Analysis of Cosmic Data

Robert Keyes

September 7, 2009

University of Victoria - ATLAS

Abstract

Data recorded in June 2009 by the ATLAS detector is analyzed with the goal to understand the energy deposits produced by cosmic muons with both calorimeter and muon spectrometer measurements. A cleaning procedure is developed to identify spurious events and energy clusters associated to problematic channels. From a clean sample the following details of the topology cosmic events in the L1Calo stream are examined: location in the detector, energy spectrum of the clusters, trigger rates and charge of the incoming muons.

Contents

1 Introduction

The ATLAS detector is made to identify events originating from the interaction point on the beam line of the LHC. It is important to know how the detector sees cosmic interactions to be sure that they can be understood and discriminated against while under normal running conditions. This is to be done by understanding how well the tracking and clustering algorithms can identify interactions not originating from the collision point, i.e. cosmic events. Combined runs were taken in June 2009 including all sub-systems of the ATLAS detector and these are examined here.

2 Strategy

Initially, work was done by creating ROOT ntuples from the large ESD (Event Summary Data) files. Once created, the ntuples were simple to analyze but often lacked useful information such as access to the detector geometry and useful tools available within Athena, for instance the track extrapolation and the identification of known bad channels. This strategy was then changed to creating thinned ESD files, meaning that the content and size of the ESD files are reduced by eliminating data which is not relevant to this particular analysis, but could still be analyzed using the Athena framework. Once created it is then possible to run on these thinned ESD files locally in a matter of minutes. Each ESD is made up of a collection of events. Each event is a snapshot of the detector triggered when sufficient energy is deposited in the detector, greater than 3 GeV [1]. In the context of cosmics data taking, two reconstructed objets are investigated in this analysis: simple calorimeter clusters (CaloTopoClusters) and tracks reconstructed by the MS alone (ConvertedMBoyTracks). The CaloTopoClusters are collections of calorimeter cells built according to a 4:2:0 algorithm: around a seed cell with S/N (signal-to-noise ratio) $> 4\sigma$, adjacent cells with S/N$> 2\sigma$ are added as well as all positive energy cells around them. The MuonBoy tracks (ConvertedMBoyTracks) identify particle paths by a series of track parameters giving the position and momentum at various space points in the detector.

The results displayed here, unless otherwise indicated, are from the run 121569 chosen for its low yield of events with corrupted Front-End Boards (FEBs) data and the good quality flags of the calorimeter and muon system. In 4h36m, 92329 events were recorded by the L1Calo trigger. The solenoidal and toroidal fields were both at their nominal values.

3 Analysis

The following outlines the steps taken in analyzing these ESD files.

[1] The L1Calo stream contains, among others, the EM3 threshold [1].

3.1 Bad channels treatment

Calorimeter channels with known problems are identified in Athena using the CaloBadChannelTool (which accesses the TileBadChannel and LArBadChannel databases [2]). When a bad channel is in a cluster, its energy is still considered when calculating the cluster's energy. To determine if bad channels are generating fake clusters, the fraction of the energy contributed to the clusters by them is shown (see figure 1). In some clusters, the bad energy cells account for the energy of the entire cluster. There is a noticeable elbow at 0.3, value adopted to discriminate fake clusters from those candidate products of cosmic muons. With a typical energy deposit in the calorimeters of the order of 4-5 GeV, 70% of the energy being good is still higher than the EM3 threshold (3 GeV).

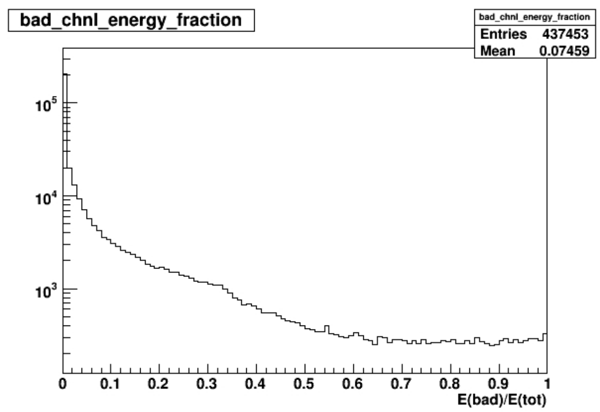

Figure 1: The ratio of the bad channels' energy to the total cluster energy displays some clusters produced by the presence of bad channels.

