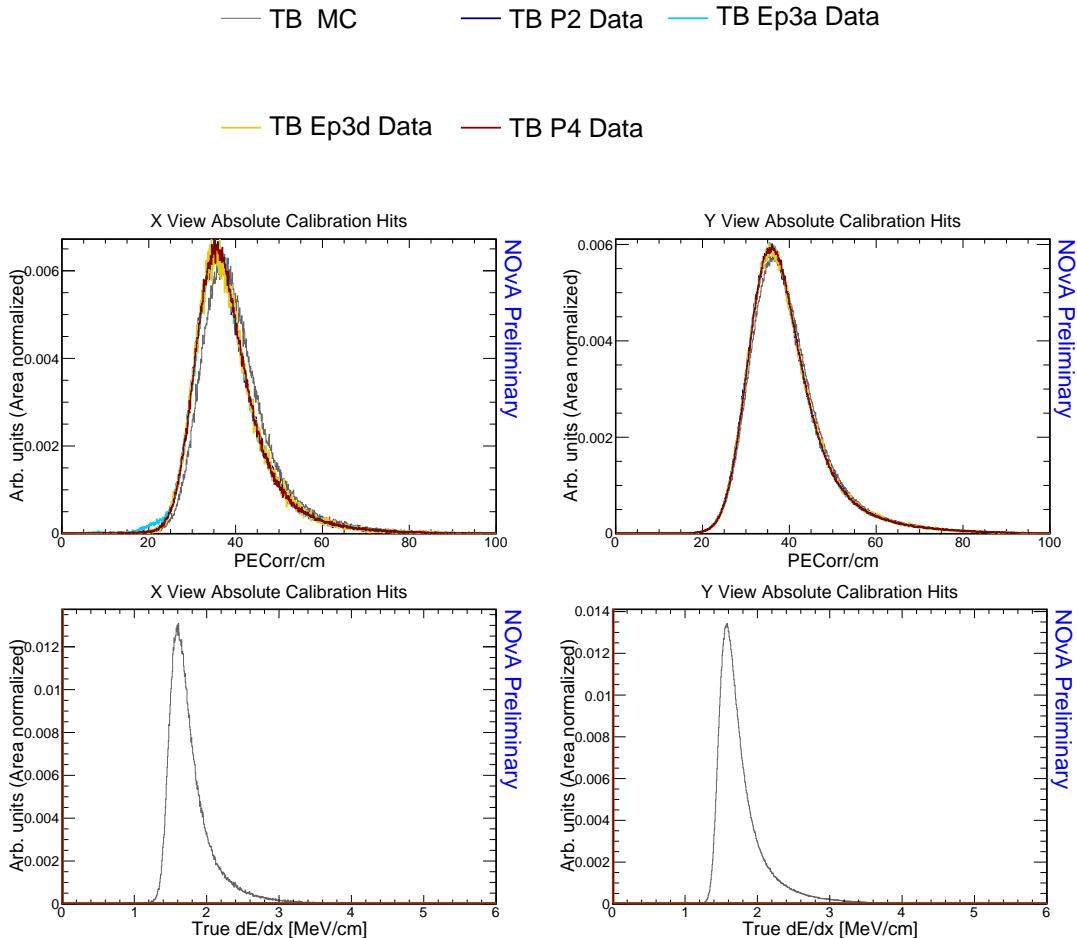


Figure 2.44: Distributions of the reconstructed (top) and true (bottom) energy response of stopping muons in the X (left) and Y (right) view within a 1 – 2 m track window from the end of their tracks. The mean of the reconstructed and true distributions of the response are the reconstructed and true MEU values respectively for the corresponding views.

rectly simulated stopping muons, resulting in higher PECorr and therefore larger MEU_{Reco} . Additionally, the true deposited energy, and therefore the MEU_{True} , which is also used for data, should be accurate.

There is a noticeable discrepancy of about 1 % in the MEU_{Reco} values between the two views. Given the minimal statistical uncertainties and the consistency of results across the samples, this is unlikely to be a random effect. Additionally, the discrepancy between the two views is in a different direction for the data samples and for simulation. The actual reason for this discrepancy is unknown; it could be due to a real difference in the stopping muon distribution between the two views that is not accounted for, or a systematic difference in the calibration treatment of the two views. This effect is observed in all NOvA detectors [104]. Since this means that the

final result is technically incorrect for both views, one possible mitigation is to apply the absolute calibration results to each view separately. However, this contradicts the logic that stopping muons can serve as a standard candle, providing a single final calibration value.

2.3.8 Validation

The initial validation of the Test Beam detector calibration is performed using the same cosmic muons that were used in the calibration process. This step is essential to ensure that the calibration performs as intended and successfully unifies the energy deposition of cosmic muons across the detector (as a function of w , cell and plane) and throughout the Test Beam detector runtime. By analysing these validation results and investigating any residual differences, we assess the stability and quality of the calibration.

After the initial validation with cosmic muons, we apply the calibration results to beam events. For the [ND](#) and the [FD](#), we use a selection of standard candles, as described in Sec. 1.8, to evaluate the performance of the cosmic-based calibration on beam events and to assess the calibration systematic uncertainties. However, Test Beam offers a unique opportunity to validate the detector calibration directly with measurements from its beamline and to reassess and potentially reduce the systematic uncertainties associated with detector calibration in [NOvA](#). This is going to be one of the main results of the Test Beam experiment, therefore, we are focusing solely on the cosmic-based validation process.

The following section highlights the most important features observed during the validation, with additional plots provided in Appendix A for the reader’s convenience.

Validation with stopping muons

First, we examine the calibration effects on the same stopping muon hits as were used for the absolute calibration (Sec. 2.3.7), including all the absolute calibration cuts. These stopping muons, which have a very reliable and stable energy deposition, are used to compare the calibration performance amongst the various data and simulation samples. However, since we require hits to be within 1 – 2 m from the end of the stopping muons’ tracks, they are not evenly distributed across the detector,

particularly absent in its bottom parts. This causes large statistical uncertainties and scattered distributions along w in the X view (for hits with $w < 0$), and across cells in the Y view (for cells < 32).

Figures 2.45-2.48 present the distributions of uncorrected (in PE/cm) and corrected (in MeV/cm) energy deposition as a function of w , cell number, plane number, and time. The uncorrected energy deposition distributions are displaying the same general attributes as were discussed for each sample in previous sections. As expected, the corrected energy deposition is generally uniform across all the studied variables. However, some residual variations can be noticed and are discussed below.

The distributions as a function of w (Fig. 2.45) illustrate the successful uniformity of energy deposition after applying calibration. Excluding the region affected by the lack of stopping muon hits, the corrected energy deposition for each sample is uniform within $\pm 0.5\%$. Additionally, all four data samples are consistent with each other, and the discrepancy between data and simulation is within $\pm 1.5\%$.

The distributions of uncorrected energy deposition as a function of w (top of Fig. 2.45) demonstrate a relative decrease in energy response over time, with period 4 data exhibiting a significantly smaller uncorrected energy response than period 2 data. This decrease, however, is corrected by calibration as expected.

The data-simulation discrepancy for the corrected energy deposition along w (bottom of Fig. 2.45) varies in the opposite direction between the X and Y views. This discrepancy arises from averaging the two view-dependent MEU_{Reco} values, which show opposite variation in simulation compared to data, as explained in Sec. 2.3.7. Ideally, there should be no data-simulation discrepancy after applying the full calibration results. Therefore, applying the view-dependent absolute calibration results separately to each view, which would likely resolve this issue, is worthwhile to consider.

The distributions of energy deposition across cells in Fig. 2.47 exhibit greater variability after calibration compared to the w dependence. This variability, particularly noticeable in the X view, is caused by issues with the threshold and shielding correction discussed in Sec. 2.3.2. The threshold and shielding correction is not applied to these distributions, just as it is not applied to the stopping muon sample for absolute calibration or to beam events, as it is not supposed to affect them. However, since the incorrect threshold and shielding correction is applied during relative calibration, it

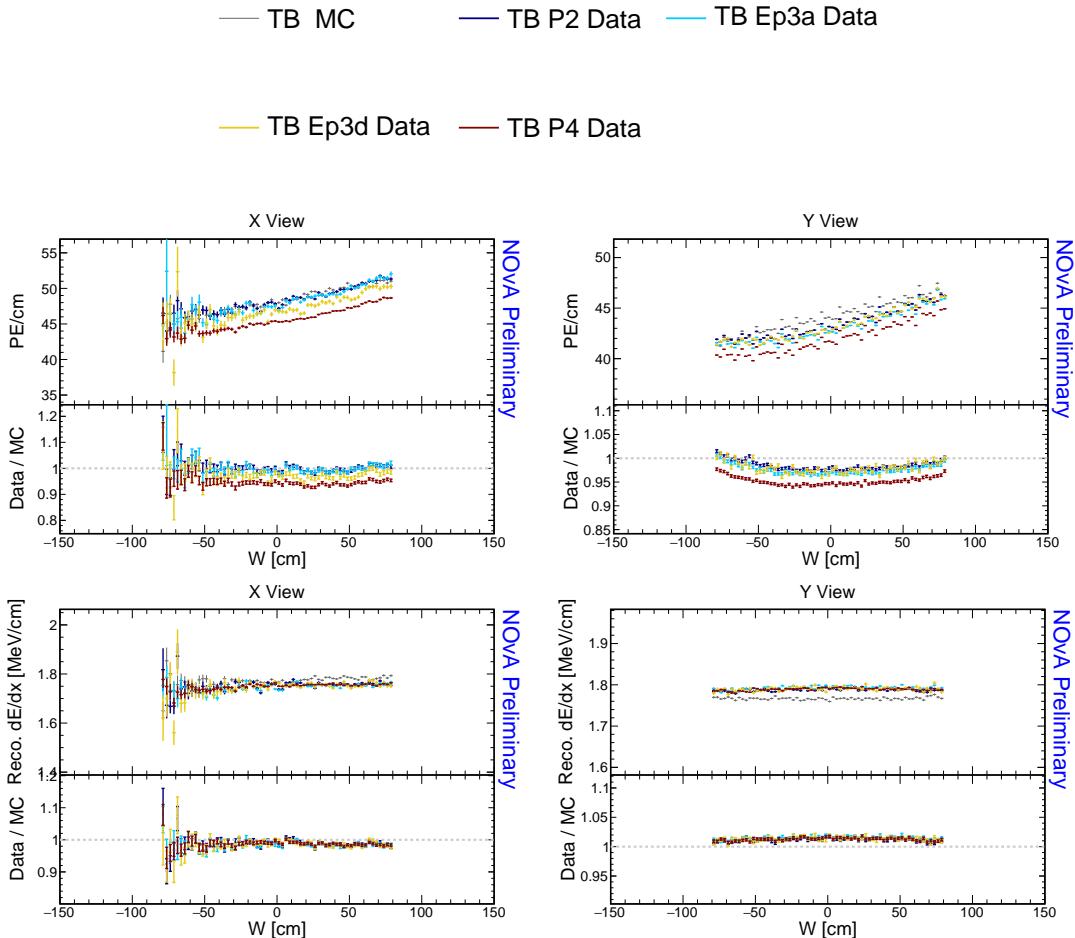


Figure 2.45: Distributions of uncorrected (top) and corrected (bottom) energy deposition for stopping MIP muons in the X view (left) and the Y view (right) as a function of the position within a cell. Bottom panel of each plot shows the ratio of the simulation sample (gray) and the four data samples, labelled at the top. Ep3a labels a combination of epochs 3a+3b+3c and Ep3d labels epochs 3d+3e. The left half ($w < 0$) of the X view distributions has large statistical uncertainties due to the low number of stopping muons at the bottom of the detector. The discrepancy between the data and the simulation samples for the corrected energy depositions is explained in text.

is incorporated into the calibration results and consequently into the reconstructed deposited energy. The variability introduced by these faulty corrections is within $\pm 3.5\%$.

The distribution of corrected energy across cells in the Y view (bottom right of Fig. 2.46) additionally shows two cells with a noticeably lower energy response ($\sim 2-4\%$) for period 2 and epochs 3a+3b+3c compared to the rest of the samples. Specifically, this is the cell 31, which was underfilled during period 2 and epoch 3a, and its neighbouring cell 32. The variation for the underfilled cell is expected. However, it is unclear why one of the neighbouring cells to the underfilled cells is miscalibrated.

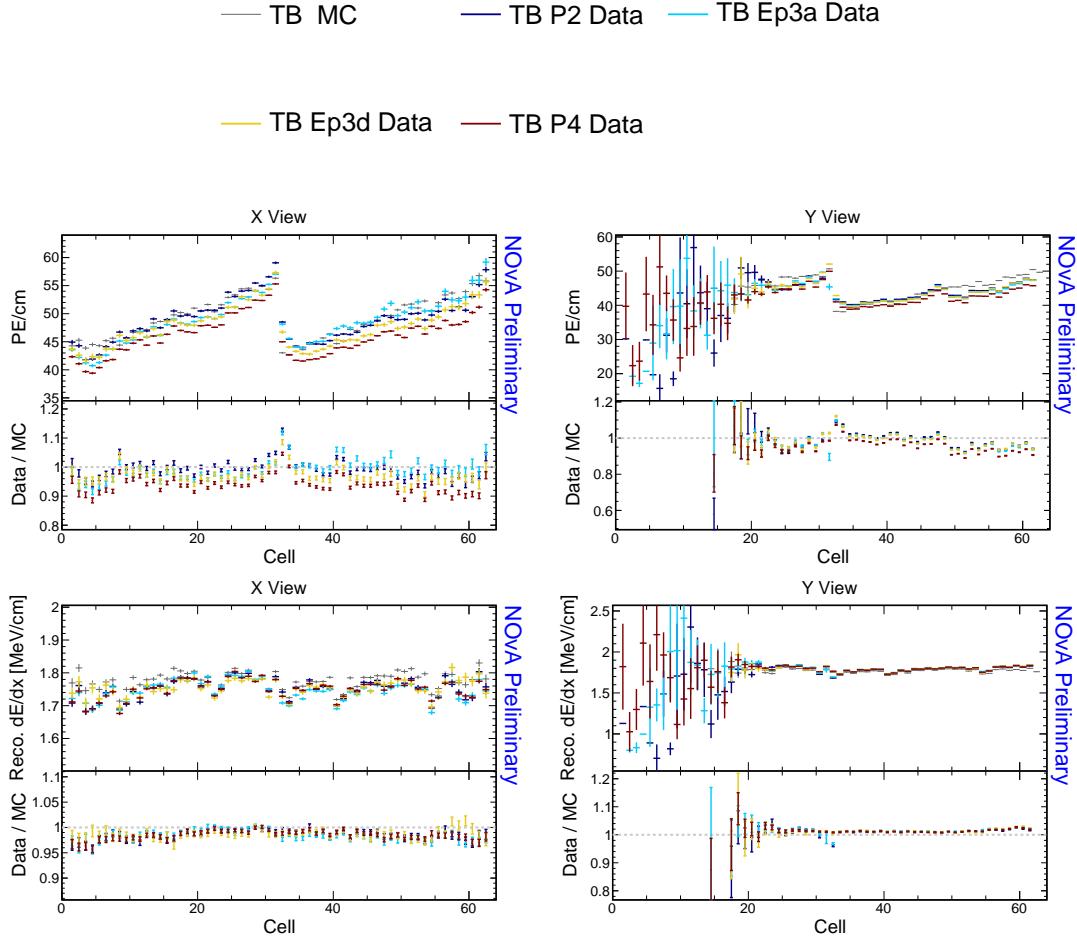


Figure 2.46: Distributions of uncorrected (top) and corrected (bottom) energy deposition for stopping MIP muons in the X view (left) and the Y view (right) as a function of the cell number. Bottom panel of each plot shows the ratio of the simulation sample (gray) and the four data samples, labelled at the top. Ep3a labels a combination of epochs 3a+3b+3c and Ep3d labels epochs 3d+3e. The left part (cell $\lesssim 25$) of the Y view distributions has large statistical uncertainties due to the low number of stopping muons at the bottom of the detector. Features described in text.

COMMENT: *Is this enough to say like this?*

The relative differences in the uncorrected energy response across planes (top of Fig. 2.47) between the three different data taking periods demonstrate that the decrease in energy deposition over time varies between planes. This variation is especially noticeable in the X view plot, but is equally present in the Y view planes. This effect is attributed to the differences in the quality of the scintillator oil used to fill the detector, as explained in Sec. 2.1. Additionally, it indicates scintillator ageing, which results in a decrease of the scintillation light produced per deposited energy over time. Specifically, the ND+NDOS scintillator oil (planes 0-31) appears to have

aged the most between periods 2 and 4, followed by the Texas NDOS oil (planes 53-62), while the Ash River oil (planes 32-52) aged the least. However, more quantitative studies are necessary to assess the ageing of the NOvA scintillator, as its details are currently not well understood.

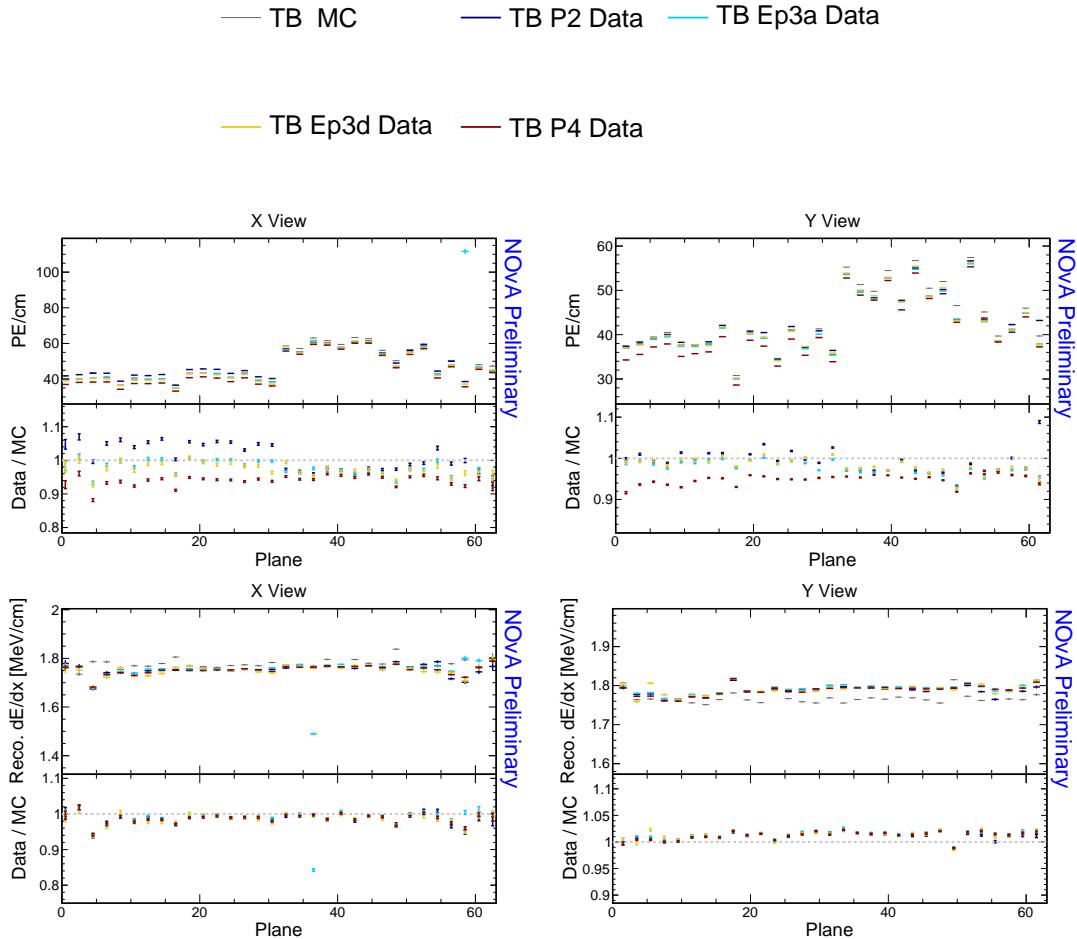


Figure 2.47: Distributions of uncorrected (top) and corrected (bottom) energy deposition for stopping MIP muons in the X view (left) and the Y view (right) as a function of the plane number. Bottom panel of each plot shows the ratio of the simulation sample (gray) and the four data samples, labelled at the top. Ep3a labels a combination of epochs 3a+3b+3c and Ep3d labels epochs 3d+3e. Features described in text.

