Analysis of DHCAL Muon Data

The CALICE Collaboration¹

Abstract

This note reports preliminary results of the large CALICE Digital Hadron Calorimeter prototype (DHCAL) from beam tests with muons. The DHCAL consists of 38 layers instrumented with Resistive Plate Chambers (RPCs) and interleaved with 2 cm thick steel absorber plates. The prototype was tested in October 2010 and January 2011 in the Fermilab FTBF test beam. In the January run part of the Tail Catcher and Muon Tracker (TCMT), located behind the DHCAL, was also instrumented with RPCs. In the final configuration the system counted 51 1x1 m^2 layers and over 470,000 readout channels. The response to muons is utilized to align the various layers of the calorimeter, to scan the response across a single pad, to measure the efficiency and pad multiplicity, and last but not least to establish the calibration constants for each RPC in the stack. The response is simulated with GEANT4 and a standalone program, RPCSIM, emulating the RPC response. The various parameters of the RPCSIM program are tuned to reproduce the measured performance of the detector.

This note contains preliminary CALICE results, and is only for the use of members of the CALICE Collaboration and for others, to whom permission has been given.

¹ Corresponding author: José Repond, repond@hep.anl.gov

1. Introduction

This note describes the analysis of muon events collected with the CALICE DHCAL in the Fermilab test beam. Muons produce straight tracks in the DHCAL, which are utilized to establish the response of the Resistive Plate Chambers (RPCs) in the stack. The response is characterized by two quantities: the efficiency of detecting Minimum Ionizing Particles (MIPs), ε , and the pad multiplicity (the number of pads firing for a single MIP traversing the chamber, requiring at least one hit), μ . In addition, the muon tracks provide an excellent tool to align the various layers in the DHCAL geometrically. The position resolution of the tracks is also sufficient to perform detailed scans of the response across the 1×1 cm² readout pads.

2. Brief description of the DHCAL

The DHCAL consists of 38 active layers, equipped with RPCs and interleaved with absorber plates. The layers are inserted into the CALICE Analog HCAL structure, placed on a movable stage. Each absorber plate consists of a 16 mm thick steel plate and the two covers of the active layers, one being 2 mm thick Copper, the other 2 mm thick Steel. A given layer contains three RPCs, each with the dimensions of 32×96 cm², and stacked on top of each other². Due to the rim of the RPCs there is a thin inactive region between the three chambers in a layer. Also the glass plates of an RPC are held apart using four fishing lines with a diameter of 1.15 mm and spaced 5 cm apart. A schematic of the cross section of a chamber is shown in Fig. 1.

Figure 1. Schematic of the cross section of an RPC (not to scale).

The RPCs are read out with readout boards, consisting of a Pad- and a Front-end board and located directly on the chambers, as shown in Fig.1. Each layer is seen by six readout boards with the dimensions of 32 x 48 cm². The readout pads are 1 x 1 cm² in size, thus a layer contains 9,216 readout channels. Each array of 8 x 8 pads is connected to a front-end ASIC, the DCAL chip. The latter applies a common threshold (variable within the range of a few fC to about 700 fC) to all 64 channels. The data from the DCAL chip consist of a timestamp (with a resolution of 100 ns) and the pattern of hits

 2^2 A right-handed coordinate system has been defined, where x is horizontal, y is vertical and z is in the beam direction. The origin of the system was placed in the right bottom corner of the first layer.

above threshold (hence the denotation of digital). The 24 chips of a board are connected to a data concentrator, located on the outer edge of the Front-end board. Figure 2 shows a photograph of the top a readout board. The Pad-board is glued onto the other side of the Front-end board and is thus not visible in this photograph.

A photograph of the DHCAL stack before cabling is shown in Fig. 3. The stack contains 114 individual RPCs and 350,208 readout channels.

Figure 2. Photograph of a DHCAL readout board. The readout pads are on the other side of the board and thus not visible.

During the course of the January 2011 test beam campaign, the Scintillator layers of the Tail Catcher and Muon Tracker (TCMT), located behind the DHCAL, have been replaced with layers identical to the ones in the DHCAL stack. By the end of the running period 13 of the 16 slots in the TCMT had been equipped with RPCs. This increased the total readout channel count to 470,016.

