

AutoDQM for DTs

Muon DPGO Meeting

April 12th 2024

**Andrew Brinkerhoff
and Caio Daumann**
for the AutoDQM team



BAYLOR



Recap: 2024 plan of action

- **1)** Decide on final set of histograms for each detector “subsystem” page in AutoDQM for Muon Doc #1.
- **2)** Generate XML files with lists of histograms for each new page.
- **3)** Run “meta-study” on 2022 data to set χ^2 and max pull value anomaly thresholds with a sufficiently low flagging rate.
- **4)** Adapt shifter script(s) to open relevant pages for each run.
- **5)** Train ML anomaly-detection algorithms (PCA, auto-encoder) for all histograms using 2022 and / or 2023 data. Repeat step #3.
- Rob White (formerly Bristol, now INFN) working on fixing some issues with the ML training code base

[1] <https://indico.cern.ch/event/1362244/#128-autodqm-tool-for-the-muon>

[2] <https://indico.cern.ch/event/1353975/#2-discussion-to-implement-auto>

AutoDQM for DT

- Caio Daumann (Aachen) has been the point man for DT AutoDQM
- DT experts expressed interest in running anomaly detection on over 1000s of low-level plots - posed new challenge for AutoDQM
- Need a very low anomaly flagging rate so shifters aren't overwhelmed by dozens of anomalies in every run
- AutoDQM GUI takes several minutes to generating 1000s of pdfs

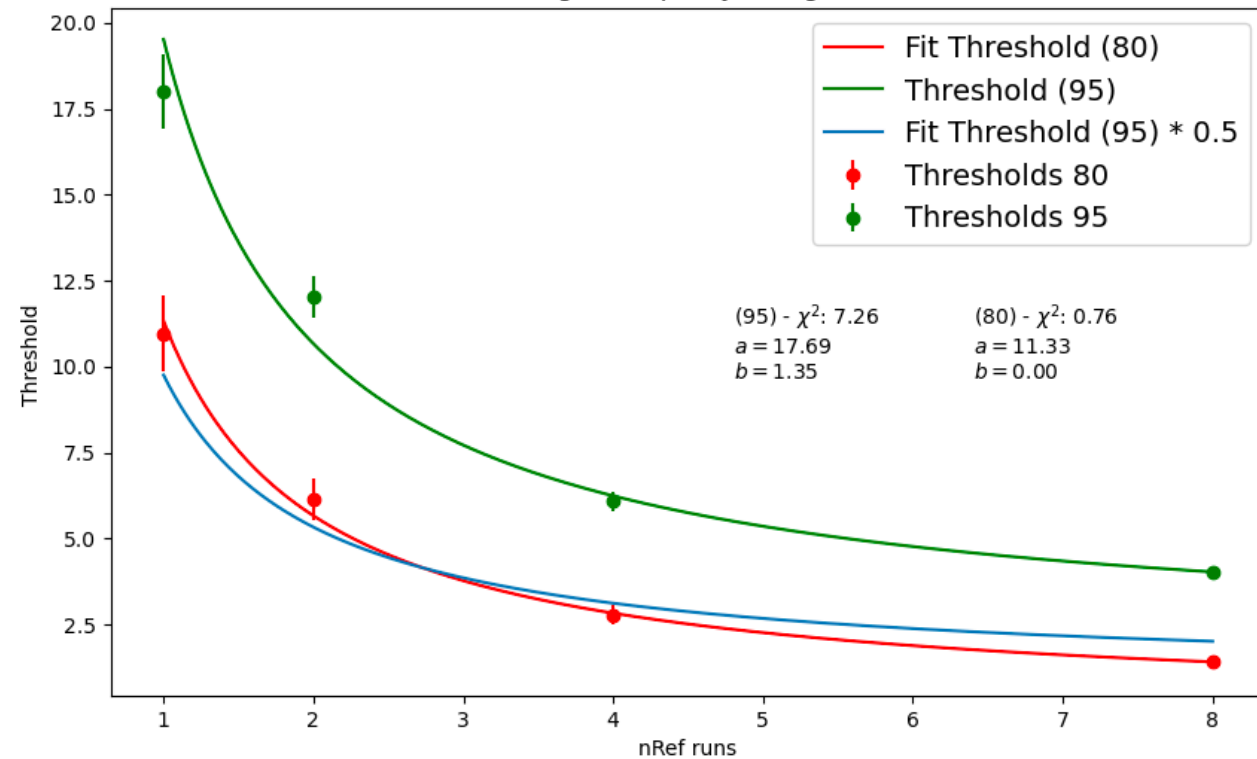
[1] https://docs.google.com/spreadsheets/d/1Bvqx-pq5qsOyRwjamzMN-_x0quZtbzxFqGzRokp6ysM

DT AutoDQM threshold studies

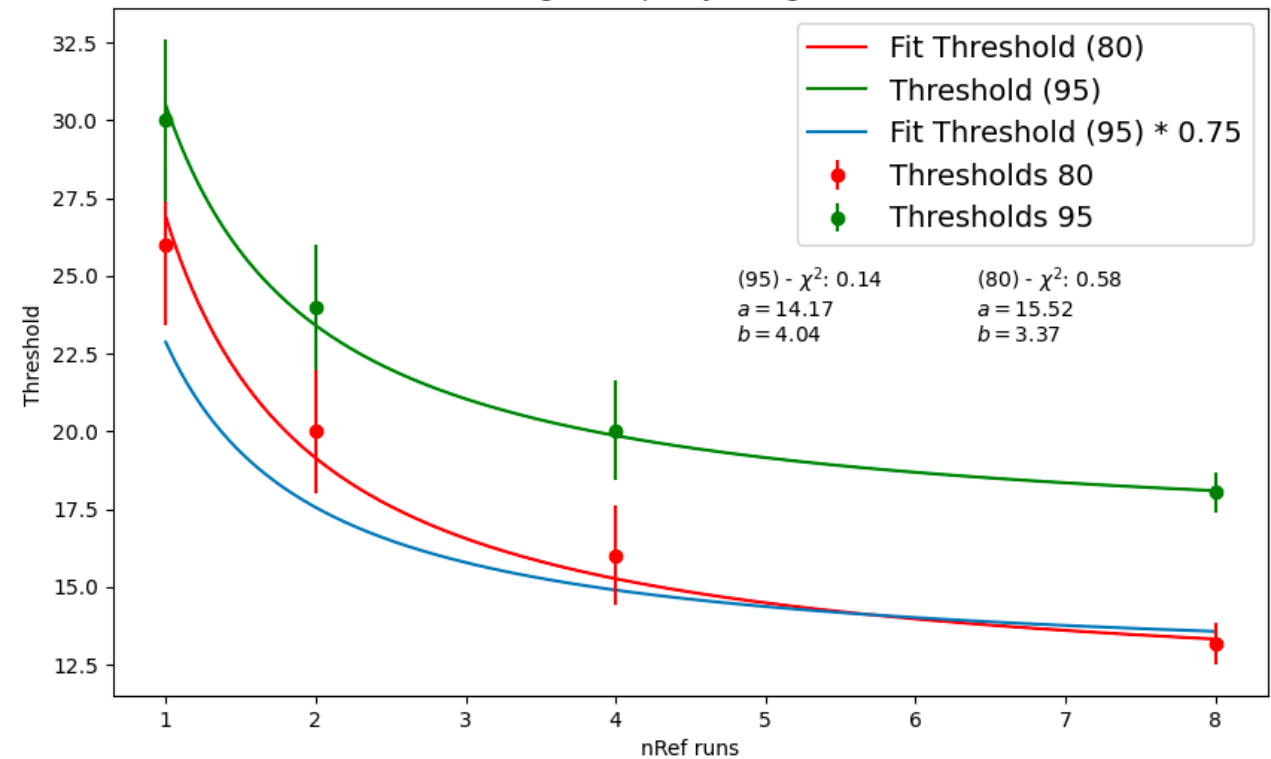
- 5 main DQM histogram classes for DT
 - **DOC 1 (Online DQM)** : Occupancy, TimeBox, and Trigger
 - **DOC 3 (Offline DQM)** : Task, Resolution
- For each class, Caio studied good runs from 2022
 - Compare each run to 1, 2, 4, or 8 previous reference runs ($nRef$)
 - Get 10 highest anomaly scores (χ^2 & max pull) for each class in each run
 - Compute separate Chi2 and maxPull thresholds for which 95% (tight) or 80% (loose) of top-10 scores from full set of runs are below threshold
 - Set anomaly flagging score to 95% threshold, and only generate plots to display for histograms failing 80% threshold (other plots discarded)

Results for the DOC1 Occupancy histograms

Curve Fitting - Occupancy histograms - Chi2



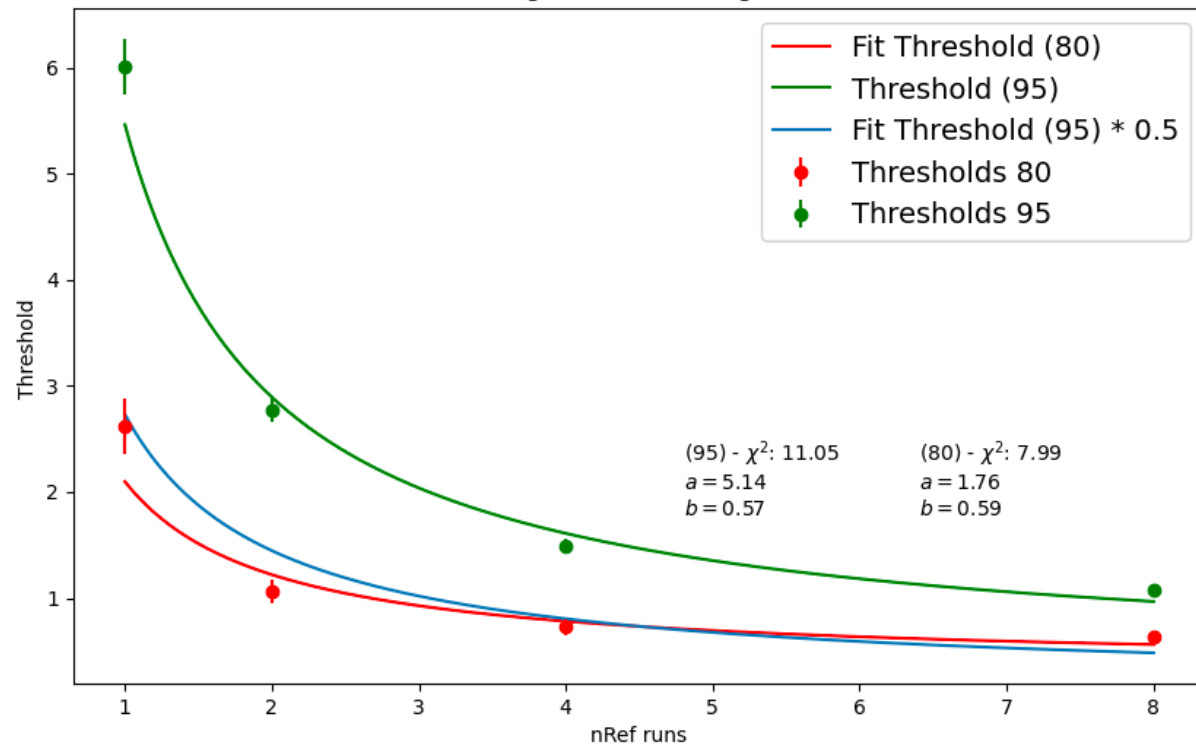
Curve Fitting - Occupancy histograms - MaxPull