The distribution of the corrected energy response across planes in the X view (bottom left of Fig. 2.47) reveals a significantly smaller response ($\sim 16\%$) for epochs 3a+3b+3c in plane 36. This indicates that the relative calibration over-corrected the energy response due to through-going muons having an unusually high energy response (as shown in Fig. 2.34), but not the selected stopping muons. The most likely explanation is that the affected FEB was ‘faulty’ only for a certain period. Consequently, the corrected energy response would be accurate for the period when the

FEB was faulty, but would be under-estimated for the period when the FEB functioned normally.

We can observe the effect of different FEB versions in both the corrected and uncorrected energy response distributions across planes, shown in Fig. 2.47. Specifically, the FEBv5 was used in planes 16, 17, 48 and 49, which is evident by the relatively larger corrected energy deposition, especially in the simulation and in the X view. Although this effect should be corrected during calibration, it is not, due to the incorrect incorporation of the variations between the FEB versions in the simulation and its impact on the threshold and shielding correction, as discussed in Sec. 2.3.2.

There are additional variations in the corrected energy deposition across planes (bottom of Fig. 2.47). In the X view, there is a significant ($\sim 6\%$) drop in corrected energy response in planes 4 and 58, along with a smaller drop in some surrounding planes. The cause of this variation is unknown. However, it appears to be due to a discrepancy between through-going and stopping muons, as it disappears when analysing through-going muons, as explained below. In the Y view, the corrected response increases with planes within the first half of the detector. This effect is only observed in data and not in simulation. The origin of this slope is unclear, however it is also absent when examining variations for through-going muons. **COMMENT:**
Should I have found this out?

The distributions of energy deposition over time (Fig. 2.48) reveal a complex dependency and significant variations. As shown, the calibration process currently only normalizes the mean of each calibrated sample, leaving time-dependent variations within the samples uncorrected. The uncorrected energy response is noticeably higher for period 2 compared to period 4, likely due to detector ageing, as previously discussed. However, there is no clear downward-going trend; the variations are larger than any consistent time-dependent pattern, except for the second half of period 4.

In the ND and FD calibration, these time-dependent variations are partially addressed by reducing the size of the calibration samples to month-long epochs. This approach reduces variations in the calibrated energy deposition but generally results in larger portions of the detector being uncalibrated. However, this did not have a significant effect on the rate of calibrated cells for the ND and FD [105] and in the future can be explored for Test Beam detector as well. The ND and FD typically exhibit

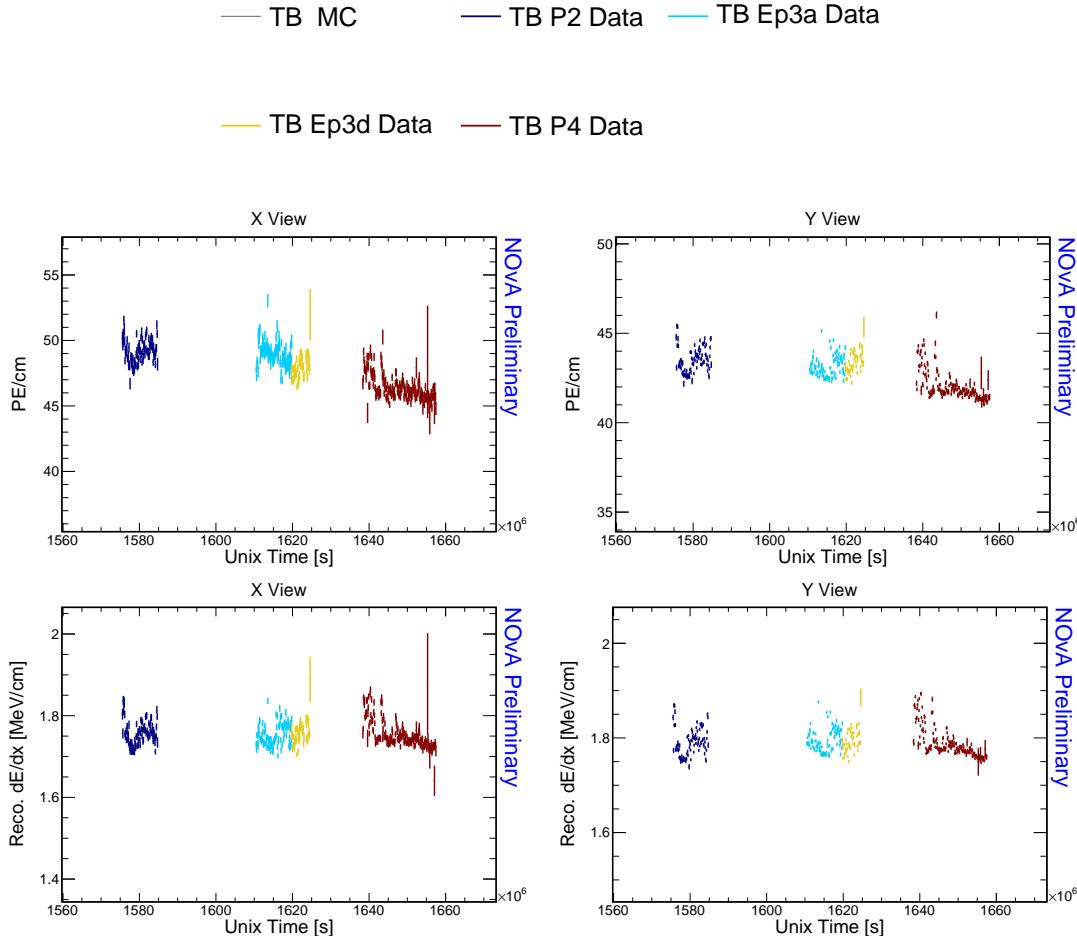


Figure 2.48: Distributions of uncorrected (top) and corrected (bottom) energy deposition for stopping MIP muons in the X view (left) and the Y view (right) as a function of the event UNIX time (which starts at the beginning of the NOvA data taking). Comparing the four data samples, labelled at the top. Ep3a labels a combination of epochs 3a+3b+3c and Ep3d labels epochs 3d+3e. Features described in text.

smaller variations in energy response over time compared to the Test Beam detector due to the Test Beam detector's less stable running conditions, such as larger fluctuations in temperature and humidity in the Test Beam hall.

The effect of these environmental factors on detector performance, along with scintillator and potential readout electronics ageing, are not well understood within NOvA. As these effects are more pronounced in the Test Beam detector and given the range of scintillator oils and readout electronics used, separating the effects of the individual factors is challenging and is currently the focus of ongoing studies [106].

Validation with the through-going muons

To validate the calibration performance without the limitations imposed by stopping muons, we examine the effects of calibration on through-going muons, which were also used for the relative calibration. Unlike for the stopping muon sample, we apply the threshold and shielding corrections for the through-going muons to truly verify the validity of the relative calibration process. Since the simulated through-going muons have incorrect incident energies, this validation focuses solely on variations in shape rather than scale.

Considering the variability of detector performance over time, we use through-going muons collected during ‘stable runs’ in period 4, as depicted in Fig. 2.49. This period was selected by visually inspecting the time dependence of corrected energy deposition and choosing runs that correspond to seemingly stable energy deposition.

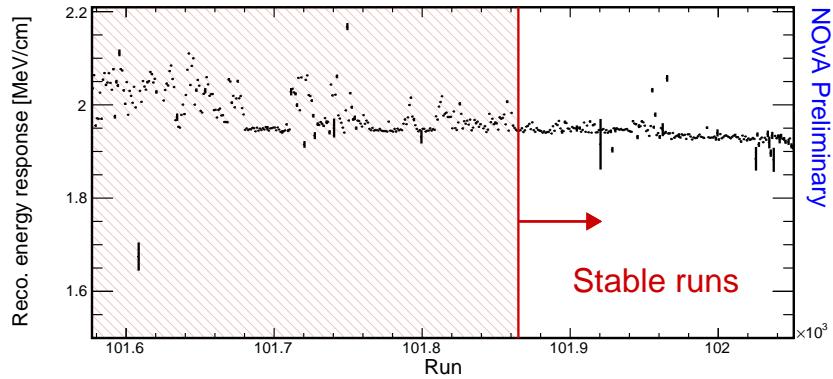


Figure 2.49: Selection of stable runs from the period 4 Test Beam data. Showing the reconstructed energy response as a function of the run number for through-going cosmic muons. Shaded red area shows rejected runs.

Furthermore, Fig. 2.50 shows significant variation of energy deposition at cell edges. This issue is addressed by the calibration shape systematic uncertainty discussed in Sec. 1.8. Therefore, for this validation study we ignore the edge variations, using the same w limit as for the stopping muon sample: $-80 < w < 80$ cm.

With the aforementioned constraints, we expect uniform distributions with limited variations. Residual variations should be scattered throughout the detector without any particular pattern, indicating potential errors in the calibration process. It is worth noting that investigating these distributions without applying the threshold and shielding correction led to the discovery of the issues discussed in Sec. 2.3.2. Any

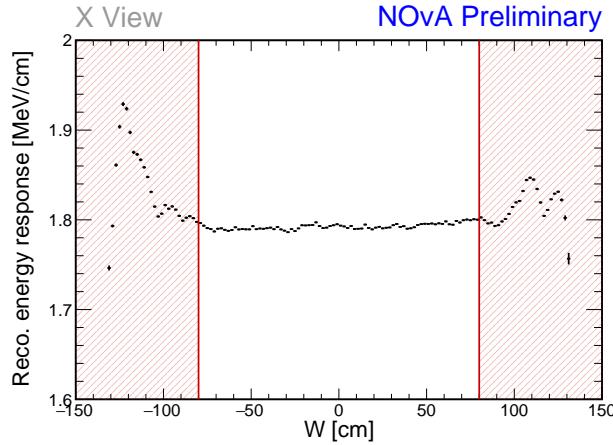


Figure 2.50: Removing edge hits for validation with through-going muons. Showing only events from the stable runs of the period 4 Test Beam cosmic data.

remaining variations point to potential systematic uncertainties in the calibration process.

Figure 2.51 displays individual distributions of corrected energy response as a function of w , plane, and cell number for both data and simulation. For through-going muons with the threshold and shielding correction applied, the calibration procedure works as expected, with residual variations around $\pm 0.2\%$ for the w and cell dependence. However, for plane dependence, these residual variations are larger: approximately $\pm 0.5\%$ in data and up to $\pm 1\%$ in simulation. Additionally, maps of corrected energy deposition as a function of both plane and cell numbers, along with their 1D projections, are shown in Fig. 2.52 for data and in Fig. 2.53 for simulation. These plots demonstrate that the final residual variations for calibration reach approximately 1% in data and 4% in simulation.

In the distributions for data, we observe a clear effect of the different scintillators used, which is related to varying levels of detector ageing among the scintillators, as previously discussed. This effect is particularly noticeable in this validation sample due to the selection of events from ‘stable runs’, while the relative calibration was calculated for the entire period 4. Therefore, calibration scales the deposited energy with respect to the mean deposited energy for the full period 4. Consequently, planes 0-31 appear to have the lowest relative corrected response in Fig. 2.51 and 2.52 because the corresponding ND+NDOS scintillator oil aged the most compared to the mean. This issue could be addressed by splitting period 4 into smaller epochs for calibration.

For simulation, the relative calibration is performed in bins of w for each cell and

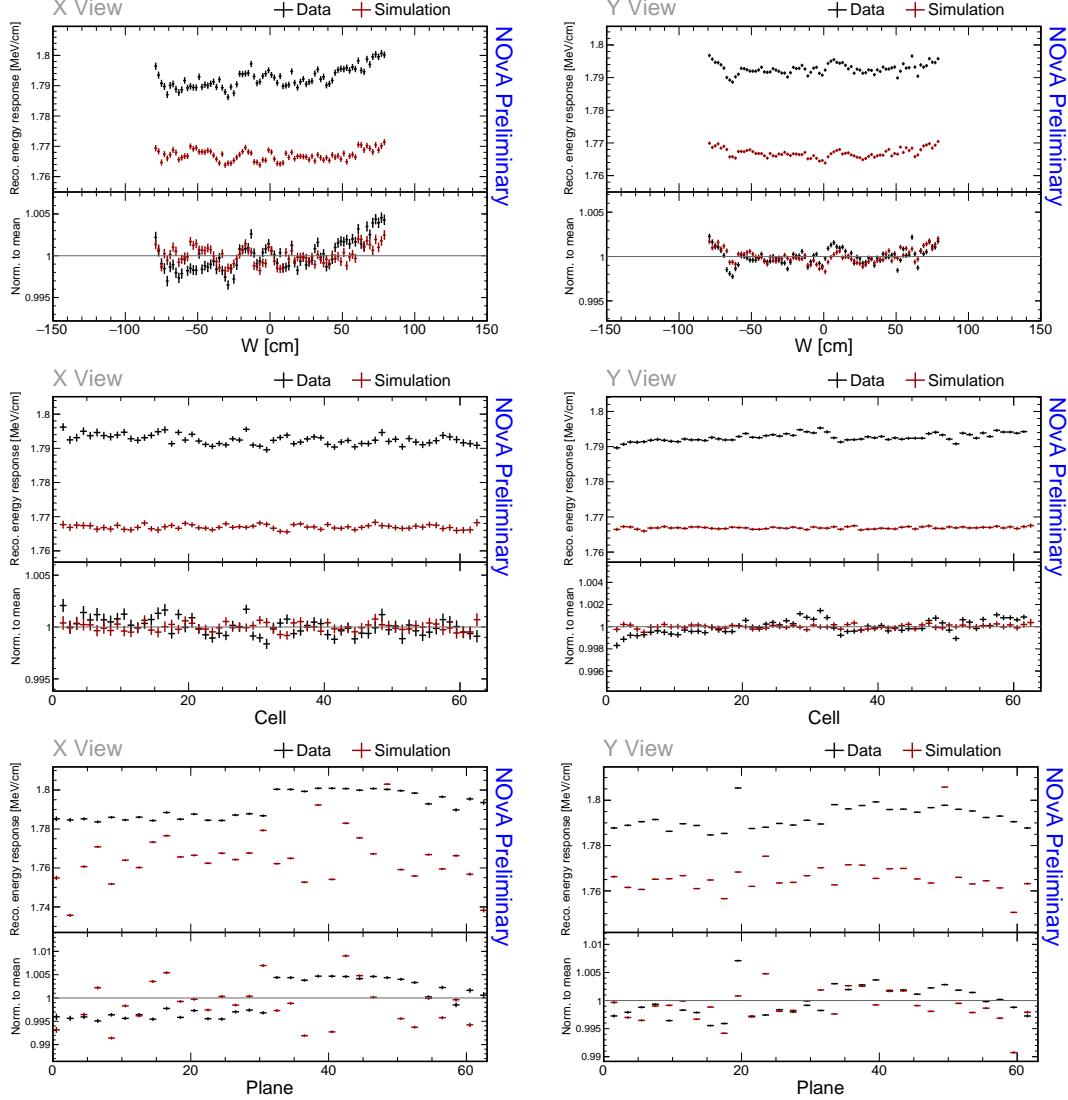


Figure 2.51: Distributions of through-going cosmic muons with $w \in (-80, 80)$ cm as a function of w , cells and planes for stable runs in the Test Beam period 4 data (black) and data-based simulation (red). Bottom panel of each plot shows the ratio between the value for each bin and the mean across the entire y axis, calculated separately for data and simulation.

each **FB** bin but importantly not for each plane. As a result, the relative calibration in simulation does not specifically correct for plane dependence, leading to larger variations between planes.

Furthermore, one half of plane 19 in data shows a corrected energy response approximately 0.7 % higher than the rest. This is due to a faulty **FEB** that was likely malfunctioning for a brief period. Since this study aims to examine purely the residual variations from the calibration procedure, cells corresponding to this **FEB** can be ignored. Additionally, planes 48 and 49 in simulation, which correspond to **FEBv5**,

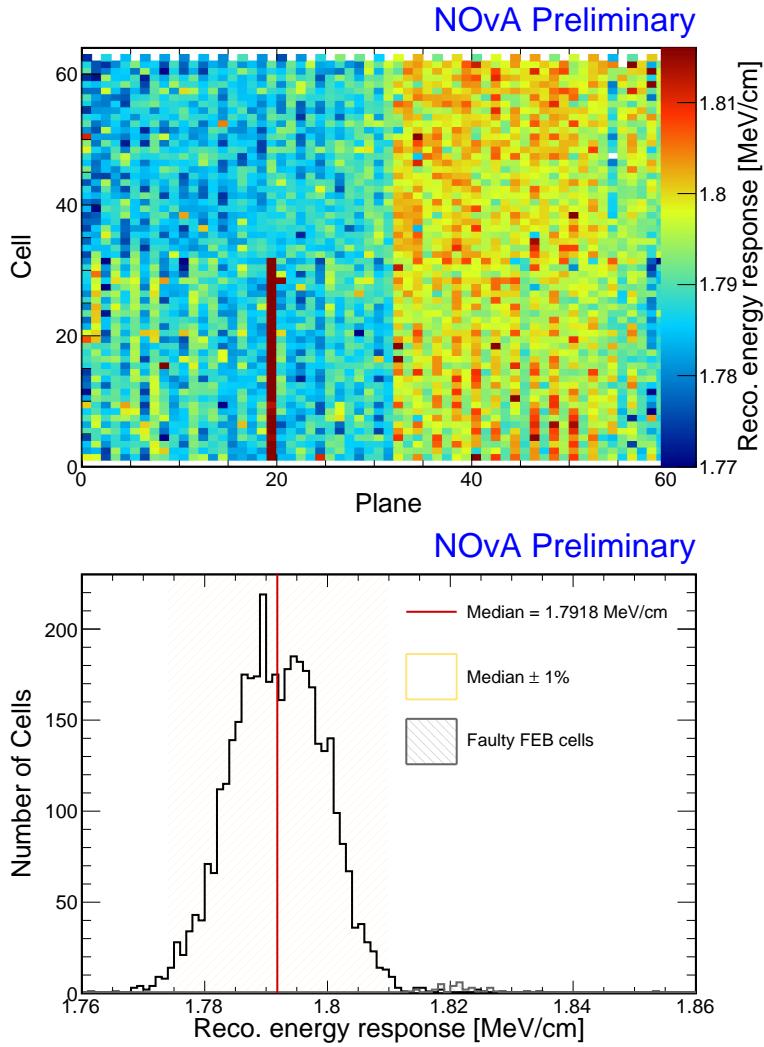


Figure 2.52: Top: Map of corrected energy deposition for through-going cosmic muons with $w \in (-80, 80)$ cm as a function of cell and plane numbers for stable runs in the Test Beam period 4 data. Bottom: Projection of each bin from the map. Red line shows the median value of the projection and red shaded area shows the $\pm 1\%$ range of the median value. Gray shaded bins show the excluded bins corresponding to faulty FEB, visible as a stark red line in plane 19 on the map. Additional features explained in text.

display a corrected energy response about 2 – 2.2 % higher than the rest. This discrepancy arises from the incorrect accounting for different FEB versions in simulation, combined with consolidating the planes into FB bins.