In order to investigate whether events are triggered in bursts or at a constant rate, the occurrence of events over time is plotted (see figure 2). This proved to be relatively constant.

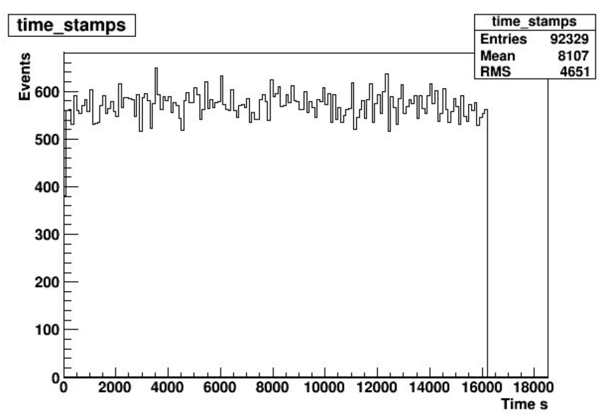

Figure 2: Events are triggered at a constant rate throughout the run.

The cluster energies are then plotted for the barrel and endcap portions of the detector (see figure 3). This is done by considering the barrel to be between $|\eta| < 1.4$ and the endcap otherwise. The cosmic spectrum creates a structure visible around 5 GeV which is not visible in the endcaps. It is not seen in the endcaps because the projectivity of the cells in the barrel align far better with the cosmic particles, while in the endcaps, the clusters are less regular in shape and size.

Throughout the cosmic runs random triggers are taken to help understand the background processes in the detector. This information is used to understand the energy spectra of the background noise (see figure 4).

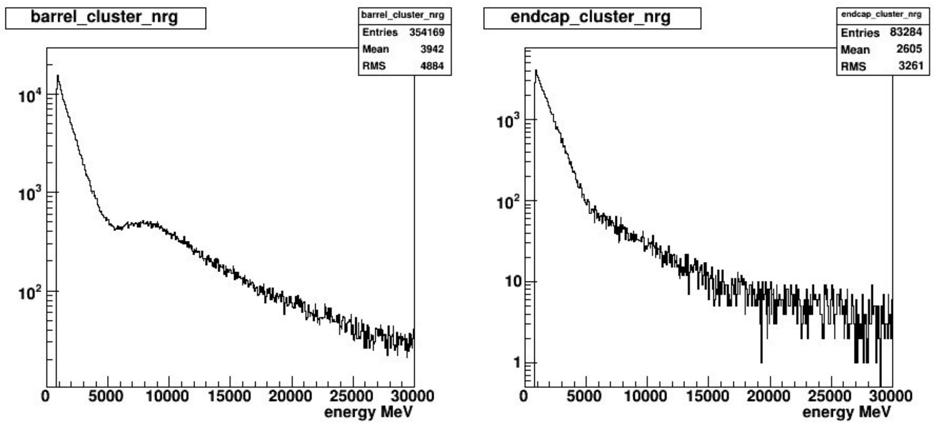

Figure 3: The spectrum of cluster energies in the barrel (left) shows muon energy deposits from around 5 GeV. This structure is not visible for the endcap (right).

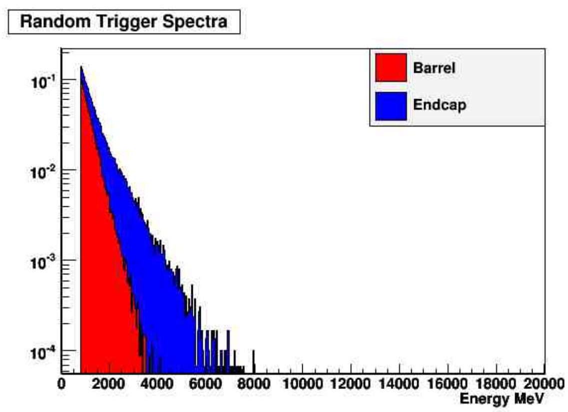

Figure 4: The spectra of cluster energies in the barrel and endcap from randomly triggered events (normalized in this plot) gives us a representation of the noise present during the run.

In the barrel the random trigger spectra ended at 4 GeV whereas in the endcap it ended at 8 GeV due to the larger noise in this region.