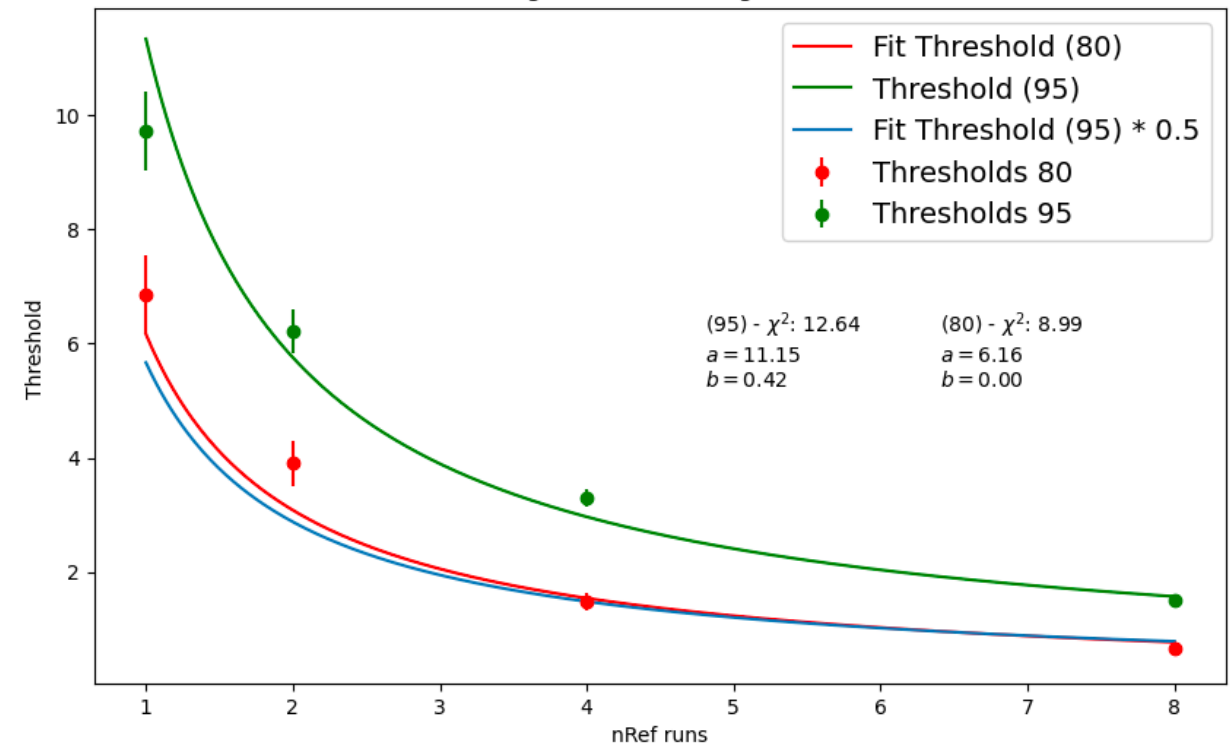
- Caio found that the thresholds can be reasonably fit by function with two free parameters: $score = a / nRef + b^2$
- Thresholds inversely related to number of reference runs

Results for the DOC1 TimeBox histograms

Curve Fitting - TimeBox histograms - Chi2



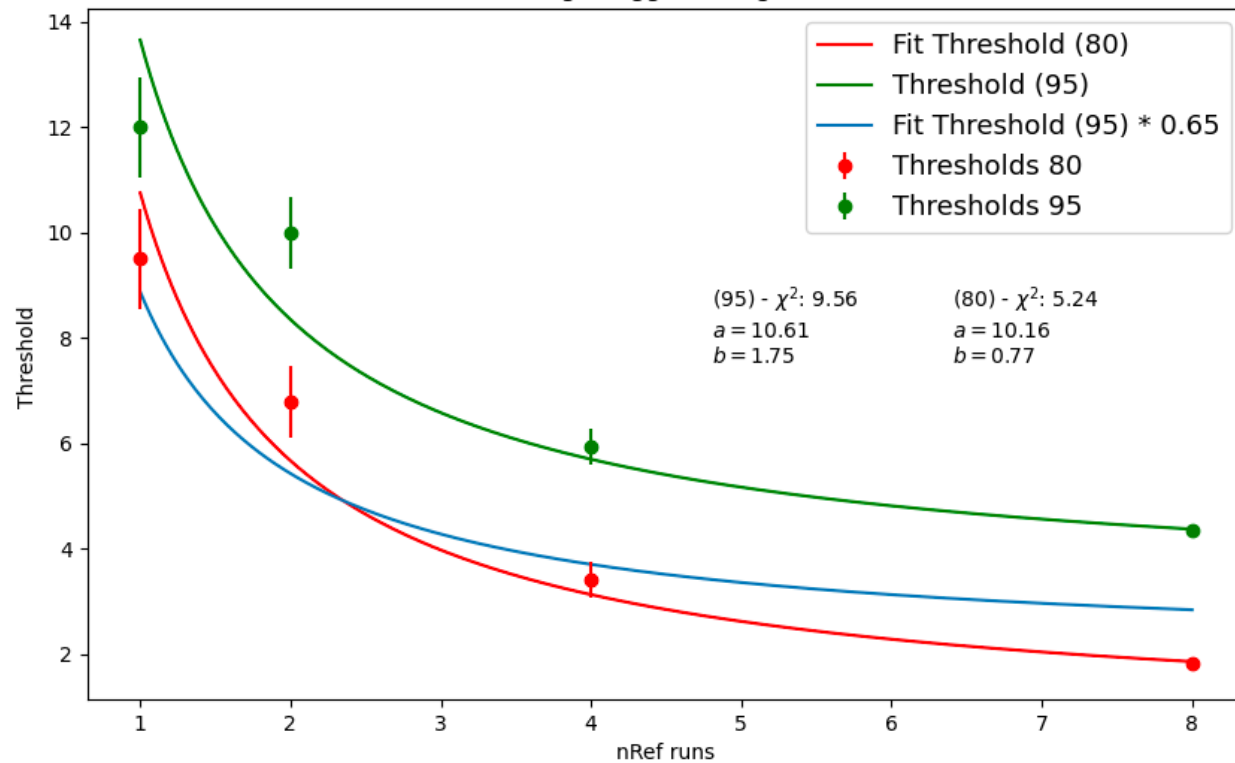
Curve Fitting - TimeBox histograms - MaxPull



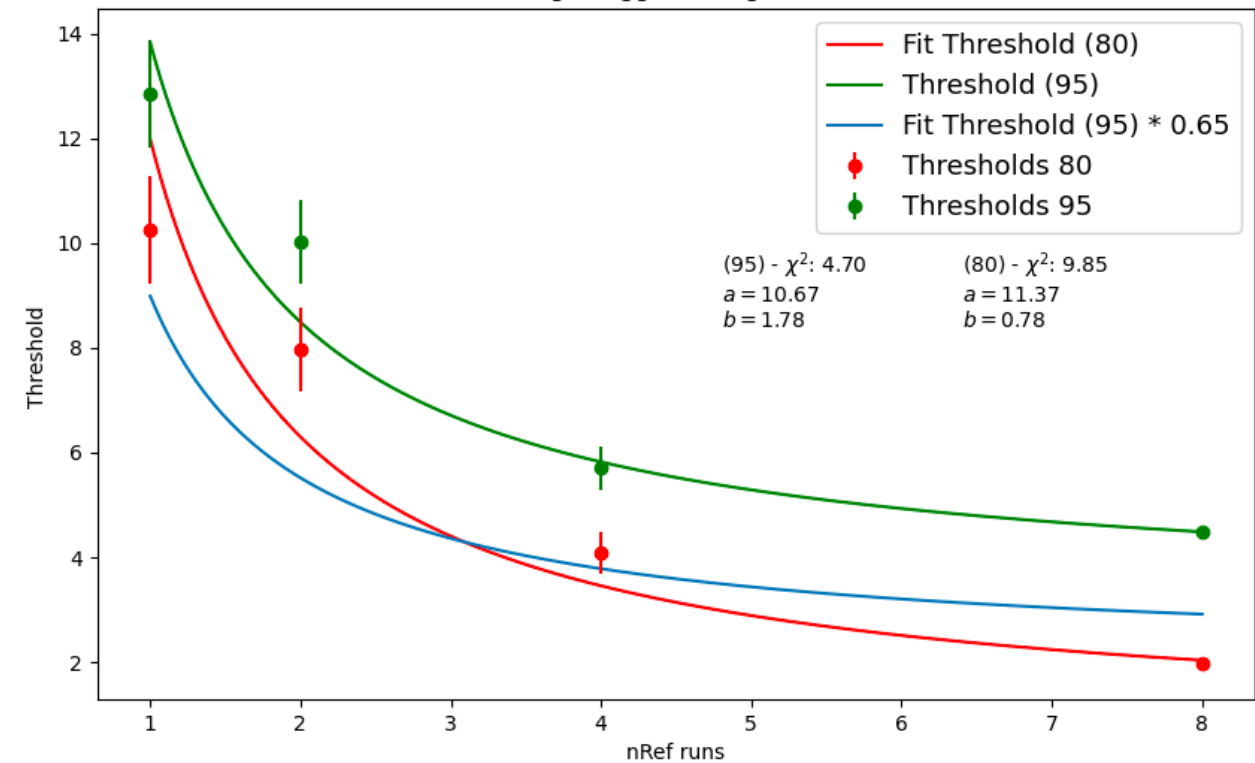
- Thresholds can be significantly different for different histogram classes, e.g. occupancy tends to have more fluctuations than time
- Also single-bin outliers (max pull) more frequent than χ^2 anomalies

Results for the DOC1 Trigger histograms

Curve Fitting - Trigger histograms - Chi2



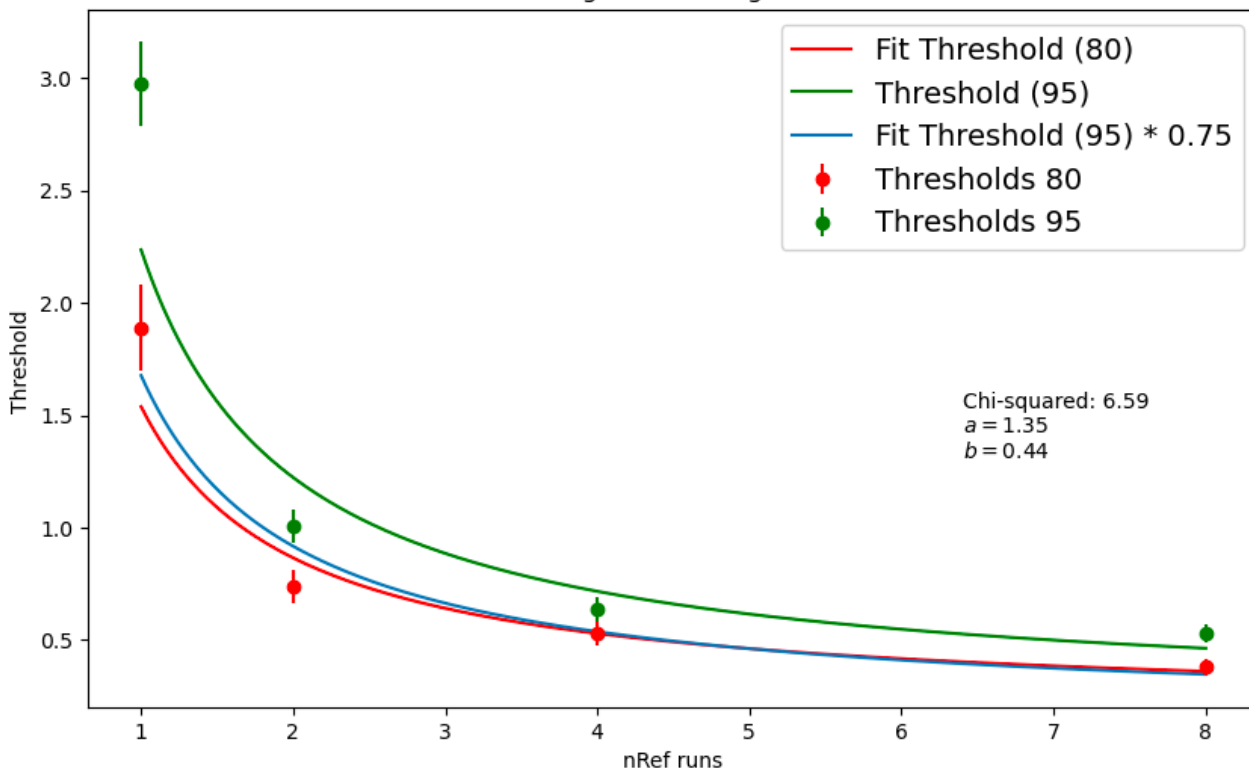
Curve Fitting - Trigger histograms - MaxPull