2.3.9 Conclusion

COMMENT: Should I create a section called discussion that would discuss all the possible improvements in the future?

We have successfully calibrated the NOvA Test Beam detector in four independent

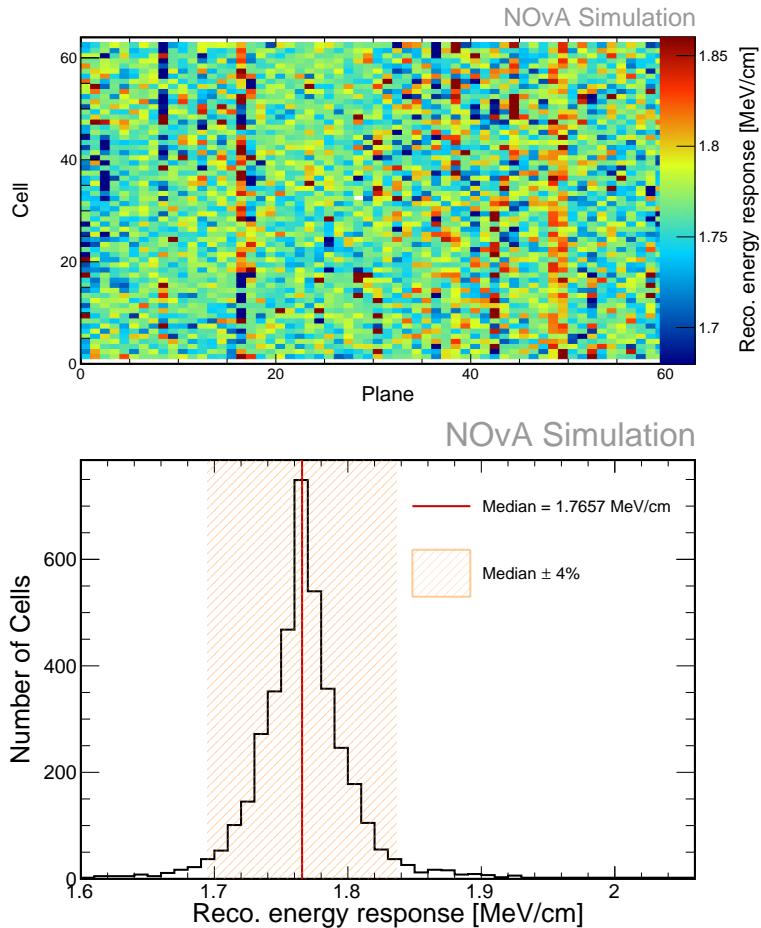


Figure 2.53: Top: Map of corrected energy deposition for through-going cosmic muons with $w \in (-80, 80)$ cm as a function of cell and plane numbers for stable runs in the Test Beam data-based simulation. Bottom: Projection of each bin from the map. Red line shows the median value of the projection and red shaded area shows the $\pm 4\%$ range of the median value. Additional features explained in text.

Test Beam data samples and in simulation. This is a critical step in order to analyse the Test Beam data, enabling crucial improvements to our understanding of the detector response and particularly of the calibration for all the NOvA detectors.

Each calibrated sample contains only a few uncalibrated cells, mainly concentrated on the periphery of the detector. Most of the other uncalibrated cells failed the calibration condition due to unavoidable circumstances, such as underfilled cells, or dead channels. However, several improvements could increase the number of calibrated cells or otherwise improve the calibration performance. Some of the improvements could benefit not only the Test Beam detector but all the NOvA detectors.

Several issues discovered during the Test Beam calibration that will enhance calibration and simulation for all the NOvA detectors include the problems found with

the threshold and shielding correction and with the readout simulation, described in Sec. 2.3.2. Ongoing work aims to address these issues by improving the logic of the threshold and shielding correction or basing it entirely on data. Additionally, plans are in place to remove the relative gain variation from the simulation, as the FB map used during simulation already sufficiently describes the relative gain variations, making the current treatment redundant.

One of the most common issues that could be resolved is the faulty FEBs, which often malfunctioned only for a limited period. This is particularly concerning, as cells belonging to faulty FEBs usually pass the calibration condition; however, if the corresponding FEB was faulty only for a limited period, the calibration results for those cells would be incorrect. This issue is currently difficult to address because the inputs for the relative calibration are organized by run and subrun numbers instead of by cells, making it impossible to calibrate cells corresponding to a single FEB separately from the rest. However, ongoing work is adapting both the input files and the calibration procedures to enable calibrating individual cells separately.

The time dependence of energy deposition in the Test Beam detector discussed in Sec. 2.3.8 is currently being studied and could help explain several contributions to the time-dependency seen in all NOvA detectors. Specifically, the Test Beam can study the environmental effects on energy deposition, as well as the ageing of different scintillators and electronics. The first results indicate that the NOvA scintillator oils age differently based on their quality.

Addressing the time-dependent variations in the energy deposition in Test Beam could include splitting the calibrated periods into smaller samples, similarly to the ND and FD. However, this approach risks insufficient statistics for attenuation fits in various cells. This could in turn be addressed by tuning the binning of the attenuation profiles to better fit the real cell dimensions that are actually being fitted in the attenuation fits.

Another technical improvements to the calibration results could stem from manually altering the χ^2 value of cells that have attenuation fits which visibly looks all right, to below 0.2. This would officially rendering these cells calibrated, similarly to how it was done for some cells in epochs 3d+3e and period 4. However, this would require manually going over all the uncalibrated cells, which is only really possible

for the Test Beam detector and only for larger calibrated samples. Reducing the size of the calibrated sample would make this extremely impractical.

Lastly, there is a possible improvement to the data-simulation discrepancy, by applying the absolute calibration results to each view independently. This would however have to be seriously considered, as the stopping muon sample used for calibration should technically not have any variation between the views and therefore this might introduce bias when applying the calibration results to beam data.

Overall, the calibration of the NOvA Test Beam detector was successful and represents a significant advancement, with ongoing efforts aimed at further refining the process and addressing the identified issues to improve calibration accuracy and reliability across all NOvA detectors. Part of the ongoing work also aims to quantify the correctness of the Test Beam calibration, devising a concrete systematic uncertainty.

Acronyms

2p2h two particle - two hole. 10

ADC Analog-to-Digital Converter. 7, 14, 17, 18, 52, 53, 63

APD Avalanche Photodiode. 5–7, 10, 17, 18, 52, 53, 55, 59, 63, 65, 69

ASIC Application-Specific Integrated Circuit. 7

BDT Boosted Decision Tree. 14

BPF Break Point Fitter. 12, 32–39, 41–46

CC Charged Current. 9–11

CMC Comprehensive Model Configuration. 10

CNN Convolutional Neural Network. 13

COH π Coherent π (production). 10

CP Charge conjugation - Parity (symmetry). 1

CRY Cosmic-Ray Shower Generator. 10, 14, 30

CVN Convolutional Visual Network. 13

DAQ Data Acquisition. 7, 11, 16

DCM Data Concentration Module. 7, 8

DIS Deep Inelastic Scattering. 10

DUNE Deep Underground Neutrino Experiment. 1, 9

FB Fibre Brightness. 17, 18, 48–50, 57, 58, 92, 93, 95

FD Far Detector. 3–7, 10, 11, 17–19, 22, 24, 25, 27–29, 32, 35, 37, 40, 47, 49, 52, 83, 88,
95

FEB Front End Board. 6–8, 17, 29, 30, 52, 53, 56, 59, 60, 62, 63, 65, 66, 68–72, 74, 75,
87, 88, 92, 93, 95

Fermilab Fermi National Accelerator Laboratory. 1, 3, 9, 27, 28

FHC Forward Horn Current (neutrino mode). 2, 3

FPGA Field Programmable Gate Array. 7

FSI Final State Interaction. 10

FTBF Fermilab Test Beam Facility. 25

LOWESS Locally Weighted Scatter plot Smoothing. 19, 20, 57

MC Monte Carlo. 8–10, 14, 30

MEC Meson Exchange Current. 10

MEU Muon Energy Unit. 21, 81, 82, 84

MI Main Injector. 1, 2, 9

MIP Minimum Ionising Particle. 14, 18, 21, 85–87, 89

MIPP Main Injector Particle Production (experiment). 9

ML Machine Learning. 13, 14

MWPC Multiwire Proportional Chamber. 26

NC Neutral Current. 11, 14

ND Near Detector. 1, 3–5, 7, 11, 12, 22, 24, 25, 27–30, 32, 35, 37, 39, 40, 47, 52, 83, 88, 95

NDOS Near Detector on the Surface. 3, 27, 28, 30, 74

NOvA NuMI Off-axis ν_e Appearance (experiment). 1, 3–14, 16–22, 24–28, 30, 31, 39, 47, 50, 53, 60, 66, 72, 80–83, 87, 89, 94–96

NuMI Neutrinos from the Main Injector. 1–3, 8, 9, 11, 26

PE Photo Electron. 6, 7, 15, 17–19, 52–54, 80, 84

PECorr Corrected Photo Electrons. 19, 21, 80–82

PID Particle Identification. [13](#)

POT Protons On Target. [2](#)

PPFX Package to Predict the Flux. [9](#), [23](#)

PVC Polyvinyl chloride. [4–6](#), [22](#)

QE Quasi Elastic (interaction). [9](#), [10](#)

ReMId Reconstructed Muon Identifier. [14](#)

Res Resonant baryon production. [9](#), [10](#)

RHC Reverse Horn Current (antineutrino mode). [2](#), [3](#)

ToF Time of Flight. [26](#)

WLS Wavelength Shifting (fibre). [5](#), [6](#), [10](#), [14](#), [16](#), [18](#), [21](#), [55](#), [60](#)

Bibliography

- [1] Nova experiments official website. URL <https://novaexperiment.fnal.gov>. Cited February 2024.
- [2] Fermi national accelerator laboratory official website. URL <https://fnal.gov>. Cited February 2024.
- [3] Gary Feldman for the NOvA Collaboration. Physics of the nova experiment. White paper, 2012. URL <https://nova-docdb.fnal.gov/cgi-bin>ShowDocument?docid=7733>. Public NOvA document: NOVA-doc-7733-v1, cited on 05.2020.
- [4] R.B. Patterson. The NOvA Experiment: Status and Outlook. *Nucl. Phys. B Proc. Suppl.*, 235-236:151–157, 2013. doi:[10.1016/j.nuclphysbps.2013.04.005](https://doi.org/10.1016/j.nuclphysbps.2013.04.005).
- [5] P. Adamson et al. First measurement of muon-neutrino disappearance in NOvA. *Phys. Rev. D*, 93(5):051104, 2016. doi:[10.1103/PhysRevD.93.051104](https://doi.org/10.1103/PhysRevD.93.051104).
- [6] M.A. Acero et al. First Measurement of Neutrino Oscillation Parameters using Neutrinos and Antineutrinos by NOvA. *Phys. Rev. Lett.*, 123(15):151803, 2019. doi:[10.1103/PhysRevLett.123.151803](https://doi.org/10.1103/PhysRevLett.123.151803).
- [7] M. A. Acero et al. Improved measurement of neutrino oscillation parameters by the NOvA experiment. *Phys. Rev. D*, 106(3):032004, 2022. doi:[10.1103/PhysRevD.106.032004](https://doi.org/10.1103/PhysRevD.106.032004).
- [8] M. A. Acero et al. Measurement of neutrino-induced neutral-current coherent π^0 production in the NOvA near detector. *Phys. Rev. D*, 102(1):012004, 2020. doi:[10.1103/PhysRevD.102.012004](https://doi.org/10.1103/PhysRevD.102.012004).
- [9] M. A. Acero et al. Measurement of the double-differential muon-neutrino charged-current inclusive cross section in the NOvA near detector. *Phys. Rev. D*, 107(5):052011, 2023. doi:[10.1103/PhysRevD.107.052011](https://doi.org/10.1103/PhysRevD.107.052011).
- [10] M. A. Acero et al. Measurement of the ν_e –Nucleus Charged-Current Double-Differential Cross Section at $\langle E_\nu \rangle = 2.4$ GeV using NOvA. *Phys. Rev. Lett.*, 130(5):051802, 2023. doi:[10.1103/PhysRevLett.130.051802](https://doi.org/10.1103/PhysRevLett.130.051802).

- [11] M. A. Acero et al. Measurement of ν_μ charged-current inclusive π^0 production in the NOvA near detector. *Phys. Rev. D*, 107(11):112008, 2023. doi:[10.1103/PhysRevD.107.112008](https://doi.org/10.1103/PhysRevD.107.112008).
- [12] P. Adamson et al. Search for active-sterile neutrino mixing using neutral-current interactions in NOvA. *Phys. Rev. D*, 96(7):072006, 2017. doi:[10.1103/PhysRevD.96.072006](https://doi.org/10.1103/PhysRevD.96.072006).
- [13] M. A. Acero et al. Search for Active-Sterile Antineutrino Mixing Using Neutral-Current Interactions with the NOvA Experiment. *Phys. Rev. Lett.*, 127(20):201801, 2021. doi:[10.1103/PhysRevLett.127.201801](https://doi.org/10.1103/PhysRevLett.127.201801).
- [14] M. A. Acero et al. Supernova neutrino detection in NOvA. *JCAP*, 10:014, 2020. doi:[10.1088/1475-7516/2020/10/014](https://doi.org/10.1088/1475-7516/2020/10/014).
- [15] M. A. Acero et al. Extended search for supernovalike neutrinos in NOvA coincident with LIGO/Virgo detections. *Phys. Rev. D*, 104(6):063024, 2021. doi:[10.1103/PhysRevD.104.063024](https://doi.org/10.1103/PhysRevD.104.063024).
- [16] M. A. Acero et al. Search for slow magnetic monopoles with the NOvA detector on the surface. *Phys. Rev. D*, 103(1):012007, 2021. doi:[10.1103/PhysRevD.103.012007](https://doi.org/10.1103/PhysRevD.103.012007).
- [17] Peter N. Shanahan and Patricia LaVern Vahle. Physics with NOvA: a half-time review. *Eur. Phys. J. ST*, 230(24):4259–4273, 2021. doi:[10.1140/epjs/s11734-021-00285-9](https://doi.org/10.1140/epjs/s11734-021-00285-9).
- [18] Peter Shanahan and Patricia Vahle. The NOvA Physics Program through 2025. Snowmass 2021 Letters of Interest NF, 135. URL https://www.snowmass21.org/docs/files/summaries/NF/SNOWMASS21-NF1_NF3_Patricia_Vahle-145.pdf.
- [19] P. Adamson et al. The NuMI Neutrino Beam. *Nucl. Instrum. Meth. A*, 806:279–306, 2016. doi:[10.1016/j.nima.2015.08.063](https://doi.org/10.1016/j.nima.2015.08.063).
- [20] Records. URL <https://operations.fnal.gov/records/>. Performance records achieved by the Fermilab accelerators. Cited May 2024.

- [21] Katsuya Yonehara. Megawatt upgrade of numi target system. 1 2022. URL <https://www.osti.gov/biblio/1844332>.
- [22] Leonidas Aliaga Soplin. *Neutrino Flux Prediction for the NuMI Beamline*. PhD thesis, William-Mary Coll., 2016.
- [23] Leonidas Aliaga Soplin. 2017-2018 Beam Plots. NOVA Document 20843. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=20843>. NOvA internal document.
- [24] D. S. Ayres et al. The nova technical design report. *NOvA Collaboration*, 2007. doi:[10.2172/935497](https://doi.org/10.2172/935497).
- [25] Yury Kudenko. Neutrino detectors for oscillation experiments. *JINST*, 12(06):C06003, 2017. doi:[10.1088/1748-0221/12/06/C06003](https://doi.org/10.1088/1748-0221/12/06/C06003).
- [26] S. Mufson et al. Liquid scintillator production for the NOvA experiment. *Nucl. Instrum. Meth. A*, 799:1–9, 2015. doi:[10.1016/j.nima.2015.07.026](https://doi.org/10.1016/j.nima.2015.07.026).
- [27] Alex Sousa. Density of NOvA Liquid Scintillator at 69° F. NOVA Document 11886, . URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=11886>. NOvA internal document.
- [28] M.J. Berger, J.S. Coursey, M.A. Zucker, and J. Chang. ESTAR, PSTAR, and ASTAR: Computer Programs for Calculating Stopping-Power and Range Tables for Electrons, Protons, and Helium Ions (version 1.2.3). URL <http://physics.nist.gov/Star>.
- [29] Luke Vinton. Calorimetric Energy Scale Calibration of the NOvA Detectors. NOVA Document 13579, document FA_Calorimetric_energy_scale.pdf. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=13579>. NOvA Technical Note.
- [30] Xinchun Tian. NOvA Data Acquisition Software System. In *Meeting of the APS Division of Particles and Fields*, 9 2011.
- [31] Joao Coelho, Barnali Chowdhury, and Ryan Murphy. Tech note: Good data selection. NOVA Document 13546. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=13546>. NOvA technical note.

- [32] Jonathan M. Paley. BadChannels Technical Note. NOVA Document 12771. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=12771>. NOvA technical note.
- [33] A. Aurisano, C. Backhouse, R. Hatcher, N. Mayer, J. Musser, R. Patterson, R. Schroeter, and A. Sousa. The NOvA simulation chain. *J. Phys. Conf. Ser.*, 664(7):072002, 2015. doi:[10.1088/1742-6596/664/7/072002](https://doi.org/10.1088/1742-6596/664/7/072002).
- [34] S. Agostinelli et al. GEANT4—a simulation toolkit. *Nucl. Instrum. Meth. A*, 506: 250–303, 2003. doi:[10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8).
- [35] Zarko Pavlovic. *Observation of Disappearance of Muon Neutrinos in the NuMI Beam*. PhD thesis, Texas U., 2008.
- [36] L. Aliaga et al. Neutrino Flux Predictions for the NuMI Beam. *Phys. Rev. D*, 94(9):092005, 2016. doi:[10.1103/PhysRevD.94.092005](https://doi.org/10.1103/PhysRevD.94.092005). [Addendum: *Phys. Rev. D* 95, 039903 (2017)].
- [37] C. Alt et al. Inclusive production of charged pions in p+C collisions at 158-GeV/c beam momentum. *Eur. Phys. J. C*, 49:897–917, 2007. doi:[10.1140/epjc/s10052-006-0165-7](https://doi.org/10.1140/epjc/s10052-006-0165-7).
- [38] Gemma Maria Tinti. *Sterile neutrino oscillations in MINOS and hadron production in pC collisions*. PhD thesis, 2010.
- [39] B Baatar, G Barr, J Bartke, L Betev, O Chvala, J Dolejsi, V Eckardt, HG Fischer, Z Fodor, A Karev, et al. Inclusive production of protons, anti-protons, neutrons, deuterons and tritons in p+ c collisions at 158 gev/c beam momentum. *The European Physical Journal C*, 73:1–66, 2013. doi:[10.1140/epjc/s10052-013-2364-3](https://doi.org/10.1140/epjc/s10052-013-2364-3).
- [40] D. S. Barton et al. Experimental study of the a dependence of inclusive hadron fragmentation. *Phys. Rev. D*, 27:2580–2599, Jun 1983. doi:[10.1103/PhysRevD.27.2580](https://doi.org/10.1103/PhysRevD.27.2580).
- [41] Andrey V. Lebedev. *Ratio of Pion Kaon Production in Proton Carbon Interactions*. PhD thesis, 2007.