Figure 3. Photograph of the DHCAL stack before cabling.

3. Monte Carlo simulation of the DHCAL

The test beam set-up has been simulated with a Monte Carlo program based on the GEANT4 package [1] and a standalone program RPCSIM for the simulation of the response of RPCs. The GEANT4 simulation includes the details of the calorimeter stack and the RPCs. The range cut in the GEANT4 program was set to 0.005 mm. Muon data are generated with a momentum of 32 GeV/c and the lateral beam profile and slope distributions of the beam (see below). The spatial coordinates of any energy deposition in the gas gap of an RPC is recorded for further analysis. In the following these energy depositions are named *points*.

For each generated *point*, the RPCSIM program generates a signal charge Q, distributes this charge over the pads, sums up all charges on a given pad, and applies a threshold *T* to identify the pads with hits.

The signal charges are generated according to the measured [2] spectrum of avalanche charges utilizing cosmic rays. The spectrum is fitted to the following functional form:

$$
N(Q) = \alpha Q^{\beta} e^{-\gamma Q},\tag{1}
$$

where the parameters α , β and γ depend on the operating high voltage. An additional parameter \mathbf{Q}_{0} was introduced to accommodate possible differences between the charge distributions as measured in the laboratory and obtained in the test beam set-up:

$$
Q' = Q + Q_0 \tag{2}
$$

As a function of lateral distance R to a given *point* the induced charge in the plane of the readout pads is assumed to decrease exponentially with a slope *a*:

$$
dQ'/dR = -(Q'/a) \cdot e^{-R/a}
$$

An additional parameter was introduced to simulate a possible inefficiency of the chambers for additional *points* within a radius *dcut* from a given *point*. Due to the low number of *points* in muon induced events the response to muons is only weakly dependent on this parameter.

The four parameters Q_0 , T , a and d_{cut} were tuned to reproduce the distribution of hits in individual chambers, see below. The Monte Carlo generated events were formatted the same way as the test beam data and were analyzed by identical programs.

4. Trigger and data taking

The stack was located in the Fermilab MTest beam line. Broadband muons were generated with the

secondary beam tuned to +32 GeV/c and impinging on a 3 m Iron beam blocker placed in the beam line upstream of the experimental area. The data acquisition rates were typically 500 events/spill, where the spills lasted approximately 3.5 seconds.

The data acquisition was triggered by two large Scintillator paddles, with an area of 1 x 1 m², located upstream of the DHCAL stack and behind the TCMT, see Fig.4. The paddle in front was placed about 4.02 ± 0.02 meters upstream of the first absorber plate of the DHCAL. The paddle in the back was implemented into the TCMT structure.

Figure 4. Schematic of the set-up for collecting muon events.

Muon events were collected in three different running periods, see Table I for details. The numbers of events entered into Table I also contain standard CALICE calibration events and a significant fraction of spurious triggers. The number of muon events is approximately 65% of the values quoted in the last column.

To illustrate the quality of the data taken, Figure 5 shows two sample events collected with the DHCAL and the 13 layer TCMT. Note the absence of random hits due to noise and distinct from the muon track. The events in Fig. 5 are shown before application of any noise filtering algorithm.

5. Clustering of hits and tracking of muons

The data are recorded as a timestamp (with a resolution of 100 ns) and the hit pattern for any ASIC with at least one hit. For a given trigger, hits in seven consecutive time bins are recorded. It is observed that the large majority of hits are contained in two time bins, whereas the other time bins only see orders of magnitude less data. A study of the geometrical location of the hits in the outer time bins shows that most of these are not related to the event and appear to be random. Therefore, in the following only hits in the central two time bins are utilized.

Figure 5. Event display of two muon tracks traversing the DHCAL and the TCMT.

All hits in a given layer are clustered using a closest-neighbor clustering algorithm. Pads hit are required to share one side to be considered part of the same cluster. The position of a cluster is determined as the unweighted average, $x_{cluster}$ and $y_{cluster}$, of its pads.

In order to measure the response in layer i, clusters in all layers (in the following named *tracking layers*), apart from layer i, are used to reconstruct straight tracks. The track reconstruction requires that there be at most one (1) cluster in any tracking layer. Only clusters with at most four (4) hits are accepted for tracking. Finally, for a track to be accepted, it is required to contain at least ten (10) tracking clusters.