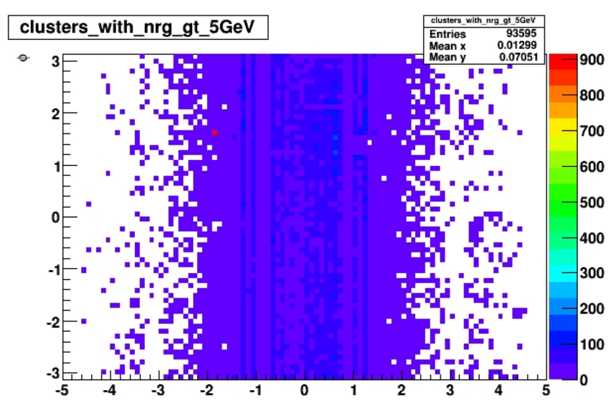

Figure 5: This $\eta - \phi$ map of the clusters with energy greater than 5 GeV indicates that some cells create frequent clusters.

To investigate the distribution of cluster energies in the detector, the cluster distribution over different energy ranges is examined. It is found regardless of the energy range there are always particular channels which would throw the range of the plots off because they are triggered on the order of 9 times more than the rest of the channels in the detector (see figure 5). It is evident that there are particular channels which are triggering extremely frequently. The number of times every cell is triggered during the run is then plotted (see figure 6).

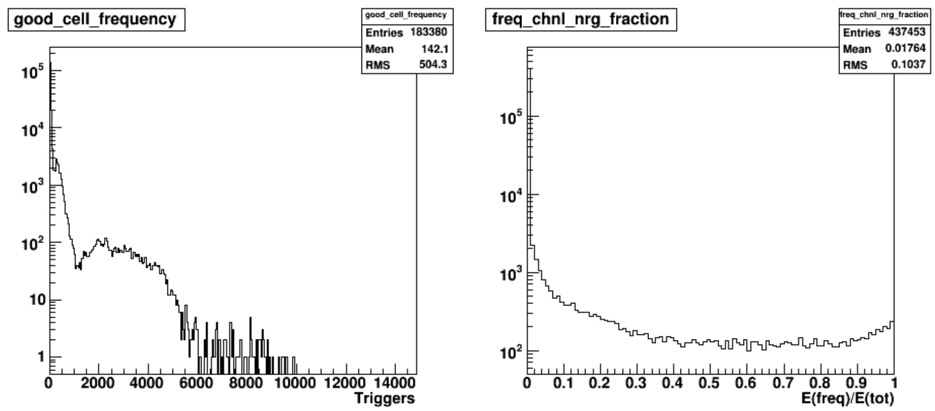

Figure 6: a)Plotting the number of times each channel in the detector is triggered led to choosing to cut channels which trigger more than 6000 times b)Plotting the fraction of each cluster's energy due to channels which triggered more than 6000 times in the run illustrated that some of these channels are indeed responsible for entire clusters.

By examining this plot it is decided to make a cut at 6000. These cells are then added to the bad channel list, meaning that their energy is taken into account when calculating the bad channel energy fraction. The frequent channel energy fraction (the fraction of the energy in a cluster due to channels tagged as being frequent, i.e. occurring in over 6000 events) is also plotted. The fact that hundreds of clusters' energy is entirely made up of frequent channels this indicated that many of these cells are indeed responsible for entire clusters.

3.2 Clusters in L1Calo events

From here on clusters will be referred to as *clean* if the fraction of their energy due to frequent channels and bad channels does not exceed 30%. Cosmic events are in general triggered by one particular cosmic interaction, though more than one could occur at once, one for instance in each hemisphere (see figure 7.

The clustering algorithm finds clusters of low energy caused by fluctuations in the detector, noise or bad channels. To avoid these and examine the energy spectrum, the maximum cluster energy for each event is plotted (see figures 8). The peak around 5-6 GeV could be totally hidden if bad channels are not treated correctly as observed with run 121275 (figure 9). The distribution of clean clusters with energy greater than 5 GeV is plotted in figure 7, along with the number of clean clusters per event greater than 5 GeV.