- [42] T.T. Böhlen, F. Cerutti, M.P.W. Chin, A. Fassò, A. Ferrari, P.G. Ortega, A. Mairani, P.R. Sala, G. Smirnov, and V. Vlachoudis. The fluka code: Developments and challenges for high energy and medical applications. 120:211–214. ISSN 0090-3752. doi:[10.1016/j.nds.2014.07.049](https://doi.org/10.1016/j.nds.2014.07.049).
- [43] Alfredo Ferrari, Paola R. Sala, Alberto Fasso, and Johannes Ranft. FLUKA: A multi-particle transport code (Program version 2005). 10 2005. doi:[10.2172/877507](https://doi.org/10.2172/877507).
- [44] A. Aduszkiewicz et al. Measurements of production and inelastic cross sections for p+C , p+Be , and p+Al at 60 GeV/c and p+C and p+Be at 120 GeV/c. *Phys. Rev. D*, 100(11):112001, 2019. doi:[10.1103/PhysRevD.100.112001](https://doi.org/10.1103/PhysRevD.100.112001).
- [45] H. Adhikary et al. Measurements of π^+ , π^- , p , \bar{p} , K^+ and K^- production in 120 GeV/c p + C interactions. *Phys. Rev. D*, 108:072013, 2023. doi:[10.1103/PhysRevD.108.072013](https://doi.org/10.1103/PhysRevD.108.072013).
- [46] H. Adhikary et al. Measurements of KS0, Λ , and Λ^- production in 120 GeV/c p+C interactions. *Phys. Rev. D*, 107(7):072004, 2023. doi:[10.1103/PhysRevD.107.072004](https://doi.org/10.1103/PhysRevD.107.072004).
- [47] A. Aduszkiewicz et al. Measurements of hadron production in $\pi^+ + C$ and $\pi^+ + Be$ interactions at 60 GeV/c. *Phys. Rev. D*, 100(11):112004, 2019. doi:[10.1103/PhysRevD.100.112004](https://doi.org/10.1103/PhysRevD.100.112004).
- [48] Matej Pavin. Hadron production measurements for neutrino experiments, 2020. URL https://indico.fnal.gov/event/43209/contributions/187869/attachments/130502/159028/HadronProduction_neutrino2020.pdf. Presented on Neutrino 2020 conference.
- [49] T. Akaishi et al. EMPHATIC: A Proposed Experiment to Measure Hadron Scattering and Production Cross Sections for Improved Neutrino Flux Predictions. 12 2019.
- [50] C. Andreopoulos et al. The GENIE Neutrino Monte Carlo Generator. *Nucl. Instrum. Meth. A*, 614:87–104, 2010. doi:[10.1016/j.nima.2009.12.009](https://doi.org/10.1016/j.nima.2009.12.009).

- [51] J. Nieves, J. E. Amaro, and M. Valverde. Inclusive quasielastic charged-current neutrino-nucleus reactions. *Phys. Rev. C*, 70:055503, Nov 2004. doi:[10.1103/PhysRevC.70.055503](https://doi.org/10.1103/PhysRevC.70.055503).
- [52] Aaron S. Meyer, Minerba Betancourt, Richard Gran, and Richard J. Hill. Deuterium target data for precision neutrino-nucleus cross sections. *Phys. Rev. D*, 93:113015, Jun 2016. doi:[10.1103/PhysRevD.93.113015](https://doi.org/10.1103/PhysRevD.93.113015).
- [53] J. Nieves, I. Ruiz Simo, and M. J. Vicente Vacas. Inclusive charged-current neutrino-nucleus reactions. *Phys. Rev. C*, 83:045501, Apr 2011. doi:[10.1103/PhysRevC.83.045501](https://doi.org/10.1103/PhysRevC.83.045501).
- [54] R. Gran, J. Nieves, F. Sanchez, and M. J. Vicente Vacas. Neutrino-nucleus quasielastic and 2p2h interactions up to 10 gev. *Phys. Rev. D*, 88:113007, Dec 2013. doi:[10.1103/PhysRevD.88.113007](https://doi.org/10.1103/PhysRevD.88.113007).
- [55] Ch. Berger and L. M. Sehgal. Lepton mass effects in single pion production by neutrinos. *Phys. Rev. D*, 76:113004, Dec 2007. doi:[10.1103/PhysRevD.76.113004](https://doi.org/10.1103/PhysRevD.76.113004).
- [56] Ch. Berger and L. M. Sehgal. Partially conserved axial vector current and coherent pion production by low energy neutrinos. *Phys. Rev. D*, 79:053003, Mar 2009. doi:[10.1103/PhysRevD.79.053003](https://doi.org/10.1103/PhysRevD.79.053003).
- [57] A Bodek and U. K. Yang. Higher twist, xi(omega) scaling, and effective LO PDFs for lepton scattering in the few GeV region. *J. Phys. G*, 29:1899–1906, 2003. doi:[10.1088/0954-3899/29/8/369](https://doi.org/10.1088/0954-3899/29/8/369).
- [58] T. Yang, C. Andreopoulos, H. Gallagher, K. Hoffmann, and P. Kehayias. A Hadronization Model for Few-GeV Neutrino Interactions. *Eur. Phys. J. C*, 63: 1–10, 2009. doi:[10.1140/epjc/s10052-009-1094-z](https://doi.org/10.1140/epjc/s10052-009-1094-z).
- [59] L.L. Salcedo, E. Oset, M.J. Vicente-Vacas, and C. Garcia-Recio. Computer simulation of inclusive pion nuclear reactions. 484(3):557–592. ISSN 0375-9474. doi:[10.1016/0375-9474\(88\)90310-7](https://doi.org/10.1016/0375-9474(88)90310-7).
- [60] Chris Hagmann, David Lange, and Douglas Wright. Cosmic-ray shower generator (cry) for monte carlo transport codes. volume 2, pages 1143 – 1146. ISBN 978-1-4244-0922-8. doi:[10.1109/NSSMIC.2007.4437209](https://doi.org/10.1109/NSSMIC.2007.4437209).

- [61] C. N. Chou. The nature of the saturation effect of fluorescent scintillators. *Phys. Rev.*, 87:904–905, Sep 1952. doi:[10.1103/PhysRev.87.904](https://doi.org/10.1103/PhysRev.87.904). URL <https://link.aps.org/doi/10.1103/PhysRev.87.904>.
- [62] M. Baird, J. Bian, M. Messier, E. Niner, D. Rocco, and K. Sachdev. Event Reconstruction Techniques in NOvA. *J. Phys. Conf. Ser.*, 664(7):072035, 2015. doi:[10.1088/1742-6596/664/7/072035](https://doi.org/10.1088/1742-6596/664/7/072035).
- [63] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, KDD’96, page 226–231. AAAI Press, 1996. doi:[10.5555/3001460.3001507](https://doi.org/10.5555/3001460.3001507).
- [64] Leandro A. F. Fernandes and Manuel Menezes de Oliveira Neto. Real-time line detection through an improved hough transform voting scheme. *Pattern Recognit.*, 41:299–314, 2008. doi:[10.1016/j.patcog.2007.04.003](https://doi.org/10.1016/j.patcog.2007.04.003). URL <https://api.semanticscholar.org/CorpusID:5996185>.
- [65] Mattias Ohlsson, Carsten Peterson, and Alan L. Yuille. Track finding with deformable templates – the elastic arms approach. 71(1):77–98. ISSN 0010-4655. doi:[10.1016/0010-4655\(92\)90074-9](https://doi.org/10.1016/0010-4655(92)90074-9). URL <https://www.sciencedirect.com/science/article/pii/0010465592900749>.
- [66] R. Krishnapuram and J.M. Keller. A possibilistic approach to clustering. *IEEE Transactions on Fuzzy Systems*, 1(2):98–110, 1993. doi:[10.1109/91.227387](https://doi.org/10.1109/91.227387).
- [67] Miin-Shen Yang and Kuo-Lung Wu. Unsupervised possibilistic clustering. 39(1):5–21. ISSN 0031-3203. doi:<https://doi.org/10.1016/j.patcog.2005.07.005>. URL <https://www.sciencedirect.com/science/article/pii/S0031320305002943>.
- [68] Jianming Bian, Hongyue Duyang, and Alexander Radovic. Blessing package for Neutrino-Electron Scattering and Coherent Pi0. NOVA Document 13862. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=13862>. NOvA internal document.

- [69] Nicholas Jacob Raddatz. *Measurement of Muon Neutrino Disappearance with Non-Fiducial Interactions in the NOvA Experiment*. PhD thesis, Minnesota U., 2016.
- [70] Michael David Baird. *An Analysis of Muon Neutrino Disappearance from the NuMI Beam Using an Optimal Track Fitter*. PhD thesis, Indiana U., 2015.
- [71] G. Lutz. Optimum track fitting in the presence of multiple scattering. 273 (1):349–361. ISSN 0168-9002. doi:[https://doi.org/10.1016/0168-9002\(88\)90836-4](https://doi.org/10.1016/0168-9002(88)90836-4). URL <https://www.sciencedirect.com/science/article/pii/0168900288908364>.
- [72] Brian Rebel. Window tracking algorithm for cosmic ray muons. NOVA Document 15977-v1. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=15977>. NOvA internal document.
- [73] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going Deeper with Convolutions. art. arXiv:1409.4842. doi:[10.48550/arXiv.1409.4842](https://arxiv.org/abs/1409.4842).
- [74] A. Aurisano, A. Radovic, D. Rocco, A. Himmel, M.D. Messier, E. Niner, G. Pawloski, F. Psihas, A. Sousa, and P. Vahle. A Convolutional Neural Network Neutrino Event Classifier. *JINST*, 11(09):P09001, 2016. doi:[10.1088/1748-0221/11/09/P09001](https://doi.org/10.1088/1748-0221/11/09/P09001).
- [75] Fernanda Psihas. Measurement of long baseline neutrino oscillations and improvements from deep learning. 1 2018. doi:[10.2172/1437288](https://doi.org/10.2172/1437288). URL <https://www.osti.gov/biblio/1437288>.
- [76] Prabhjot Singh. *Extraction of Neutrino Oscillation Parameters using a Simultaneous Fit of ν_μ Disappearance and ν_e Appearance data with the NOvA Experiment*. PhD thesis, University of Delhi, Dept. of Physics and Astrophysics, India, Delhi U., 9 2019.
- [77] Ryan J. Nichol. Fibre brightness from cosmic muon data. NOVA Document 34909. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=34909>. NOvA technical note.

- [78] Prabhjot Singh. Attenuation Calibration of the NOvA Detectors. In *The 38th International Conference on High Energy Physics (ICHEP 2016)*, 2024. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=15840>. (cited on 03.2024).
- [79] Michael Wallbank. The NOvA Test Beam Program. ICHEP2020:188, . doi:[10.22323/1.390.0188](https://doi.org/10.22323/1.390.0188).
- [80] 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.
- [81] 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.
- [82] 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.
- [83] Fermilab test beam facility official website. URL <https://ftbf.fnal.gov/>. Cited March 2024.
- [84] 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.
- [85] 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.
- [86] 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.
- [87] 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.

- [88] 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.
- [89] Miniboone experiments official website. URL <https://www-boone.fnal.gov/>. Cited March 2024.
- [90] 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.
- [91] 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.
- [92] 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.
- [93] 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.
- [94] 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.
- [95] Teresa Megan Lackey. *Proton Scattering in NOvA Test Beam*. PhD thesis.
- [96] David Northacker, Alex Sousa, and Yagnur Torun. Test Beam - Overfilling Horizontal Planes. NOVA Document 49439. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=49439>. NOvA internal document.
- [97] 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.

- [98] R. L. Workman and Others. Review of Particle Physics. 2022:083C01. doi:[10.1093/ptep/ptac097](https://doi.org/10.1093/ptep/ptac097).
- [99] Leon M. Mualem. Hamamatsu QC data for ND APD arrays. NOVA Document 5239. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=5239>. NOvA internal document.
- [100] Hayes D Merritt. DAQChannelMap Terms and Definitions. NOVA Document 11570. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=11570>. NOvA internal document.
- [101] Tian Xin. Technote for data-driven threshold and shadow correction. NOVA Document 15223. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=15223>. NOvA technical note.
- [102] Matthew Strait. Update on light level tuning. NOVA Document 43249. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=43249>. NOvA internal document.
- [103] Artur Sztuc. DCM/FEB Shut-Off Studies. NOVA Document 53658. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=53658>. NOvA internal document.
- [104] Adam Lister. Period 12-14 Calibration Validation Document. NOVA Document 60709. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=60709>. NOvA technical note.
- [105] Adam Lister and Jozef Trokan-Tenorio. Calibration of Periods 12-14 Data in the NOvA Near and Far Detectors. NOVA Document 60838. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=60838>. NOvA technical note.
- [106] Randeeth L Dasanayaka. TB response with UnixTime(S). NOVA Document 59591. URL <https://nova-docdb.fnal.gov/cgi-bin/sso>ShowDocument?docid=59591>. NOvA internal document.

APPENDIX A

Test Beam Calibration Validation Plots

A.1 Distributions for Stopping Muons

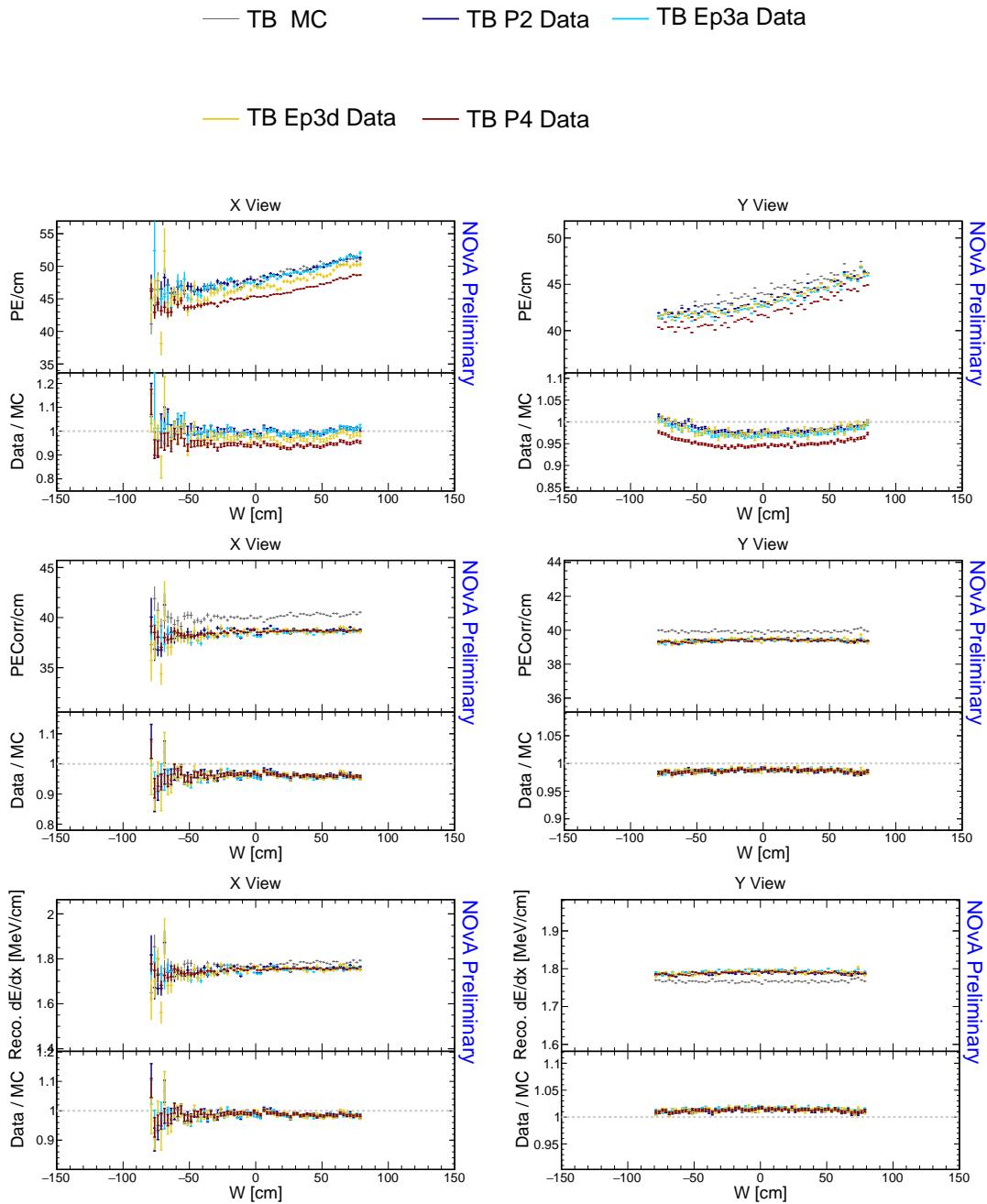


Figure A.1: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the position within a cell.

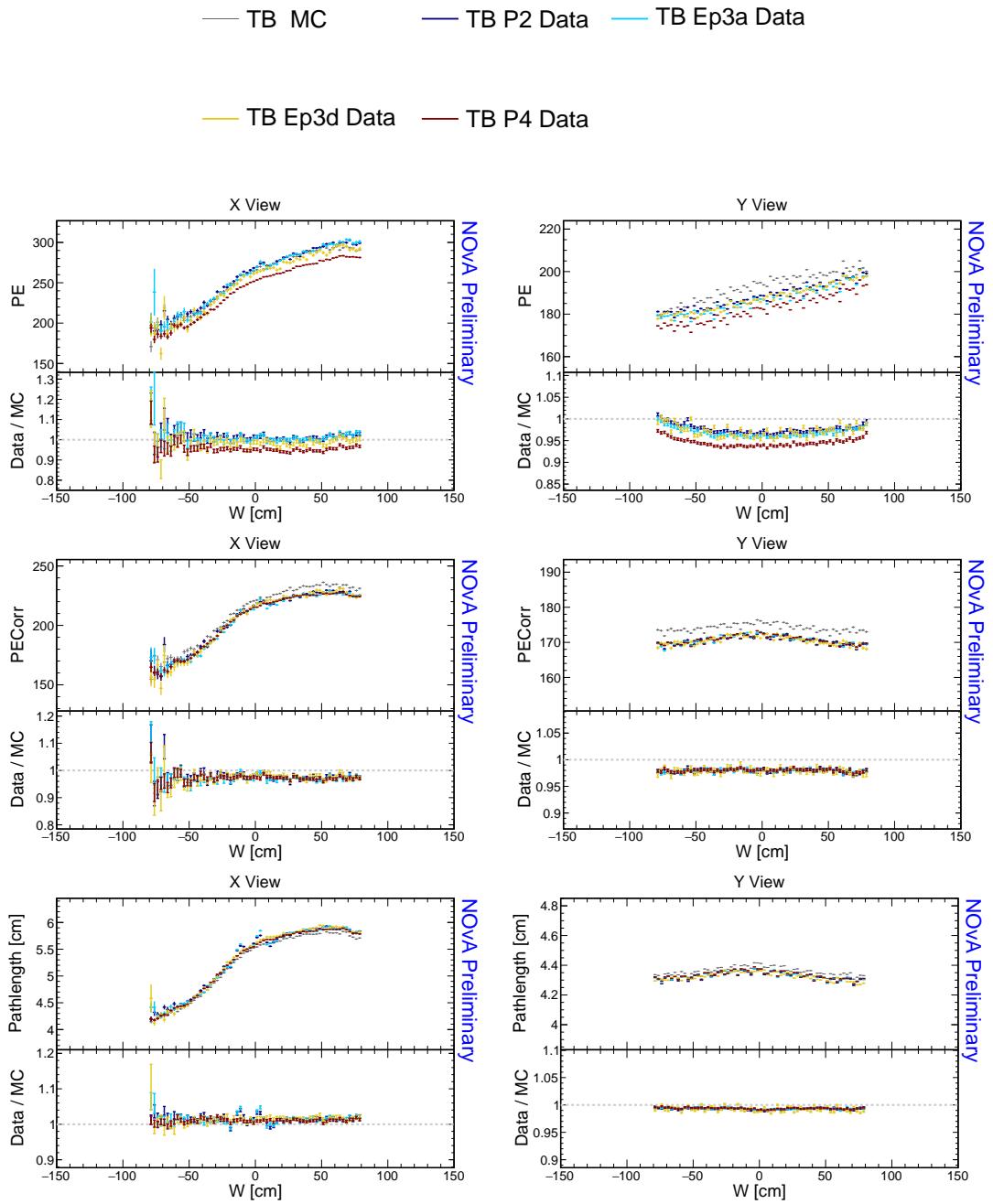


Figure A.2: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the position within a cell.