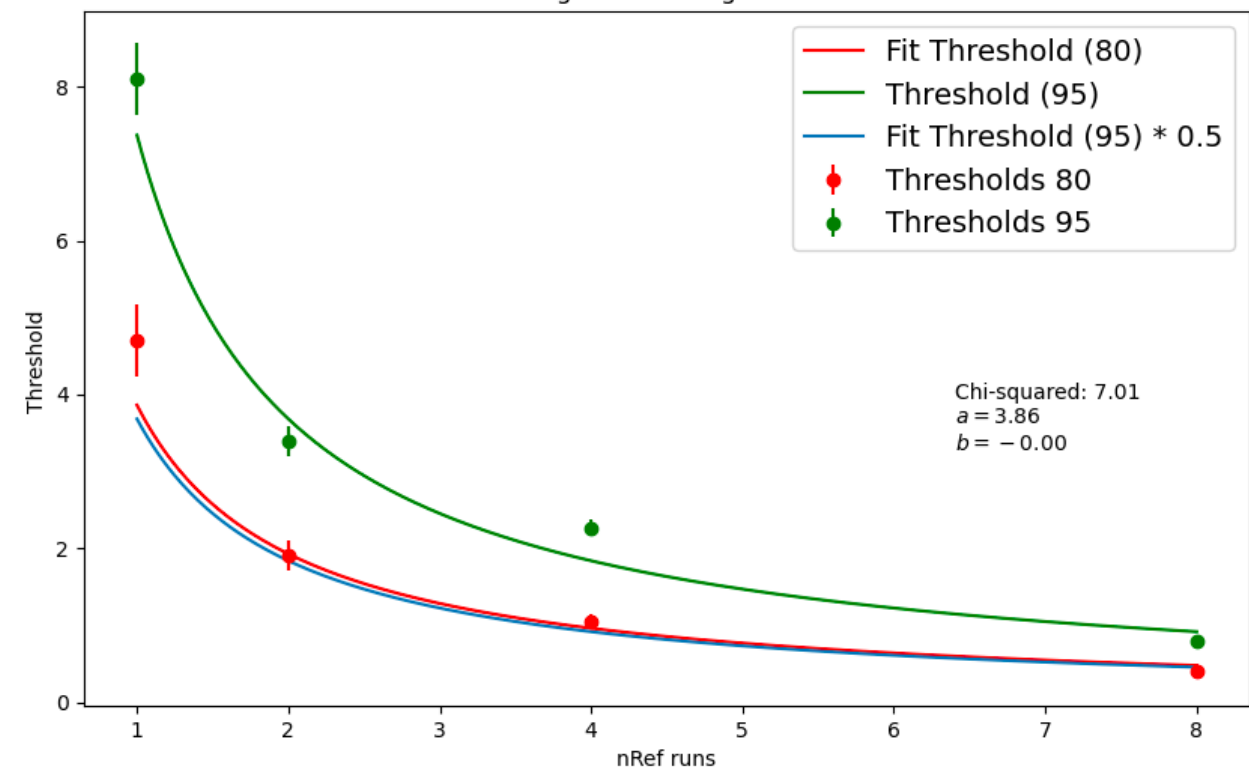
- Implemented thresholds based on these fits in AutoDQM code
- Decided to just use 50% of “tight” flagging threshold (i.e. 95% WP) to define “loose” plot-generation / discard threshold

Results for the DOC3 Task histograms

Curve Fitting - Task histograms - Chi2

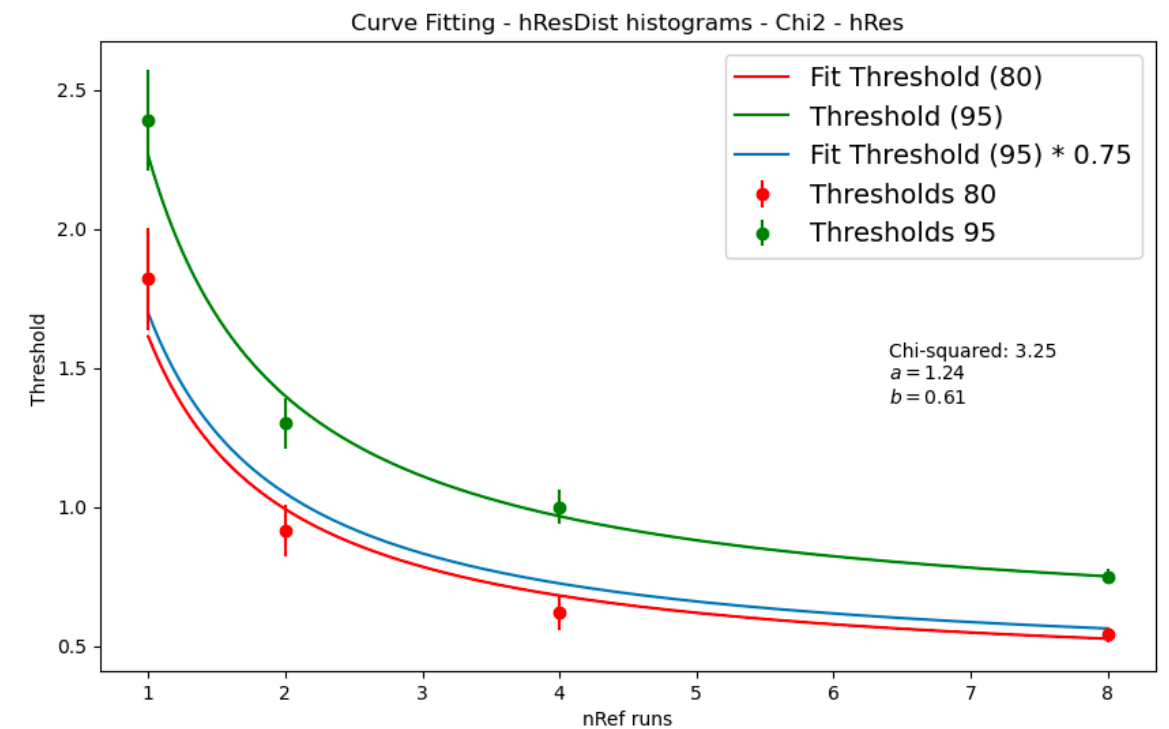
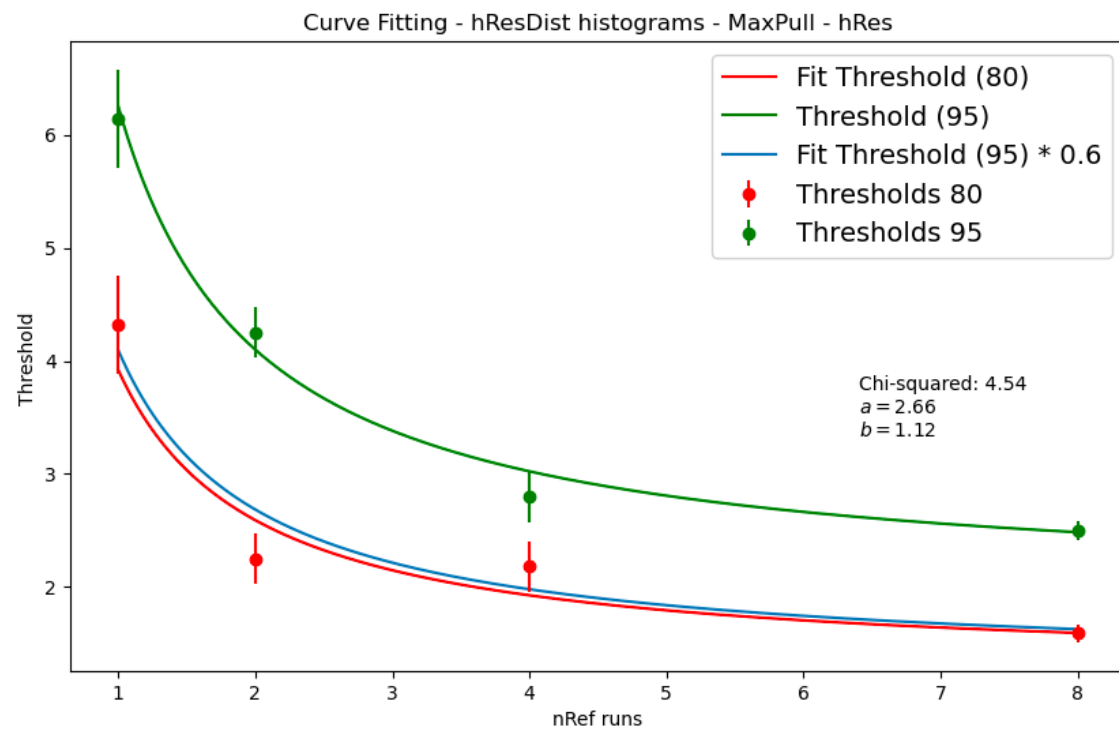


Curve Fitting - Task histograms - MaxPull



- In a couple cases manually adjusted flagging threshold fit so that it would never flag χ^2 or max pull scores < 1.0

Results for the DOC3 hResDist histograms



- Full set of plots is deployed in two separate AutoDQM workspaces, DT_DOC1 and DT_DOC3 — and seems to be working well!

AutoDQM | **Plots**

Filter plots

Select a different data run

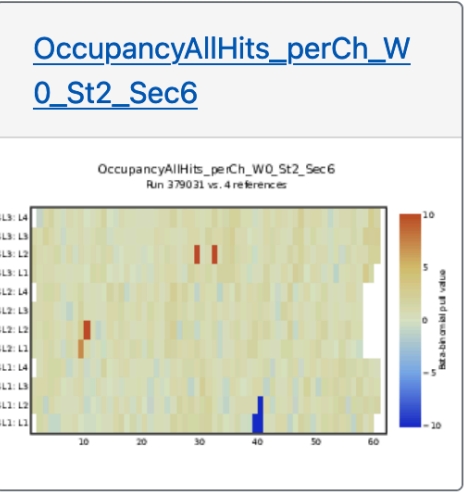
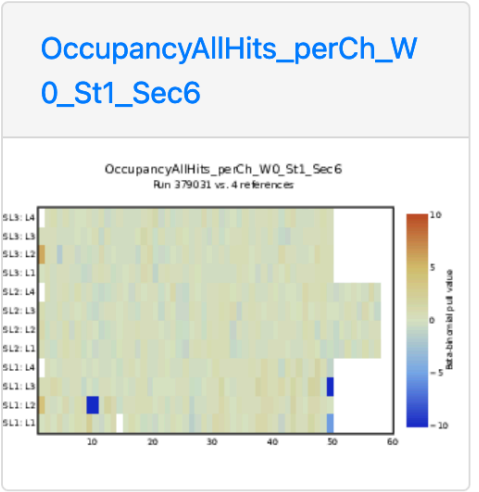
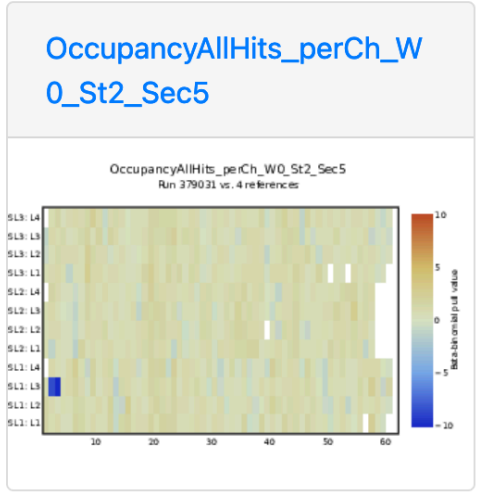
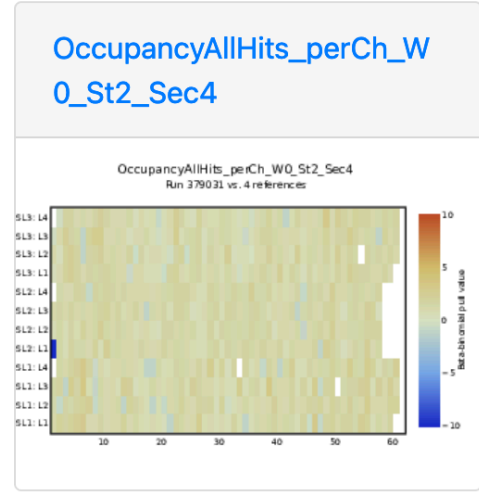
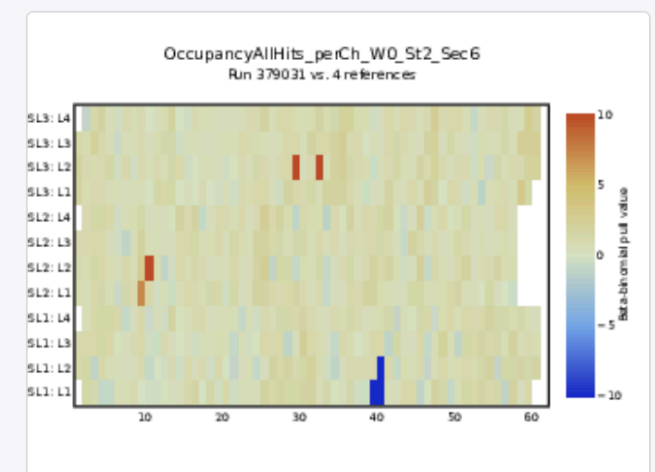
[Prev](#) | [Select...](#) | [Next](#)

Show hidden plots

AutoDQM Report

DQM source: Online
Subsystem: DT_DOC1
Fri, 12 Apr 2024 04:29:42 GMT

Details	Data Run	Ref Run
Series	00037xxxx	00037xxxx
Sample	0003790xx	0003790xx
Run	379031	379028_379011_378993_378985



NameOccupancyAllHits_perCh_W0_St2_S...

