

Pion Data Quality

Erik Buhmann¹



Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG



Bundesministerium
für Bildung
und Forschung



Friedrich Naumann
STIFTUNG

FÜR DIE FREIHEIT

¹ eMail: erik.buhmann@desy.de

Links & Paths

Run list (read-only link):

https://docs.google.com/spreadsheets/d/1nZiltumo3yqxcxpDWMMRuEHWBUFFrsaTvkxp_ODQGBf0/edit?usp=sharing

Path to reconstructed root files:

May: /nfs/dust/ilc/group/flchcal/AHCAL_Testbeam_SPS_May2018/reco_rootfiles

June: /nfs/dust/ilc/group/flchcal/AHCAL_Testbeam_SPS_June2018/reco_rootfiles

Git repository with root macros used to create plots for this quality check:


[calice_ROOTmacros](#)

Confluence documentation: <https://confluence.desy.de/display/Calice/Run+List>



Summary of first talk: “Good run” criteria

Proposal:

- eSum peak bin within 5 % of each other (same energy)
 - Except 10 GeV, \bc of large e- contamination
- nHit peak bin within 5 % of each other (same energy)

Open for discussion!  → went with 6 %, as binning gives a ~1% error anyway

Afterwards:

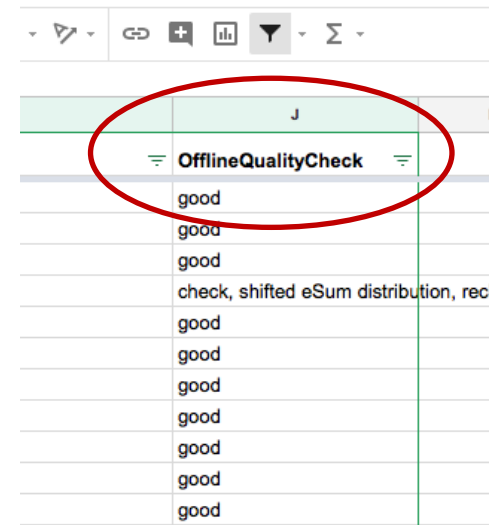
- New column in run list with flag ‘good run’ or flag ‘need-to-check run’ 
- Document criterias for ‘good run’ on confluence 

Outlook of first talk

- Fix criteria for ‘good runs’ ✓
 - Document on confluence ✓
- Add quality flag in run list ✓
- ‘Need-to-check’ runs need to be investigated
 - + Comparison between PP and No_PP mode → Naoki
- Quality check for June pion data ✓
- Quality check for electron data & muon data → Amine & Daniel
- Move run list away from Google sheet ✗ (not yet)

OfflineQualityCheck flags

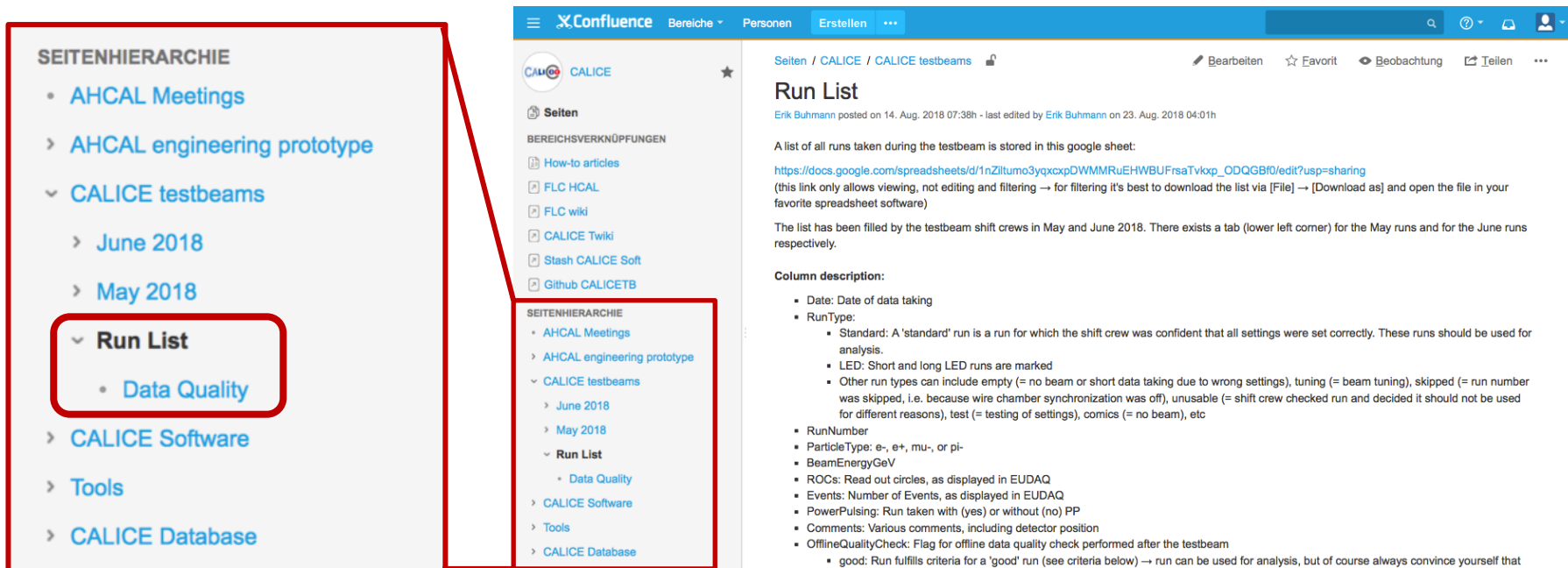
- Introduced new column for quality check in run list ('OfflineQualityCheck')
- Three flags:
 - good
 - check
 - bad (not yet used)
- Checked all 'standard' runs (*Runtype* in runlist):
 - May: 150 runs
 - June: 121 runs
- Checked only the HCAL part of the runs (no PS or TC for June data as calibration constants are off at the moment)



J		K
OfflineQualityCheck		
good		
good		
good		
check, shifted eSum distribution, rect		
good		
good		
good		
good		
good		
good		
good		

Confluence documentation

- More detailed description can be found on [confluence](#):



The screenshot displays the Confluence interface for the CALICE project. On the left, a sidebar navigation menu titled 'SEITENHIERARCHIE' (Side Navigation) lists various pages. The 'Run List' page is highlighted with a red box, and its sub-page 'Data Quality' is also highlighted. The main content area shows the 'Run List' page, which includes a description of the testbeam runs, a link to a Google Sheet, and a detailed column description for the data.

SEITENHIERARCHIE

- [AHCAL Meetings](#)
- › [AHCAL engineering prototype](#)
- ▼ [CALICE testbeams](#)
 - › [June 2018](#)
 - › [May 2018](#)
 - ▼ [Run List](#)
 - [Data Quality](#)
- › [CALICE Software](#)
- › [Tools](#)
- › [CALICE Database](#)

Confluence Page: Run List

Seiten / CALICE / CALICE testbeams

Run List

Erik Buhmann posted on 14. Aug. 2018 07:38h - last edited by Erik Buhmann on 23. Aug. 2018 04:01h

A list of all runs taken during the testbeam is stored in this google sheet:
https://docs.google.com/spreadsheets/d/1nZilumo3yqxcxpDWMuEHwBFrsaTvKxp_ODQGBf0/edit?usp=sharing
(this link only allows viewing, not editing and filtering → for filtering it's best to download the list via [File] → [Download as] and open the file in your favorite spreadsheet software)

The list has been filled by the testbeam shift crews in May and June 2018. There exists a tab (lower left corner) for the May runs and for the June runs respectively.

Column description:

- Date: Date of data taking
- RunType:
 - Standard: A 'standard' run is a run for which the shift crew was confident that all settings were set correctly. These runs should be used for analysis.
 - LED: Short and long LED runs are marked
 - Other run types can include empty (= no beam or short data taking due to wrong settings), tuning (= beam tuning), skipped (= run number was skipped, i.e. because wire chamber synchronization was off), unusable (= shift crew checked run and decided it should not be used for different reasons), test (= testing of settings), comics (= no beam), etc
- RunNumber
- ParticleType: e-, e+, mu-, or pi-
- BeamEnergyGeV
- ROCs: Read out circles, as displayed in EUDAQ
- Events: Number of Events, as displayed in EUDAQ
- PowerPulsing: Run taken with (yes) or without (no) PP
- Comments: Various comments, including detector position
- OfflineQualityCheck: Flag for offline data quality check performed after the testbeam
 - good: Run fulfills criteria for a 'good' run (see criteria below) → run can be used for analysis, but of course always convince yourself that