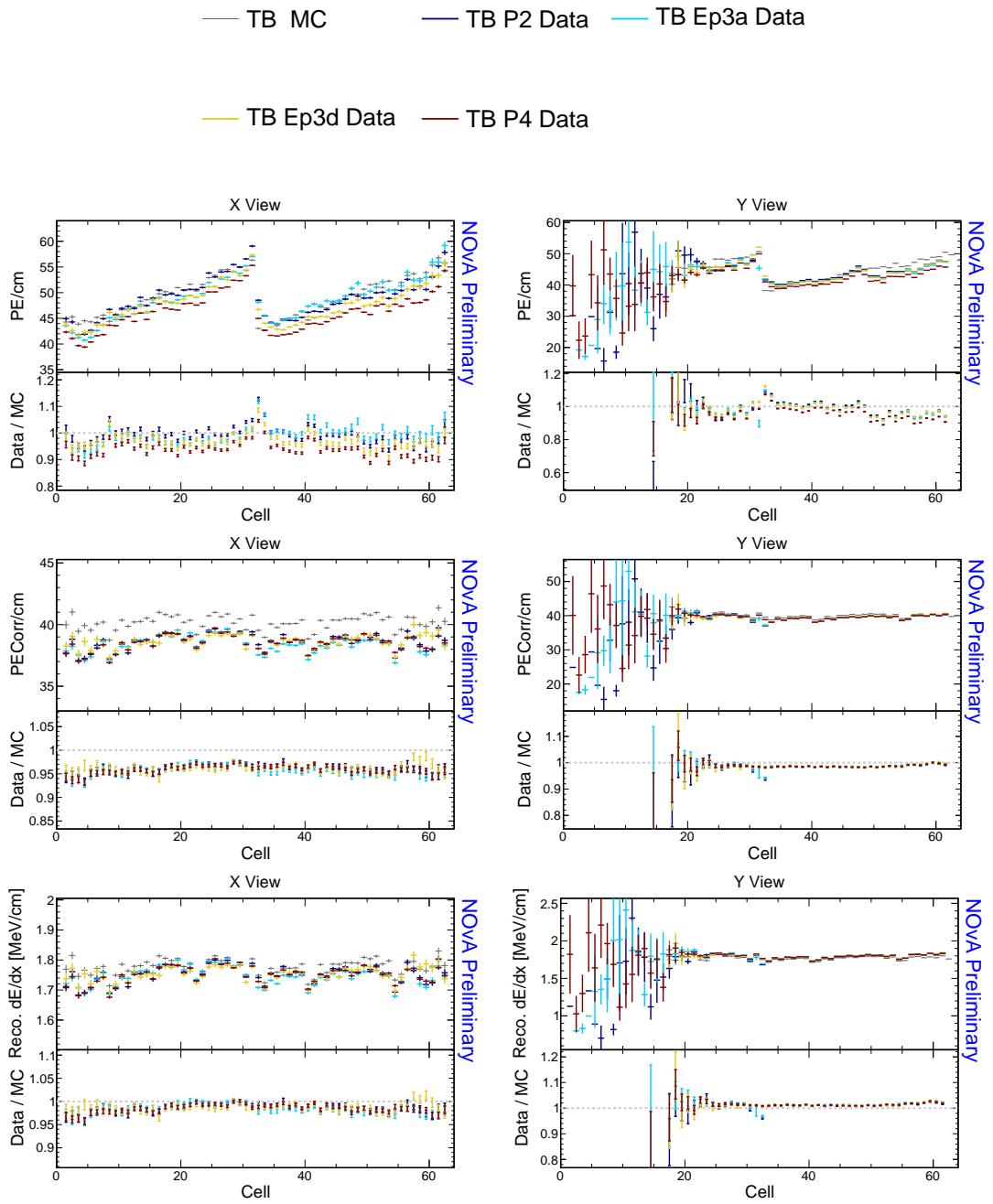


Figure A.3: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the cells of the detector.

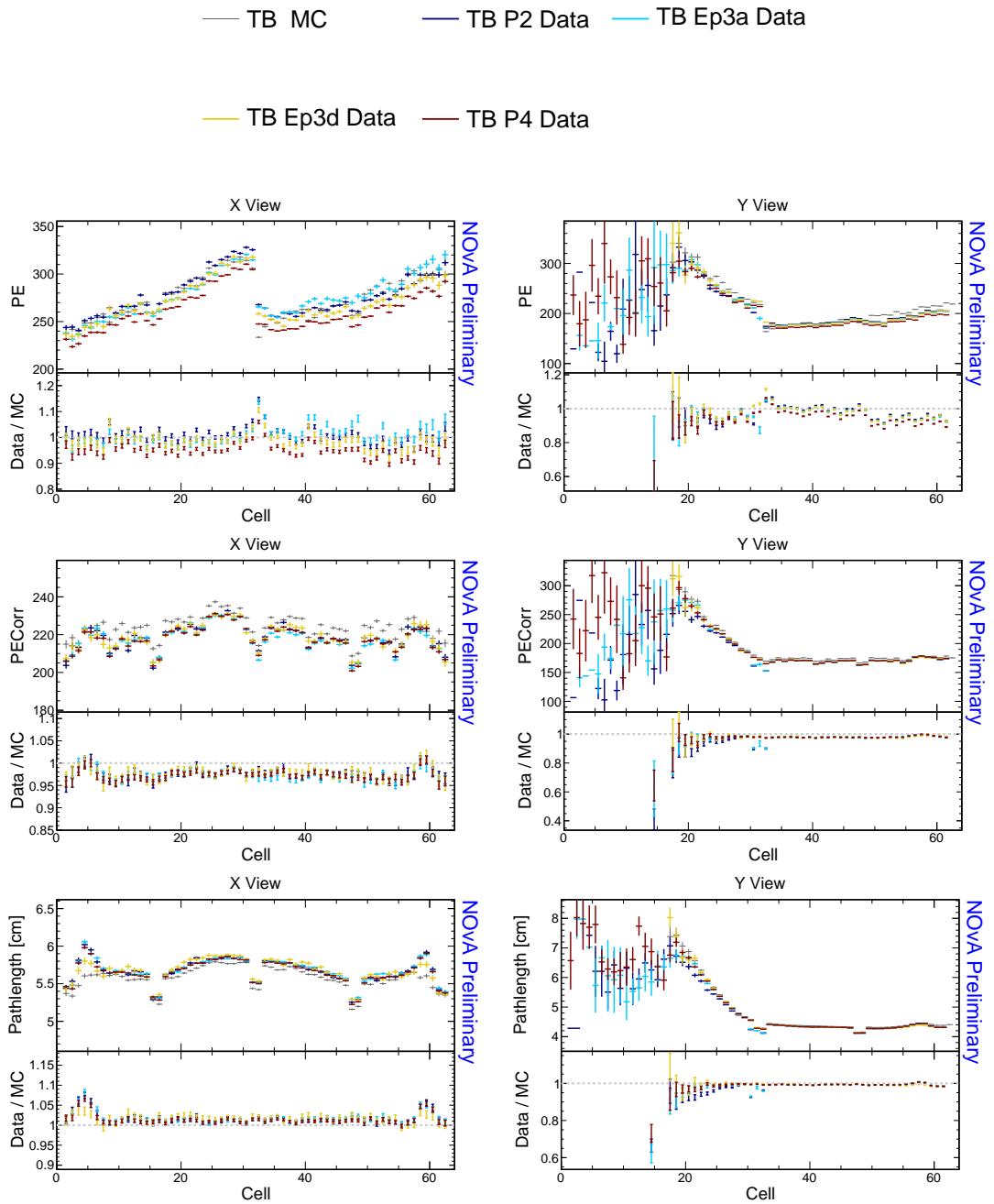


Figure A.4: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the cells of the detector.

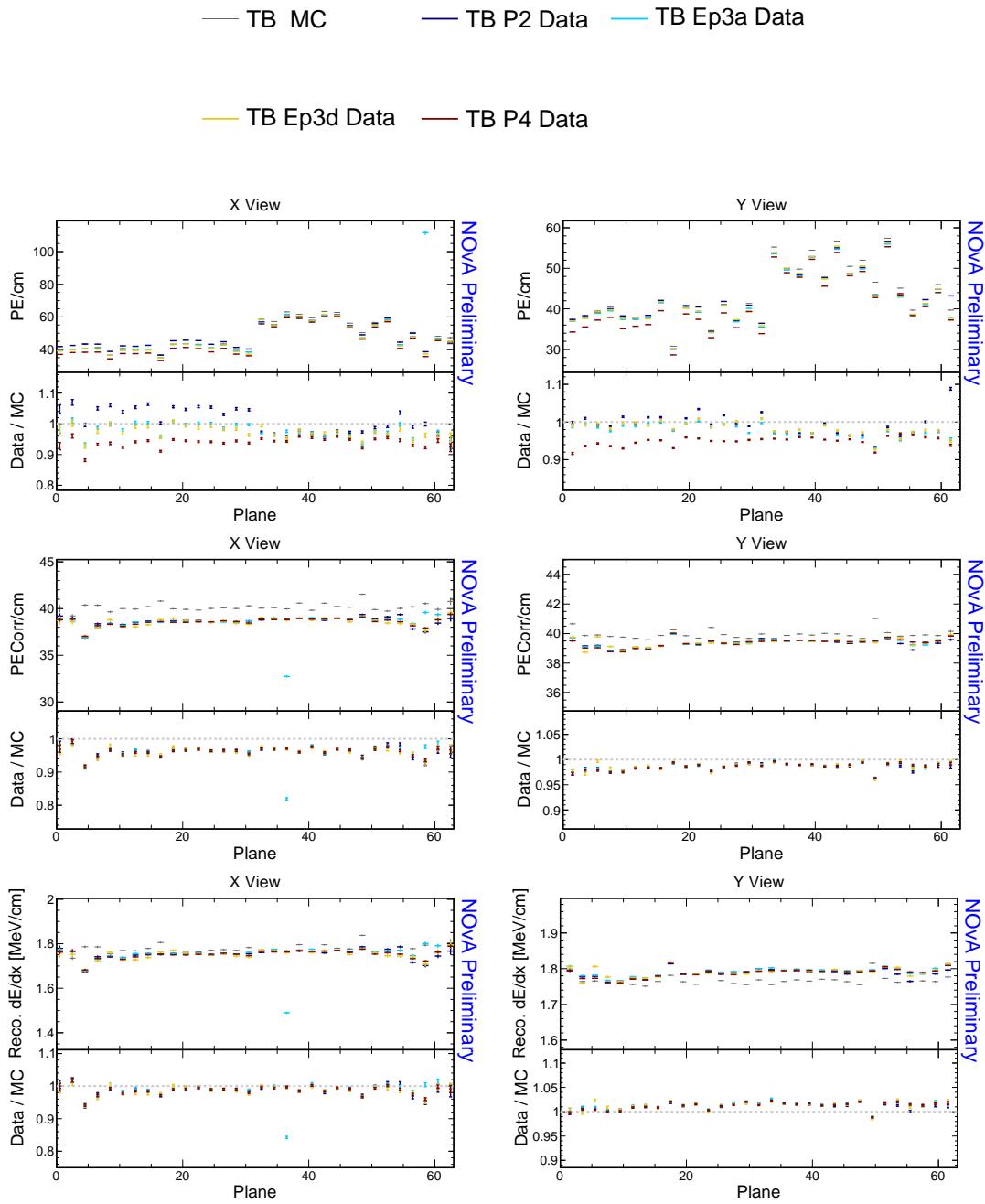


Figure A.5: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the planes of the detector.

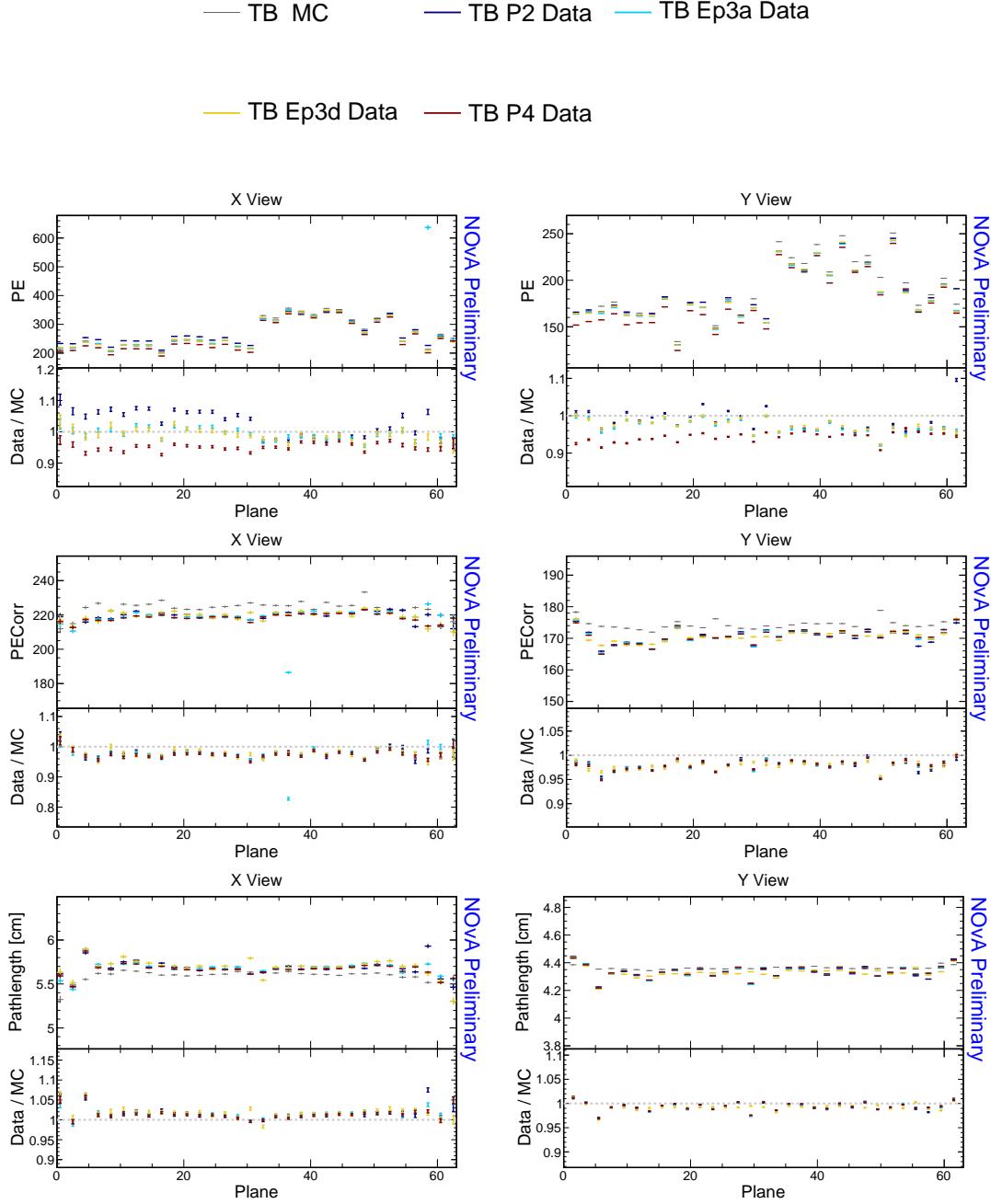


Figure A.6: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the planes of the detector.

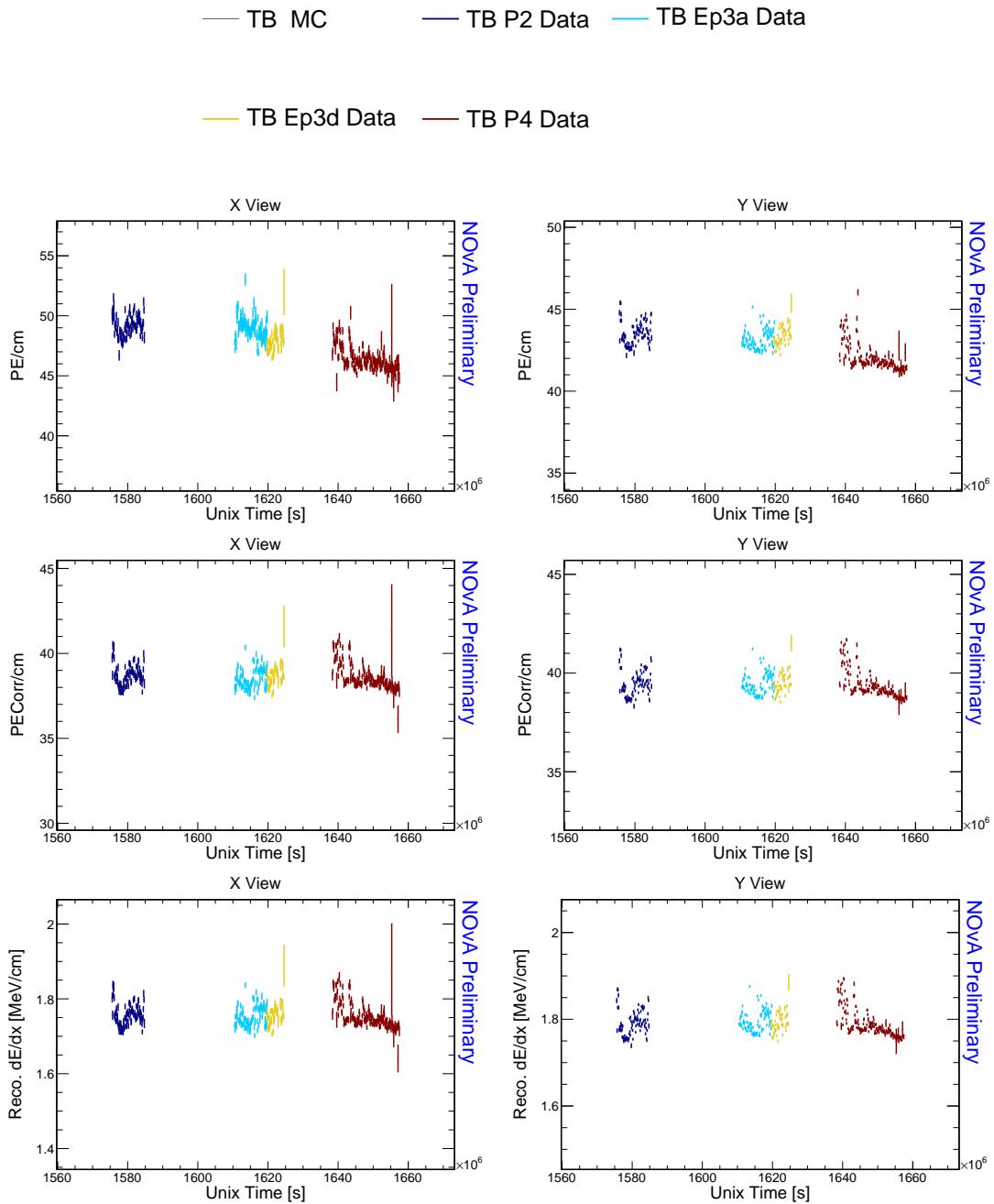


Figure A.7: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the event UNIX time.