Tracks are reconstructed by performing two least-square linear fits in x/z and z/y, respectively. Fits use the unweighted average position of the clusters in the tracking layers. The errors on the positions are assumed to be 1 cm.

A combined, reduced χ^2/N_{track} of the fits is calculated as

$$
\chi^2 / N_{\text{track}} = \sum_{j \neq i} \frac{(x_{\text{cluster}}^j - x_{\text{track}}^j)^2}{1} + \sum_{j \neq i} \frac{(y_{\text{cluster}}^j - y_{\text{track}}^j)^2}{1}
$$

where the sums run over all tracking clusters, $x_{\text{track}}^j / y_{\text{track}}^j$ is the inter/extrapolated position of the track in layer j, and N_{track} is the number of clusters used in the fitting. Only tracks with χ^2/N_{track} smaller than 1.0 are retained for further analysis.

Once a track is reconstructed the track position is inter/extrapolated to layer i, the layer being

investigated and not used for tracking. Looping over all clusters in layer i, the closest cluster within a radius $R = 2.5$ cm is considered a match. The number of hits contained in the matching cluster is retained for measuring the pad multiplicity, μ . If no cluster is found within R, the layer is considered inefficient for this track.

6. Geometrical Alignment of the Layers

The horizontal and vertical alignment of the layers is investigated using the residuals, i.e. differences of the inter/extrapolated track and the matching cluster position in both x and y:

$$
R^{i}_{x} = x^{i}_{cluster} - x^{i}_{track}
$$

$$
R^{i}_{y} = y^{i}_{cluster} - y^{i}_{track}
$$

Figure 6 shows a typical distribution of such residuals. A small fraction of the distributions, related to layers with larger off-sets, were seen to display double peaks. The two peaks can be understood as a binning effect related to the finite1x1 cm² pad size of the readout. This has been confirmed by a simple toy Monte Carlo, which generated hits at a particular position followed by the simulation of the RPC response using RPCSIM. The mean of the residual distributions versus layer number and for each readout board in a given layer individually is shown in Fig.7. The variations in y are seen to be significantly smaller than in x. This can be explained as a result of the fact that the bottom plate of the CALICE stage provides a natural reference point in y, whereas the layers were only aligned in x 'by eye'. The systematically higher values of the residuals in y in the center of the stack are presumably due to the weight of the absorber plates bowing the bottom plate of the center of the stage. Also, notice the strong correlations between the positions of the different boards in a given layer.

Figure 6. Typical distribution of the residual of the position of the cluster and the extrapolated track in a given readout board.

The means of the residual distributions are used to align the layers. Due to the observed strong

correlation between the positions of individual readout boards in a given layer, the position of the layers as a whole are aligned. Figure 8 shows the mean of the residuals as function of layer number after alignment. Notice the distinct improvement, compared to Fig.7.

The distribution of the means of the residuals for the readout boards is shown in Fig. 9 after alignment in both x and y. The alignment procedure reduced the RMS of the distributions from 1.4 mm (0.23 mm) to 0.57 mm (0.13 mm) in x (y). The corresponding RMS for the entire layers is, as expected, significantly smaller: 70 (14) μ m in x (y).

As a cross check of the procedure, the alignment obtained with one data run was applied to other runs. The method was shown to provide similar reductions in the value of the residuals.

Figure 7. Mean of the position residuals in x (left) and y (right) before alignment. The different colors indicate the six readout boards in a layer. The units are in [cm].

Figure 8. Mean of the residual distributions after alignment: in x (left) and in y (right). The different colors indicate the six readout boards in a given layer. The units are in [cm].

Figure 9. Distribution of the means of the residuals for the readout boards in the DHCAL after alignement: in x (left) and in y (right).

7. Tuning of the beam simulation

Using reconstructed tracks obtained with the first ten layers of the stack (to minimize multiple scattering effects) both the position of the clusters in the first layer and the slope of the tracks in x/z and y/z was determined. The slopes in x/z, y/z and in both dimensions combined are shown in Fig. 10 for both data and Monte Carlo simulation. After adjusting the slopes of the generated muons in GEANT4, adequate agreement is observed between data and simulation.

