

# Muon flux unfolding: update on systematics

Pascal Gutjahr

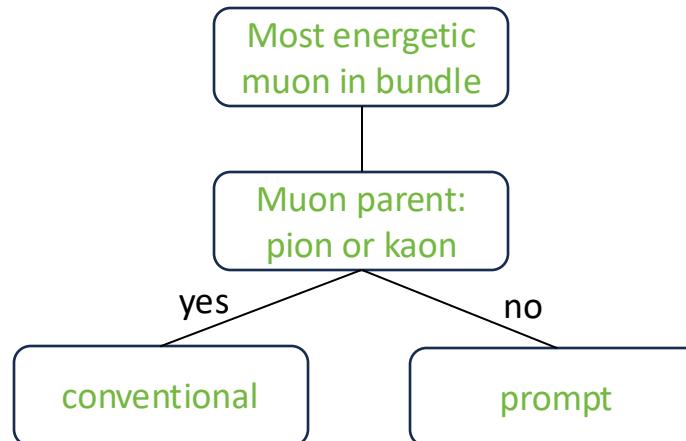
WG reviewer: Dennis Soldin  
Technical reviewer: Karolin Hymon  
Wiki: [prompt wiki](#)  
Last update: [madison\\_talk](#)

CR-Call

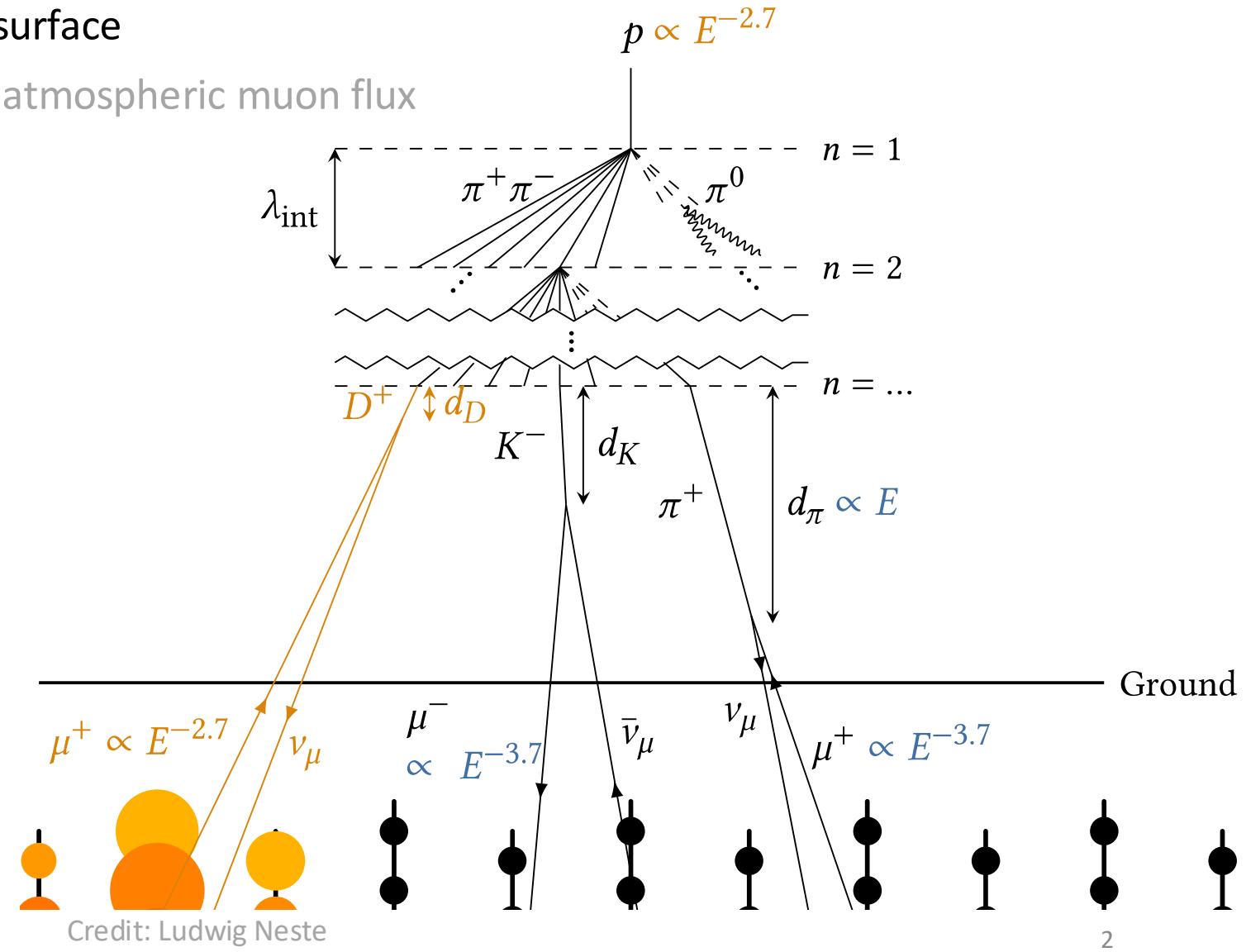
February 7, 2025

# Analysis Goals: Detect Prompt Muons

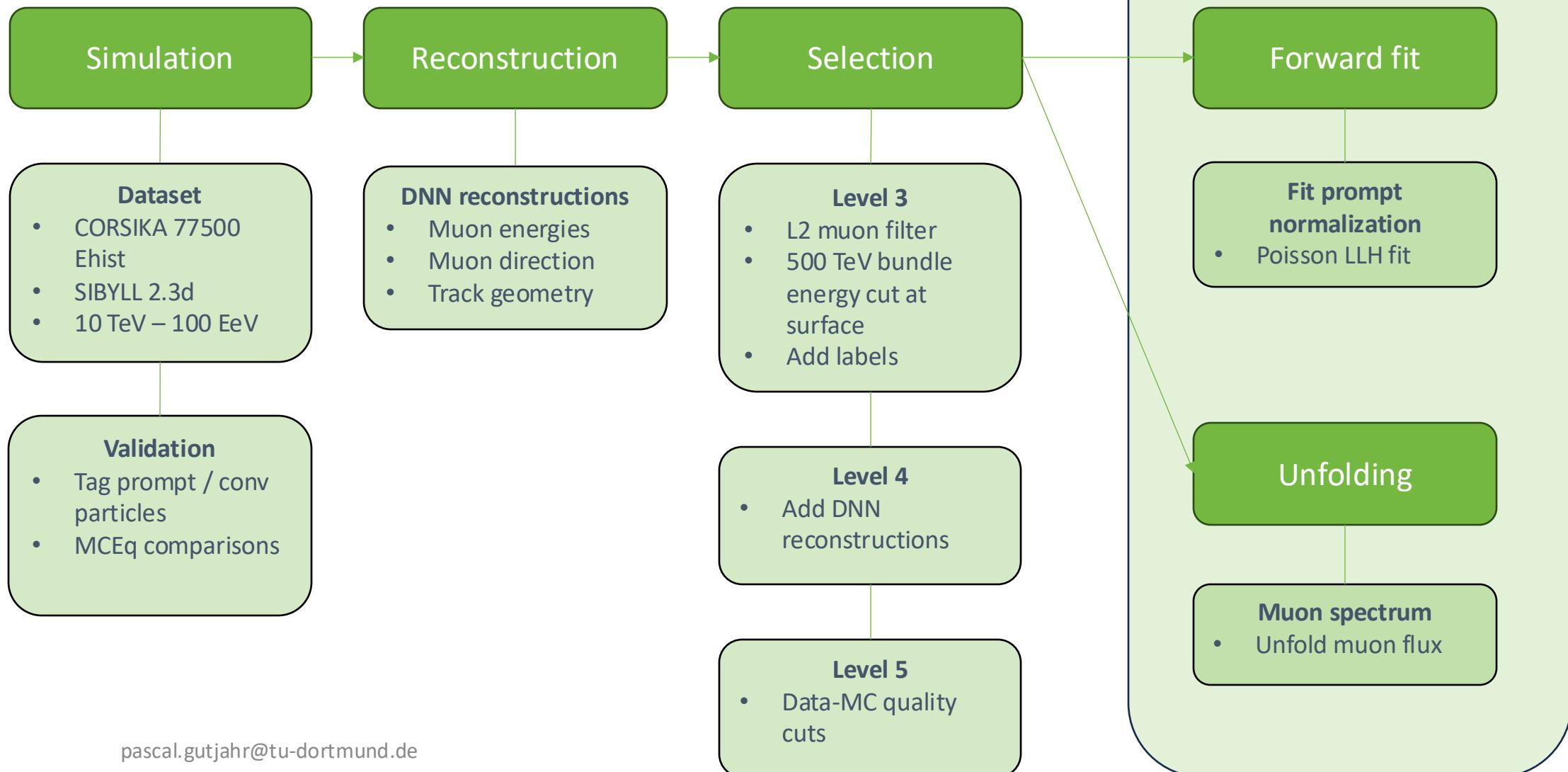
- 1) Unfold a muon energy spectrum at surface
- 2) Measure prompt component of the atmospheric muon flux



pascal.gutjahr@tu-dortmund.de



# Overview



# Unfolding

# Unfolding in a nutshell

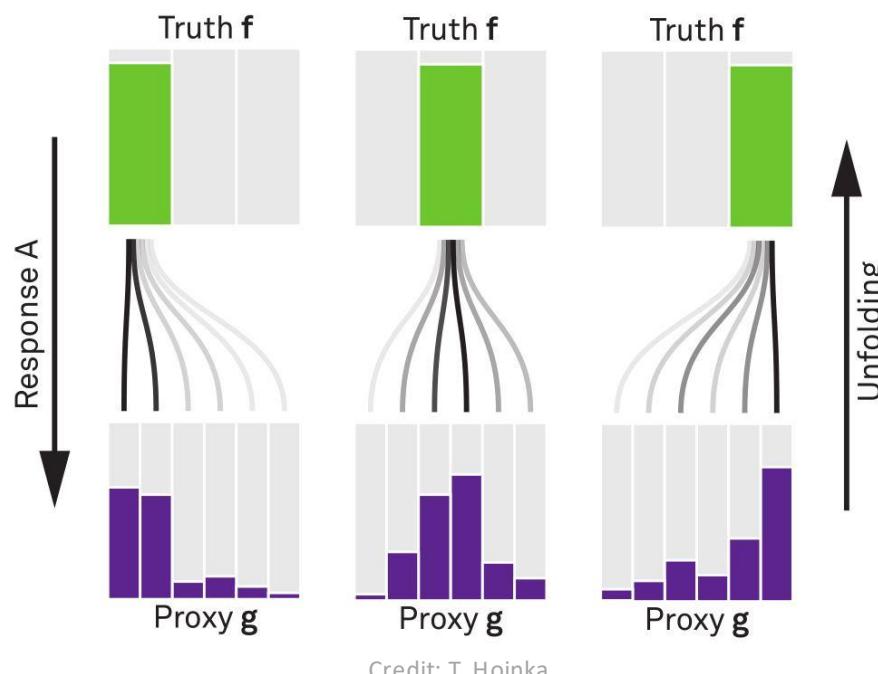
measured proxy

$$g(y) = \int_{E_0}^{E_1} A(E_\mu, y) f(E_\mu) dE_\mu + b(y)$$

detector response

background

true energy distribution



folding

unfolding

1. discretized form:  $\vec{g} = A\vec{f} \leftrightarrow \vec{f} = A^{-1}\vec{g}$

2. maximum likelihood method:

3. Tikhonov regularization:

$$\mathcal{L}(\vec{g}|\vec{f}) = \prod_{j=1}^M \frac{\lambda_j^{g_j}}{g_j!} \exp(-\lambda_j)$$

$$= \prod_{j=1}^M \frac{(A\vec{f})_j^{g_j}}{g_j!} \exp(-(A\vec{f})_j)$$

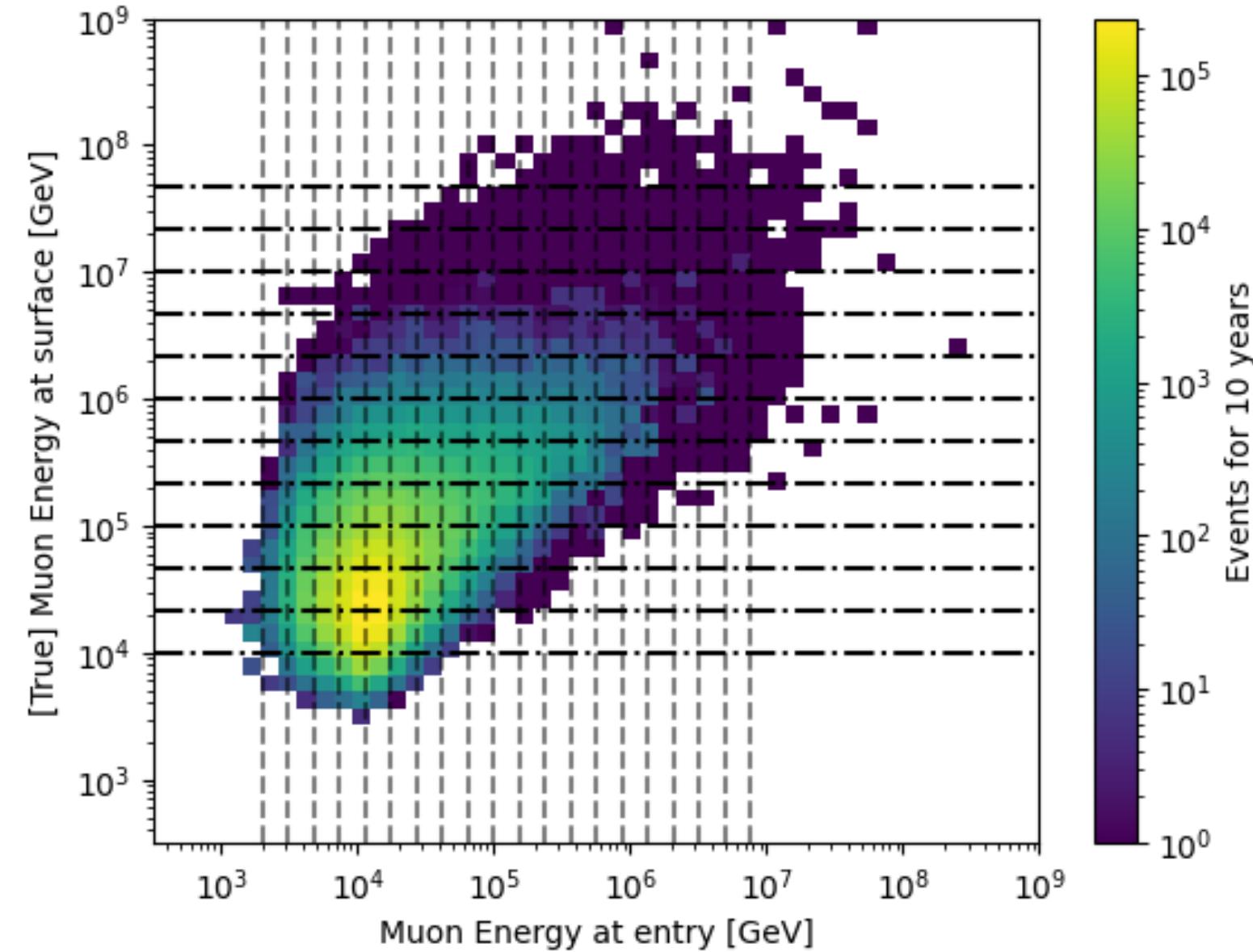
$$t(\vec{f}) = -\frac{1}{2} (\vec{C}\vec{f})^T (\tau_1)^{-1} (\vec{C}\vec{f})$$

4. maximize  $\log(\mathcal{L}(\vec{g}|\vec{f})) + t(\vec{f})$   
with respect to  $\vec{f}$  using  
Markov Chain Monte Carlo (MCMC)  
or Minuit



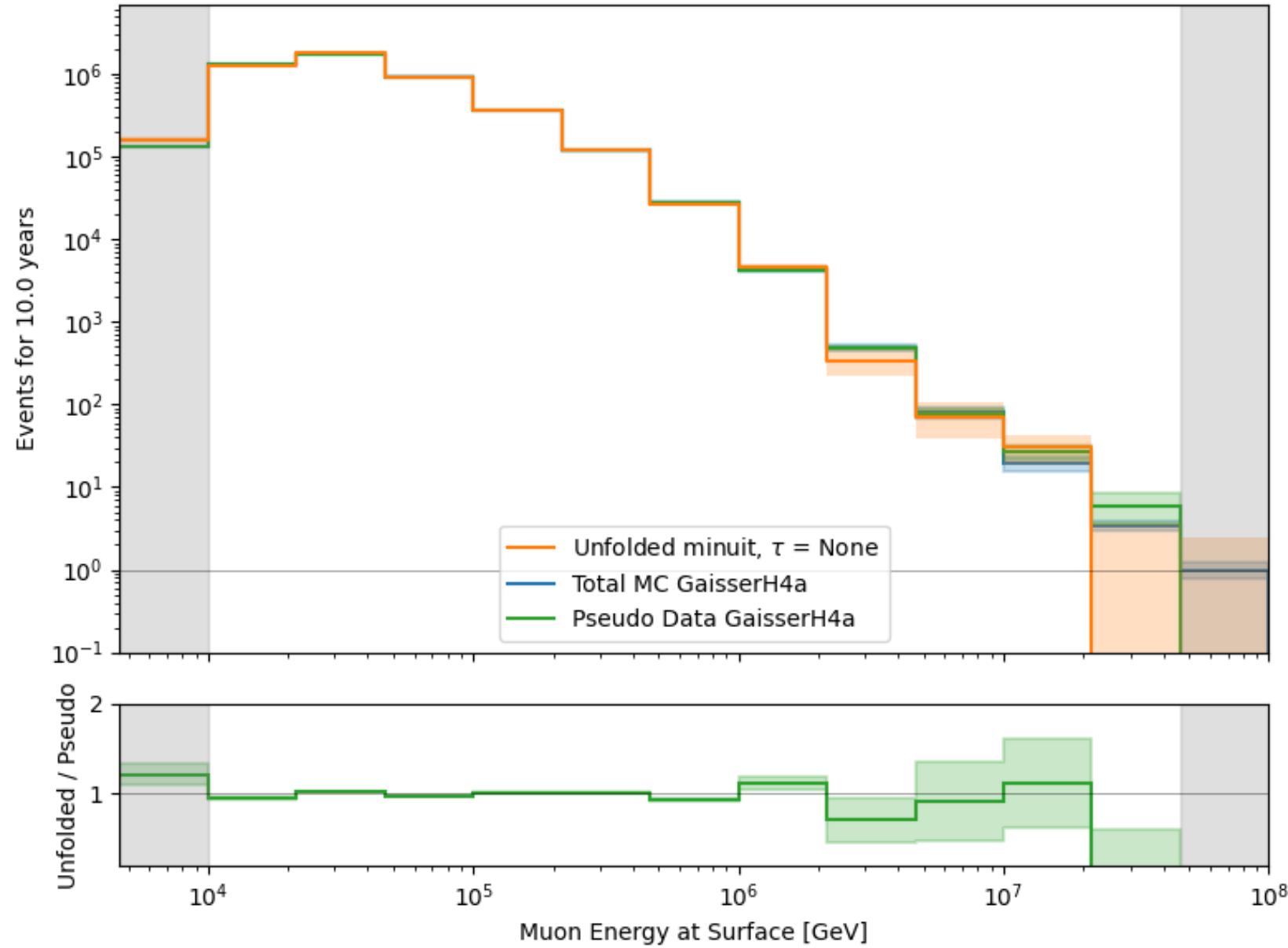
funfolding  
by M. Börner

## Proxy vs Target



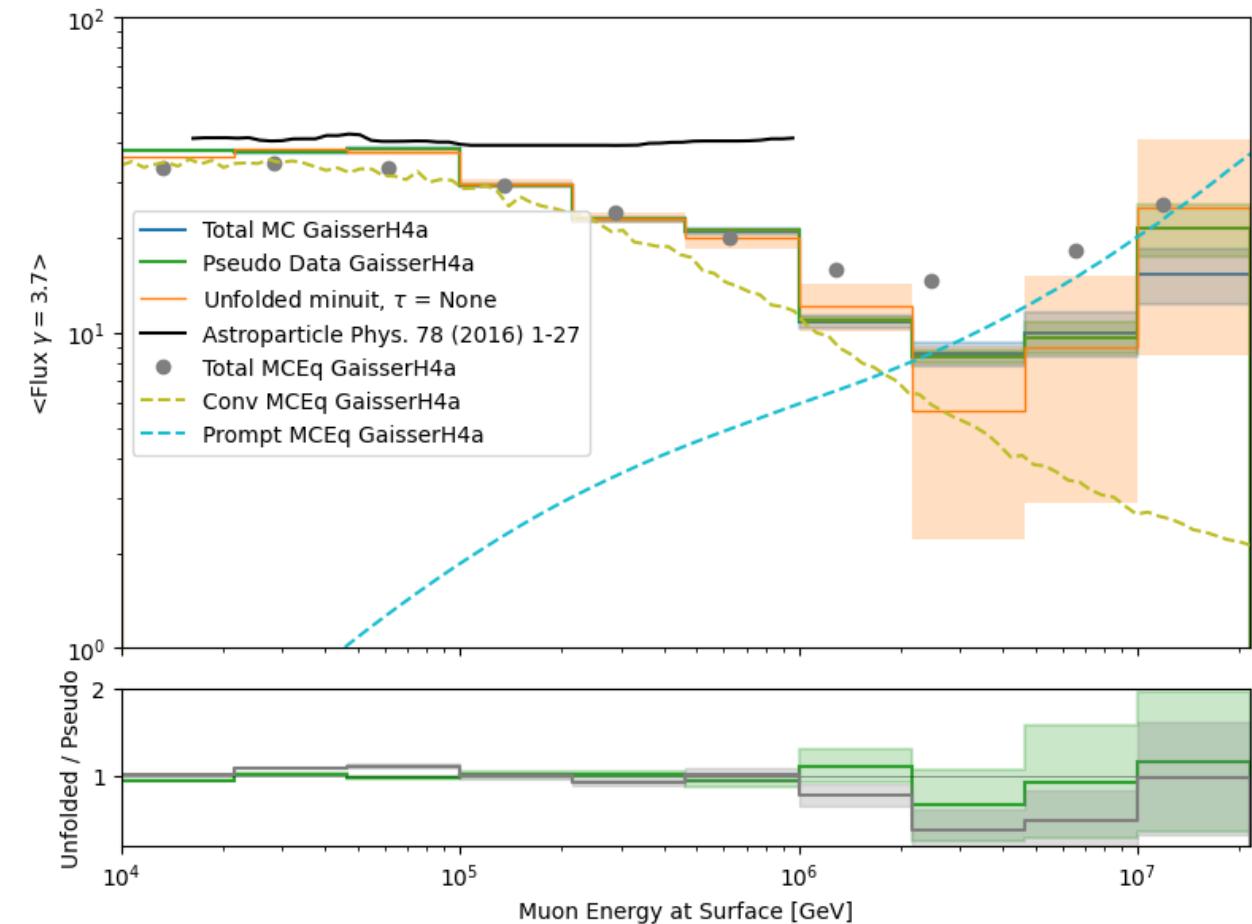
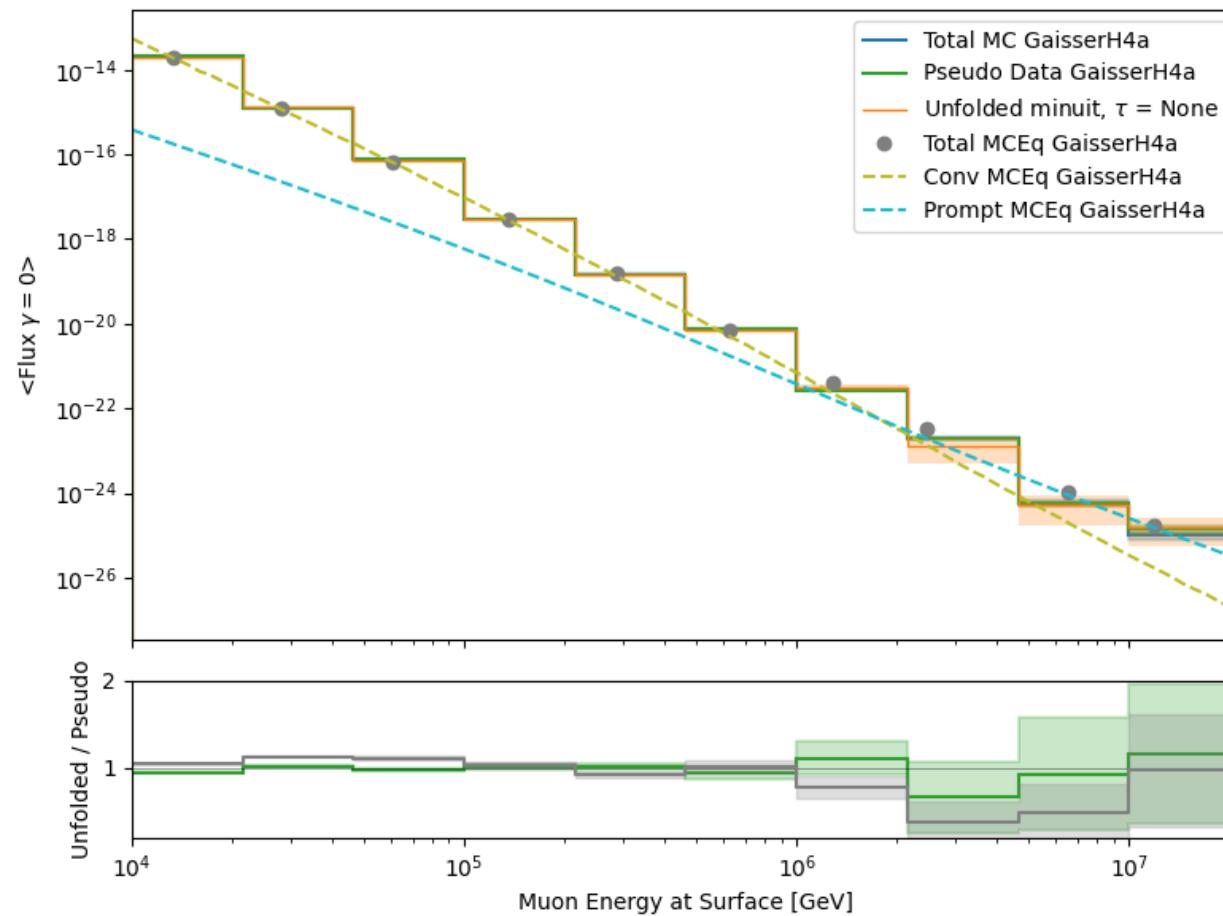
# Unfold Event Rate

## Event Rate



# Unfold Muon Flux

# Muon Flux at Surface

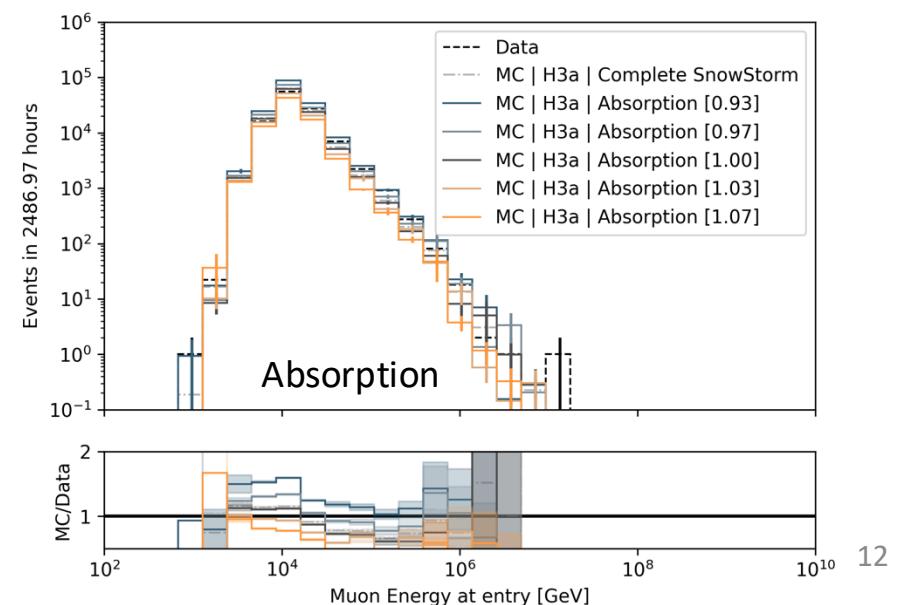
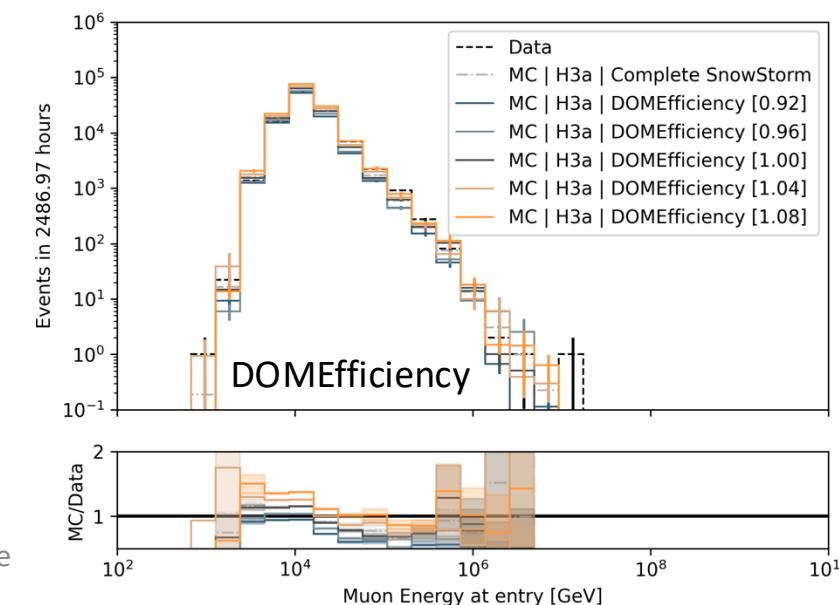
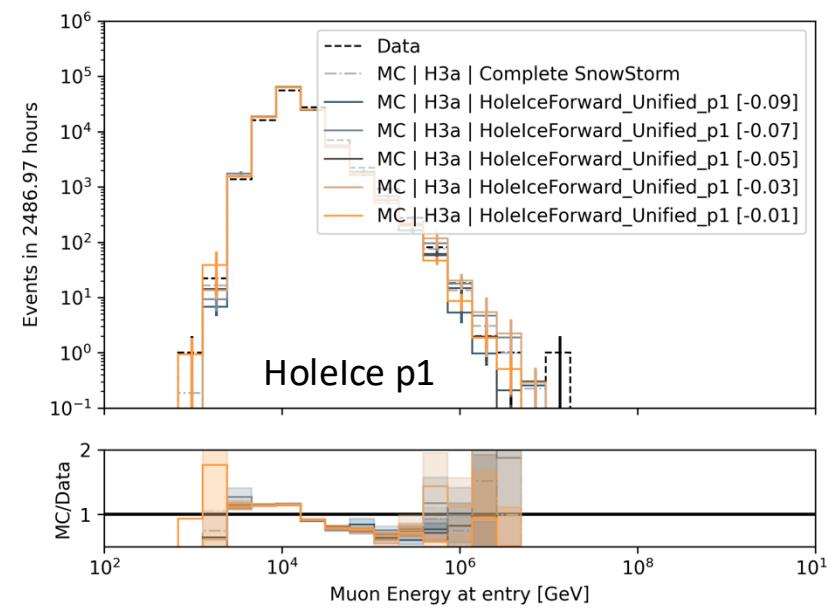
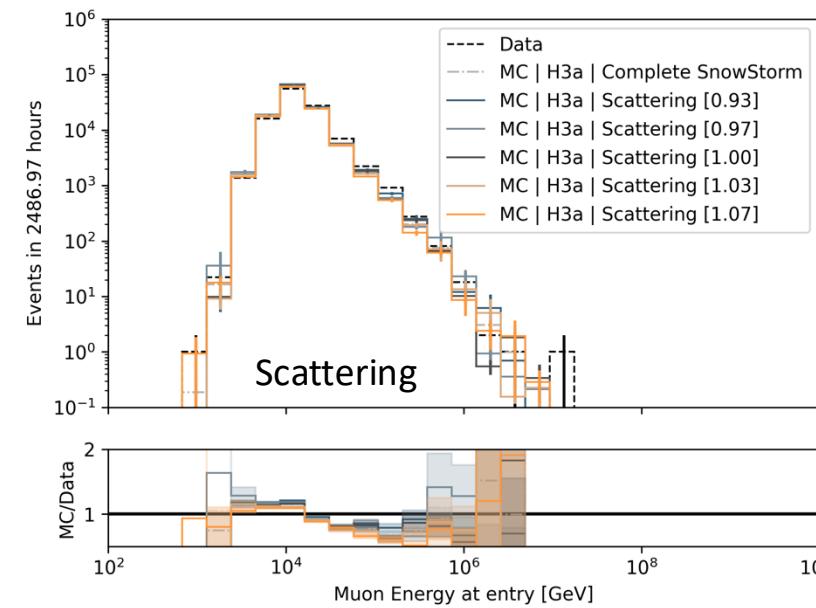
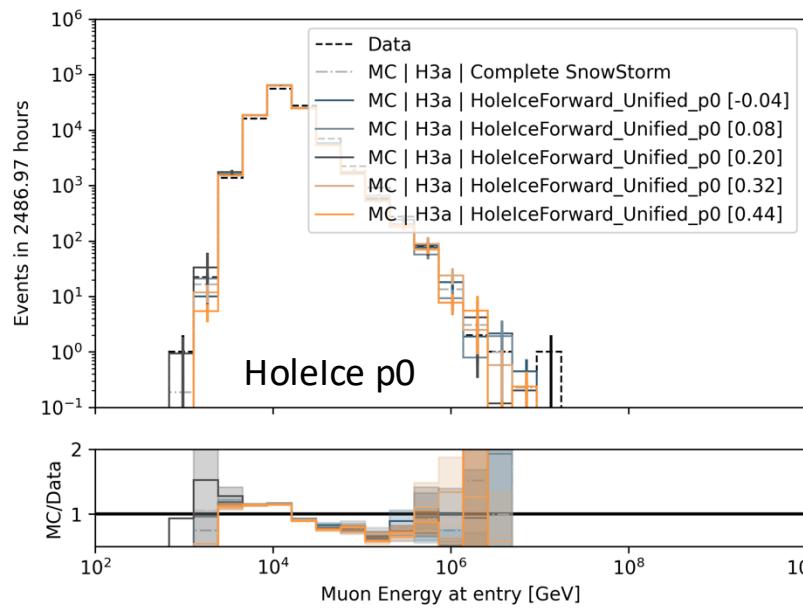


# Ice Systematics

## Snowstorm

1. Scattering
2. Absorption
3. DOMEfficiency
4. Hole Ice Forward p0
5. Hole Ice Forward p1

# Systematics Impact on Muon Energy at Entry (Proxy)

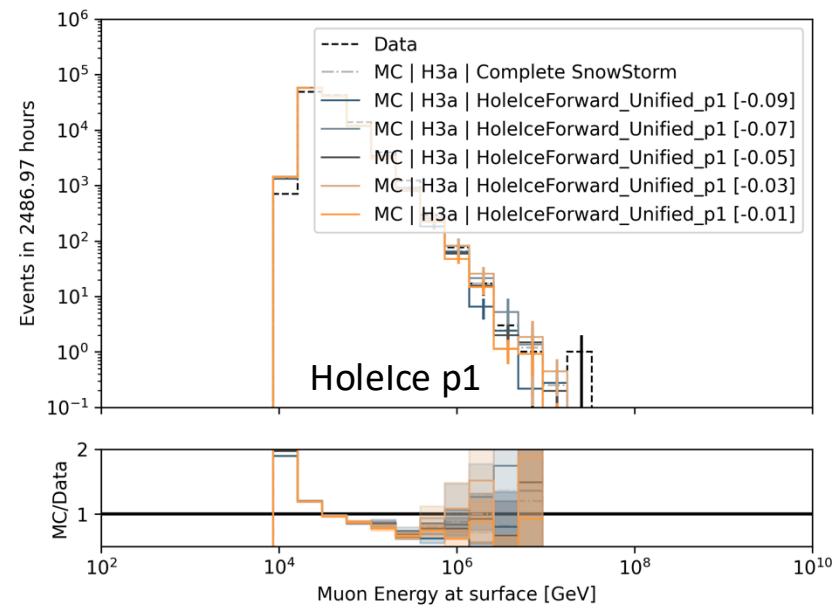
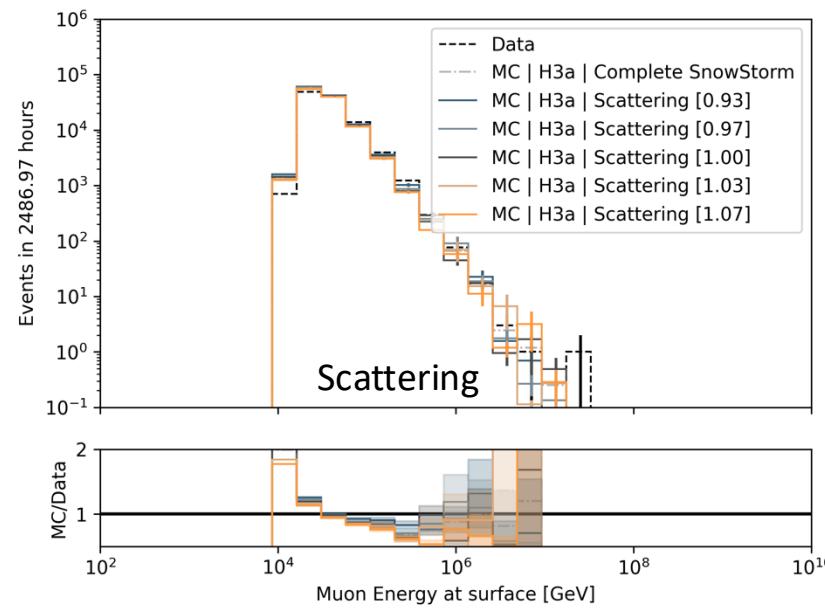
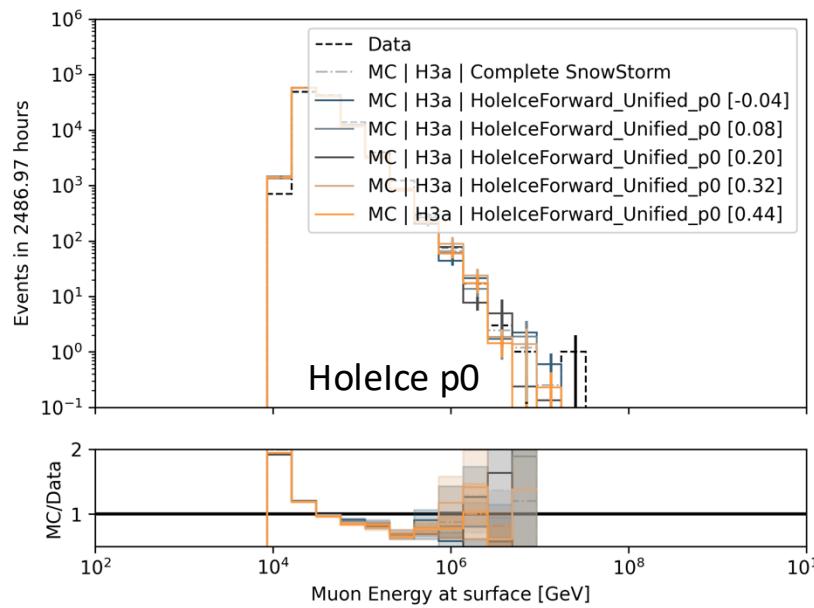


- Divide systematic range into 5 equidistant bins
- Largest impact by DOMEfficiency and Absorption

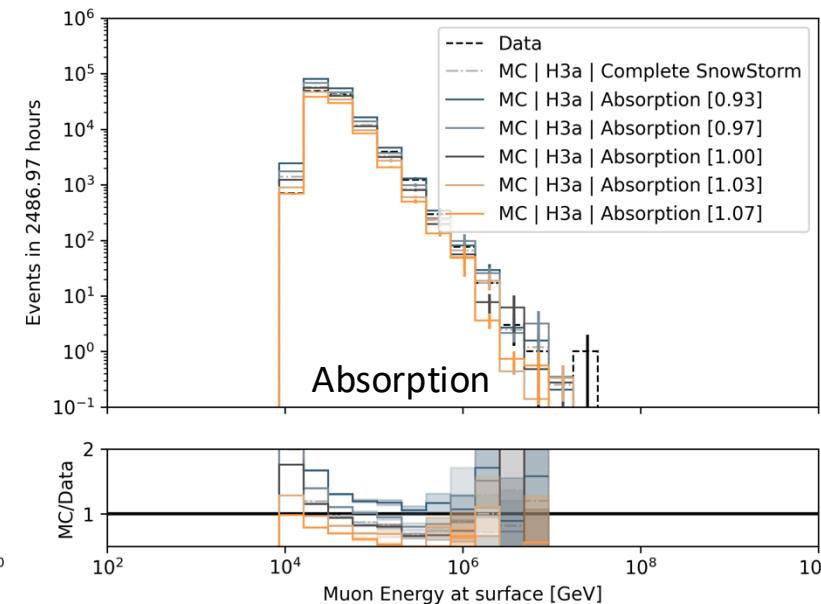
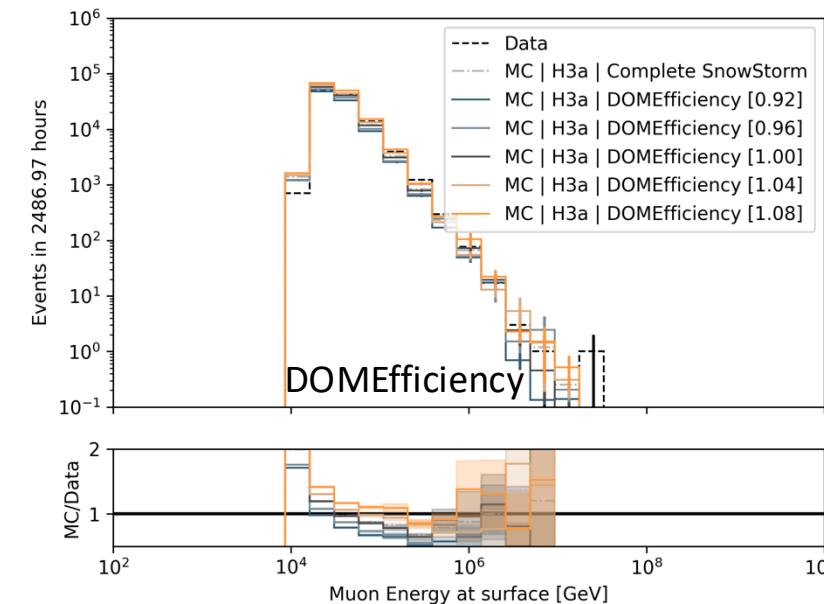
pascal.gutjahr@tu-dortmund.de

12

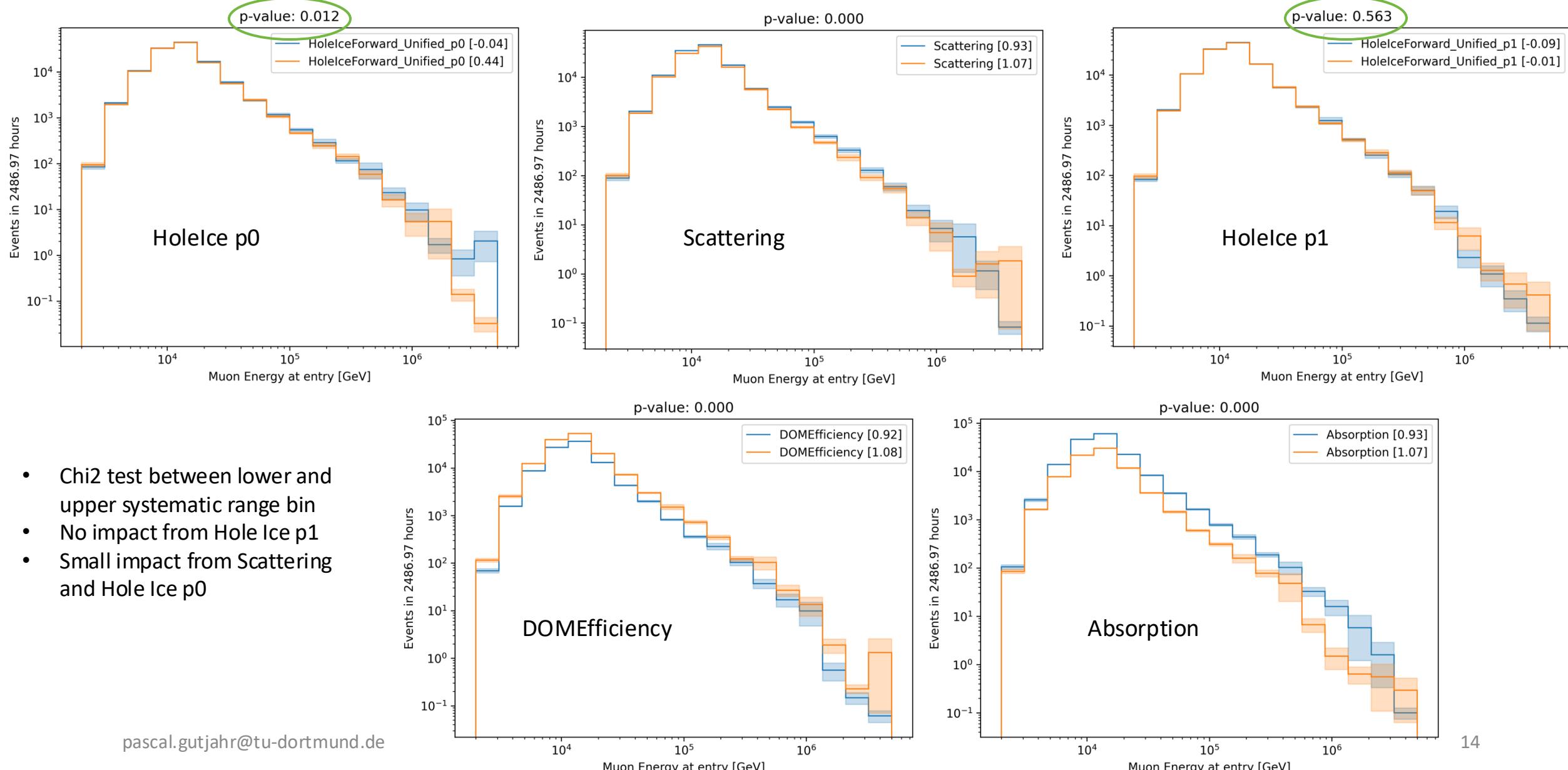
# Systematics Impact on Muon Energy at Surface (Target)



- Divide systematic range into 5 equidistant bins
- Largest impact by DOMEfficiency and Absorption



