Quick look at DATA - MC comparisons

David Paneque *(dpaneque@slac.stanford.edu)*

The aim is to have an "quick evaluation" of the agreement/differences between data runs and MC runs

Quick comparisons of some selected data-mc runs.

Table relating MC with Data runs:

Energy(GeV)	Position	Angle	MC Run	Data Run
282	201,40,-47	0	166	1922
282	idem	30	167	1944
200	idem	0	164	1911
200	idem	30	168	1902
100	idem	0	162	1981
100	idem	30	169	1999
50	idem	0	172	2039
50	idem	30	173	2054
20	idem	0	176	2082
20	idem	30	177	2096

Quick look at DATA - MC comparisons

Comparison of some selected parameters for several Energies (282, 200, 100, 50, 20) and 2 angles (0, 30)

Tracker
TkrTotalHits
TkrNumTracks
Tkr1X0, Tkr1Y0 (*beam dimensions*)

Calorimeter
CalZEcntr
CalTransRms
CalLyr7Ratio
CalEnergyRaw
CalCfpEnergy (*other E reconstructions currently not available for tower 2*)

Important remark

Beam **incidence angle**, **impact position** and **dimensions** in the Montecarlo are NOT exactly the same to those ones of the data.

Therefore, **we do expect differences**. Specially, we'll see big differences in those parameters used to estimate the "non-measured" energy in the calorimeter.

In order to make these comparisons more meaningful, it is required to extract these parameters from the experimental runs, and then produce simulations with exactly these parameters.

Besides, it would be convenient to increase the statistics of the "detailed" MC runs to (at least) match the data runs.

This is an ongoing work

Important remark

The only cuts applied to the data are :

1 - CalEnergyRaw > 10 MeV (No-empty events)

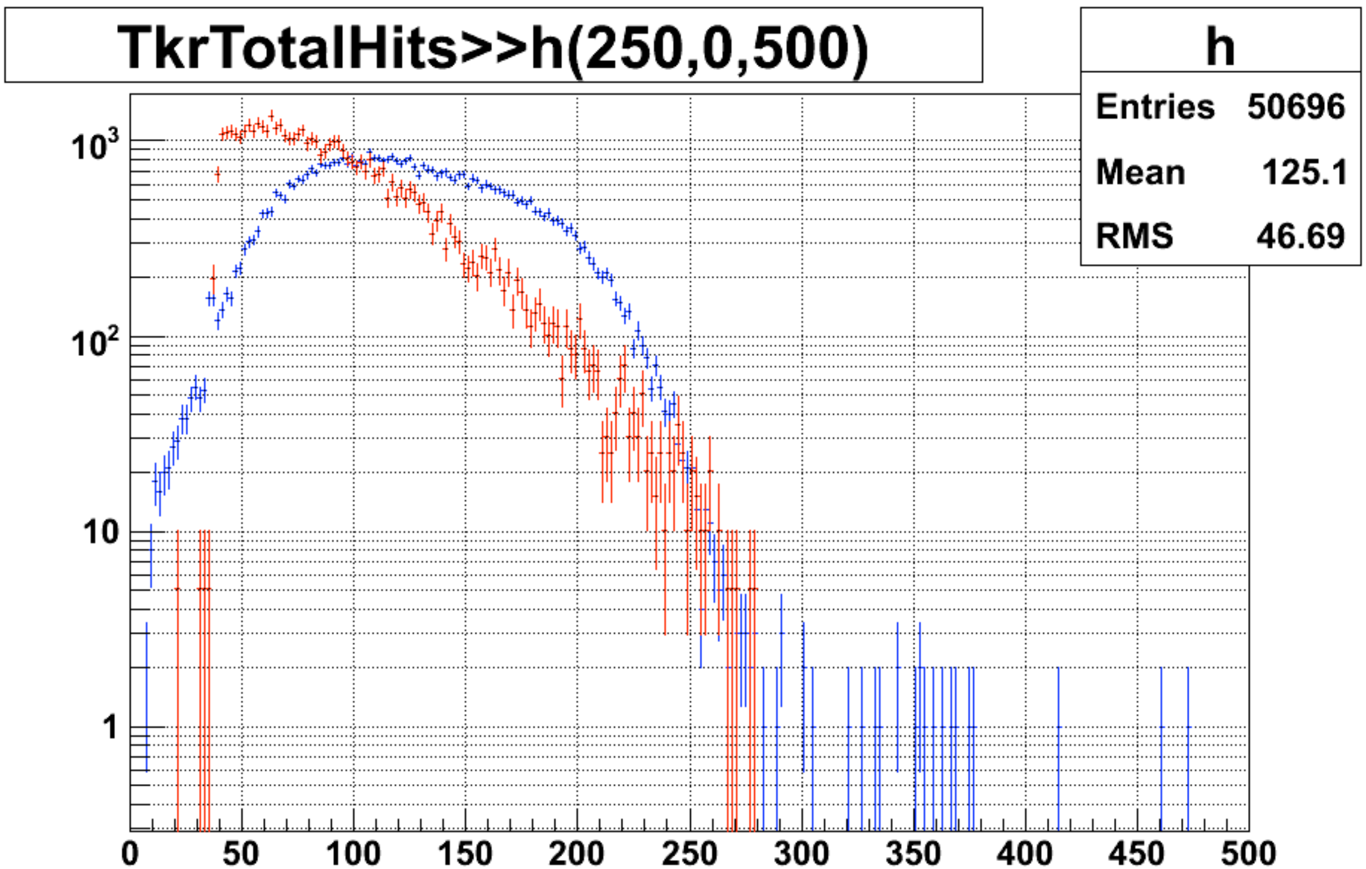
2 - TkrNumTracks > 0.5 (events with at least 1 track)

These are very simple cuts which are expected to be fulfilled by all the electrons (>20 GeV) entering in the calibration unit.

More sophisticated cuts (e.j. removing events crossing cracks, removing MIPs…) which might improve the agreement data-mc are NOT applied. These additional cuts must be applied with care, since they might also bias the comparison if not carefully done

E = 282 GeV , 0 deg: Data 1922, MC 166

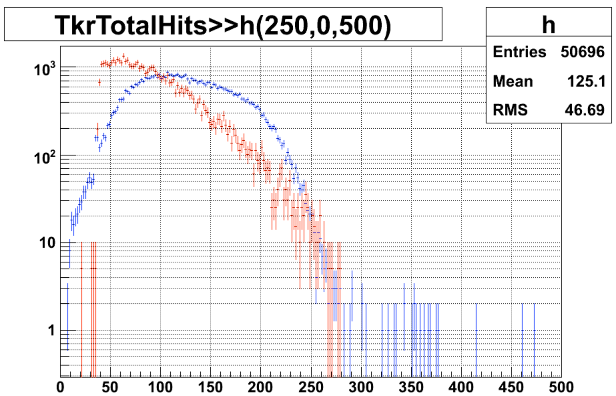
MC in red; **Data in blue**



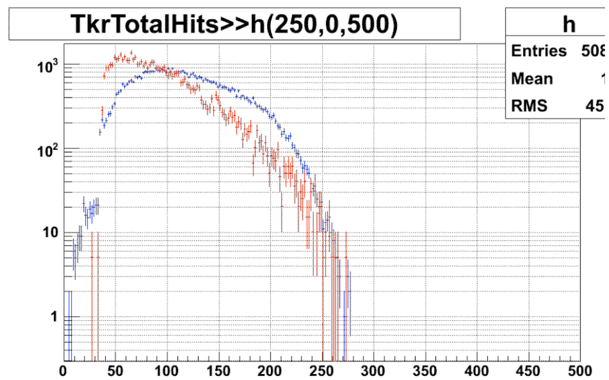
Incidence angle, 0 deg

MC in red; **Data in blue**

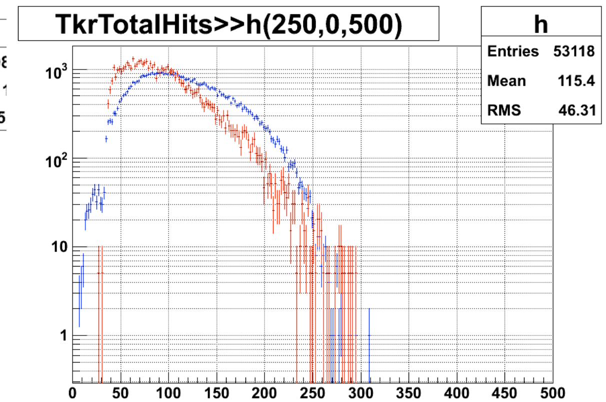
E = 282 GeV



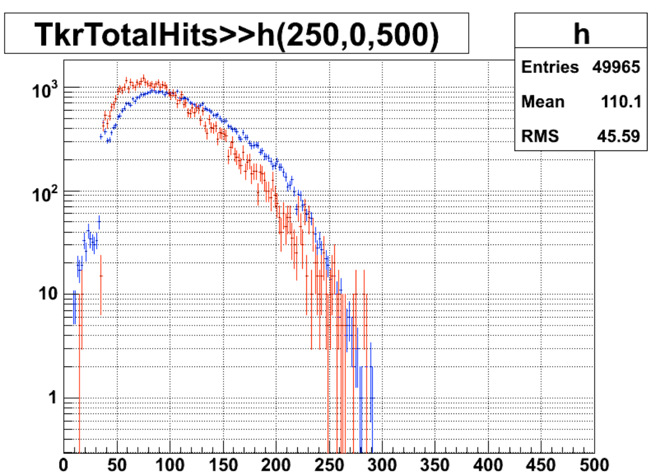
E = 200 GeV



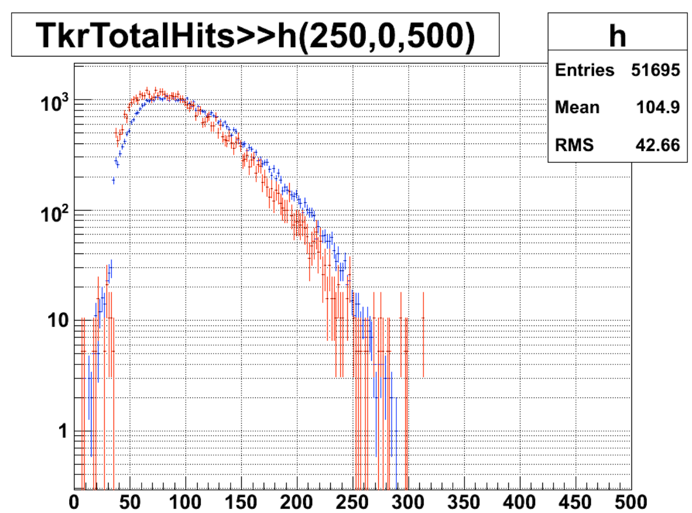
E = 100 GeV



E = 50 GeV



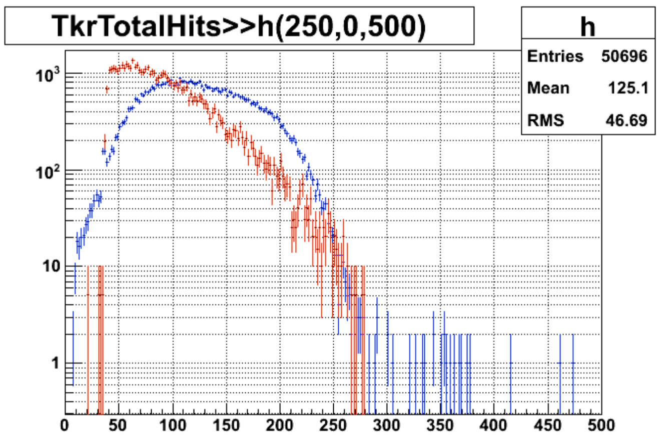
E = 20 GeV



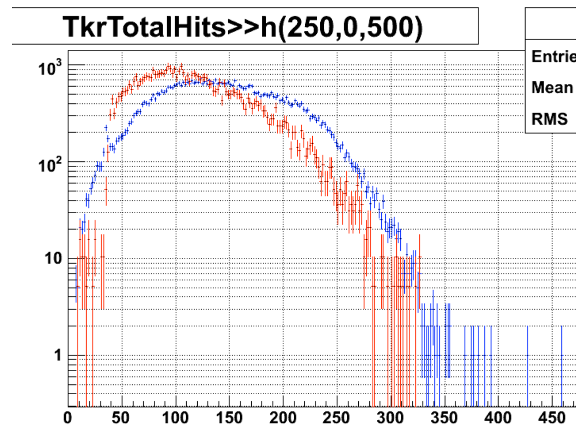
Incidence angle, 30 deg

MC in red; **Data in blue**

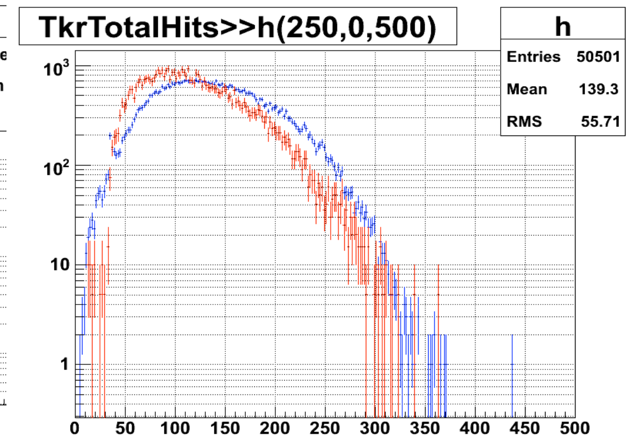
E = 282 GeV



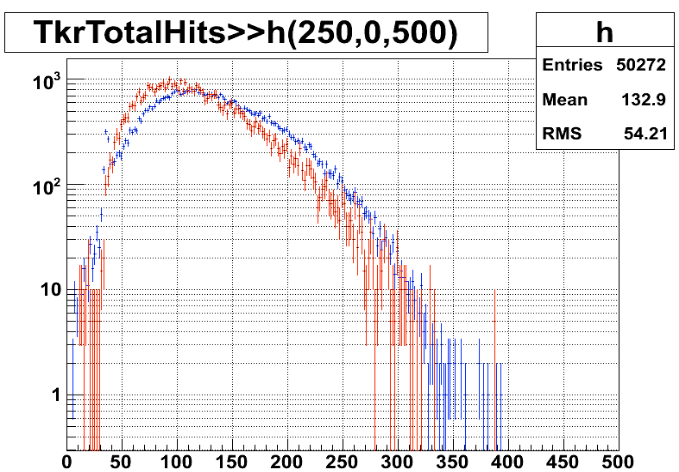
E = 200 GeV



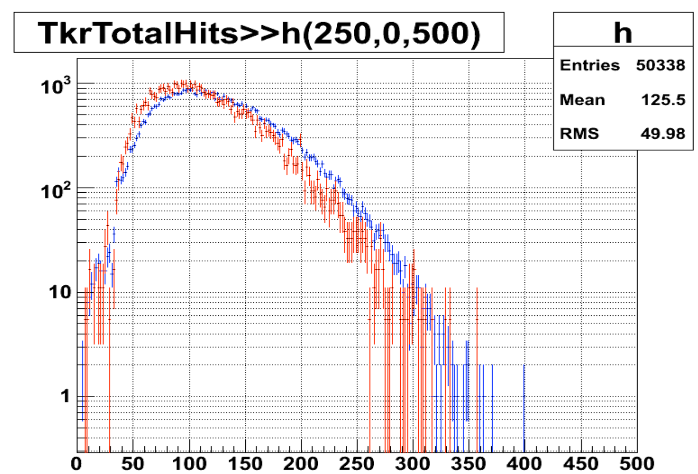
E = 100 GeV



E = 50 GeV



E = 20 GeV



Quick evaluation of agreement in TkrTotalHits

In general the events in the data do have more hits than the events in the MC. Perhaps this is due to “more backsplash“ in the data than in the MC

Above 250 hits, the distribution falls down quicker in the data than in the MC (probably due to event read-out limitations in data)

When the Energy decreases, the distributions of TkrTotalHits become more equal.