3.3 Tracks in L1Calo events

The detector is designed to identify events originating from the interaction point. The efficiency of the cluster and tracking algorithms in identifying cosmic particles is also evaluated by determining the distance between a cluster and the nearest track. This is then plotted vs ϕ, η and ρ (distance from the z axis in the xy plane) to see if the tracking system is more efficient in particular parts of the detector (see figure 10).

The MS tracks are extrapolated to a cylinder of 4.25 m corresponding to the entrance point of the calorimeters. Inside the calorimeter volume, no magnetic field is expected and tracks could be treated as straight lines. The entrance and exit points of the tracks are then plotted on the surface of a 2 m

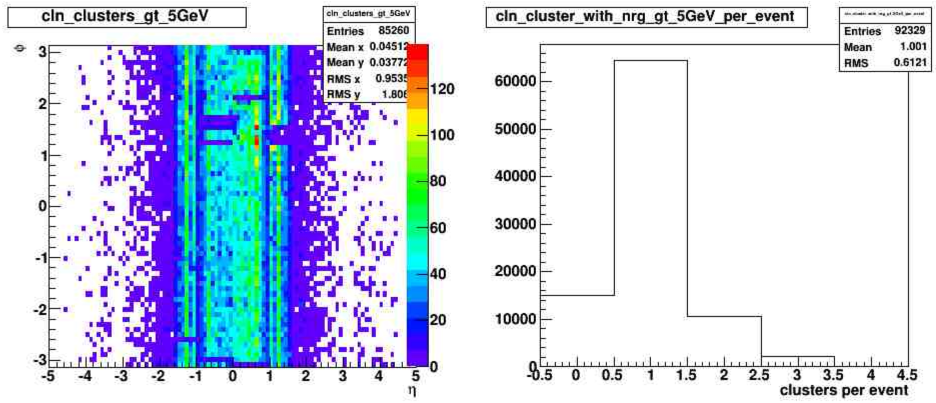

Figure 7: a)The map of maximum clean clusters with energy greater than 5 GeV exhibits a flatter distribution than in figure 5 b)The number of clean clusters with energy greater than 5 GeV per event distributed closely around 1.

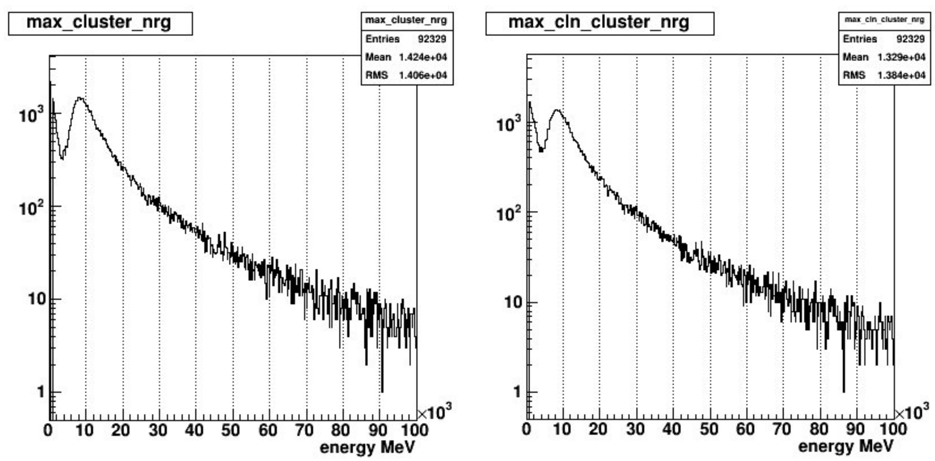

Figure 8: Plotting the spectrum of the maximum cluster energy in each event, and then the same with clean clusters, there is little difference for run 121569.

cylindrical surface to view the flux of cosmic particles at the level of the EM calorimeter (see figure 11). The distribution of tracks in z is plotted in a 1D histogram to see if the position of the two shafts is visible. A slightly greater muon flux in the positive z direction is observed as expected from the presence above of the larger shaft.

In order to assess the efficiency of the tracking and clustering algorithms, the position of clusters without tracks associated with them are plotted in η and ϕ (see figure 12). The first plot illustrates that the majority of the tracks occur in the barrel, and that beyond 200 mm there is a uniform distribution of trackless clusters. The second plot, which is in ϕ, shows there is a symmetry due to the entry and exit of the tracks and otherwise little phi dependence. The final plot in rho shows that there is a high concentration of close clusters around 1700, 2400, 3000 and 3500 mm, which is where the EM calorimeters are located in the barrel.