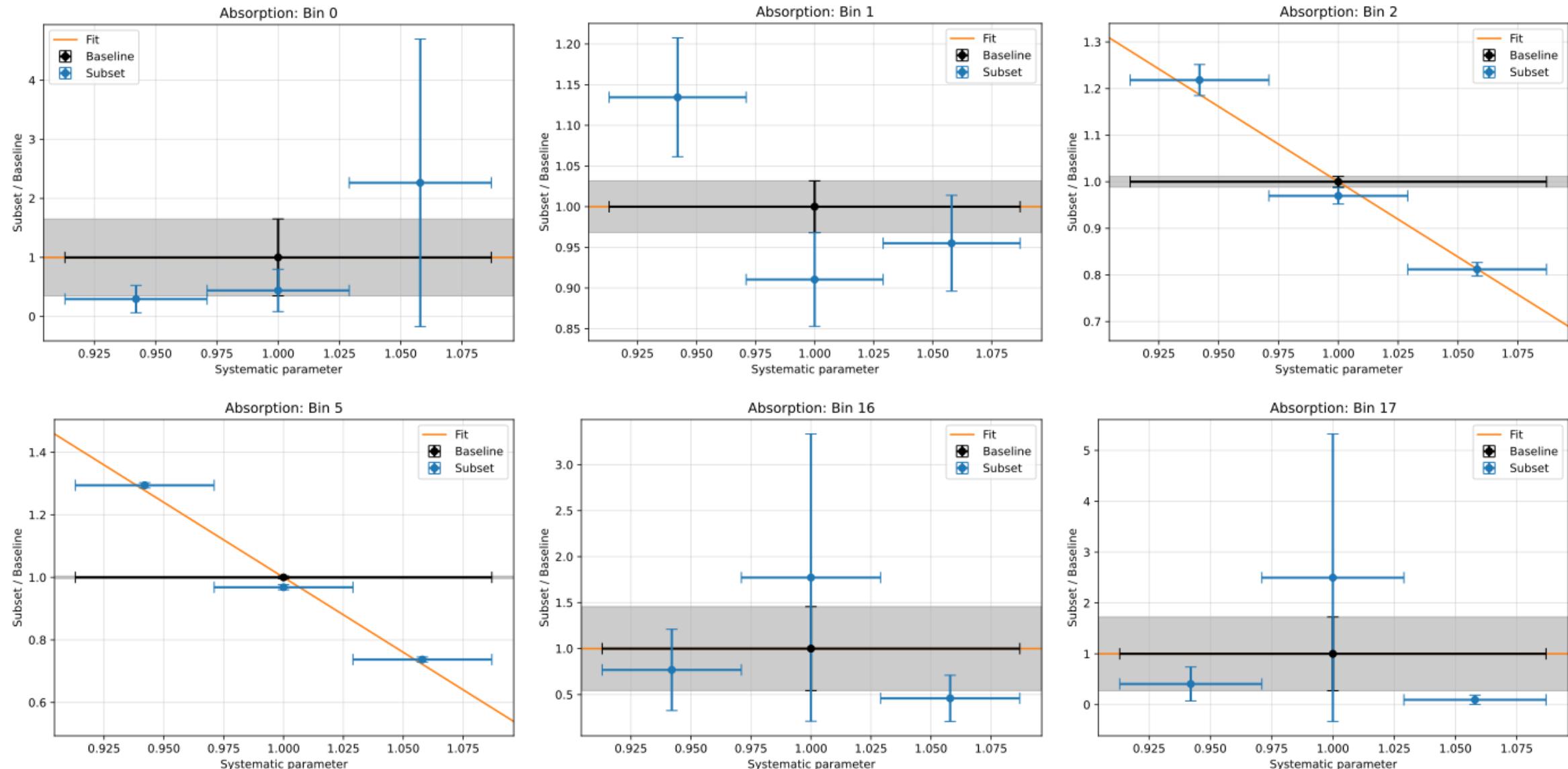
# Chi2 Test: Muon Energy at Entry (Proxy)



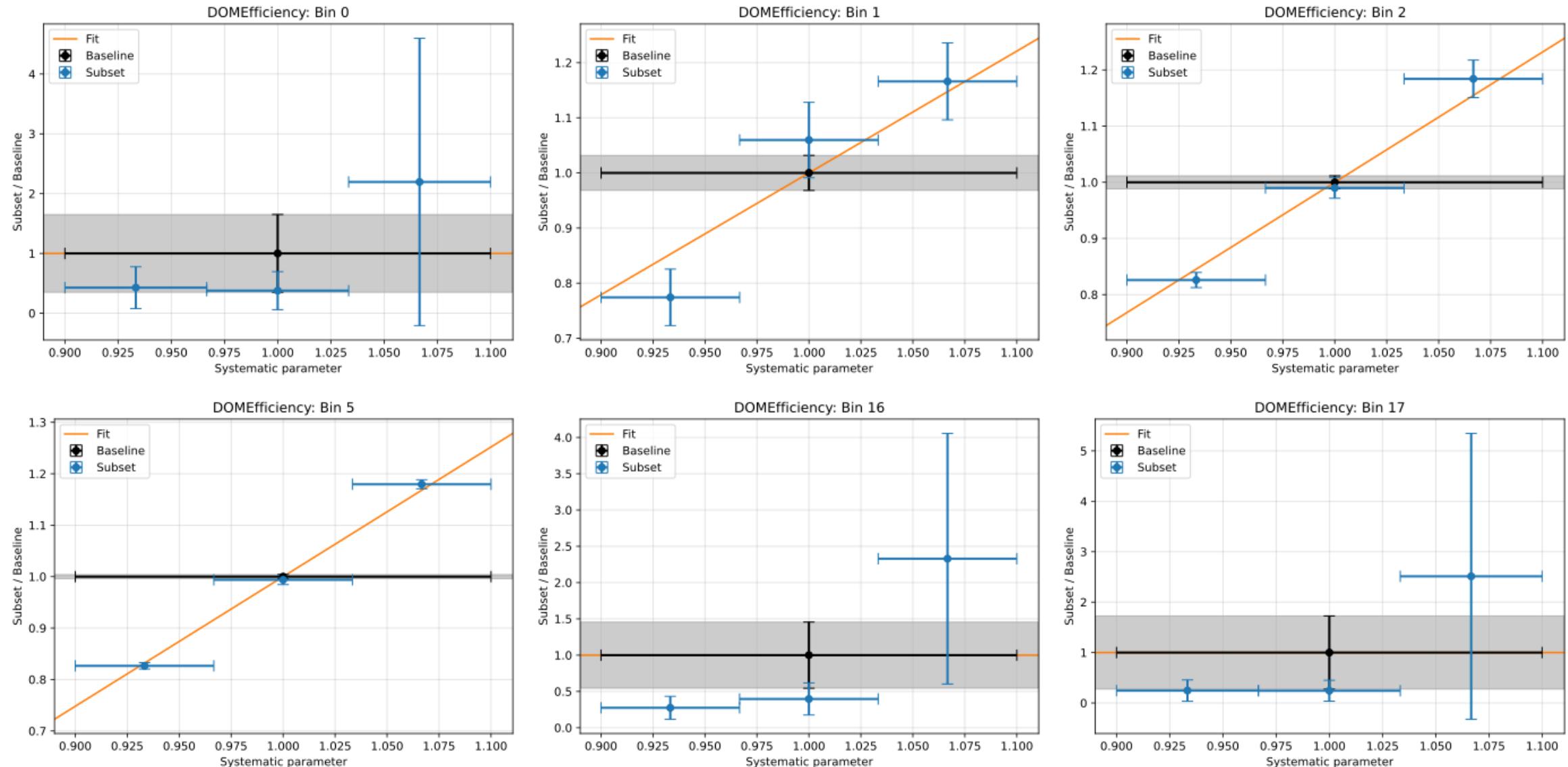
# Parameterize Systematics

- Vary each systematic parameter for each proxy bin  
(18 proxy bins, 3 systematic bins)
  - Fit the relative change of the bin content
1. Chi2 test, if constant  $y = 1$  is compatible with data points ( $p\text{-value} = 0.05$ )
  2. Find linear fit with the lowest slope that is compatible with  $p\text{-value} = 0.2$ 
    1. Require fit is going through (center, 1)

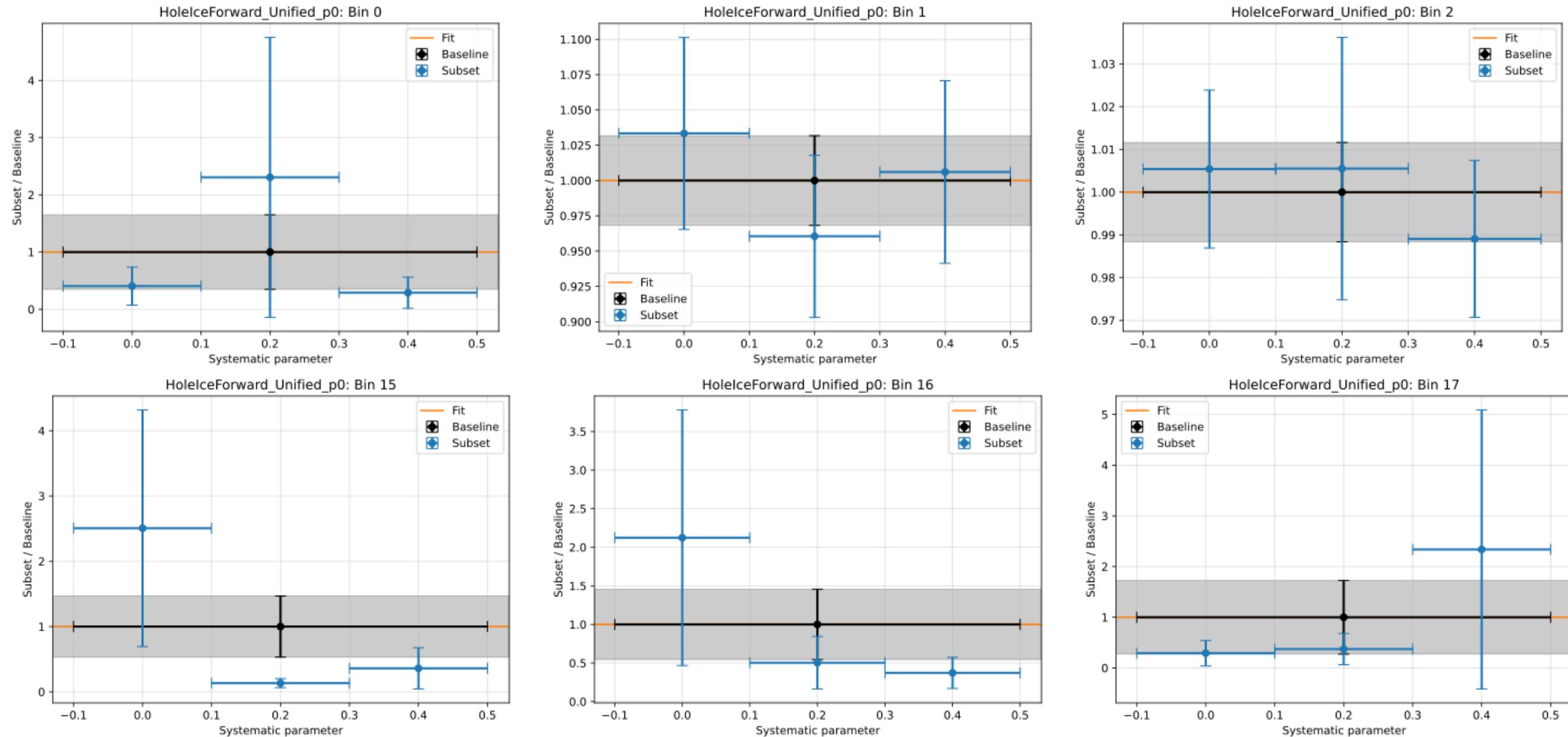
# Absorption



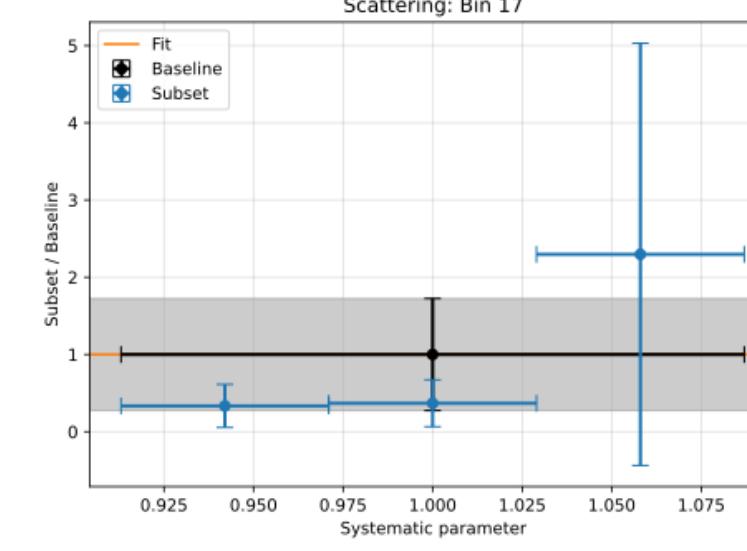
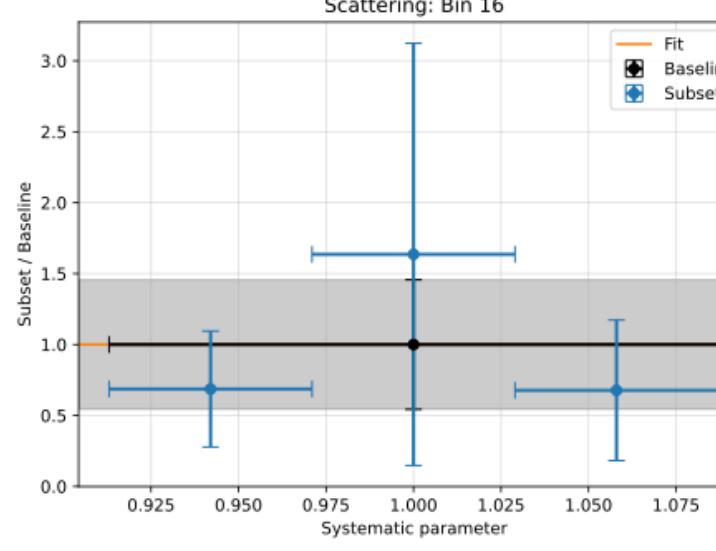
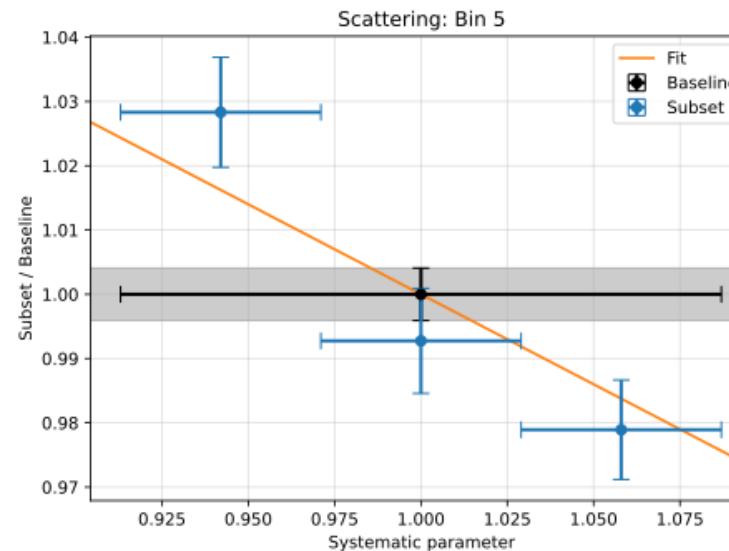
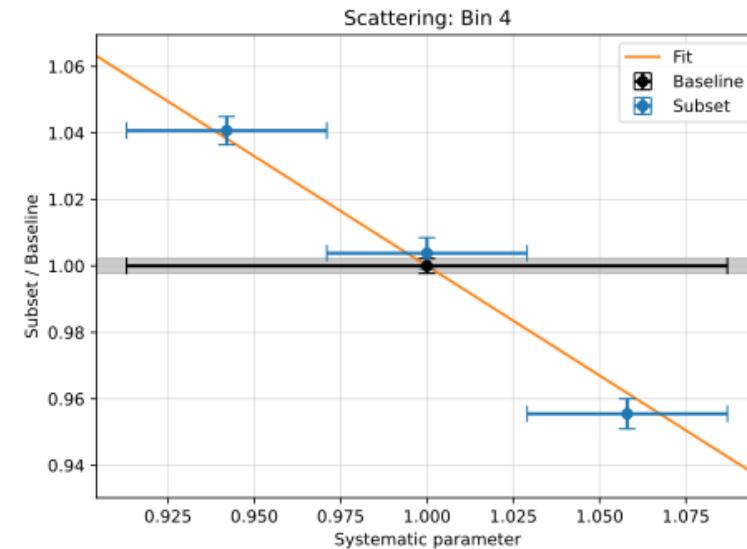
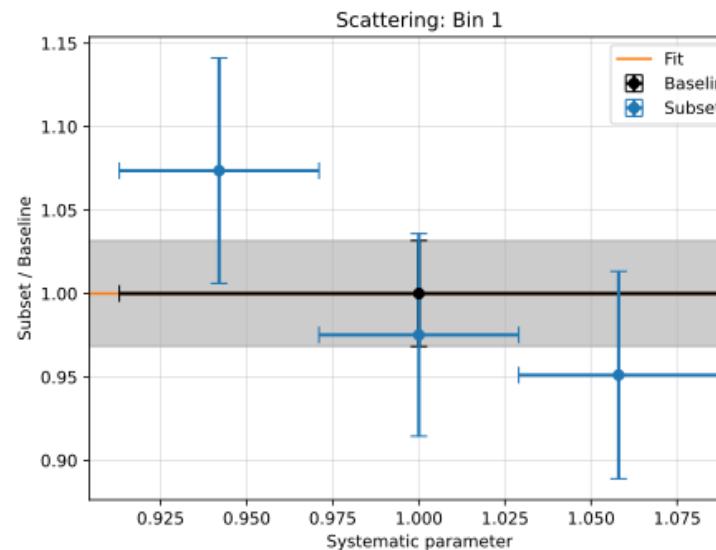
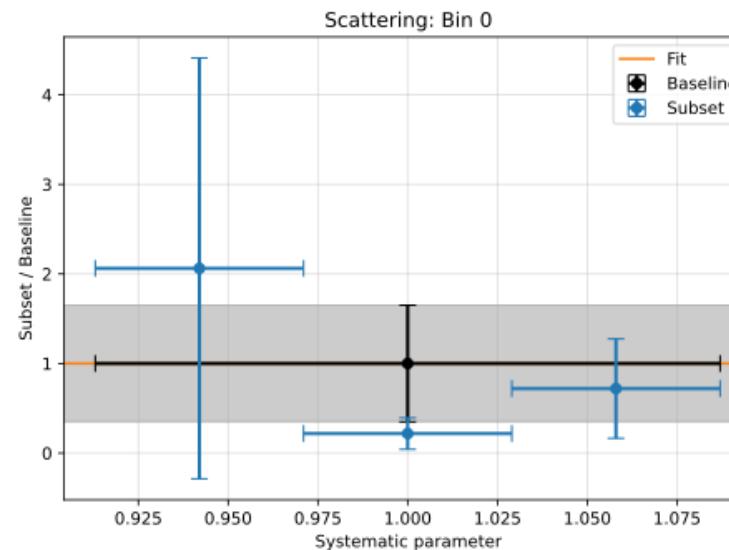
## DOMEfficiency



# Hole Ice p0 → not sensitive enough (same for Hole Ice p1)



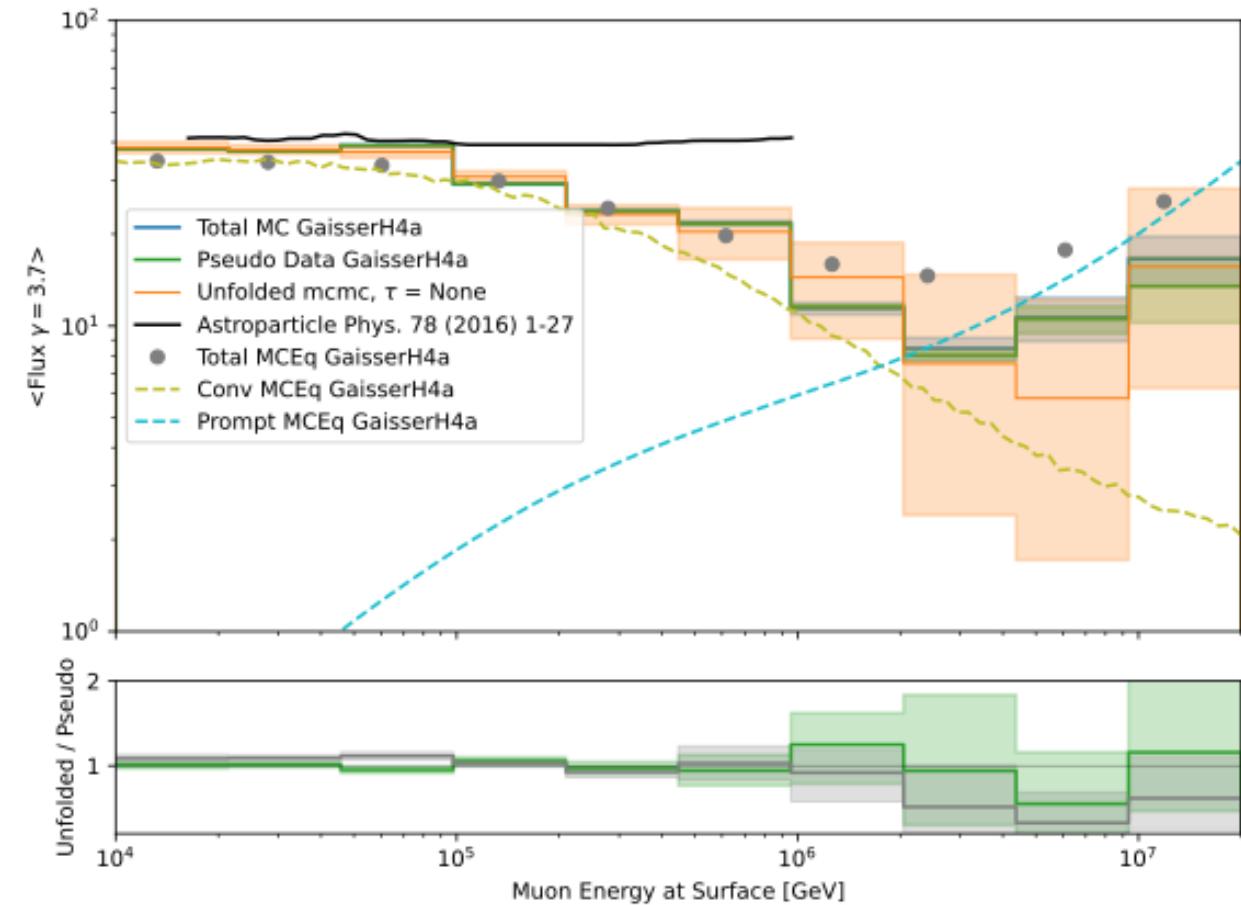
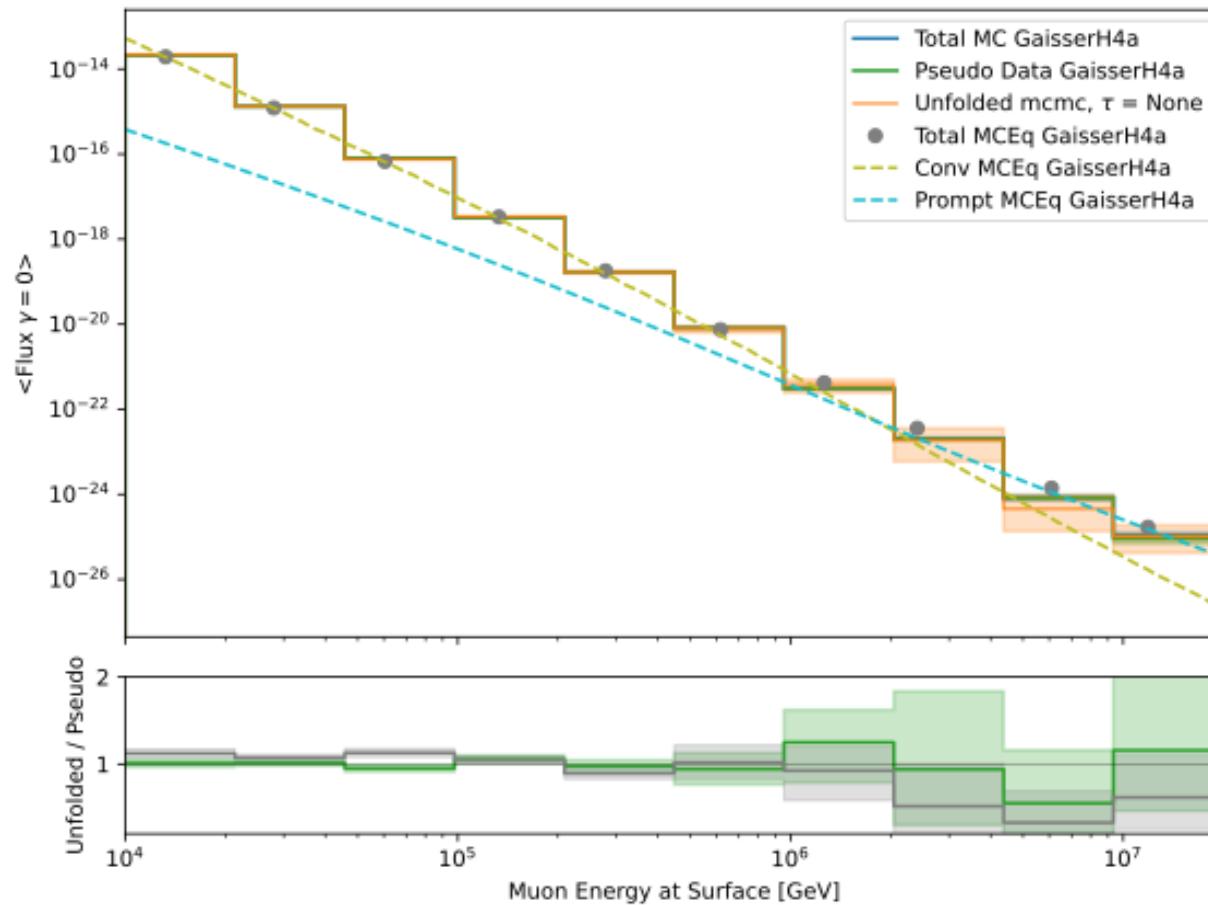
# Scattering



# Unfold Muon Flux with Systematics

No regularization

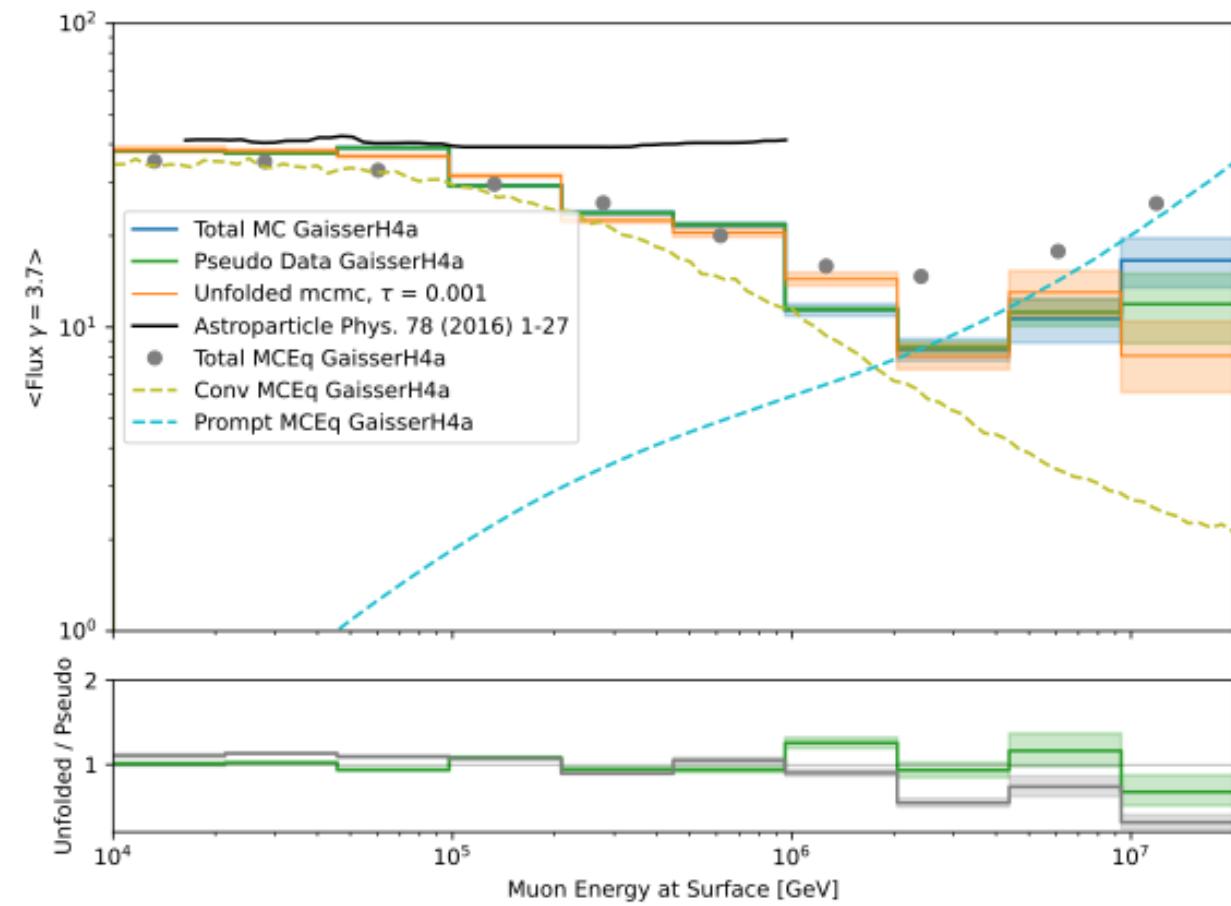
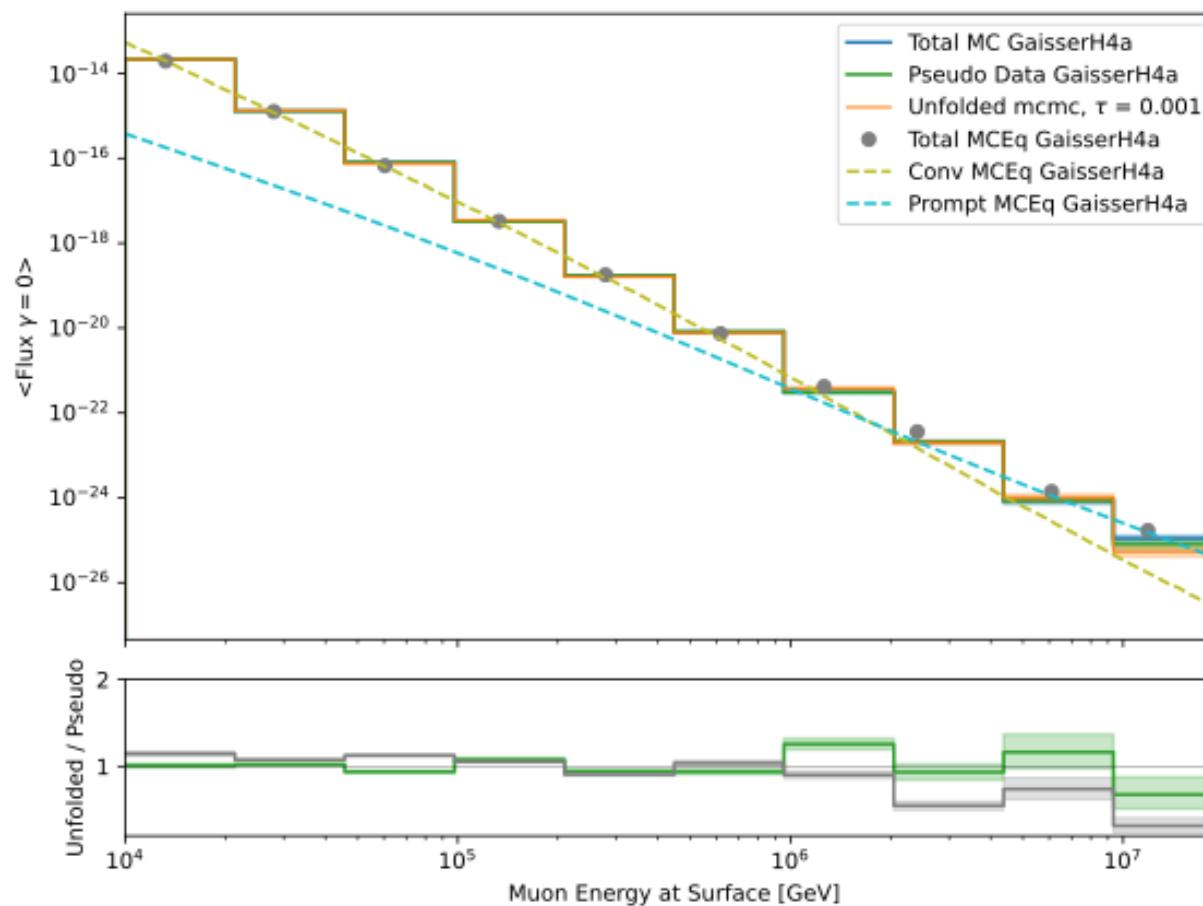
# Muon Flux at Surface (no regularization)

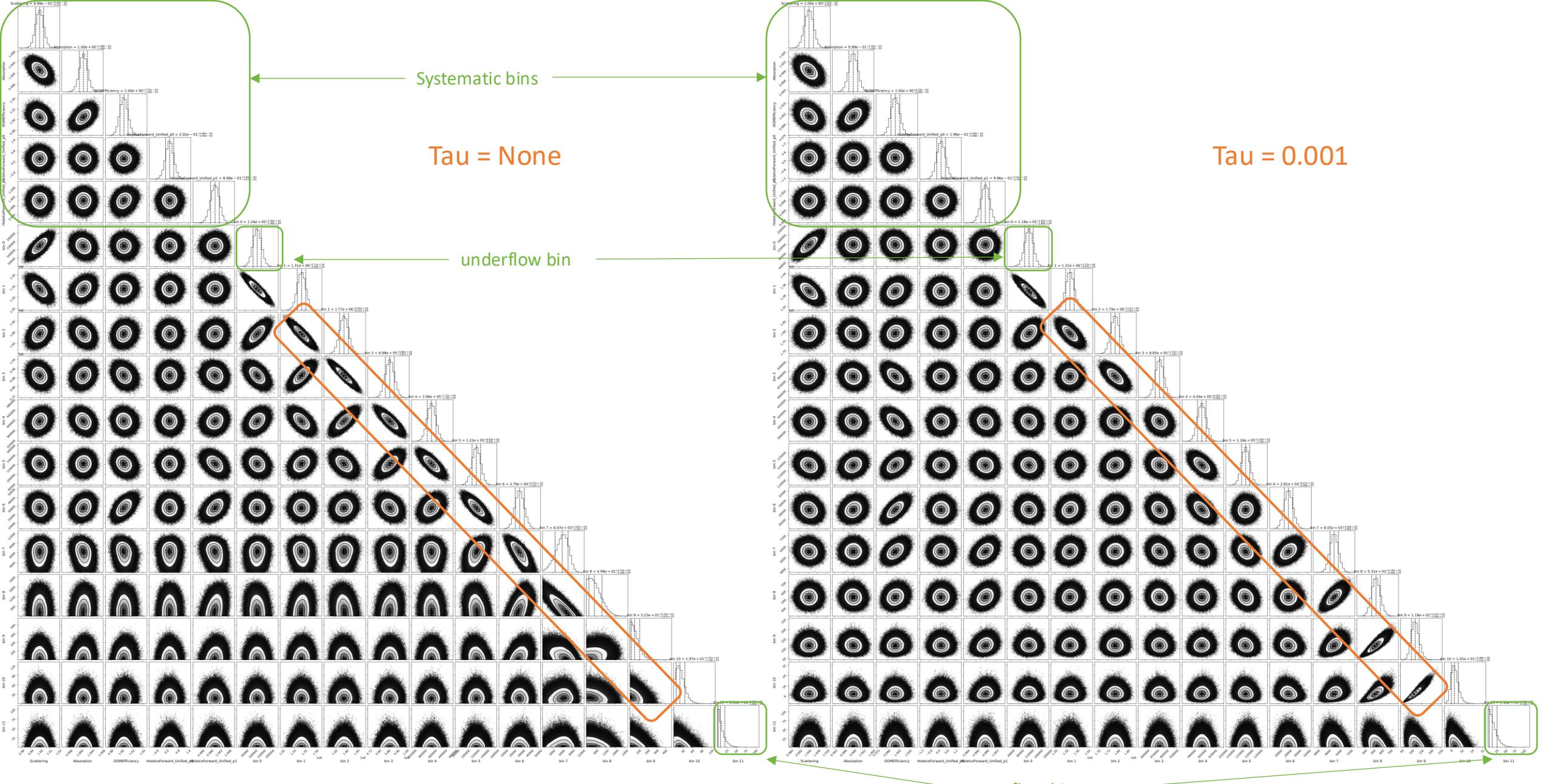


# Unfold Muon Flux with Systematics

With regularization

# Muon Flux at Surface ( $\tau = 0.001$ )





- Regularization minimizes the correlation
- Find tau with the minimum correlation

# Conclusion

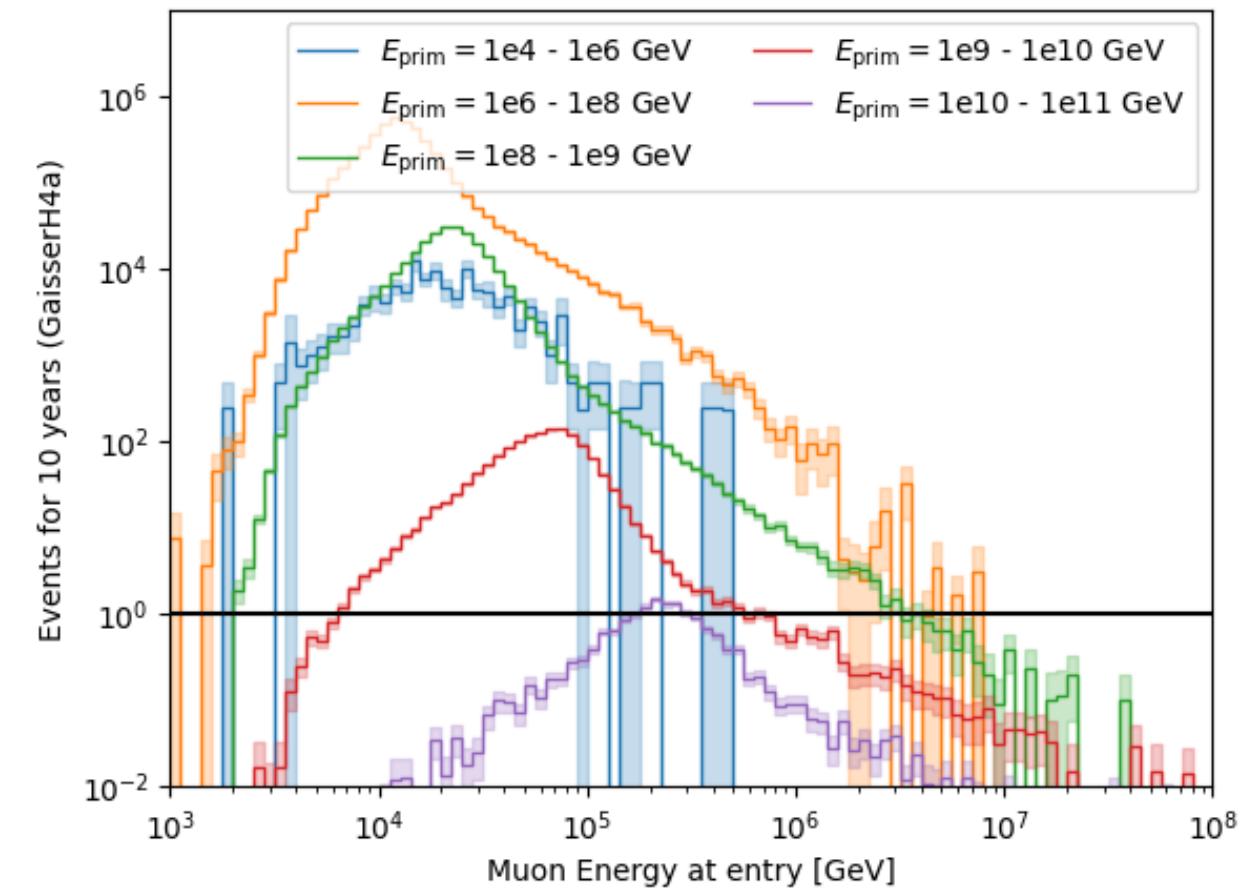
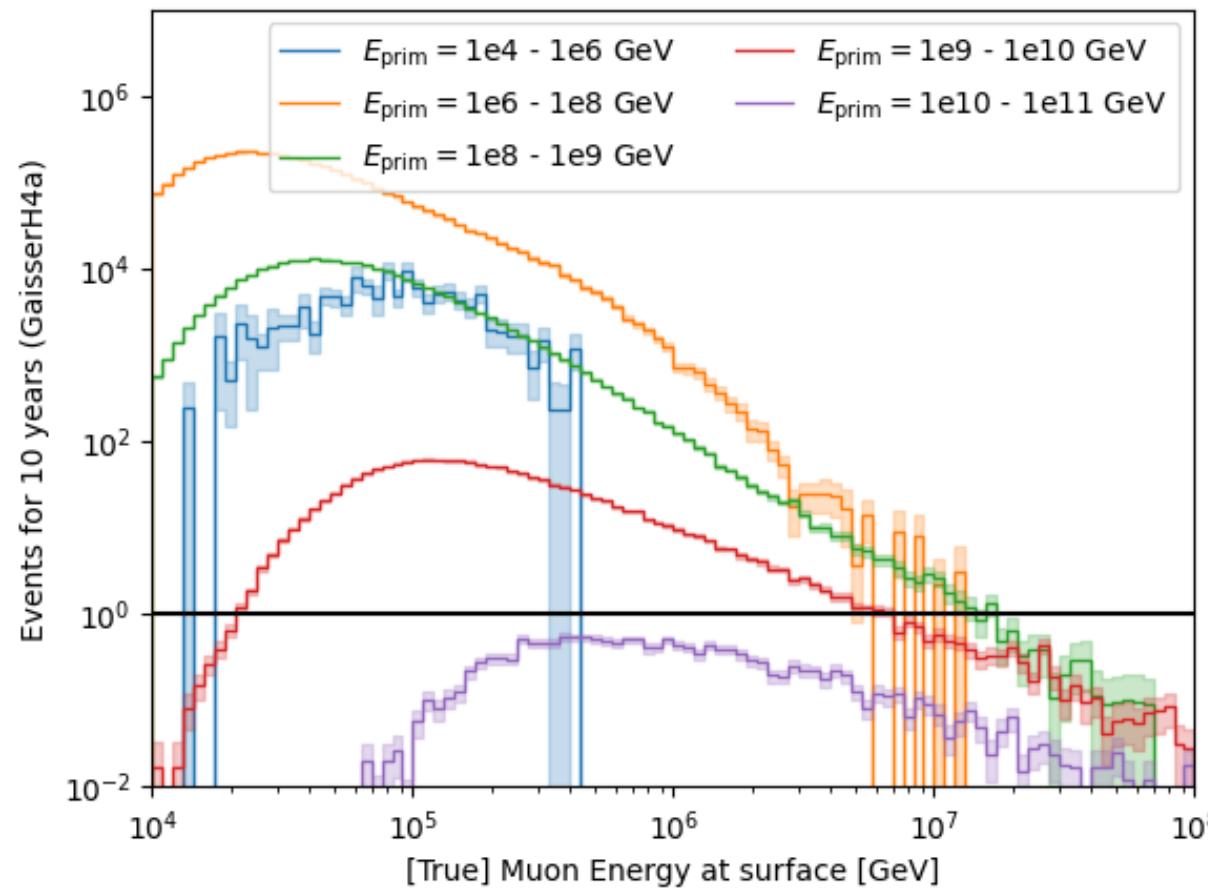
- Unfold muon event rate 
- Unfold muon flux at surface 
- Parameterize ice systematics 
  - Small impact from Scattering, Hole Ice p0 , and Hole Ice p1
  - Linear parameterization of Absorption and DOMEfficiency
- Unfold muon flux with systematics 

## Next steps

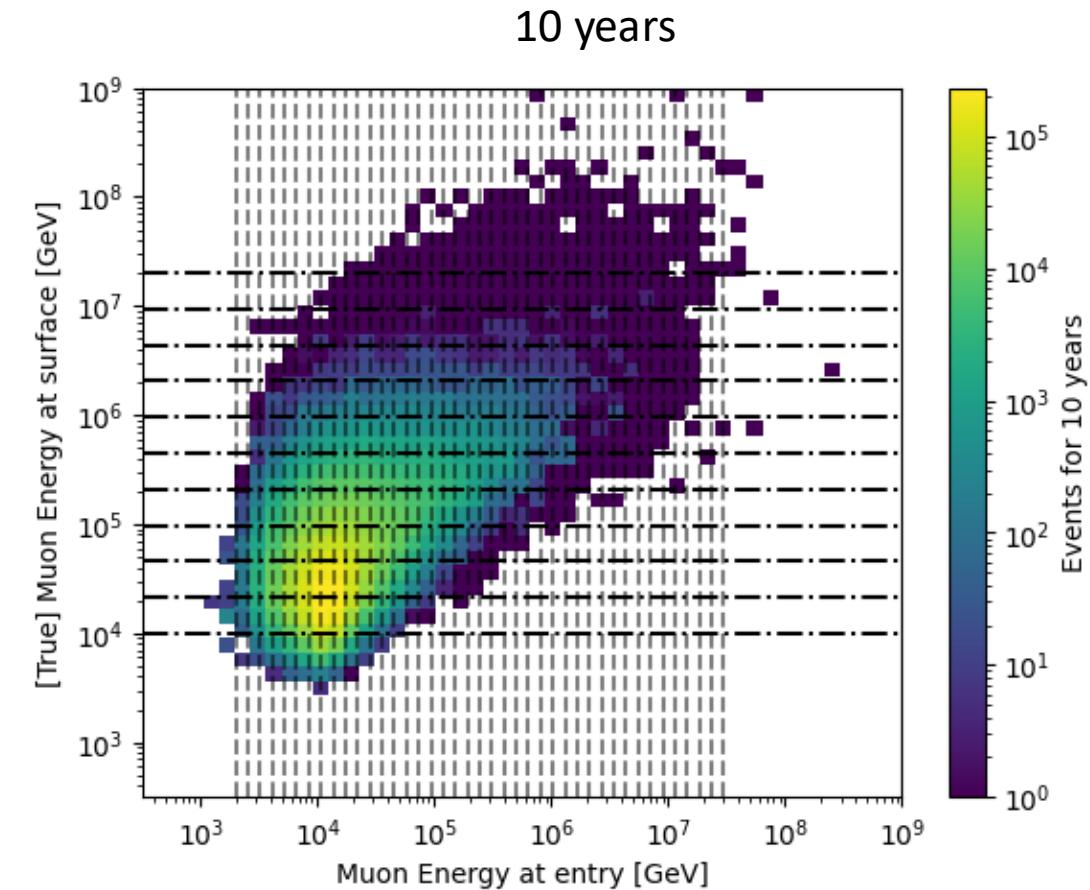
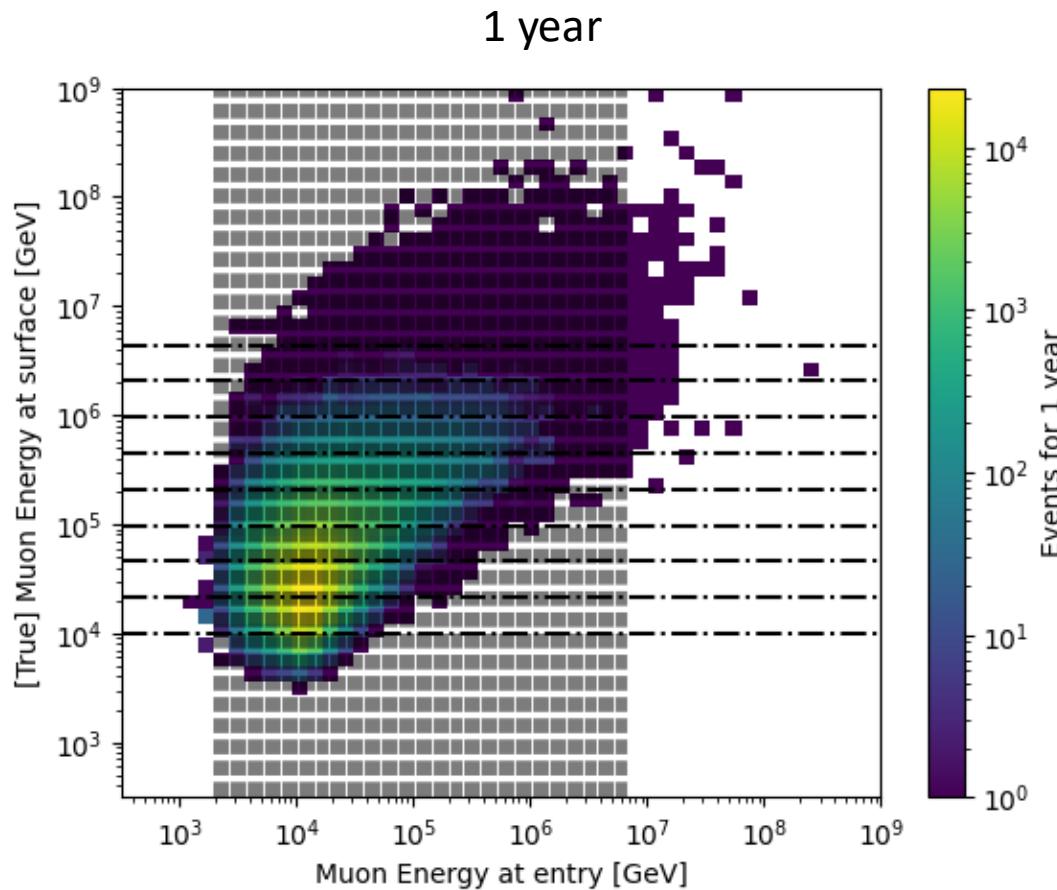
- Test impact of different primary models (spectral index)
- Choose optimal regularization