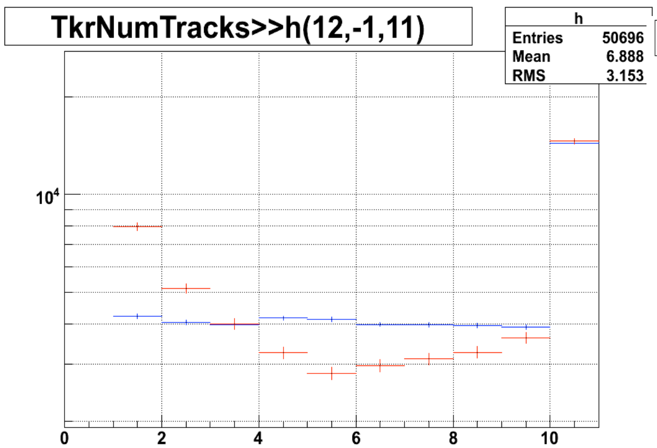
That is consistent with the hipothesis of non-properly simulated backsplash, since *backsplash goes with sqrt(Energy);* and thus, the lower the energy, the lower the effect of backsplash.

No big differences between 0 and 30 degrees Incident angle

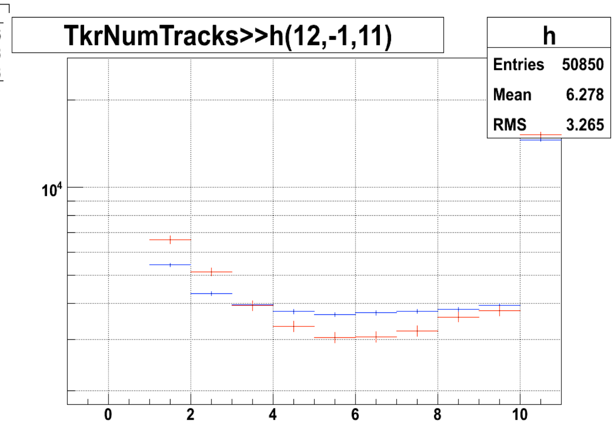
Incidence angle, 0 deg

MC in red; Data in blue

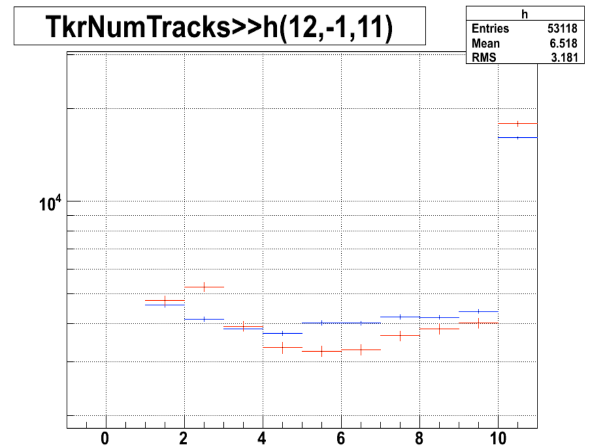
E = 282 GeV



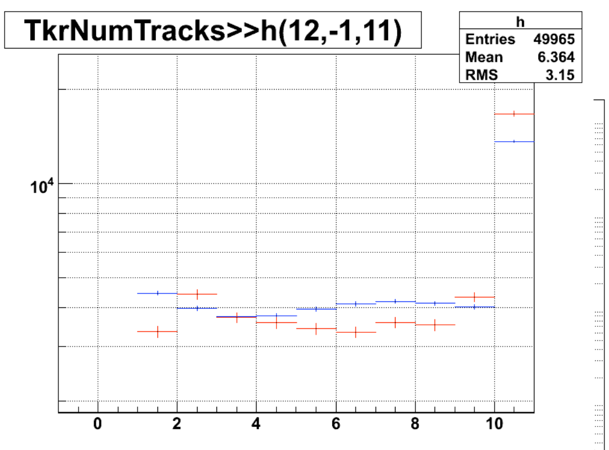
E = 200 GeV



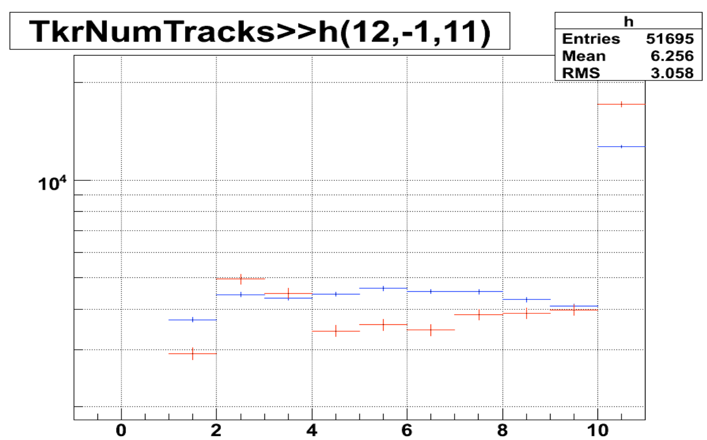
E = 100 GeV



E = 50 GeV



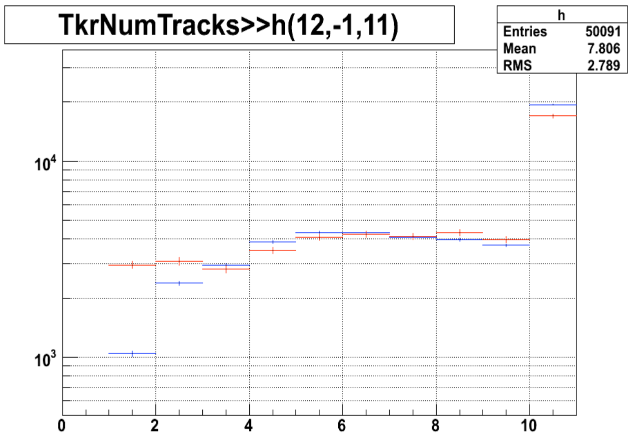
E = 20 GeV



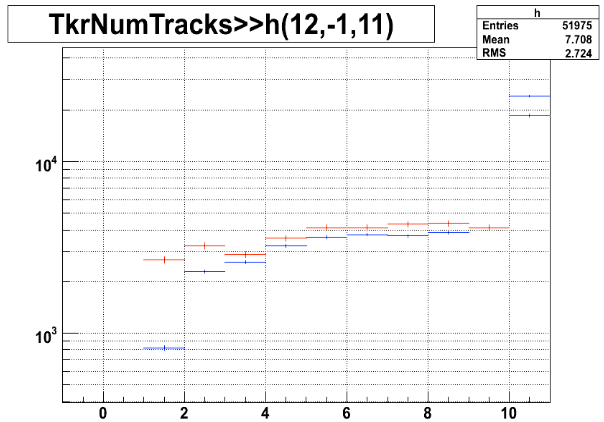
Incidence angle, 30 deg

MC in red; Data in blue

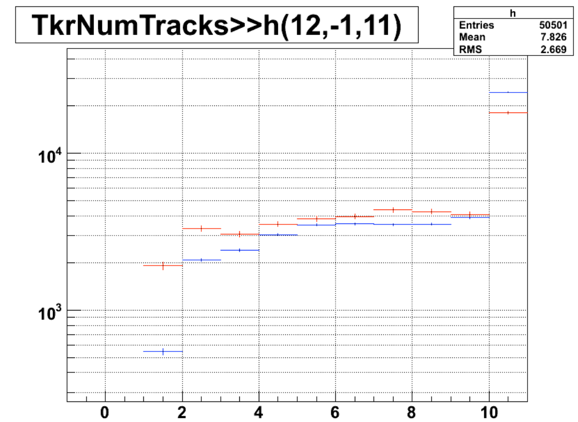
E = 282 GeV



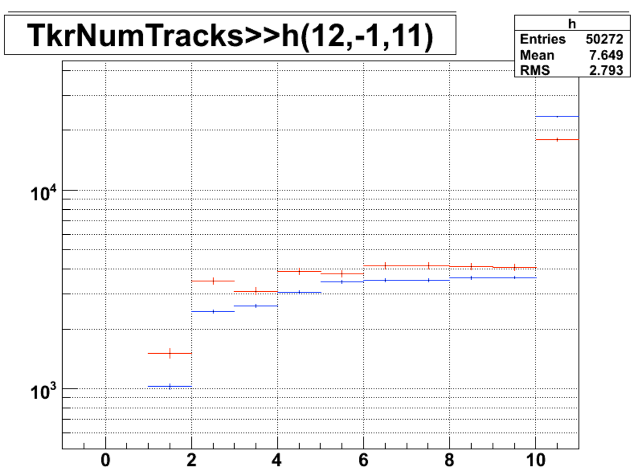
E = 200 GeV



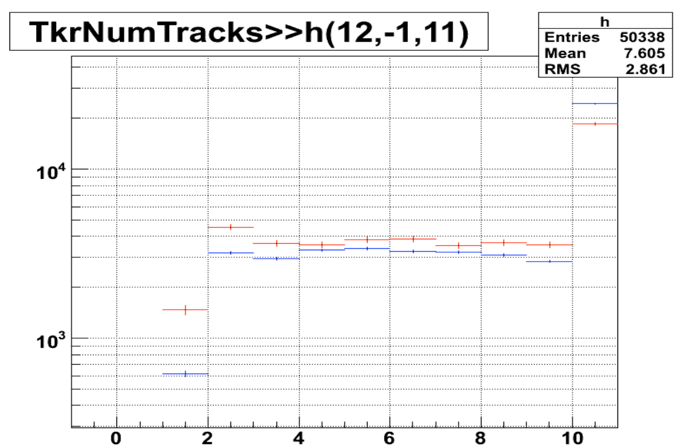
E = 100 GeV



E = 50 GeV



E = 20 GeV



Quick evaluation of agreement in TkrNumTracks

Largest differences at 1, 2 and >=10 tracks. The number of events with >=10 tracks dominates the normalization.

The distribution of TkrNumTracks in the MC is more “curved” than in the data at 0 degrees, and “flatter” than in the data at 30 degrees.

(normalized) number of events with >= 10 tracks is larger in the MC (than in the data) for all energies at 0 degrees, yet it is smaller at 30 degrees for all energies

How can we interpret this ??

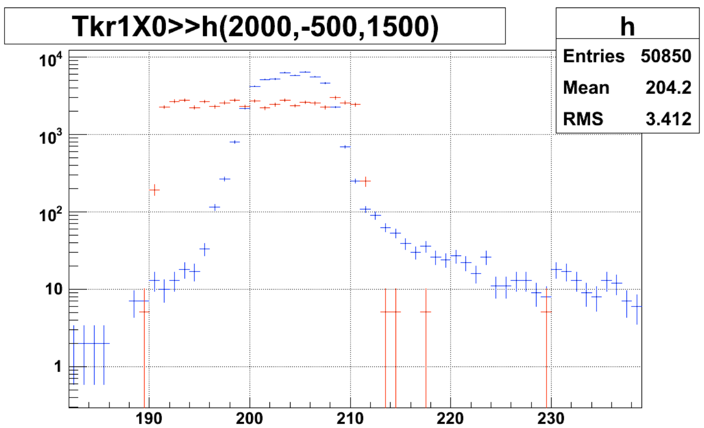
Incidence angle, 00 deg

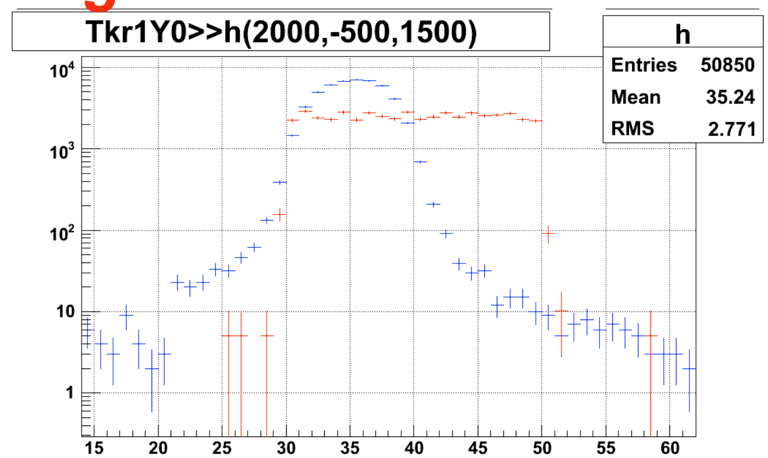
MC in red; Data in blue

E = 200 GeV

MC RUN 164 (2006/09/**07**)

MC beam larger than data





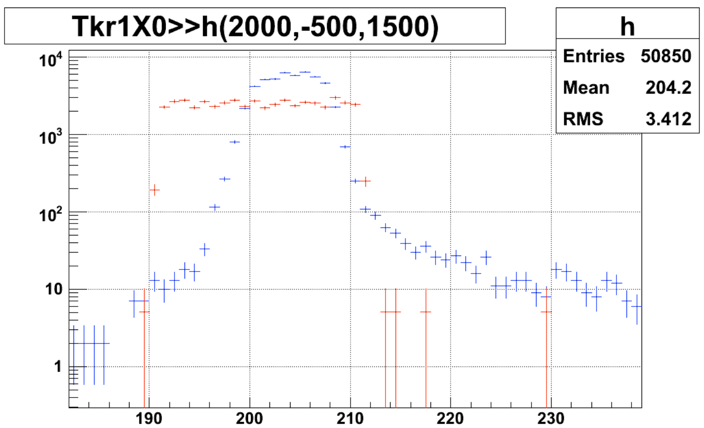
Incidence angle, 00 deg

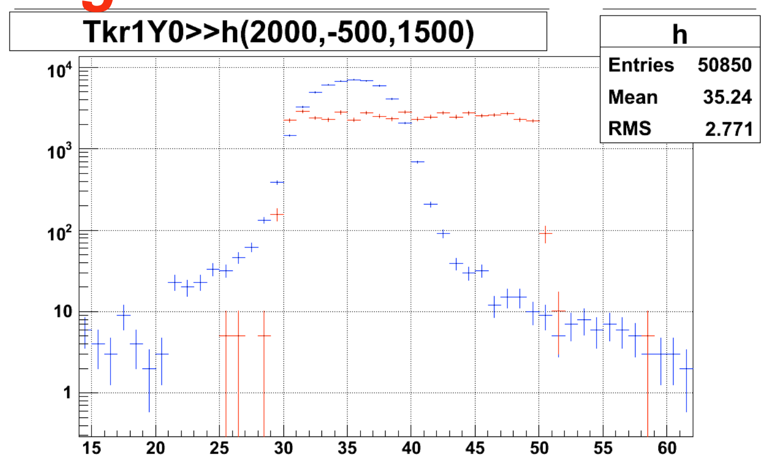
MC in red; Data in blue

E = 200 GeV

MC RUN 164 (2006/09/07)

MC beam larger than data





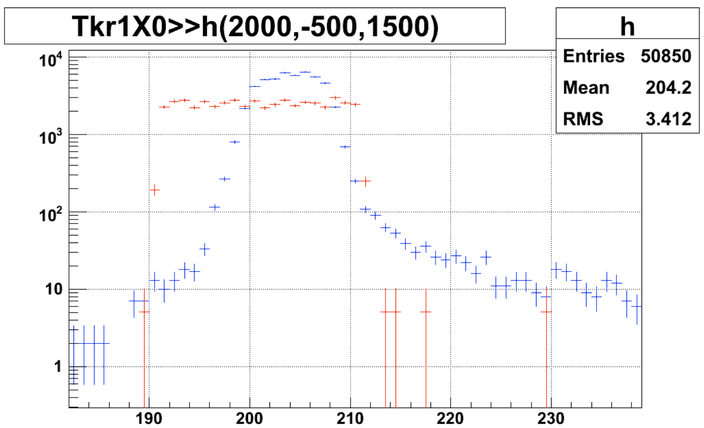
I reported (2006/09/08) Francesco Longo that the MC beam was wider than the beam from the experimental data….