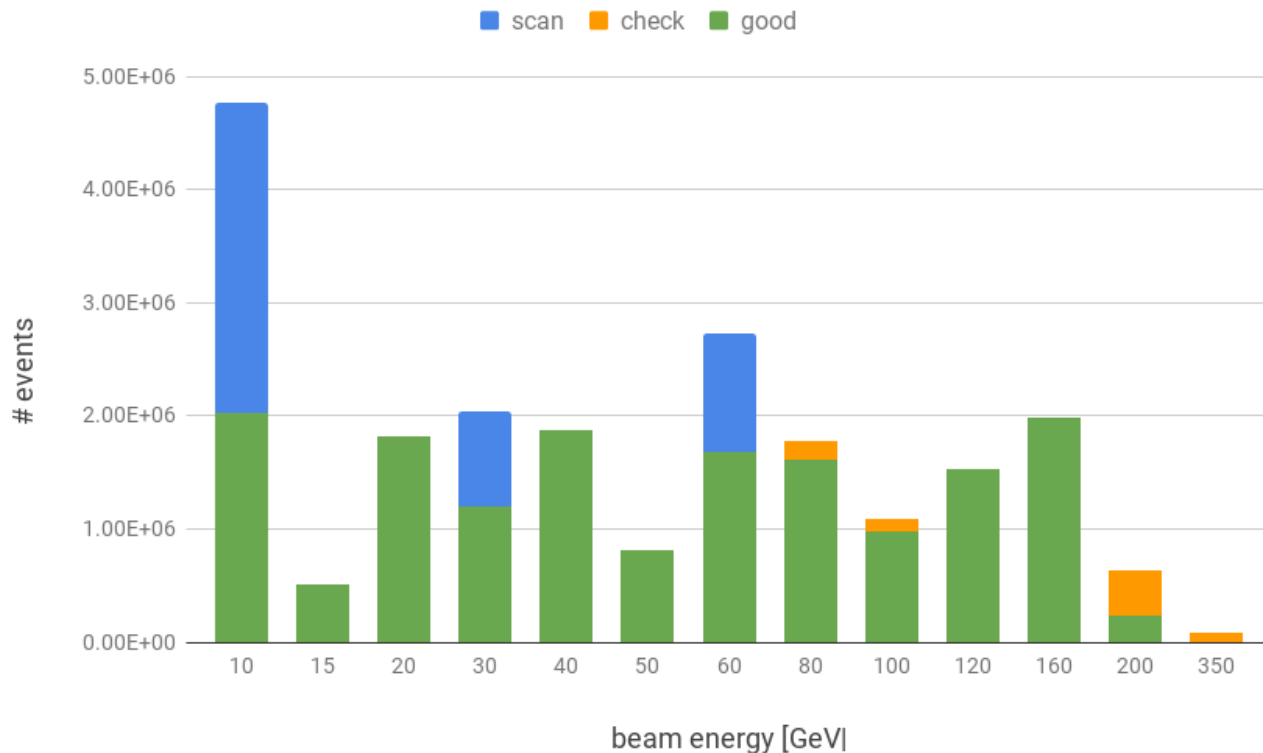
Quality Flag overview (count of runs for each flag)

Energy	May		June		Energy	May		June	
[GeV]	Good	Check	Good	Check	[GeV]	Good	Check	Good	Check
10	28	-	5	27 (scan)	80	14	1	5	-
15	18	-	-	-	100	11	1	-	-
20	15	-	5	-	120	10	-	5	-
30	11	-	5	16 (scan)	160	11	-	12	-
40	11	-	5	-	200	-	-	5	4
50	7	-	-	-	350	-	-	-	4
60	12	-	5	18 (scan)	Total	148	2	52	69

Quality Flag overview (total events in runs for each flag)

Energy	May		June		Energy	May		June	
[GeV]	Good	Check	Good	Check	[GeV]	Good	Check	Good	Check
10	1.5E+06	-	0.5E06	2.7E06	80	1.1E06	0.16E06	0.5E06	-
15	0.5E06	-	-	-	100	1.0E06	0.1E06	-	-
20	1.3E06	-	0.5E06	-	120	1.0E06	-	0.5E06	-
30	1.0E06	-	0.2E06	0.8E06	160	0.9E06	-	1.0E06	-
40	1.4E06	-	0.5E06	-	200	-	-	0.2E06	0.4E06
50	0.8E06	-	-	-	350	-	-	-	0.08E06
60	1.1E06	-	0.4E06	1.0E06	Total	11.6E06	0.26E06	4.3E06	5.0E06

Quality Flag overview (total events in runs for each flag)



Features noticed during check

May data:

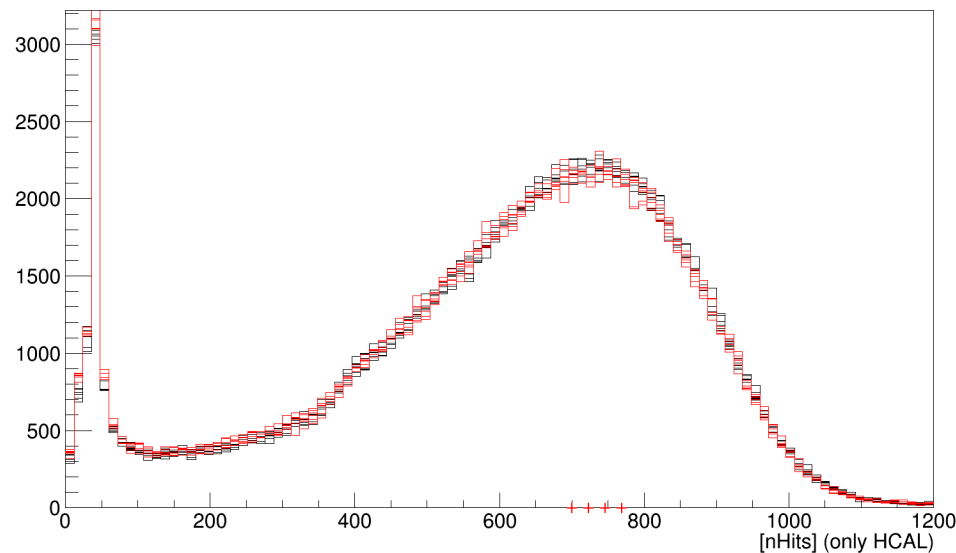
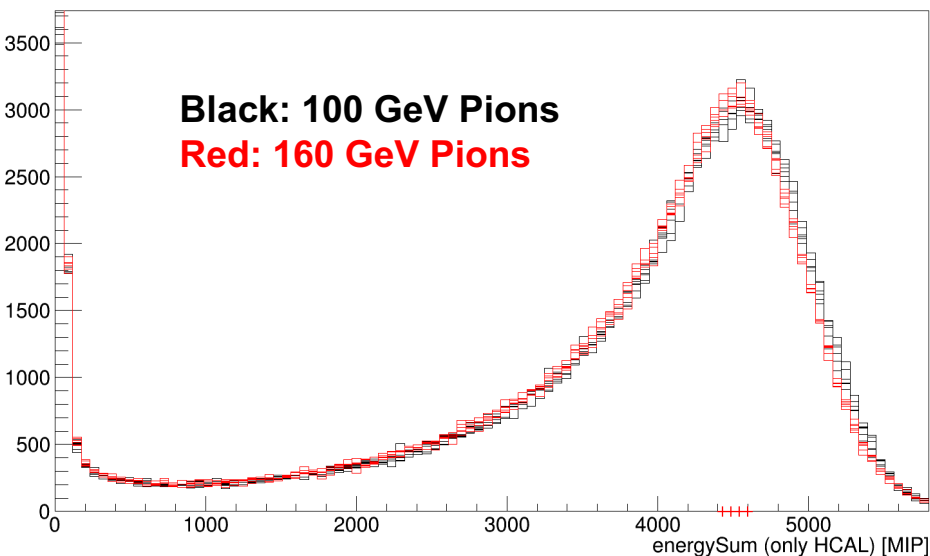
- 100 GeV with contamination due to open collimator
- 10 & 15 GeV show large electron contamination (was known already during testbeam and we'll still flag all runs as 'good')

June data:

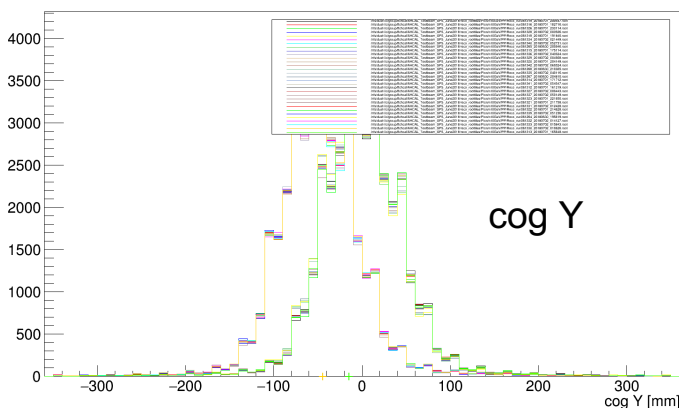
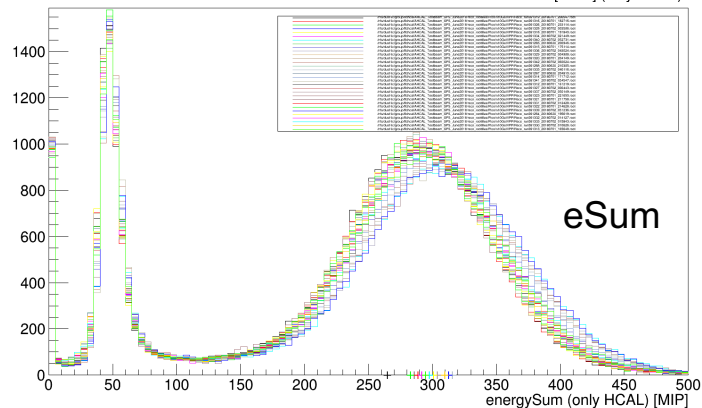
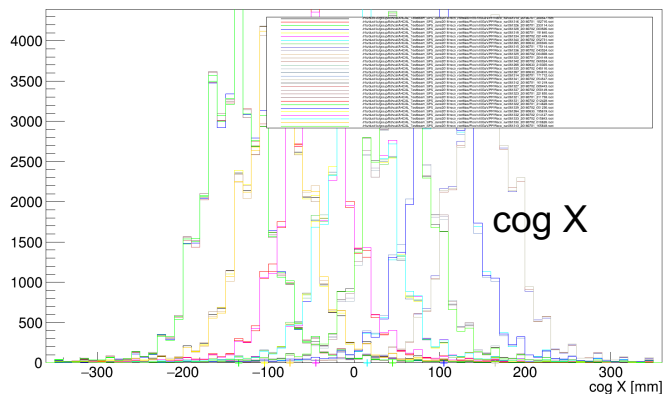
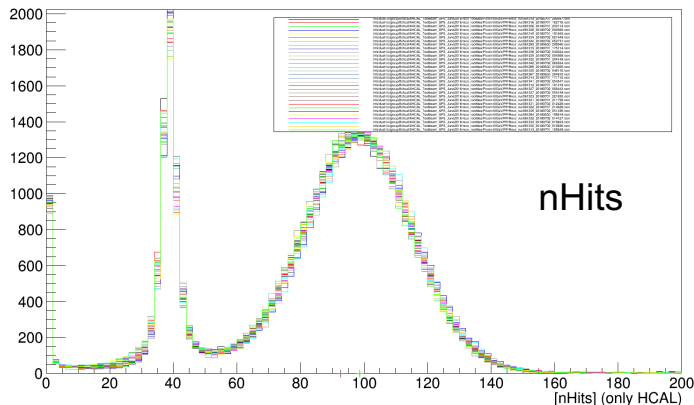
- Runs marked as '100 GeV' were actually taken with 160 GeV beam
- 10, 30 & 60 GeV runs were scan runs and need to be checked again as the calibrations constants are improving (currently flagged as 'good; SCAN', except center position is flagged as 'good')