# Backup

# Relation between primary energy and muon energy at surface and entry

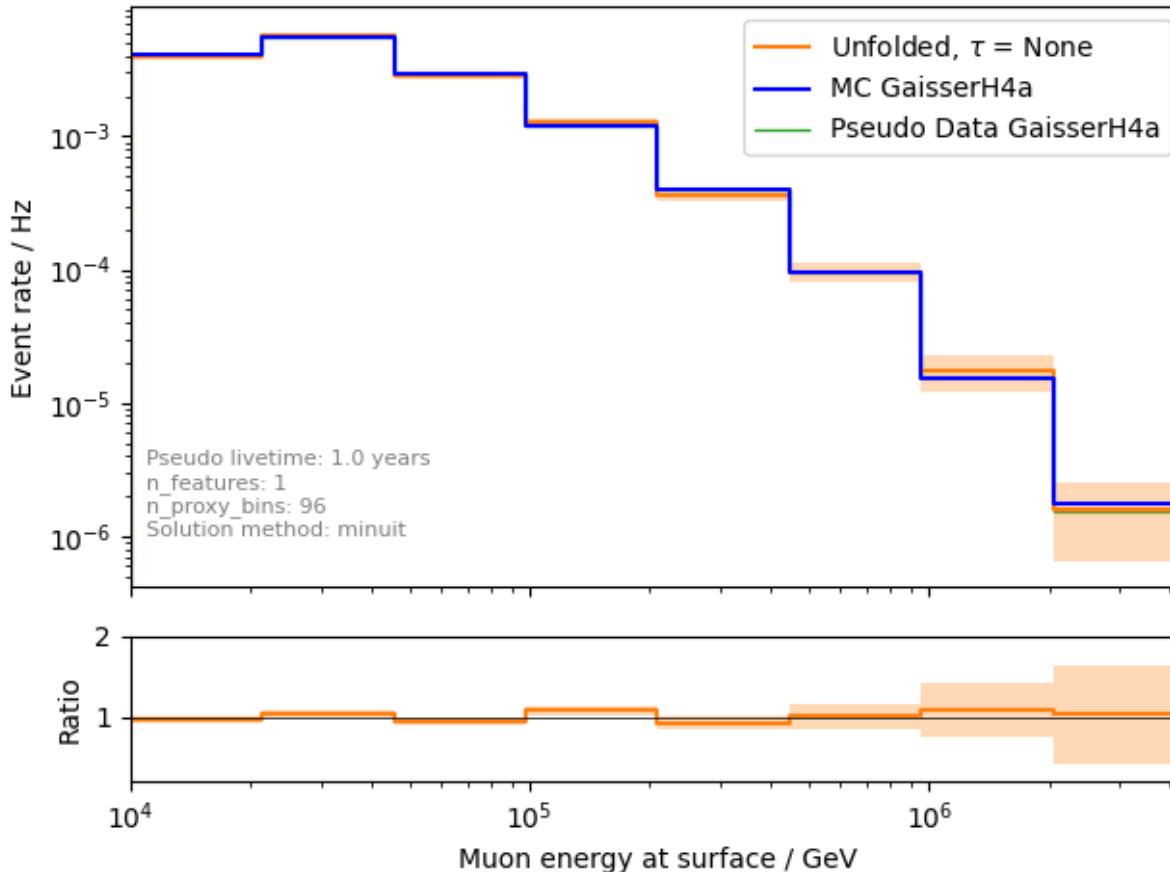


## Proxy vs. target

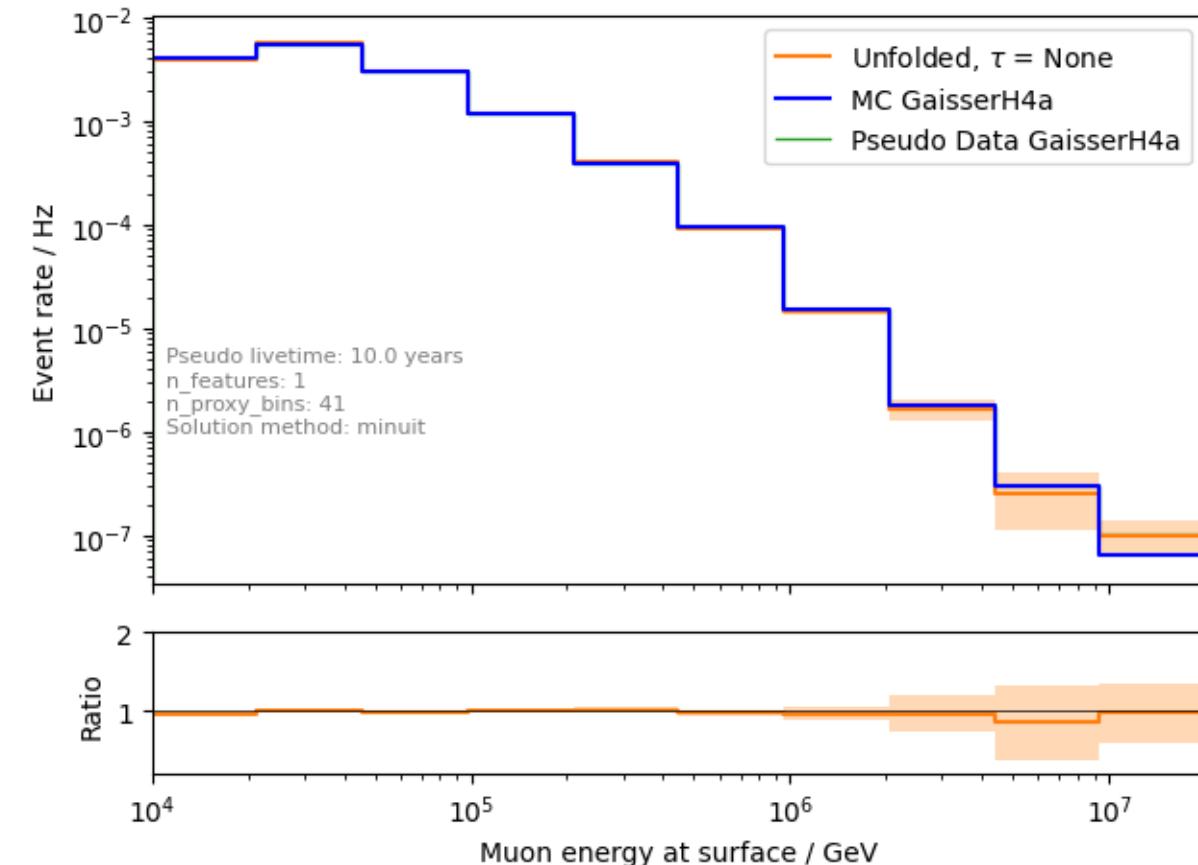


# Unfolding event rate

1 year

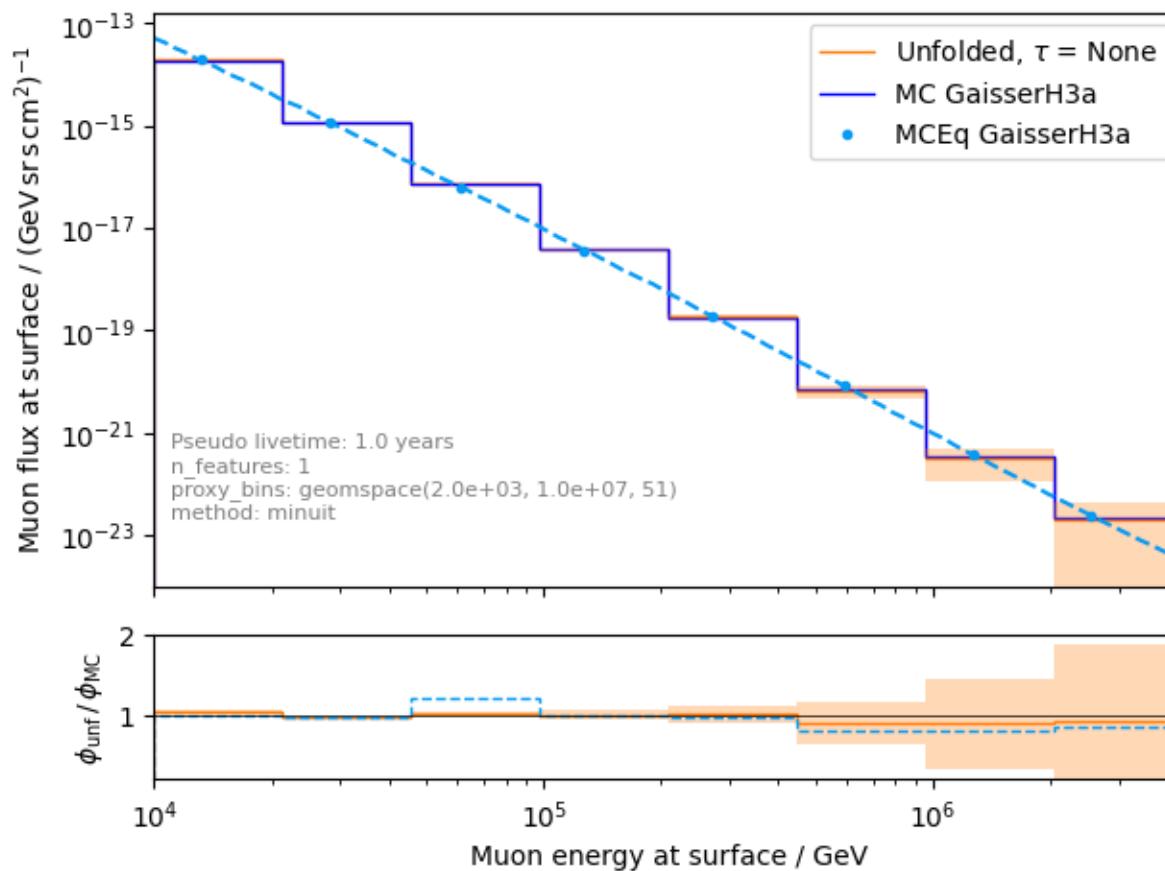


10 years

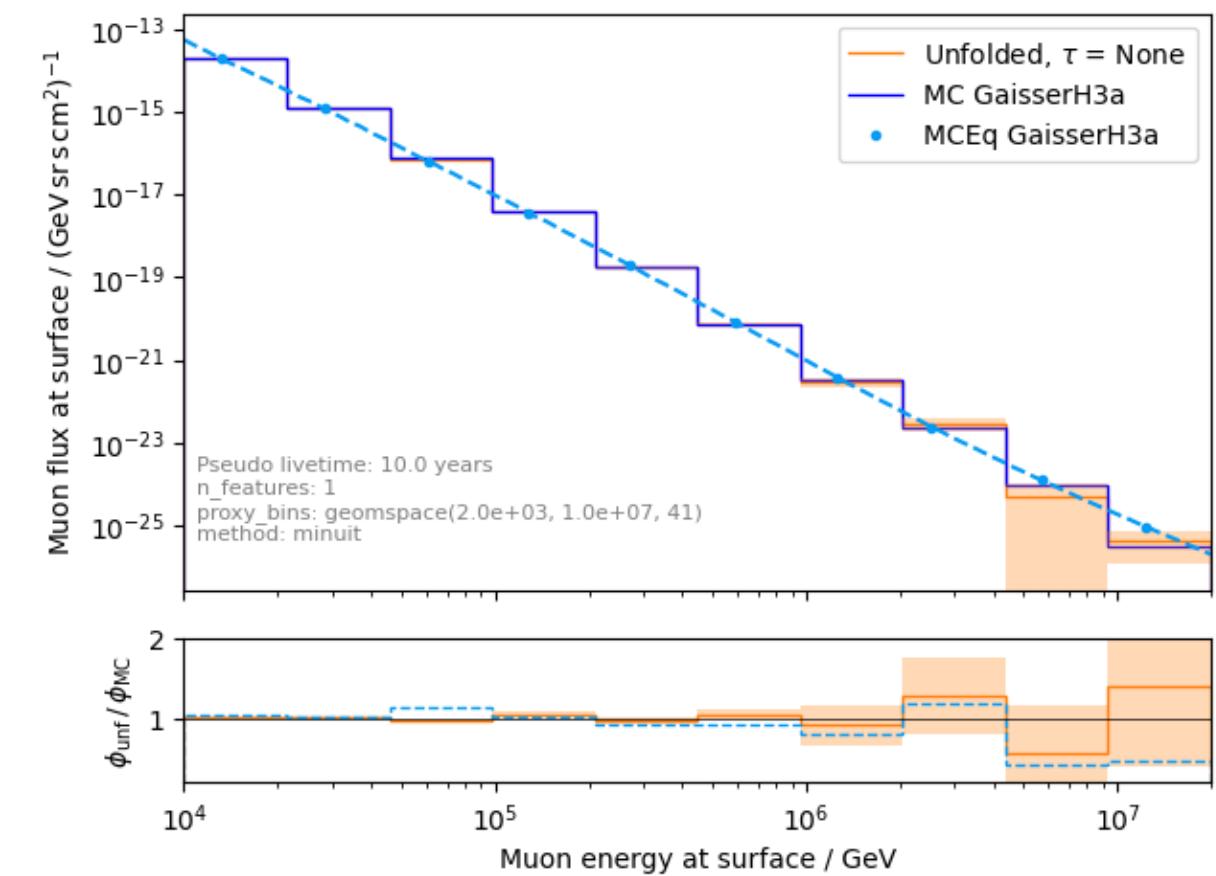


## unfolding of muon flux at surface

1 year



10 years



# Ice systematics

Snowstorm

Absorption

Scattering

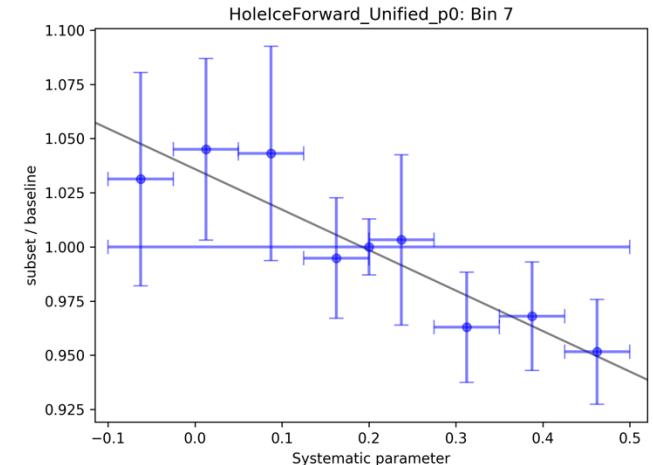
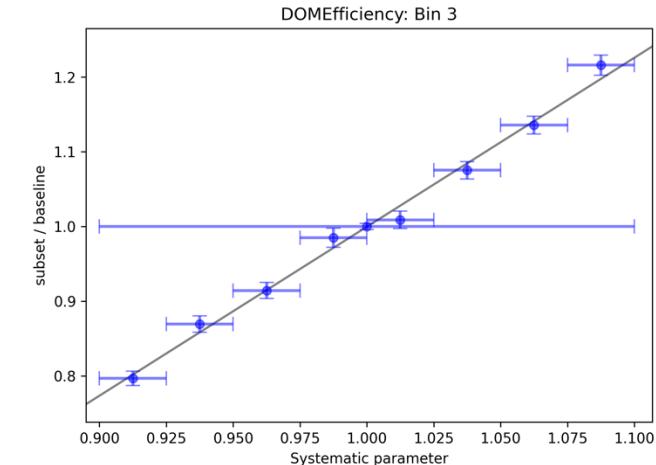
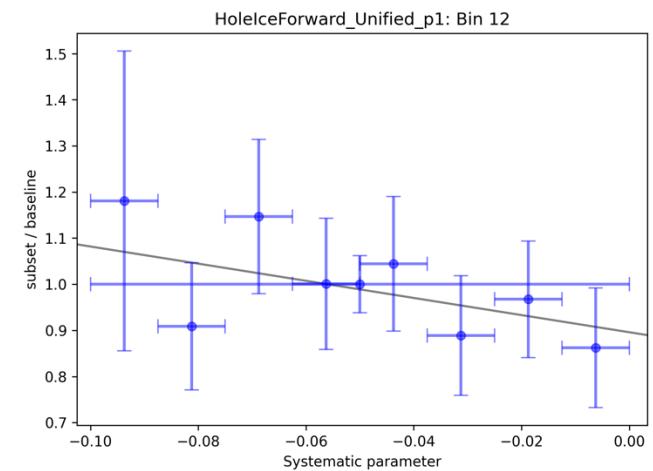
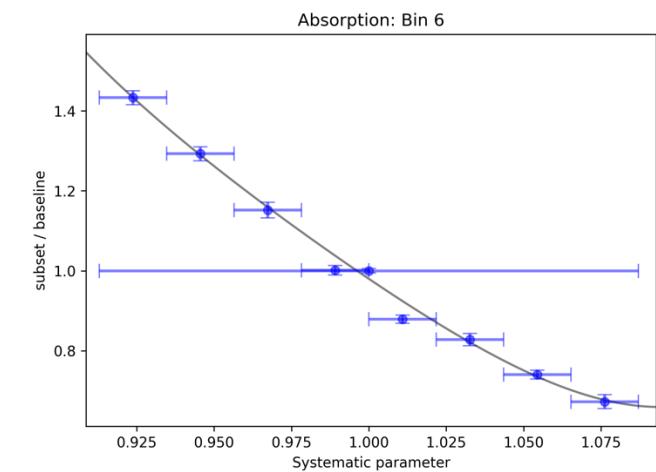
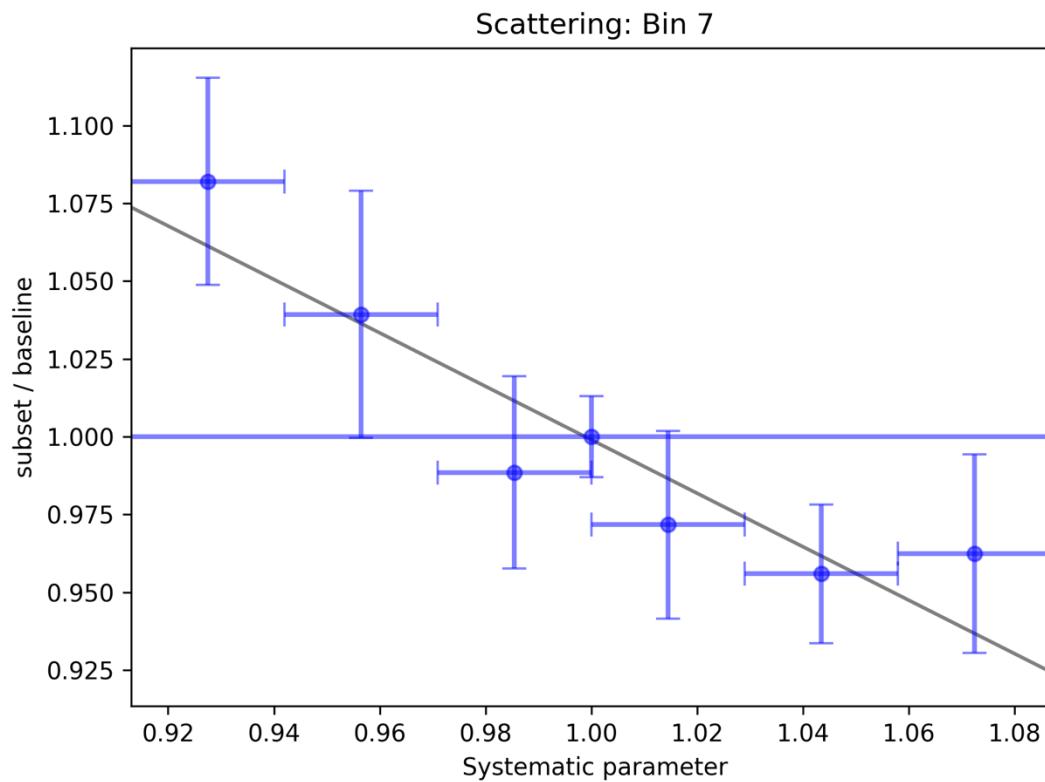
DOMEfficiency

HoleIce 0

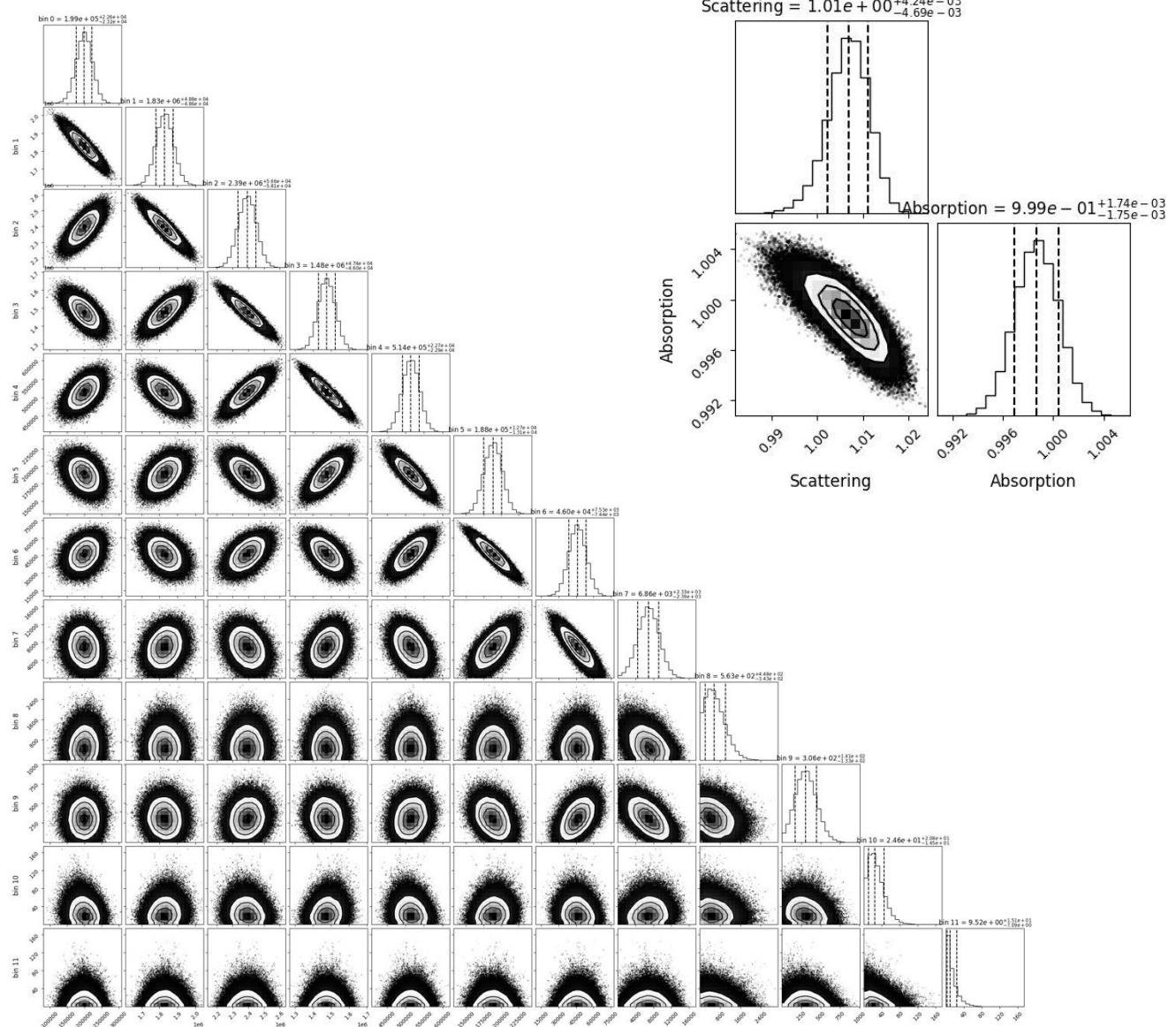
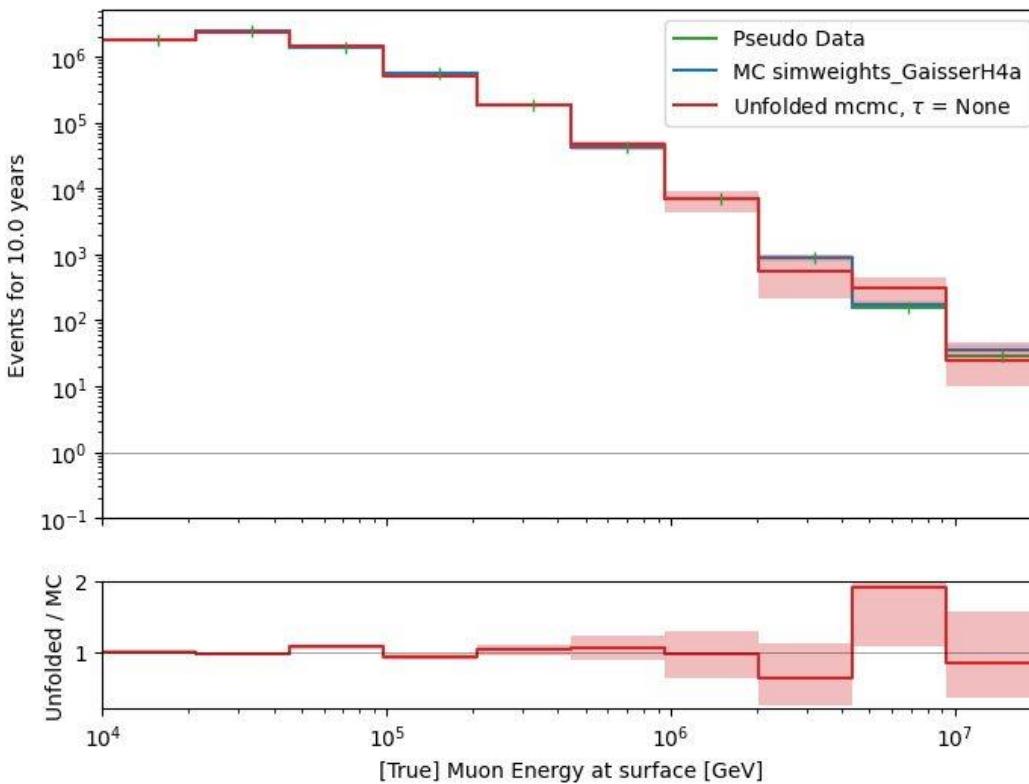
HoleIce 1

# Fit ice systematics:

- Parametrization for systematics in each proxy bin
- Proxy: leading muon energy at entry



# Unfolded event rate with systematics



➤ Use regularization to reduce unphysical oscillations

# Conclusion & outlook

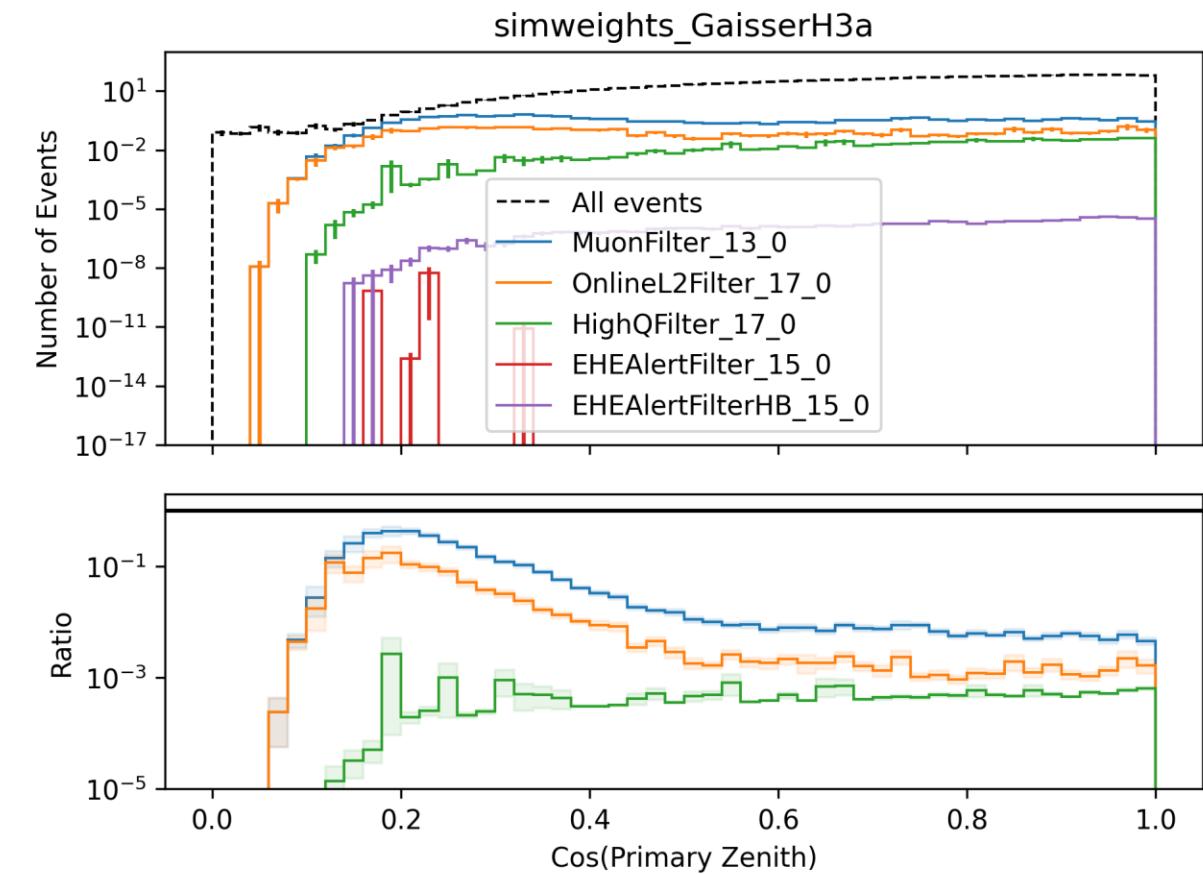
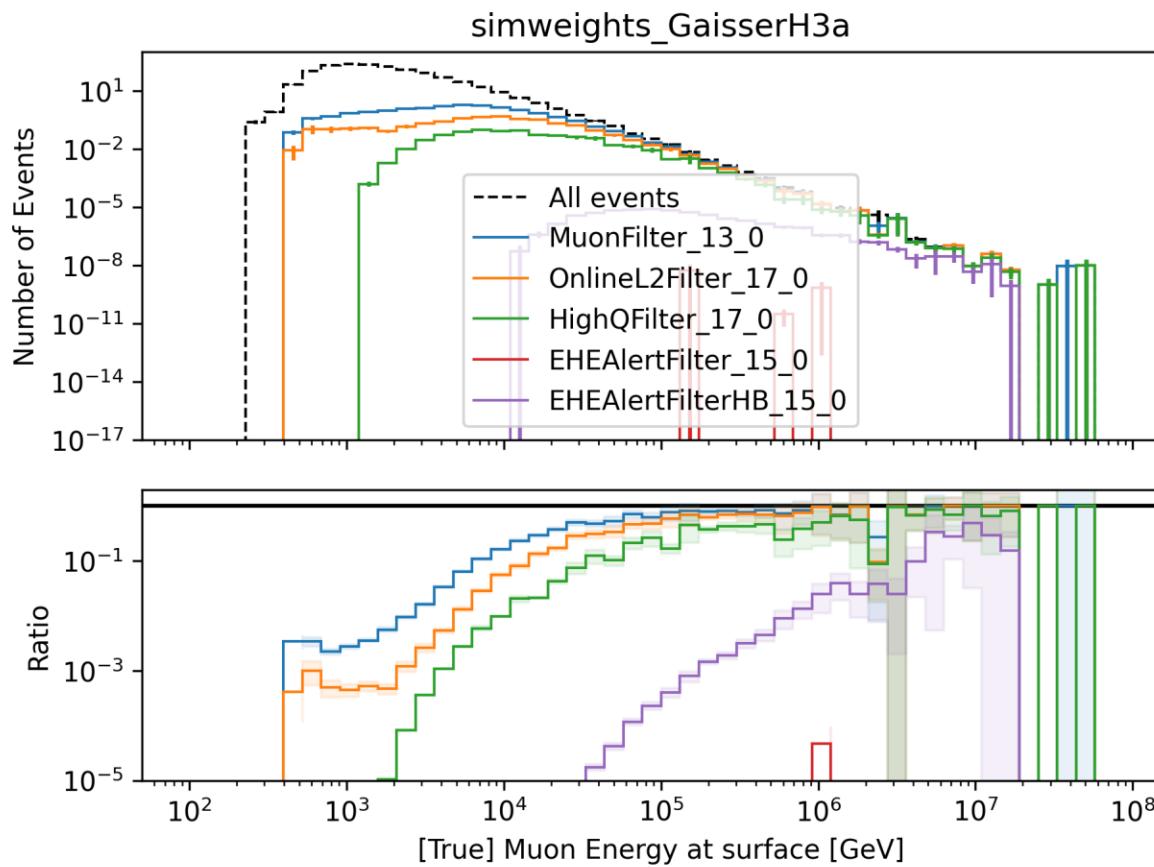
- New CORSIKA simulations with parent information
  - tag prompt and conventional muons
  - validation: agreement with MCEq
- Add DNN-based reconstructions
- Good data-MC agreement
  - quality cuts defined
  - good agreement in  $\cos(\text{zenith})$  distributions
  - mismatch in z—position → re-weight distributions
- Unfolding is set up
  - systematics can be included, minor investigations
- Fit of prompt normalization is promising
  - systematics need to be included



How ChatGPT illustrates the detection of a prompt muon with IceCube.

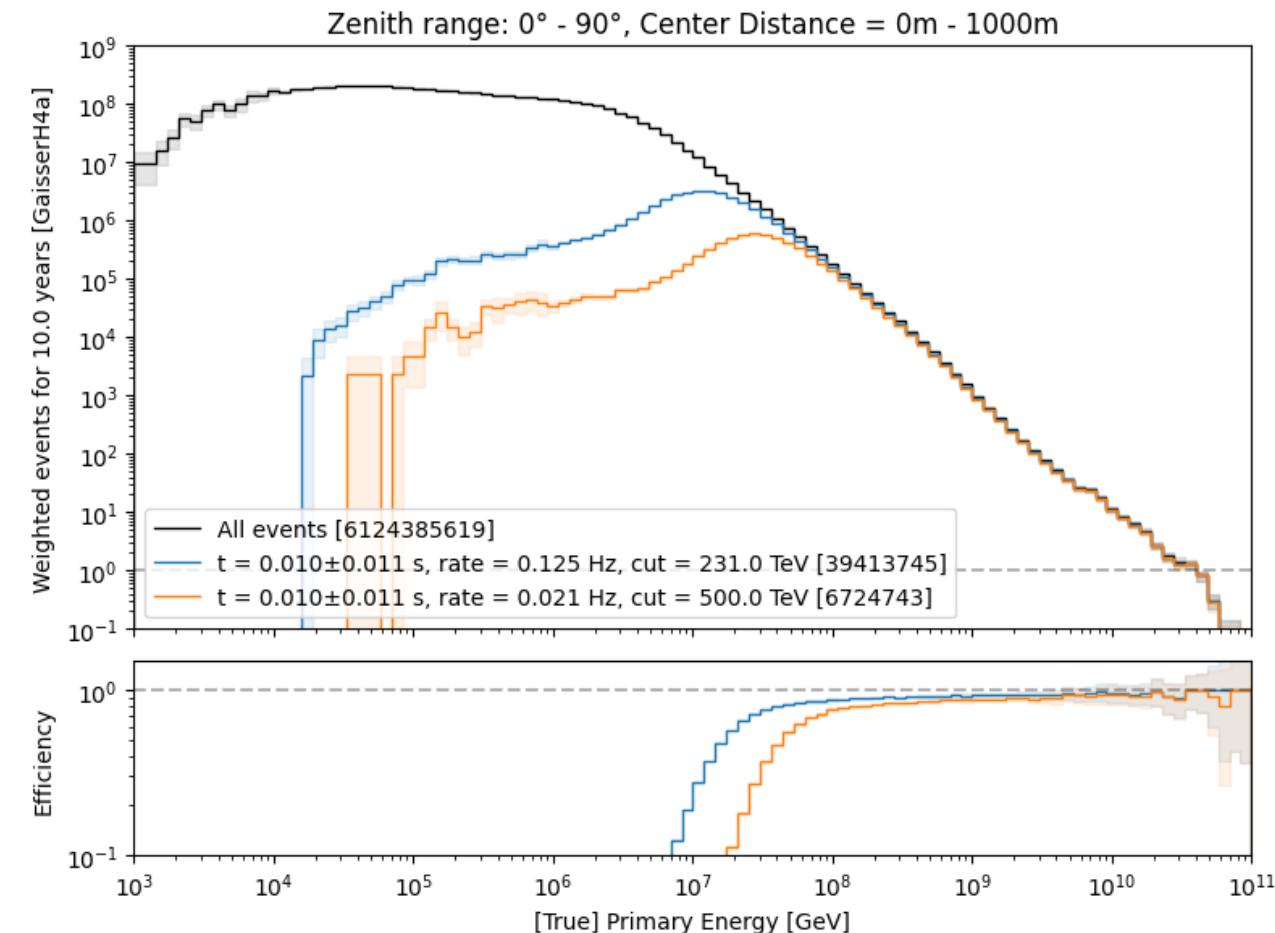
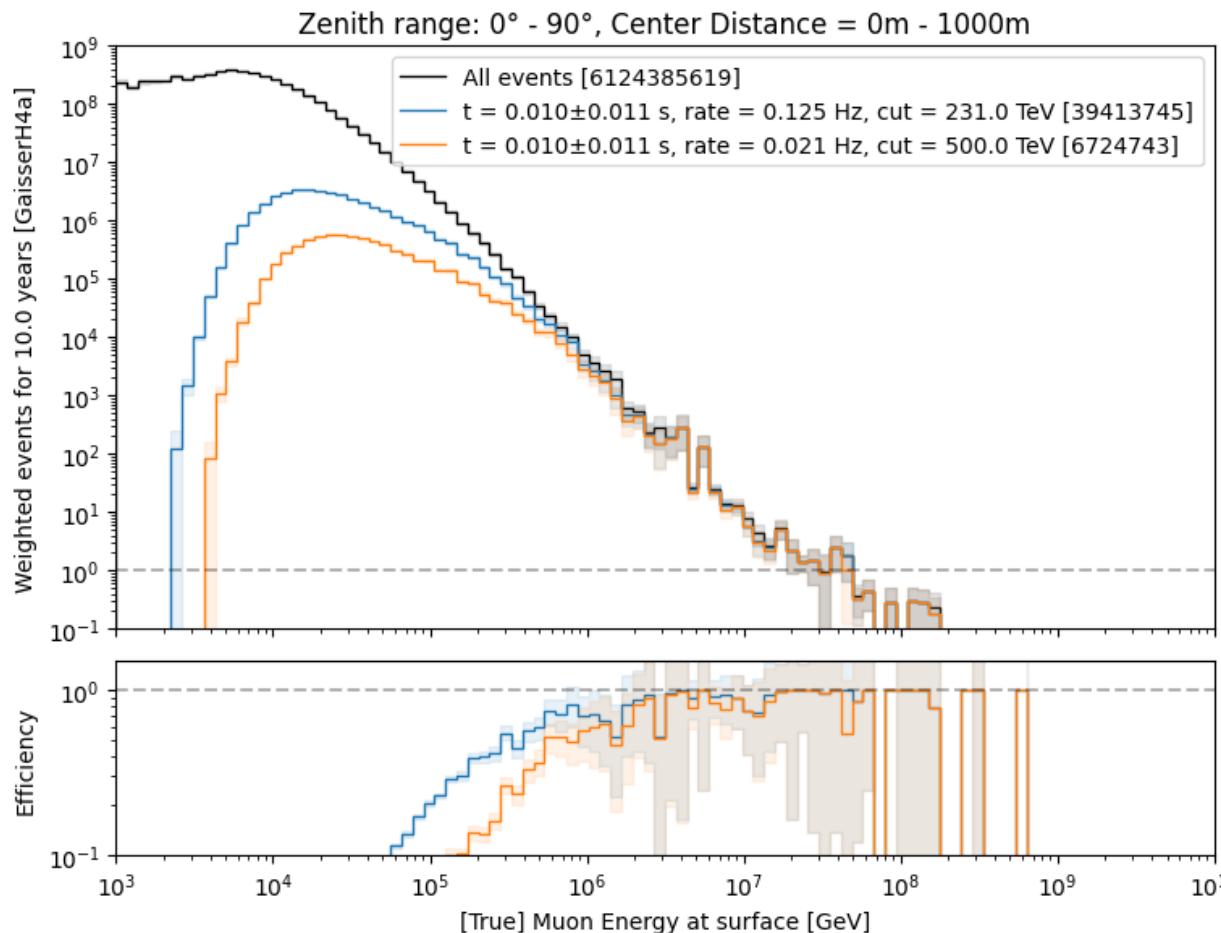
# Selection

# level3: muon filter



- HighQFilter: removes more horizontal events → muons travel large distances, not enough energy left to pass high-charge filter
- Goal: keep AMAP high-energy events at surface → muon filter

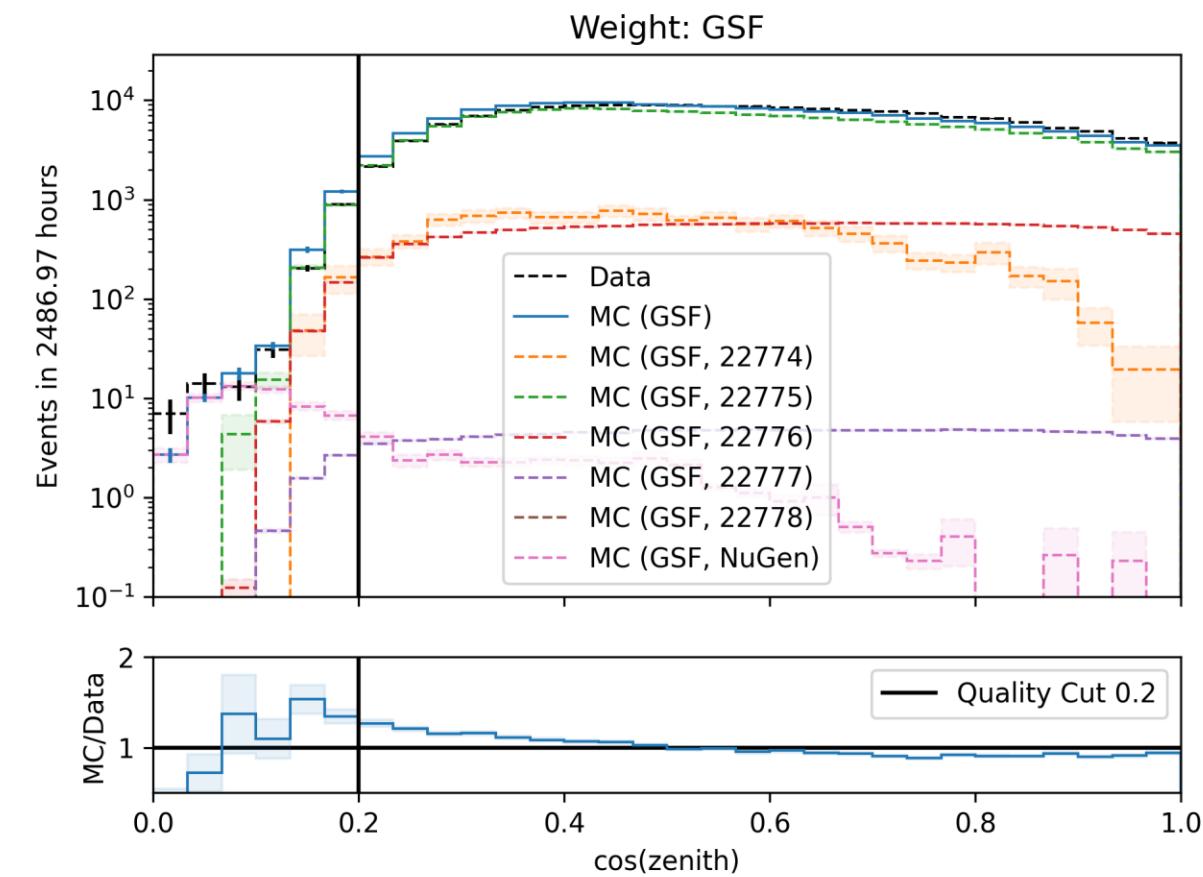
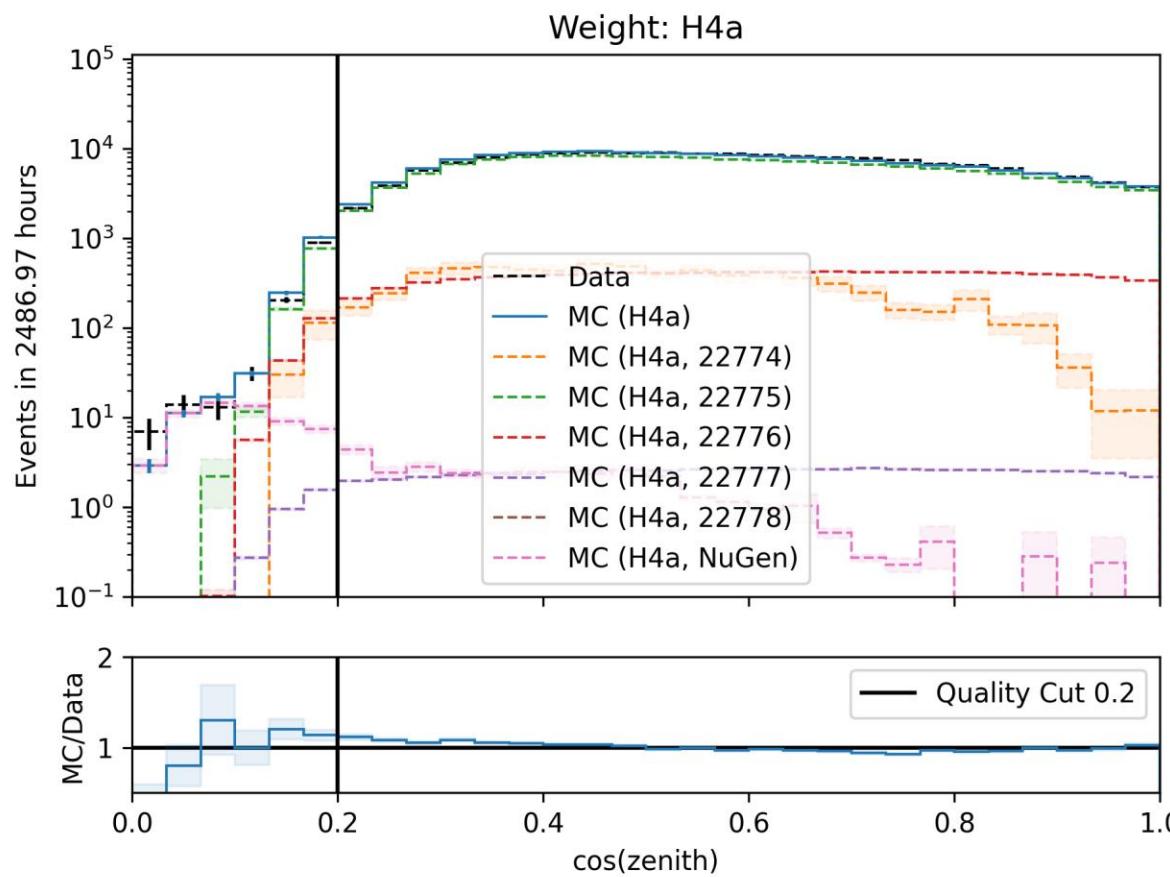
# level3: 500 TeV bundle cut at surface