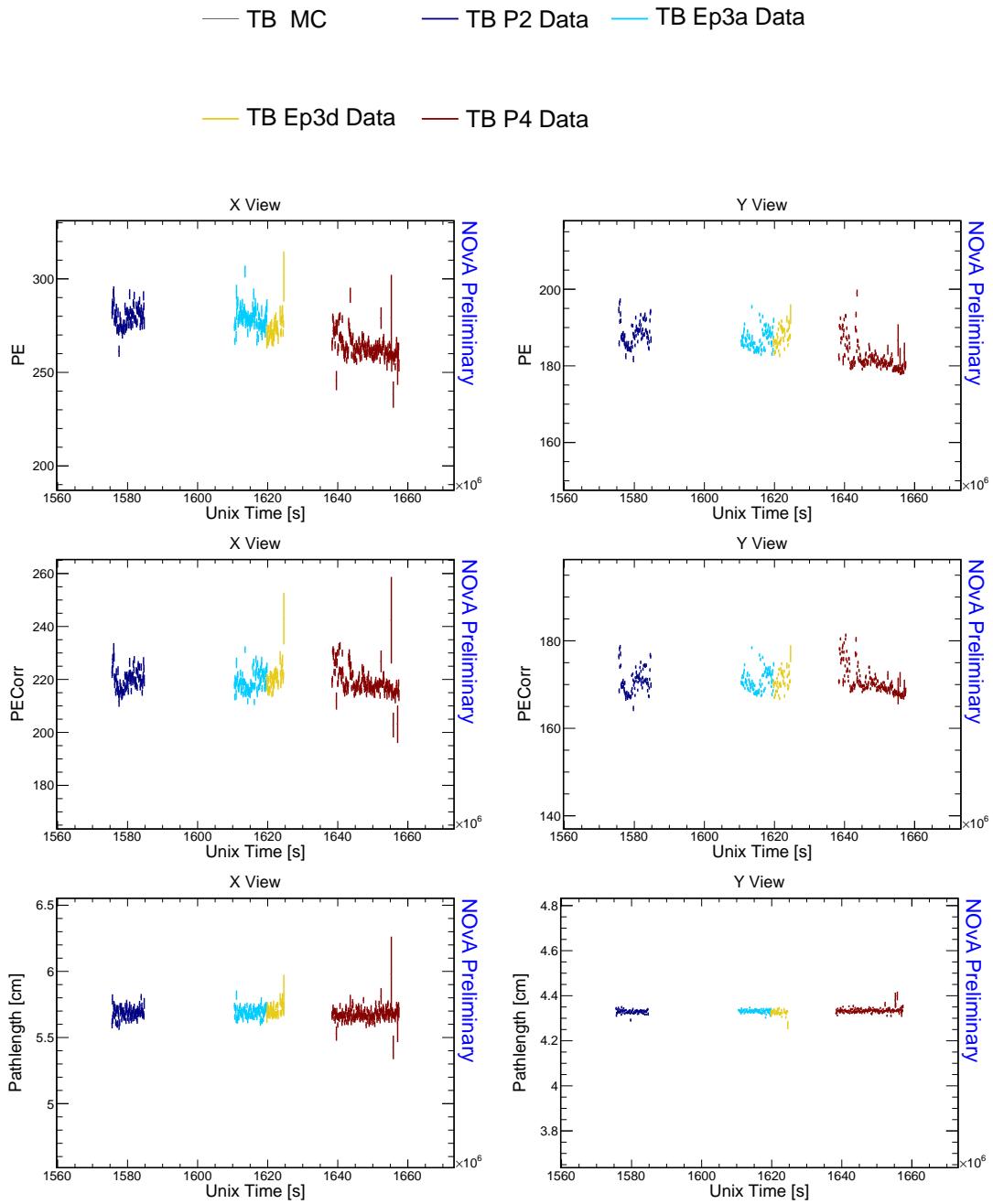


Figure A.8: Distributions of stopping muons within a 1-2 m track window from the end of their tracks across the event UNIX time.

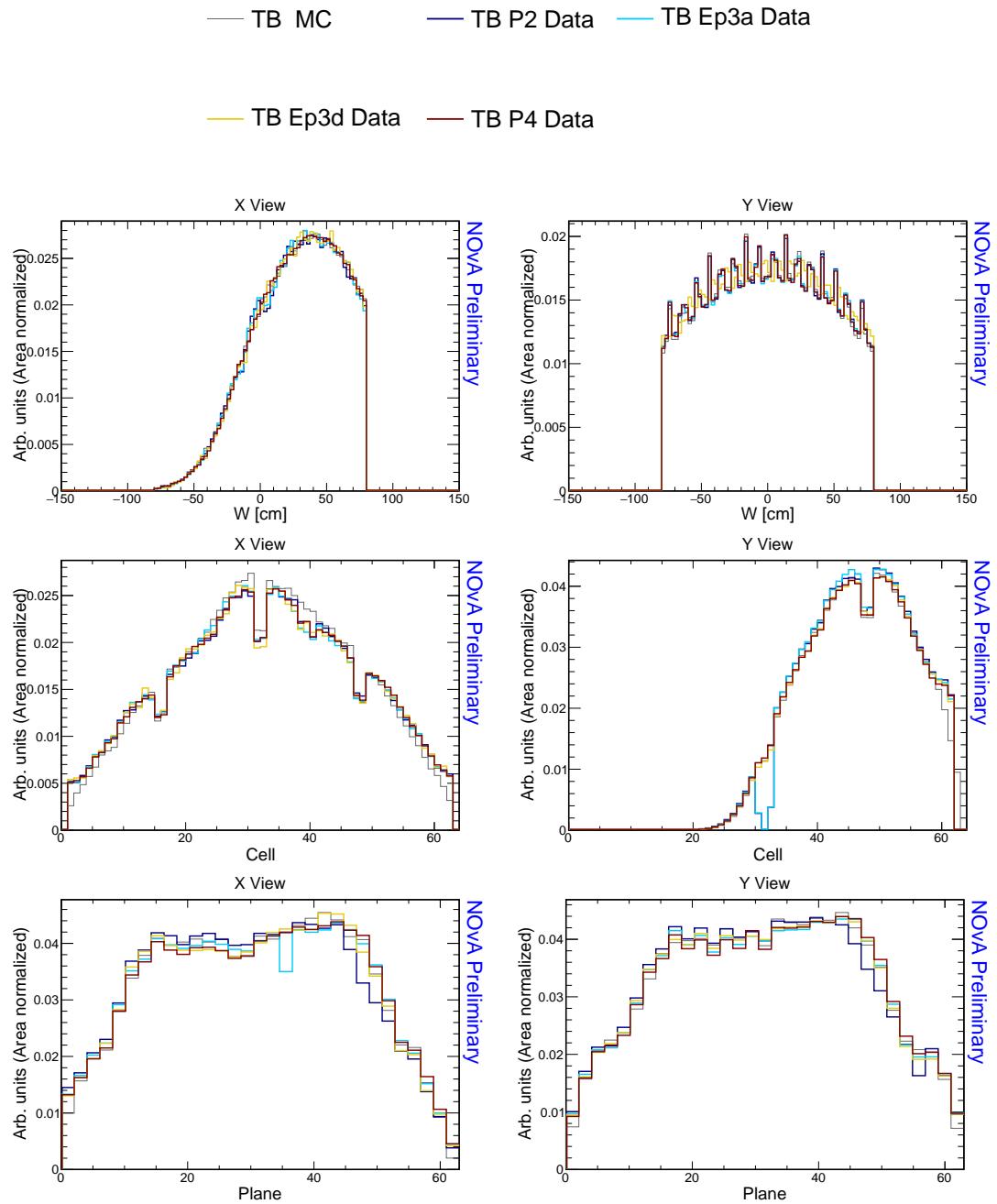


Figure A.9: Distributions of stopping muons within a 1-2 m track window from the end of their tracks.

A.2 Distributions for Through-going Muons

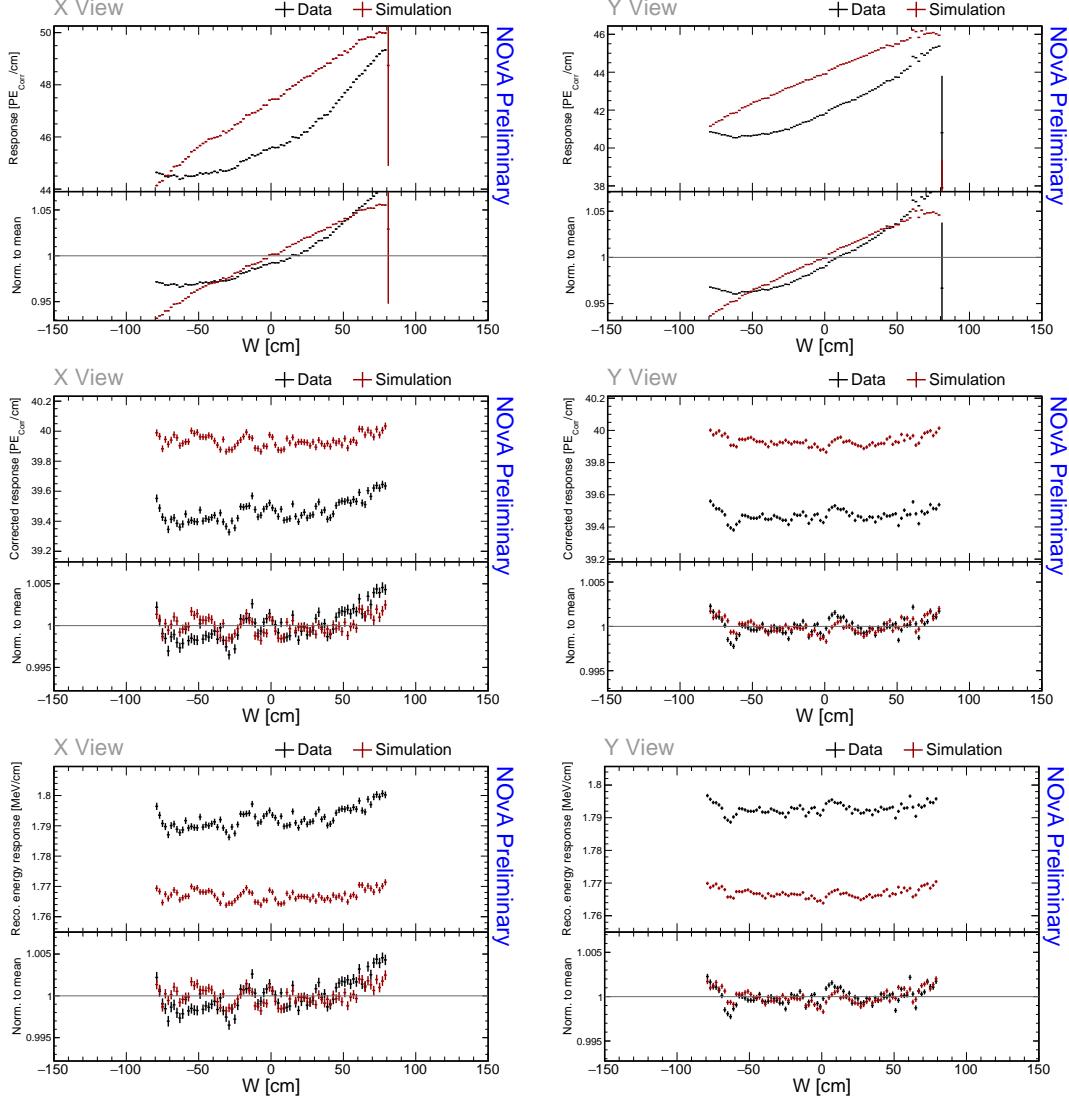


Figure A.10: Distributions of through-going cosmic muons with $w \in (-80, 80)$ cm as a function of w for stable runs in the Test Beam period 4 data (black) and data-based simulation (red). Bottom panel of each plot shows the ratio of each bin and the mean y axis, separately for data and simulation. Discrepancy in the right-most bin is solely due to binning.

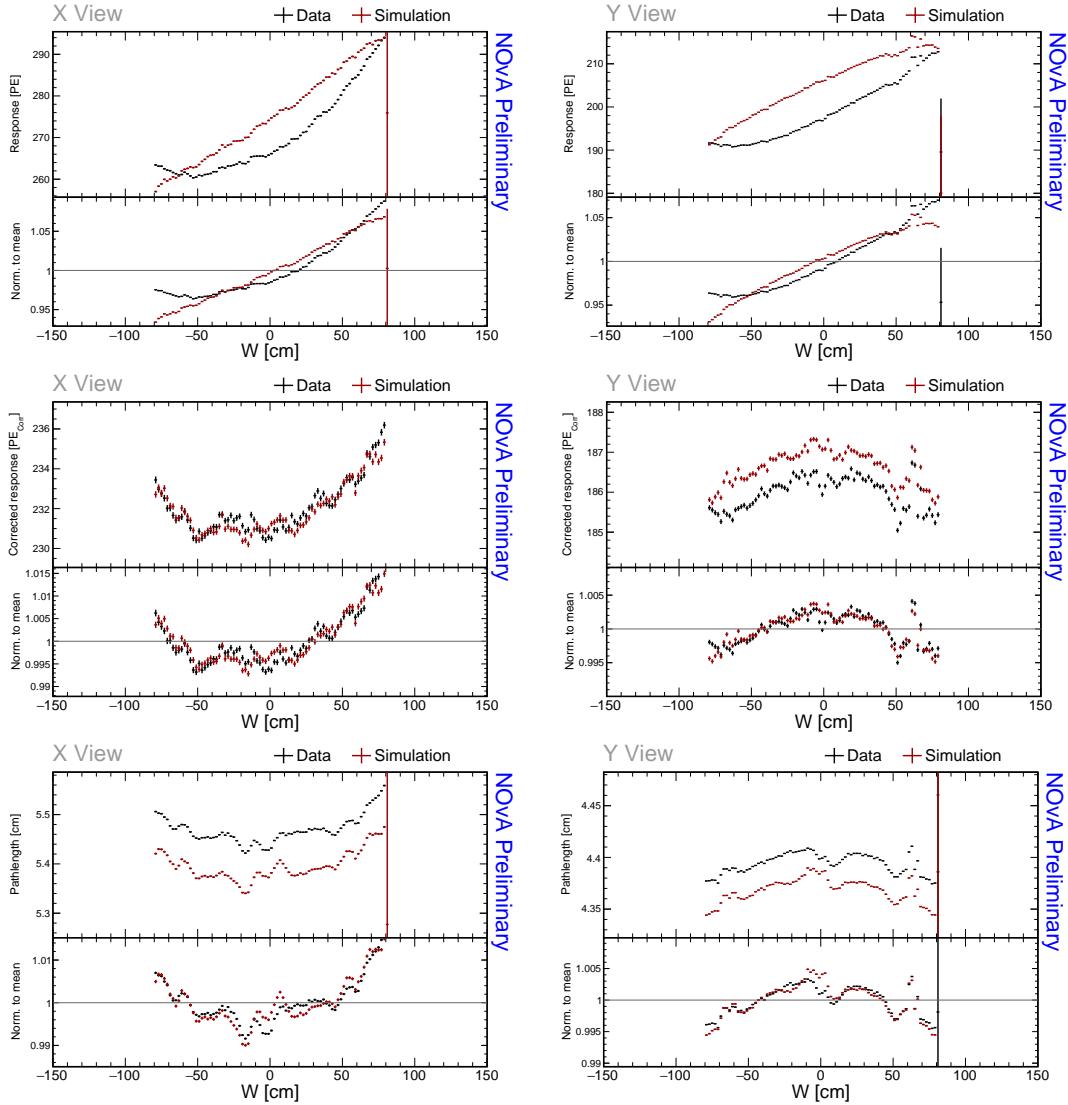


Figure A.11: Distributions of through-going cosmic muons with $w \in (-80, 80)$ cm as a function of w for stable runs in the Test Beam period 4 data (black) and data-based simulation (red). Bottom panel of each plot shows the ratio of each bin and the mean y axis, separately for data and simulation. Discrepancy in the right-most bin is solely due to binning.

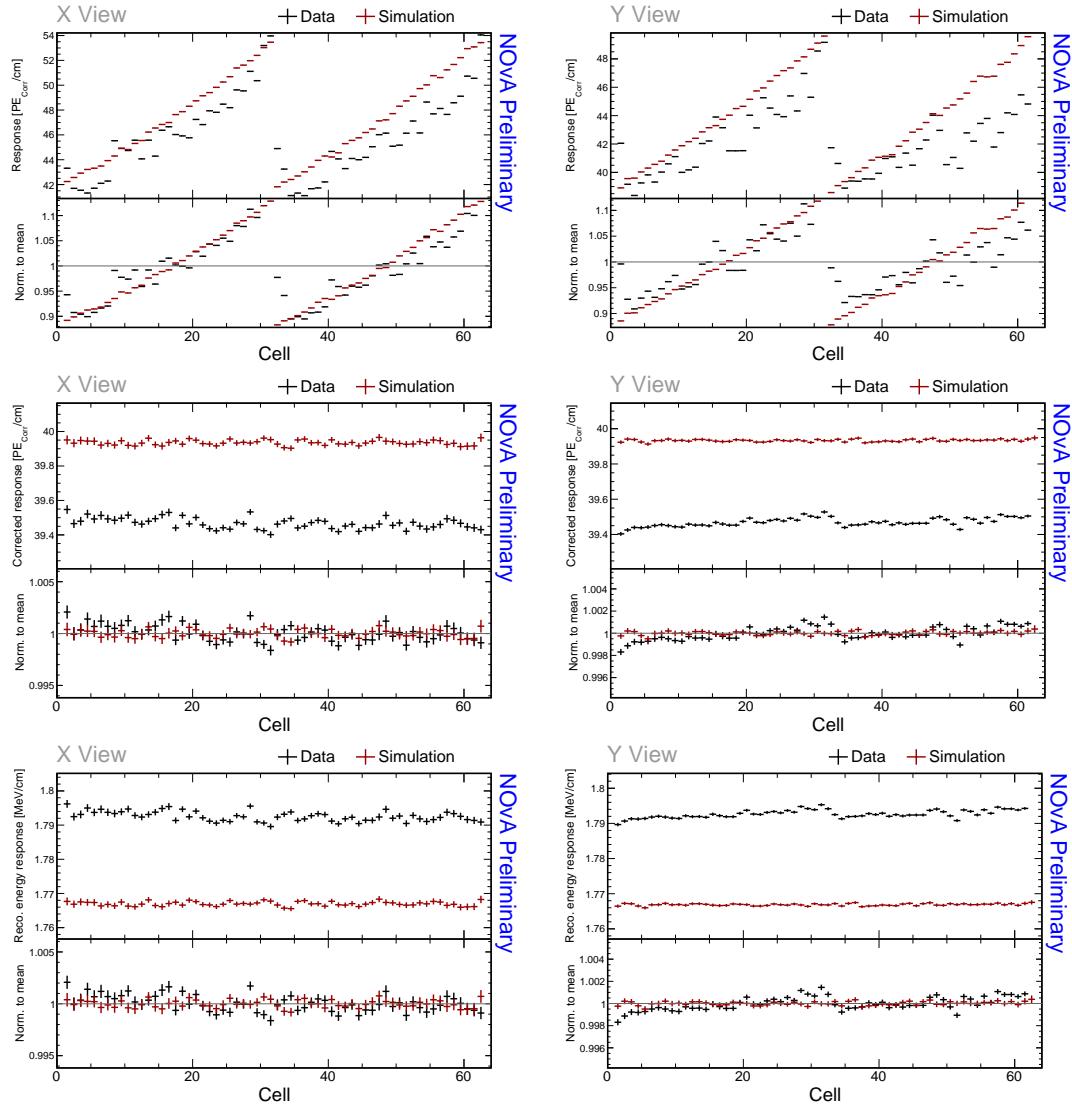


Figure A.12: Distributions of through-going cosmic muons with $w \in (-80, 80)$ cm as a function of cell number for stable runs in the Test Beam period 4 data (black) and data-based simulation (red). Bottom panel of each plot shows the ratio of each bin and the mean y axis, separately for data and simulation.