June: "100" and 160 GeV runs

- CESAR crashed and did not load 100 GeV beamfile correct
- 100 GeV beam was actually a 160 GeV beam

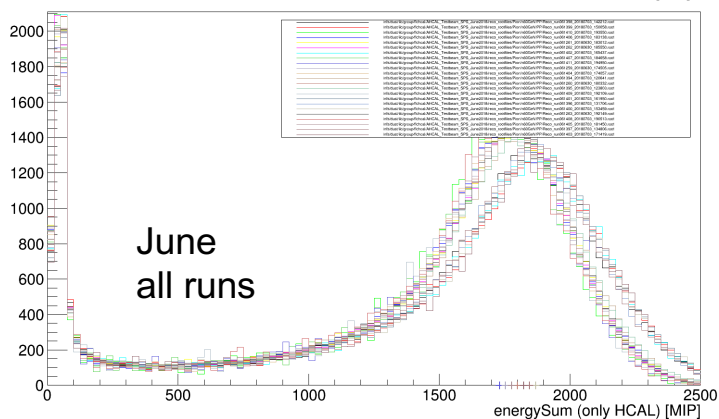
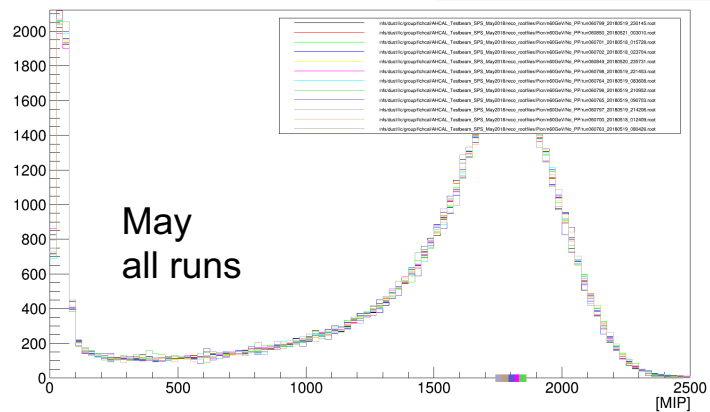


Pion Scan Runs (i.e. June, 10GeV)

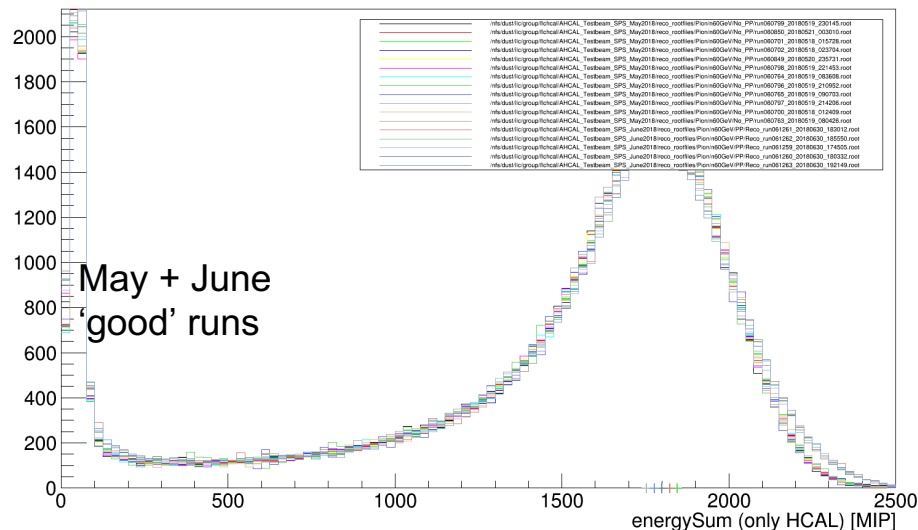


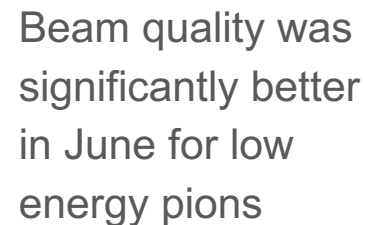
- Scan runs taken for 10, 30 & 60 GeV in June
- Calibration constants will be improved soon
- Need to recheck the eSum distribution
- For now only center position marked as 'good'

60 GeV Scan Runs (eSum)

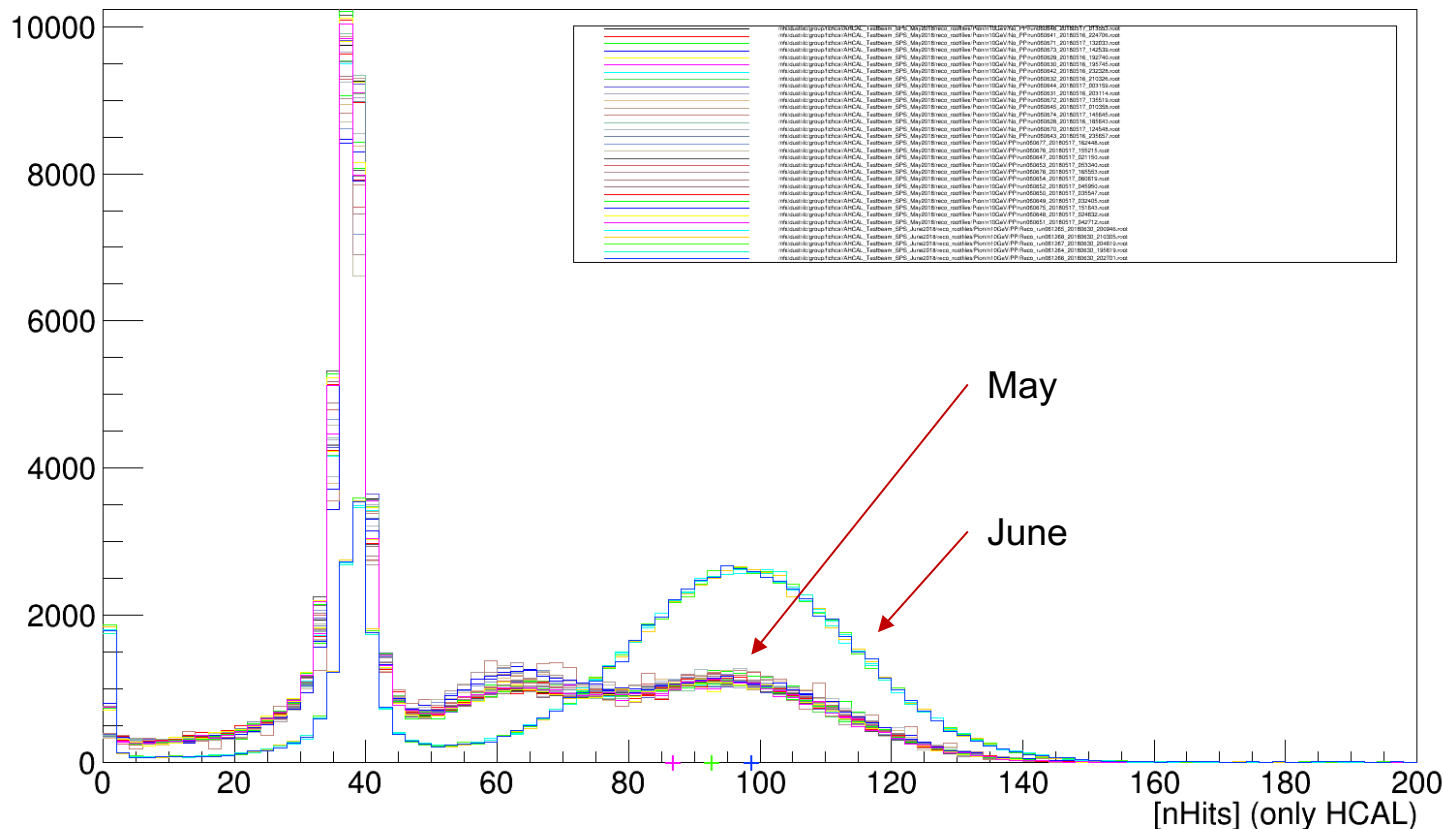


- Energies with scan runs the center position was flagged as 'good'
- The shifted positions as 'good; SCAN; to be checked'
- Comparing center position May & June runs:





Pion 10 GeV Runs (May + June 'good' runs)