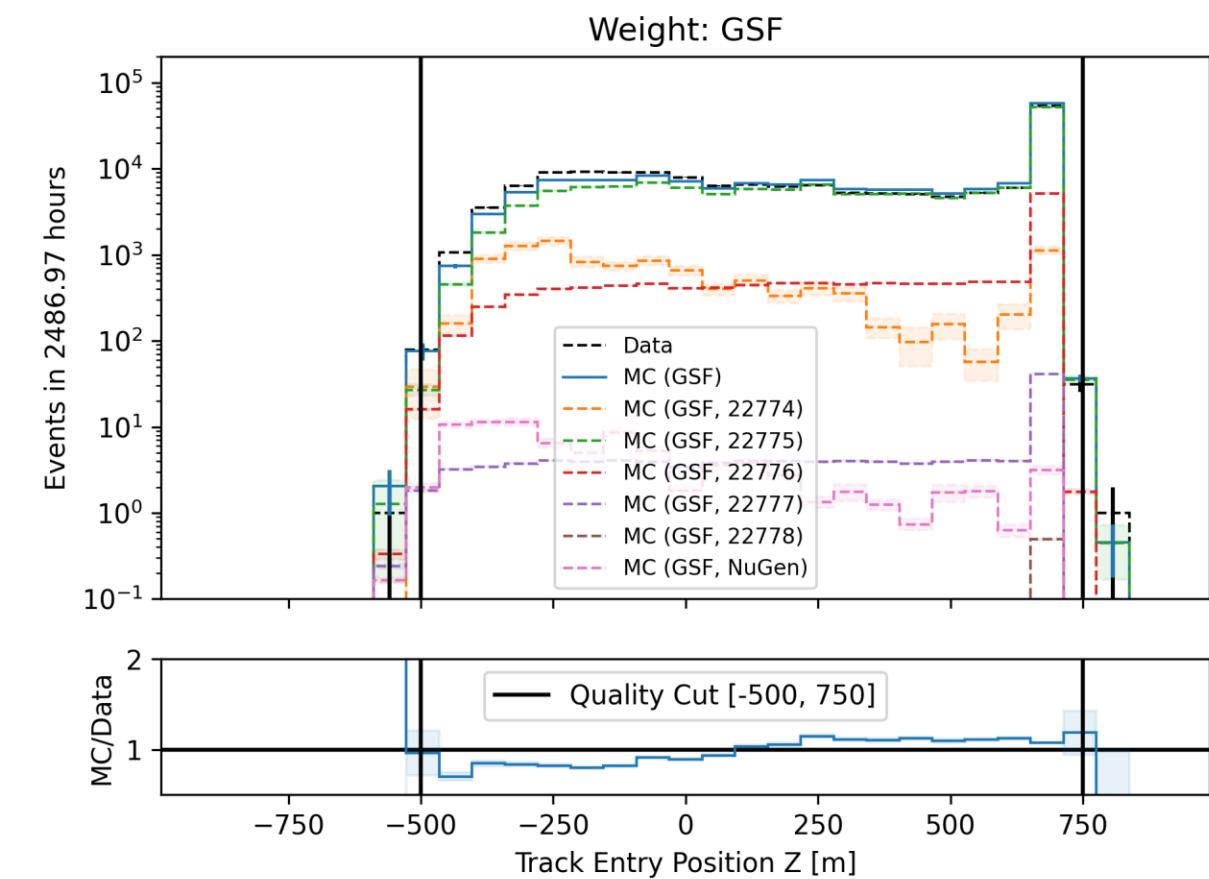
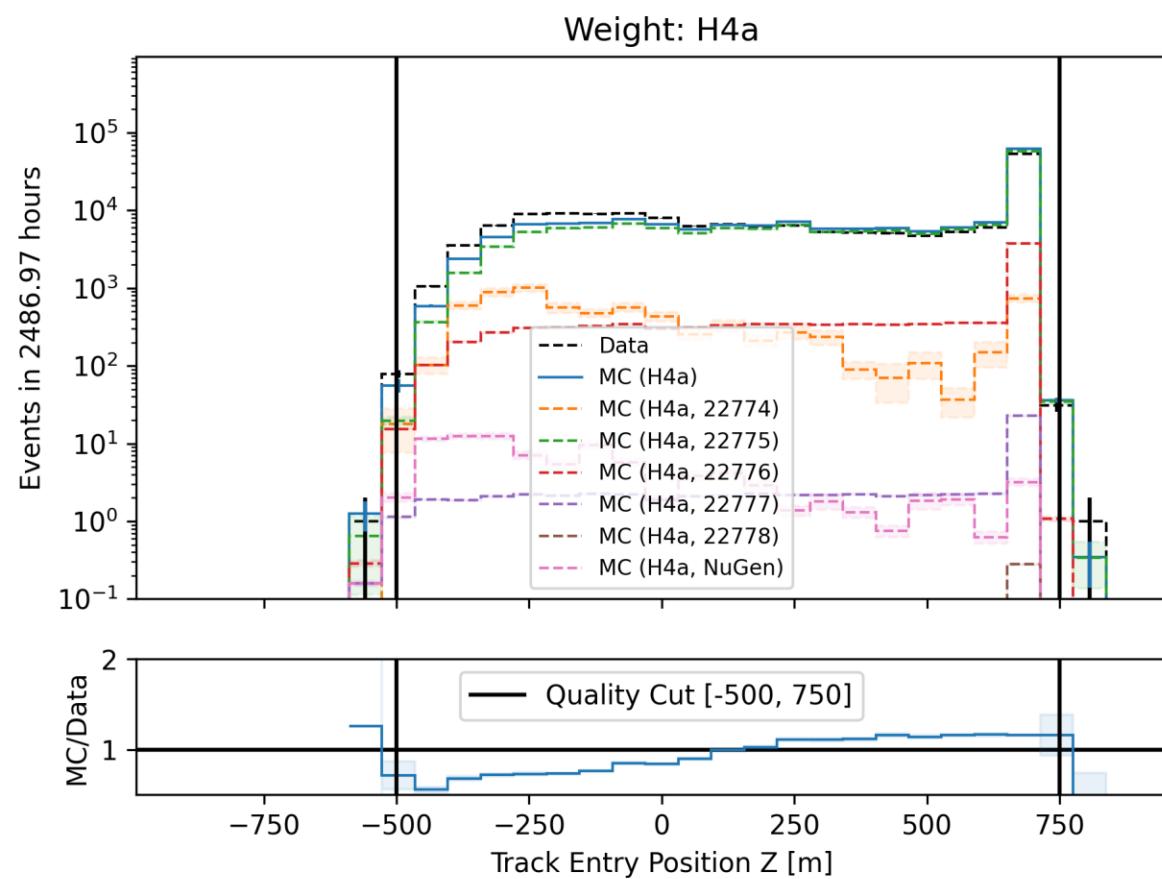
- 6 billion events expected: computationally not feasible → focus on high-energetic events
  - 500 TeV cut: 6.7 Mio events left

# Data-MC

## Cos(zenith)

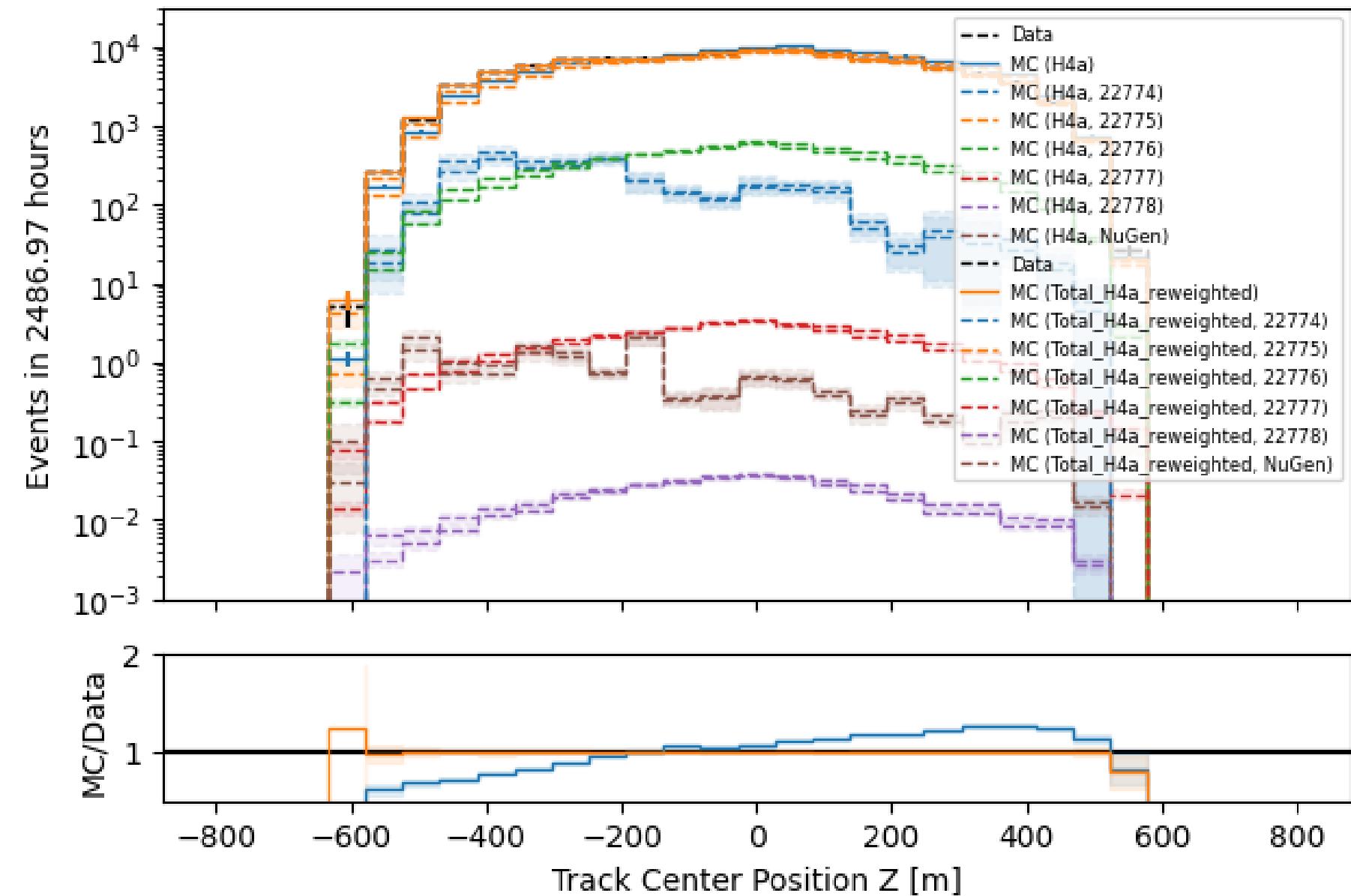


# Entry position z—vertex

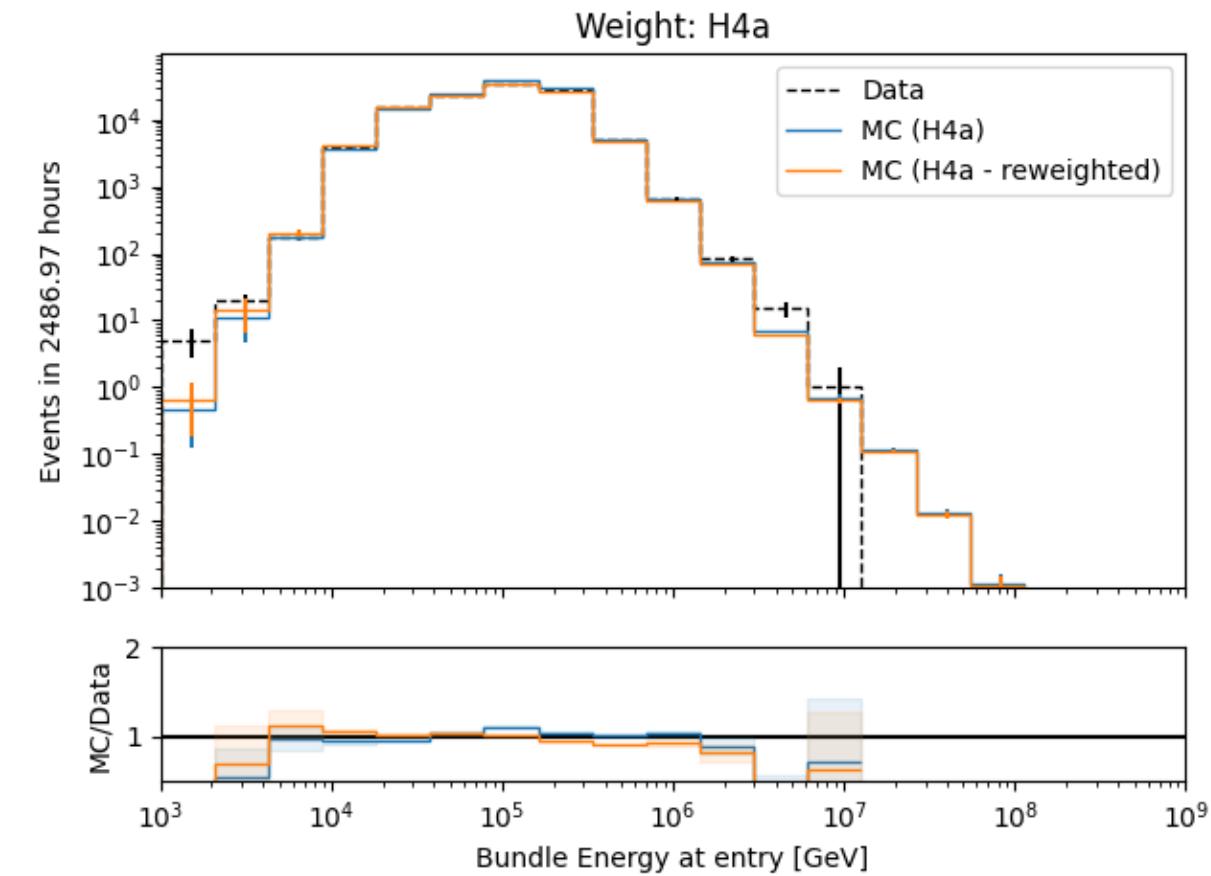
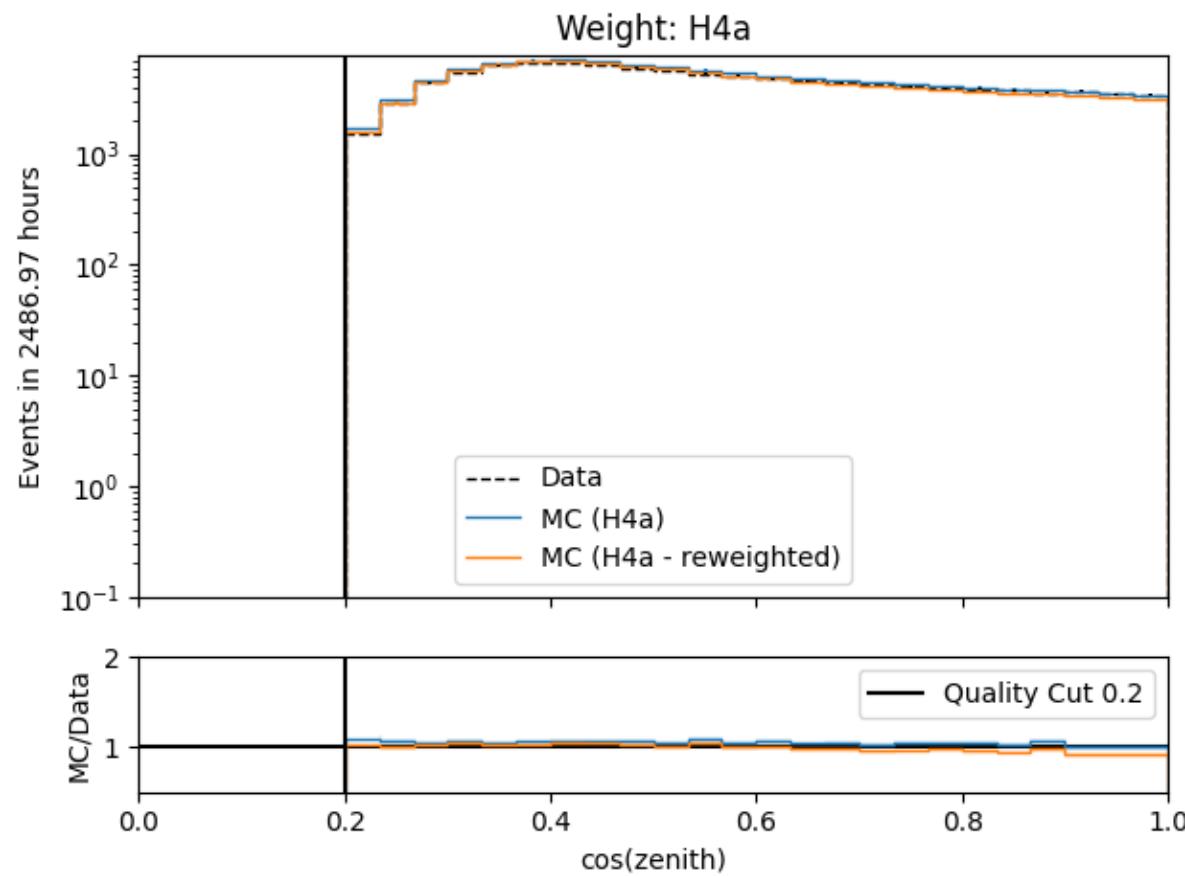


## z—reweighting

- re-weight events to enforce a perfect agreement in the center z—position



# Check z—reweighting



➤ re-weighting has no large effect on cos(zenith) and energy

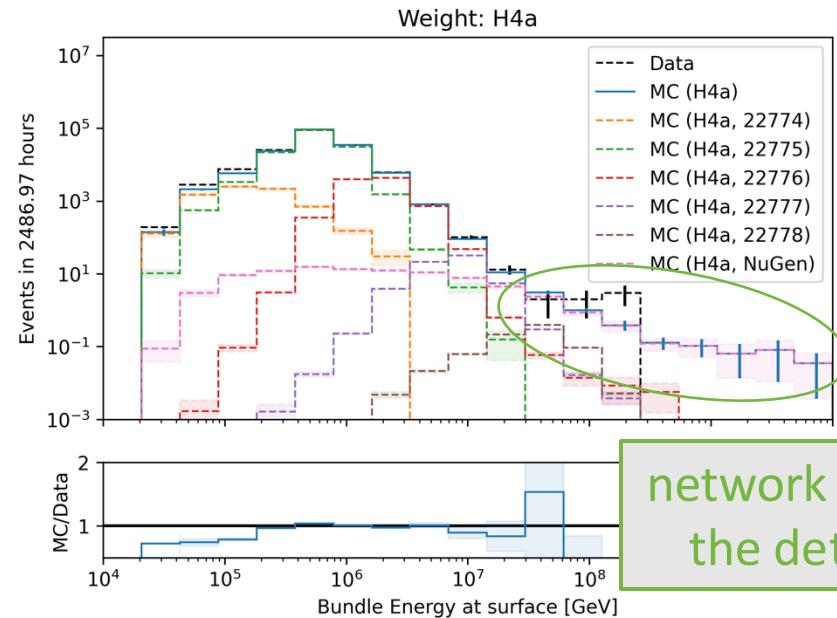
# Level5: quality cuts

containment cuts	>	<
length in detector	1000 m	2000 m
entry pos x, y	-750 m	750 m
entry pos z	-500 m	750 m
center pos x, y	-550 m	550 m
center pos z	-650 m	650 m

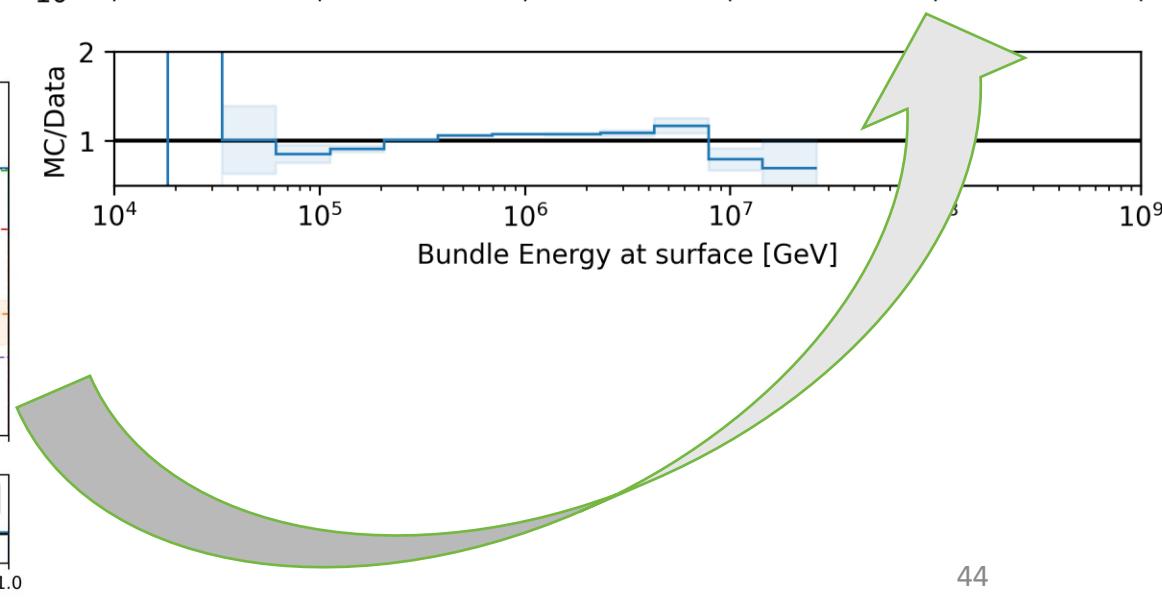
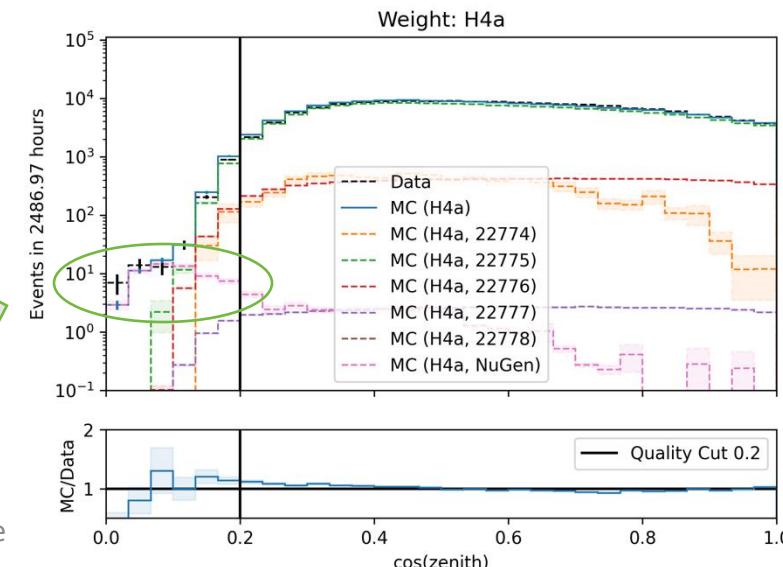
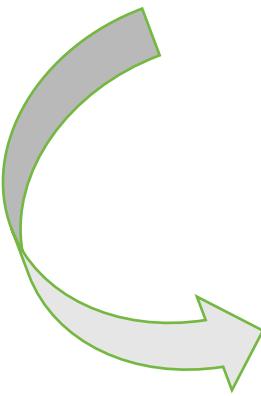
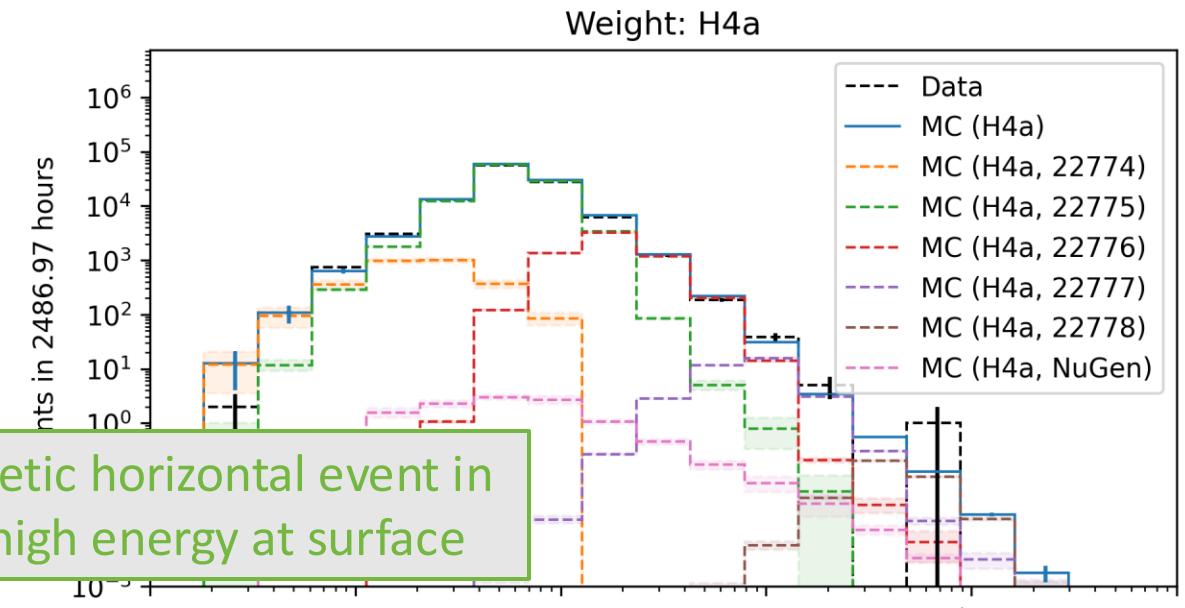
neutrino cuts	>	<
$\cos(\text{zenith})$	0.2	
length	5000 m	15000 m

uncertainty cuts	<
bundle energy at entry	$0.9 \log_{10}(\text{GeV})$
bundle energy at surface	$2.0 \log_{10}(\text{GeV})$
zenith	0.1 rad
azimuth	0.2 rad
entry pos x, y, z	42 m
center pos x, y, z	50 m
entry pos time	200 ns
center pos time	600 ns
length in detector	160 m
length	2000 m

# Level5: quality cuts → removes neutrinos



network learns: a high-energetic horizontal event in  
the detector needs a very high energy at surface

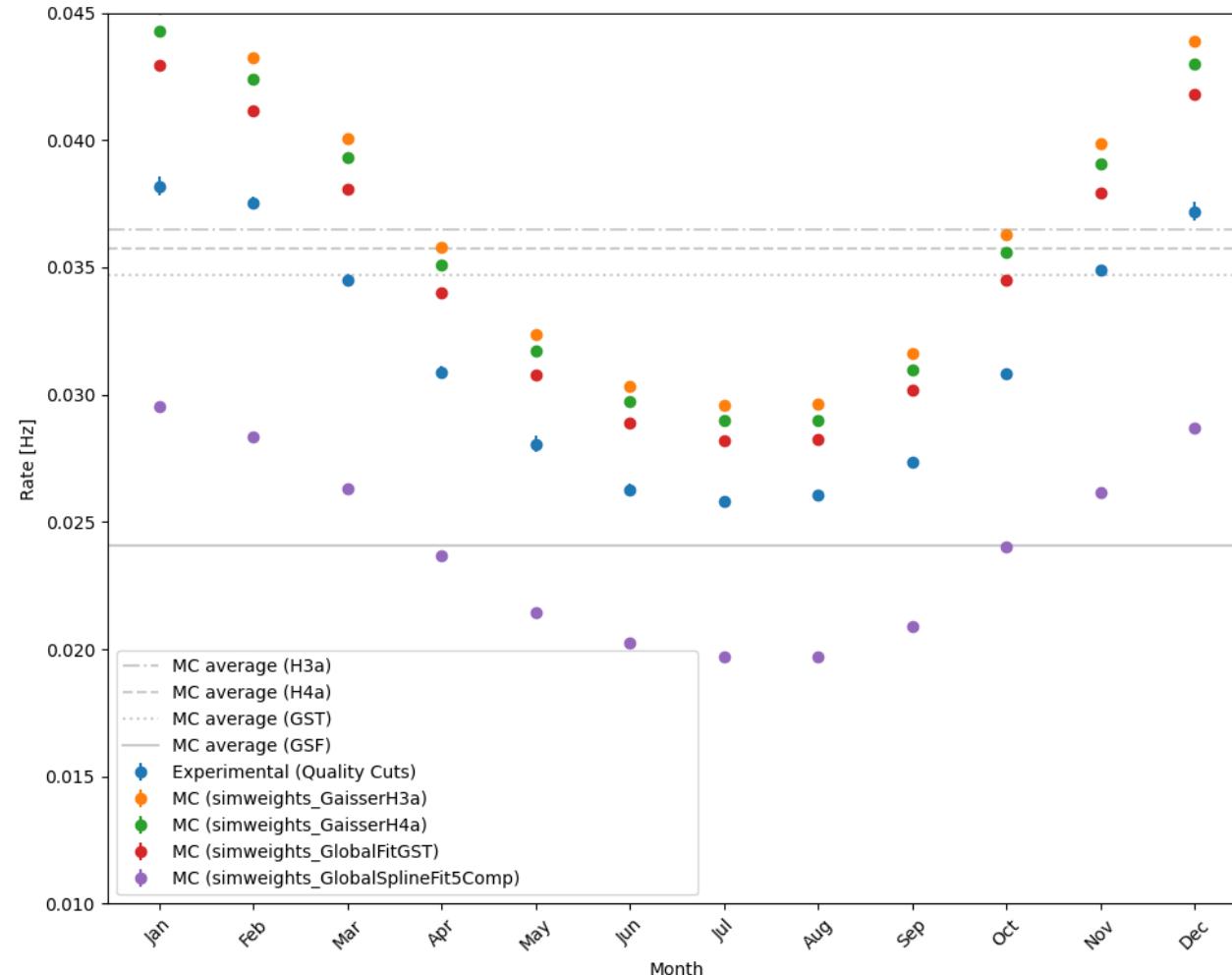


# Rates per month

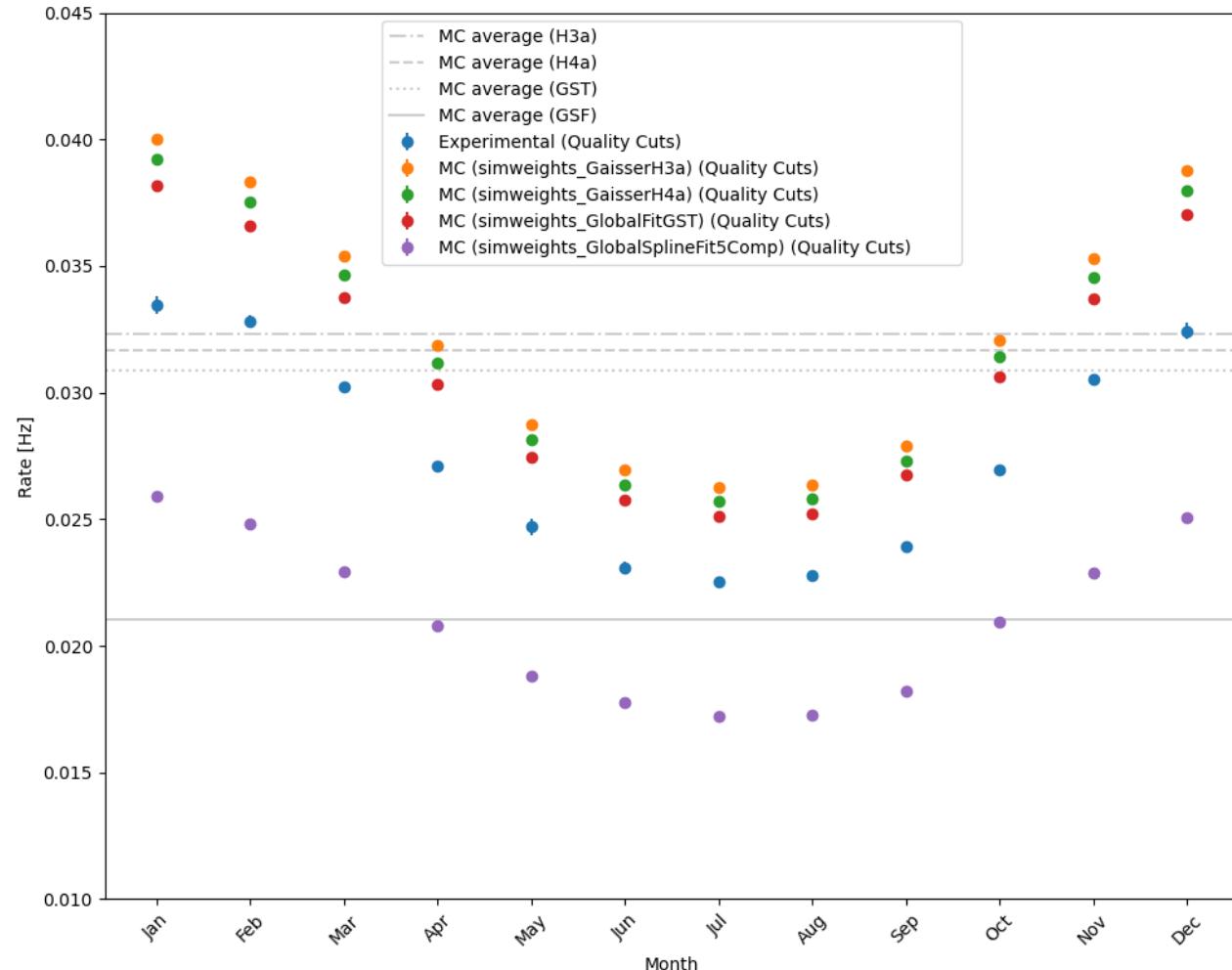
Data-MC

# Rates per month

before quality cuts

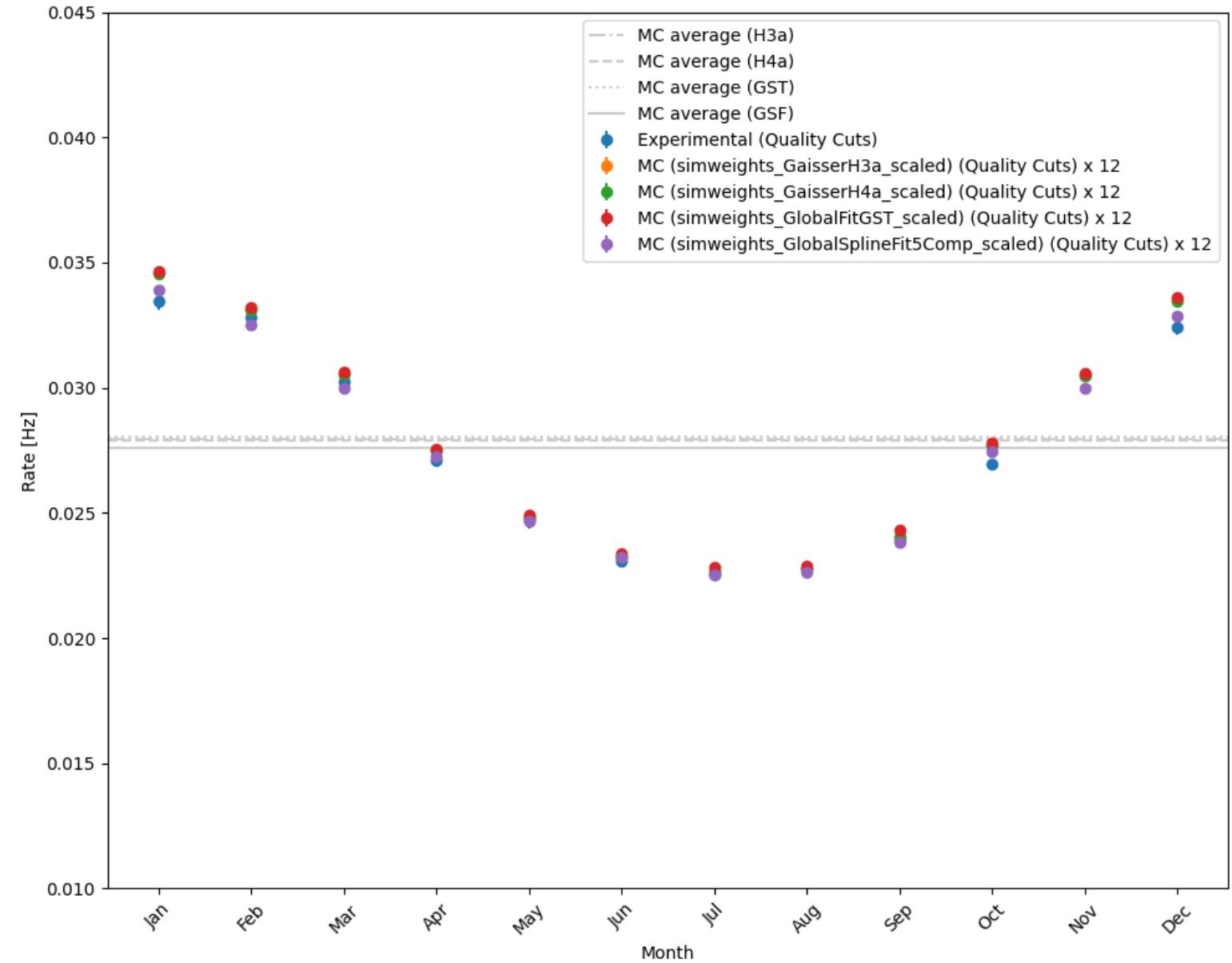


after quality cuts



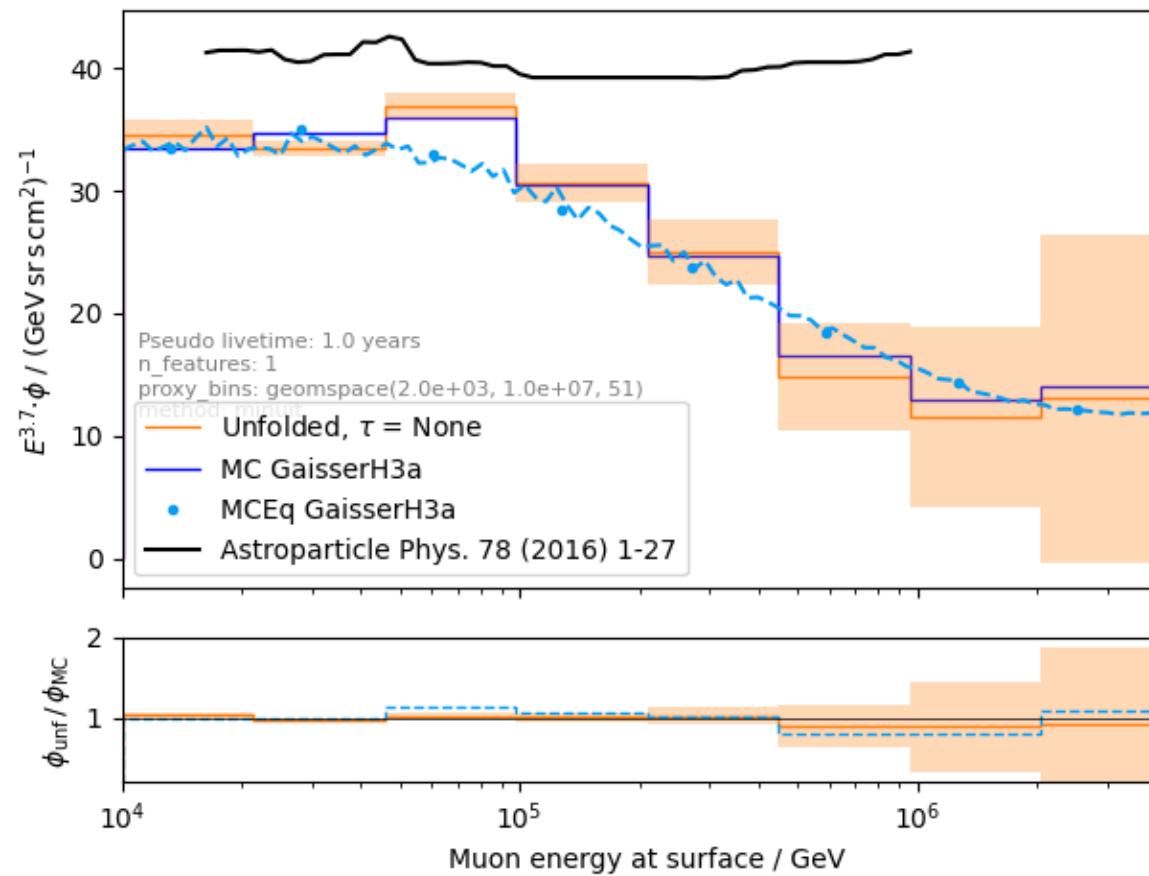
# Rates per month: scaled

- Rates are scaled to the experimental rate to compare the shapes
- Good agreement per month

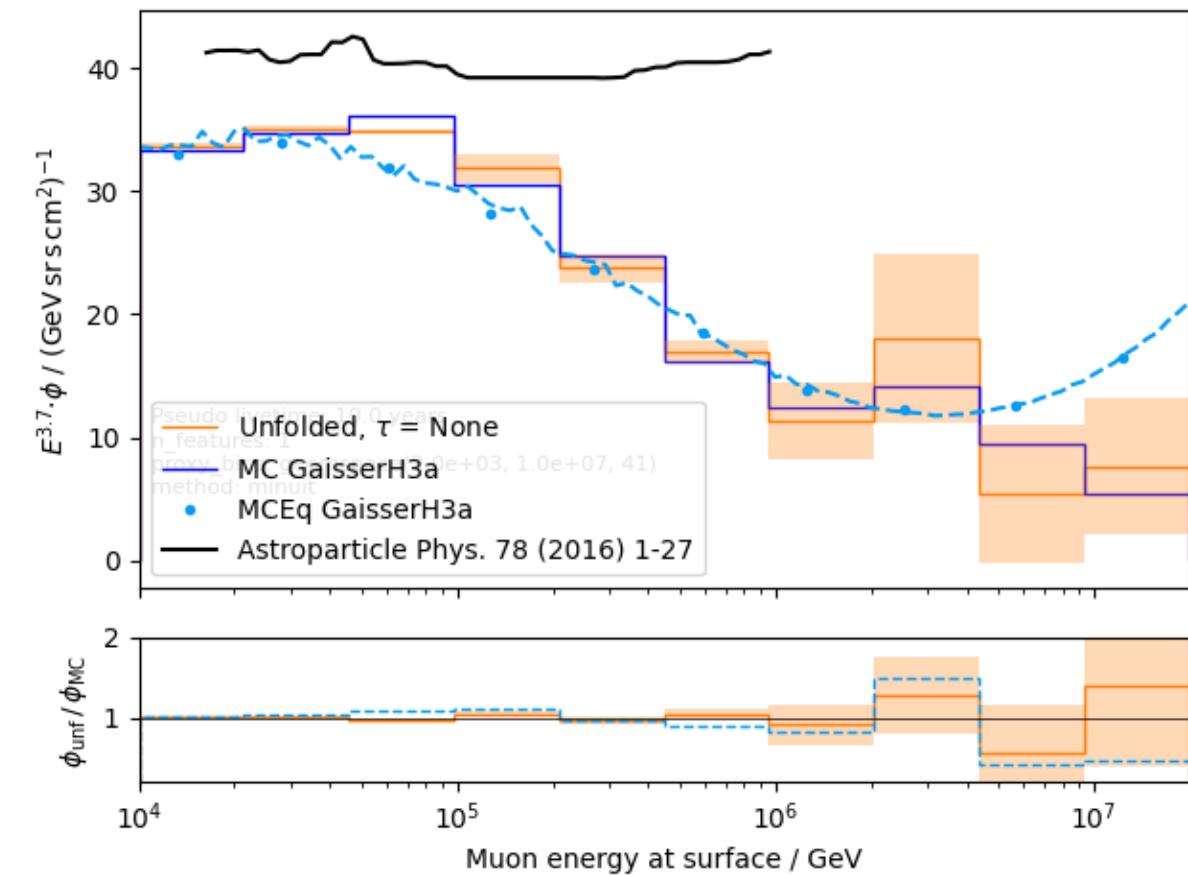


unfolding of muon flux at surface – weighted to  $E^{3.7}$ 

1 year



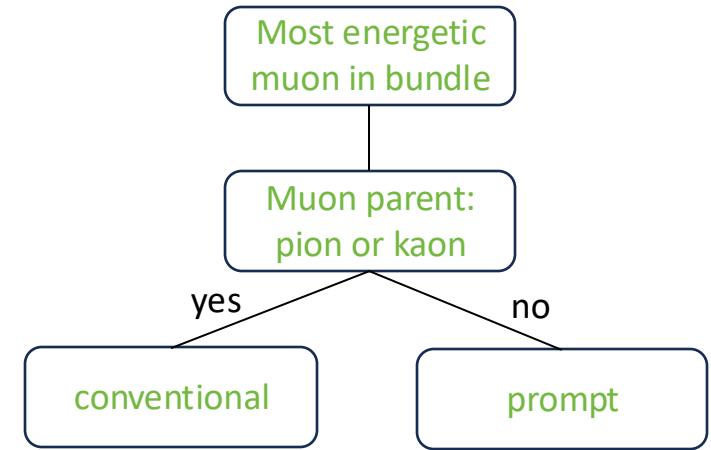
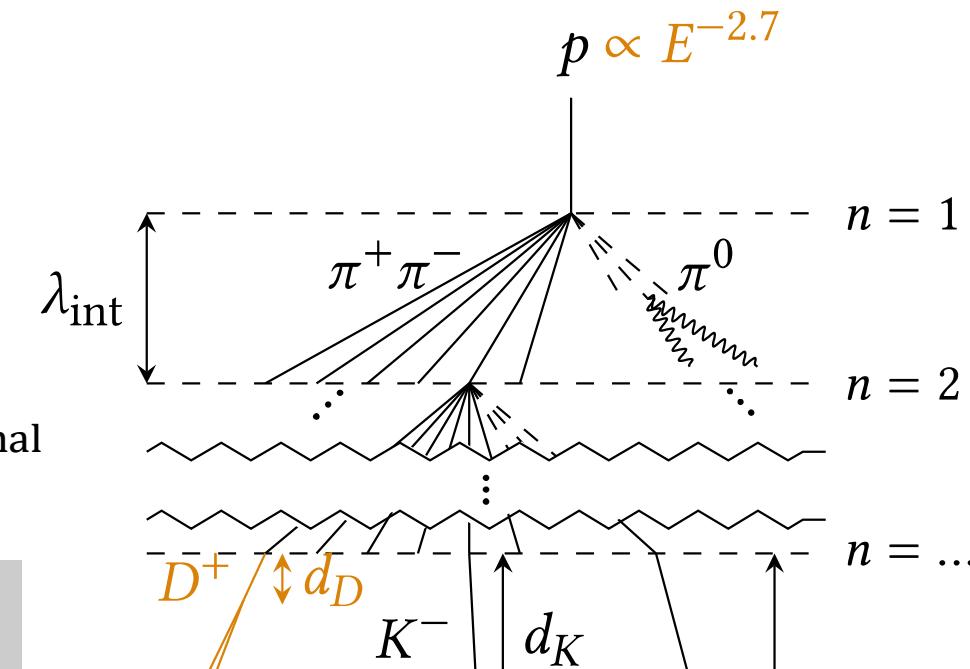
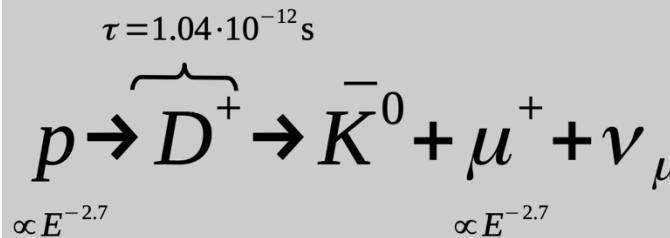
10 years