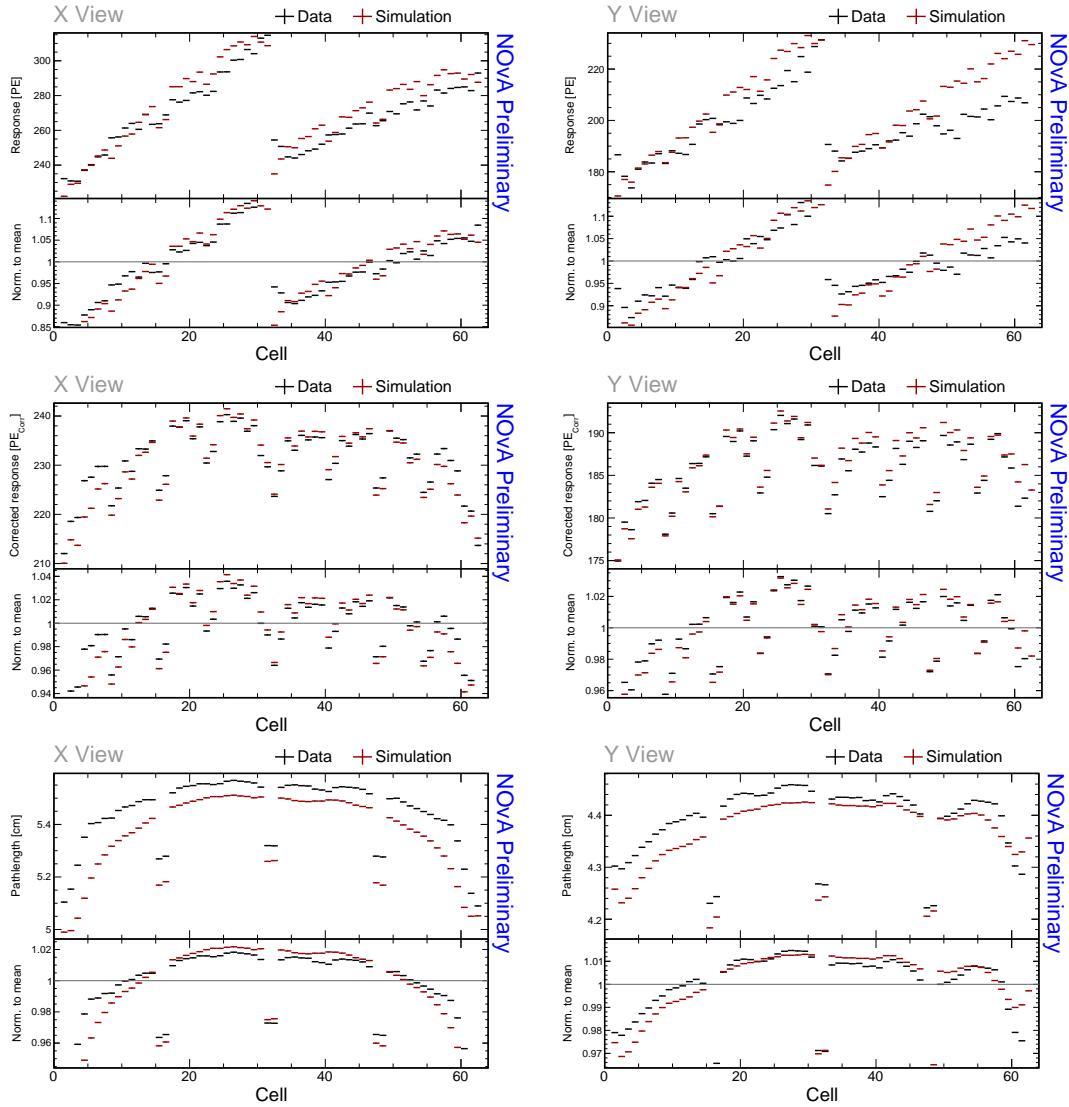


Figure A.13: Distributions of through-going cosmic muons with $w \in (-80, 80)$ cm as a function of cell number for stable runs in the Test Beam period 4 data (black) and data-based simulation (red). Bottom panel of each plot shows the ratio of each bin and the mean y axis, separately for data and simulation.

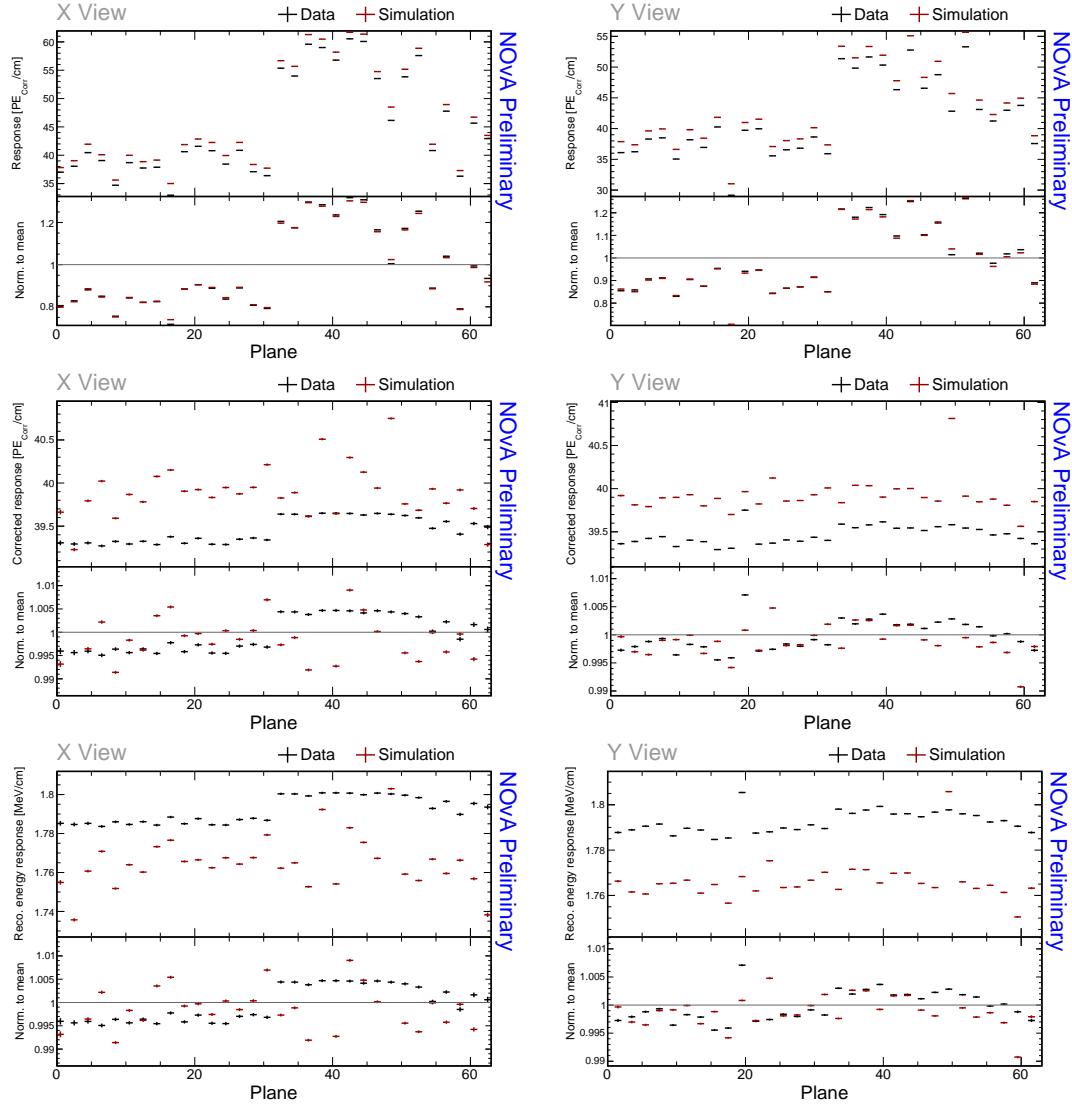


Figure A.14: Distributions of through-going cosmic muons with $w \in (-80, 80)$ cm as a function of plane number for stable runs in the Test Beam period 4 data (black) and data-based simulation (red). Bottom panel of each plot shows the ratio of each bin and the mean y axis, separately for data and simulation.

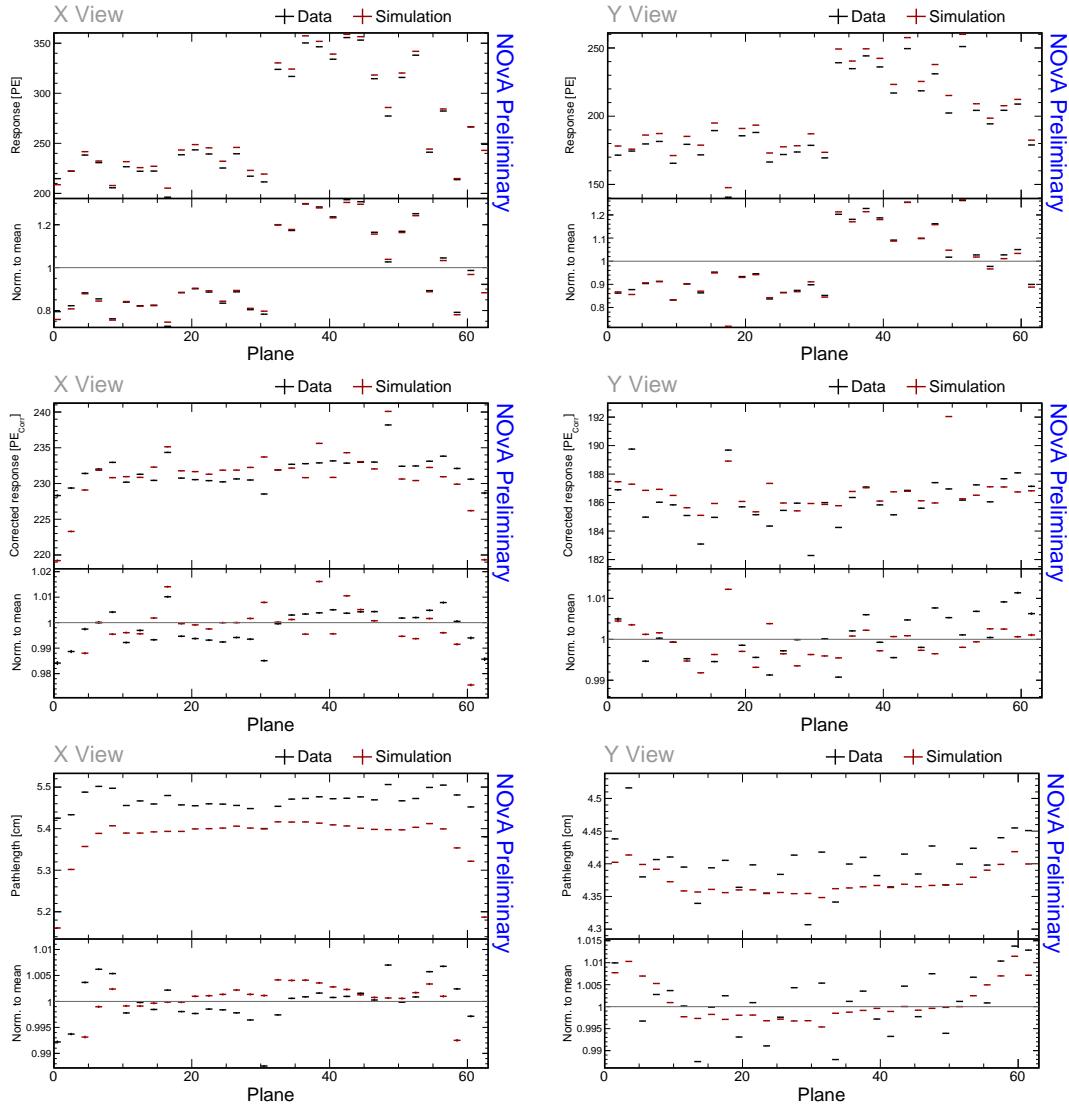


Figure A.15: Distributions of through-going cosmic muons with $w \in (-80, 80)$ cm as a function of plane number for stable runs in the Test Beam period 4 data (black) and data-based simulation (red). Bottom panel of each plot shows the ratio of each bin and the mean y axis, separately for data and simulation.

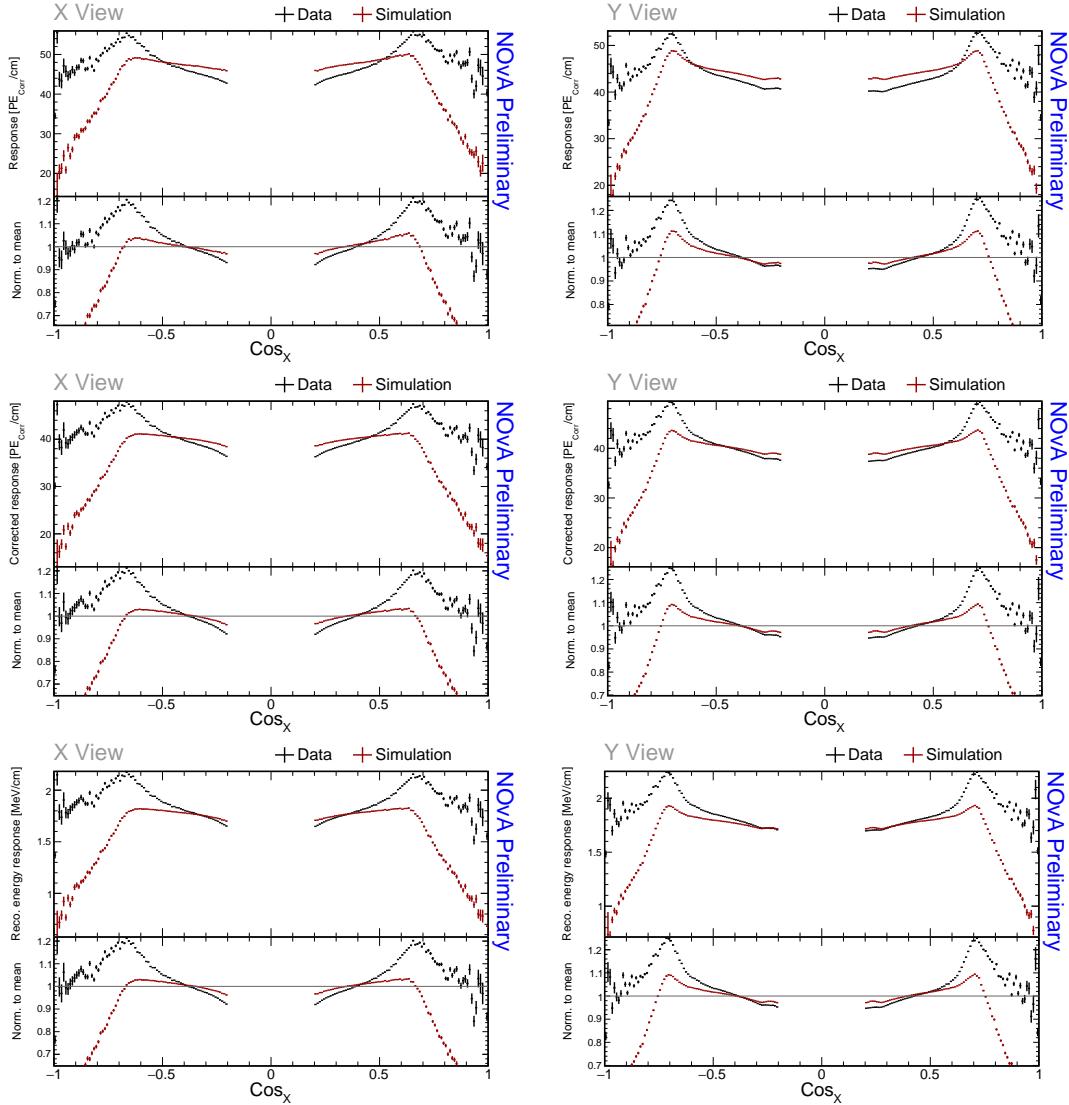


Figure A.16: Distributions of through-going cosmic muons with $w \in (-80, 80)$ cm as a function of the cosine of the angle from the x axis for stable runs in the Test Beam period 4 data (black) and data-based simulation (red). Bottom panel of each plot shows the ratio of each bin and the mean y axis, separately for data and simulation.