Comparator: beta_binomial

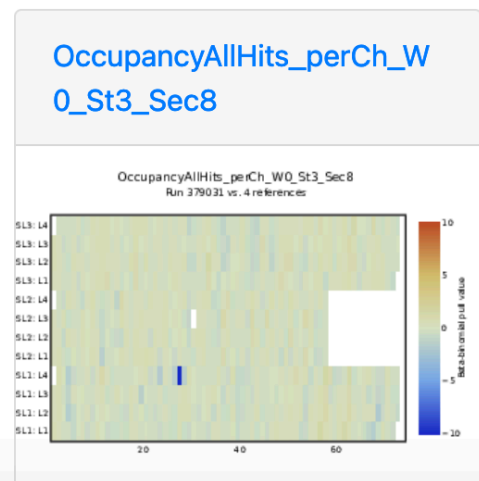
Anomalous: true

Chi_Squared: 6.93

Max_Pull_Val: -36.82

Data_Entries: 15020

Ref_Entries: 2400, 20000



- Recent run (379031) has just 5 occupancy plots flagged in Doc 1 workspace, and only 5 plots generated — all others have scores < 50% of flagging threshold

https://cmsweb-testbed.cern.ch/dqm/autodqm/results/pdfs/DATA-00037xxxx-0003790xx-379031_REF-00037xxxx-0003790xx-379028_379011_378993_378985_OccupancyAllHits_perCh_W0_St2_Sec6_COMP-beta_binomial_69092.pdf

[1] https://cmsweb-testbed.cern.ch/dqm/autodqm/plots/Online/DT_DOC1/00037xxxx/0003790xx/379028_379011_378993_378985/00037xxxx/0003790xx/379031



Filter plots

Select a different data run

Prev Select... Next

Show hidden plots



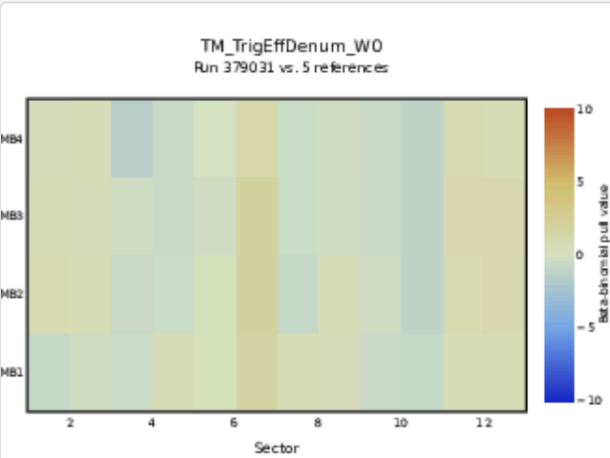
AutoDQM Report

DQM source: Offline

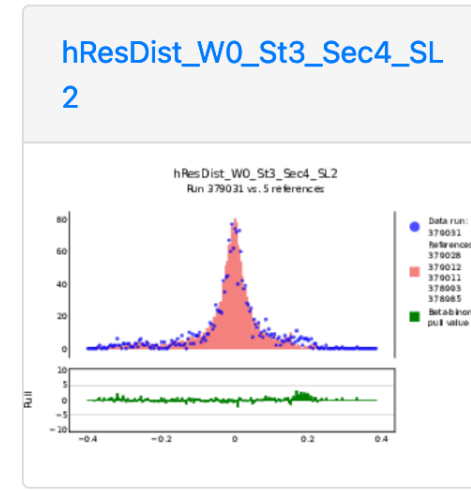
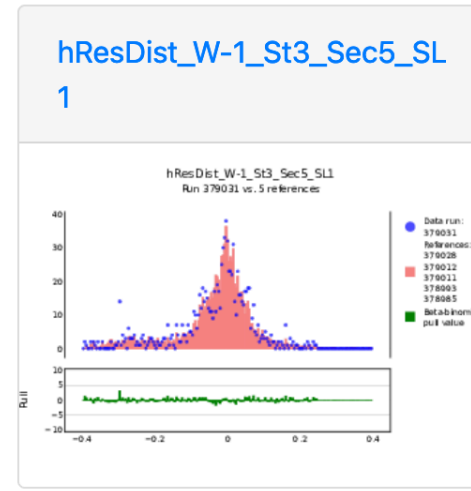
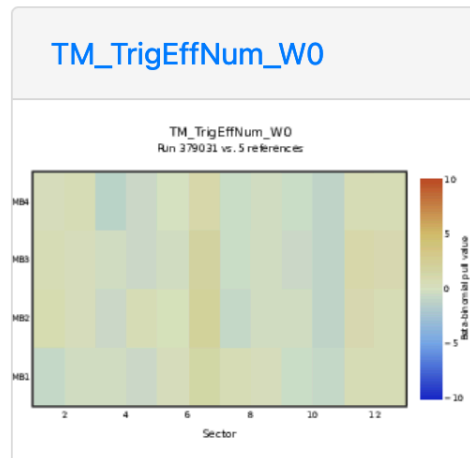
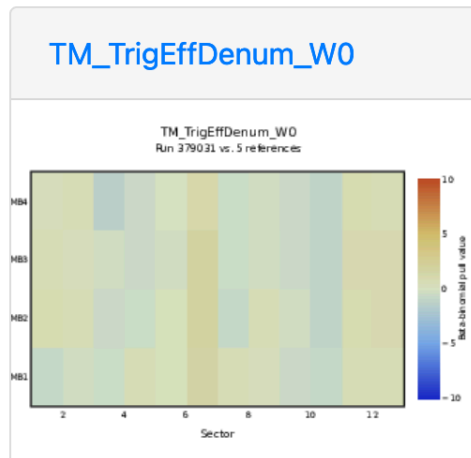
Subsystem: DT_DOC3

Fri, 12 Apr 2024 04:33:15 GMT

Details	Data Run	Ref Run
Series	Run2024	Run2024
Sample	Muon0	Muon0
Run	379031	379028_379012_379011_378993_378985



Name	TM_TriggerEffDenum_W0
Comparator	beta_binomial
Anomalous	false
Chi_Squared	0.68
Max_Pull_Val	0.2
Data_Entries	5040
Ref_Entries	2665 - 9850



- No plots flagged for same run in Doc 3 workspace, and only 4 plots generated (click “Show hidden plots”)

[1] https://cmsweb-testbed.cern.ch/dqm/autodqm/plots/Offline/DT_DOC3/Run2024/Muon0/379028_379012_379011_378993_378985/Run2024/Muon0/379031

Filter plots

Select a different data run

Prev Select... Next

Show hidden plots

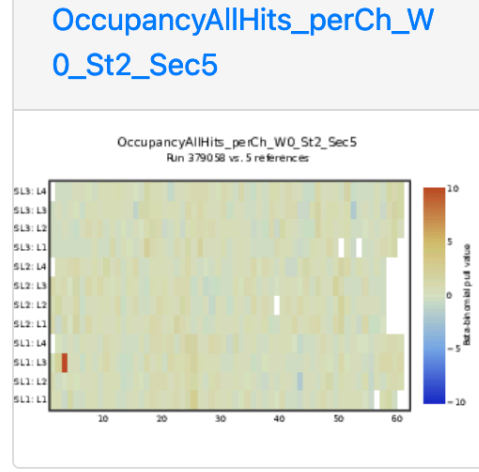
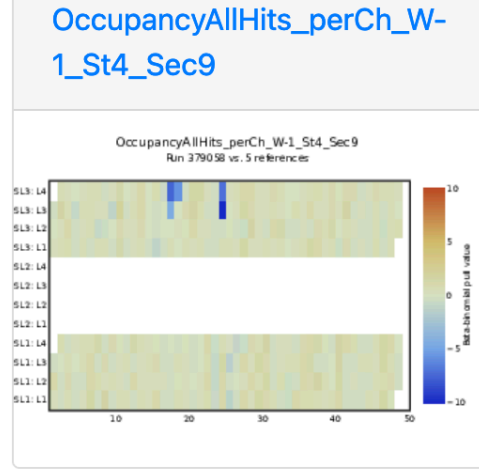
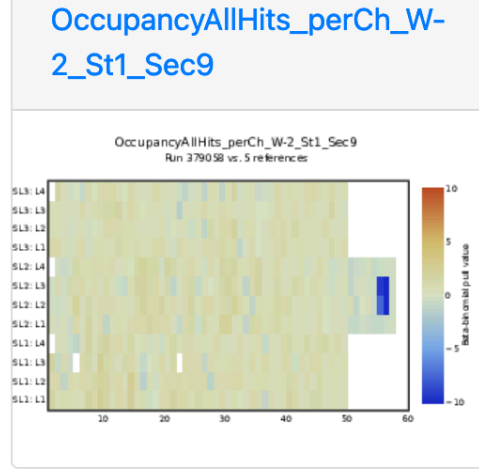
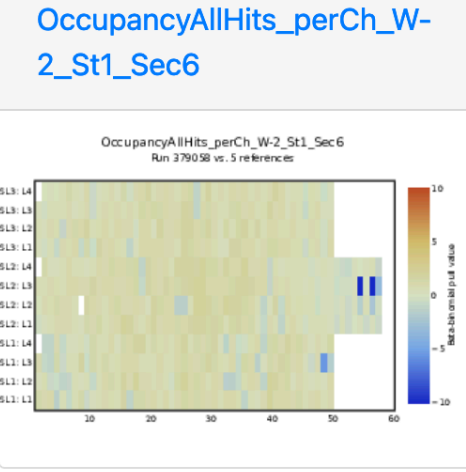
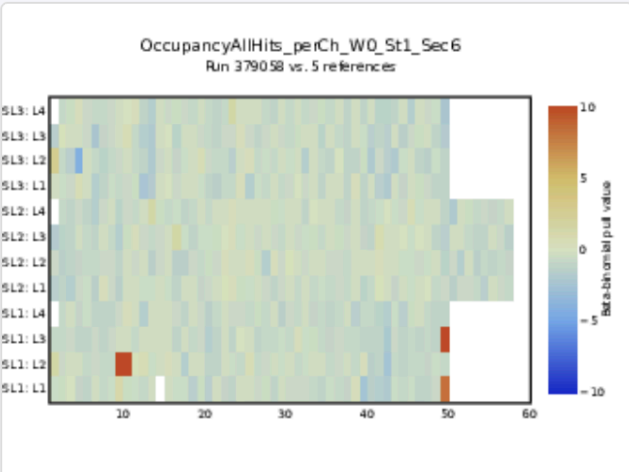
AutoDQM Report