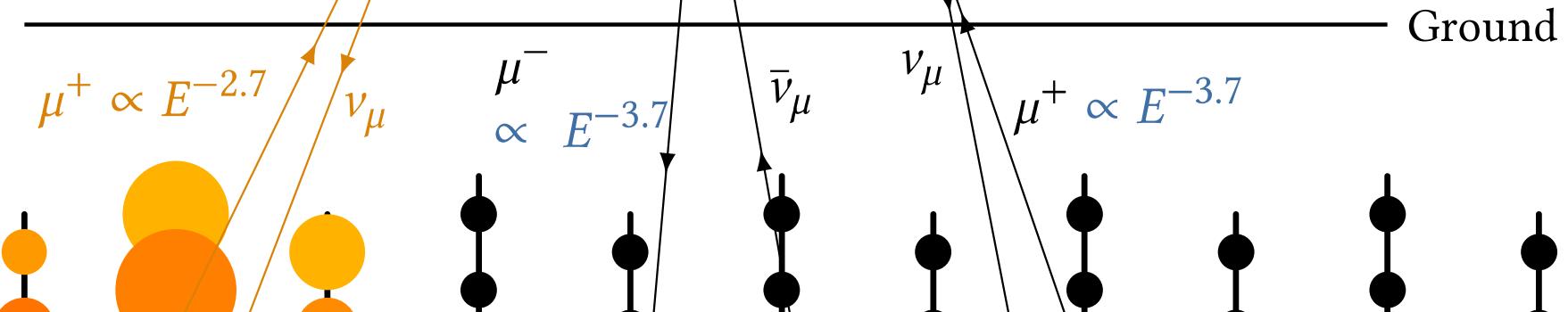
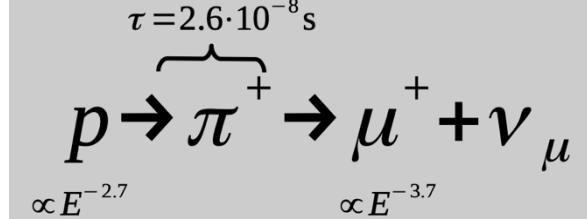
# Muon flux

$$\Phi_{\text{tot}} = \Phi_{\text{prompt}} + \Phi_{\text{conventional}}$$

**prompt component:**

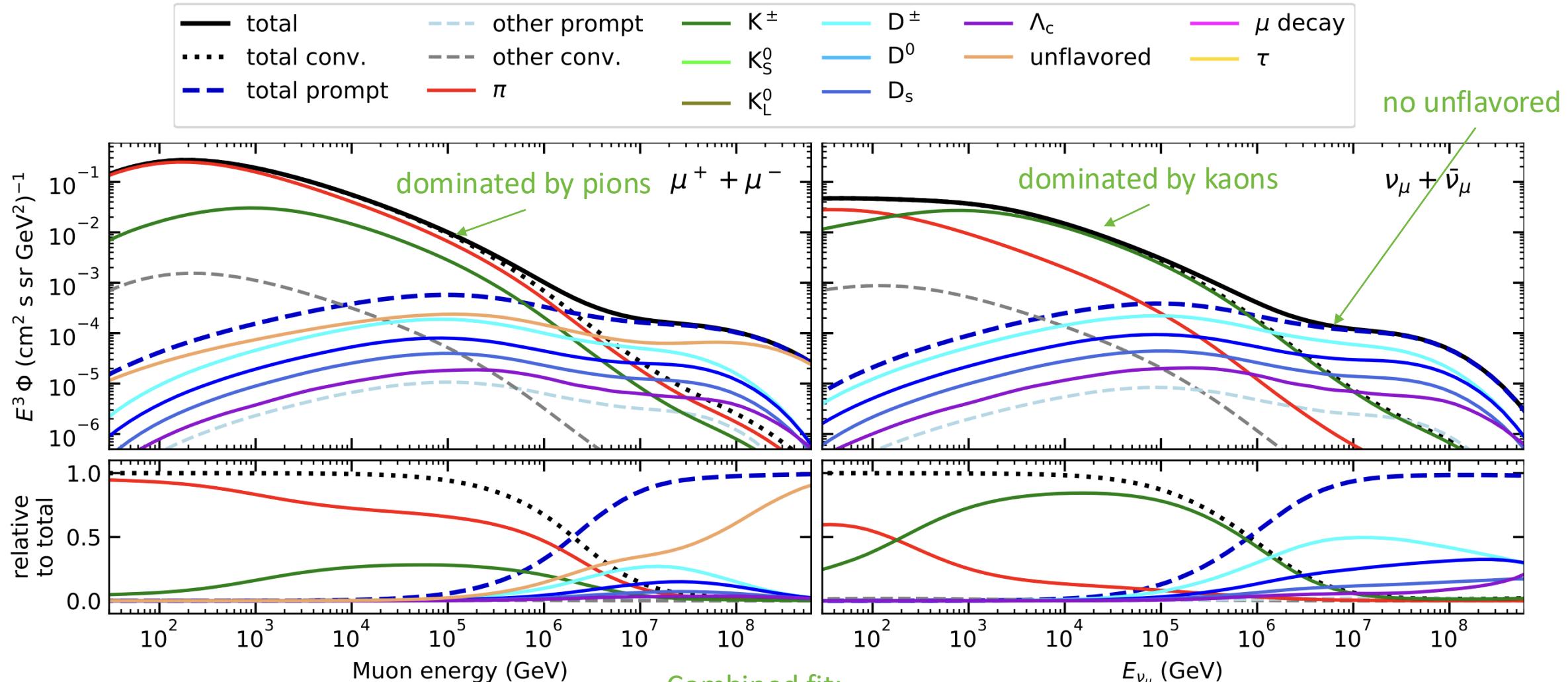


**Conventional component:**



# Prompt atmospheric muons and neutrinos

10.1103/PhysRevD.100.103018

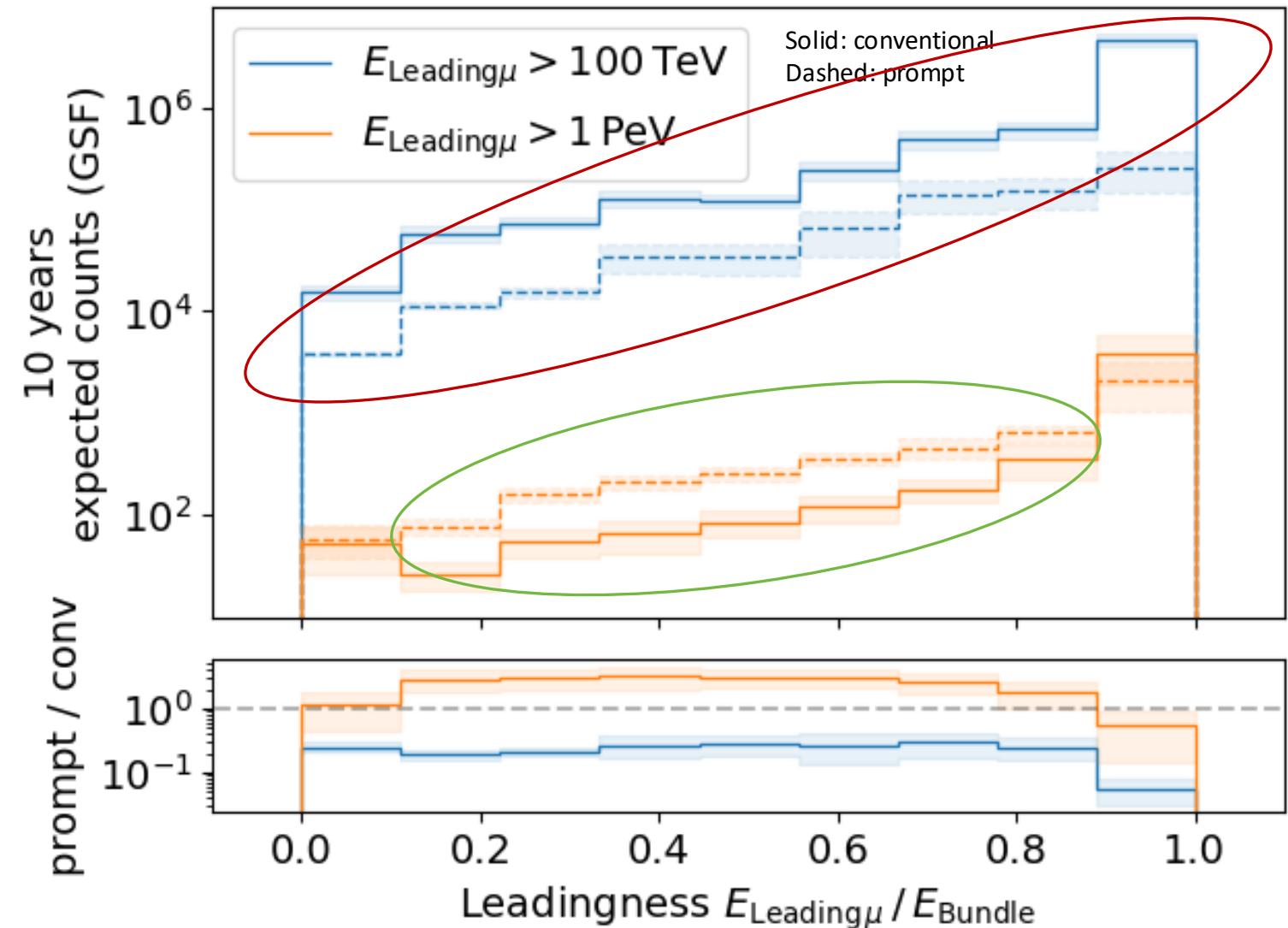


Combined fit:

- handle on pion/kaon ratio
- handle on charmed mesons

# Leading muon energy fraction

- Prompt dominates for energies  $> 1 \text{ PeV}$
- Leading energy sweet spot:  $0.1 - 0.9$



# Leading muon contribution

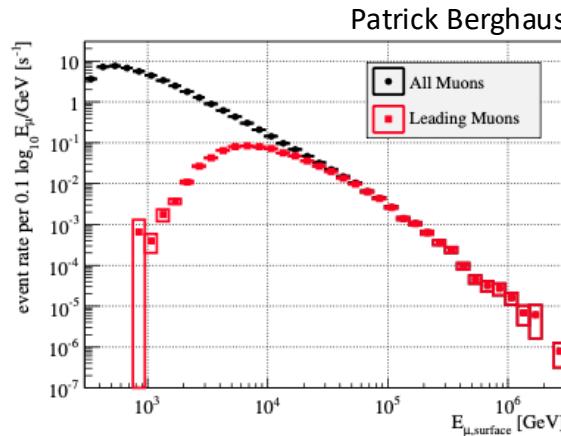
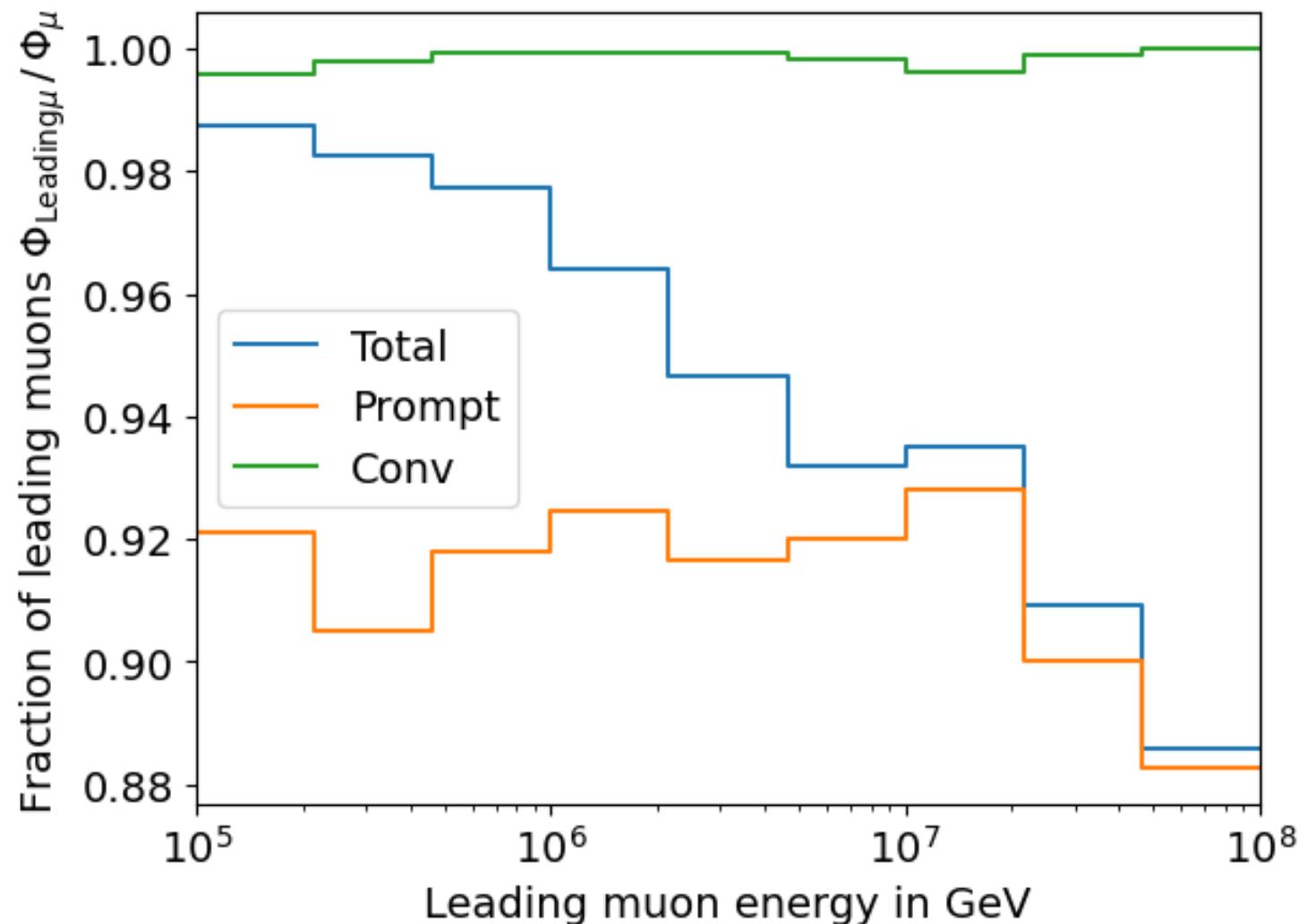


Figure 10: Surface energy distribution for all and most energetic (“leading”) muons in simulated events with a total of more than 1,000 registered photo-electrons in IceCube.

- Muons with energies between 100 TeV and 50 PeV dominate the bundle by more than 90%
  - In average conventional muons are more dominant than prompt
  - But: at high energies, there are more prompt than conventional events
- High leading energy fraction does not lead to more sensitivity to detect prompt

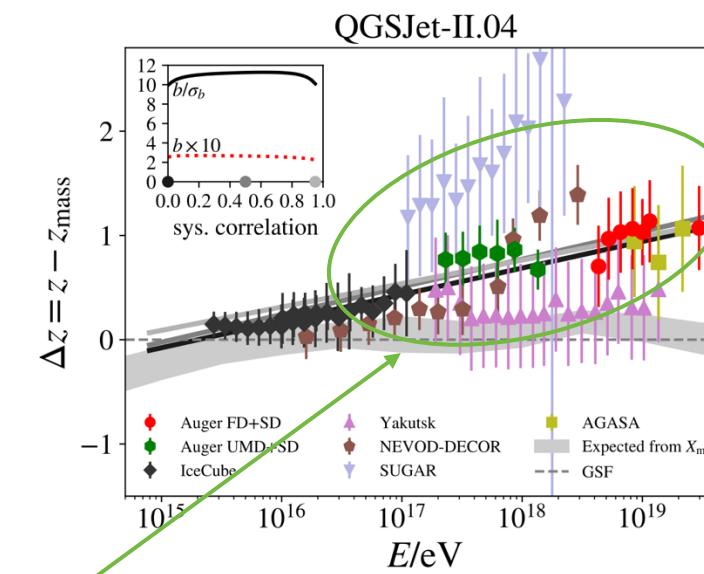
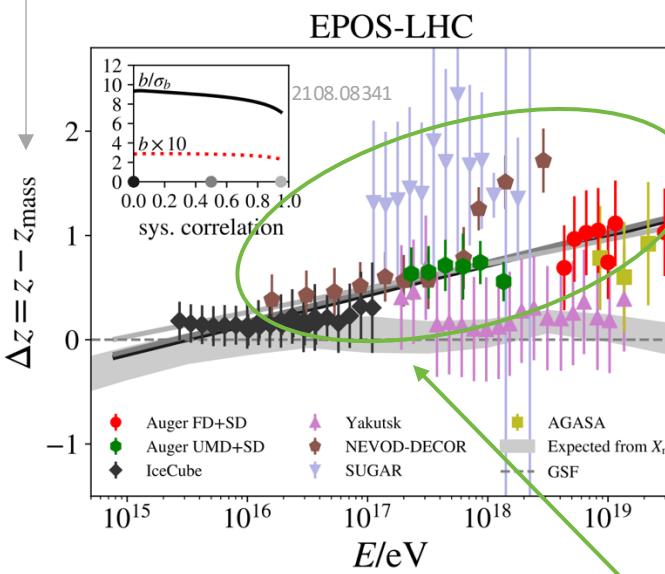


# Muon puzzle and hadronic uncertainties

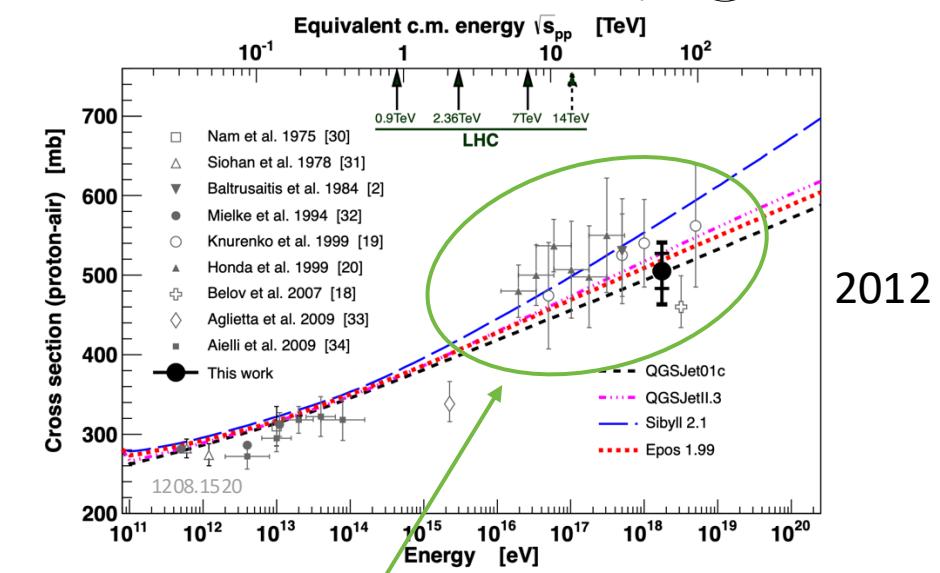
"muon number"

$$z = \frac{\ln\langle N_\mu \rangle - \ln\langle N_\mu \rangle_p}{\ln\langle N_\mu \rangle_{\text{Fe}} - \ln\langle N_\mu \rangle_p}$$

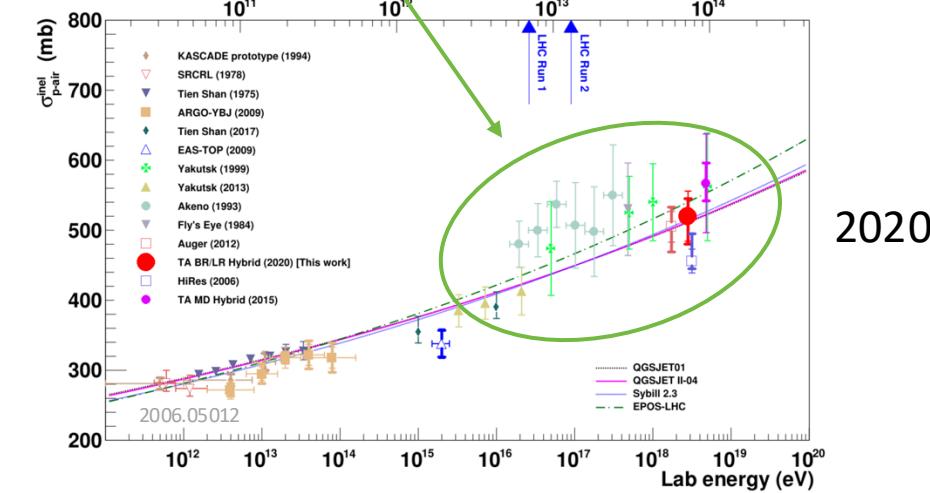
Expected  $z$   
("muon number")



➤ More muons measured than simulated for  $E > 40 \text{ PeV} \sim \text{ cms } 8 \text{ TeV}$



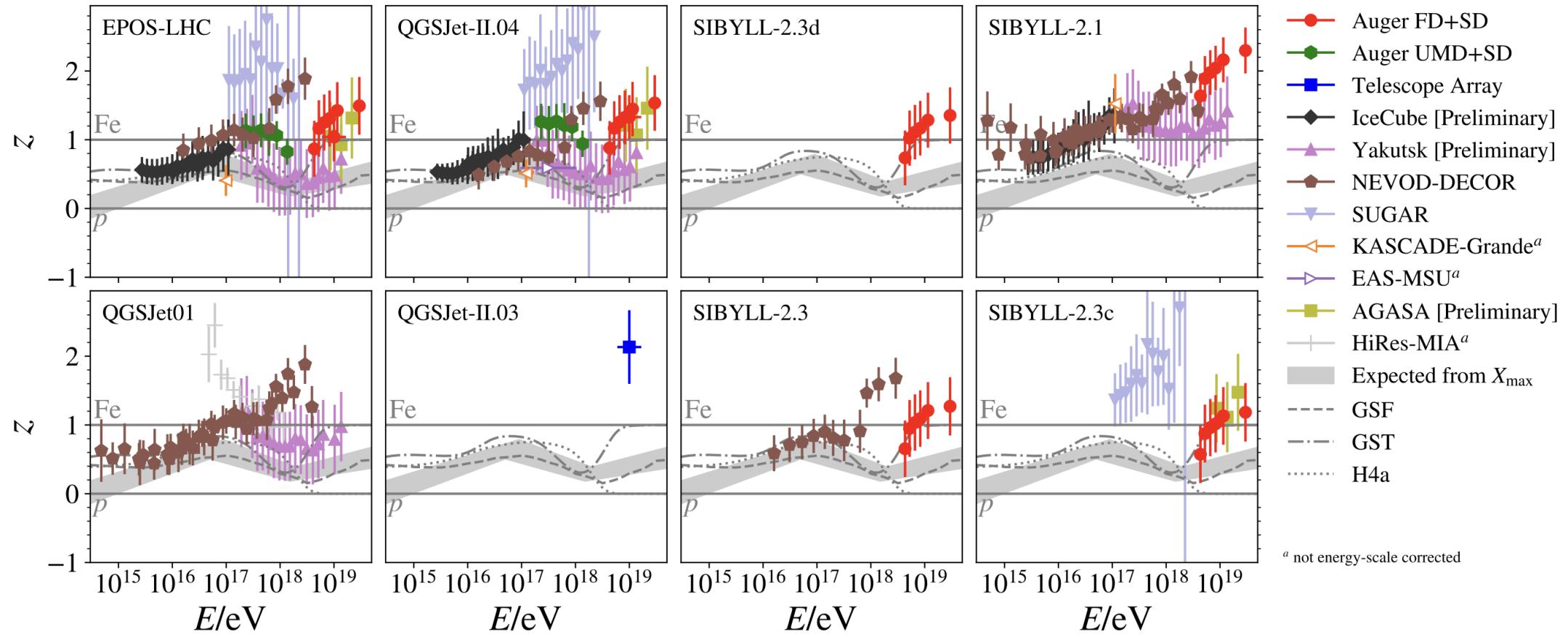
➤ Uncertainties at  $E > 10 \text{ PeV}$



# The Muon Puzzle

"muon number"

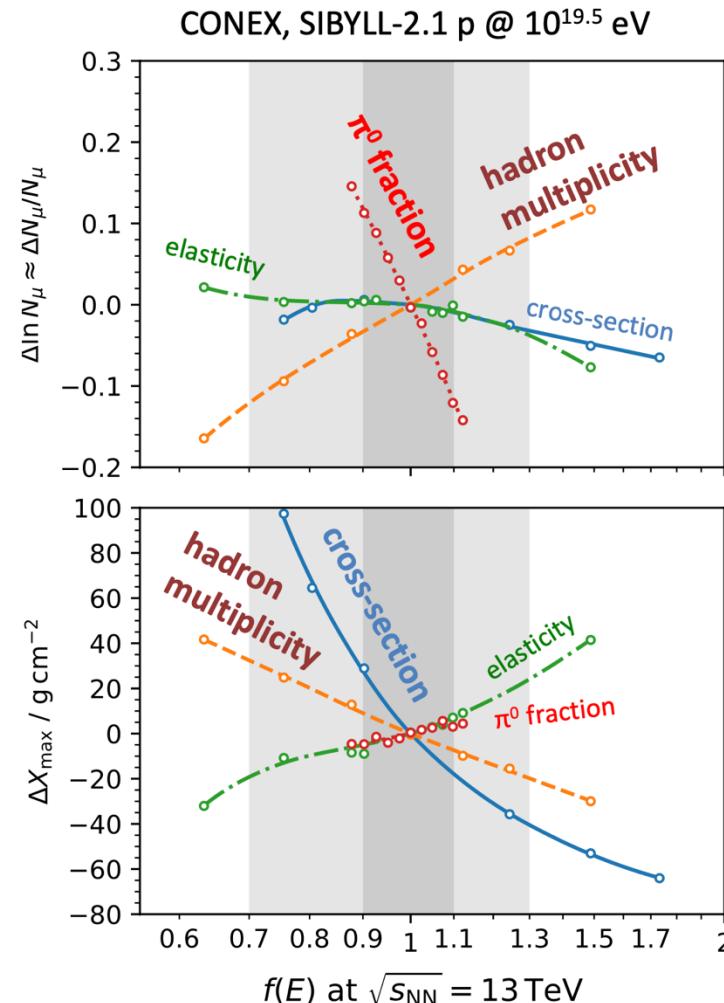
$$z = \frac{\ln\langle N_\mu \rangle - \ln\langle N_\mu \rangle_p}{\ln\langle N_\mu \rangle_{\text{Fe}} - \ln\langle N_\mu \rangle_p}$$



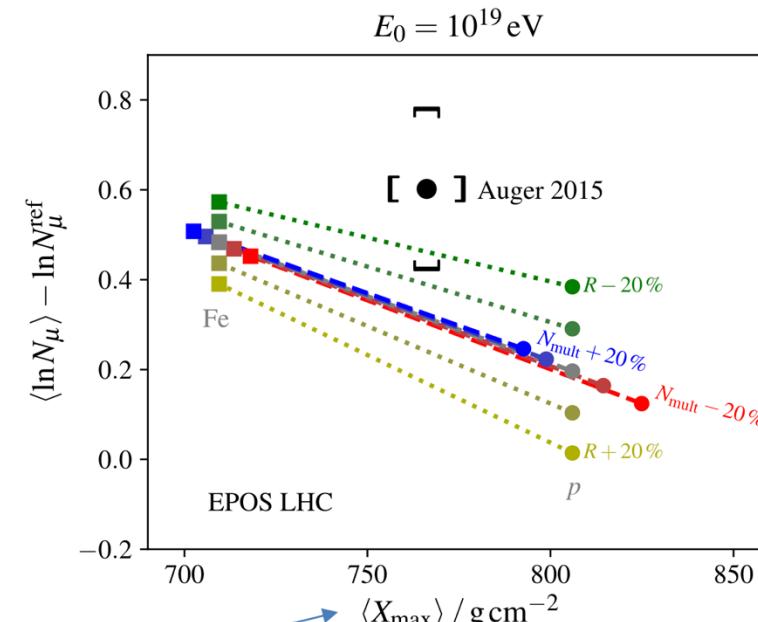
2108.08341

# Possible Solutions

R. Ulrich, R. Engel, M. Unger, PRD 83 (2011) 054026



S. Baur, HD, M. Perlin, T. Pierog, R. Ulrich, K. Werner,  
arXiv:1902.09265

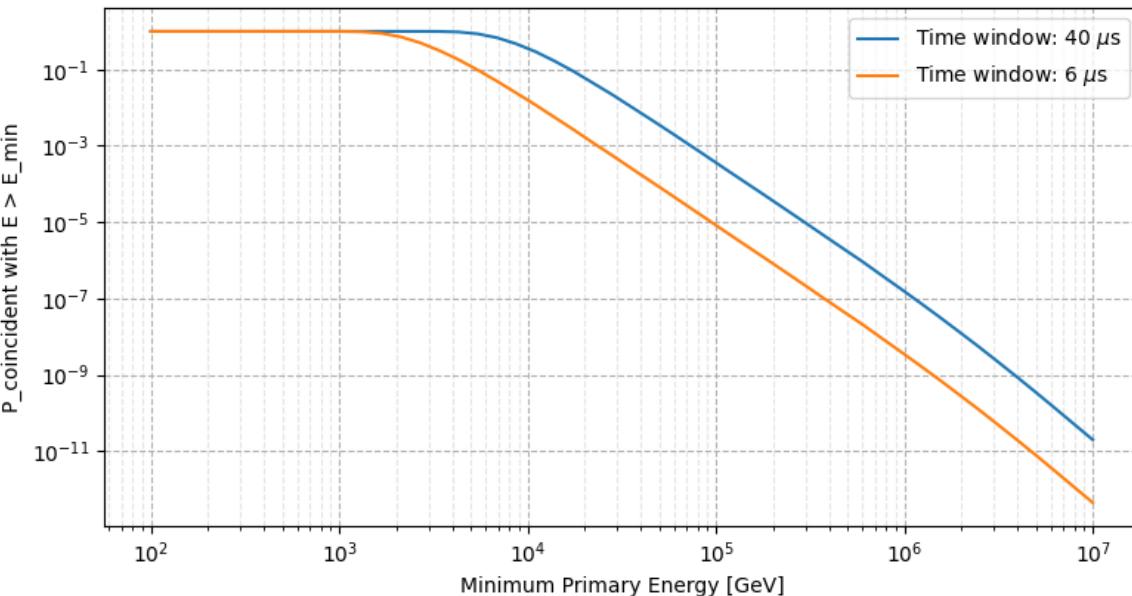
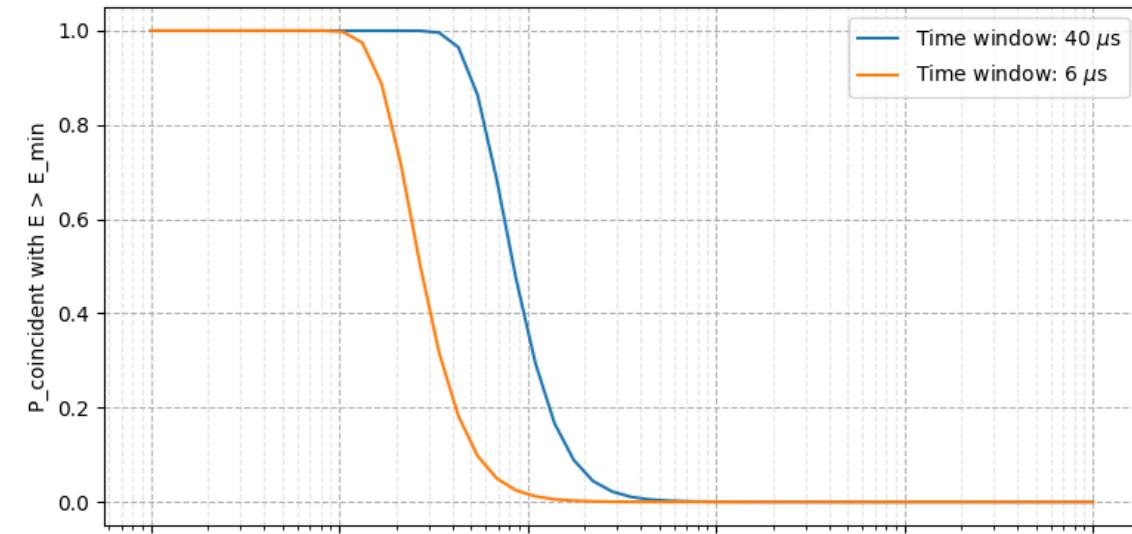


$$R = \frac{E_{\pi^0}}{E_{\text{other hadrons}}}$$

- Only changes to  $R$  can solve muon puzzle
- Small changes have large effect,  
 $R$  needs to be known to about 5 %

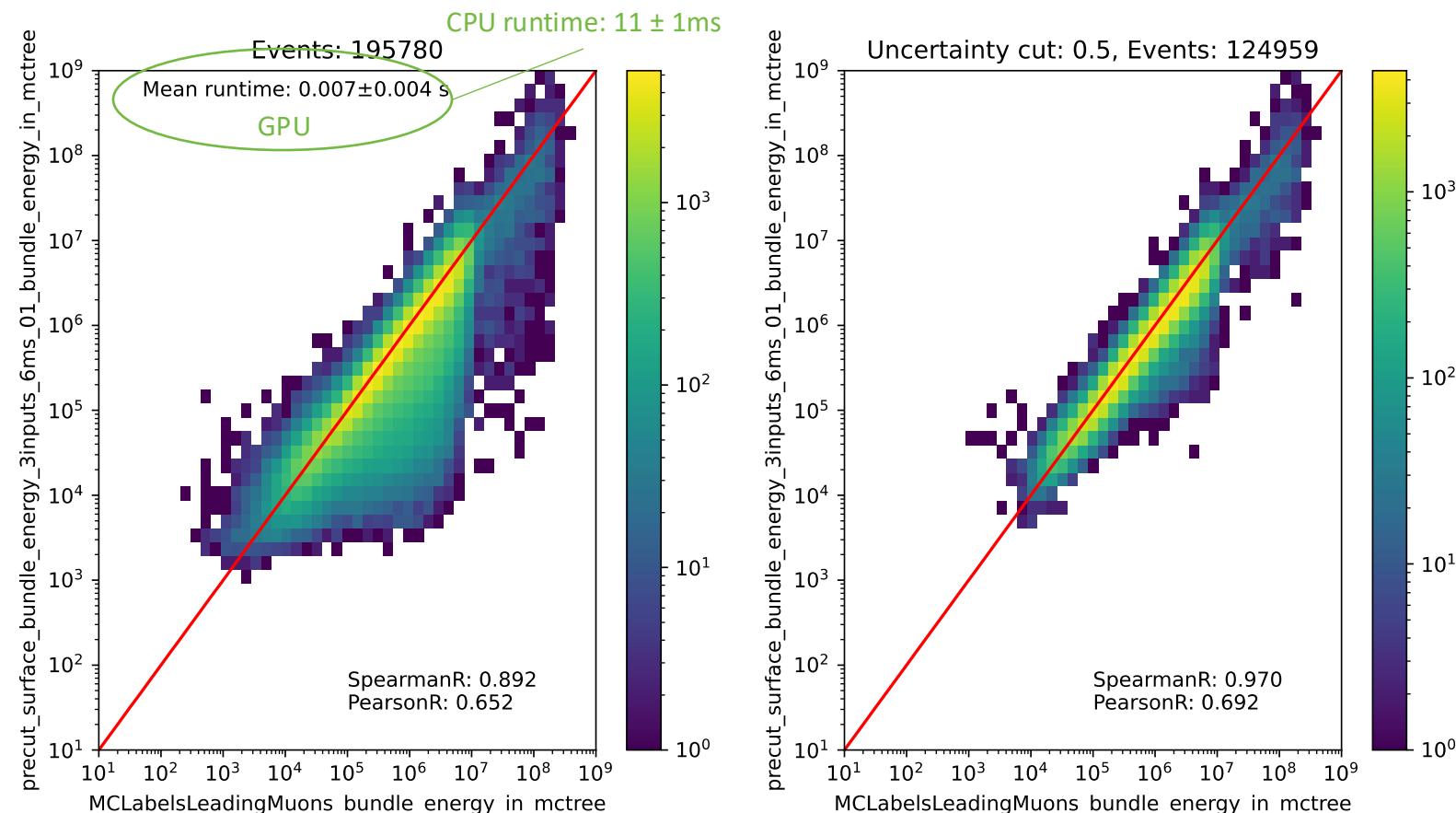
# Simulation w/o coincident events

- for primary energies above 1 TeV, using a time window of 6  $\mu$ s, the chance for a coincident event is  $\leq 1\%$



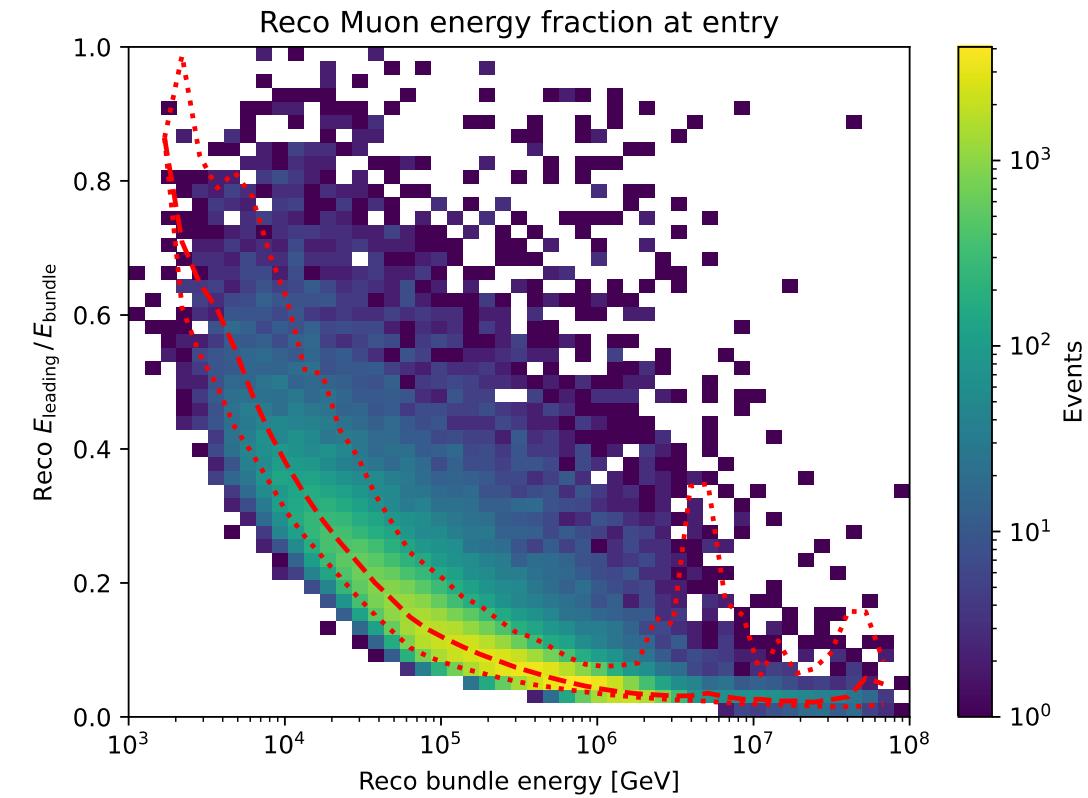
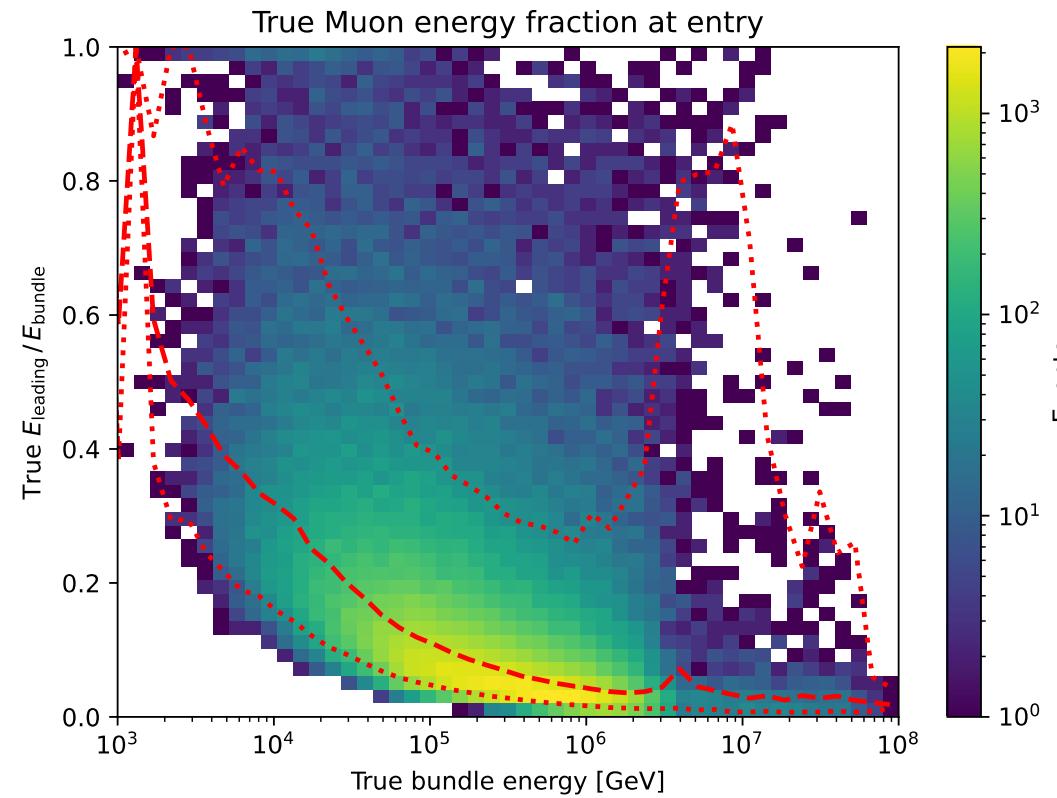
# Bundle energy cut

- Rate after muon filter: 24.62 Hz
- If the process of 1 event needs 1 second, 8h run takes 200h -> needs to be reduced!
- Use small, fast network to remove low energy events -> target rate 125 mHz

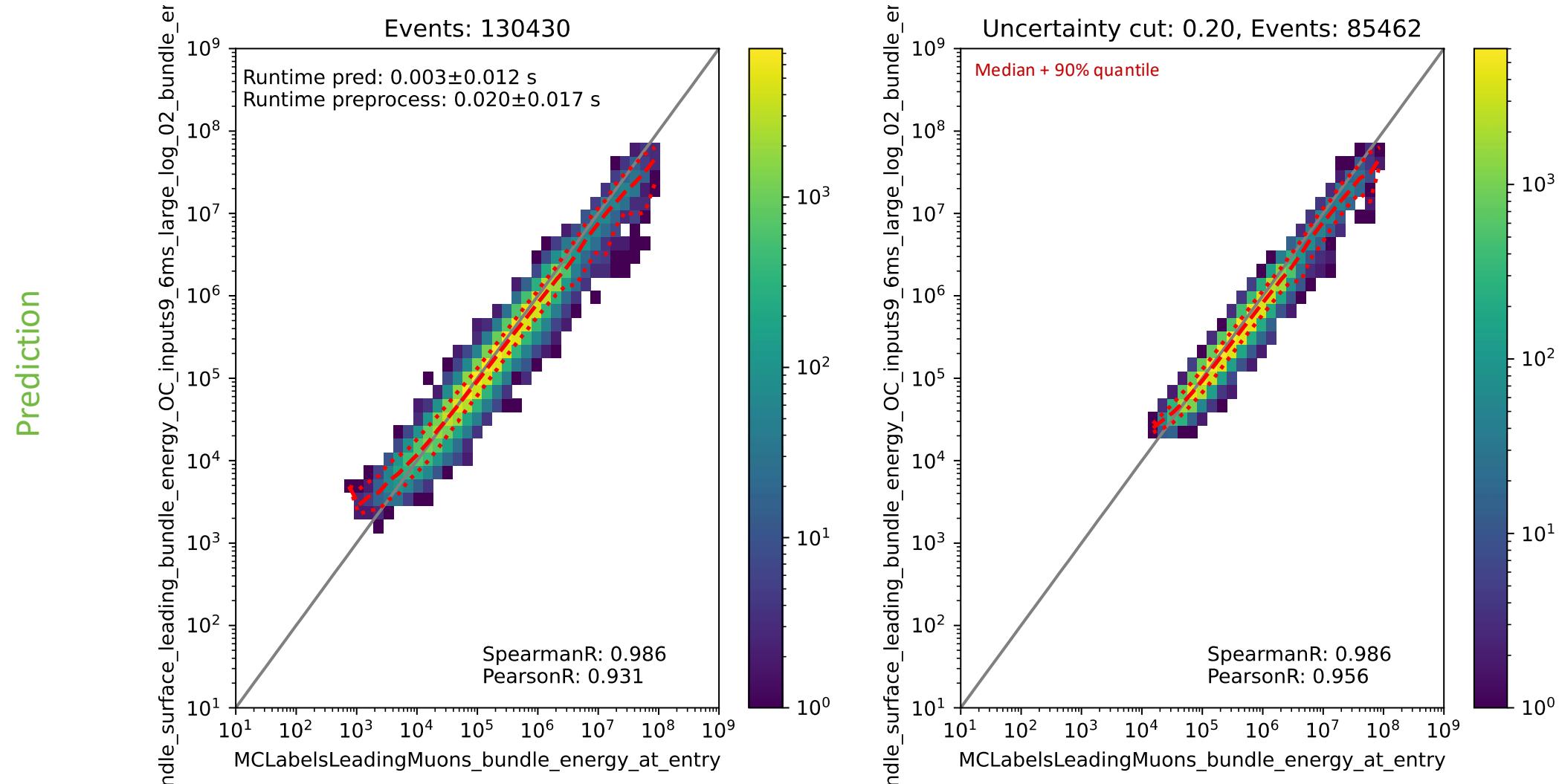


# Leading muon energy fraction - leadingness

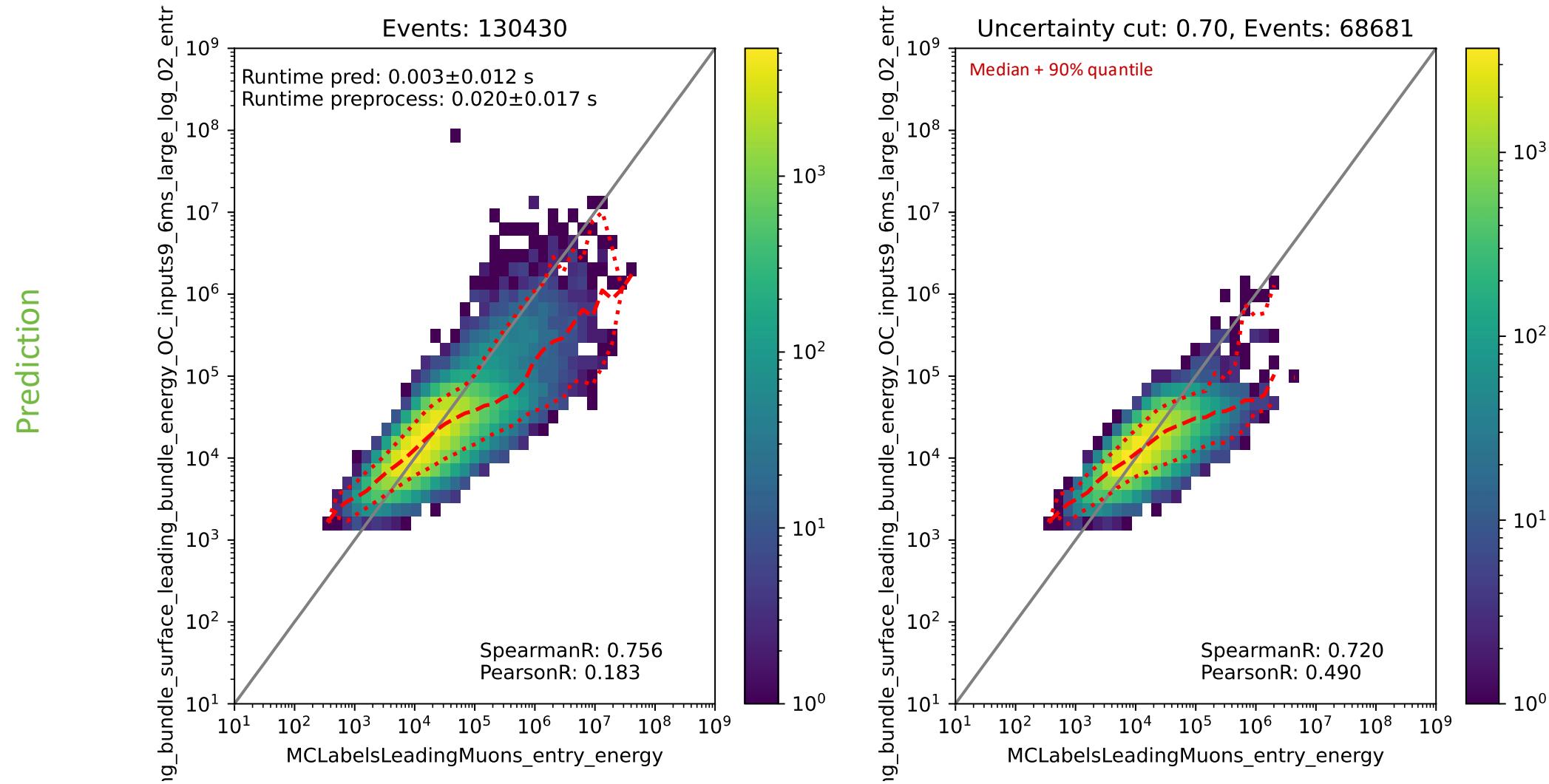
- True muon energy fraction is smeared
- Network tries to predict the median of the distribution