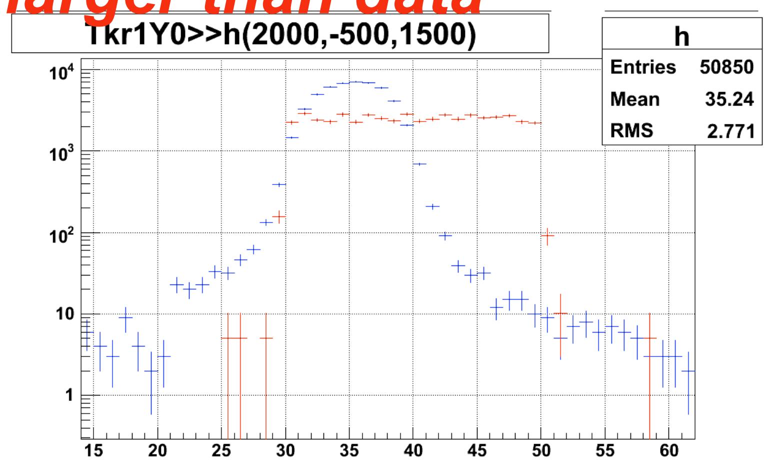
Incidence angle, 00 deg **MC in red**; **Data in blue**

E = 200 GeV

MC RUN 164 (2006/09/**07**)

MC beam larger than data

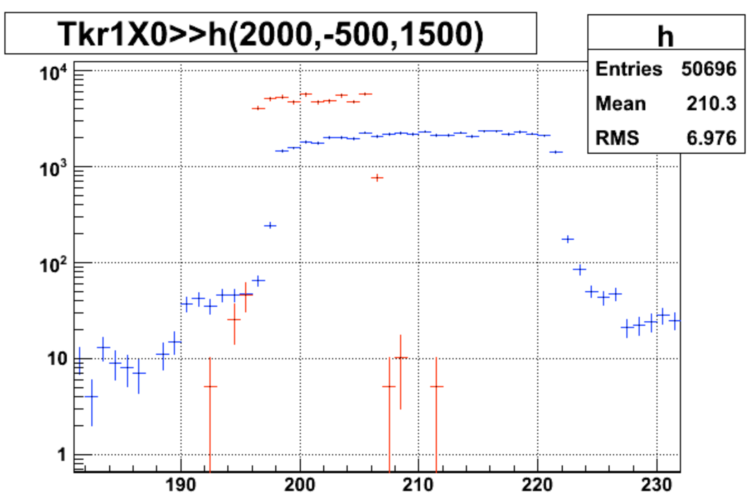


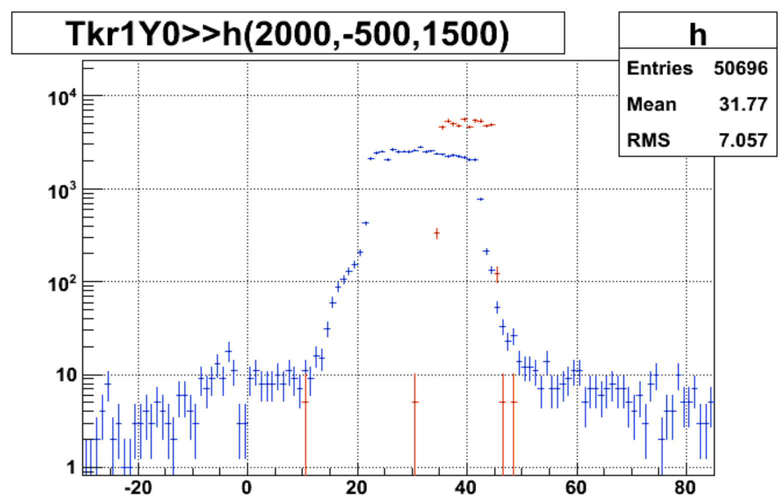


E = 282 GeV

MC RUN 166 (2006/09/**12**)

MC beam smaller than data



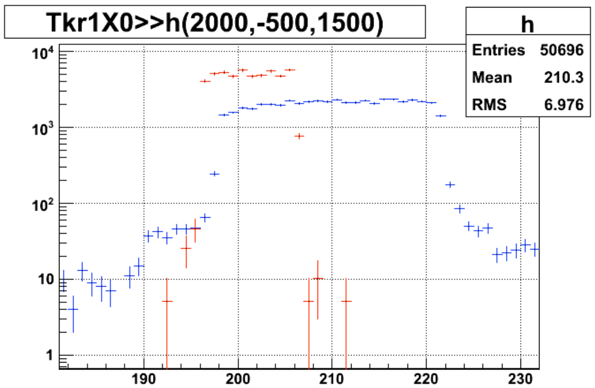


In order to do a good job, we really need to go over the data, quantify the beam (*individually for each data run*), and then produce MC simulations specifically for those runs

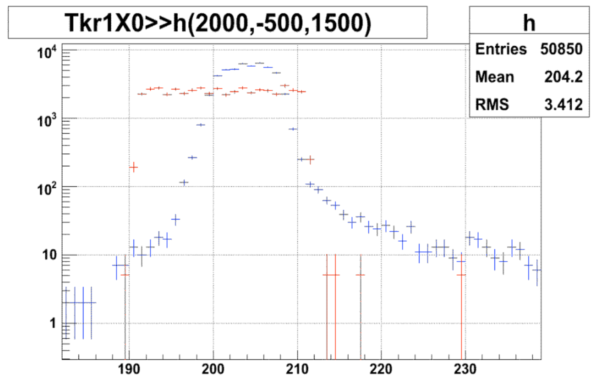
Incidence angle, 00 deg

MC in red; Data in blue

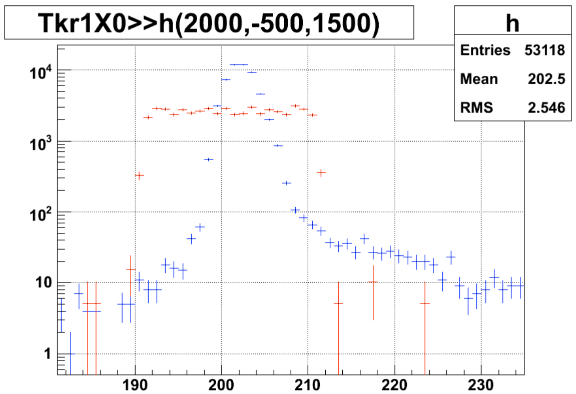
E = 282 GeV



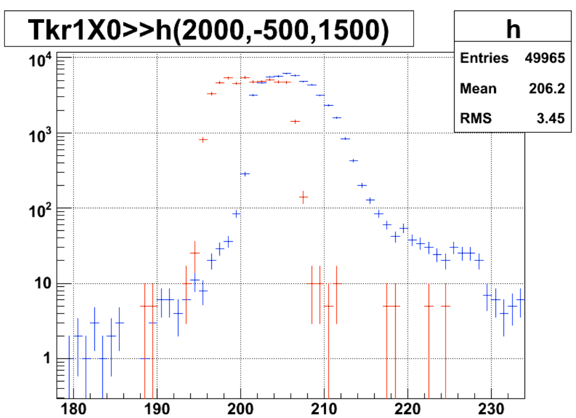
E = 200 GeV



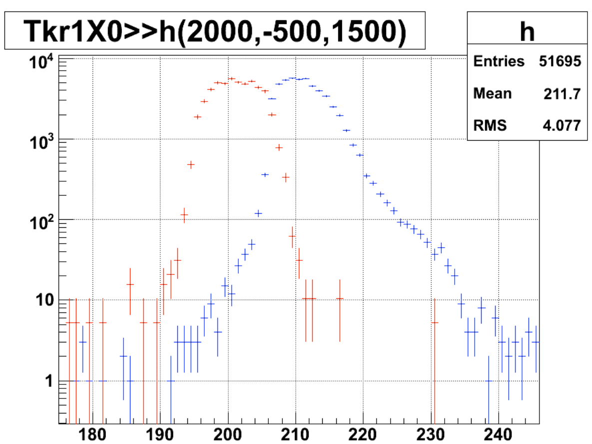
E = 100 GeV



E = 50 GeV



E = 20 GeV

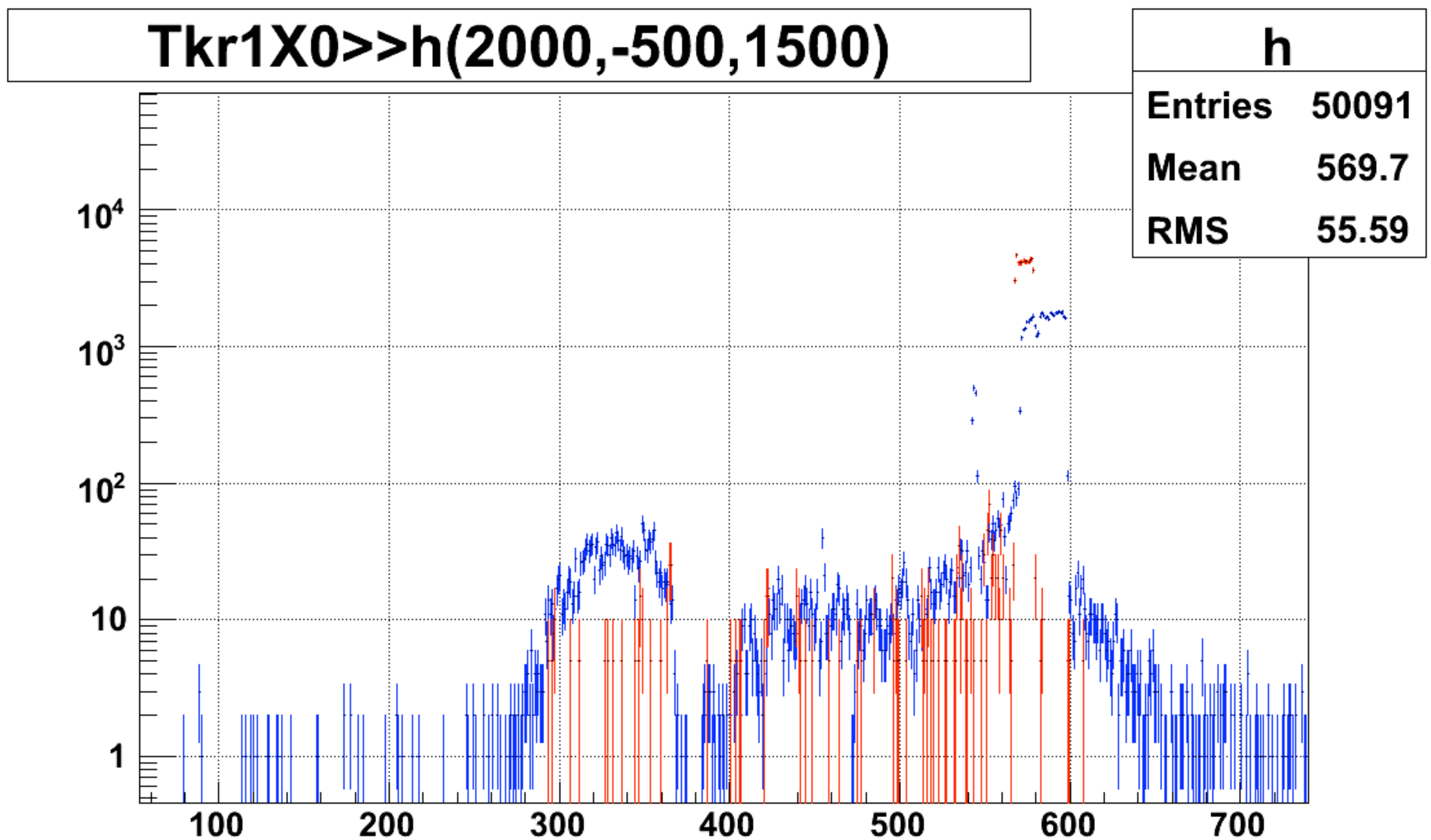


Incidence angle, **30 deg**

MC in red; **Data in blue**

E = 282 GeV

Data run 1942, MC run 167



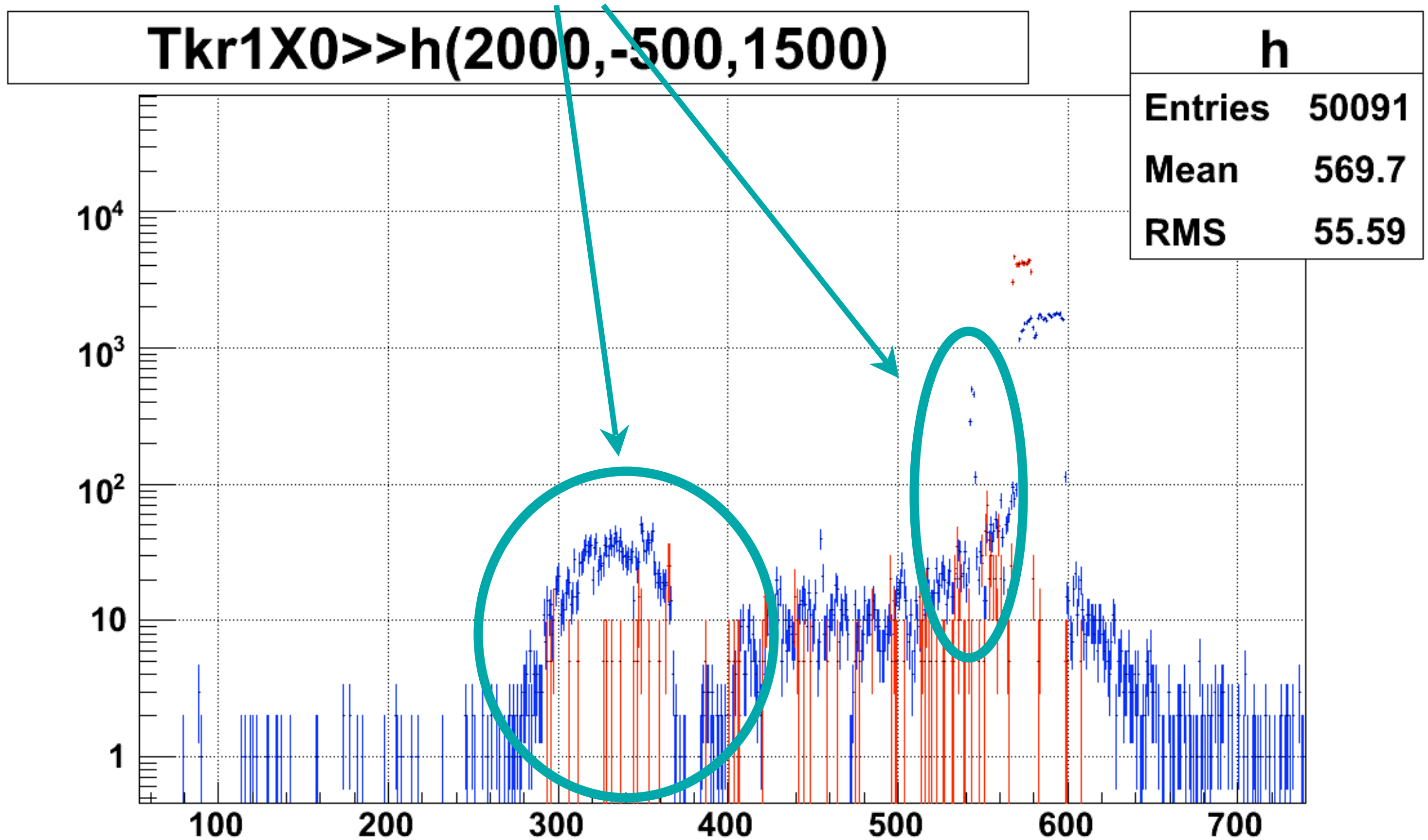
Incidence angle, 30 deg

MC in red; **Data in blue**

E = 282 GeV

Data run 1942, MC run 167

What is this ????