Beam quality was significantly better in June for low energy pions

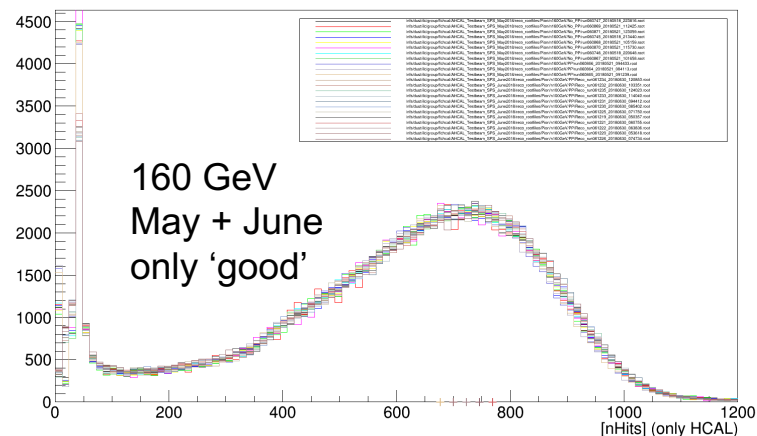


Updated Run List

A	B	C	D	E	F	G	H	I	J
Date	RunType	RunNuml	ParticleType	BeamEnergyGeV	ROCs	Events	PowerP1	Comments	OfflineQualityCheck
2018-05-21	standard	60854	pi-	80	8013	82400	no	2 spills, SC: 36.0s	good
2018-05-21	standard	60855	pi-	80	2310	27324	no	2 spills, SC: 36.0s	good
2018-05-21	standard	60856	pi-	80	12632	150916	yes	2 spills, SC: 36.0s	good
2018-05-21	standard	60857	pi-	80	13895	166725	yes	2 spills, SC: 36.0s	check, shifted eSum distribution
2018-05-21	standard	60858	pi-	120	9793	110542	yes	2 spills, SC: 36.0s	good

For analysis:

- You now can filter for
 - Runtype = 'standard'
 - OfflineQualityCheck = 'good'
 - + ParticleType, BeamEnergyGeV, PowerPulsing, ...



Summary & Outlook

- Overall the data quality for pions is very good (only few 'check' runs, no 'bad' runs)
 - The shift crews did a fantastic job (quasi-)online monitoring the run quality
- Check-runs need to be checked
 - Best once new calibrations constants in the database
 - Difference between PP and no_PP needs to be understood and corrected for

Summary & Outlook

- Overall the data quality for pions is very good (only few 'check' runs, no 'bad' runs)
 - The shift crews did a fantastic job (quasi-)online monitoring the run quality
- Check-runs need to be checked
 - Best once new calibrations constants in the database
 - Difference between PP and no_PP needs to be understood and corrected for



Thank you for a fantastic and very productive workshop!!

Bonus slides



Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG



Bundesministerium
für Bildung
und Forschung

Friedrich Naumann
STIFTUNG

FÜR DIE FREIHEIT

First talks slides



Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG



Bundesministerium
für Bildung
und Forschung

Friedrich Naumann
STIFTUNG

FÜR DIE FREIHEIT

May & June run list

[Click here to access run list google sheet](#)

Thanks to all the shifters for filling in the list!!
(If you notice mistakes, please report and amend)

- How to use the list:
 - Filter for '*standard*' run (those runs the shifters considered as taken with correct settings)
 - Filter for particle type, beam energy, PP or no_PP
- Now: How to proceed with the list?
 - .tsv in stash?
 - Table in Confluence?
 - Other ideas? Preferably an option that makes it easy to filter the list

Pion data quality check

Checking all pion ‘standard runs’ (according to run list) for outliers

Creating lists:

- “Good” runs: ?
- “Need-to-check” runs: ?

All plots only for May data so far

→ current reco files: `/nfs/dust/ilc/group/flchcal/AHCAL_Testbeam_SPS_May2018/reco_rootfiles/`

June plots did not make sense yet as calibrations constants are off for tail catcher and Tokyo Module

Pion data quality check

Checking all pion 'standard runs' (according to run list) for outliers

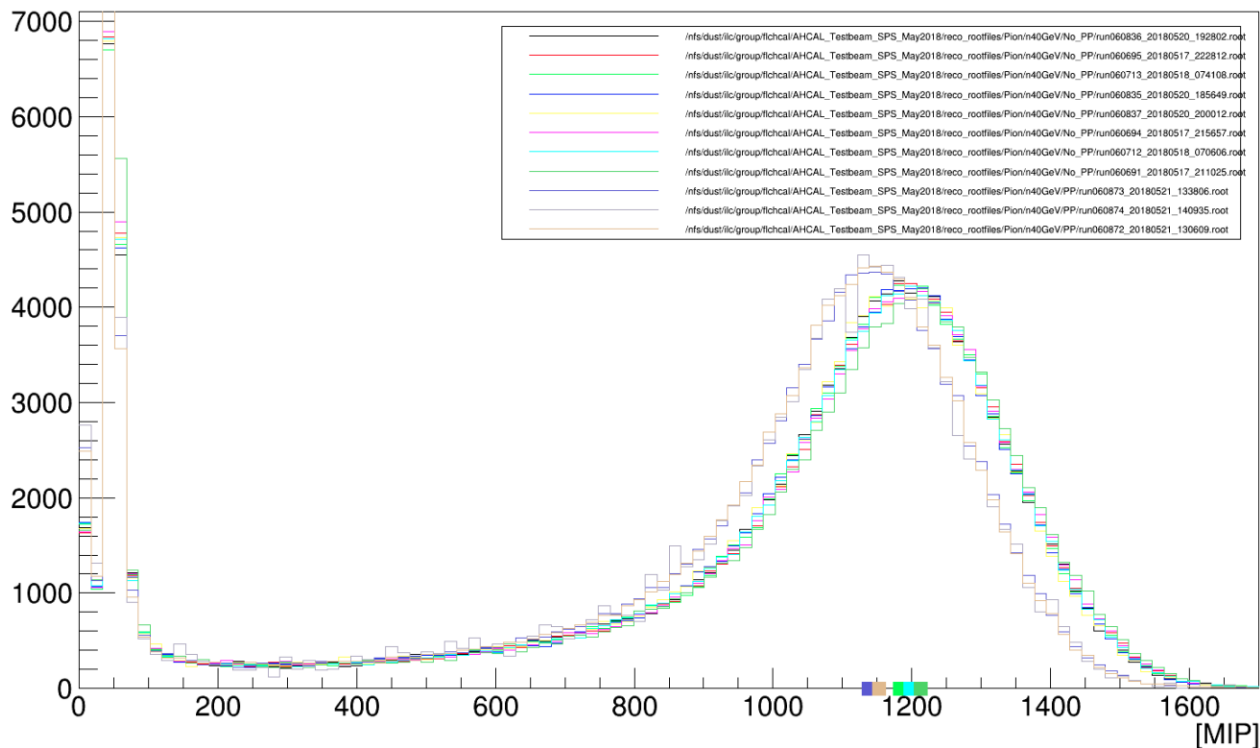
Looked at:

- Energy Sum
- nHits
- Center of Gravity in X & Y
- Ratio nPions vs nMuons based on energy cut @ 200 MIP

Root macros to create all plots can be found here:

`/nfs/dust/ilc/user/buhmae/tokyoWorkshop/macros_PionQuality/`

Energy sum histograms 40 GeV



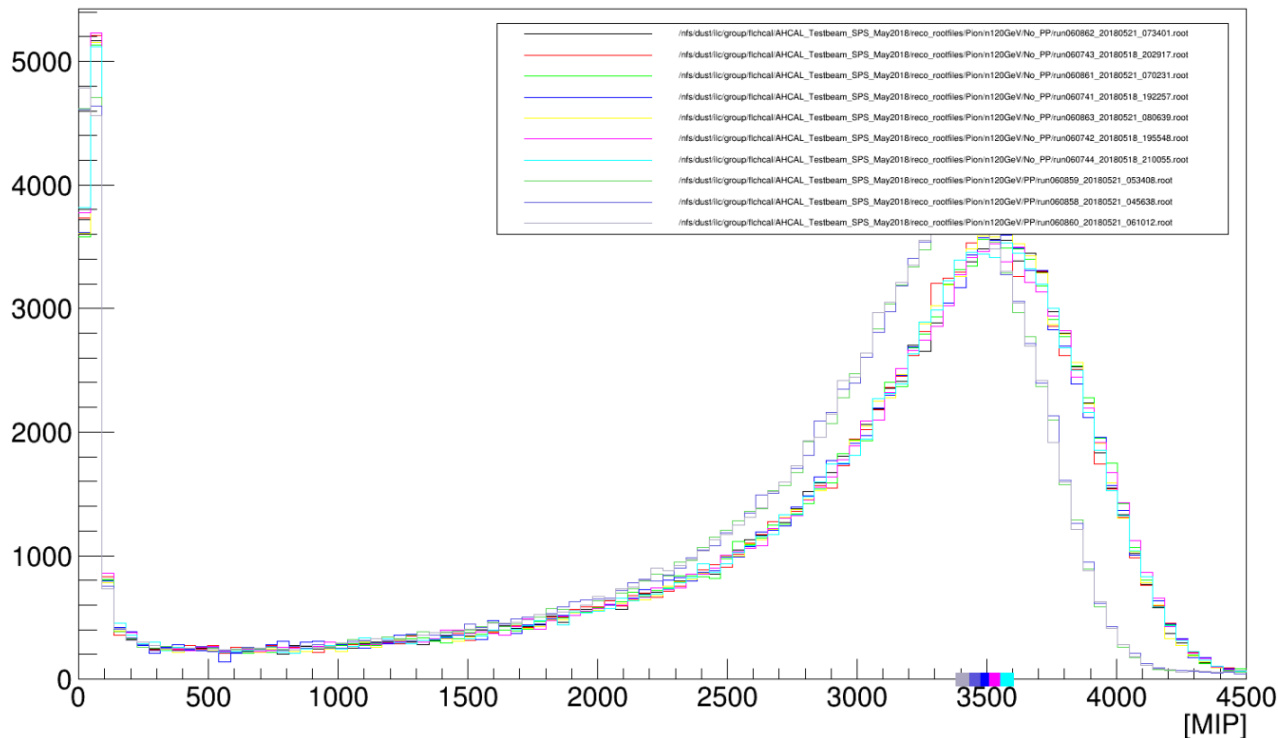
Normed histograms of energy sums of each 40 GeV run

Binning = 100
(for all histograms)

Marked bin with peak position for comparison

Systematic difference between PP & No_PP ?