Figure 10. Slope of muon tracks reconstructed using the clusters of the ten foremost layers of the DHCAL stack: data (left) and MC (right). Note the logarithmic scales.

8. Response in 'clean' areas

In order to provide a data set to tune the RPCSIM parameters, the response of the RPCs has been determined away from special regions, such as

- Dead ASICs (cut away an area of $8 \times 8 \text{ cm}^2$ plus a rim of 1 cm)
- Edges in x (cut away 0.5 cm from the 2 edges in x)
- Edges in y (cut away 0.5 cm from the 6 edges in y)
- Fishing lines (cut away ± 1 cm around fishing line)
- Layer 27 (which showed unusually high pad multiplicity)³

The list of dead ASICs as a function of time (or run number) has been determined using the nontriggered data runs [3].

Figure 11 shows the distribution of the number of pads for the 'clean' regions of all layers combined. The data are displayed as histogram and the simulation as data points. The statistics is very high, with approximately twice the number of entries in the data compared to the Monte Carlo.

Figure 11. Distribution of number of hits in 'clean' areas of the detector: linear y scale (left) and logarithmic y scale (right). Histogram is data and data points are simulation based on the parameters established in the Vertical Slice Test [4].

The simulated distribution shown in Fig. 11 was obtained with the RPCSIM parameters tuned to the Vertical Slice Test (VST) data [4]. Note that the DHCAL efficiency is significantly higher and that the pad multiplicity has decreased compared to the VST. This is in part due to the use of thinner glass on the anode side.

The four parameters of the RPC simulation were tuned using a χ^2 comparison between the histograms of the distributions of the measured and simulated number of hits. Figure 12 shows this χ^2 as function

 3 In later runs the RPC with the high pad multiplicity in layer 27 was replaced.

of the different tuning trials. Note that due to the high statistics of both samples, the χ^2 values are in general quite large. After trial #70 no further improvement could be obtained by varying the four parameters of the RPC response emulation.

Figure 12. χ^2 values obtained from the comparison of the histograms of the measured and simulated distribution of the number of hits in 'clean' regions of the detector. The first point (trial #1) corresponds to the parameter set obtained with the Vertical Slice Test [4].

Table II lists the values of the four RPCSIM parameters yielding the minimal χ^2 and Fig.13 shows the corresponding distribution of number of hits compared to data. The changes in parameter values between the current data and the VST can be explained as due to changes in the operating point of the RPCs, the use of thinner glass, and the higher values of the surface resistivity of the resistive paint on the anode side.

Parameter	VST value	DHCAL value
a	0.170	0.084
\mathbf{r}	0.600	0.400
d_{cut}	0.100	0.090
	0.200	-0.062

Table II. List of the parameters and their values used in the simulation of the RPC response.

A larger tail extending to higher multiplicities is observed in the data compared to the simulation. This excess amounts to about 0.6%. The reasons for this excess are currently not understood. However, it appears unlikely as due to a shortcoming of the RPC simulation. Pad multiplicities above five are expected from events where the muon interacted, generating a hard δ ray. Using a different track selection, where events containing clusters with higher multiplicities, except in the layer under investigation, are eliminated, reduces this discrepancy significantly. Further studies, such as variations of the muon momentum in the simulation or studies of the location of these high multiplicity events in the detector planes, are necessary to shed further light onto this difference.

Figure 13. Distribution of number of hits in 'clean' areas of the detector: linear y scale (left) and logarithmic y scale (right). The histogram represents the data and the data points are simulation after tuning of the RPCSIM parameters using the present data set.

Table III summarizes the performance variables for the RPCs in 'clean' regions. Adequate agreement between data and simulation is observed.

Variable	Data	Monte Carlo
Efficiency $[\%]$	93.6	
Average pad multiplicity	1.563	
Mean of the multiplicity distribution	1.461	.443

Table III. Comparison of the performance variables of RPCs in 'clean' regions between data and simulation.

Note that the present values for the performance variables are somewhat arbitrary as they depend on the track selection. A comprehensive systematic study of these dependencies is still forthcoming.