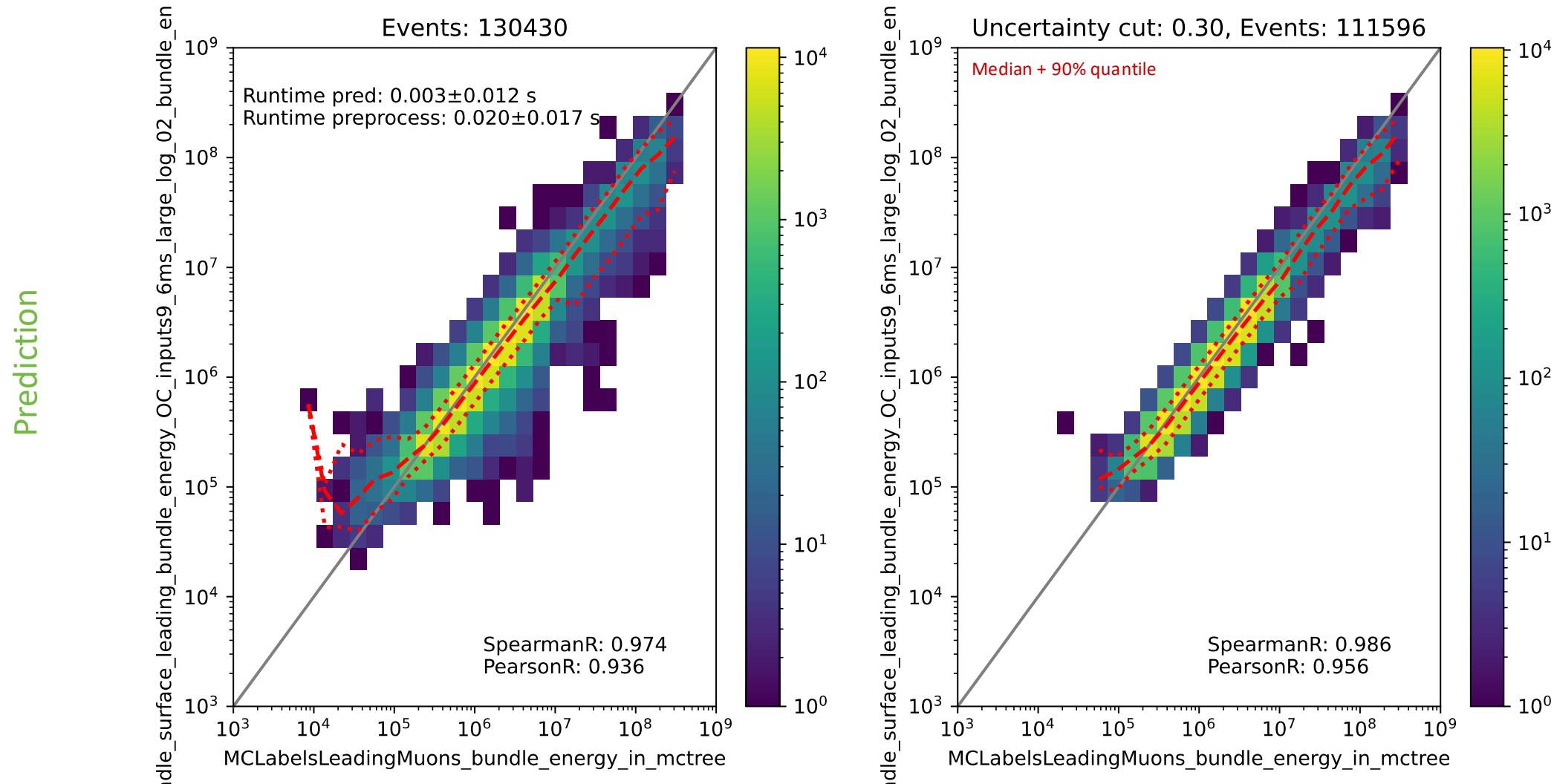
# Bundle energy at entry – 6 $\mu$ s cleaned pulses



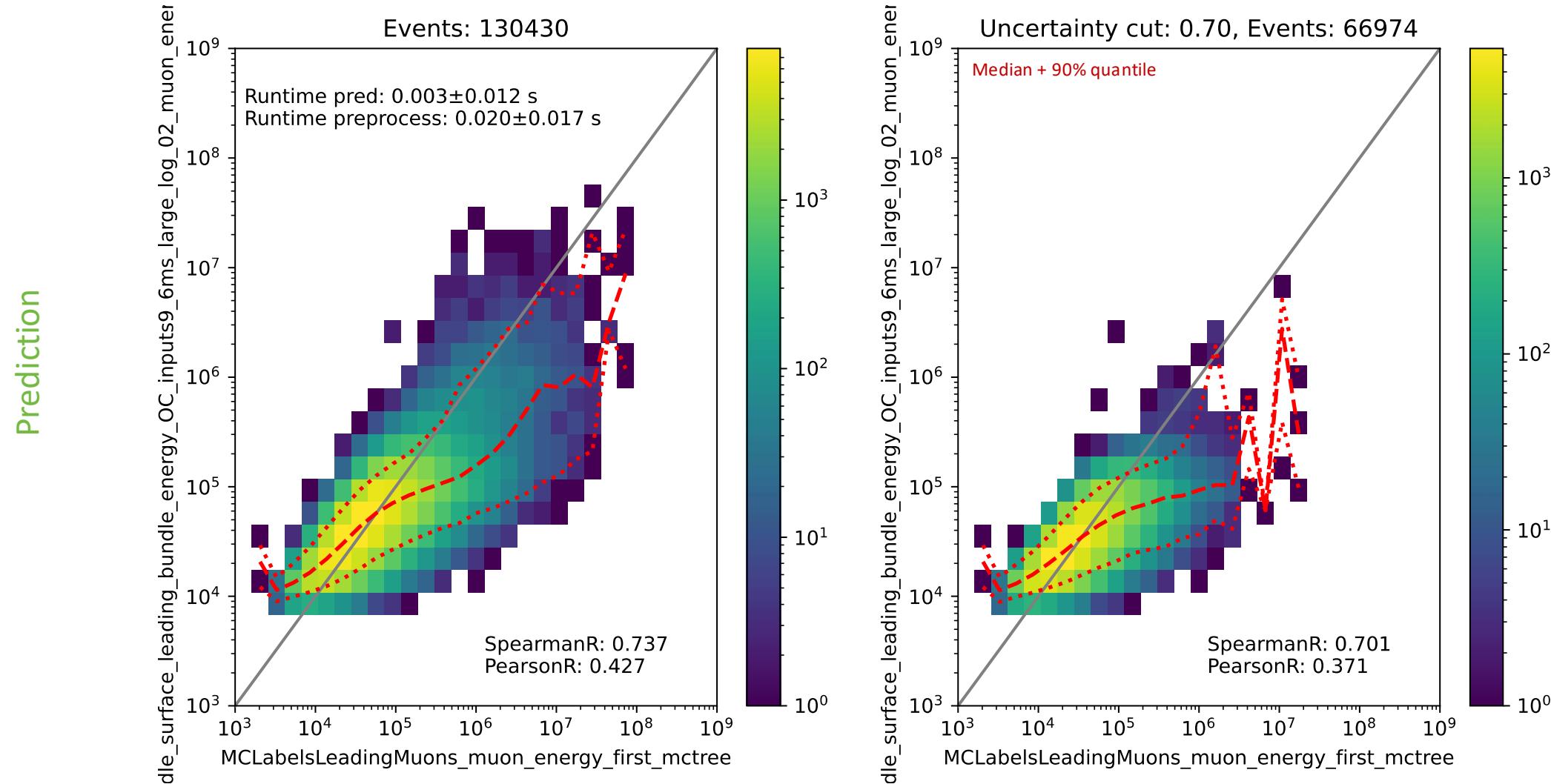
# Leading muon energy at entry – 6 $\mu$ s cleaned pulses



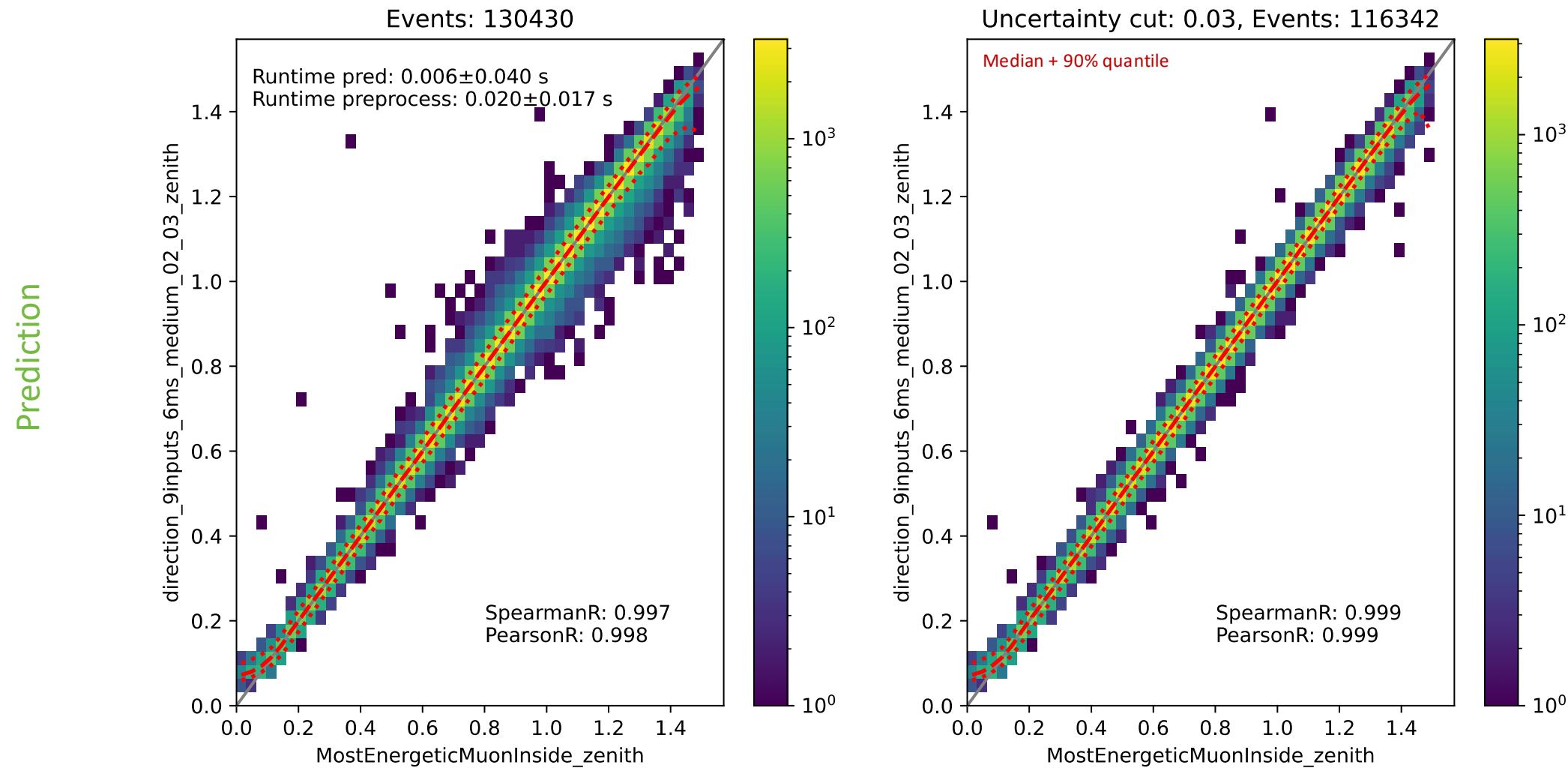
# Bundle energy at surface – 6 $\mu$ s cleaned pulses



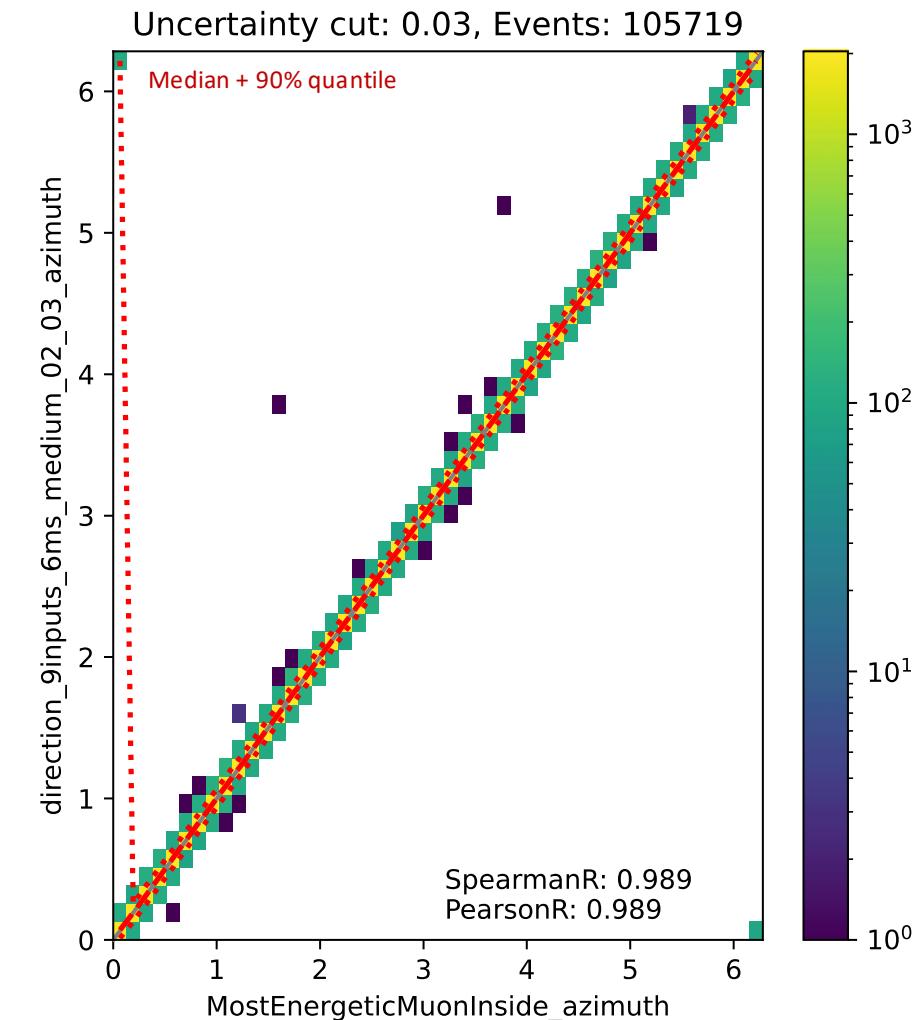
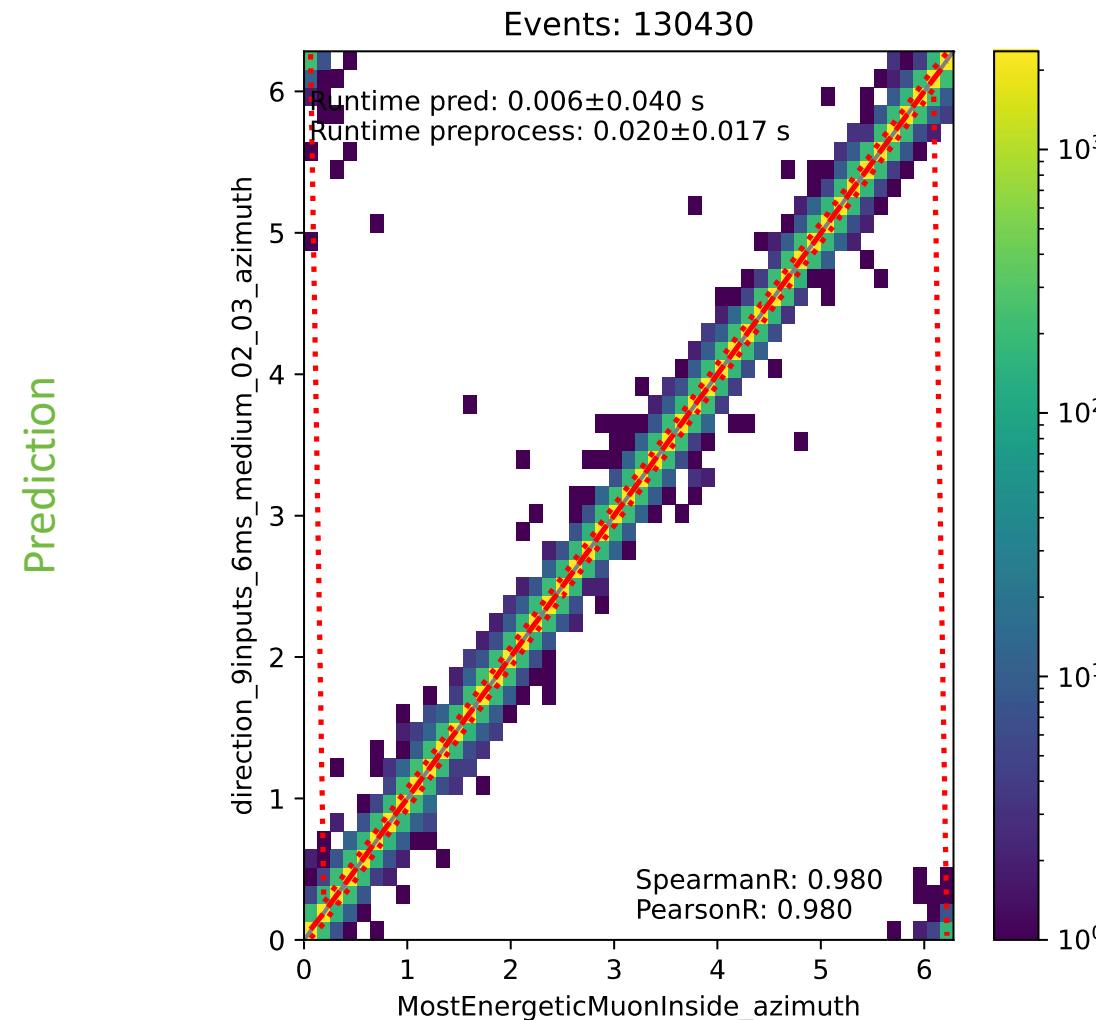
# Leading muon energy at surface – 6 $\mu$ s cleaned pulses



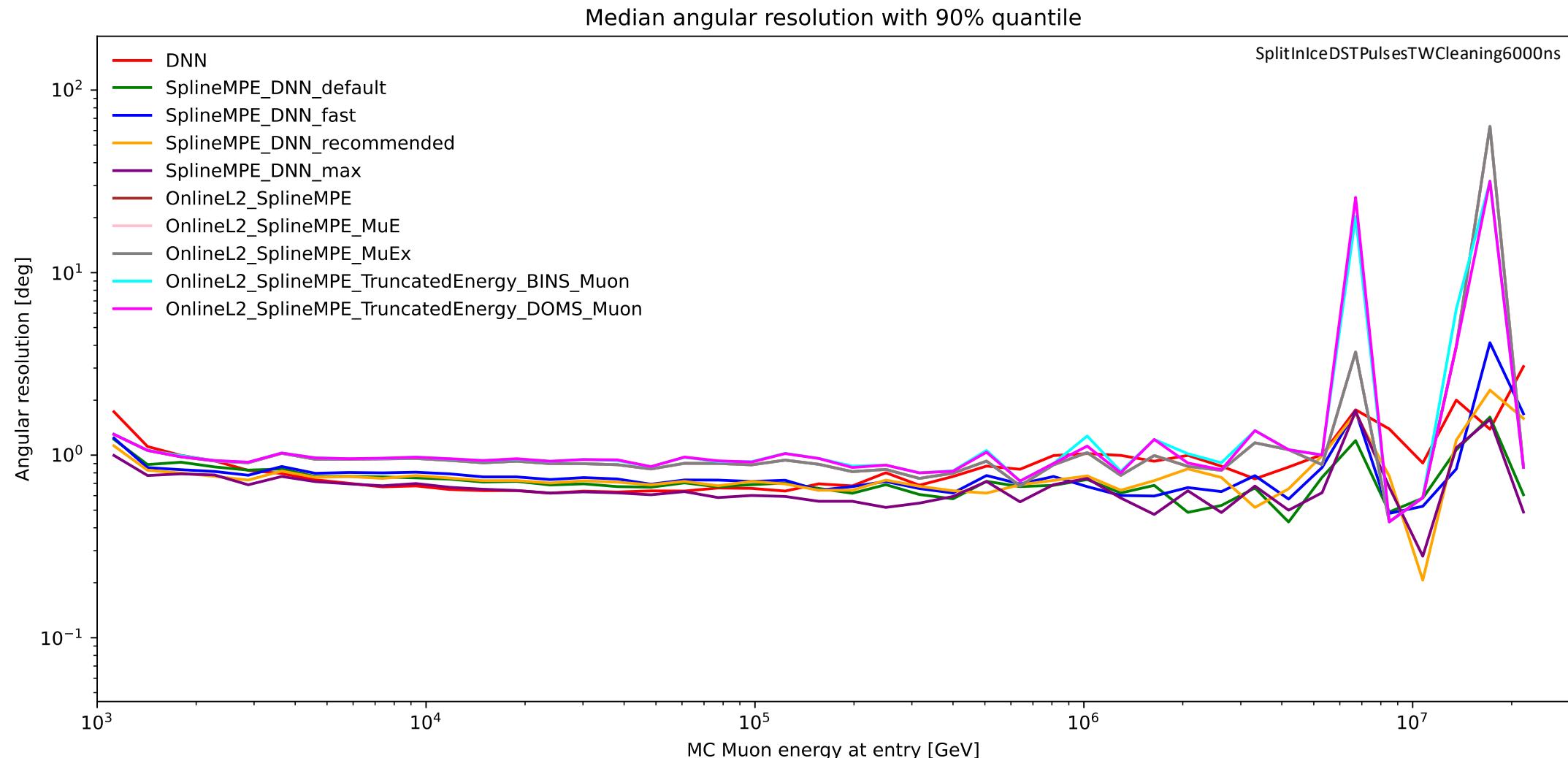
## Zenith – 6 ms cleaned pulses



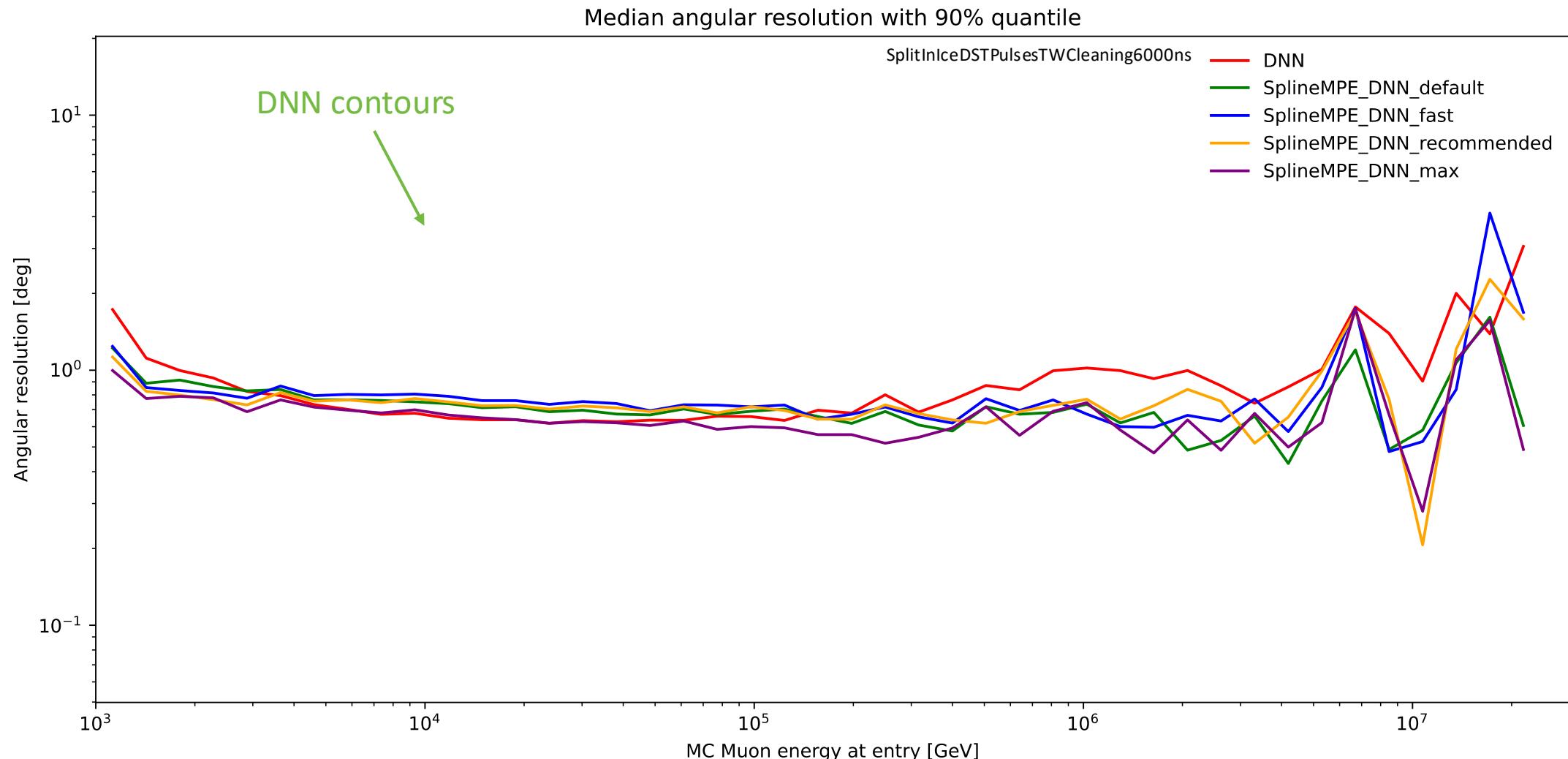
# Azimuth – 6 ms cleaned pulses

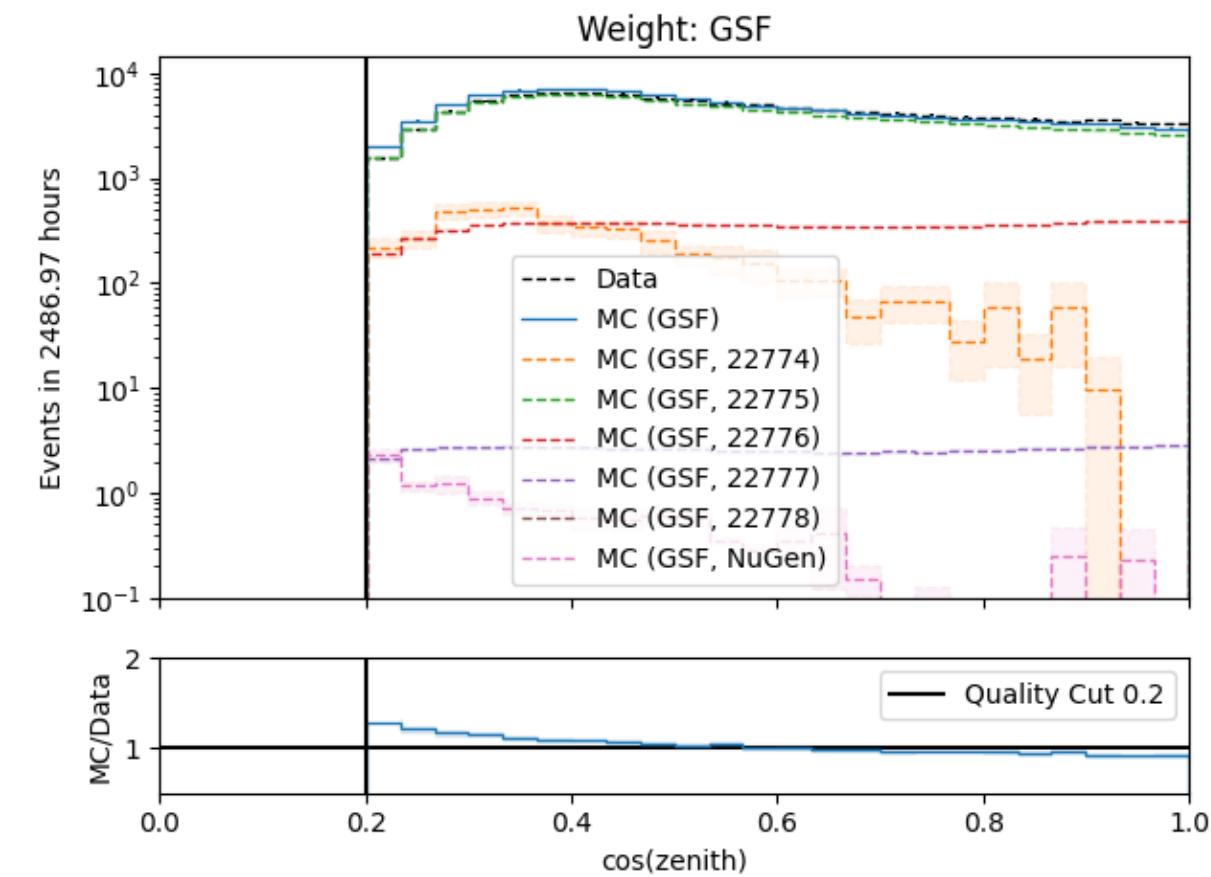
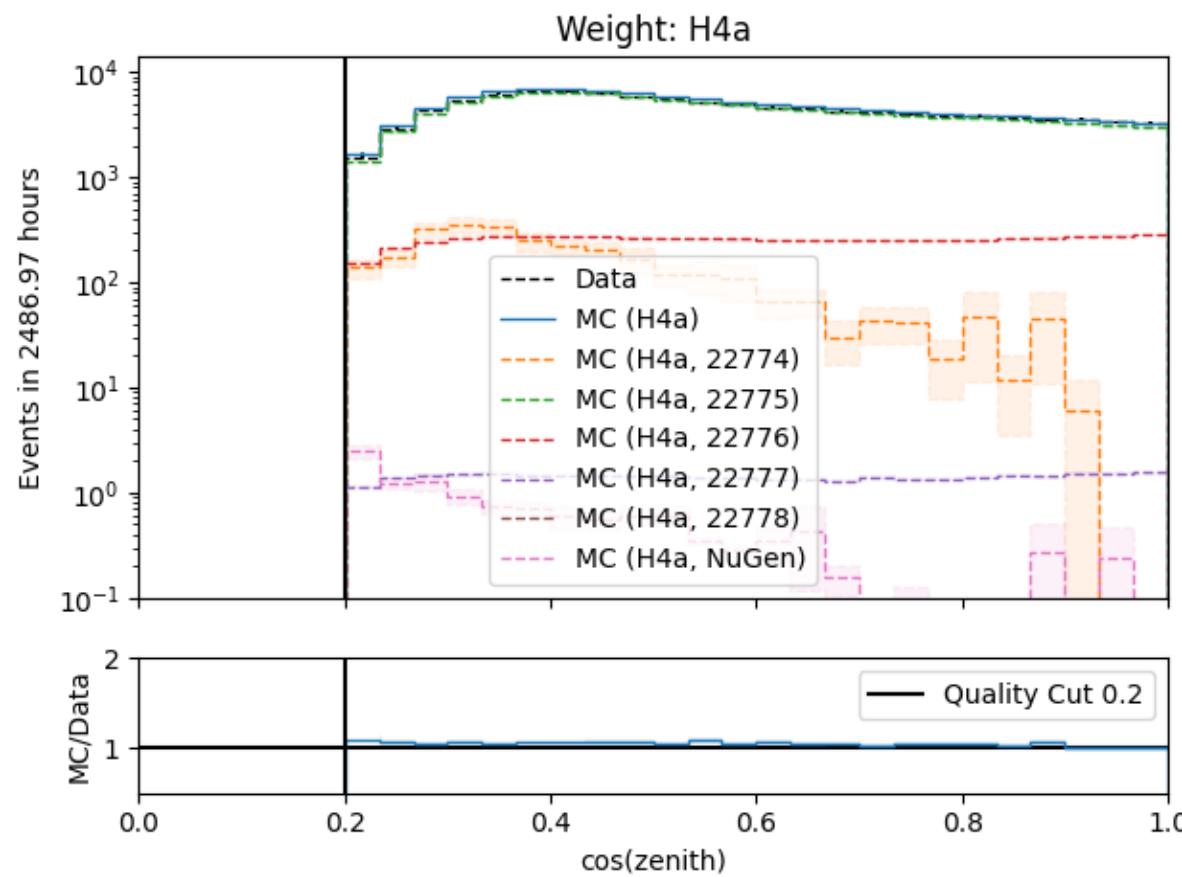


# SplineMPE – DNN and conventional seeds

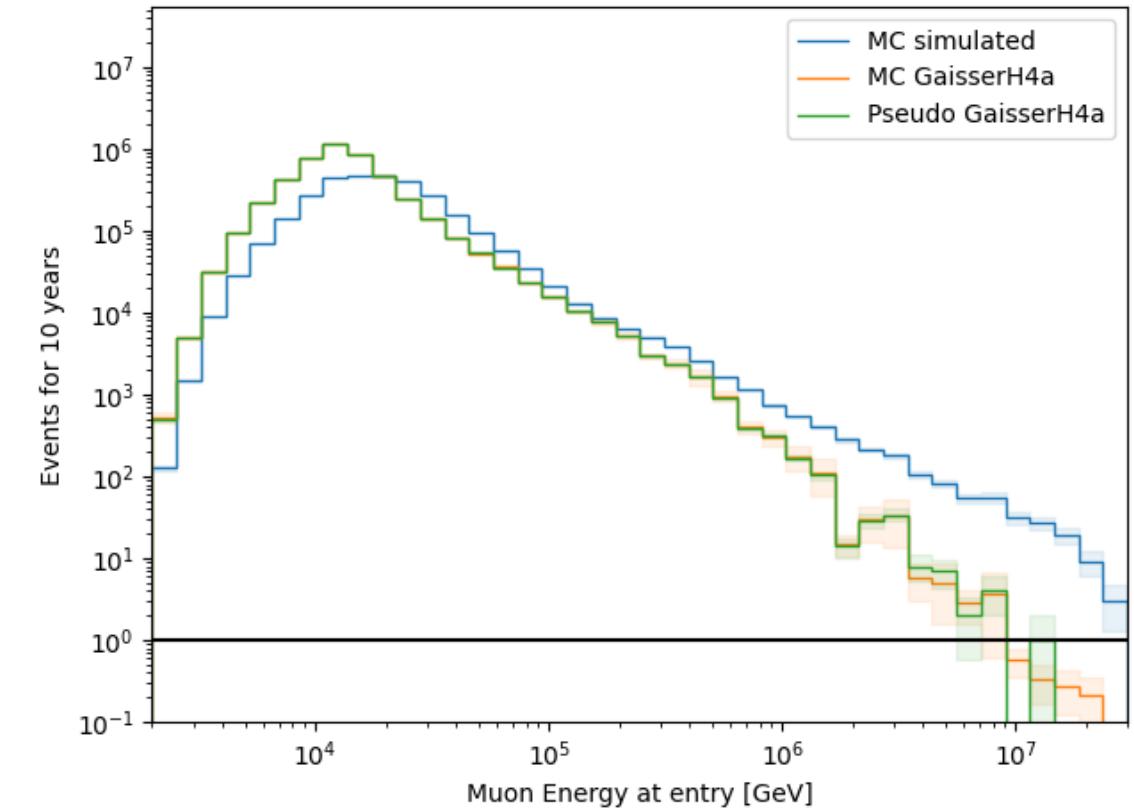
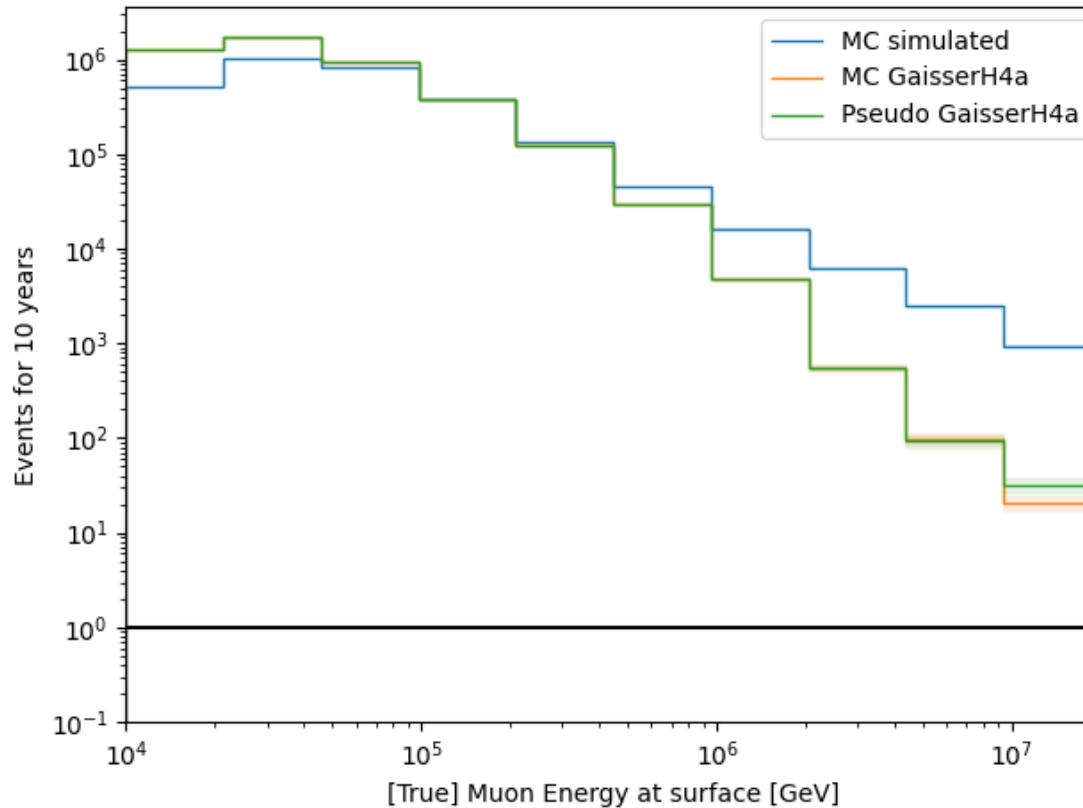


# SplineMPE – DNN seeds

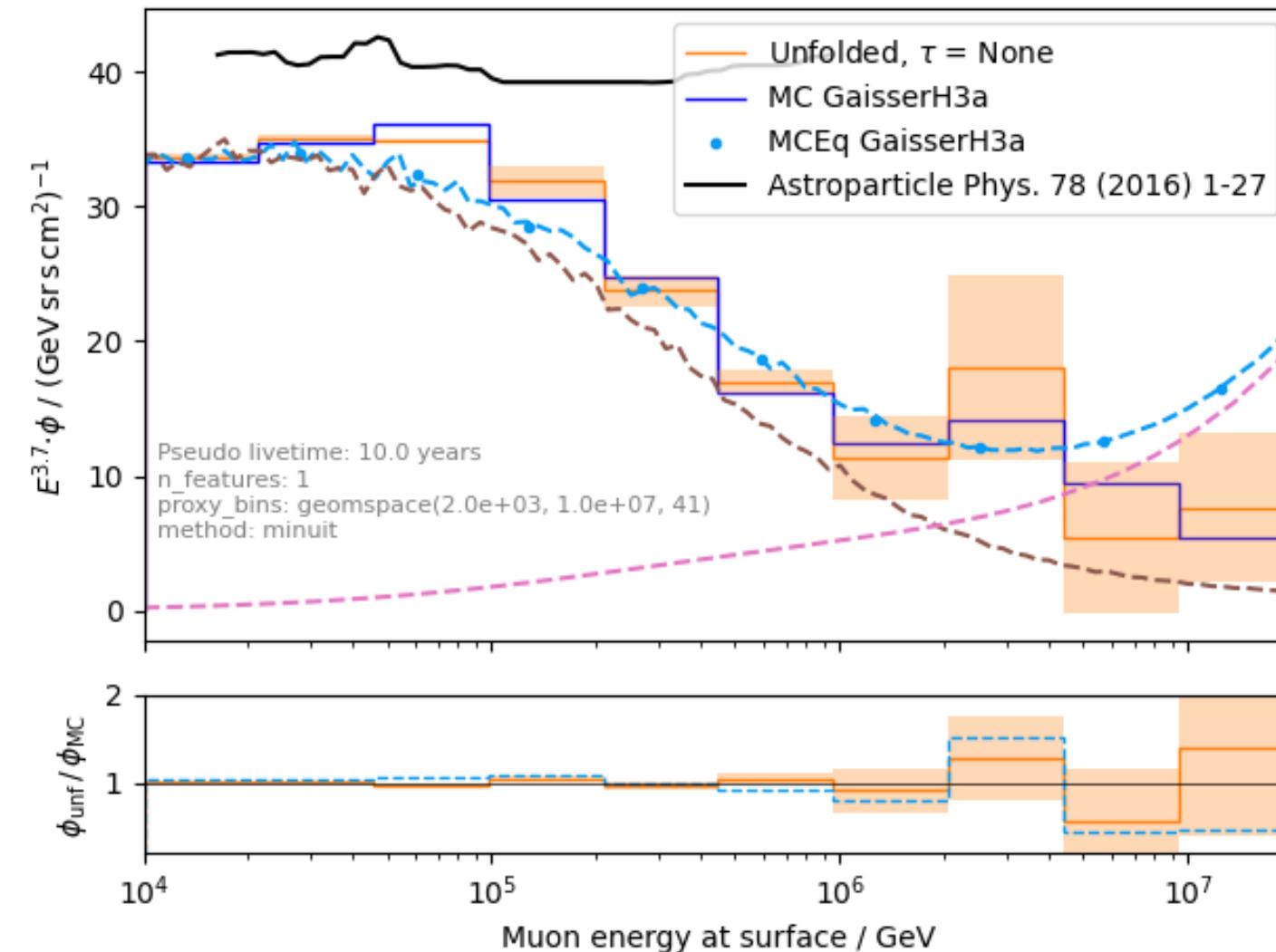


Level5:  $\cos(\text{zenith})$ 

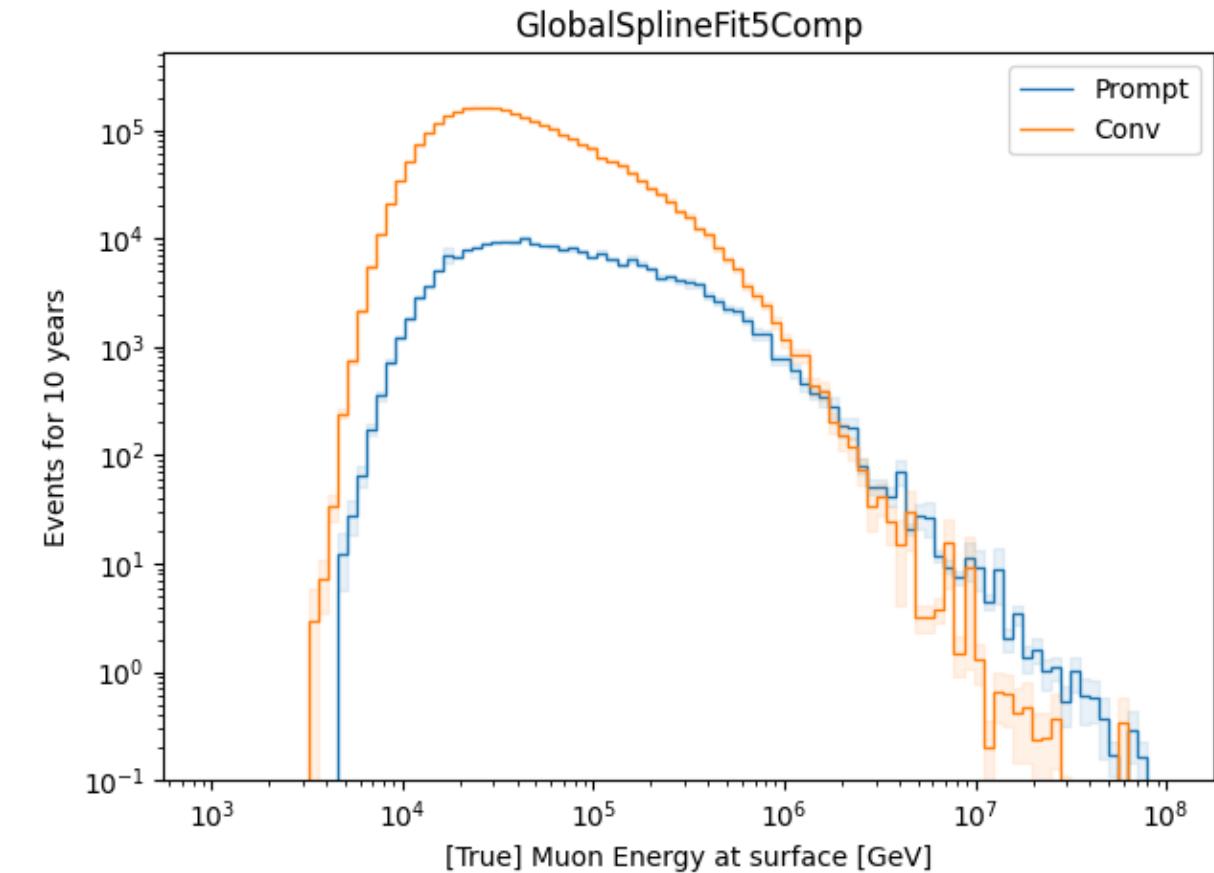
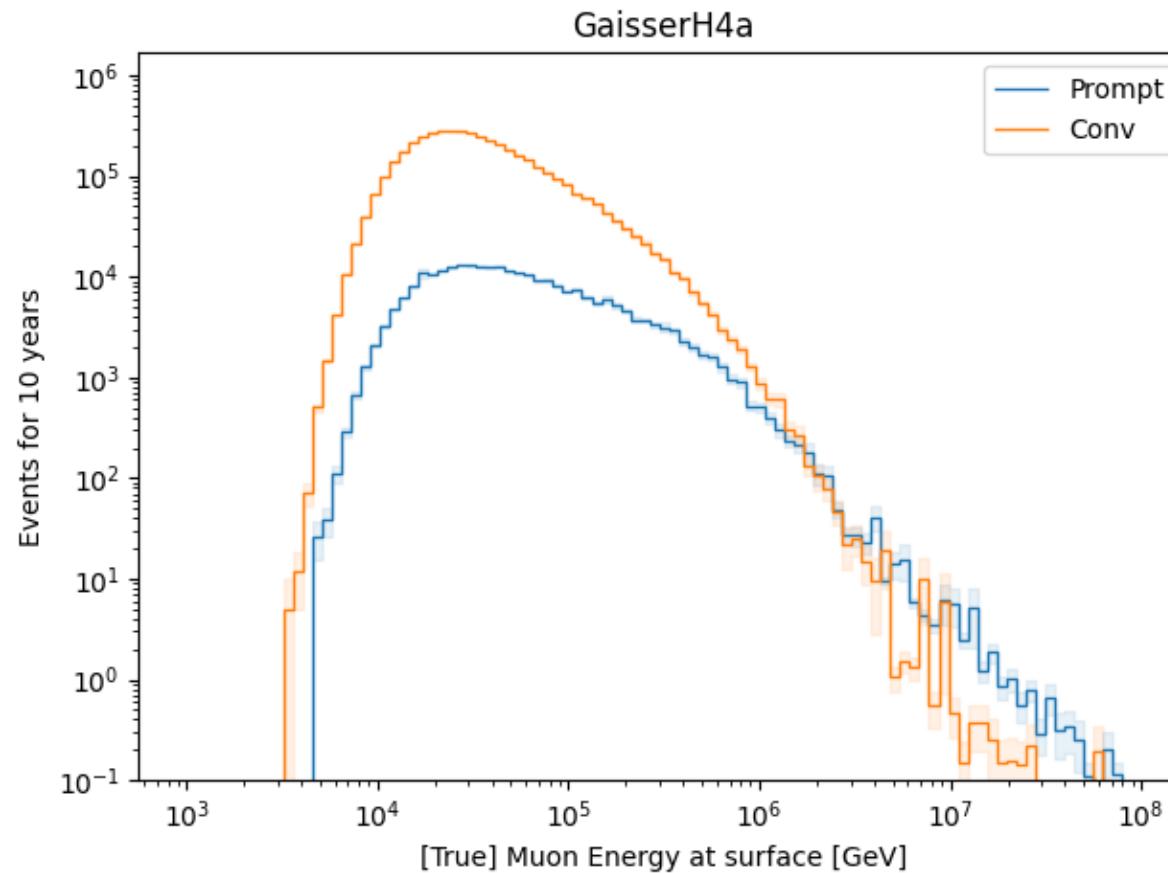
## Proxy and target variable for unfolding – 10 years [H4a]