Energy sum histograms 120 GeV

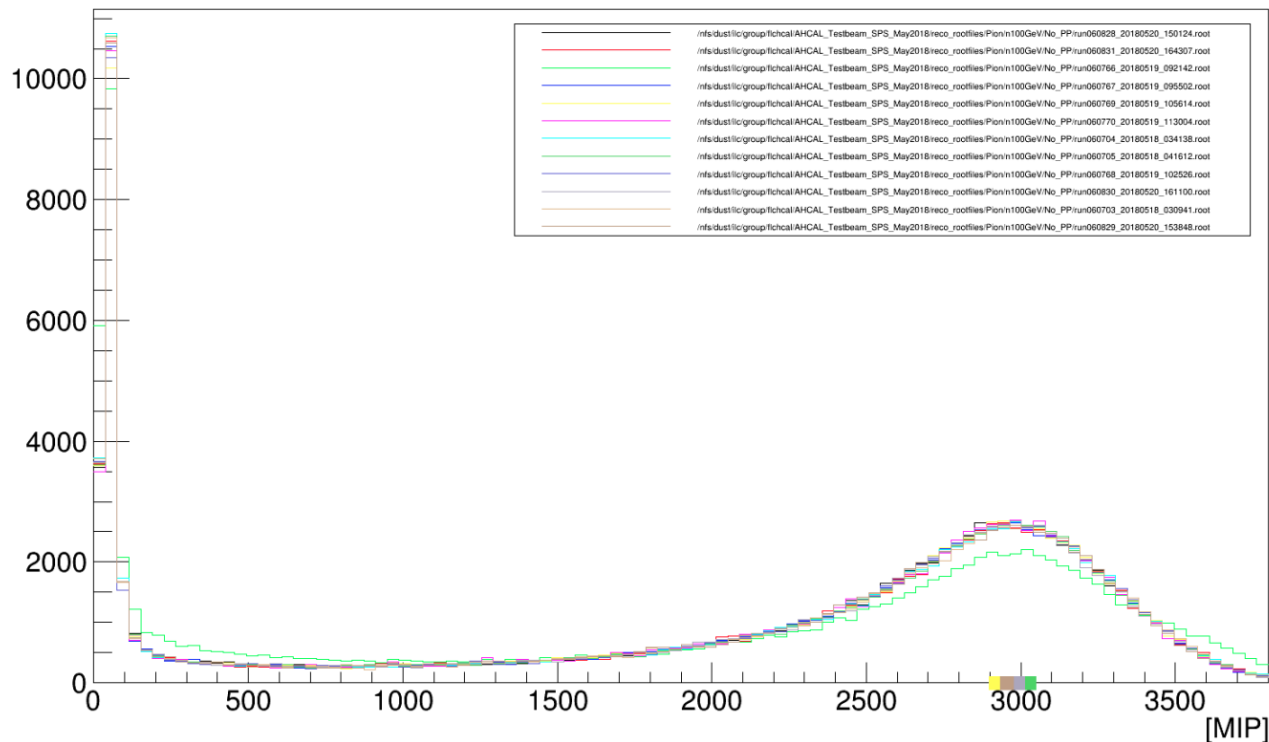


Similar distributions for
120 GeV and 160 GeV

Distribution shift
between PP and
No_PP mode

Energy sum histograms

100 GeV

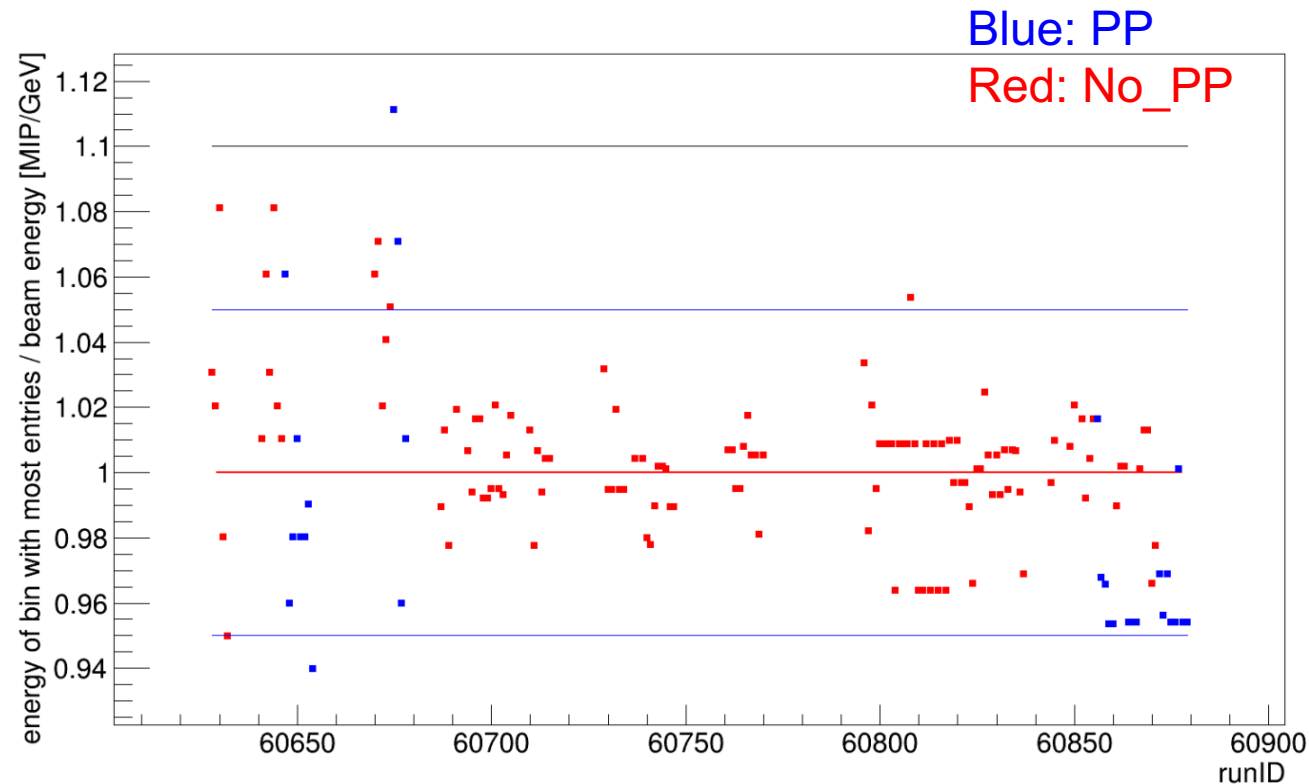


Clear outlier:

Run 60766

(missing absorber according to eLog)

E_sum peaks vs Run ID

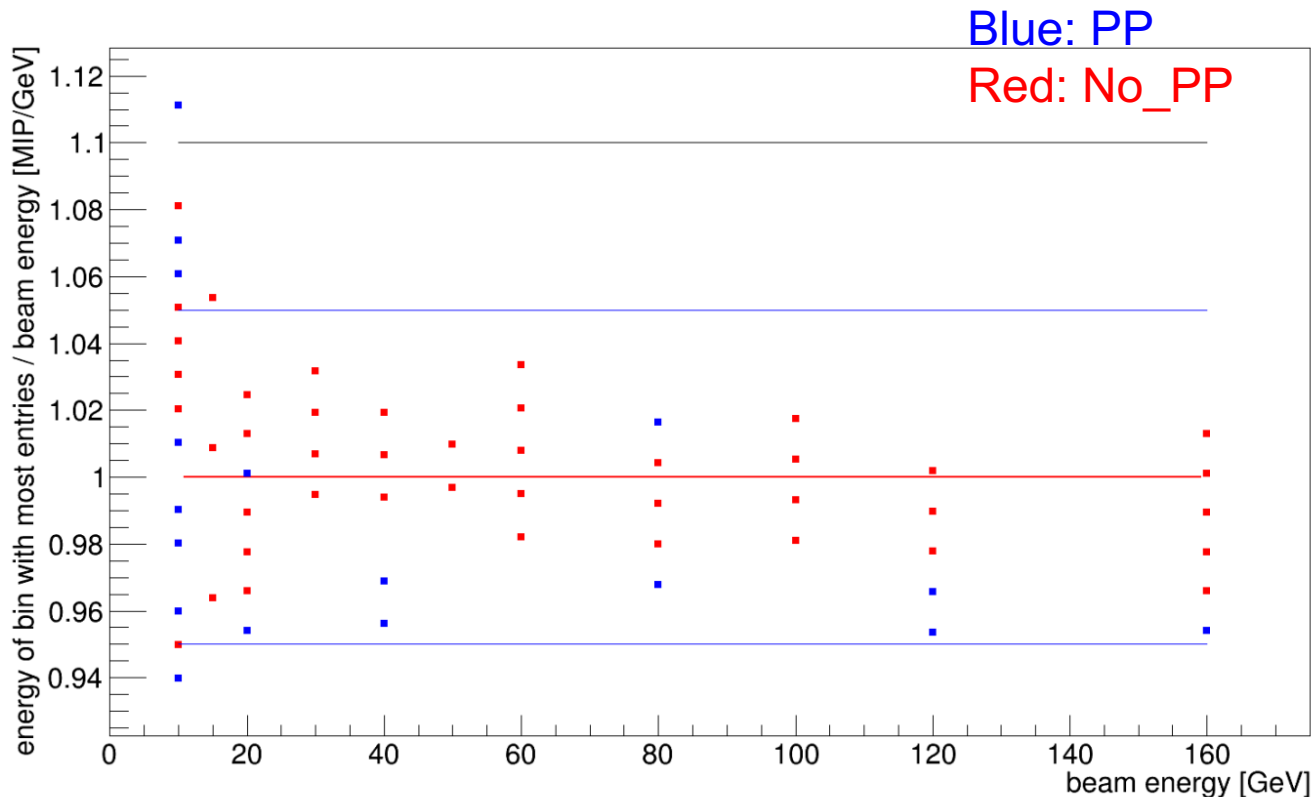


E_sum peak / beam energy
→ Detektor response
surprisingly linear (!)