Finally, assuming that each event is triggered by one cosmic muon, if there is found a track within 40 cm of a clean cluster in the event, the charge of this track is recorded with its time stamp. This is then used to determine the rate at which positive and negative muons are seen by the detector (see figure 13). A back of the envelope calculation reveals that the observed ratio of positive to negative muons is 1.21. This result agrees well with the ratio in the PDG[3], being around 1.23 . Combined average rate of muons being detected is then 1.8 ± 0.2Hz.

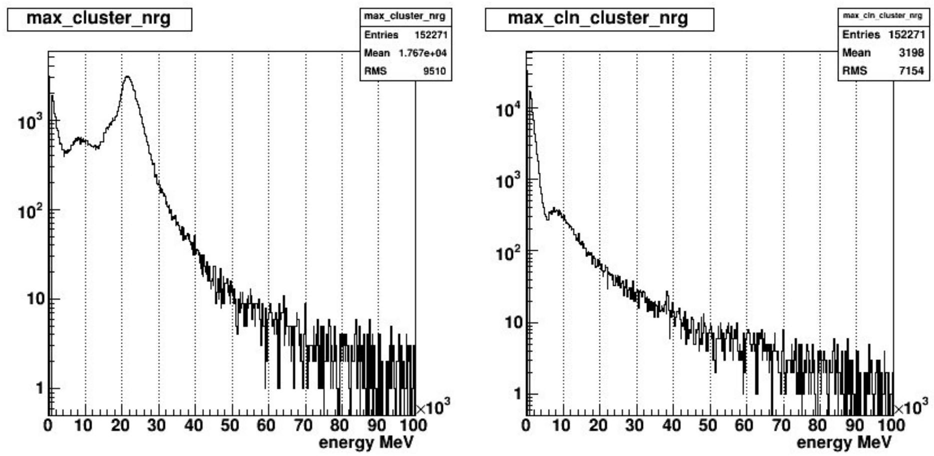

Figure 9: Plotting the spectrum of the maximum cluster energy in each event, and then the same with clean clusters, many spurious processes due to noise are filtered out leaving behind what appears to be the proper spectrum for run 121275.

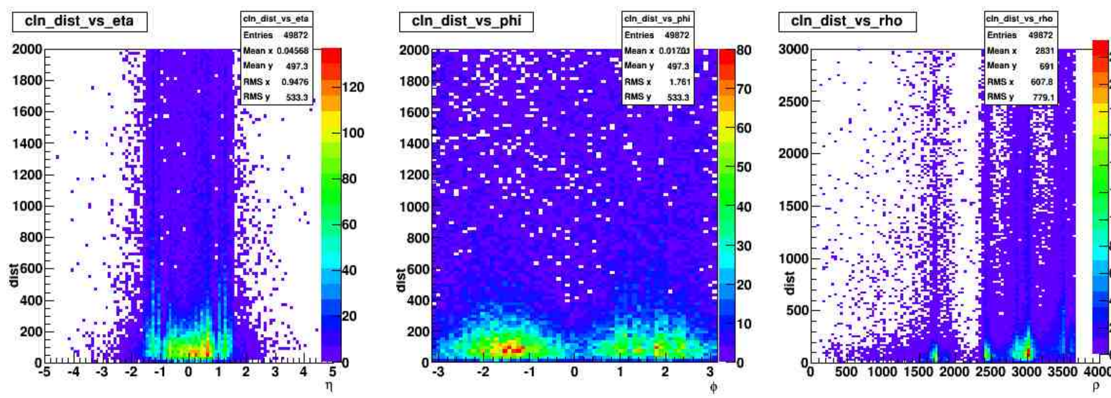

Figure 10: The distance from a cluster to the nearest track, in η, ϕ and ρ (distance from the z axis in the xy plane).

3.4 Further steps

The next steps which are to be taken would have included investigating the correlation of the cluster moments with the tracks. In theory a cluster ought to be aligned with the track, this might be an efficient way of discriminating relevant clusters. In addition, an analysis of comparing track timing to nearest cluster timing could be done.