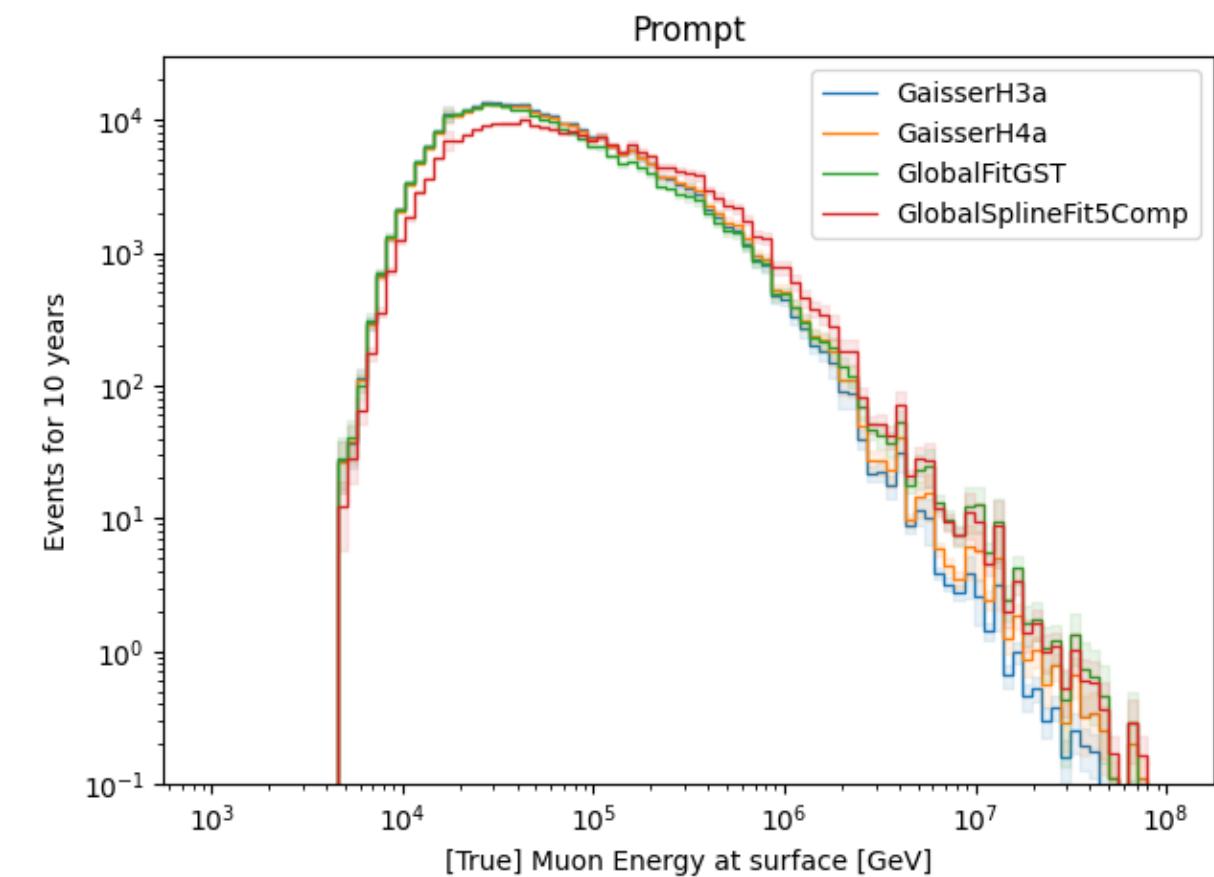
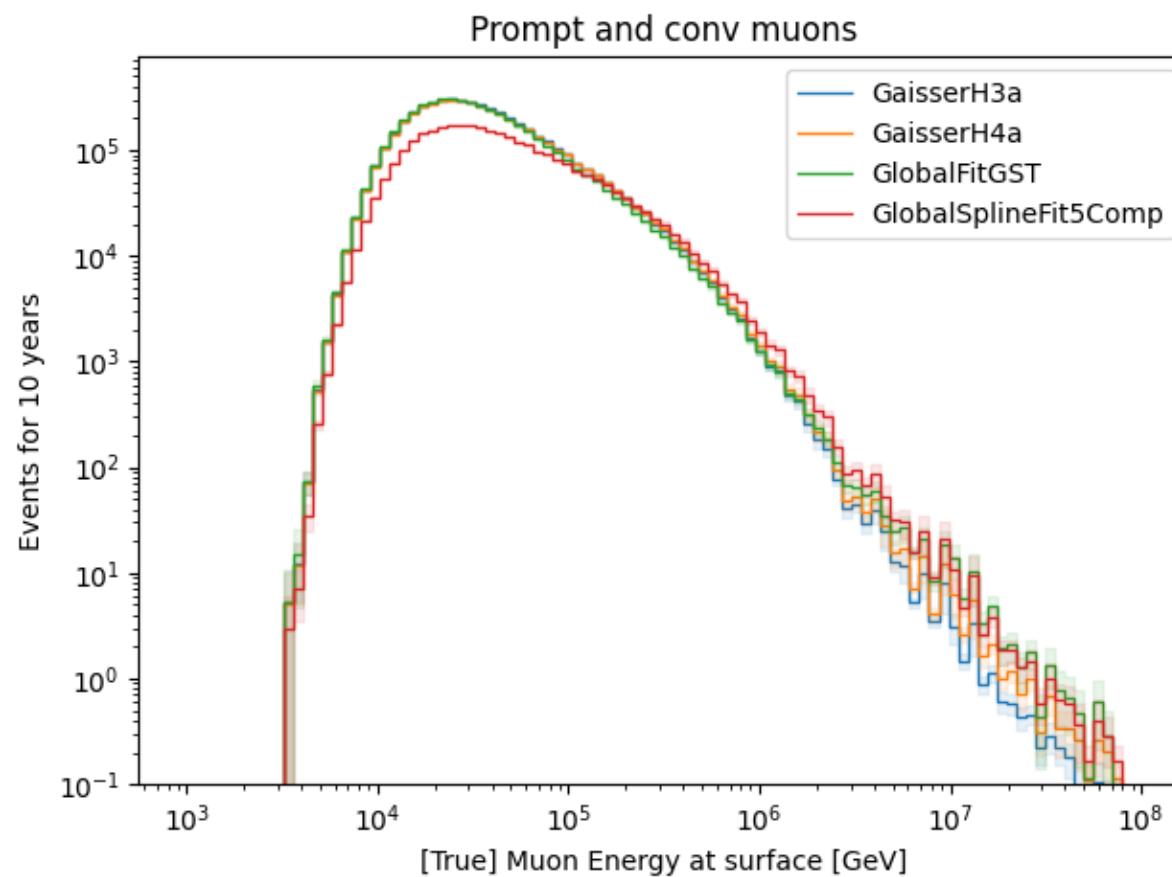
# Unfolding with prompt and conv MCEq



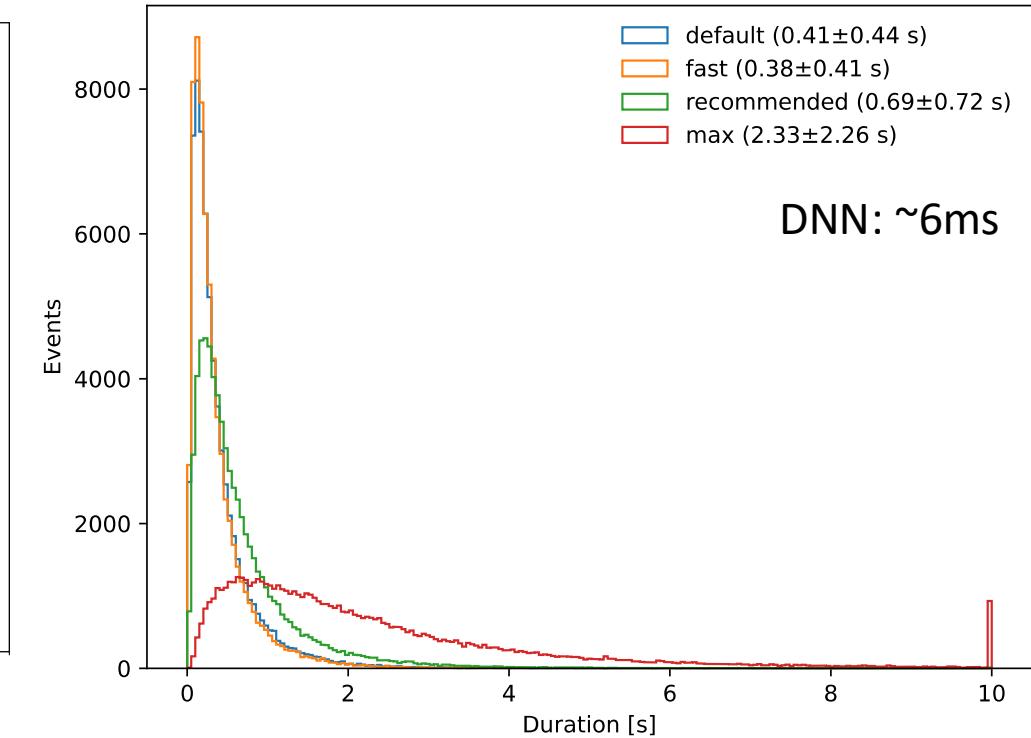
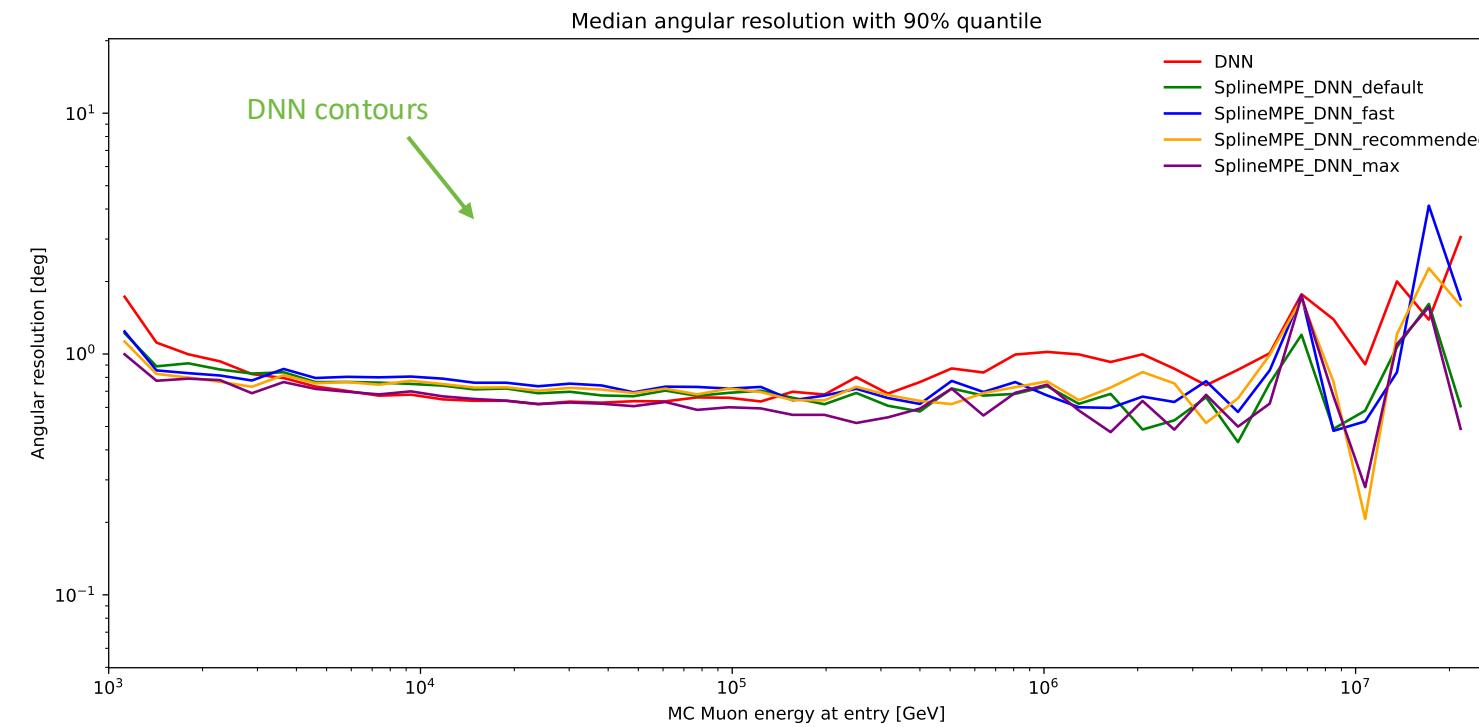
# Prompt and conv: muon energy at surface for 10 years



# Compare all primary fluxes



# Angular reconstructions: DNN vs. SplineMPE

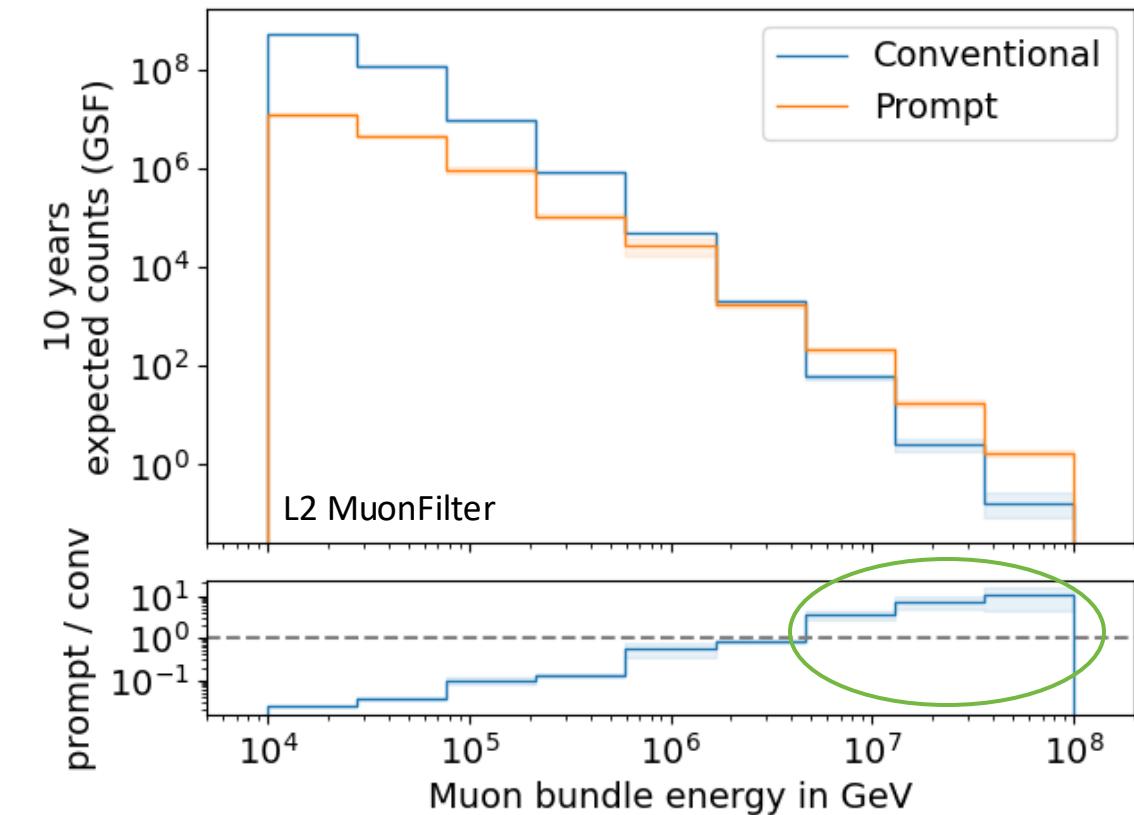
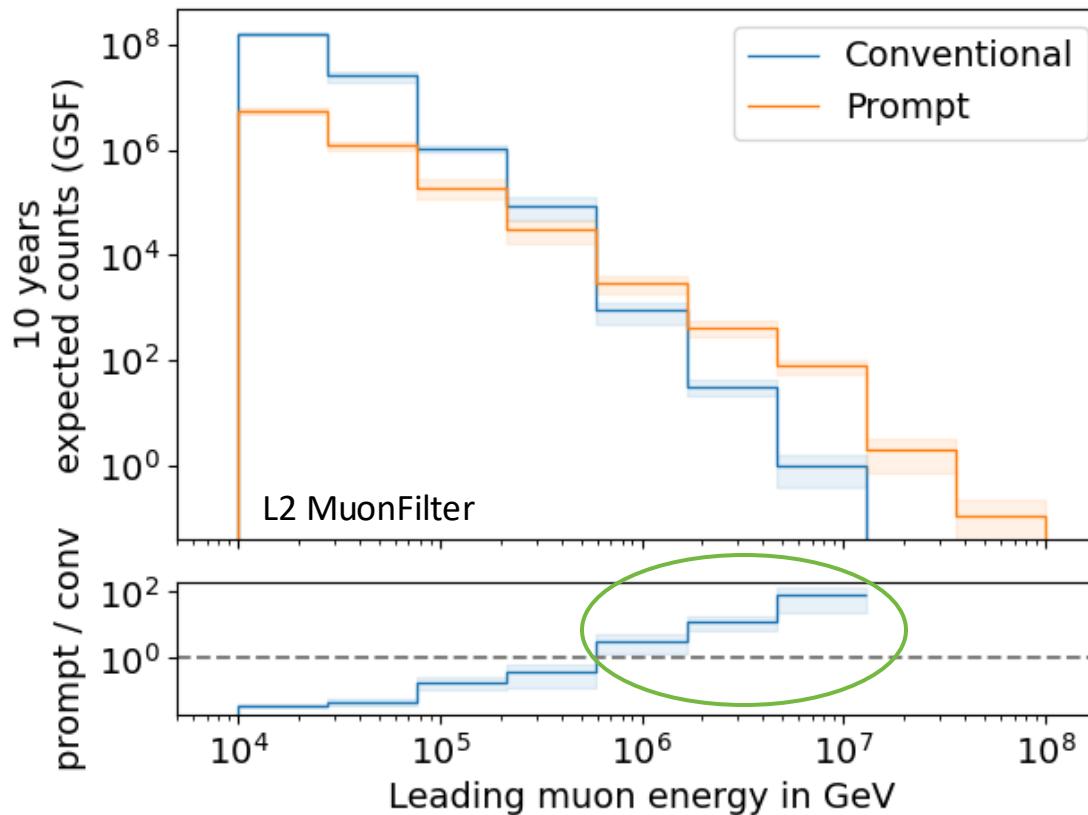


- Only small improvement at energies around 1 PeV
  - Contours are larger
  - Additional runtime
- > Use only DNN reconstruction, since we do not need the best angular resolution

# Forward fit

# Expected muons for 10 years: leading vs. bundle energy (GSF)

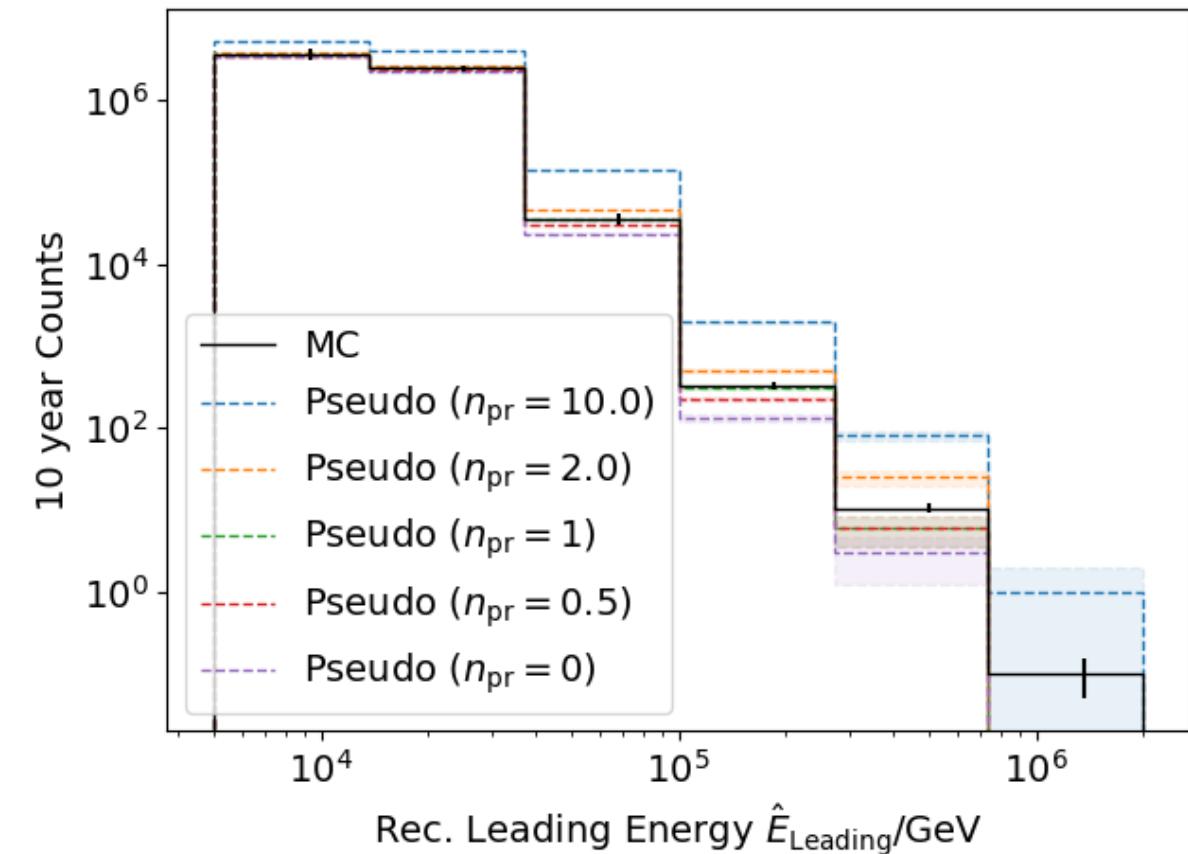
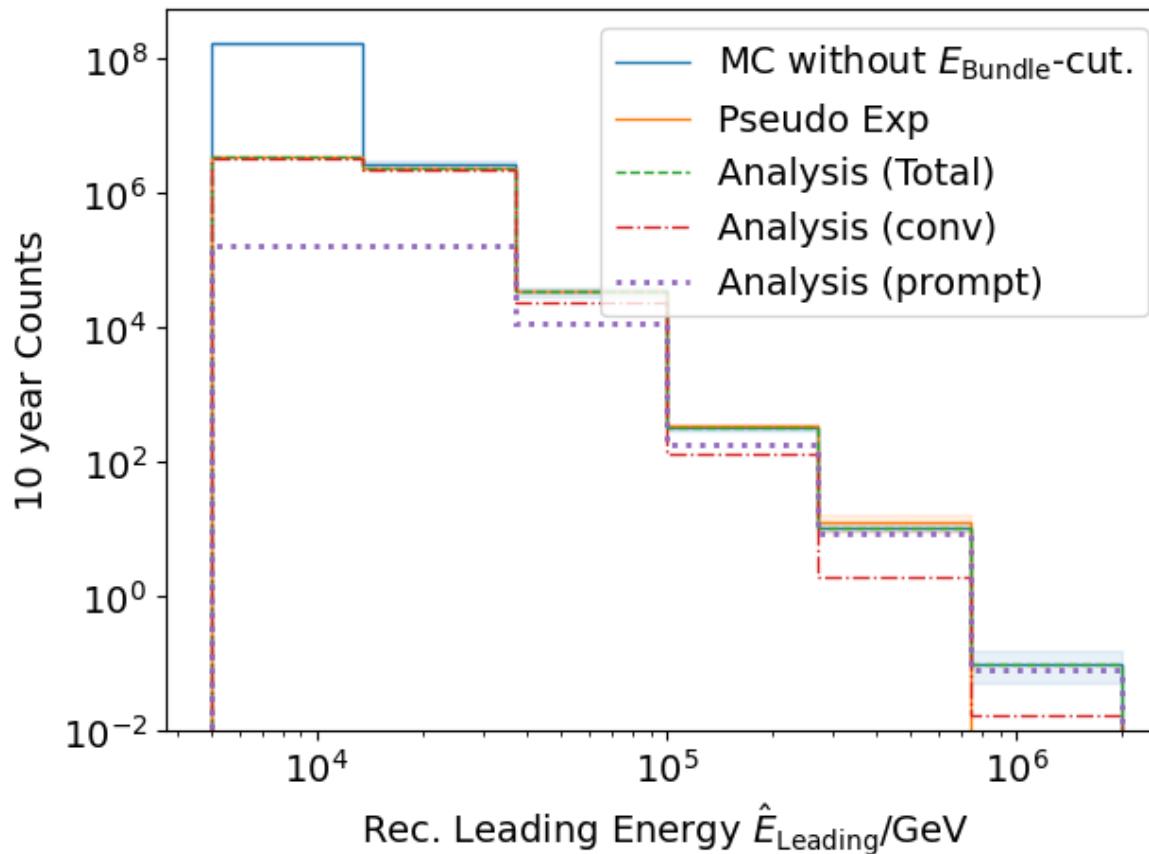
- leading: energy of most energetic muon in a muon bundle
- bundle: sum of energies of all muons of the bundle



- Both leading and bundle energy are sensitive to detect prompt
- Leading muon energy is more sensitive

# Pseudo data sampling

Cuts:  
L2 MuonFilter  
Bundle energy at entry > 100 TeV



➤ Tagging allows scaling of prompt by factor  $n_{\text{pr}}$

# Poisson likelihood fit performed in leading muon energy

Cuts:  
 L2 MuonFilter  
 Bundle energy at entry > 100 TeV

Prompt scaling/normalization

$$C_1^{\text{MC}} = n_{\text{pr}} C_1^{\text{MC,pr}} + n_{\text{conv}} C_1^{\text{MC,conv}}, \dots, C_M^{\text{MC}} = n_{\text{pr}} C_M^{\text{MC,pr}} + n_{\text{conv}} C_M^{\text{MC,conv}}$$

MC counts per bin i

Conv norm = 1

Experimental counts

$$p(C_i) = p_{\text{poisson}}(C_i; \lambda(n_{\text{pr}}) = C_i^{\text{MC}}(n_{\text{pr}})) = \frac{\lambda(n_{\text{pr}})^{C_i} e^{-\lambda(n_{\text{pr}})}}{C_i!}$$

Maximize likelihood

$$\mathcal{L}(n_{\text{pr}}) = \prod_{i=1}^M p(C_i; n_{\text{pr}})$$

Easier:  
 minimize negative  
 log-likelihood

$$-\ln \mathcal{L} = -\sum_{i=1}^M C_i \ln \lambda(n_{\text{pr}}) - \lambda(n_{\text{pr}}) - \ln C_i!$$

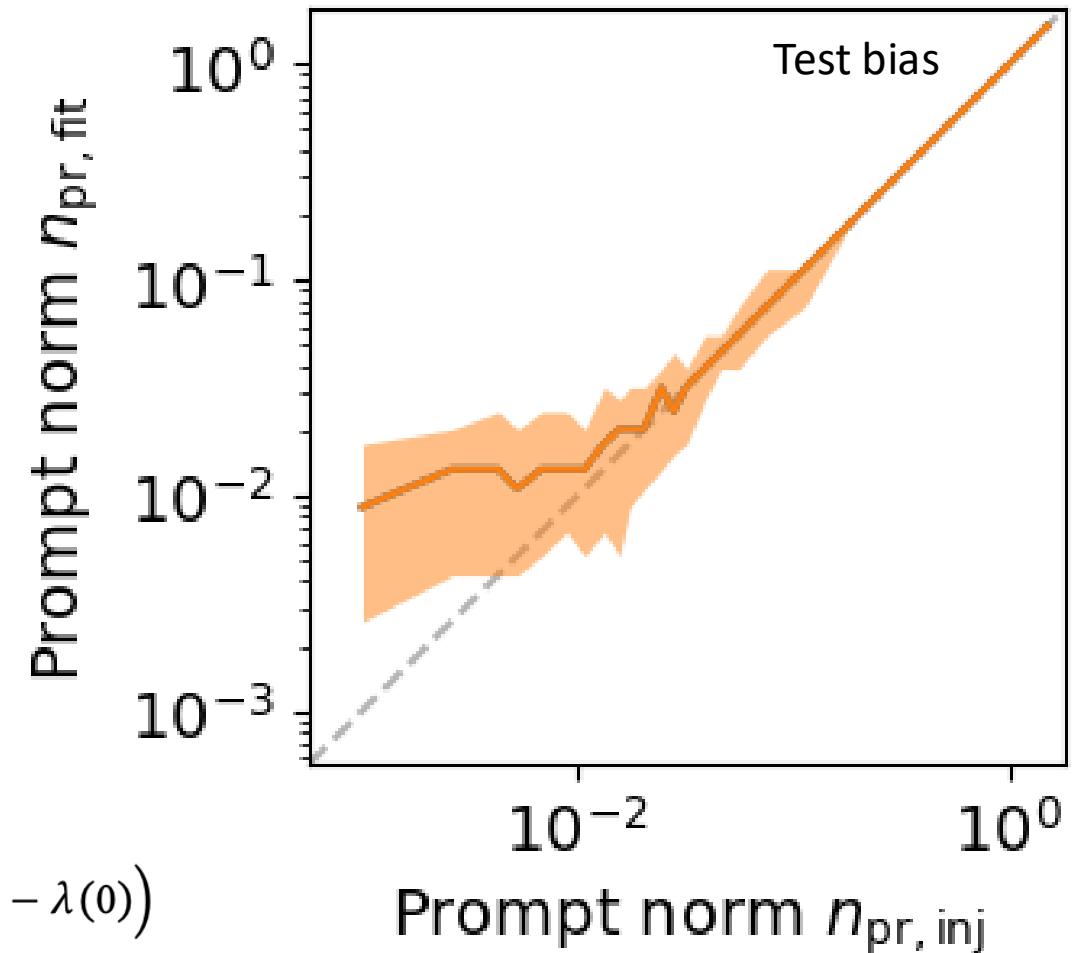
With a constant conv norm:  
 bin counts depend only on prompt norm  
 = expectation value per bin

$\Lambda = -2 \ln \frac{\mathcal{L}(n_{\text{pr}} = \hat{n}_{\text{pr}})}{\mathcal{L}(n_{\text{pr}=0})} = -2 \sum_{i=1}^M C_i (\ln \lambda(\hat{n}_{\text{pr}}) - \ln \lambda(0)) - (\lambda(n_{\text{pr}}) - \lambda(0))$

Test statistic for Wilks' theorem

Null hypothesis: no prompt

pascal.gutjahr@tu-dortmund.de



➤ Bias starts at a prompt  
 normalization of ~0.1

# Discovery potential and sensitivity

Expectation for 1 year:

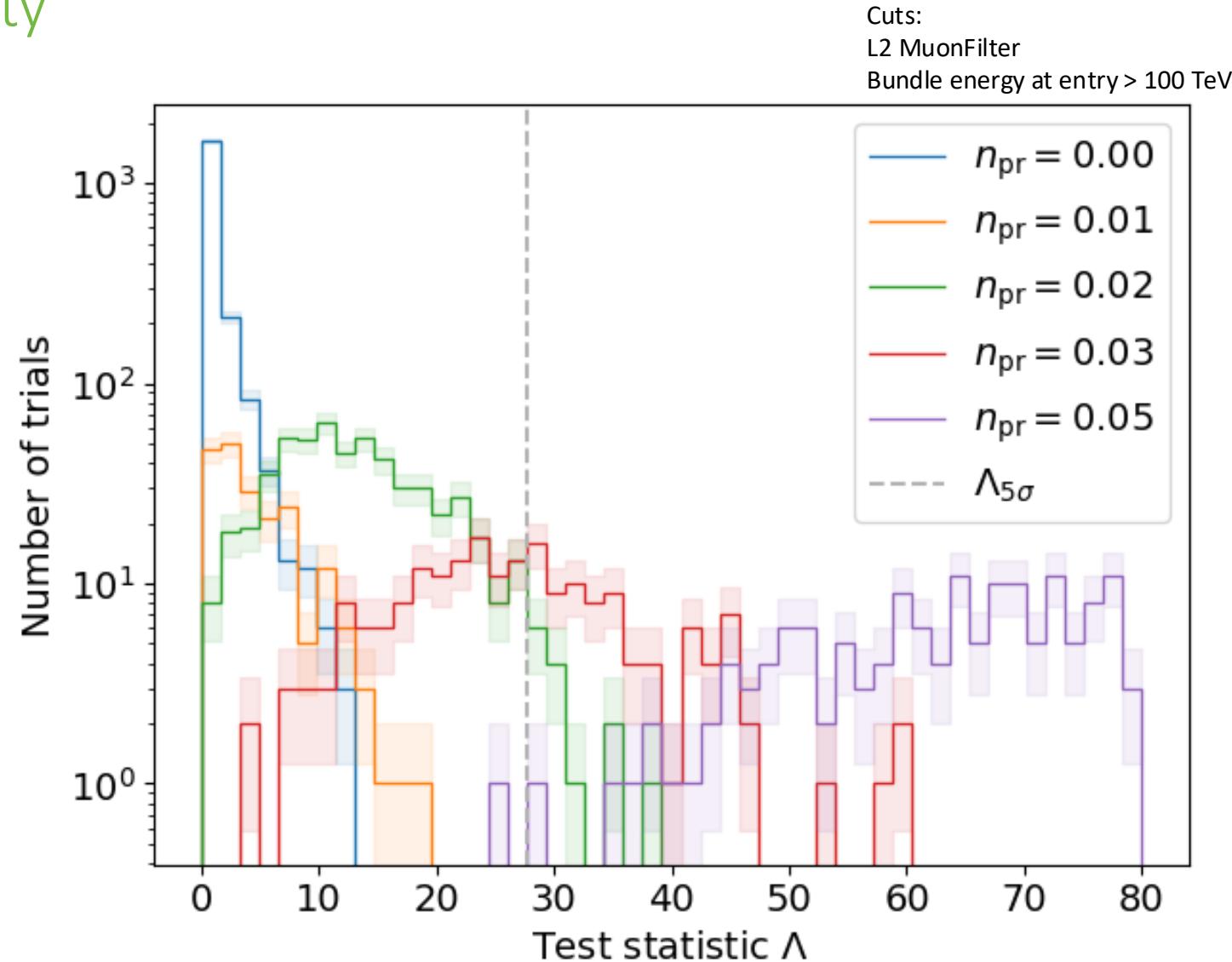
- 5 sigma discovery potential:  $0.102 \pm 0.005$
- Sensitivity:  $0.024 \pm 0.001$

Expectation for 10 years:

- 5 sigma discovery potential:  $0.032 \pm 0.001$
- Sensitivity:  $0.007 \pm 0.000$

Caution:

- Limited MC statistics -> events are oversampled in pseudo dataset
- No systematic uncertainties



# Input data per DOM

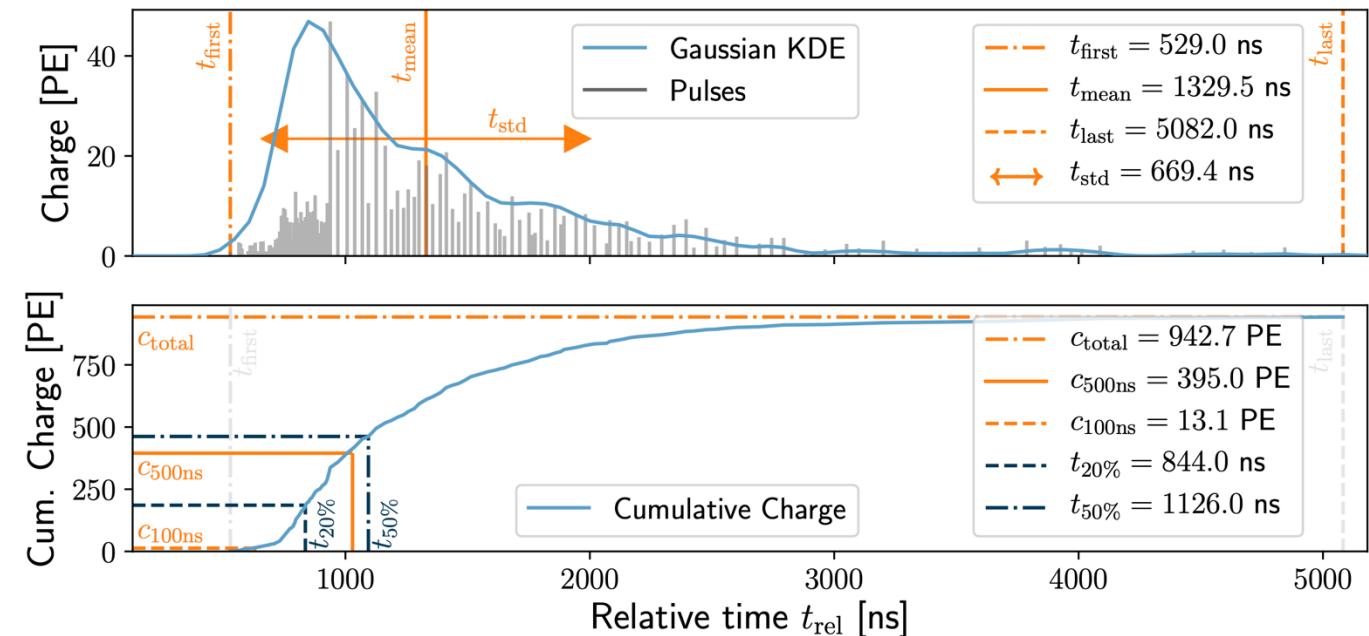
10.1088/1748-0221/16/07/P07041

## 3 inputs

- $c_{\text{total}}$ : Total charge
  - Sum of charge
- $t_{\text{first}}$ : Relative time of first pulse
  - Relative to total time offset, calculated as the charge weighted mean time of all pulses
- $t_{\text{std}}$ : Standard deviation of first pulse
  - Charge weighted standard deviation of pulse times relative to total time offset

## 9 inputs

- $t_{\text{last}}$ : Relative time of last pulse
  - Relative to total time offset, calculated as the charge weighted mean time of all pulses
- $t_{20\%}$ : Relative time of 20% charge
  - Relative to total time offset, calculated as the charge weighted mean time of all pulses
- $t_{50\%}$ : Relative time of 50% charge
  - Relative to total time offset, calculated as the charge weighted mean time of all pulses
- $t_{\text{mean}}$ : Mean time
  - Charge weighted mean time of all pulses relative to total time offset
- $c_{500\text{ns}}$ : Charge at 500ns
  - Sum of charge after 500ns
- $c_{100\text{ns}}$ : Charge at 100ns
  - Sum of charge after 100ns



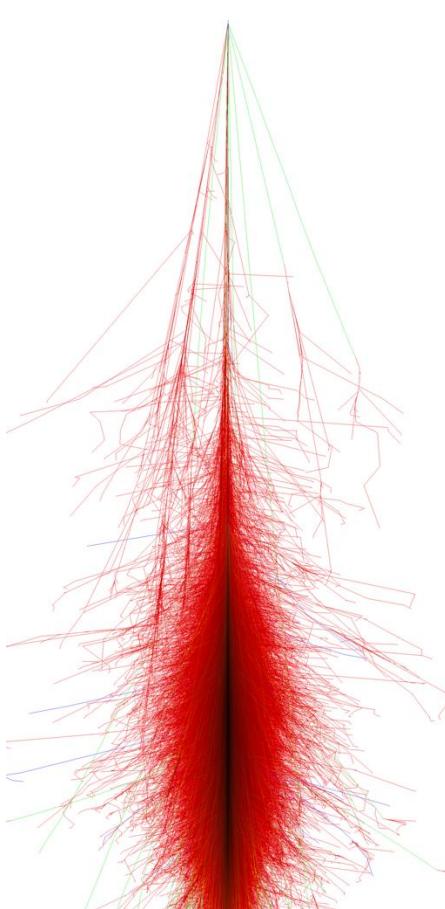
## Input pulses

- SplitInIceDSTPulses
- SplitInIceDSTPulsesTWCleaning6000ns
- (DNN framework performs an internal cleaning)

## Training datasets

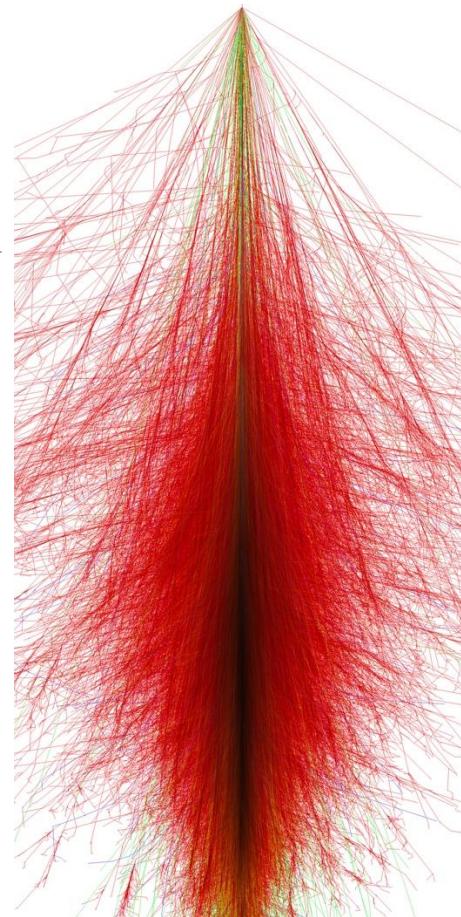
- 20904
- 21962
- 22020
- 22187

## Air shower – 10 TeV

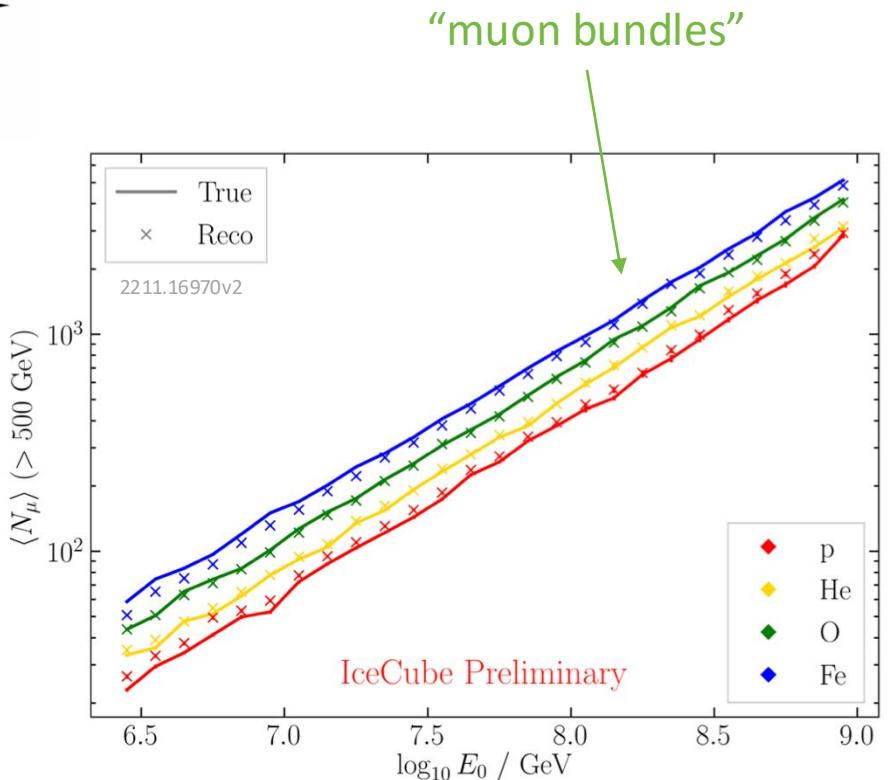
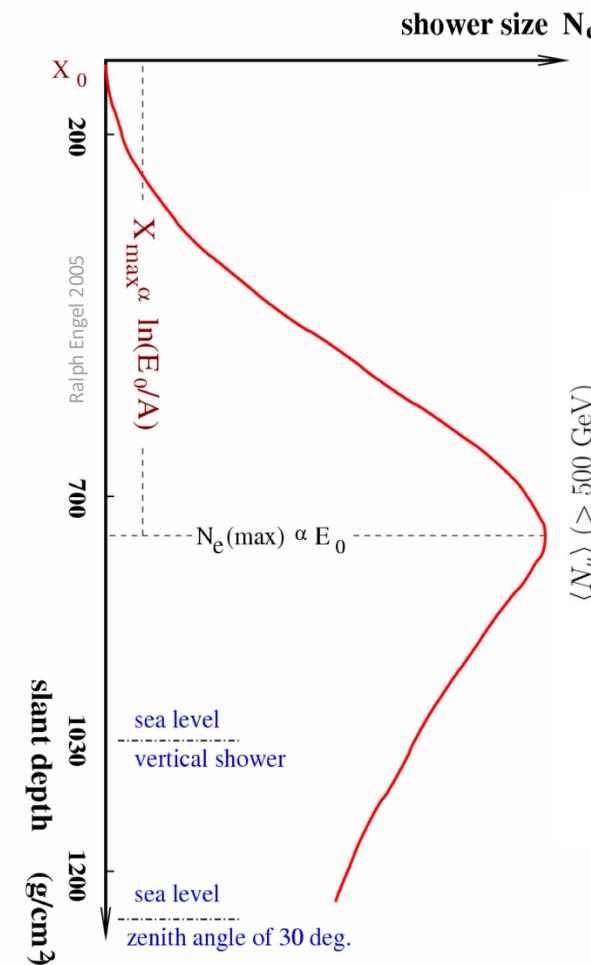