Checking time dependence
with Run ID

Few outlier off > 5%,
just one off > 10%

E_sum peaks vs beam energy



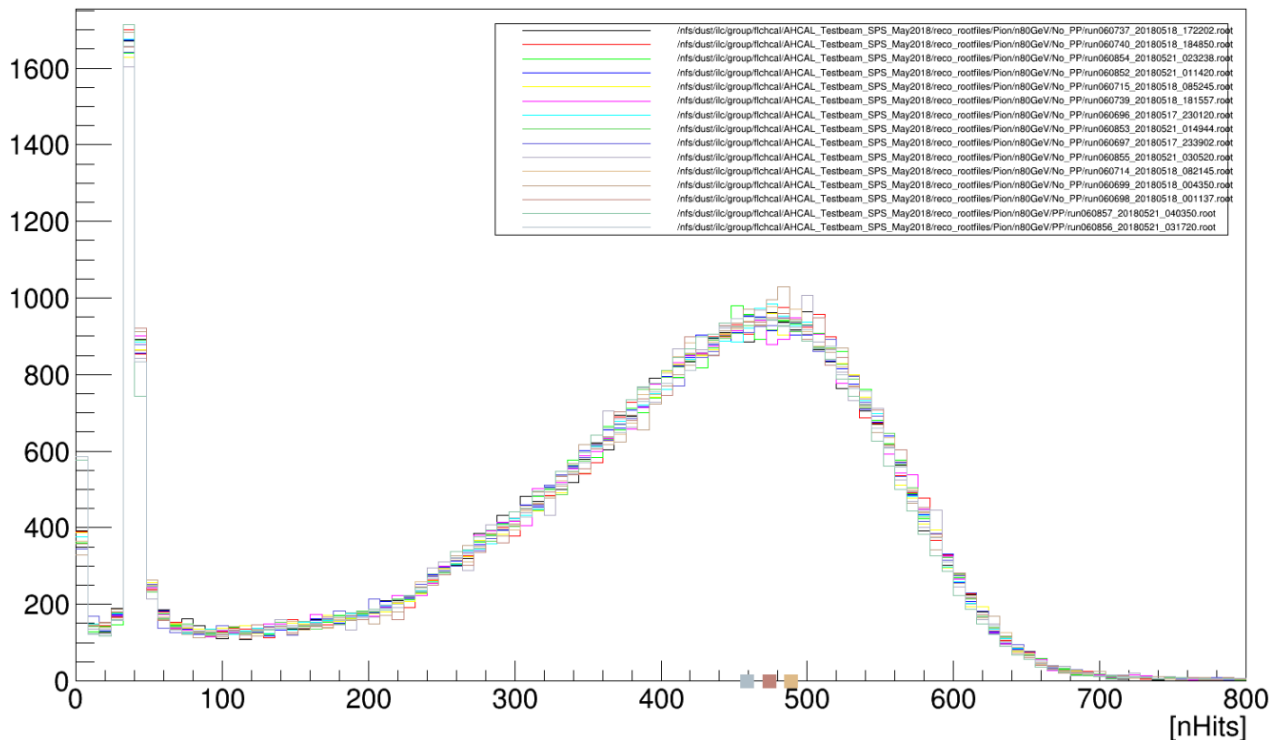
Same representation, now sorted by beam energy

(!) markers are overlaid, for full story look at this plot and the former one

All peak position within 5% (except 10 GeV)

Systematic peak shift PP vs no_PP? At high energies? (needs to be investigated)

nHits histograms 80 GeV



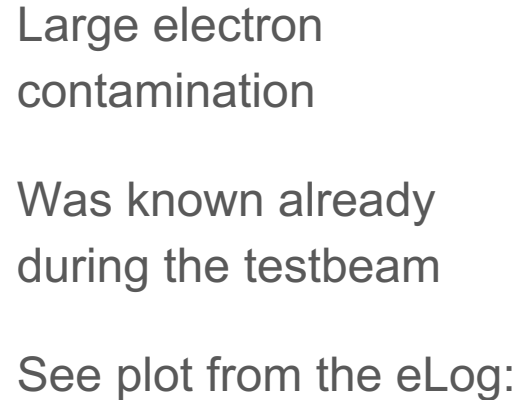
Normalized histograms of all 80 GeV runs

Marked bin with peak position for comparison

All energies look similar (except 10 GeV and 100 GeV)

No systematic difference between PP and No_PP visible

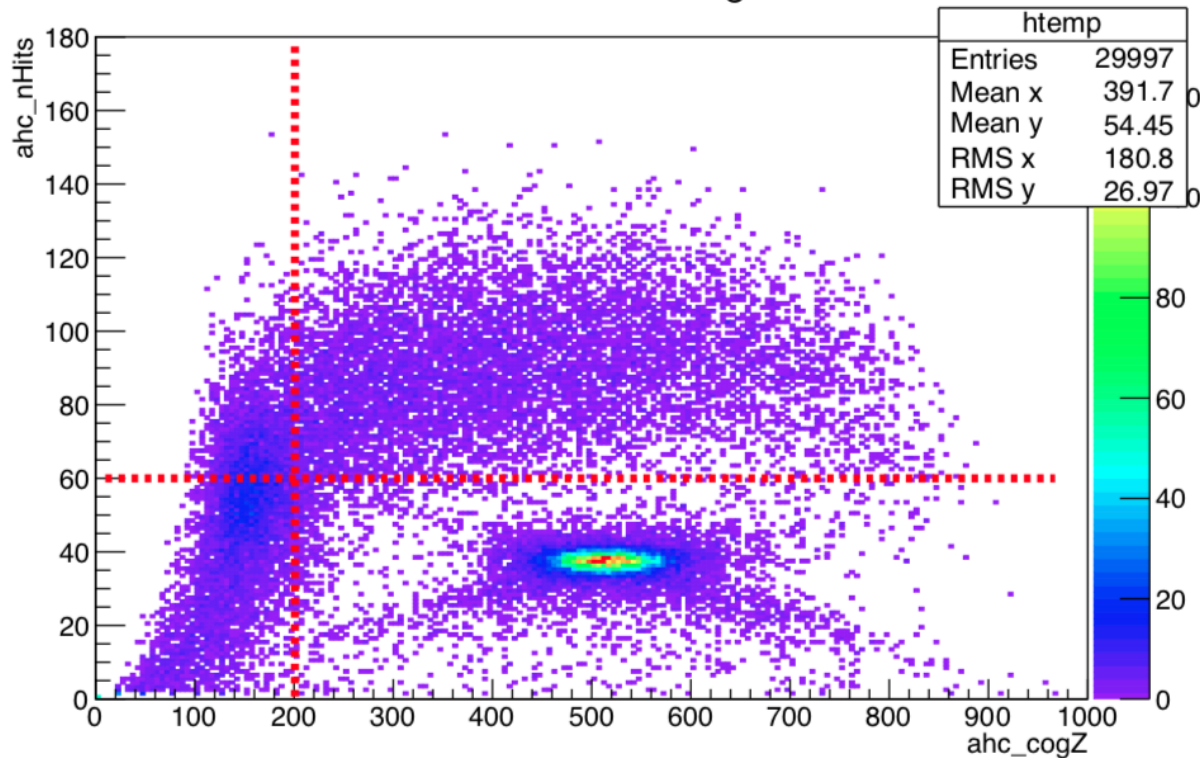
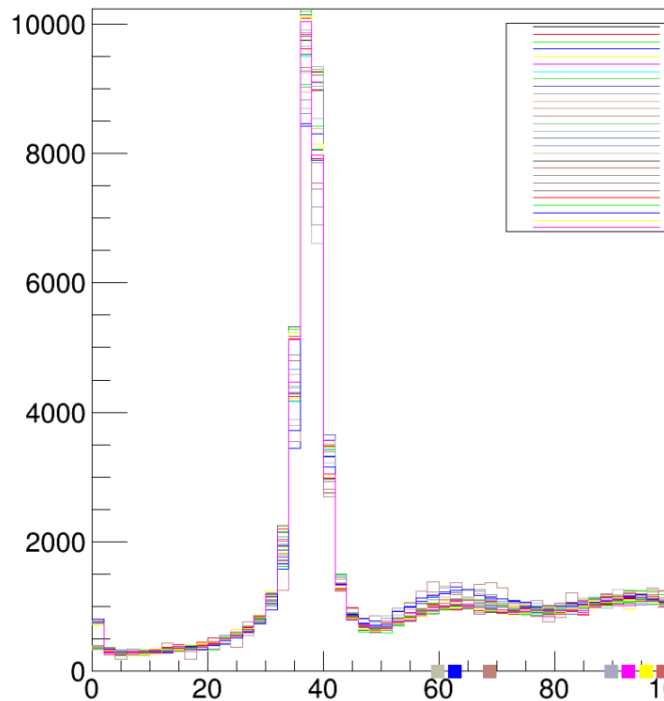
Binning = 100 (for peak bin with 1000 entries)