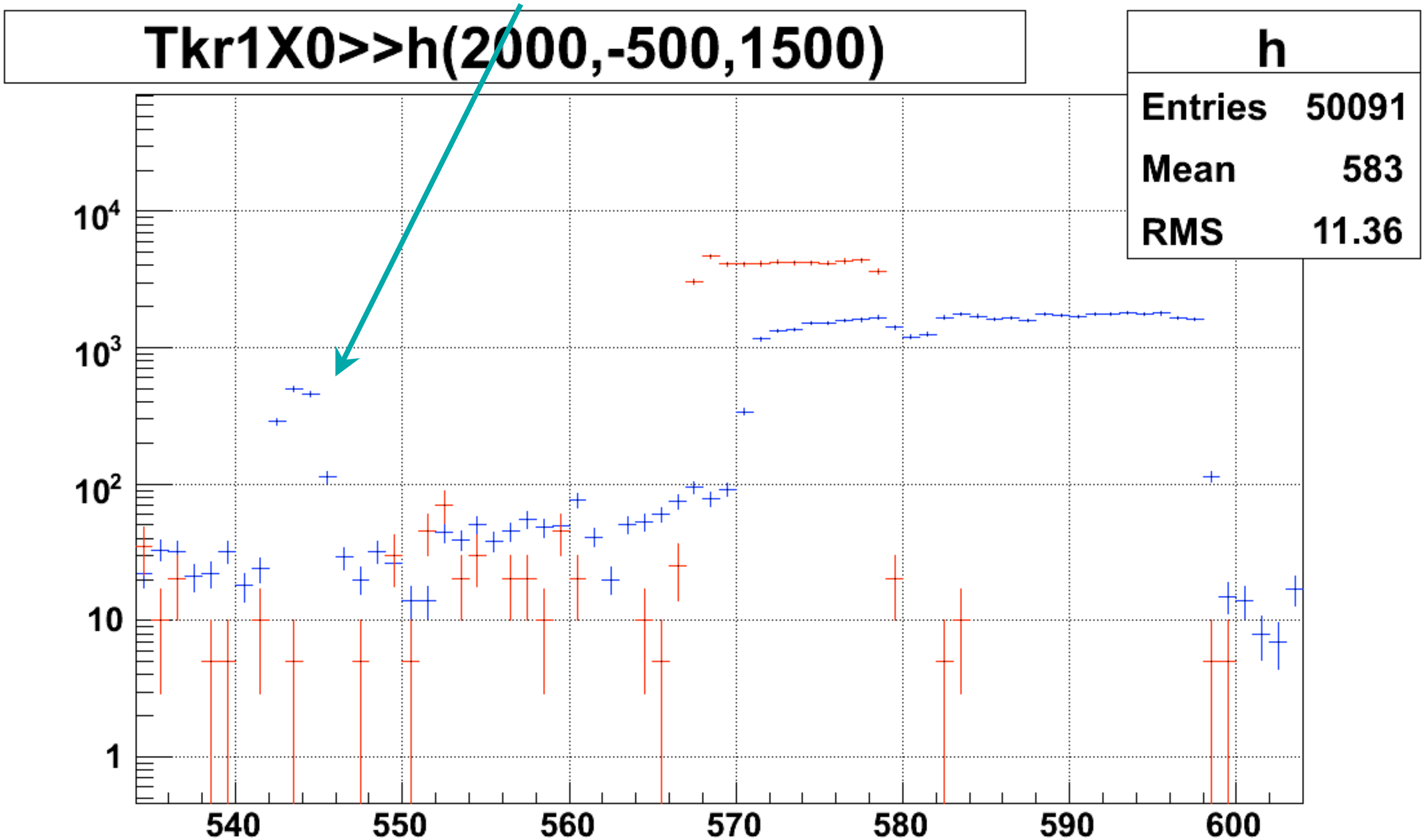
Incidence angle, 30 deg

MC in red; Data in blue

E = 282 GeV

Data run 1942, MC run 167

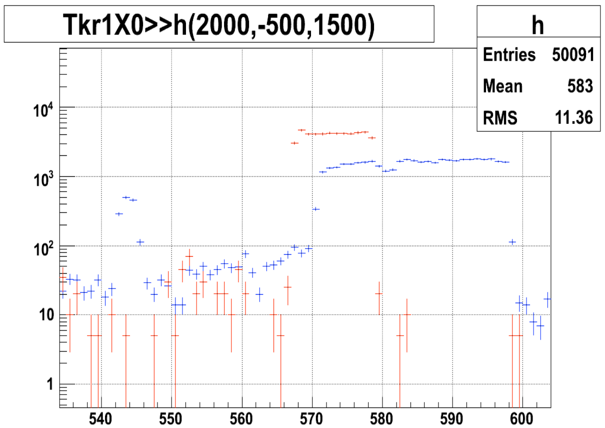
Blow up of previous image



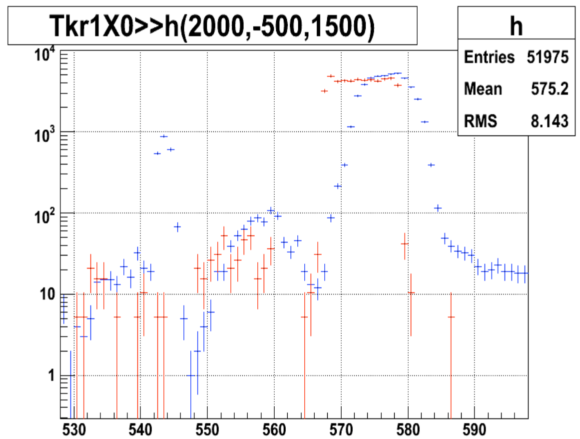
Incidence angle, 30 deg

MC in red; Data in blue

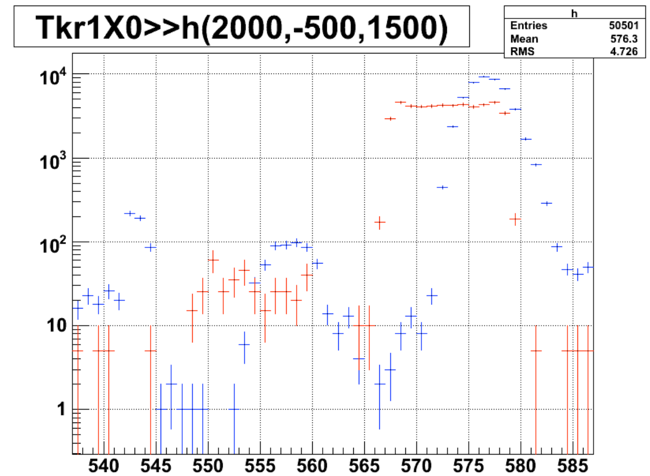
E = 282 GeV



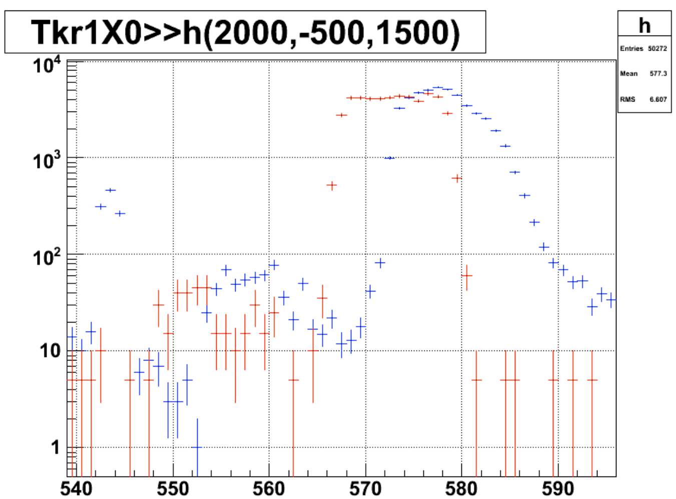
E = 200 GeV



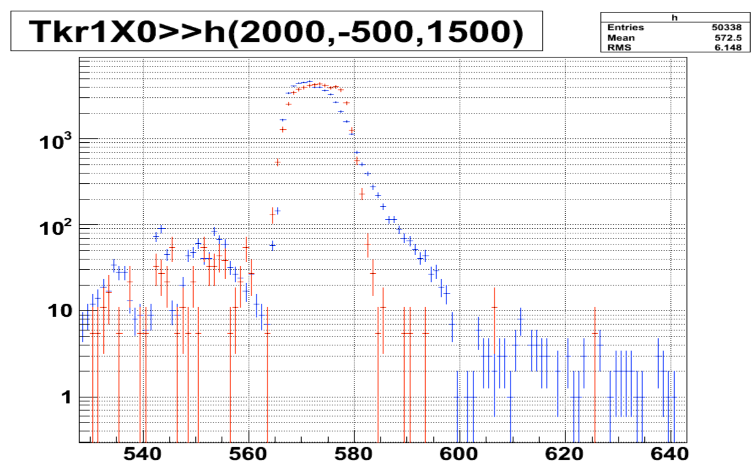
E = 100 GeV



E = 50 GeV

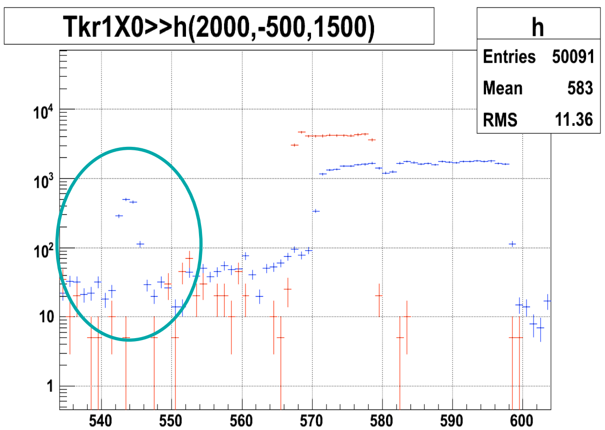


E = 20 GeV

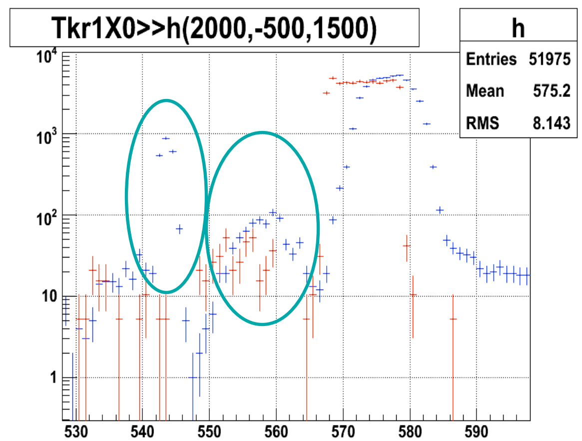


Incidence angle, 30 deg **MC in red; Data in blue**

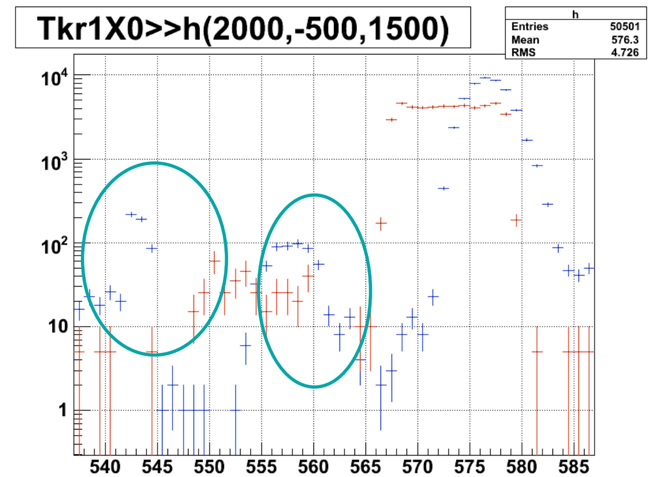
E = 282 GeV



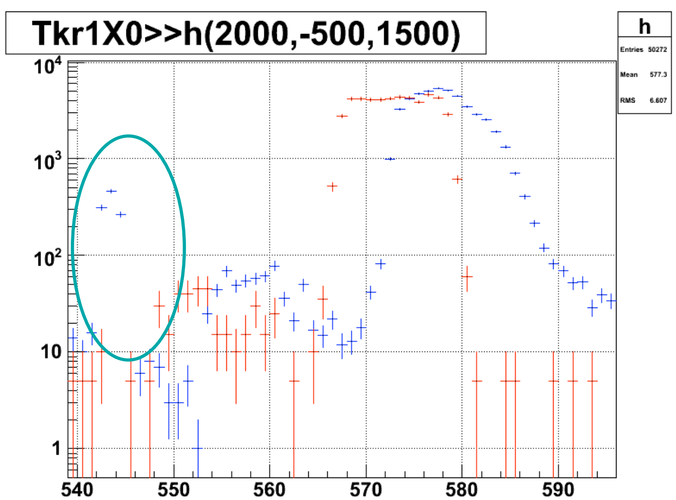
E = 200 GeV



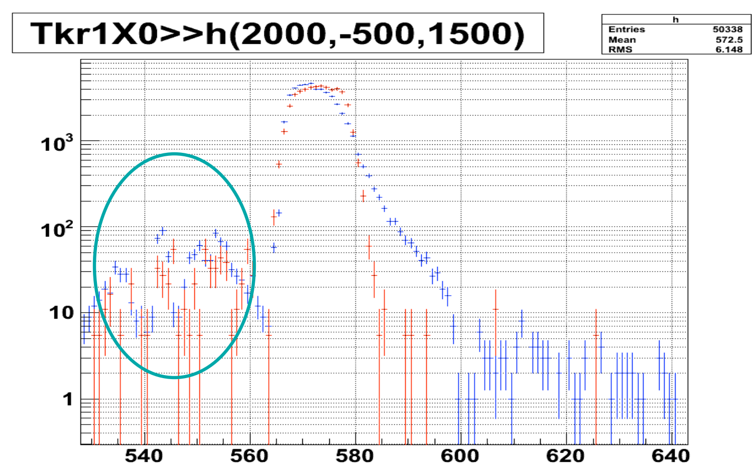
E = 100 GeV



E = 50 GeV



E = 20 GeV



Quick evaluation of agreement in Tkr1X0

Beam dimensions, and exact impact points are different.This also applies to Tkr1Y0. *This was indeed expected*

Secondary peaks in the data show up at 30 degrees, But NOT at 0 degrees

These peaks DO NOT exist in Tkr1Y0

What are they ????

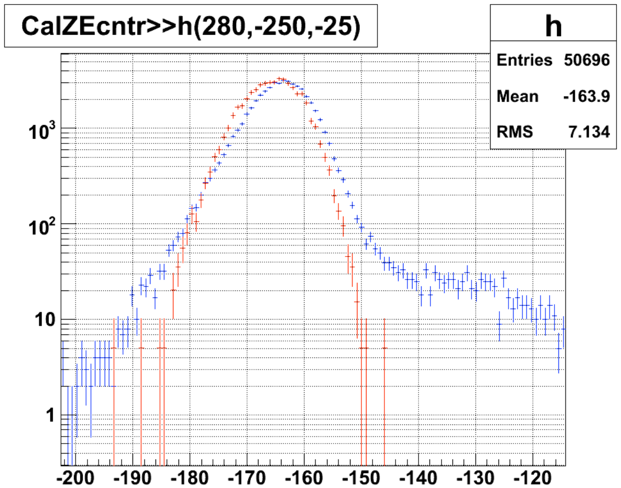
Inicially I thought they were electrons not interacting in the first layer… but this hypothesis is NOT correct (at least for the prominent secondary peak) due to:

1 - Dimensions of the secondary peak, which are significantly smaller than the main one
2 - Location of the peak, which is ~4 cm, instead of the ~2cm that would correspond to the 30 degrees and the separation between layer