9. Scan across readout pads

Tracks inter/extrapolated into 'clean' regions are used to scan the response across a $1x1$ cm² pad. In order to increase the statistics the x/y – positions of all pads are overlaid. Figure 14 shows the average number of hits as function of x, y, and R on the pad for both data and simulation,. Here R is defined as

$$
R = \sqrt{(x - x_{center})^2 + (y - y_{center})^2}
$$

and x_{center} and y_{center} denote the center of the pad. To better illustrate the $x(y)$ dependence, for these plots only tracks in a band of $0.25 < y < 0.75$ ($0.25 < x < 0.75$) have been utilized. The agreement between data and simulation is quite remarkable, given the fact that the present results were obtained by implementing the response on a first principle basis and not explicitly as function of position on the pad.

Figure 14. Average number of hits (in 'clean' regions of the detector) as function of position on the readout pads: data (left) and simulation (right). The red lines are empirical fits to 2^{nd} second order polynomials.

10. Response over the entire plane

Next, the response of the detector over the entire x/y – plane is investigated. Only the regions around dead ASICs are omitted, as the same dead areas are also implemented into the RPCSIM program by eliminating hits in the corresponding regions.

Figure 15 compares the measured and simulated inefficiency as function of x/y . Here the inefficiency is defined as the number of tracks where there is no matching cluster in a given layer divided by the total number of tracks. Note the following features:

- The simulation does not extend as far to the edge of the chambers in x and y as the data. The simulated efficiency vanishes about 3 cm from the edges.
- Apart from the edges the distribution in x is featureless and well simulated.
- Apart from the structure at the edges, the distribution in y shows the effects of the fishing lines and the gaps between RPCs. The effects are adequately reproduced by the simulation. The gaps between RPCs seem to be somewhat wider in the data than in the simulation.

Figure 15. Comparison of the measured (black) and simulated (red) inefficiency as function of x(left) and y (right).

The distribution of the number of hits averaged over the entire plane of the detector is shown in Fig. 16. Note that the tail towards higher values increased in both the data and in the simulation compared to the analysis of the 'clean' regions. Table IV compares the performance variables in data and simulation obtained over the entire detector plane. In general, satisfactory agreement between measurement and simulation is observed.

Figure 16. Distribution of number of hits over the entire x/y-plane of the detector: linear y scale (left) and logarithmic y scale (right). The histogram represents the data and the data points show the results of the simulation after tuning of the RPCSIM parameters using the 'clean' regions of the detector.

Variable	Data	Monte Carlo
Efficiency $[\%]$	90.9	
Average pad multiplicity	1.611	1.535
Mean of the multiplicity distribution	1.464	Δ

Table IV. Comparison between data and simulation of the performance variables of RPCs averaged over the entire detector plane.

11. Calibration of the DHCAL

In order to calibrate the response of the various layers of the DHCAL and the TCMT we revert back to the track selection in 'clean' areas of the detector. The distribution of number of hits versus layer number in the DHCAL stack is shown in Fig.17.

Figure 17. Distribution of number of hits versus layer number using only the 'clean' areas of the detector shown from different angles: logarithmic y-scale (left) and linear y-scale (right).

The distributions to the left (logarithmic y scale) and right (linear y scale) are identical, but shown from a different angle. The one to the right shows in particular the inefficiency of the chambers.

It is straightforward to show that the means of the hit distributions are identical to the products of efficiency and average pad multiplicity. For a given RPC, the mean of the hit distribution divided by the average mean of the distribution obtained in simulation $(=1.443)$ is considered to be the calibration factor of the response.

The performance variables, i.e. the efficiency, the average pad multiplicity and the calibration factors,

are shown in Fig. 18 versus layer number and separately for each RPC. Note the uniform performance from layer to layer of the DHCAL (layers 0-37). The TCMT (operating at a lower temperature) shows a lower efficiency and pad multiplicity.

Figure 18. Performance variables as function of layer number in the DHCAL (layers 0-37) and the TCMT (38-51). The values were obtained using the 'clean' regions of the detector.

The average pad multiplicity versus efficiency is shown in Fig.19 for each RPC in the DHCAL and TCMT stacks. As observed previously [5], higher efficiencies entail in general higher pad multiplicities.

Figure 20 shows the distributions of the three performance variables (efficiency, average pad multiplicity and calibration factor) for all RPCs in the DHCAL and the TCMT. Again the lower values found in all three variables correspond to the RPCs in the TCMT.