nHits histograms

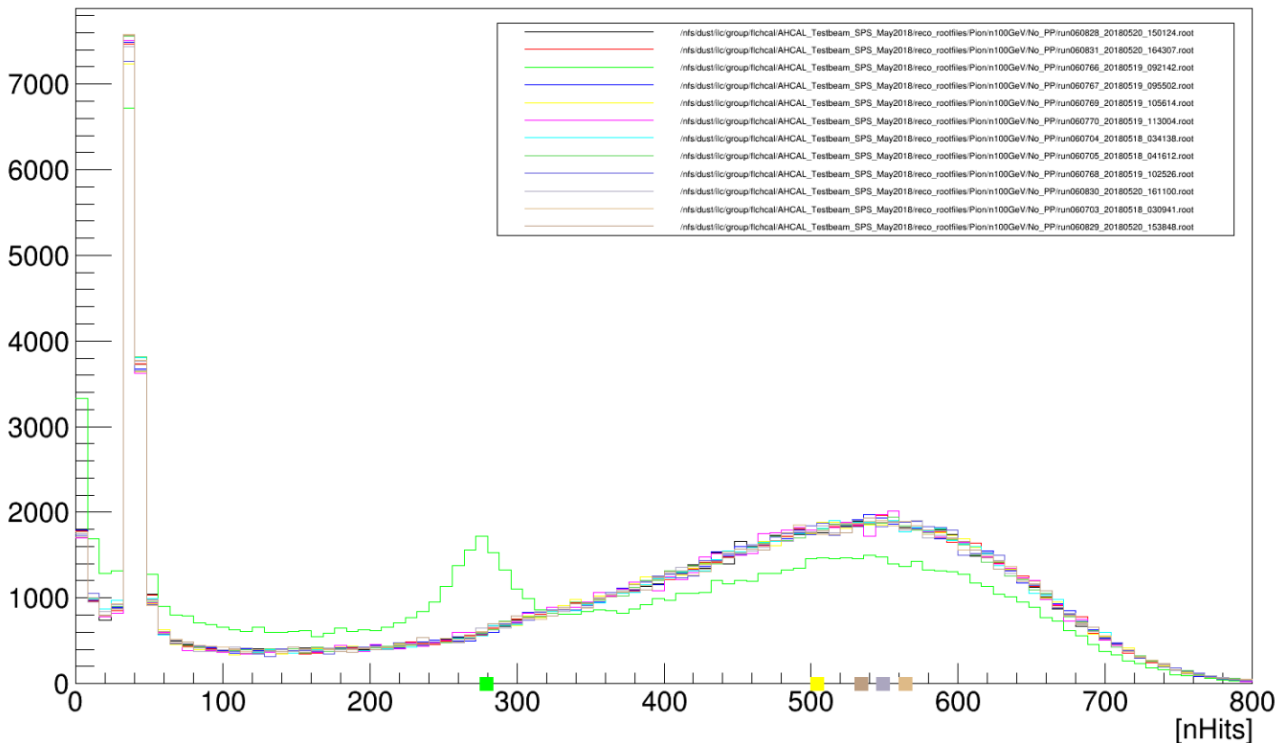
10 GeV

ahc_nHits:ahc_cogZ



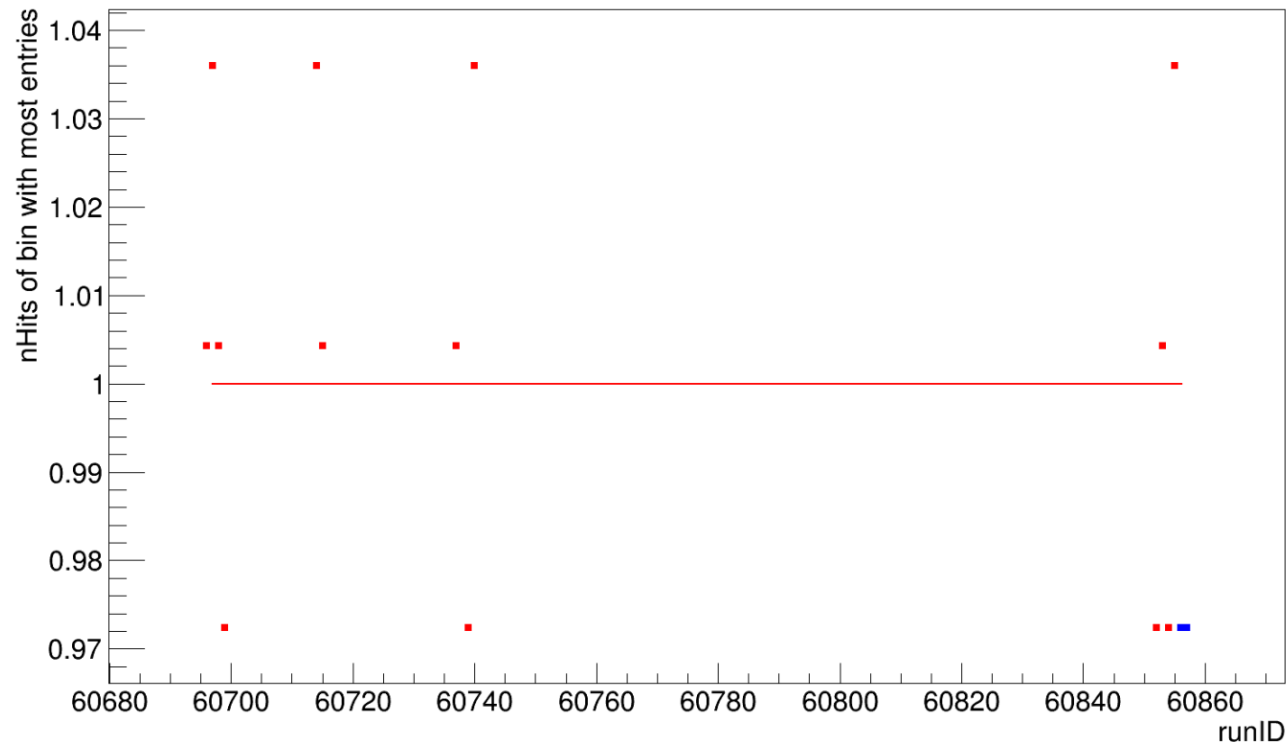
nHits histograms

100 GeV



Already noticed this run in the energy sum histograms

nHit peak bin vs run ID 80 GeV

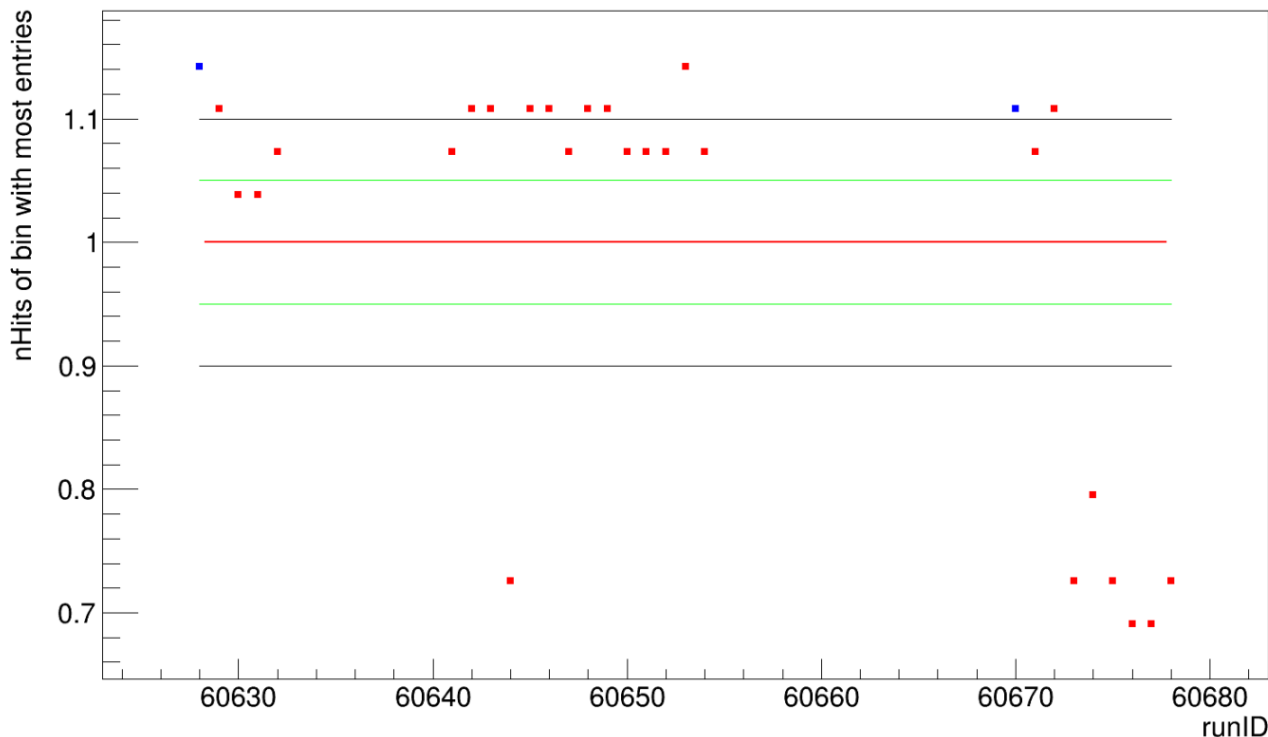


Distribution of peak bin positions for runs of single energy

Most peaks within 5%,
all peaks within 10%

(except 10 GeV and
the one 100 GeV run)