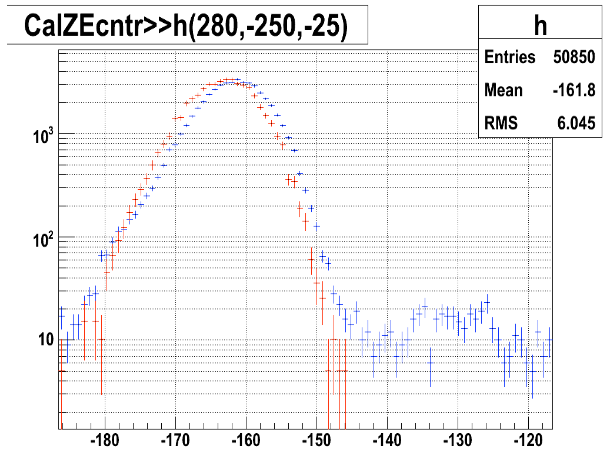
Incidence angle, 00 deg

MC in red; **Data in blue**

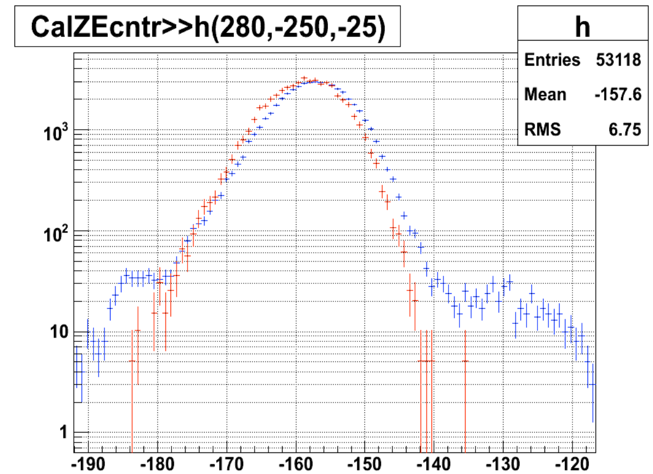
E = 282 GeV



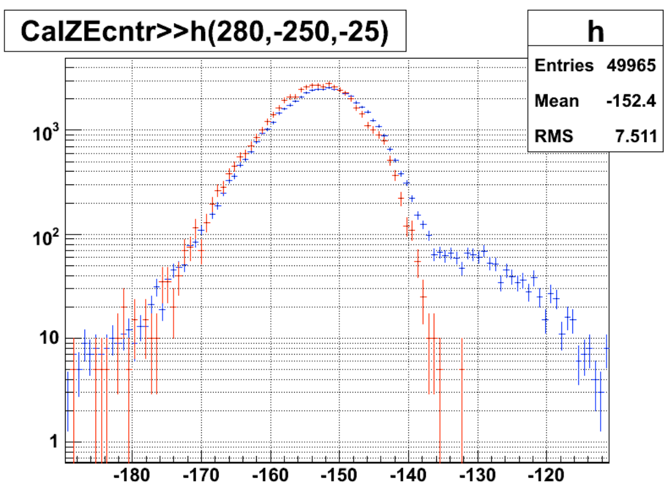
E = 200 GeV



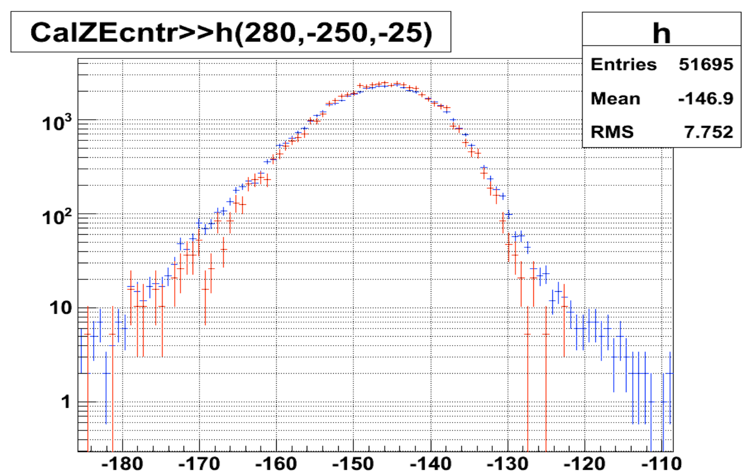
E = 100 GeV



E = 50 GeV



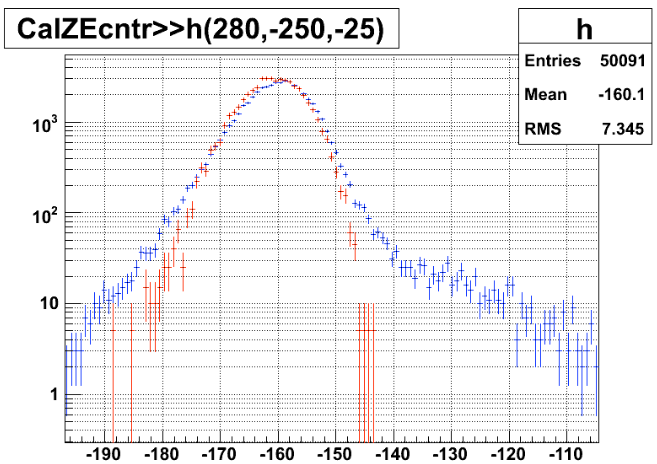
E = 20 GeV



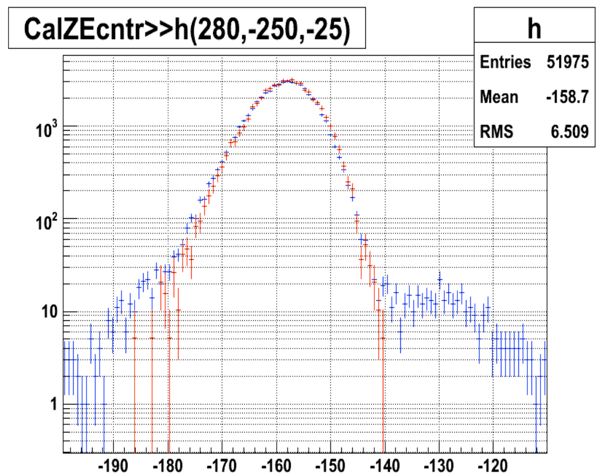
Incidence angle, 30 deg

MC in red; Data in blue

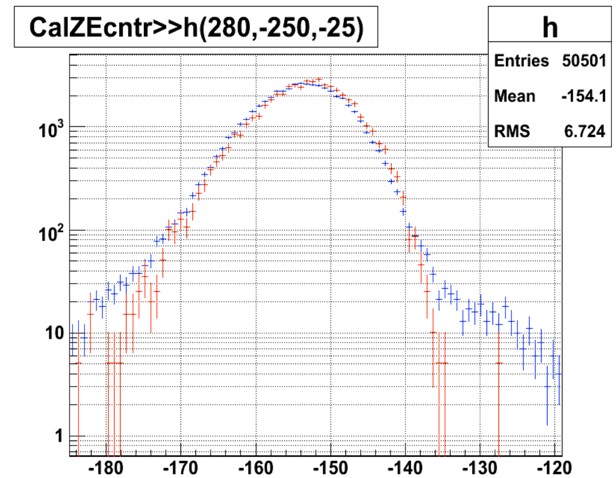
E = 282 GeV



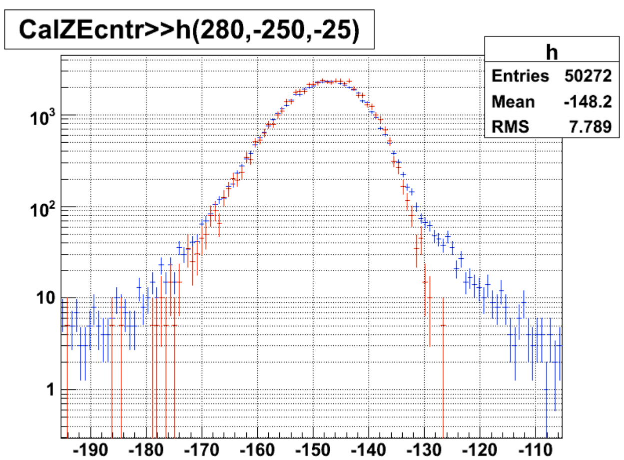
E = 200 GeV



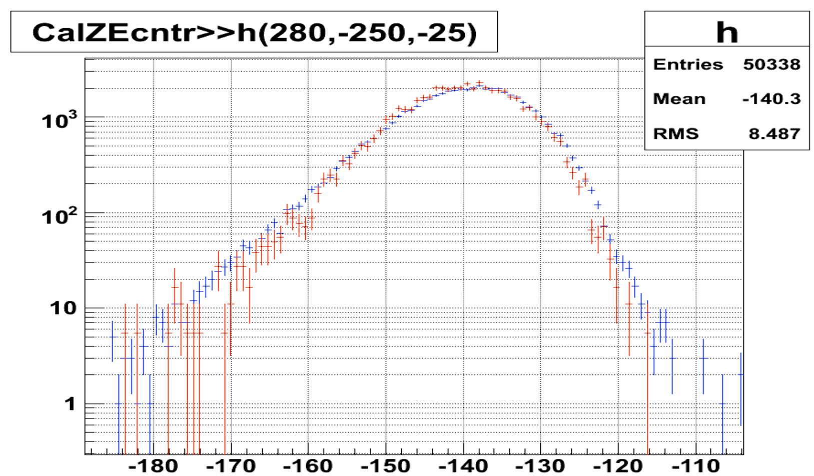
E = 100 GeV



E = 50 GeV



E = 20 GeV



Quick evaluation of agreement in CalZEcntr

Tail in the Data distribution at low values is larger than that of the MC. This might be due to the presence of MIPs in the data, which are NOT present in the MC. *This still needs to be confirmed*

At 0 degrees,
The main peak of the distributions are “shifted” (MC goes deeper in the calorimeter) at the highest energies. *When energy decreases, the distributions get more similar*

At 30 degrees,

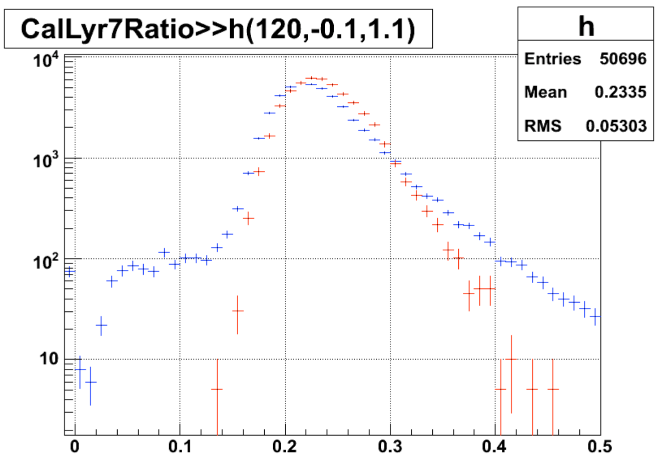
The main peak of the distributions agree very well basically for all energies

How can we explain that ??

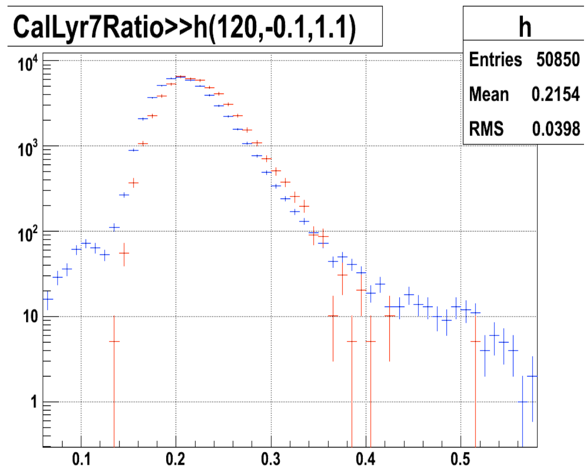
Incidence angle, 0 deg

MC in red; Data in blue

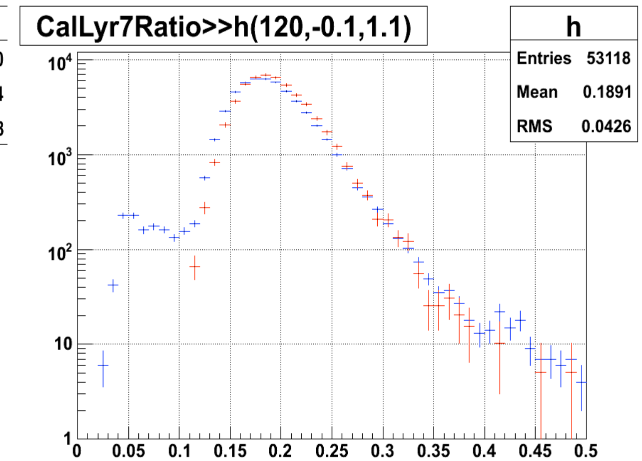
E = 282 GeV



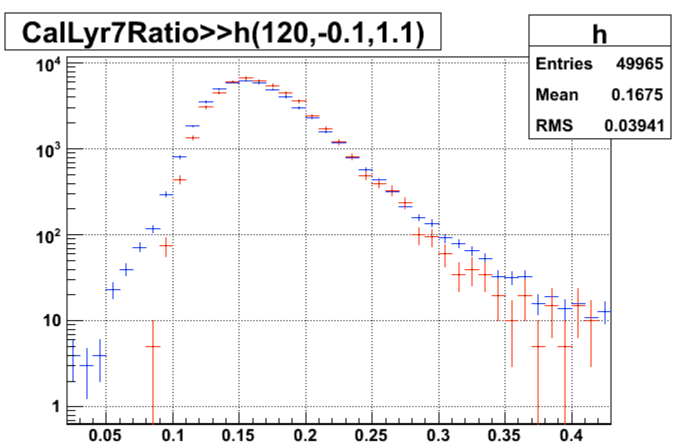
E = 200 GeV



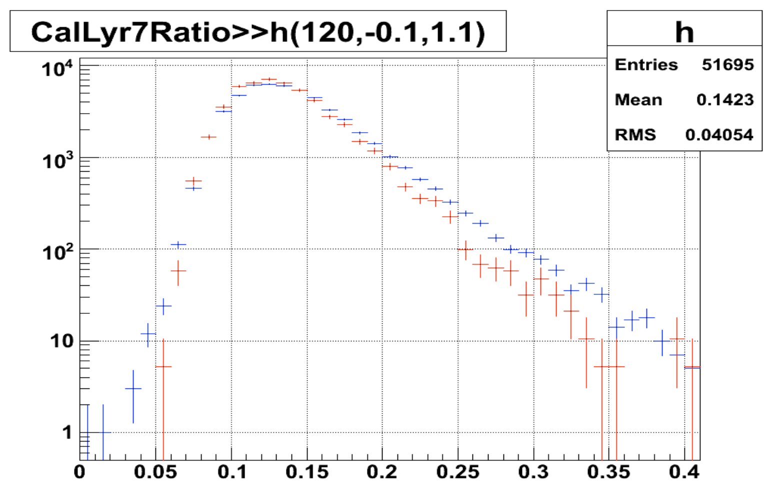
E = 100 GeV



E = 50 GeV



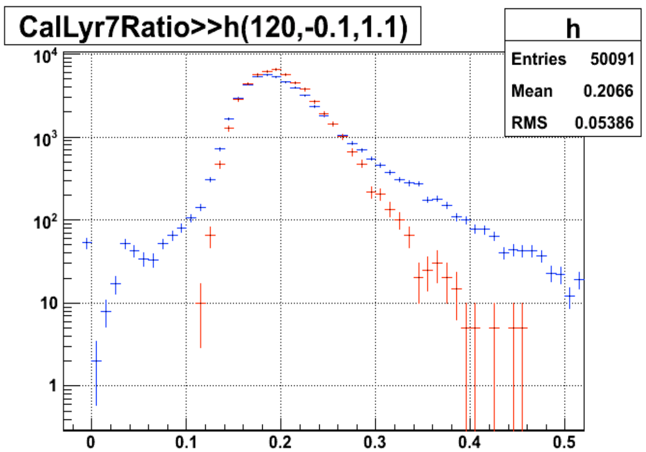
E = 20 GeV



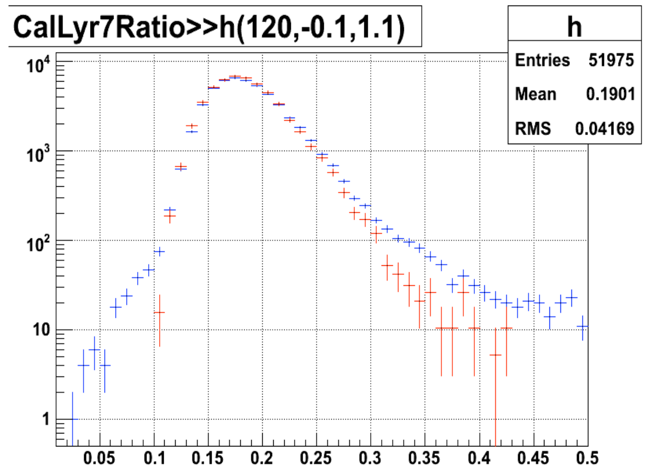
Incidence angle, 30 deg

MC in red; Data in blue

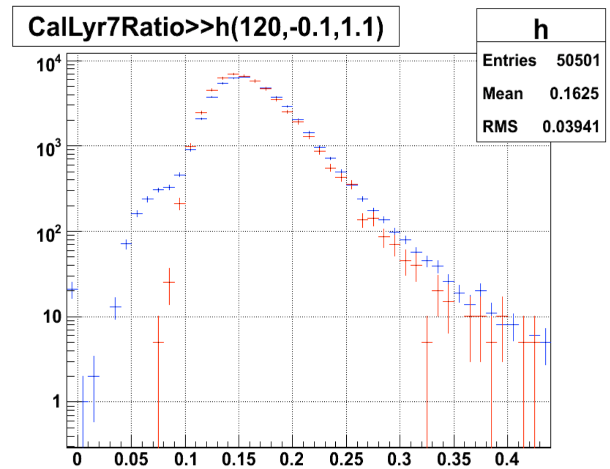
E = 282 GeV



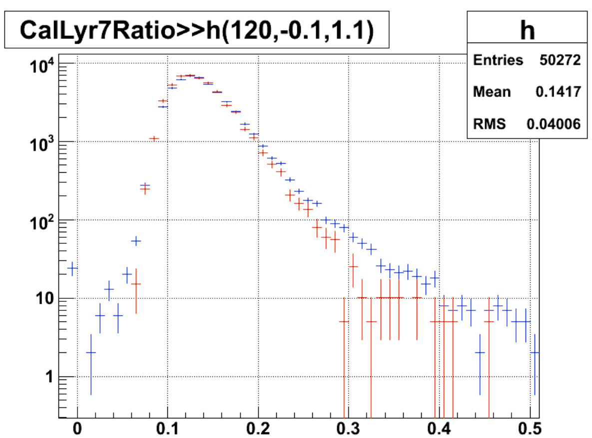
E = 200 GeV



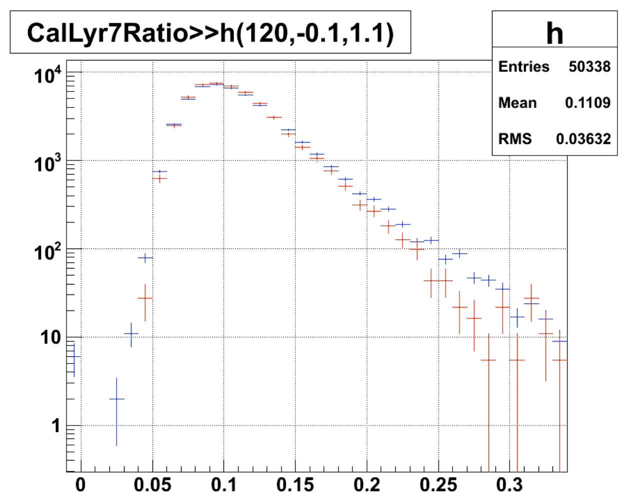
E = 100 GeV



E = 50 GeV



E = 20 GeV



Quick evaluation of agreement in CalLyr7Ratio

Tails of distributions are larger in the data, and specially at high energies. At the lowest energies the tails agree better