4 Conclusion

Tools have been created to investigate cosmic data which should be helpful in this pursuit. It is observed that there are many clusters which do not match up with a track and that the endcaps are not as proficient at identifying cosmic events as the barrel. An additional cut is made based on the number of times a channel triggers in a run, and this appears to be an efficient discriminatory measure. The ratio of positive to negative muons is calculated and found to agree with known results.

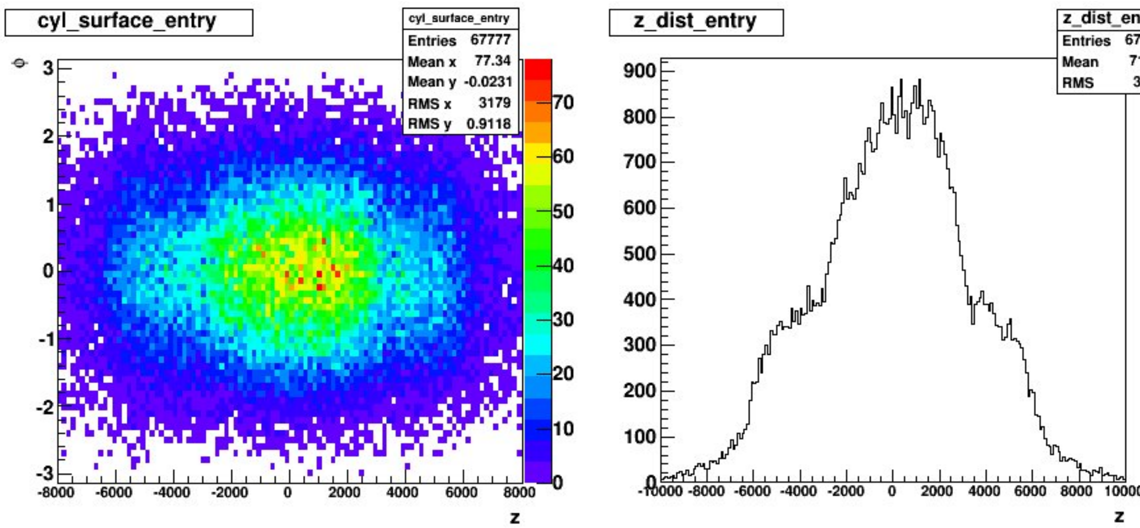

Figure 11: Track entry points on 2 m radius cylinder. Slightly more tracks in positive z explained by the large shaft above the detector.

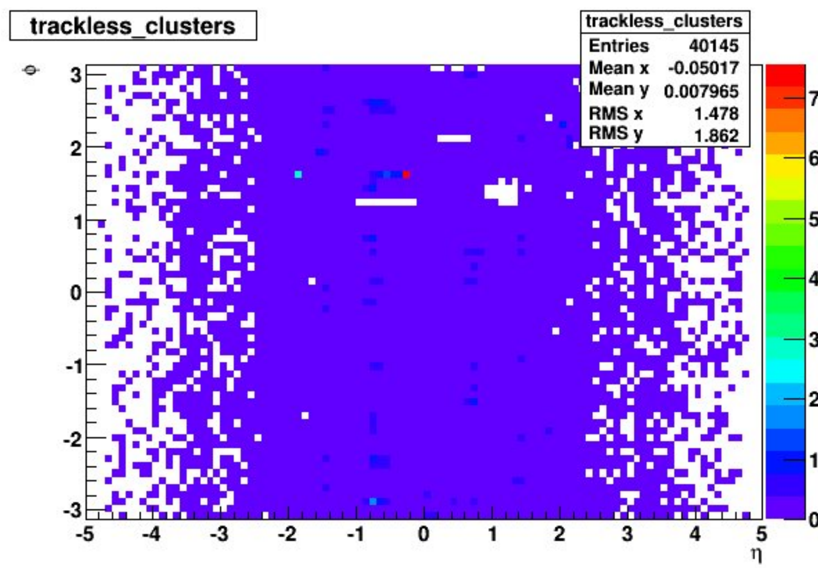

Figure 12: Location of trackless clusters in detector. Observed to be quite evenly distributed with one particularly frequent cell.

References

[1] https://twiki.cern.ch/twiki/bin/view/Atlas/StartupTriggerMenus.

[2] https://twiki.cern.ch/twiki/bin/view/Atlas/LArP3CommissioningUnhappyChannels.