DQM source: Online

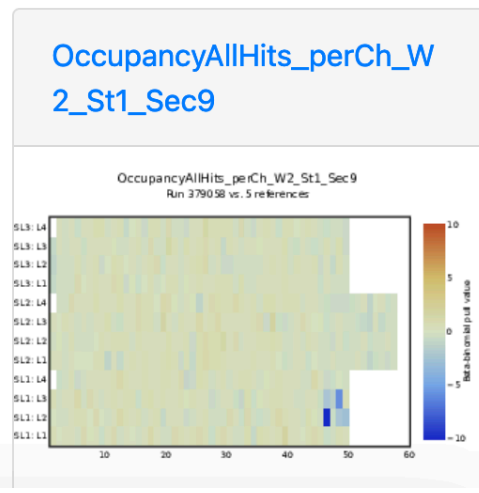
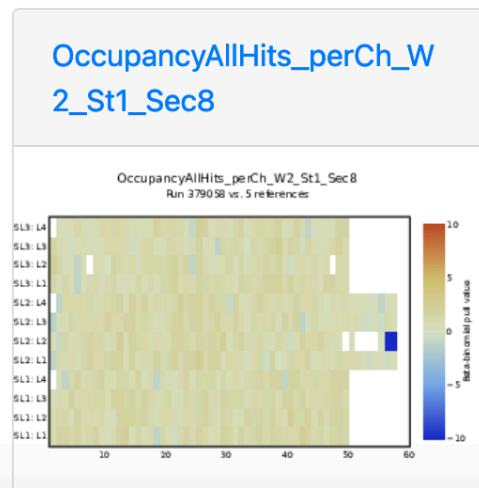
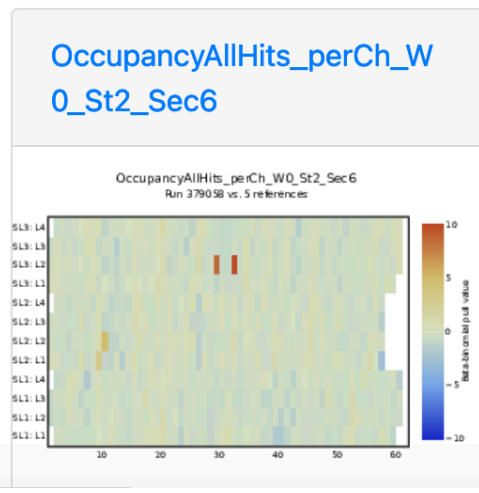
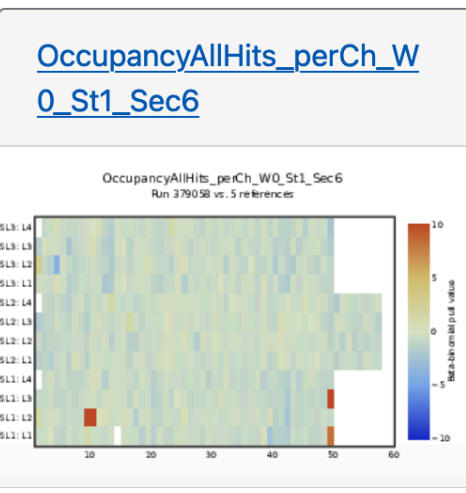
Subsystem: DT_DOC1

Fri, 12 Apr 2024 04:37:08 GMT

Details	Data Run	Ref Run
Series	00037xxxx	00037xxxx
Sample	0003790xx	0003790xx
Run	379058	379031_379028_379011_378993_378985



Name	OccupancyAllHits_perCh_W0_St1_S...
Comparator	beta_binomial
Anomalous	true
Chi_Squared	1.63
Max_Pull_Val	15.58
Data_Entries	12990
Ref_Entries	1572 14500



By contrast, run 379058 from the next fill has dozens of flagged anomalies!

[1] https://cmsweb-testbed.cern.ch/dqm/autodqm/plots/Online/DT_DOC1/00037xxxx/0003790xx/379031_379028_379011_378993_378985/00037xxxx/0003790xx/379058



Filter plots

Select a different data run

Prev Select... Next

Show hidden plots

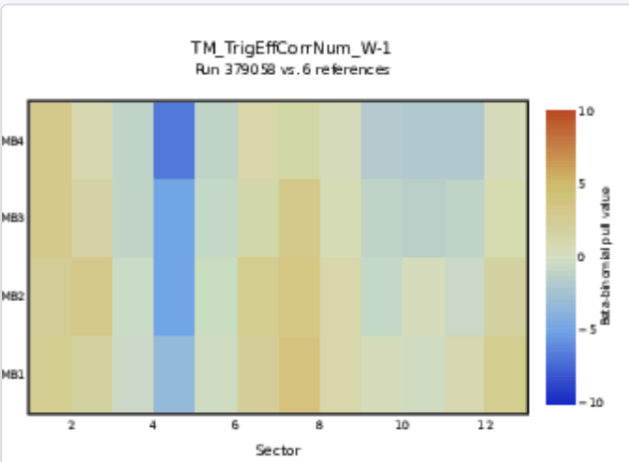
AutoDQM Report

DQM source: Offline

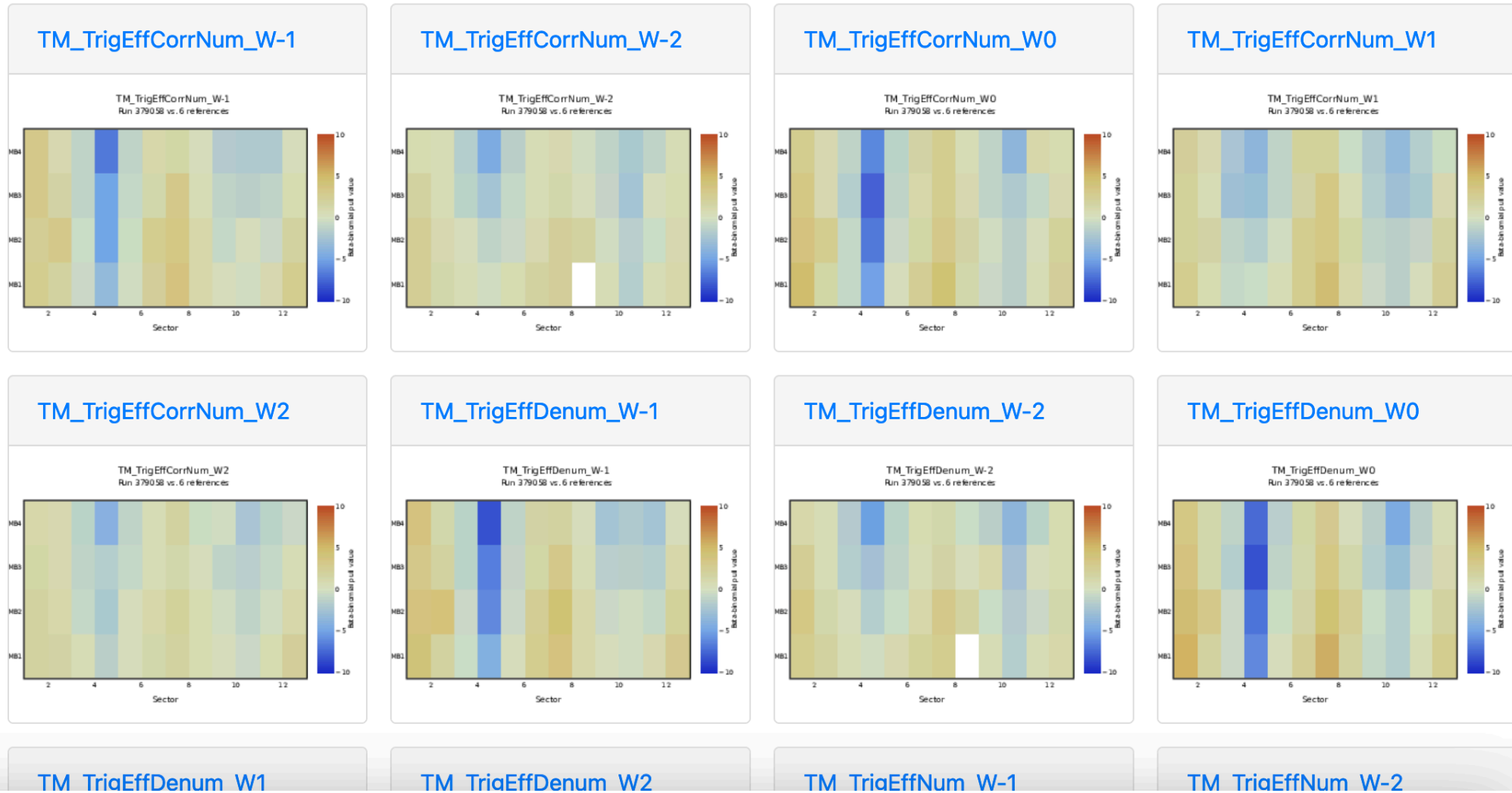
Subsystem: DT_DOC3

Fri, 12 Apr 2024 04:36:45 GMT

Details	Data Run	Ref Run
Series	Run2024	Run2024
Sample	Muon0	Muon0
Run	379058	379031_379028_379012_379011_378993_378985



Name	TM_TrigEffCorrNum_W-1
Comparator	beta_binomial
Anomalous	true
Chi_Squared	5.21
Max_Pull_Val	-6.14
Data_Entries	50822
Ref_Entries	1536 - 6523



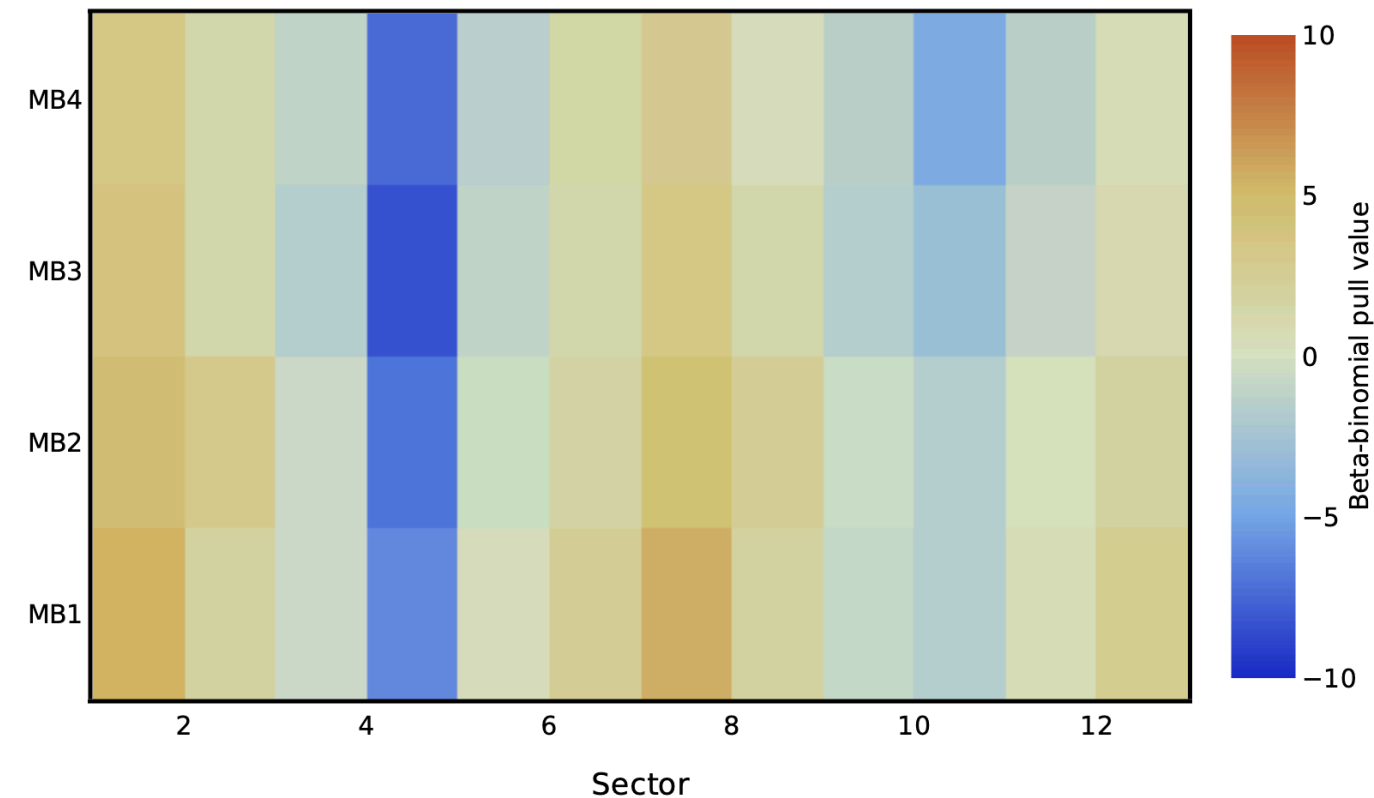
• Dozens of flagged anomalies in Offline DQM histograms as well

[1] https://cmsweb-testbed.cern.ch/dqm/autodqm/plots/Offline/DT_DOC3/Run2024/Muon0/379031_379028_379012_379011_378993_378985/Run2024/Muon0/379058