The lower tail might be due to MIPs, which are in the data and (obviously) not in the MC. Currently, I do not have hypothesis for the high tail…

At 0 degrees, the main peaks of the distributions are "shifted" (MC has a larger fraction of energy in the last layer) at the highest energies. *When energy decreases, the distributions get more similar*

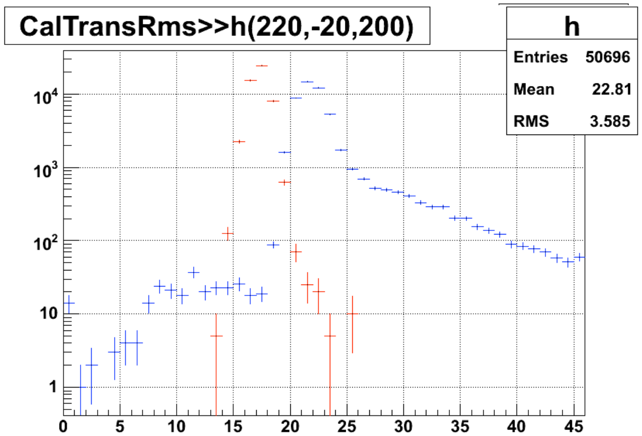
At 30 degrees,The main peak of the distributions agree very well basically for all energies

Exactly the same occurred with CalZEcntr Currently, the behaviour of these two variables is (for me) a mistery…

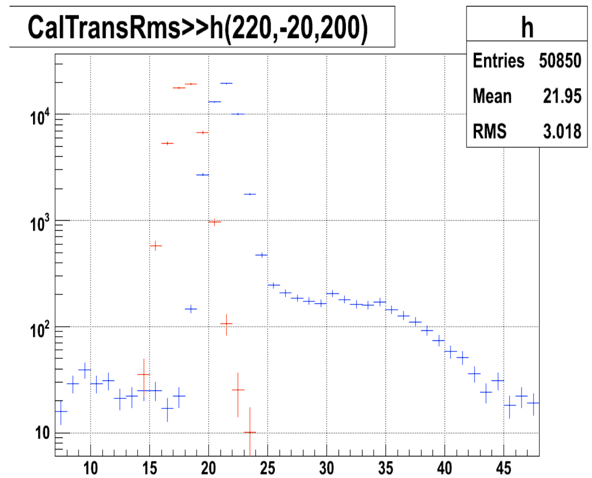
Incidence angle, 00 deg

MC in red; Data in blue

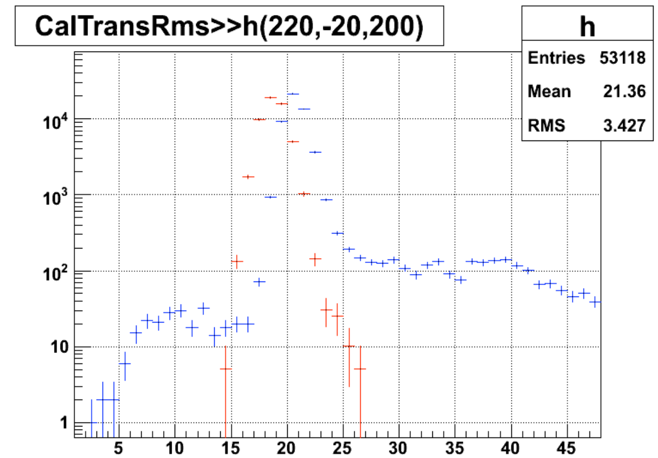
E = 282 GeV



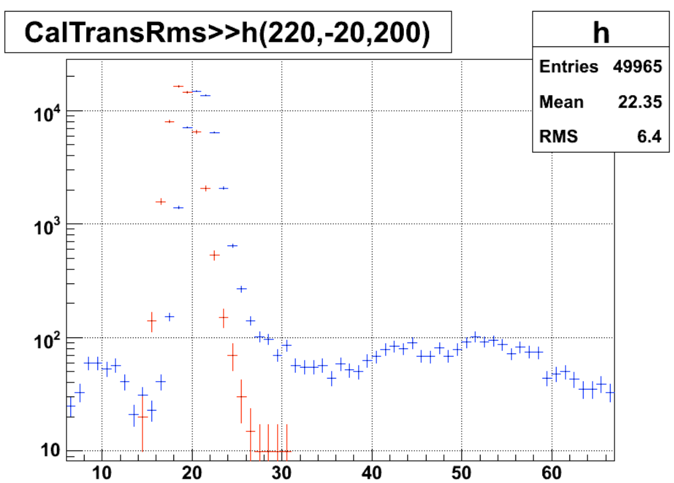
E = 200 GeV



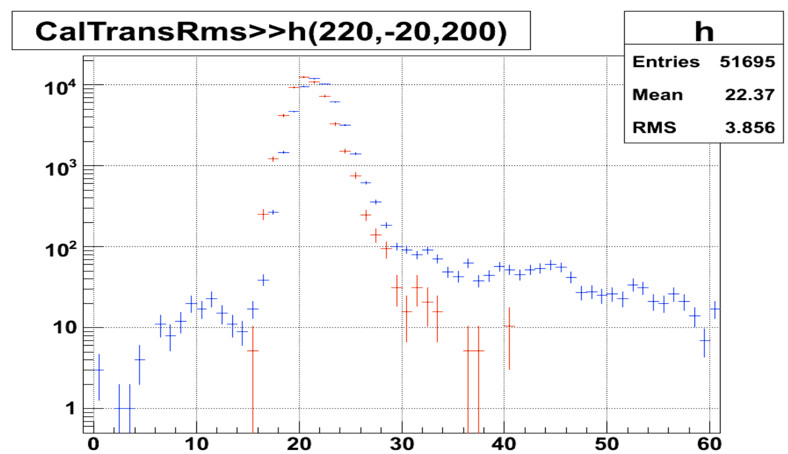
E = 100 GeV



E = 50 GeV

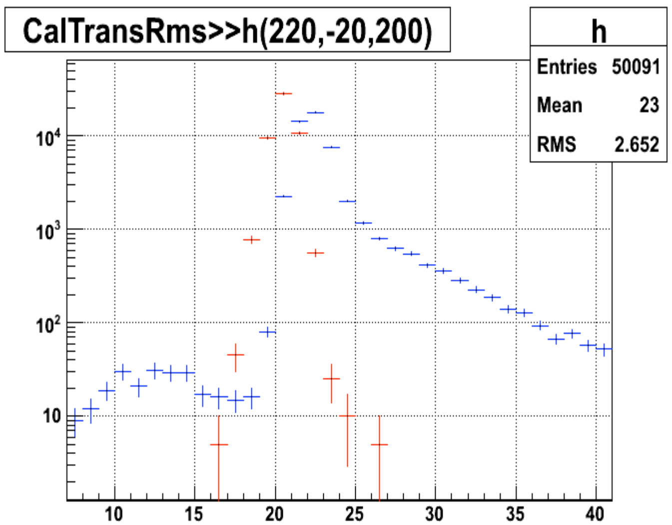


E = 20 GeV

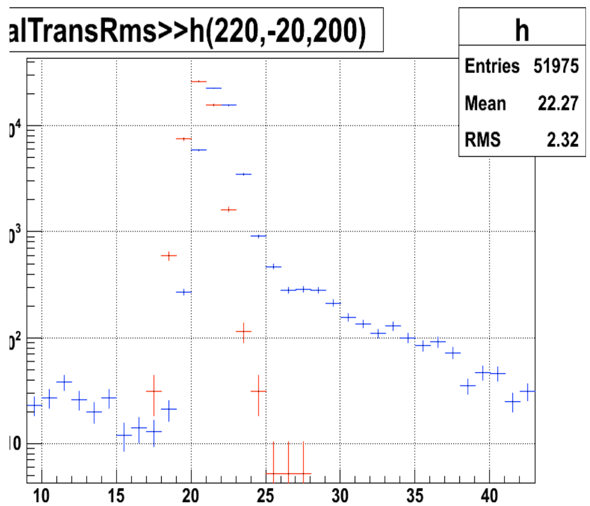


Incidence angle, 30 deg MC in red; Data in blue

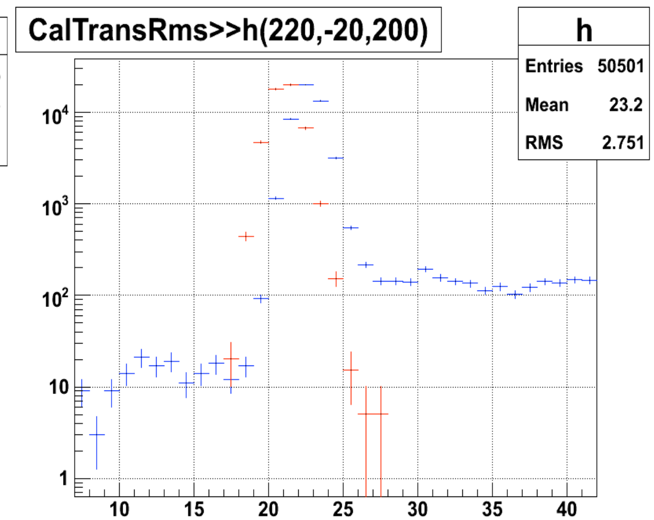
E = 282 GeV



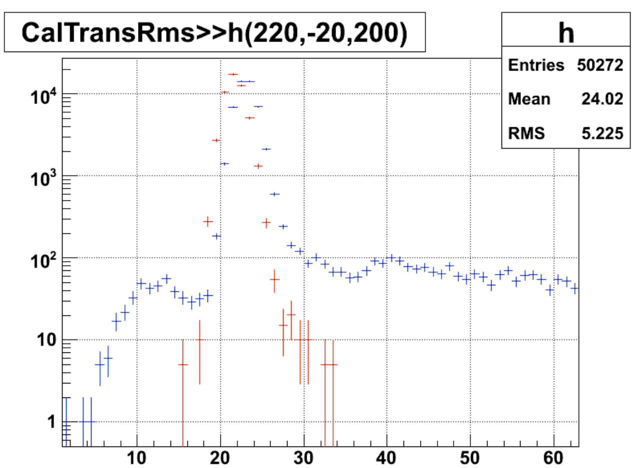
E = 200 GeV



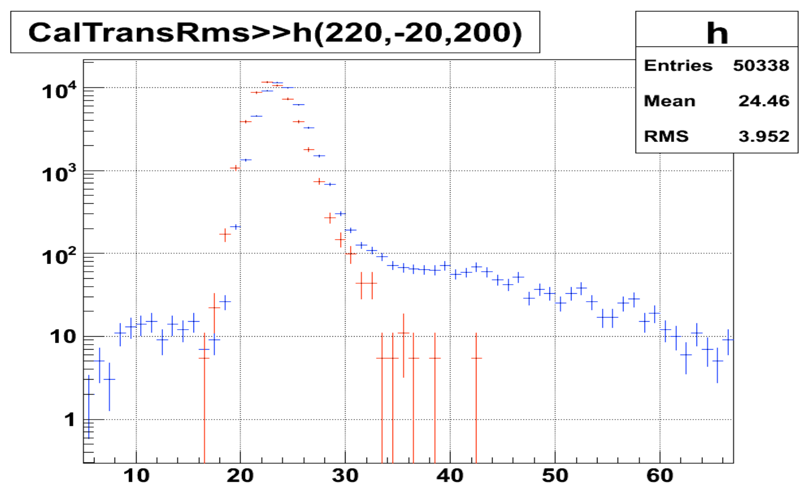
E = 100 GeV



E = 50 GeV



E = 20 GeV



Quick evaluation of agreement in CalTransRms

Tails of distributions are larger in the data for all energies

The lower tail might be due to MIPs, which are in the data and (obviously) not in the MC. Currently, I do not have hypothesis for the high tail… it might be related to the high tail in CalLyryRatio

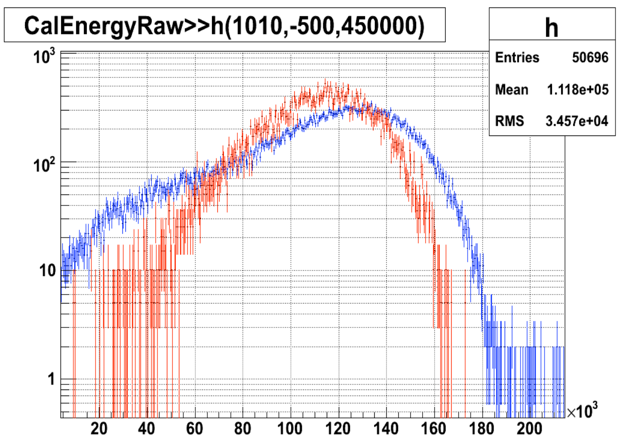
At 0 degrees, the main peaks of the distributions are "shifted" (MC showers are narrower) at the highest energies. *When energy decreases, the distributions get more similar*

At 30 degrees, the same occurs; yet the differences data-MC in the distributions peak seem to be a bit smaller