[3] W.-M. et al Yao. Review of Particle Physics. *Journal of Physics G*, 33:1+, 2006.

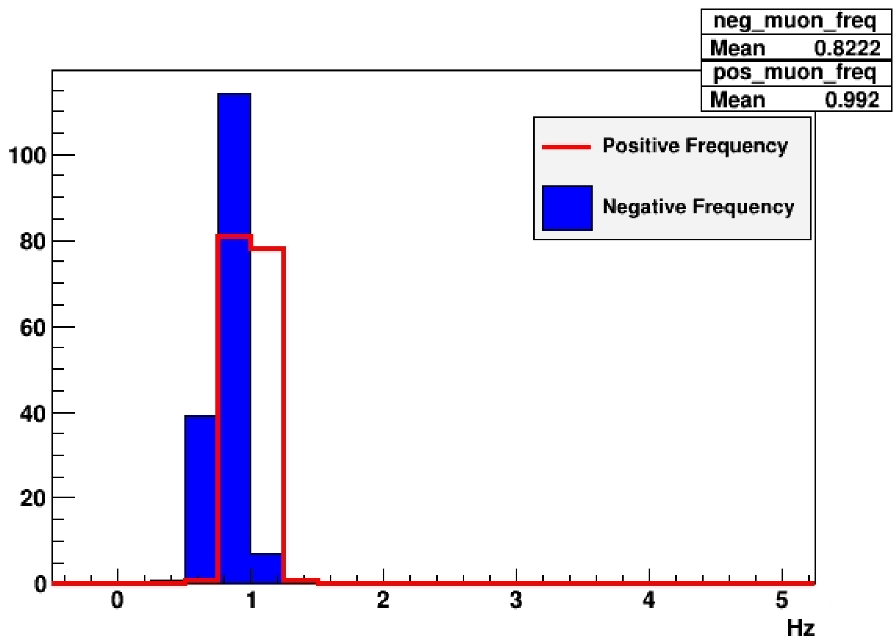

Figure 13: The frequency of positive muons is distributed around 1.00 Hz while for negative muons, it is distributed around 0.82 Hz .

ECal Team Comparison Tool

Robert Keyes

September 4, 2009

UVic - ATLAS

Abstract

A python script was written to facilitate the monitoring of the stability of the LAr calorimiters by the LAr ECAL team. The script was translated from a Perl script and then modified to give a summarized output of the comparison of a list of calibration runs to a set of reference runs. The final output of this script would summarize the types of runs which were compared (ie the gain and run type), list missing front end boards in each run, display global shifts with respect to the reference run in ramp, delay, pedestal and autocorrelation coefficients, and list noisy channels in order of the most frequently occurring channels. This code is now available on the ECAL team CVS group area.

1 Introduction

It is the job of the ECAL team to analyze calibration run data in order to monitor the performance and stability of the ATLAS liquid argon calorimeters. The different partitions of the calorimeter are delegated to the different members of the team. Calibration runs allow us to convert from ADC (measured) to MeV (deposited energy). There are three types of calibration runs: ramp, pedestals and delay. Ramp runs pulse the channels with a scan input function, measuring the gain and slope and thus allowing for the conversion from ADC to MeV. Pedestal runs have no input signal, thus the offset, noise and autocorrelation of the noise are measured. Delay runs pulse the channels with a known signal and permits the calculation of the optimal filtering coefficients. These coefficients, which exists for all 200 000 channels, must be monitored and any instability must be understood to the best of our ability. To understand the finer points of this analysis refer to chapter 4 of Marco Delmastro's thesis.[1]

2 Motivation

During the spring shutdown in 2009 repairs were being done to the calorimeter, which meant the ECAL team was working through a period of unstable running. This also meant that much of the work being done by the team was to validate new hardware. Calibration analysis in normal running is always done with respect to some pre-determined reference run. During this time the task was just as much finding a good reference run as it was observing and investigating detector fluctuations. The goal of my work was to write a code which summarizes and prioritizes the problems observed in a set of calibration runs with respect to a chosen set of references.