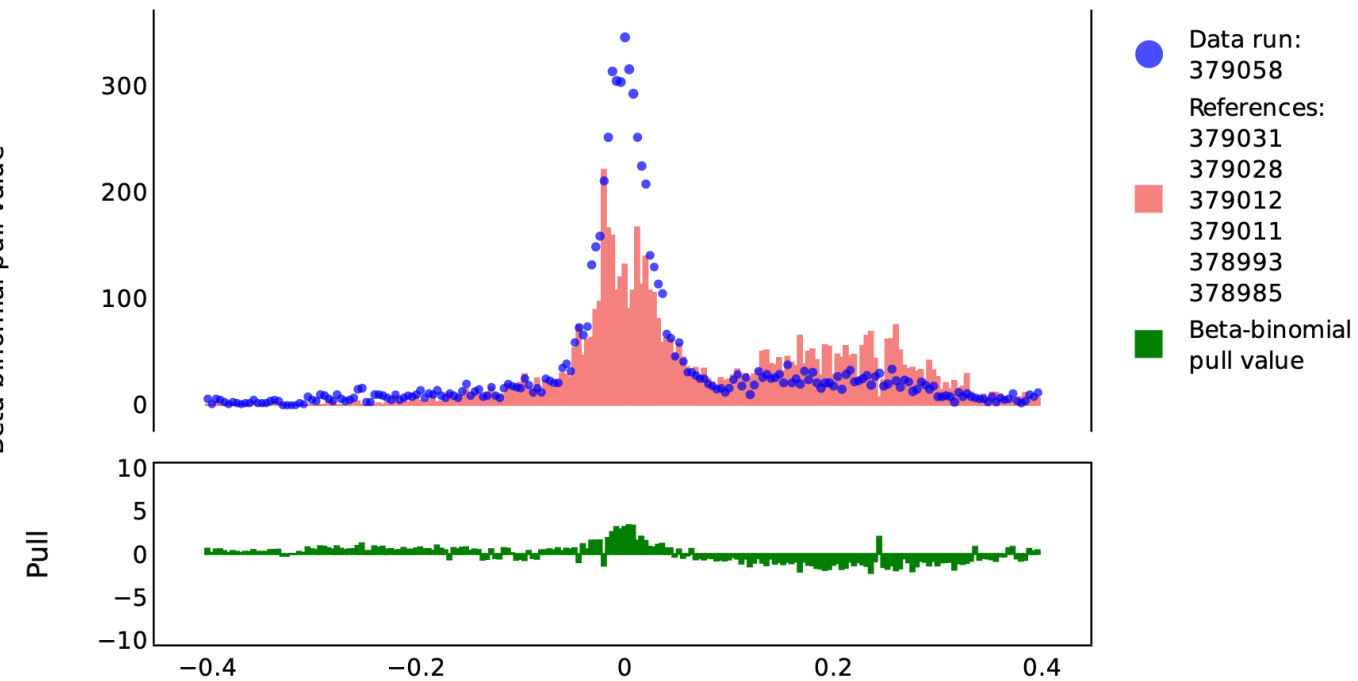


Taking a closer look

TM_TrigEffDenum_W0
Run 379058 vs. 6 references

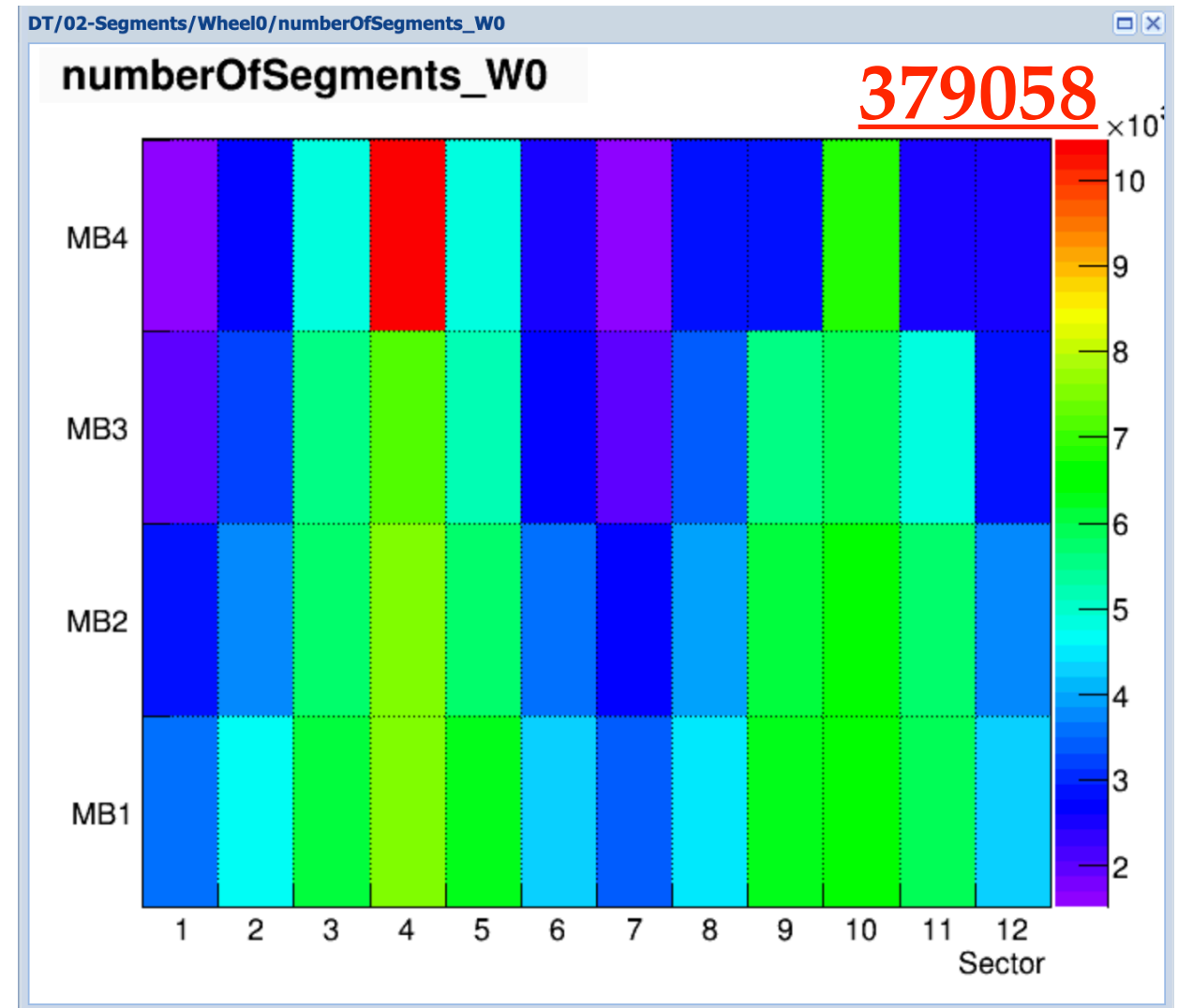
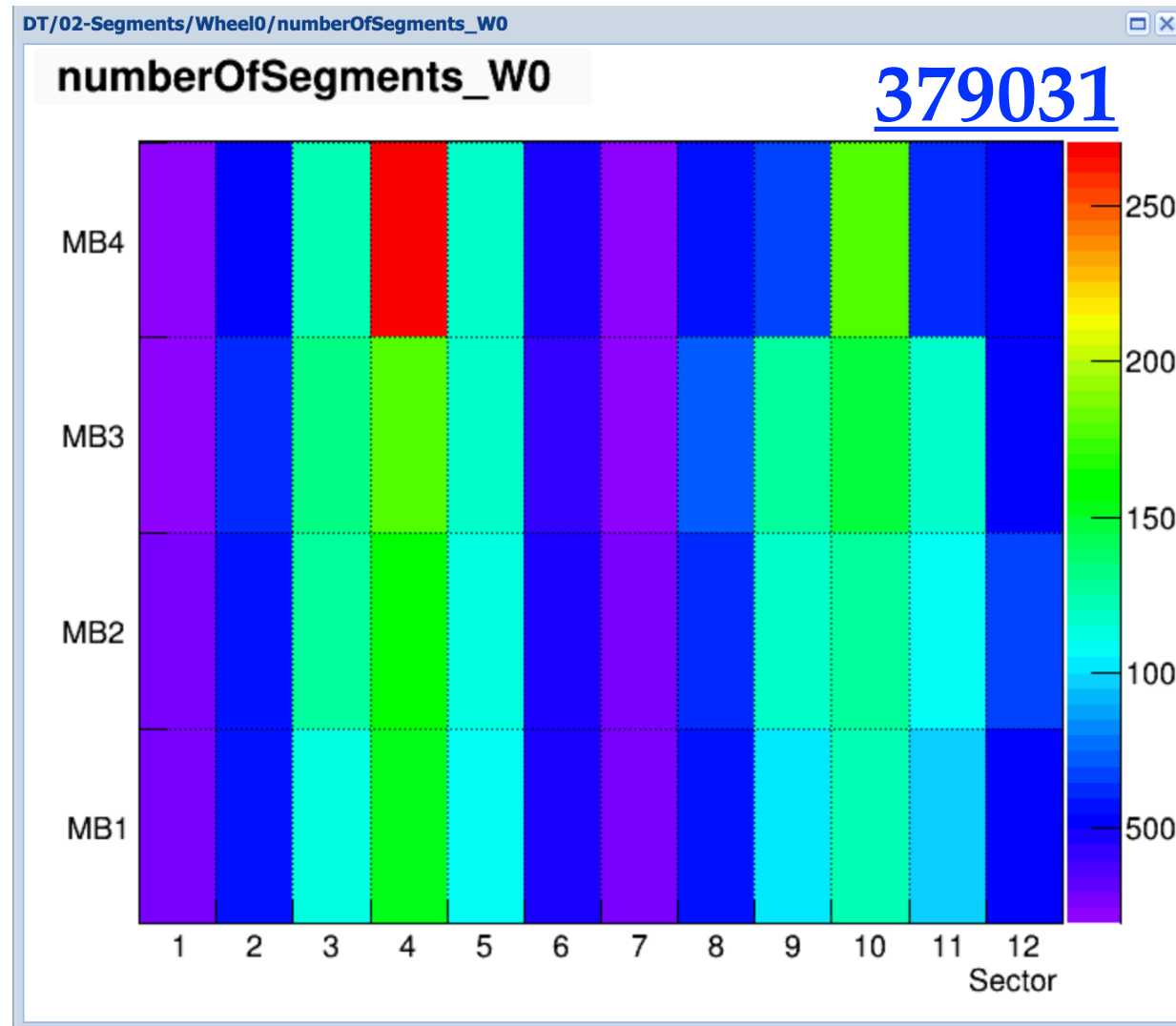


hResDist_W-2_St2_Sec9_SL3
Run 379058 vs. 6 references



- Deficits in sector 4 in multiple wheels: what could explain this?
- More in-time segments in later run: probably because LHC went from 8 to 62 colliding bunches

Cosmic explanation



- Collision rate small enough that most muons in DT are still comics
- “Deficit” in sector 4 really just an increase in collision muons, spread evenly among all sectors. Impressive flag by AutoDQM!

Shifter script

- Stand-alone script (run from laptop!) to automatically form web addresses and open AutoDQM pages for multiple data runs
- Especially useful when using multiple reference runs, since this can only be done via web address right now
- Options to open relevant Online DQM and OMS pages for run
 - Very configurable, could add Offline DQM pages, other Muon resources?

```
python3 Downloads/URL_AutoDQM_Muon.py --subsystem DT_DOC1 --source Online
--data_runs 379031,379058 --ref_runs 379028,379011,378993,378985 --
multiref --max_ref 8 --dqm_data --oms_data
```

```
python3 Downloads/URL_AutoDQM_Muon.py --subsystem DT_DOC3 --source Offline
--data_runs 379031,379058,379028,379011,378993,378985 --recursive
--prev_ref --multiref --max_ref 8
```


Next steps

- **Please have shifters / DOCs try out DT AutoDQM pages!**
- For DT, Caio working on flagging thresholds for 26 “DT Online” histograms from spreadsheet, which feed the DT summary plots
 - May make a “DT_summary” page with no plots thrown away
- Ryan Nie (Boston) working on flagging thresholds for CSC plots
- Could use a volunteer to help with final selection of RPC histograms to monitor, setting flagging thresholds


BACKUPS

Muon Doc3 documentation

- Muon Run Coordination Office (RCO) twiki
 - <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuRCO>
- Muon Doc3 #1 twiki and tutorial (2022)
 - <https://twiki.cern.ch/twiki/bin/view/CMS/MuonDPGDataCertificationShiftInstructions2022>
 - <https://indico.cern.ch/event/1164944/>
- Muon Doc3 #2 twiki, proposal (2022), and tutorial (2023)
 - <https://twiki.cern.ch/twiki/bin/view/CMS/MuonDOC3TrendsMonitoring>
 - <https://indico.cern.ch/event/1169570/#3-muon-doc3-2-trends-monitorin>
 - <https://indico.cern.ch/event/1283847/#1-muon-doc3-2-trends-monitorin>