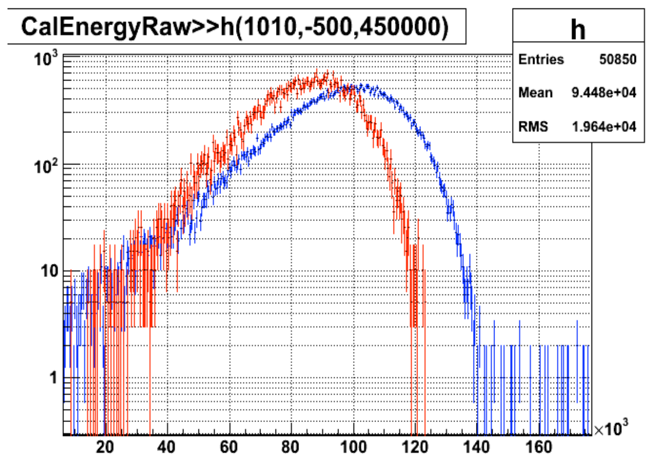
Incidence angle, 00 deg

MC in red; Data in blue

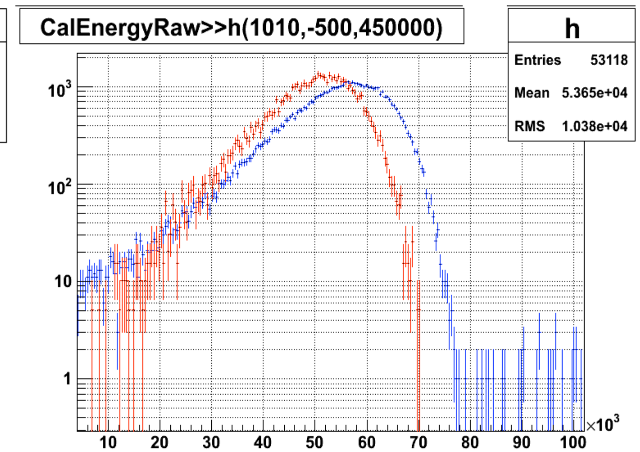
E = 282 GeV



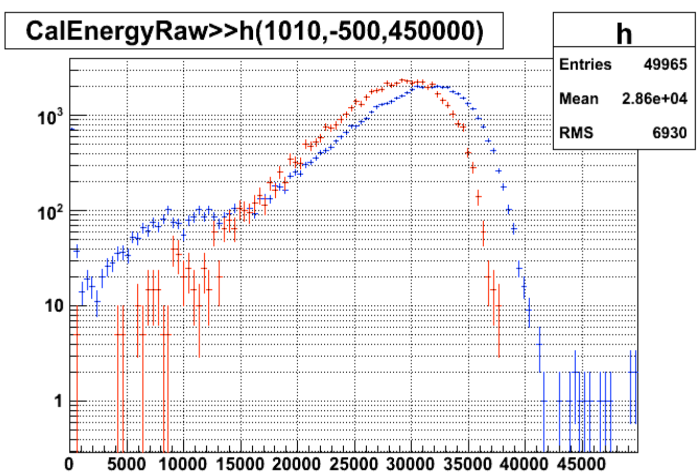
E = 200 GeV



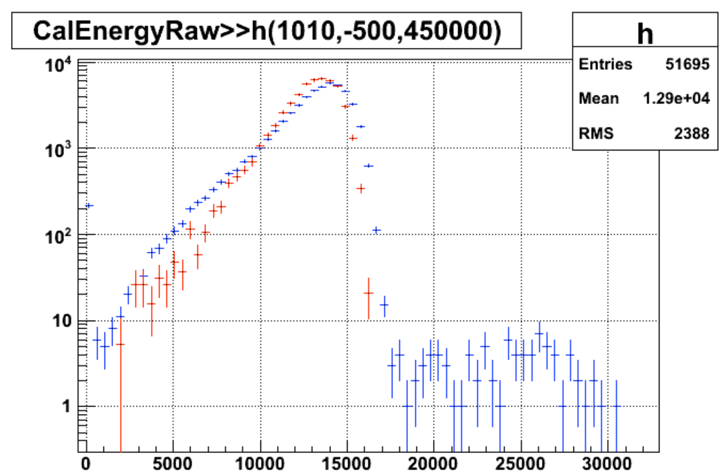
E = 100 GeV



E = 50 GeV



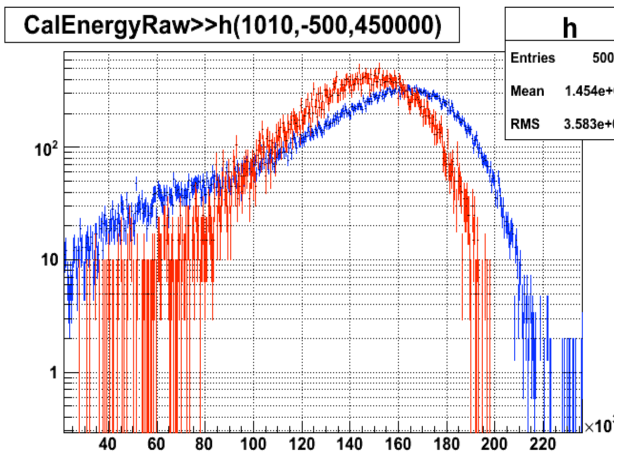
E = 20 GeV



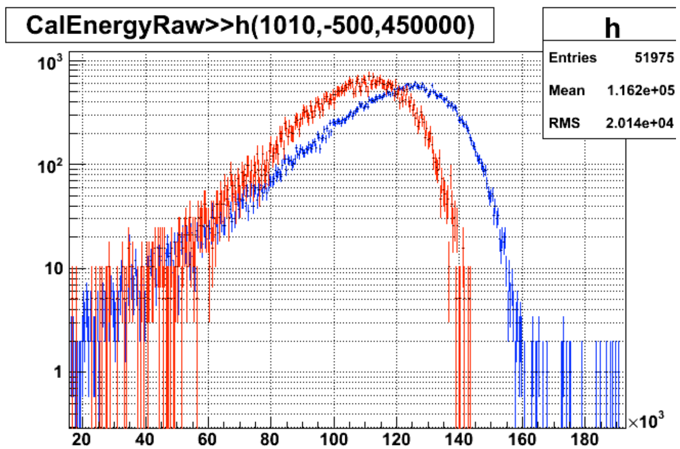
Incidence angle, 30 deg

MC in red; Data in blue

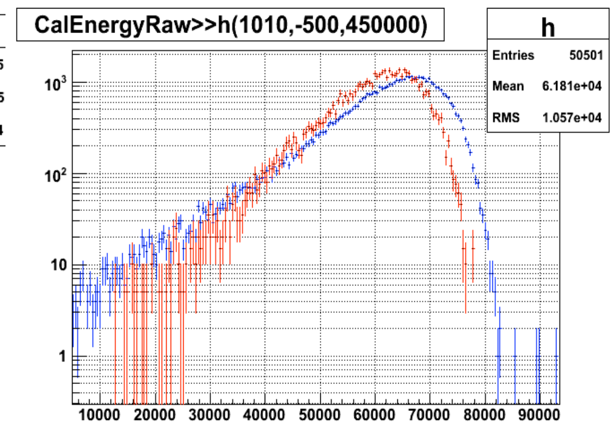
E = 282 GeV



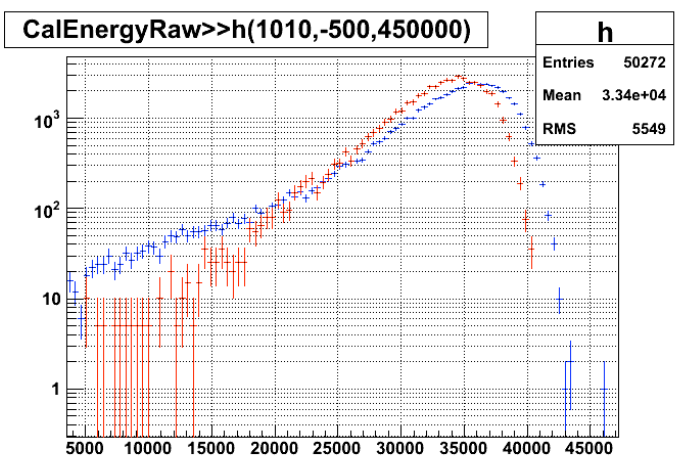
E = 200 GeV



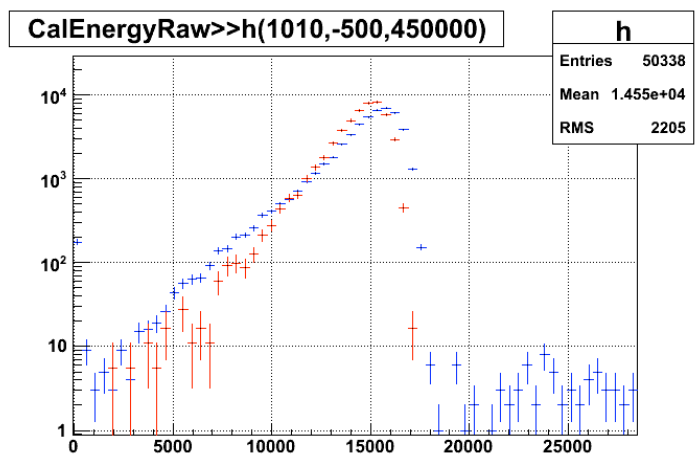
E = 100 GeV



E = 50 GeV



E = 20 GeV



Quick evaluation of agreement in CalEnergyRaw

The main peaks of the distributions are “shifted”; MC events have larger measured energy. *When energy of the electron beam decreases, the distributions get more similar. At the higher energies the peak is shifted by about 10 %, and the lowest by <=5%*

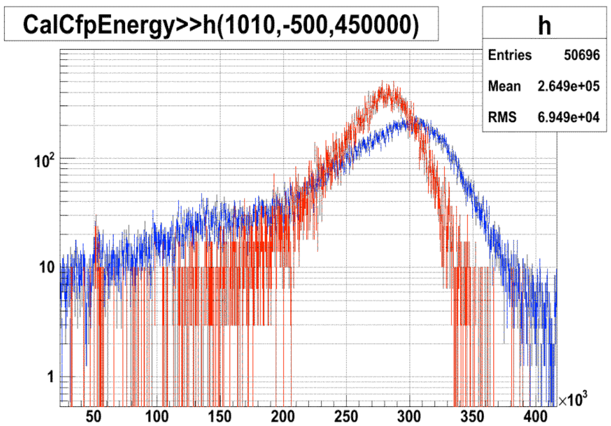
The same behaviour is observed at 0 and 30 degrees incidence angle

The distribution for data has a larger tail at lower energies. This CANNOT be due to the presence of MIPs in the data, since the energy deposited is far larger than the expected ~100 MeV for single MIPs