3 The Code

The code, adapted from a Perl script and written in Python, employed previously existing c++ code to compare a list of runs to a list of reference runs. The code summarized the results of this comparison in a simple text file, and output a root file for each comparison with the results for each channel, as well as an average over each front end board. An example of this summer is as follows.

```
Run Details                                                 Global Shift    Missing FEBs- FT_SLT
EndCap_HEC---_HIGH--_Pedestal_COMP_00116185_REF_00116539FEB 0.00587252      None
EndCap_HEC---_HIGH--_Ramps---_COMP_00116188_REF_00116541FEB 0.0147593       None
EndCap_HEC---_HIGH--_AutoCorr_COMP_00116185_REF_00116539FEB -6.59079e-05    None

Run Breakdown            Detctr  Total  LOW  MEDIUM  HIGH  Ramps  Ped  ACorr  Delay
                         HEC---  3      0    0       3     1      1    1      0
                         FCAL--  0      0    0       0     0      0    0      0
                         EMECA-  0      0    0       0     0      0    0      0
                         EMECC-  0      0    0       0     0      0    0      0
                         EMBPSA  0      0    0       0     0      0    0      0
                         EMBPSC  0      0    0       0     0      0    0      0
                         EMBA--  0      0    0       0     0      0    0      0
                         EMBC--  0      0    0       0     0      0    0      0

BE_AC_FT_SLT_CHL_CAL#    Detctr  Total  LOW  MEDIUM  HIGH  Ramps  Ped  ACorr  Delay
1_0_16_09_035_000        HEC---  2      0    0       2     1      1    0      0
1_0_16_07_051_002        HEC---  2      0    0       2     1      1    0      0
1_0_16_07_048_006        HEC---  2      0    0       2     1      1    0      0
1_0_16_07_036_012        HEC---  2      0    0       2     1      1    0      0
1_0_16_07_035_002        HEC---  2      0    0       2     1      1    0      0
1_0_10_08_102_101        HEC---  1      0    0       1     0      0    1      0
1_1_22_10_119_115        HEC---  1      0    0       1     1      0    0      0
1_1_22_10_117_125        HEC---  1      0    0       1     1      0    0      0
1_1_22_10_116_125        HEC---  1      0    0       1     1      0    0      0
1_1_22_10_115_099        HEC---  1      0    0       1     1      0    0      0
1_1_22_10_114_099        HEC---  1      0    0       1     1      0    0      0
1_1_22_10_112_109        HEC---  1      0    0       1     1      0    0      0
1_1_22_10_108_119        HEC---  1      0    0       1     1      0    0      0
1_1_22_10_107_096        HEC---  1      0    0       1     1      0    0      0
1_1_22_10_105_103        HEC---  1      0    0       1     1      0    0      0
1_1_22_10_104_103        HEC---  1      0    0       1     1      0    0      0
```

Figure 1: .

The first section of the output lists for each run comparison the global shift in parameter values and a list of any missing FEBs . The next section summarizes the types of runs which were compared. The third and final section of the output lists all the channels, which are not on the bad channel list, whose deviation was over threshold while displaying during what types of run this occurred. This list is organized according the number of times the channel was over threshold from most to least. The thresholds are predetermined acceptable levels of fluctuation for each component in the calorimeter.

The code is available on CVS at: http://atlas-sw.cern.ch/cgi-bin/viewcvs-atlas.cgi/groups/LArElecCalib/python/

4 Scope

This code is useful when one runs over a handful of particular runs to investigate some strange behavior. This is not in the scope of the ECAL team under normal running conditions though will remain to be useful during and after the shutdown period. Under constant running conditions the analysis of calibration runs is automated as possible to efficiently deal with the incoming volume of calibration data. In order to do so analysis algorithms have to be more rigorously developed.

5 Stability

The ECAL team is responsible for understanding the long term stability of the detector. This poses many challenges and questions. For instance, if a channel is not a bad channel but exhibits some strange behavior is this recorded? This is currently not the case and to implement another channel status requires adding another level of flagging to the database. In order to run over the database to create a history of a particular channel's behavior would require the development of a new tool which does not currently exists.

6 Conclusion

This code is useful for a specific investigation into a list of runs in situations where the ECAL team has to investigate abnormalities in depth, ie during hardware repairs or if the database reference was found to be bad. The logic in the code or something similar could be implemented into the automated calibration run analysis as the summary was seen as a useful tool.

References

[1] http://cdsweb.cern.ch/record/953119/files/thesis 2003-033.pdf.