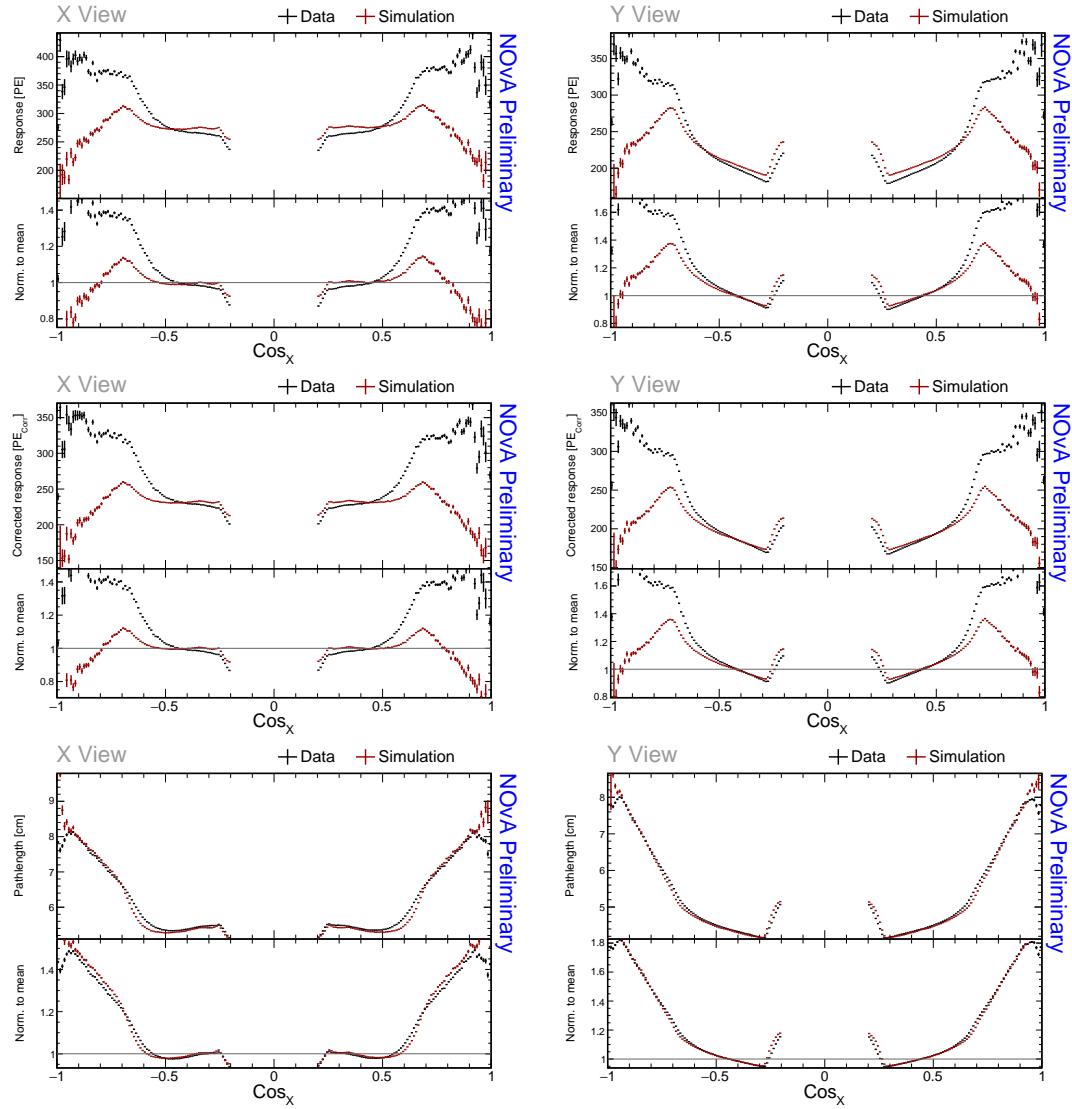


Figure A.17: Distributions of through-going cosmic muons with $w \in (-80, 80)$ cm as a function of the cosine of the angle from the x axis for stable runs in the Test Beam period 4 data (black) and data-based simulation (red). Bottom panel of each plot shows the ratio of each bin and the mean y axis, separately for data and simulation.

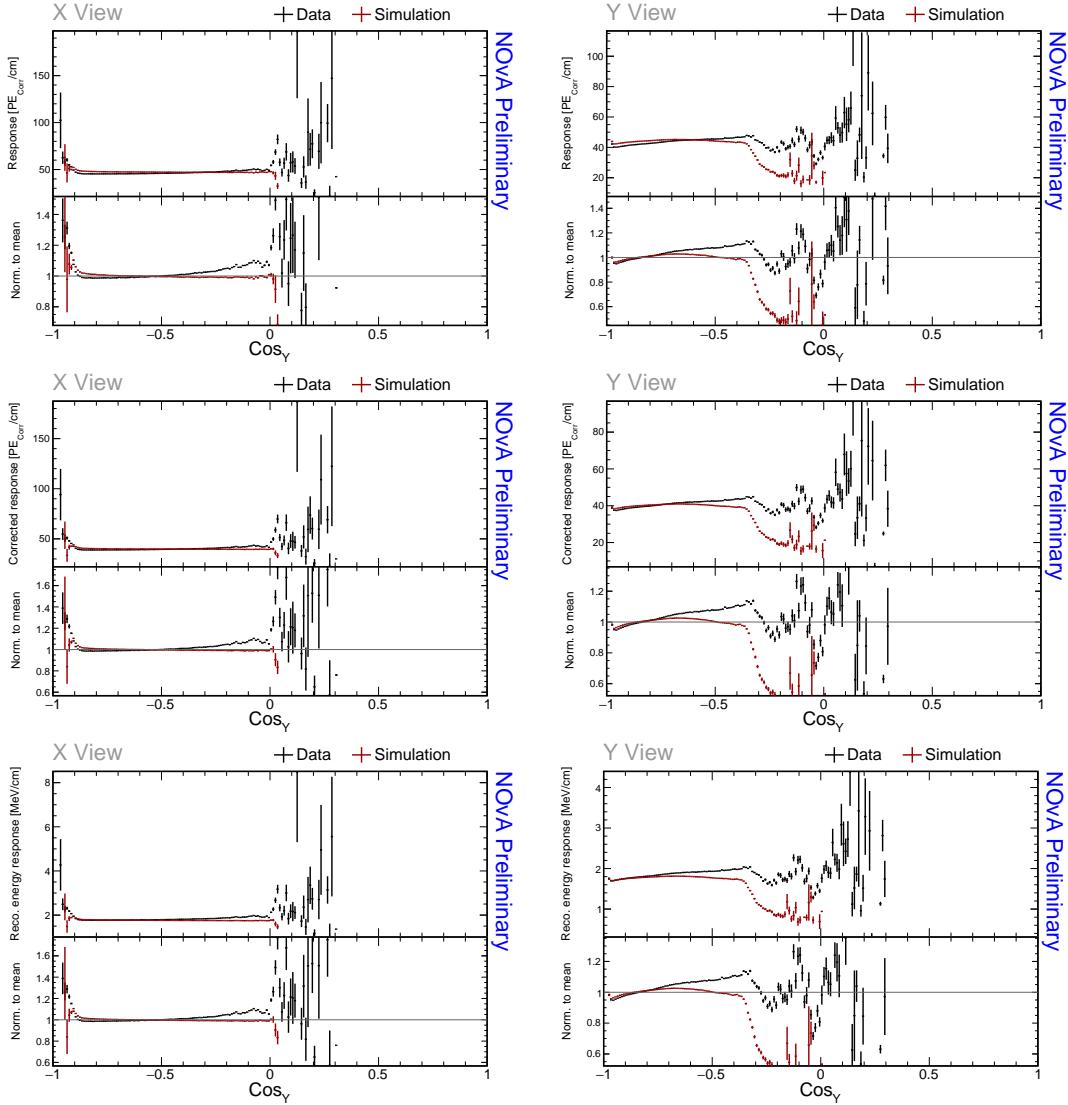


Figure A.18: Distributions of through-going cosmic muons with $w \in (-80, 80)$ cm as a function of the cosine of the angle from the y axis for stable runs in the Test Beam period 4 data (black) and data-based simulation (red). Bottom panel of each plot shows the ratio of each bin and the mean y axis, separately for data and simulation.

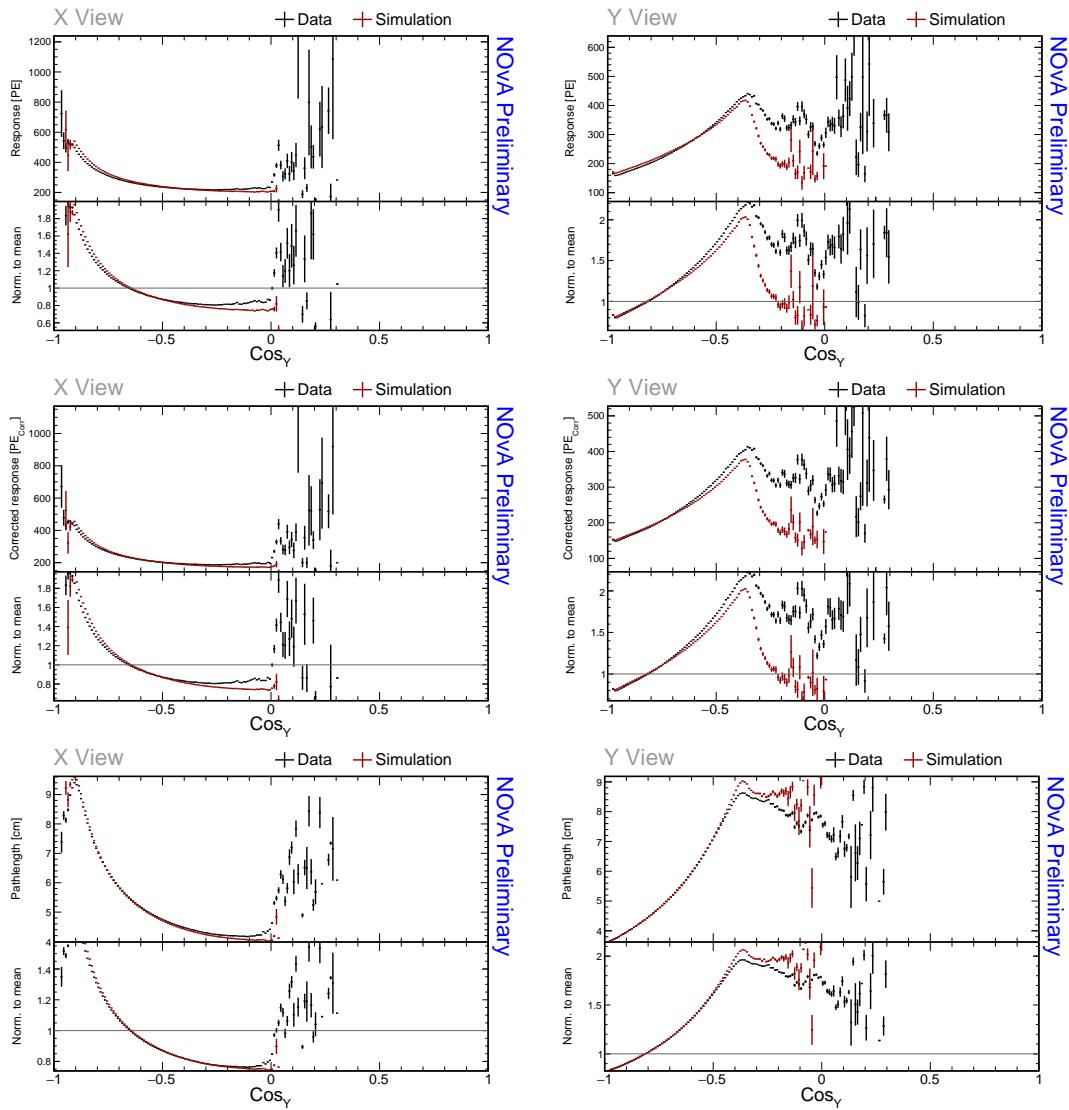


Figure A.19: Distributions of through-going cosmic muons with $w \in (-80, 80)$ cm as a function of the cosine of the angle from the y axis for stable runs in the Test Beam period 4 data (black) and data-based simulation (red). Bottom panel of each plot shows the ratio of each bin and the mean y axis, separately for data and simulation.

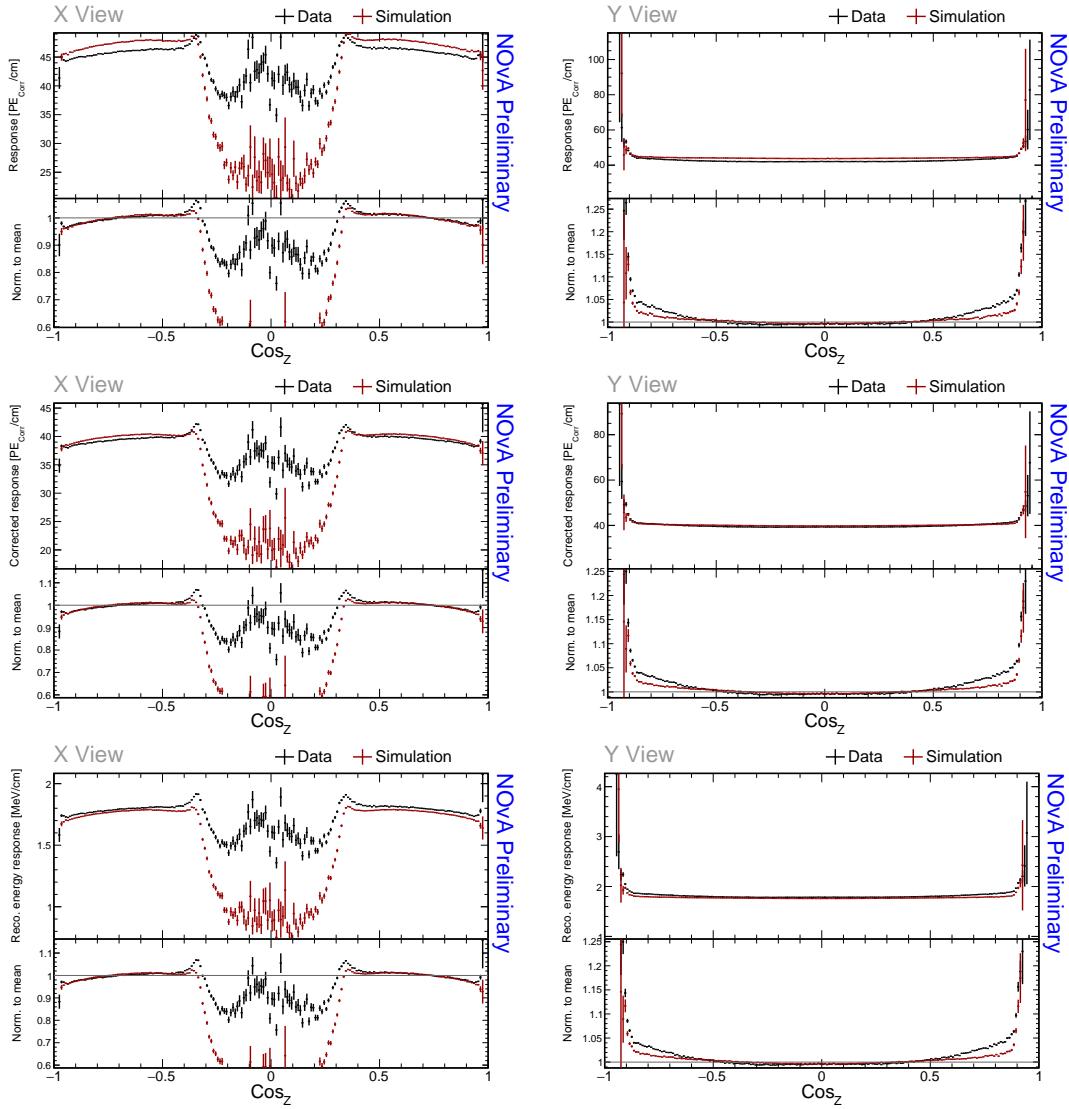


Figure A.20: Distributions of through-going cosmic muons with $w \in (-80, 80)$ cm as a function of the cosine of the angle from the z axis for stable runs in the Test Beam period 4 data (black) and data-based simulation (red). Bottom panel of each plot shows the ratio of each bin and the mean y axis, separately for data and simulation.

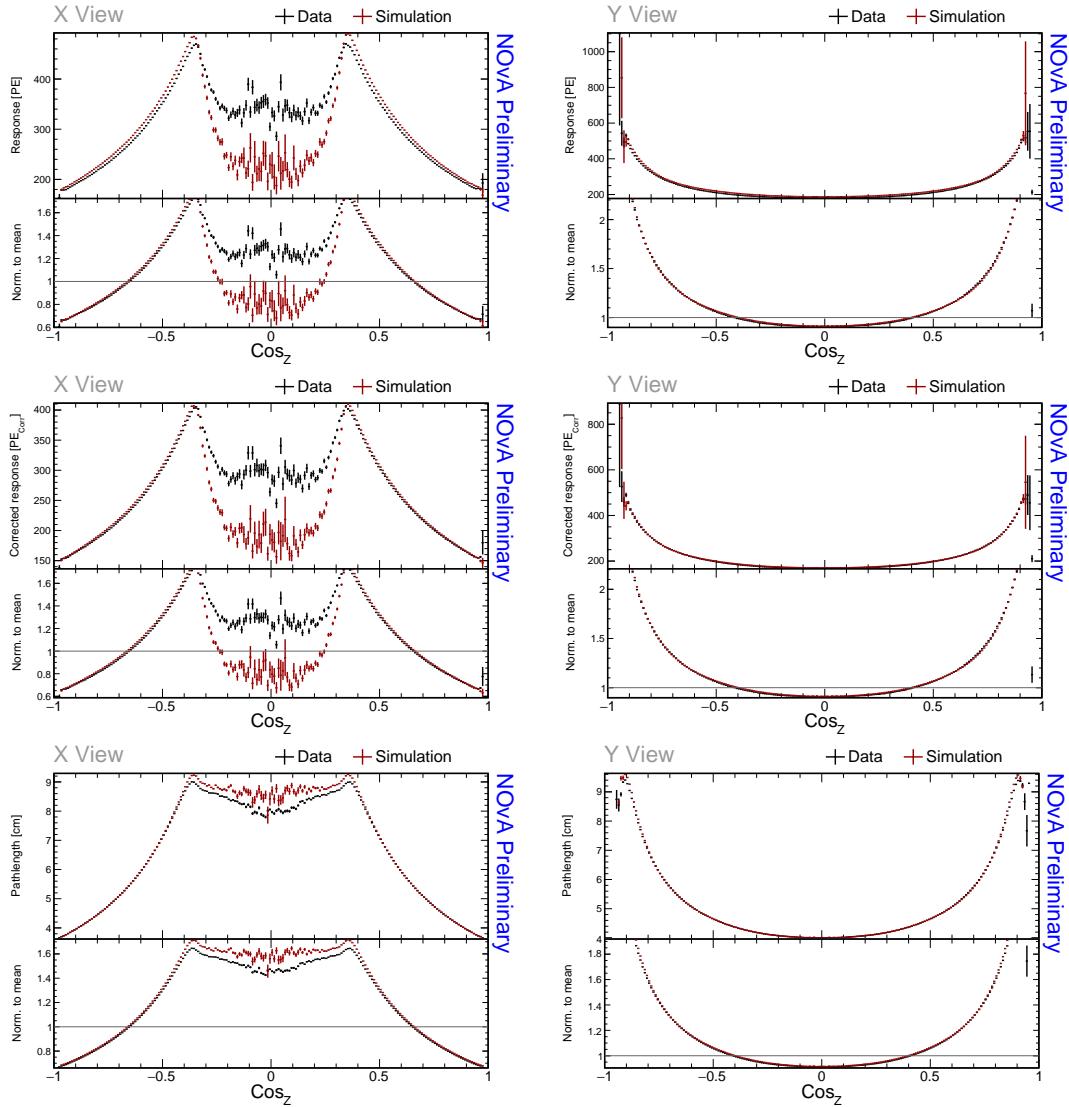


Figure A.21: Distributions of through-going cosmic muons with $w \in (-80, 80)$ cm as a function of the cosine of the angle from the x axis for stable runs in the Test Beam period 4 data (black) and data-based simulation (red). Bottom panel of each plot shows the ratio of each bin and the mean y axis, separately for data and simulation.