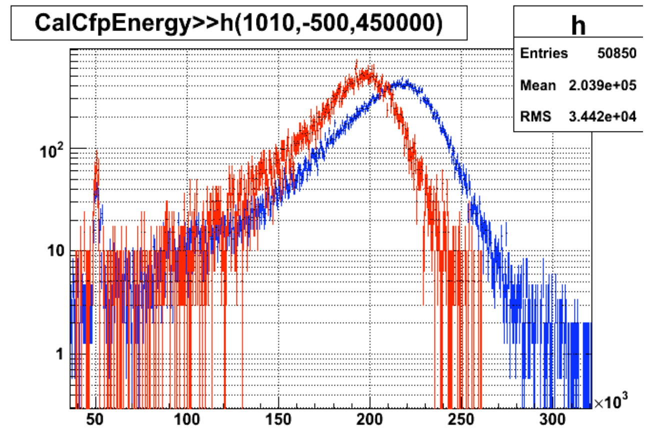
Incidence angle, 0 deg

MC in red; Data in blue

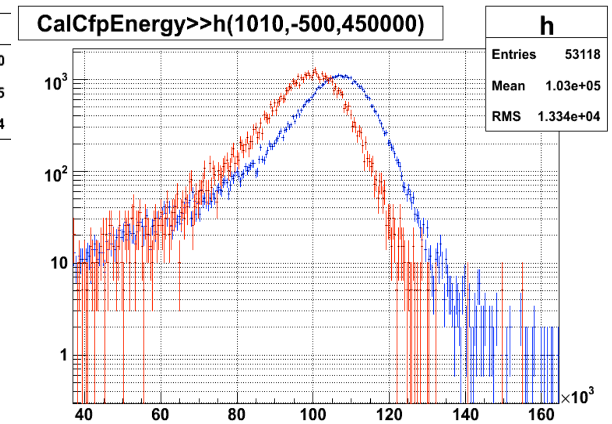
E = 282 GeV



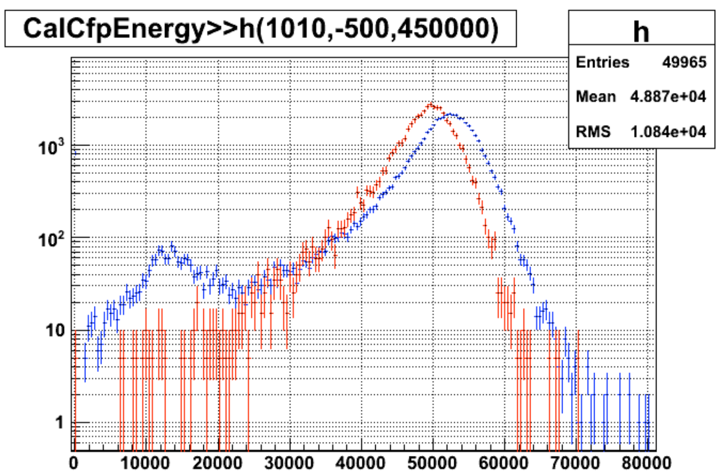
E = 200 GeV



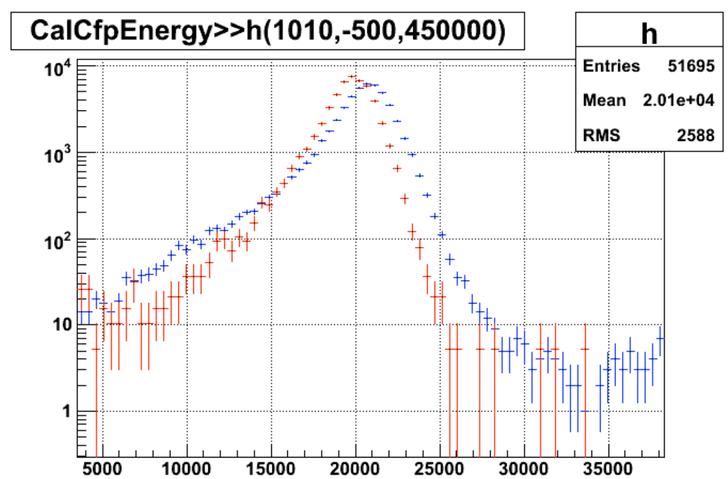
E = 100 GeV



E = 50 GeV



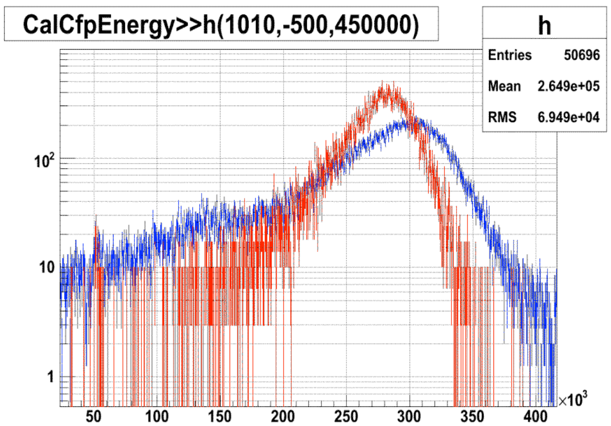
E = 20 GeV



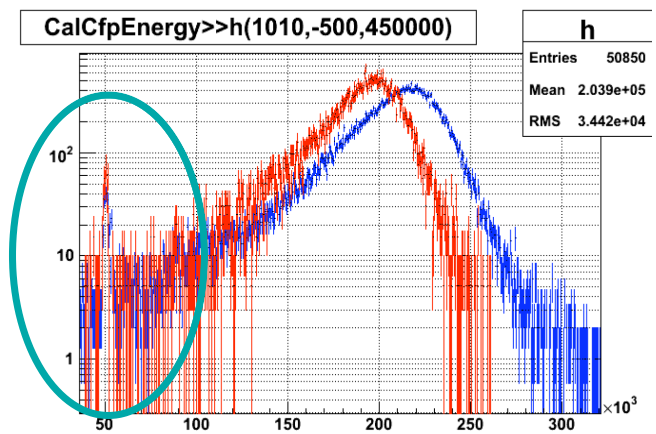
Incidence angle, 0 deg

MC in red; **Data in blue**

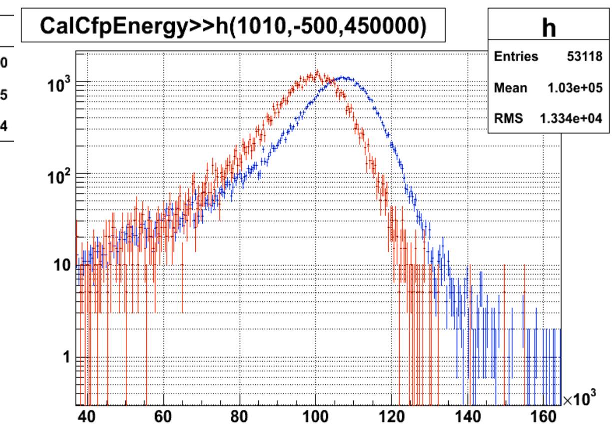
E = 282 GeV



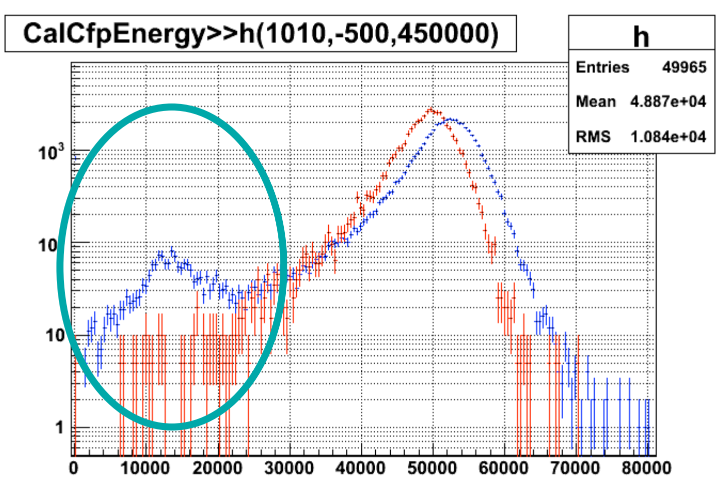
E = 200 GeV



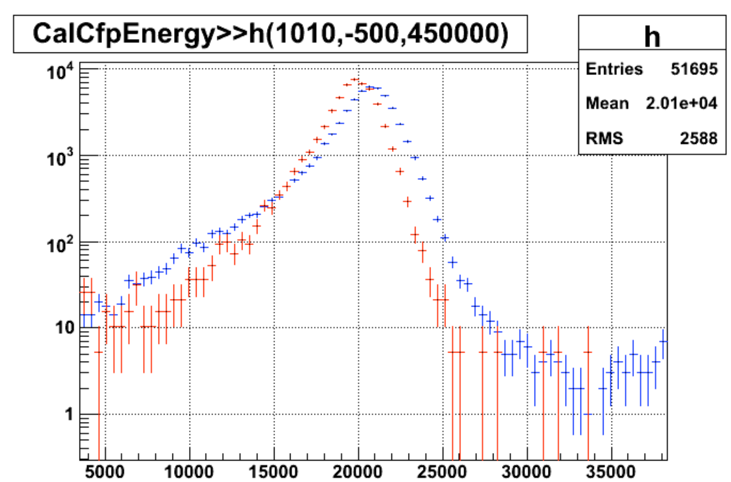
E = 100 GeV



E = 50 GeV



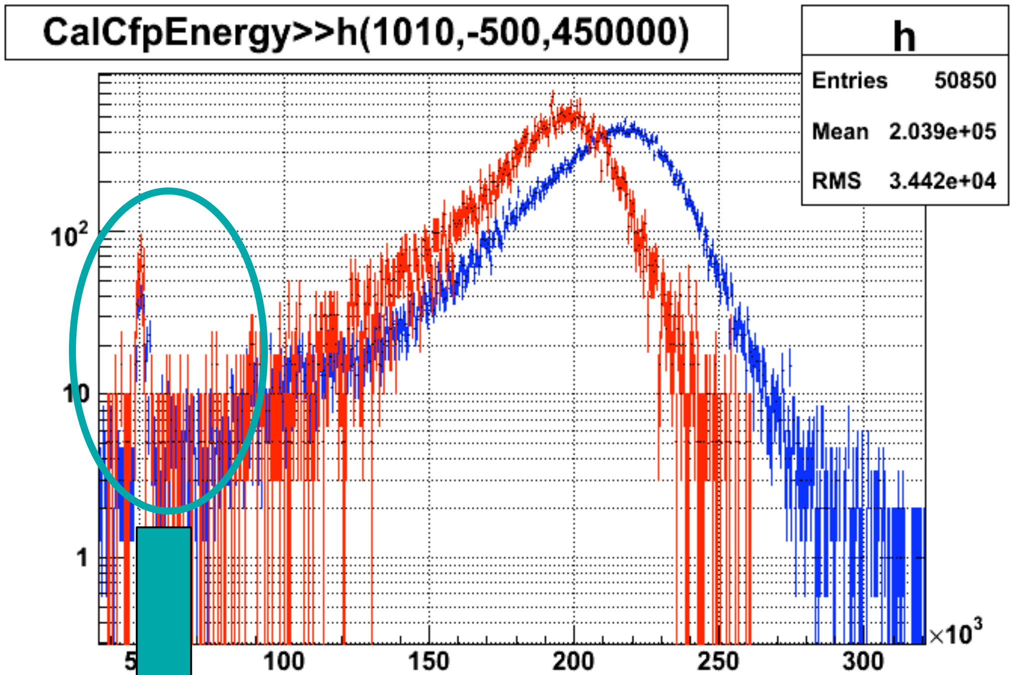
E = 20 GeV



Incidence angle, 0 deg

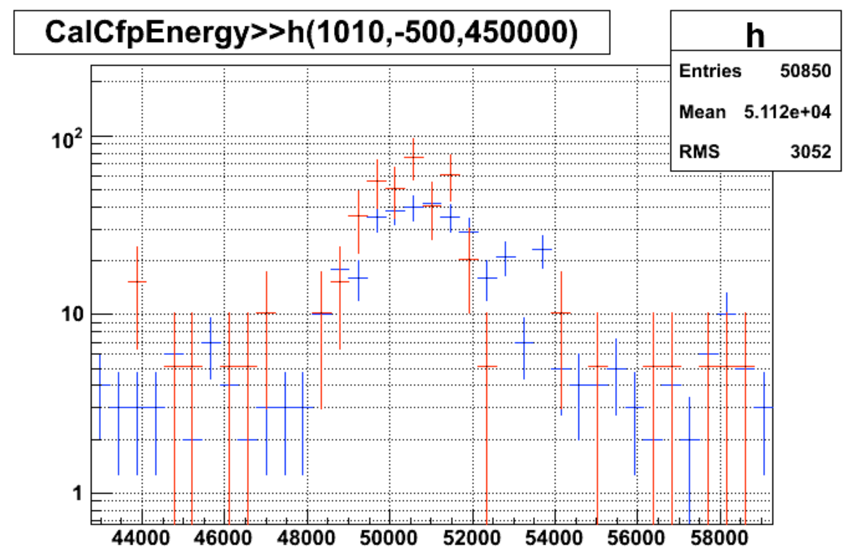
MC in red; Data in blue

E = 200 GeV; Data run 1911, MC run 164



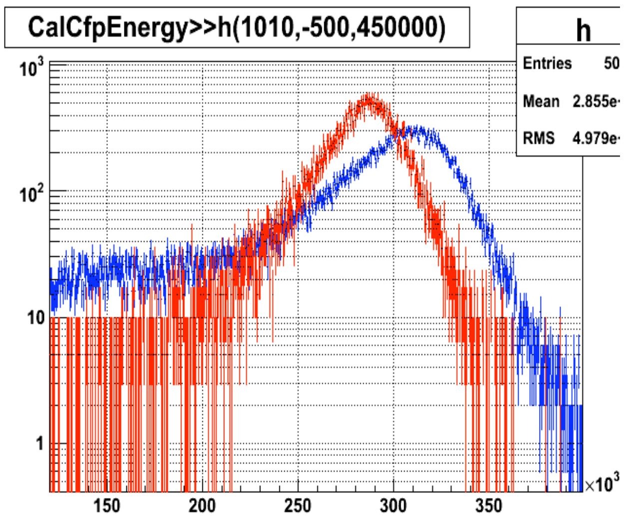
- **This feature is not visible in CalEnergyRaw**
- **It occurs for Data & MC**

What is that ?

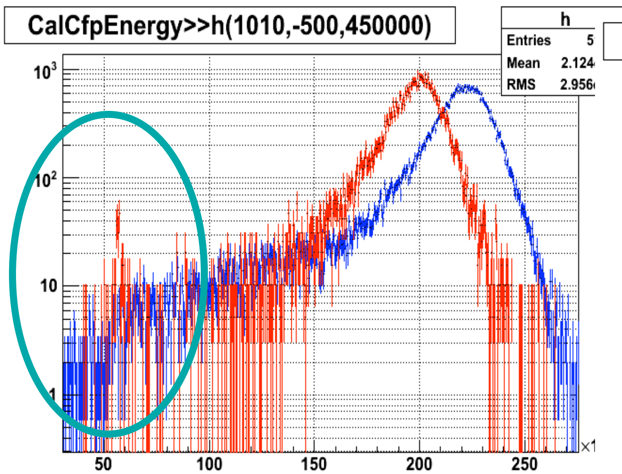


Incidence angle, 30 deg   **MC in red; Data in blue**

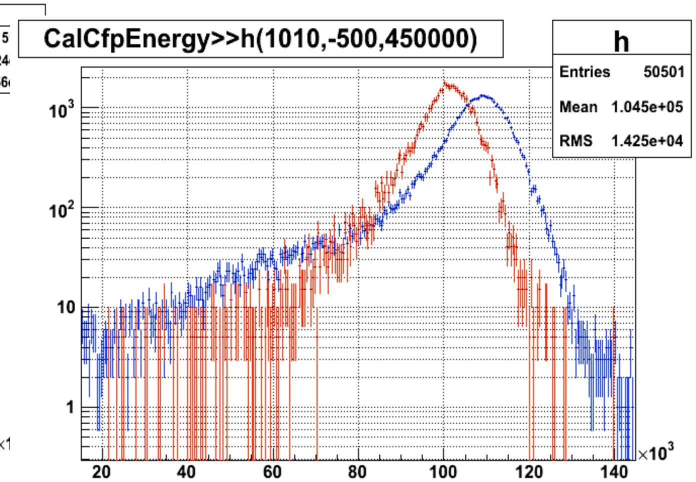
E = 282 GeV



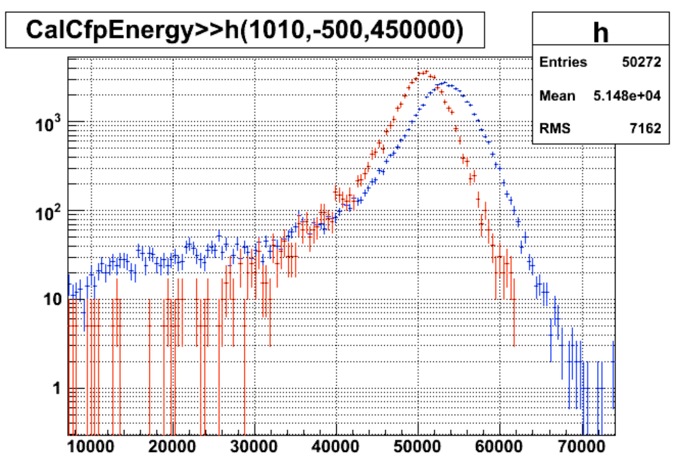
E = 200 GeV



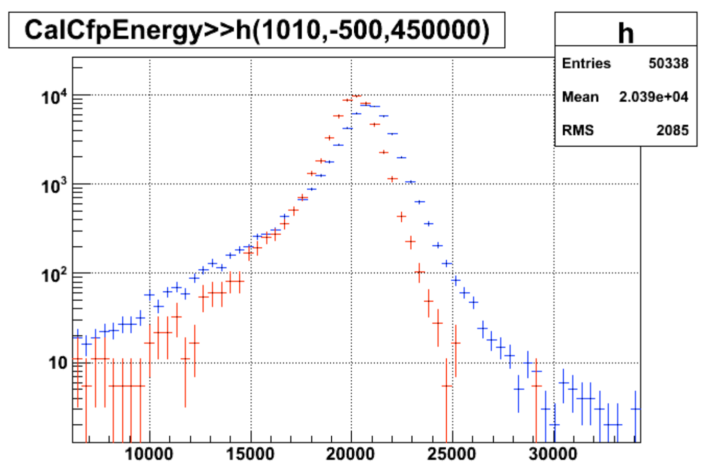
E = 100 GeV



E = 50 GeV

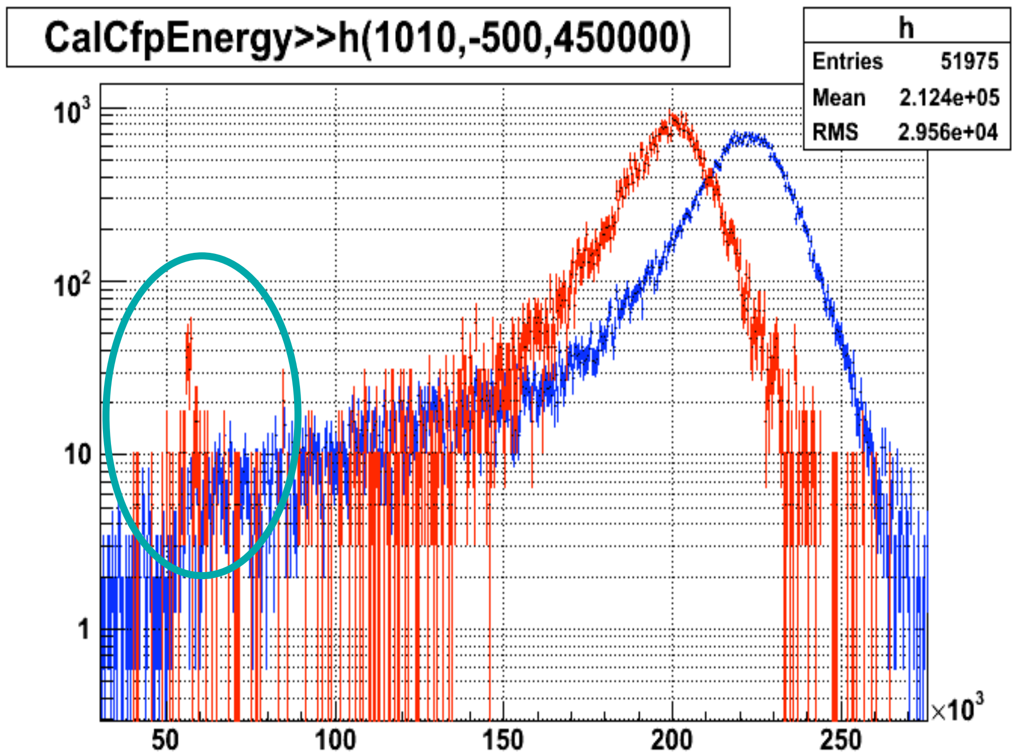


E = 20 GeV



Incidence angle, 30 deg **MC in red**; **Data in blue**

E = 200 GeV; Data run 1902, MC run 168



This time the feature only shows up in the MC, and NOT in the data…

Can it be related to the path of the secondary particles through a given region in the calorimeter ??

Quick evaluation of agreement in CalCfpEnergy

The main peaks of the distributions are “shifted”; MC events have larger measured energy. *When energy of the electron beam decreases, the distributions get more similar. At the higher energies the peak is shifted by about 10 %, and the lowest by <=5%*

The same behaviour is observed at 0 and 30 degrees incidence angle

The distribution for data has a larger tail at lower energies. This CANNOT be due to the presence of MIPs in the data, since the energy deposited is far larger than the expected ~100 MeV for single MIPs

These are the same features observed in CalEnergyRaw

Presence of secondary peaks on the low energy tail, which are still not well understood

CONCLUSIONS

There is an overall rough good agreement Data-MC

There are differences in several parameters. The differences data-mc depend on the (electron beam) energy and (sometimes) also on the beam incidence angle

Some differences (not very much discussed in this presentation) are due to the (currently) different beam dimensions an exact incidence angle and impact point in MC

Work is ongoing to make a table (run by run) with estimates (from the data itself) for

- beam dimensions and impact point in the CU
- beam incidence angle in the CU and beam divergence (convoluted with CU PSF…)

New MC runs will be produced with this info. This will fit our data better

CONCLUSIONS

Some differences data-mc might be coming from the cross-talk in the diode read out signals; which might not be (yet) properly calibrated (*see presentations by Benoit and Philippe about energy measured per layer…)*

Other differences might coming by the presence of MIPs in the data, which are (obviously) not present in the MC. These guys should be properly removed for further comparisons…

Some other differences are still not understood. I did not check events with FRED (yet). This might clarify many of the issues… this work is to be done

Opinions/Suggestions/Remarks by more experienced people are very much welcome