Muon Doc3 #1 DQM plots

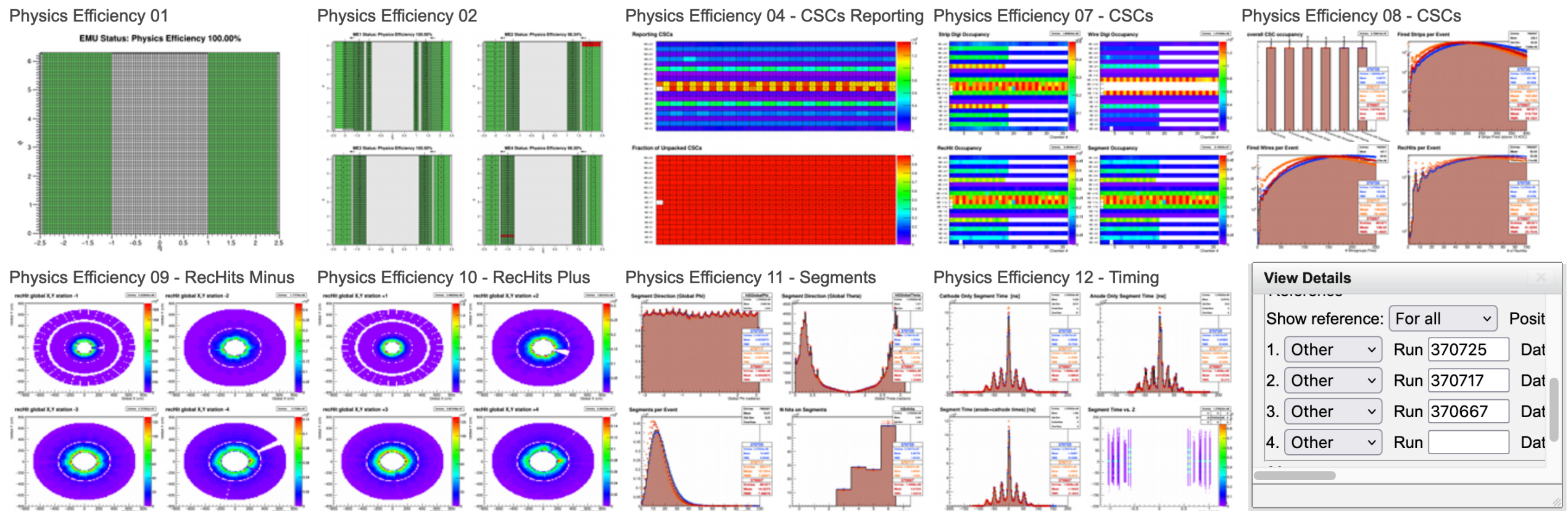
CSC Offline DQM quick collection


Service ▾ **Workspace** ▾ **Run #** ▾ LS # Event # Run started, UTC time
Offline: CSC . 370'749 . 1'460 . (None) . Sun Jul 16, 02:29
 Data: /Muon0/Run2023D-PromptReco-v2/DQMIO

CMS DQM GUI (vocms0738.cern.ch)
 Dec 13, 2023 at 22:27.49 UTC
 Andrew Brinkerhoff, [View details](#)

Size: **Medium** ▾ **Play** **Reset Workspace** **Describe** **Customise** **Layouts** [\(Top\)](#) / Quick collection


[ROOT File](#) JSROOT mode JSON data [Link-Me](#)



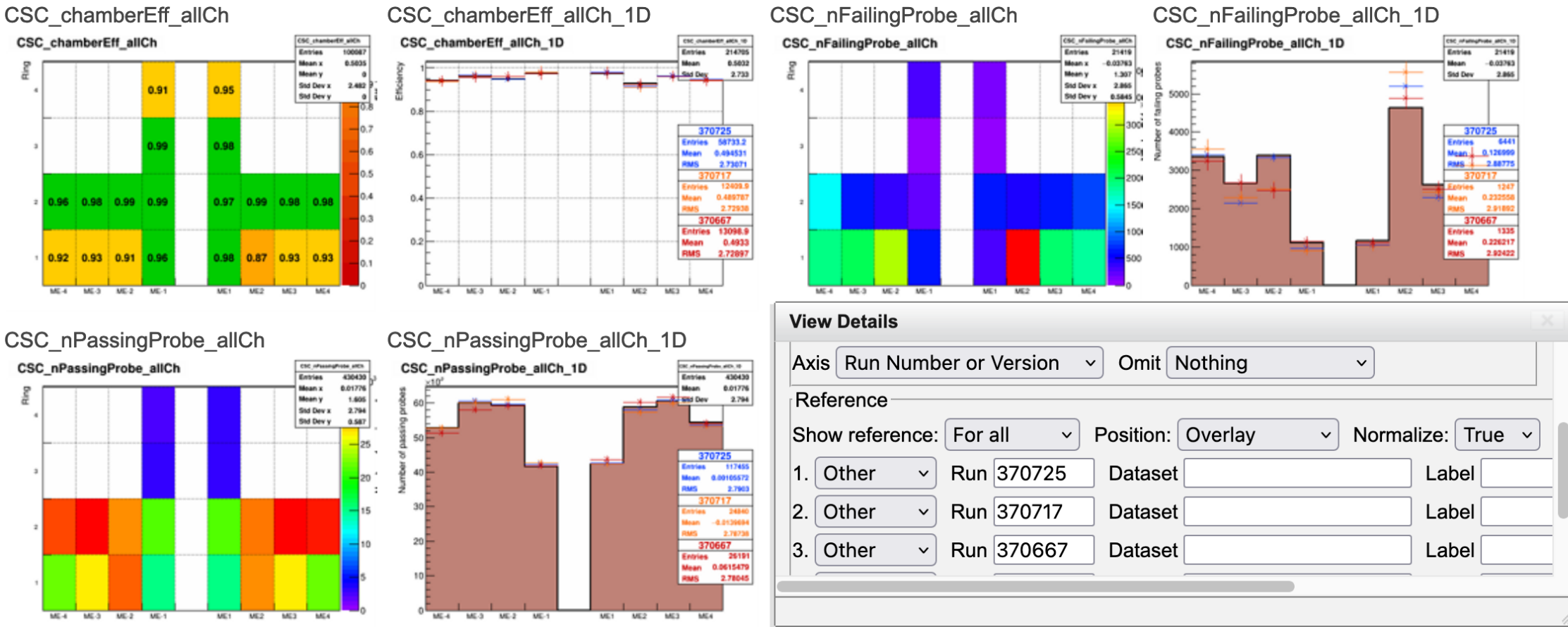
- CSC quick collection contains 31 plots : 6 2D efficiency, 14 2D + 11 1D occupancy** plots (**occupancy = any histogram filled with integer entries)
- Currently can only overlay reference runs for 1D plots

[1] <https://twiki.cern.ch/twiki/bin/view/CMS/CSCDPGQuickCert>

CSC Offline DQM segment T&P



Service ▼ **Workspace** ▼ **Run #** ▼ **LS #** **Event #** **Run started, UTC time**
Offline: CSC . 370'749 . 1'460 . (None) . Sun Jul 16, 02:29
Data: /Muon0/Run2023D-PromptReco-v2/DQMIO

Size: Medium ▼ Play Reset Workspace Describe Customise Layouts (Top) / CSC / Segment_TnP / Task
ROOT File JSROOT mode JSON data [Link-Me](#)



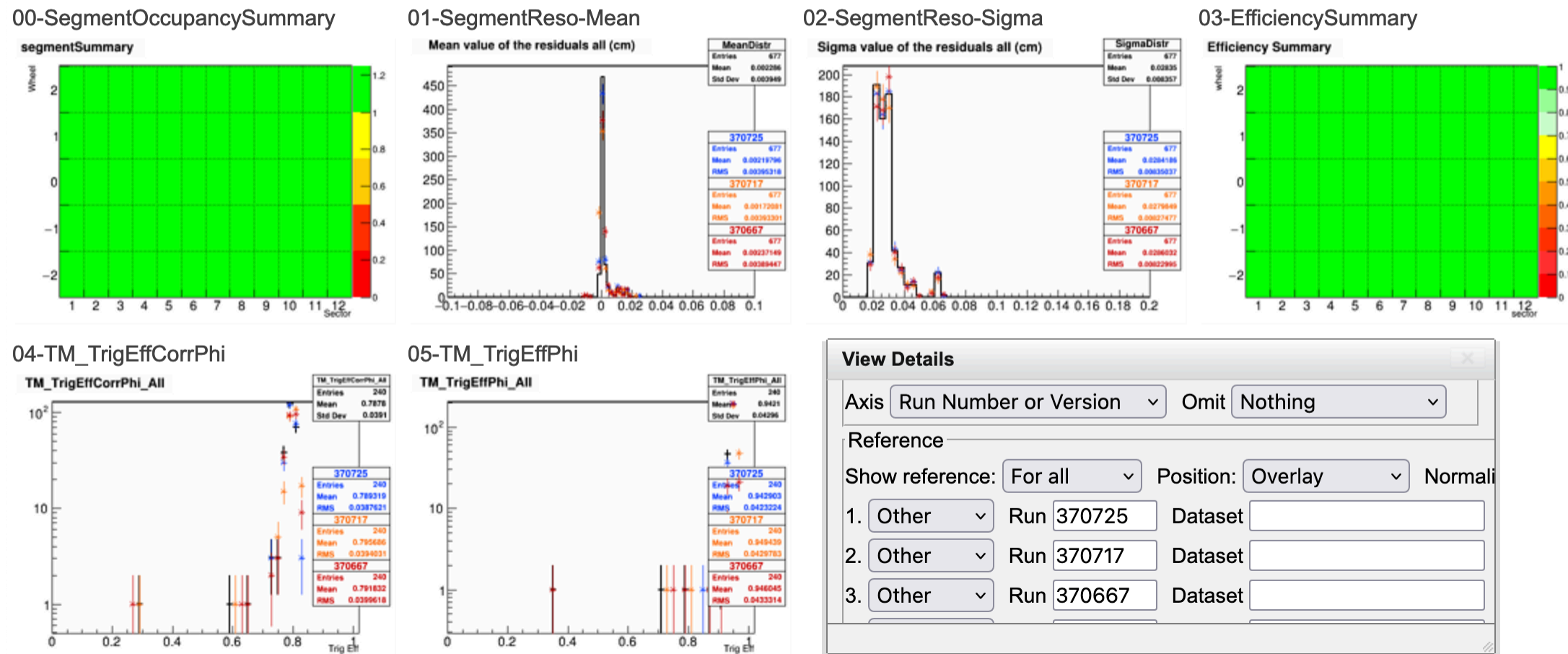
- CSC segment tag-and-probe contains one 1D and one 2D efficiency plot, plus the corresponding passing / failing occupancy plots
- AutoDQM cannot directly test efficiency plots, but could use “nFailingProbe” histograms -- should be sensitive to any decrease in efficiency

DT Offline DQM shift workspace


Service ▾ **Workspace** ▾ **Run #** ▾ LS # Event # Run started, UTC time
Offline: Shift . 370'749 . 1'460 . (None) . Sun Jul 16, 02:29
CMS DQM GUI (vocms0738.cern.ch)
Dec 14, 2023 at 05:45.53 UTC
Andrew Brinkerhoff, [View details](#)

Data: /Muon0/Run2023D-PromptReco-v2/DQMIO


Size: Medium ▾ Play Reset Workspace Describe Customise Layouts (Top) / 00 Shift / DT [ROOT File](#) JSROOT mode JSON data [Link-Me](#)



- DT shift workspace contains 6 plots : 2 “summary”, 4 1D occupancy
- AutoDQM can test 1D plots and 2D occ. plots “underneath” summary plots

[1] <https://twiki.cern.ch/twiki/bin/view/CMS/DTCertification>

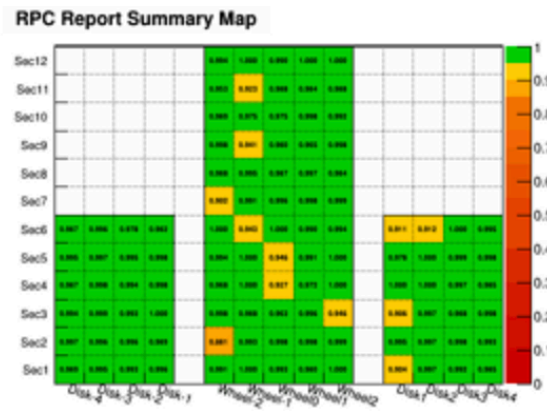
RPC Offline DQM Layouts


Service ▾ **Workspace** ▾ **Run #** ▾ LS # Event # Run started, UTC time
Offline: RPC . 370'749 . 1'460 . (None) . Sun Jul 16, 02:29
 Data: /Muon0/Run2023D-PromptReco-v2/DQMIO