Overall the detector performance was remarkably stable during the test beam campaigns. Figure 21 compares the calibration factors averaged over all layers versus run number in the January – February, 2011 run period. The various colors indicate runs taken on the same date. The error bars correspond to the RMS of the distributions with one entry per layer. Note that in this time period the calibration fluctuated between -2.5 and +7.7%. These fluctuations are consistent with the expected change in RPC performance related to changes in the environmental conditions, such as ambient Temperature and air pressure [6].

Figure 19. Average pad multiplicity versus efficiency for the RPCs in the DHCAL and the TCMT.

Figure 20. Distribution of the performance variables obtained using the 'clean' regions of the detector. The red histograms indicate the subset of RPCs in the TCMT.

Figure 21. Calibration constants for the entire DHCAL stack (excluding layer #2, 36, and 37 which showed readout problems for part of the runs) versus run number for the January-February 2011 test beam campaign.

12. Track segment analysis

In the January/February 2011 run period, the majority of the data collected with the secondary beam was taken between run 138 and run 213, during which time period no muon calibration runs were taken. Therefore, other methods to determine the overall calibration of the detector need to be explored to cover periods without muon runs. Such methods might utilize the hit rate in self-triggered noise runs [3], the analysis of muon events produced with the secondary beam, and the analysis of track segments in the pion data. The latter promises to be a powerful tool to monitor the calibration, in particular so when operating this type of detector in a colliding beam environment. A first look at this technique is presented in the following. For more information on the secondary beam, see [7].

The track segment method uses clusters, in the following named *source clusters*, in two layers to measure the response in a third layer (*target cluster*). For instance, source clusters in layers 2 and 3 are used to measure layer 1, or source clusters in layers (i-1) and (i+1) are used to measure layer i. The source clusters are required to have less than four (4) hits and their lateral distance in the x/y-plane to be less than 3 cm. There should be no additional hits in the same layer within a radius of 7 cm around the source clusters. Target clusters are searched for within a radius of 2 cm from the inter/extrapolated line between the two source clusters. To avoid gaps due to inefficient regions (dead ASICs), source clusters may be used in layers $(i-2)$ or $(i+2)$ instead. Apart from a minimum requirement of at least three layers with hits, there are no additional event selection criteria.

Comparison of the efficiency and average pad multiplicity obtained with track segments in muon and

in pion data and with fits to muon tracks is shown in Fig. 22 versus layer number in the DHCAL. A clear correlation between the three measurements is evident. However, as expected, the different methods produce different systematic shifts. Further studies of this potentially powerful calibration tool are forthcoming.

Figure 22. Efficiency (left) and average pad multiplicity (right) versus layer number obtained with muon tracks (red histogram), track segments in muon data (blue) and track segments in pion data (green).

13. Conclusions

A preliminary analysis of muon data collected with the DHCAL has been presented in this note. In general the DHCAL data collected in October 2010 and January-February 2011 appears to be of high quality. Muon data are used to

- Align the layers in the x/y-plane,
- Establish the response in 'clean' areas away from edges or other problematic areas,
- Tune the RPC response simulation in these 'clean' areas,
- Establish the response over the entire x/y plane of the detector,
- Establish the response over the surface of a single pad,
- Determine the calibration factors for each layer,
- Determine the overall calibration of the detector as function of time.

In addition, a first comparison of the measurements of the DHCAL performance parameters as established with muon tracks, muon track segments and track segments in pion events has been presented. The results provide a first indication of the potential of track segments to monitor this type of detector in a colliding beam environment.

References

- **1.** S.Agostinelli et al., Nucl. Instrum. Meth. **A506**, 250 (2003).
- **2.** G.Drake et al., Nucl. Instrum. Meth, **A578**, 88-97 (2007).
- **3.** L.Xia, CAN entitled 'Noise studies with the DHCAL'.
- **4.** B. Bilki et.al., "Measurement of positron showers with a digital hadron calorimeter", JINST 4 P04006, 2009.
- **5.** B. Bilki et.al., "Calibration of a digital hadron calorimeter with muons", JINST 3 P05001, 2008.
- **6.** Q.Zhang et al., "Environmental dependence of the performance of resistive plate chambers", JINST 5 P02007, 2010.
- **7.** B.Bilki, CAN entitled "DHCAL response to positrons and pions".