nHit peak bin vs run ID 10 GeV

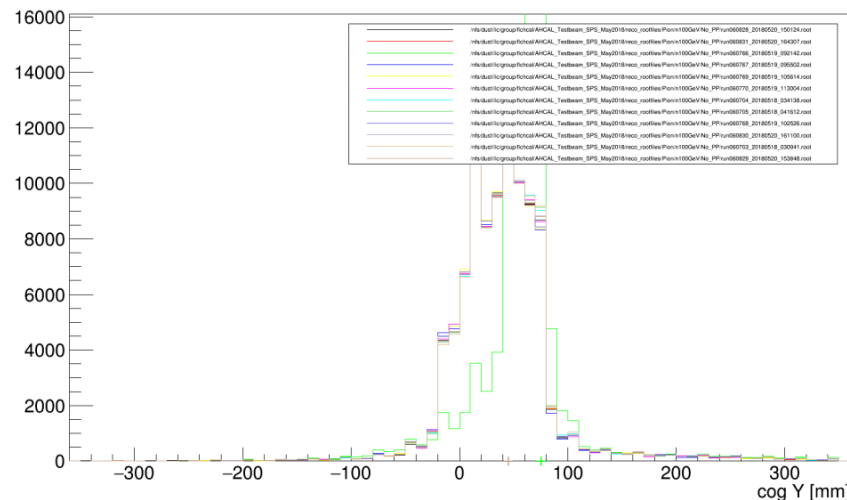
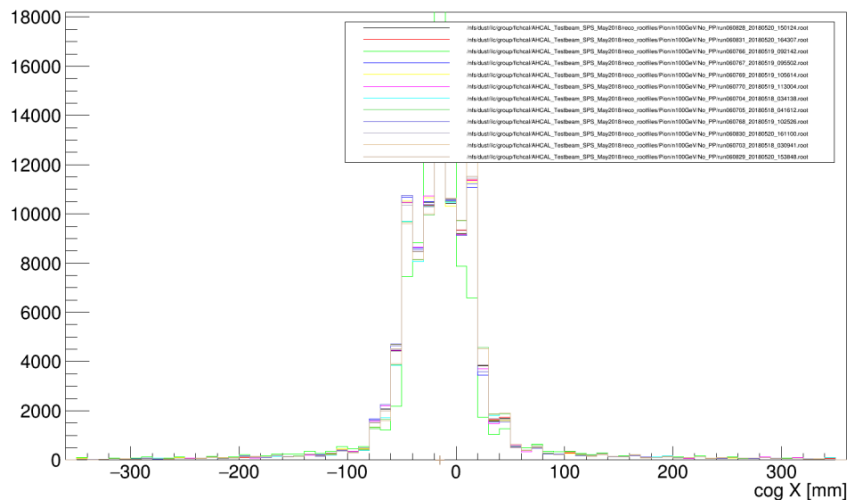


We see here the wide spread due to the electron contamination

→ Quality criteria should take exception for our 10 GeV pions into account

Center of Gravity Plots

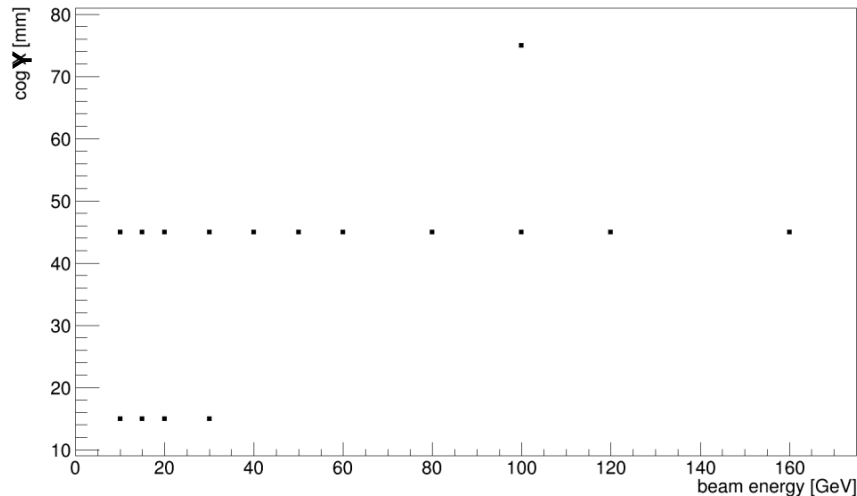
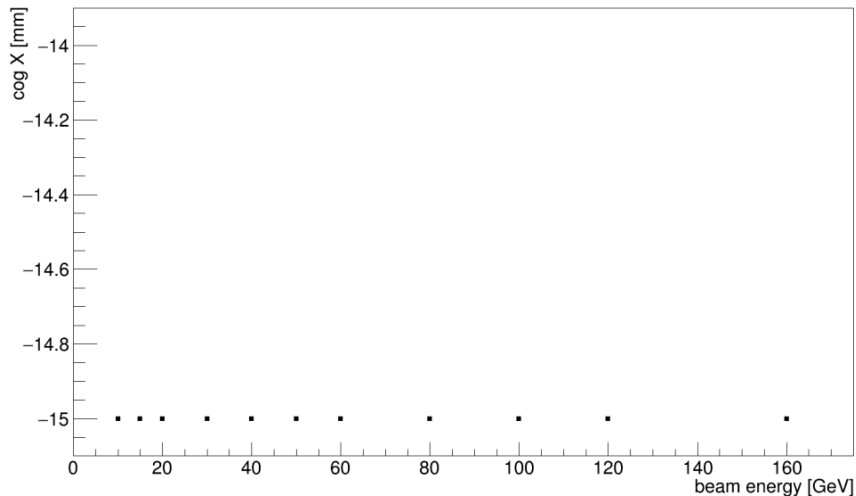
X-axis



Beam was well centered for all runs
 (Y-axis outliers because of Muons - except the 'special' 100 GeV run) (1 cm binning)

Center of Gravity Plots

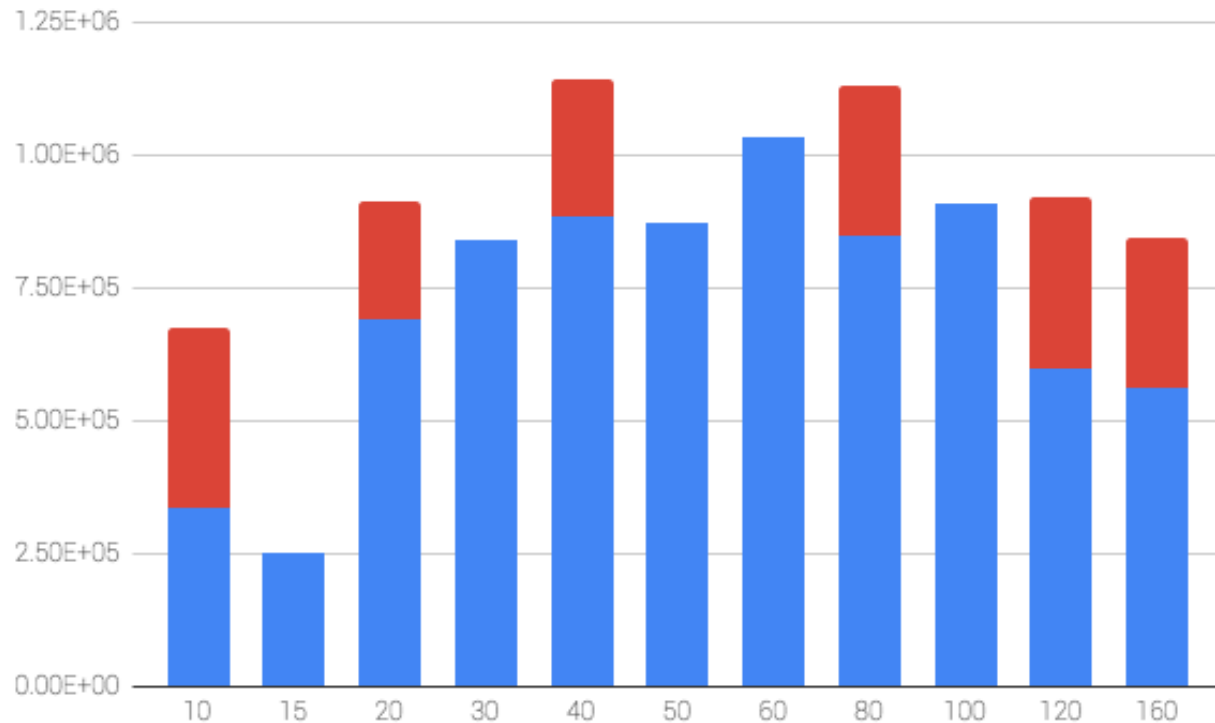
X-axis



Beam was well centered for all runs

(Y-axis outliers because of Muons - except the 'special' 100 GeV run) (1 cm binning)

Pion Candidates (cut: $e\text{Sum} > 200 \text{ MIP}$)



Blue: No_PP

Red: PP

Based on very simple
energy cut

Expect less pions for
analysis

→ Vladimir's Particle ID

+ June Pion data
(only with PP)

'Pion' / 'Muon' Ratio



Blue: No_PP

Red: PP

Do we understand this behavior?

100 GeV raises concern:
 Checked used beam file
 → XCHV.021.133 wide open!

Tail at low energy in the
 E_{sum} in comparison to
 other 60 / 60 / 120 GeV

Summary: “Good run” criteria

Proposal:

- eSum peak bin within 5 % of each other (same energy)
 - Except 10 GeV, \bc of large e- contamination
- nHit peak bin within 5 % of each other (same energy)

Open for discussion!

Afterwards:

- New column in run list with flag ‘good run’ or flag ‘need-to-check run’
- Document criterias for ‘good run’ on confluence

Outlook

- Fix criteria for ‘good runs’
 - Document on confluence
- Add quality flag in run list
- ‘Need-to-check’ runs need to be investigated
 - + Comparison between PP and No_PP mode
- Quality check for June pion data
 - Box-and-whisker plots might be helpful
- Quality check for electron data (& muon data?)
- Move run list away from Google sheet

- Fix criteria for ‘good runs’
 - Document on confluence
- Add quality flag in run list
- ‘Need-to-check’ runs need to be investigated
 - + Comparison between PP and No_PP mode
- Quality check for June pion data
 - Box-and-whisker plots might be helpful
- Quality check for electron data (& muon data?)
- Move run list away from Google sheet

Thank you!

Bonus slides



Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG



Bundesministerium
für Bildung
und Forschung

Friedrich Naumann
STIFTUNG

FÜR DIE FREIHEIT

All No_PP runs in May Noticeable tail at low energies for 100 GeV