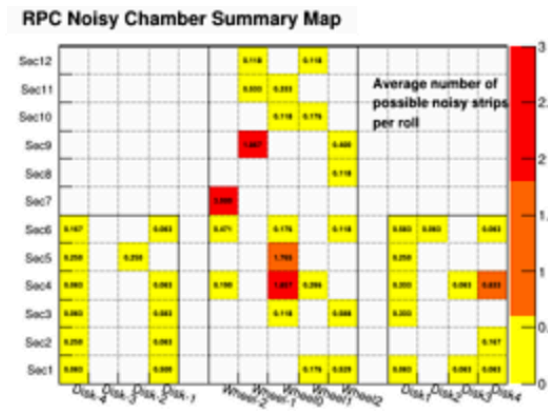
CMS DQM GUI (vocms0738.cern.ch)
 Dec 14, 2023 at 05:56.22 UTC
 Andrew Brinkerhoff, [View details](#)

Size: Medium ▾ Play Reset Workspace Describe Customise Layouts (Top) / RPC / Layouts [ROOT File](#) JSROOT mode JSON data [Link-Me](#)

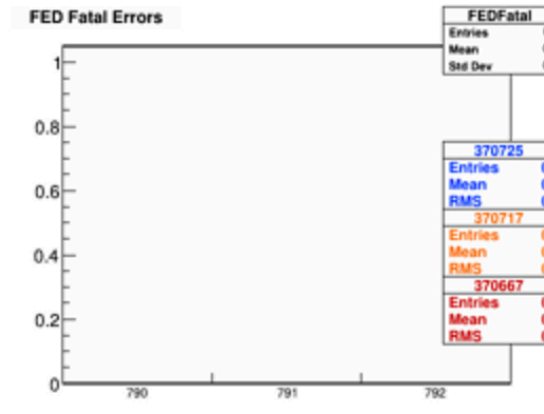
00-Summary_Map



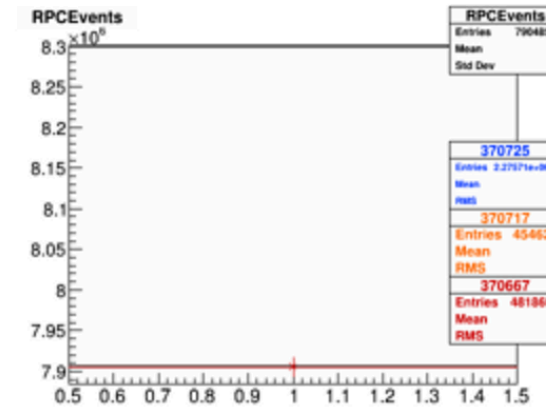
01-Noisy_summary_Map



02-Fatal_FED_Errors



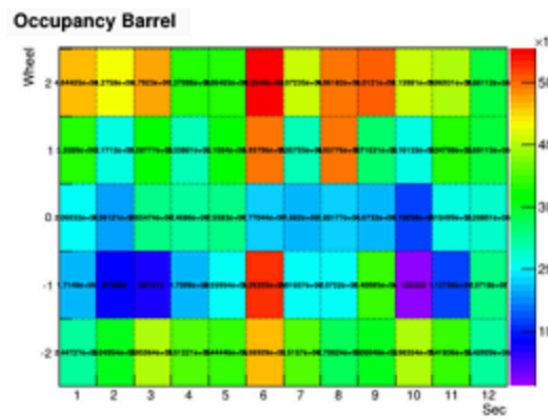
03-RPC_Events



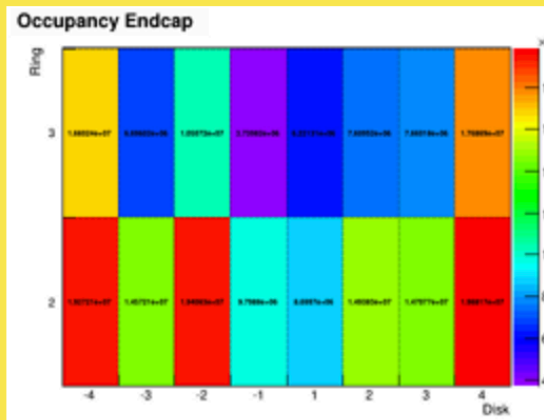
04-RPC_HV_Status



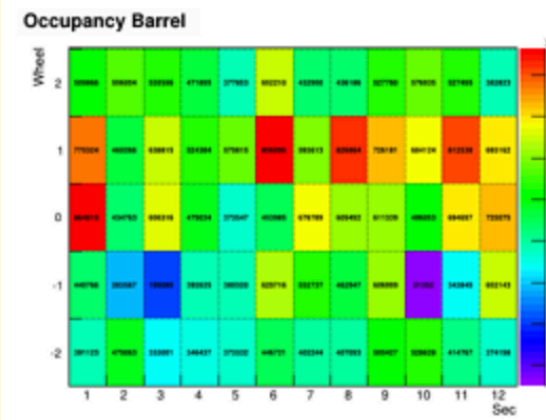
05-Barrel_Occupancy



06-Endcap_Occupancy




07-Barrel_Occupancy_Muon



- RPC layouts contain 50 plots, including 17 2D and 18 1D occupancy plots
- AutoDQM cannot test summary maps or mean cluster size plots

[1] <https://twiki.cern.ch/twiki/bin/view/CMS/RpcMuonDoc3Instructions>

GEM Offline DQM quick collection

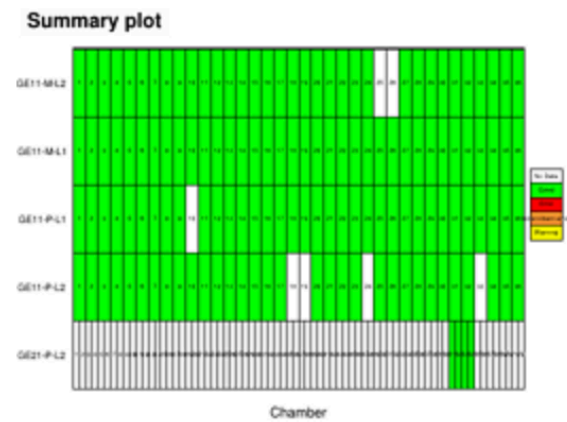

 Service ▾ Workspace ▾ Run # ▾ LS # Event # Run started, UTC time
Offline: GEM . 370'749 . 1'460 . (None) . Sun Jul 16, 02:29
 Data: /Muon0/Run2023D-PromptReco-v2/DQMIO

CMS DQM GUI (vocms0738.cern.ch)
 Dec 14, 2023 at 06:11.28 UTC
 Andrew Brinkerhoff, [View details](#)

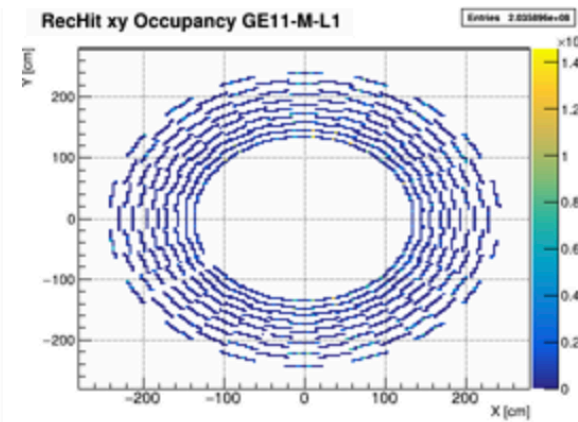
Size: Medium ▾ Play Reset Workspace Describe Customise Layouts (Top) / Quick collection

[ROOT File](#) JSROOT mode JSON data [Link-Me](#)

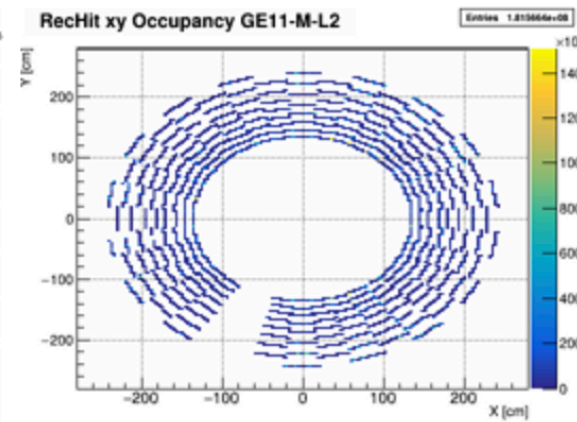
00 Summary Map



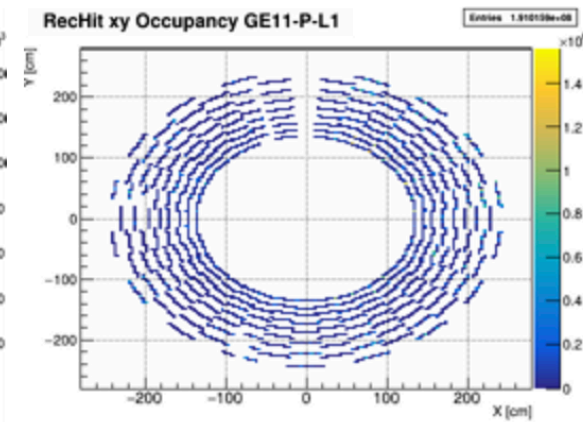
01 GE11-M-L1 RechHit Occupancy



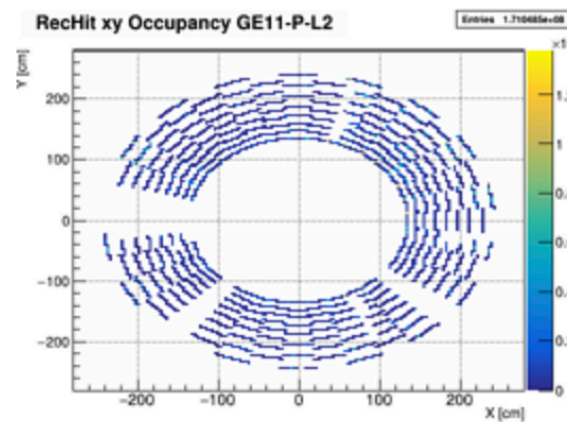
02 GE11-M-L2 RechHit Occupancy



03 GE11-P-L1 RechHit Occupancy



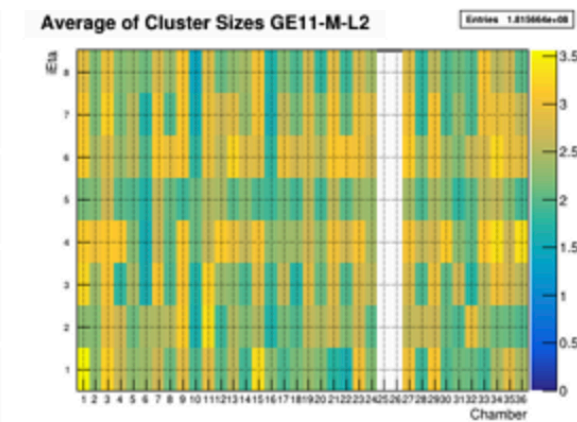
04 GE11-P-L2 RechHit Occupancy



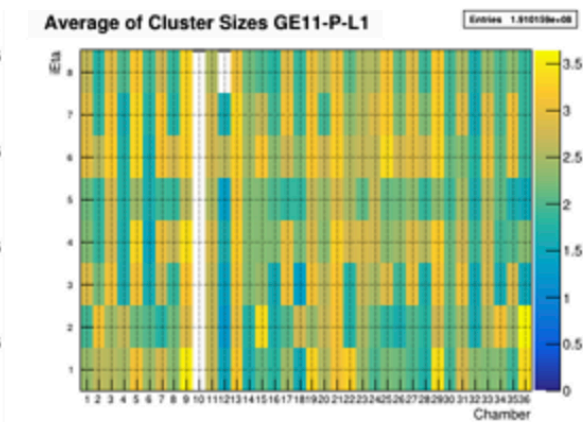
05 GE11-M-L1 RechHit Average Cluster



06 GE11-M-L2 RechHit Average Cluster



07 GE11-P-L1 RechHit Average Cluster



- GEM quick collection contains 14 plots, but only 5 occupancy plots
- Plots in quick collection / layouts don't match instructions; is one out of date?

[1] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/GEMDataCertificationAndDOC3>