Proton

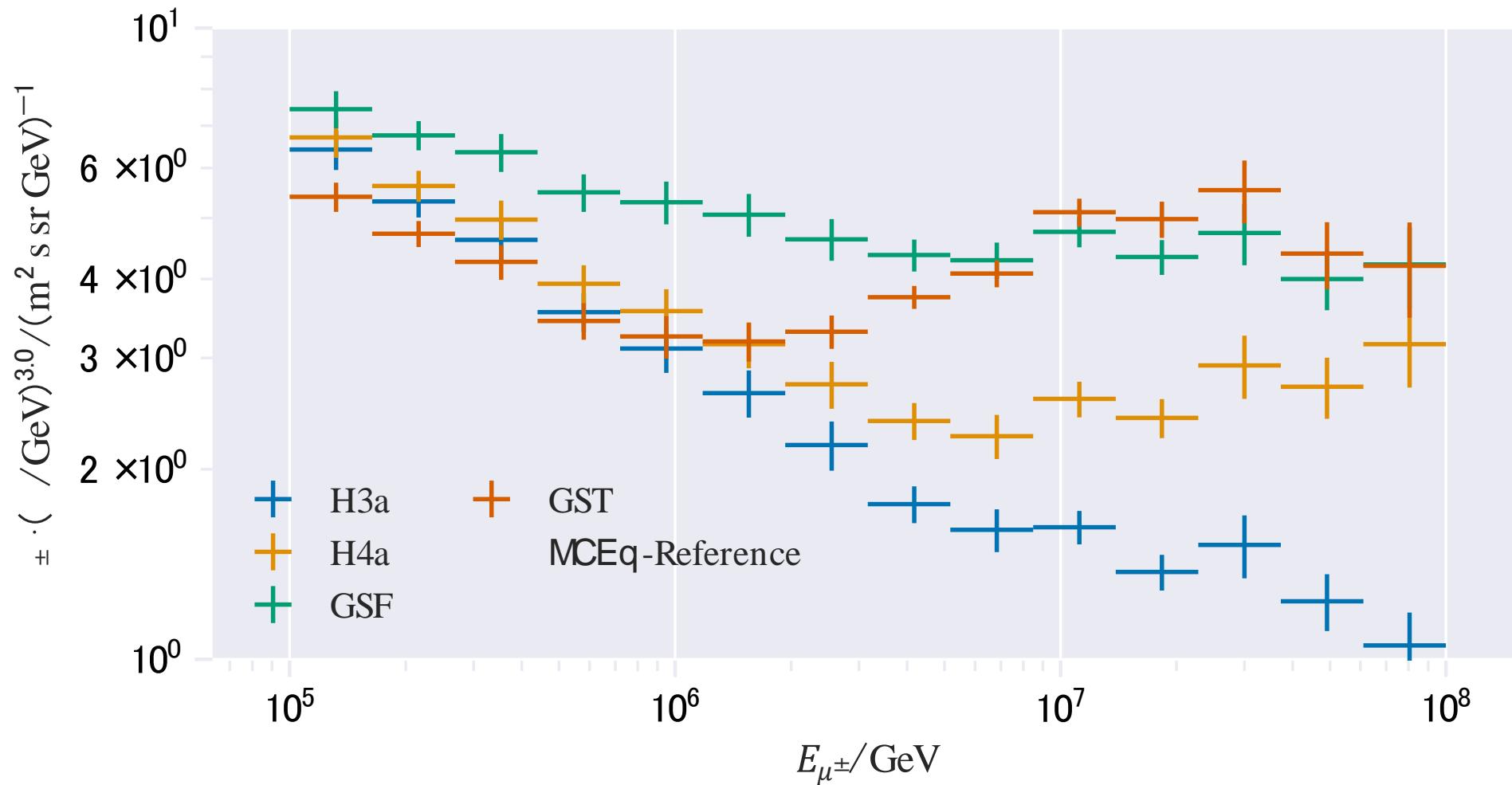
zeuthen.desy.de



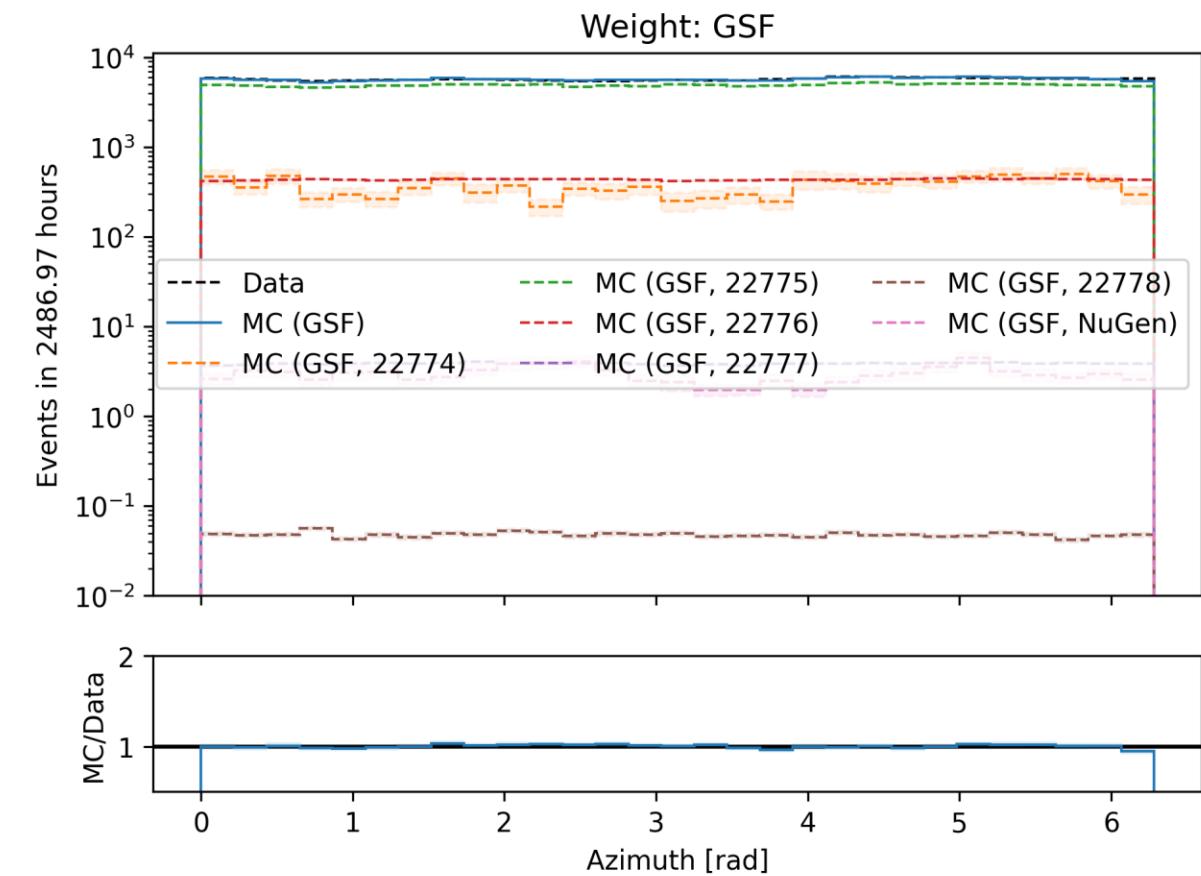
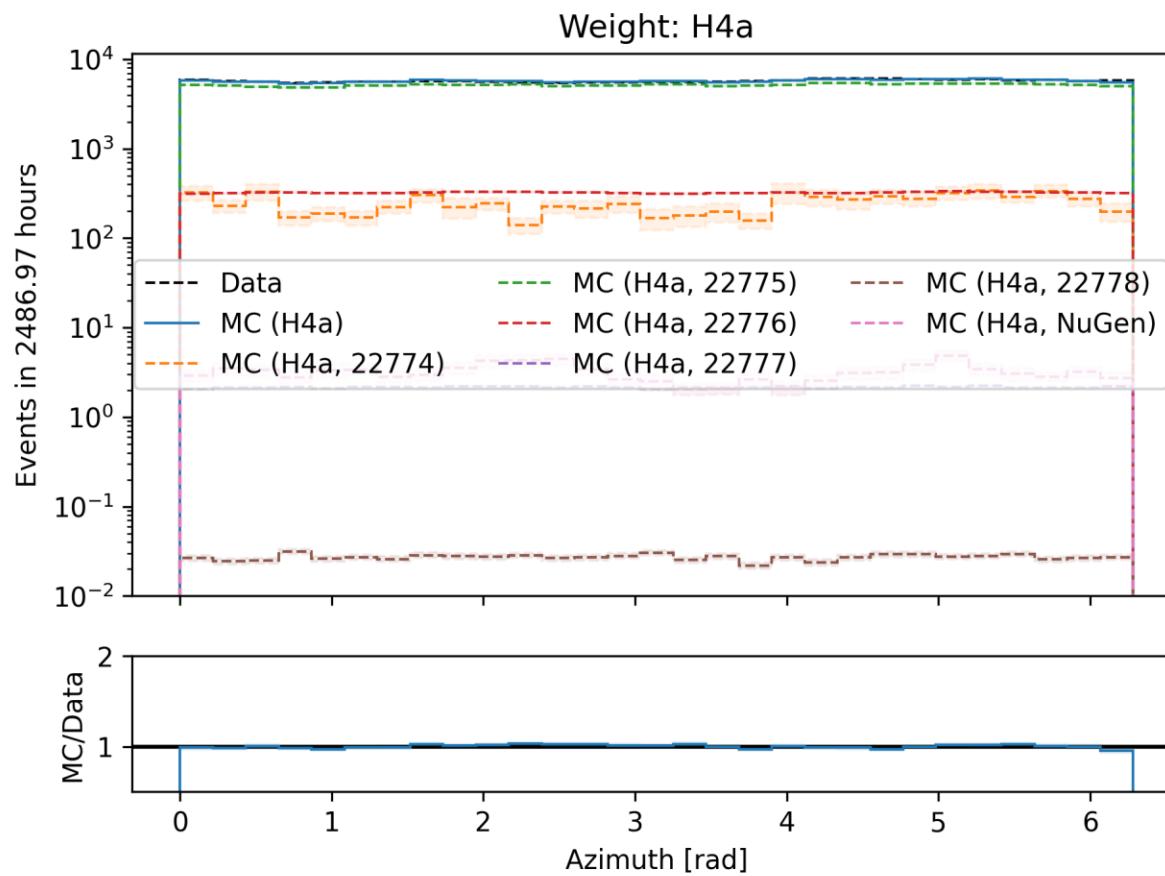
Iron



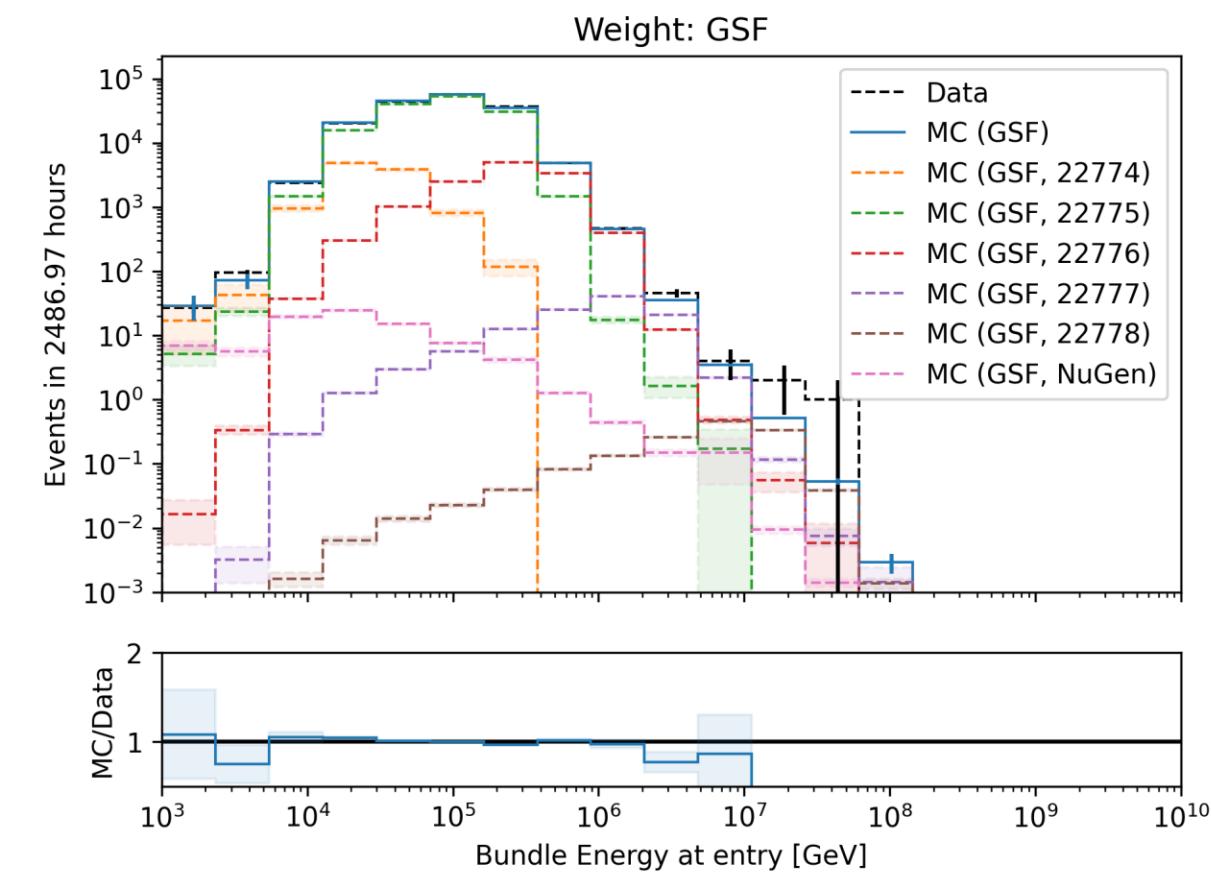
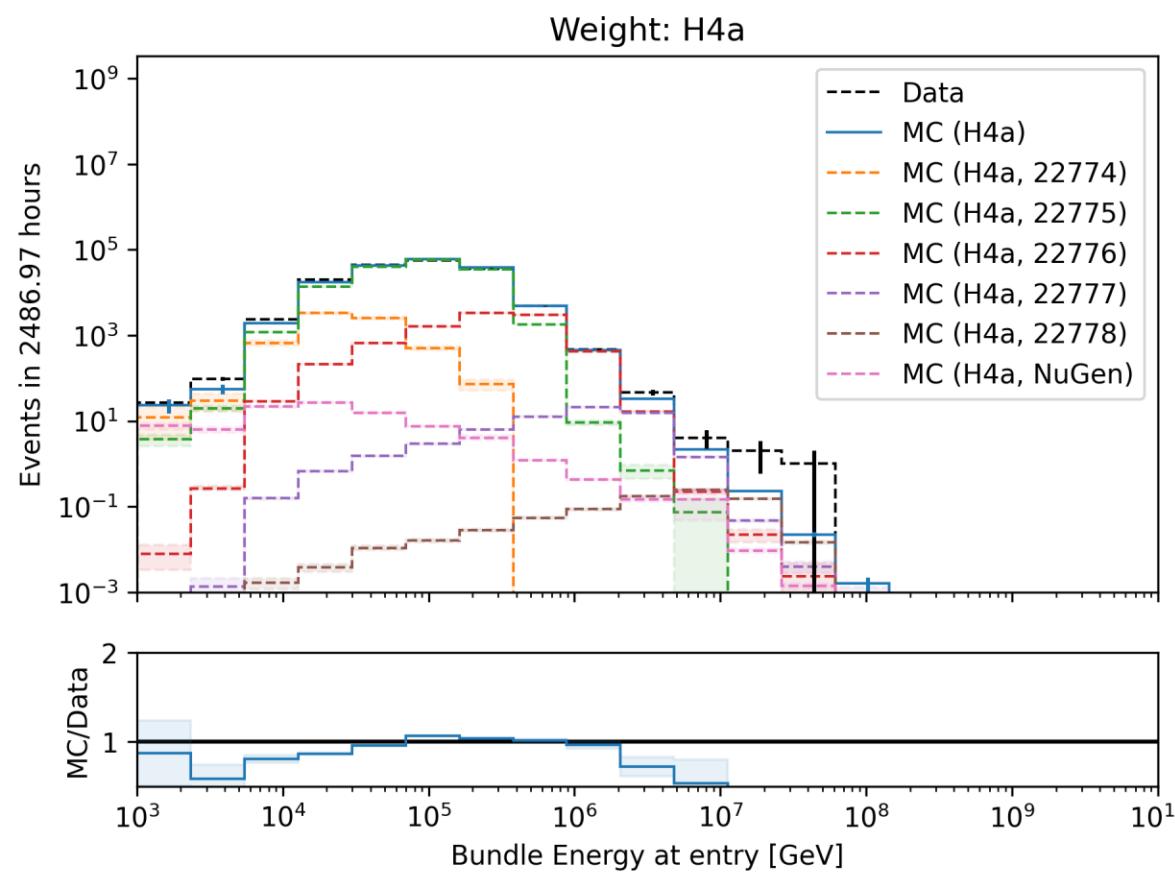
## Agreement for different primary models



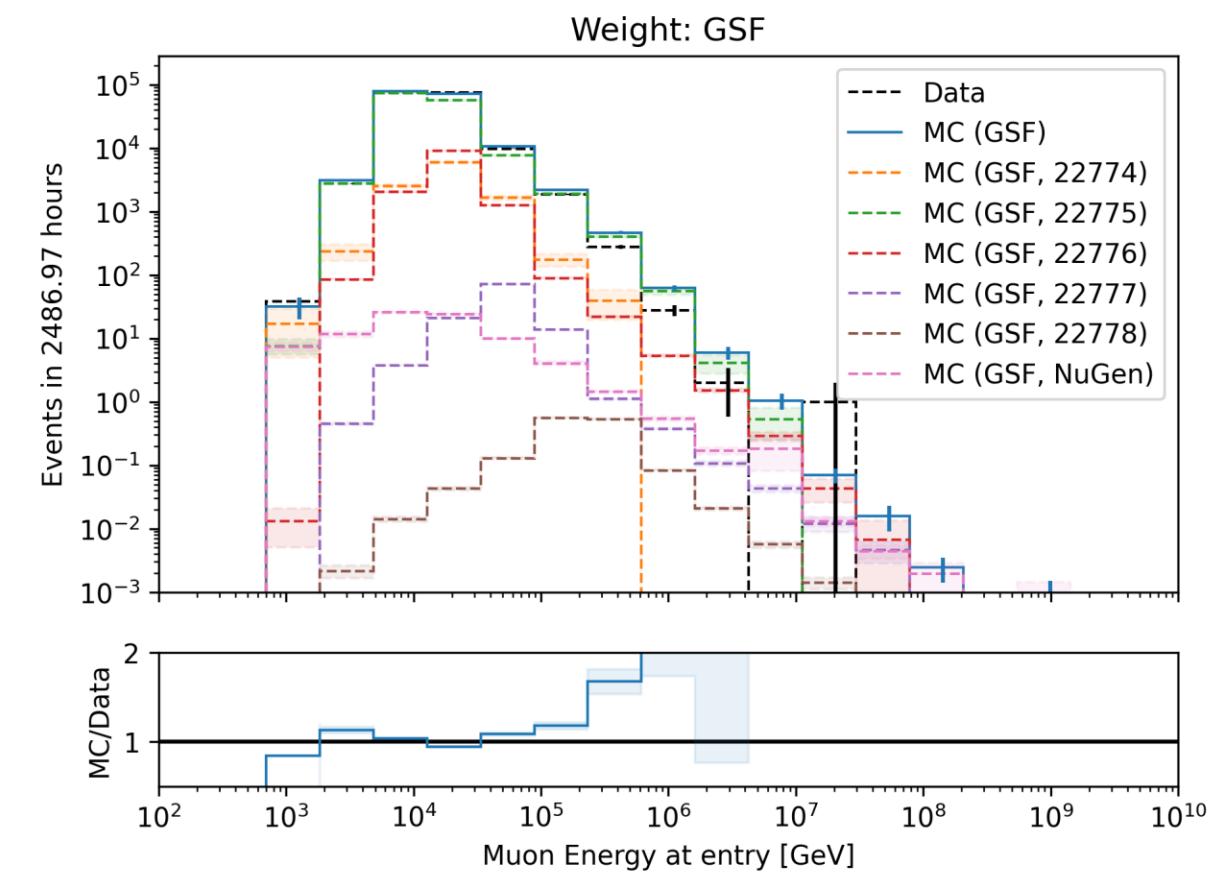
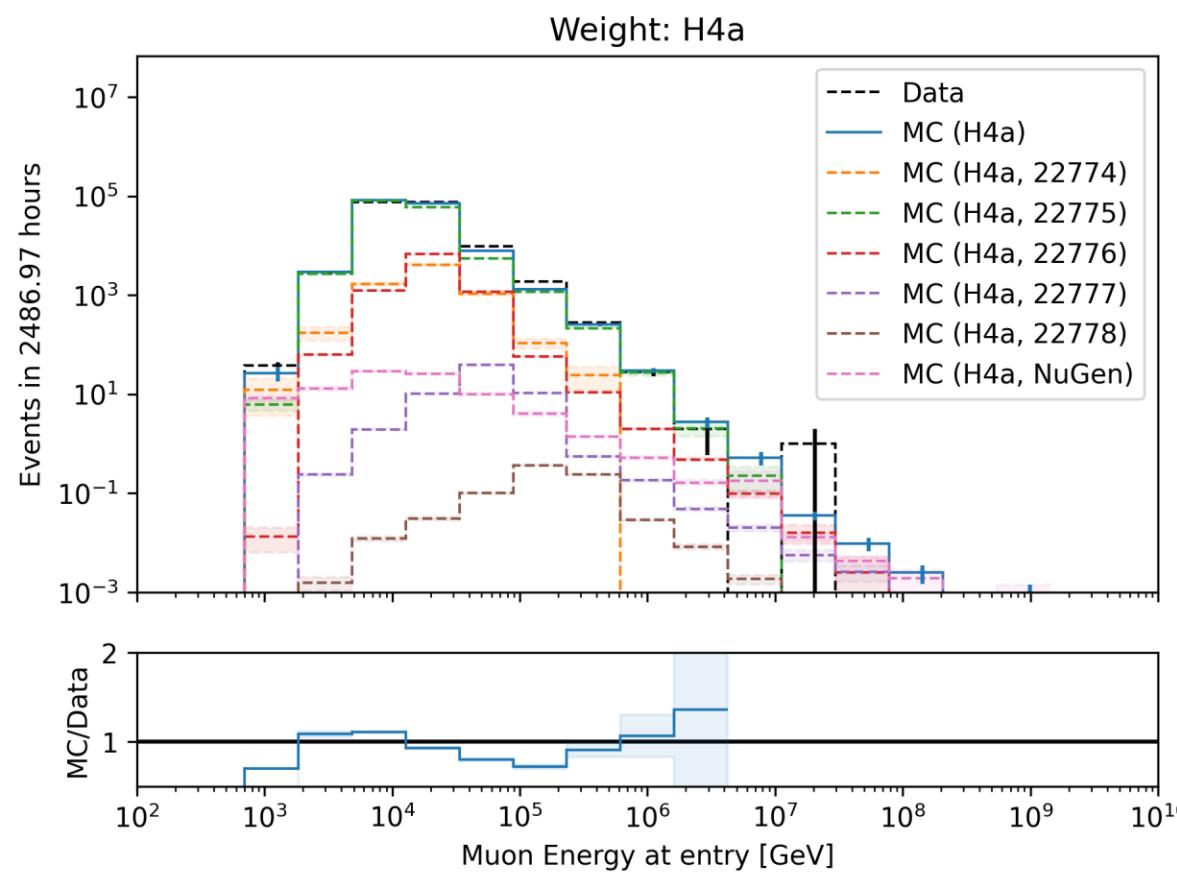
## Azimuth



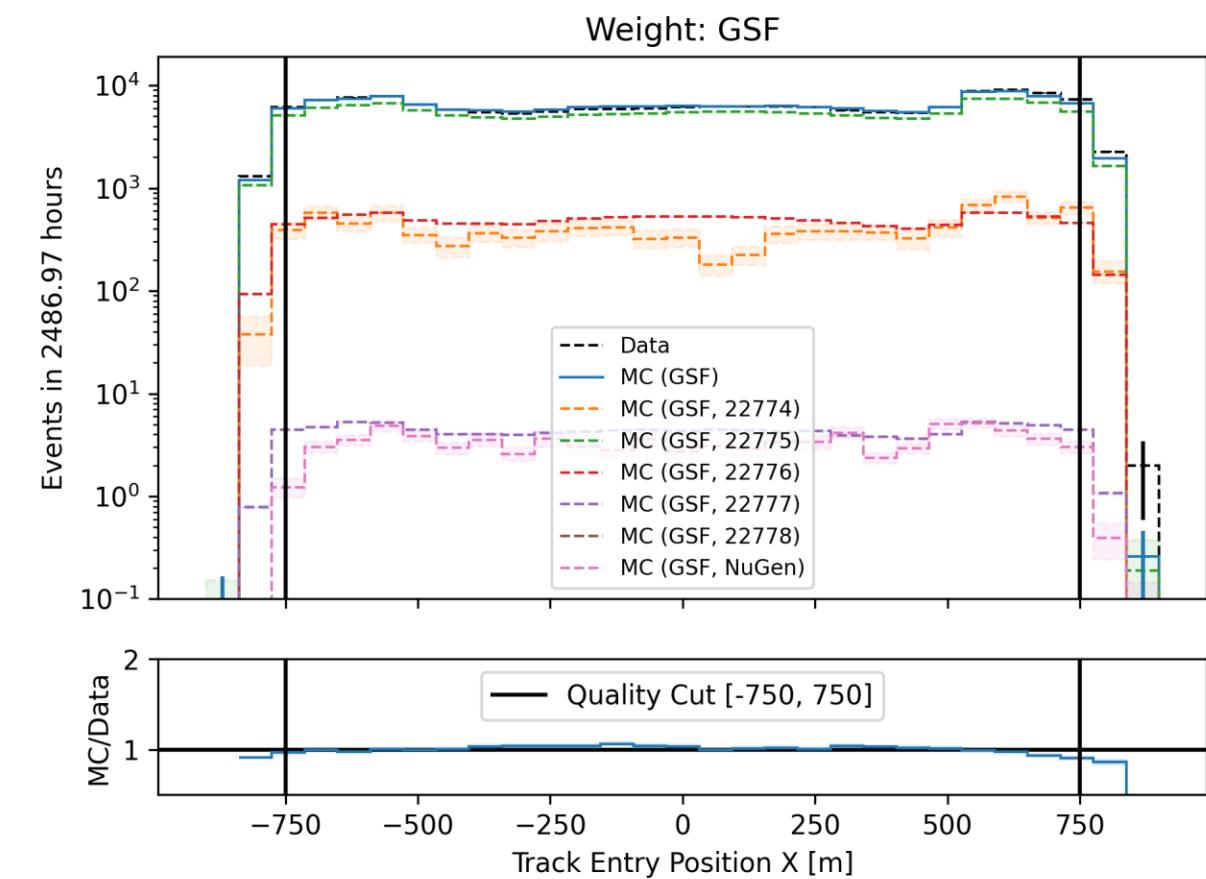
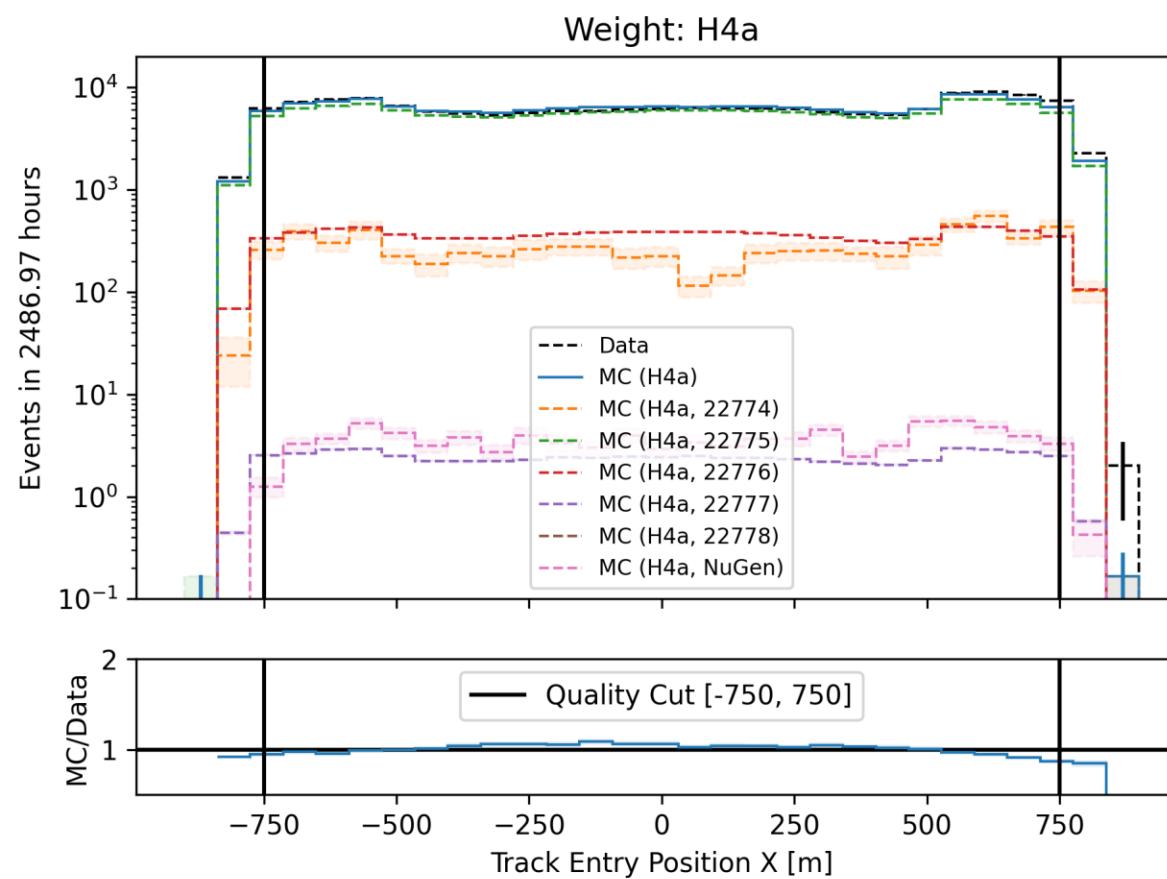
# Bundle energy at entry



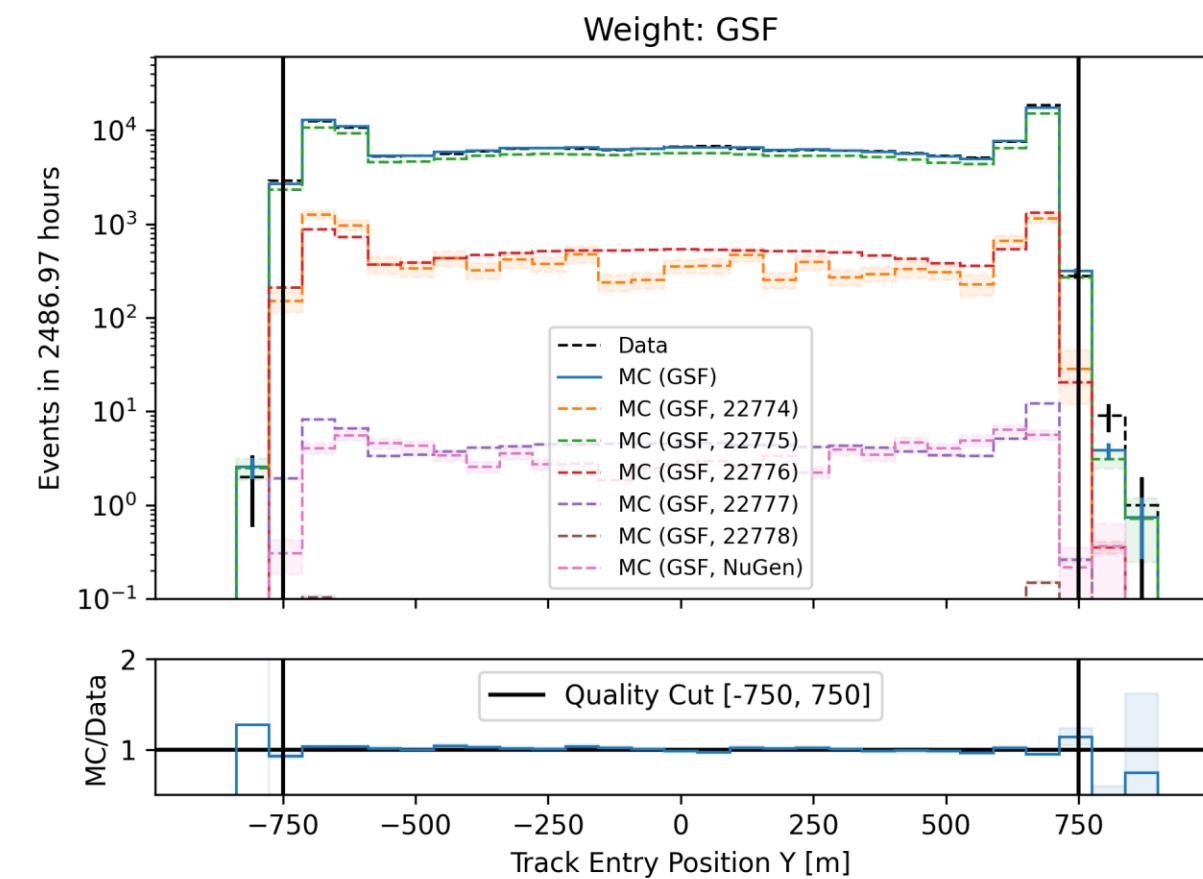
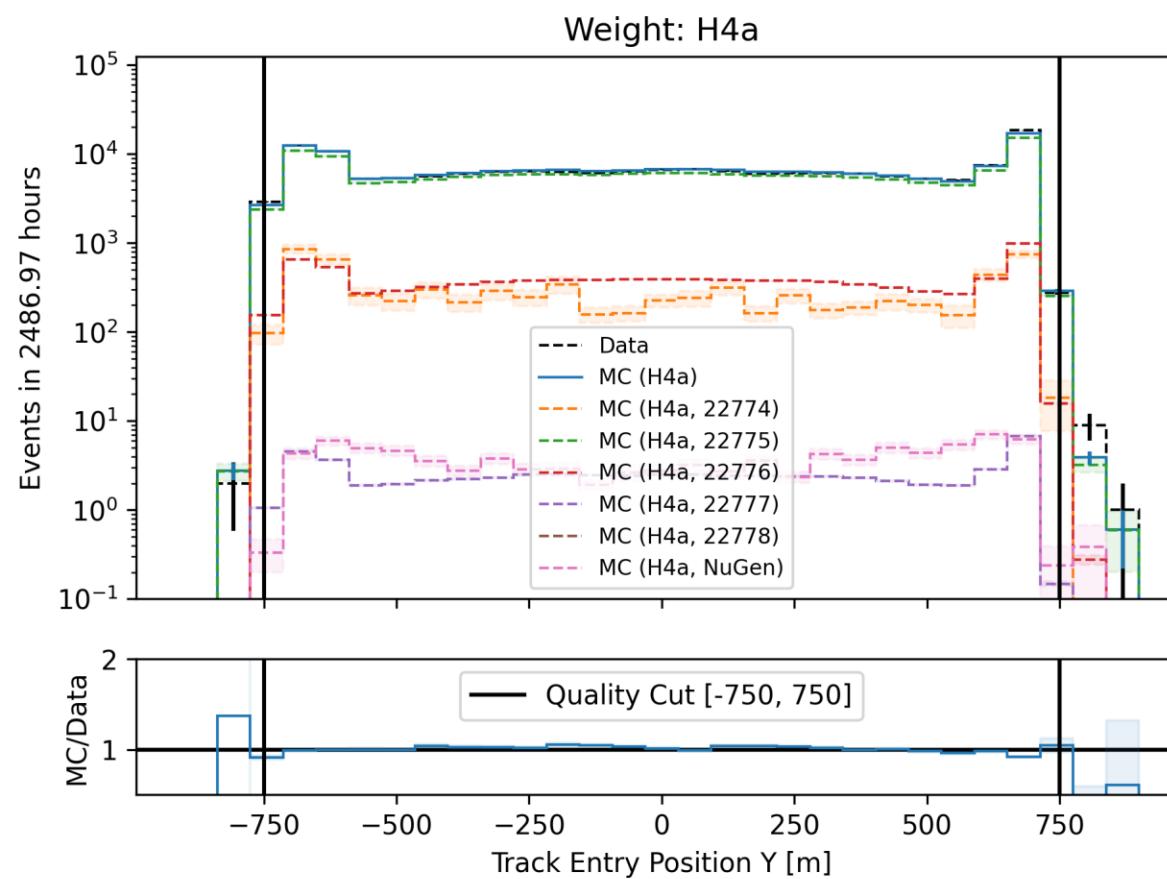
# Leading muon energy at entry



# Entry position x–vertex



# Entry position y–vertex



# DNN reconstructions

# Reconstructed properties

## Energy

- `entry_energy`: Leading muon energy at the detector entry
- `bundle_energy_at_entry`: Muon bundle energy at the detector entry
- `muon_energy_first_mctree`: Leading muon energy at surface
- `bundle_energy_in_mctree`: Muon bundle energy at surface



## Track geometry

- `Length`: Propagation length of muon in the ice
- `LengthInDetector`: Propagation length of muon in the detector
- `center_pos_x`: Closest x position of muon to center of the detector
- `center_pos_y`: Closest y position of muon to center of the detector
- `center_pos_z`: Closest z position of muon to center of the detector
- `center_pos_t`: Time of closest approach to the center of the detector
- `entry_pos_x`: x position of muon at the detector entry
- `entry_pos_y`: y position of muon at the detector entry
- `entry_pos_z`: z position of muon at the detector entry
- `entry_pos_t`: Time of muon at the detector entry

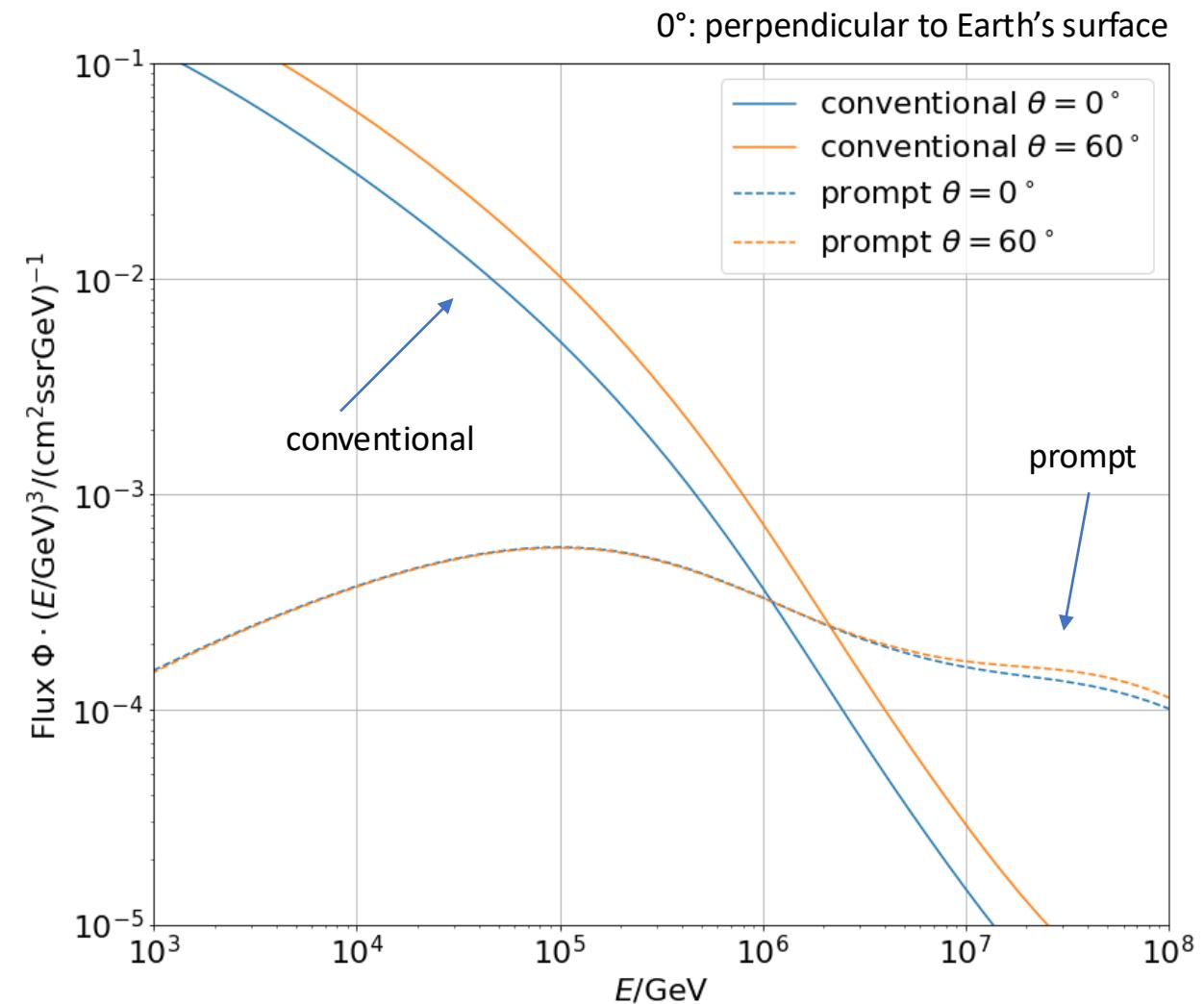
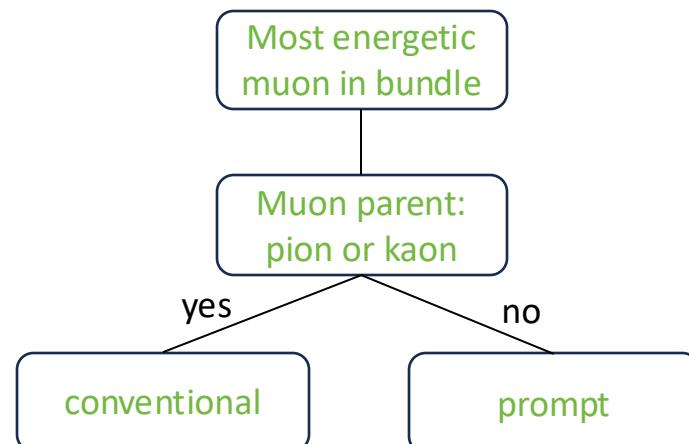
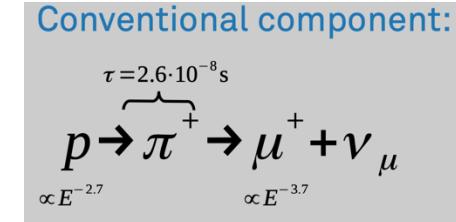
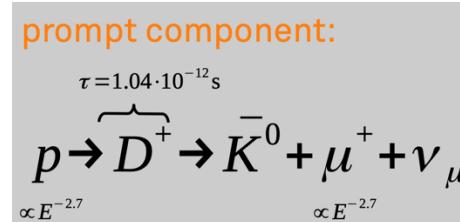


## Direction

- `zenith`: Zenith angle of muon
- `azimuth`: Azimuth angle of muon

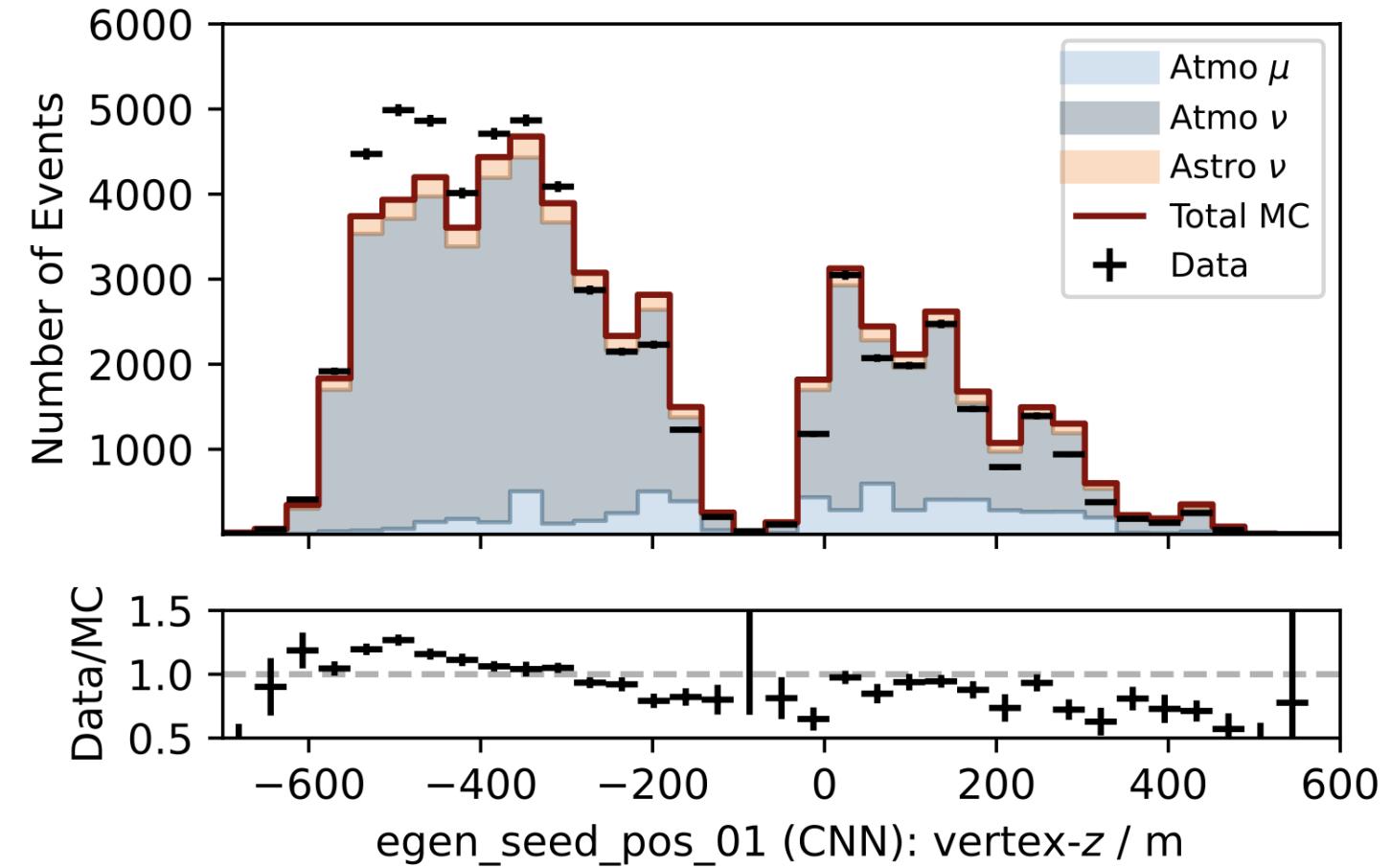
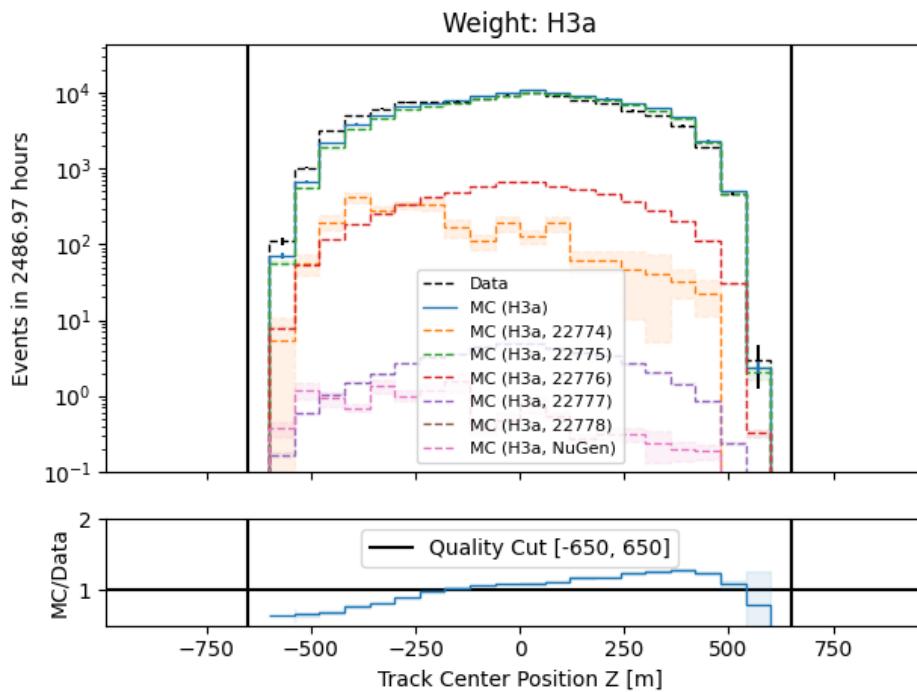
# Muon flux: Prompt spectrum

$$\Phi_{\text{tot}} = \Phi_{\text{prompt}} + \Phi_{\text{conventional}}$$

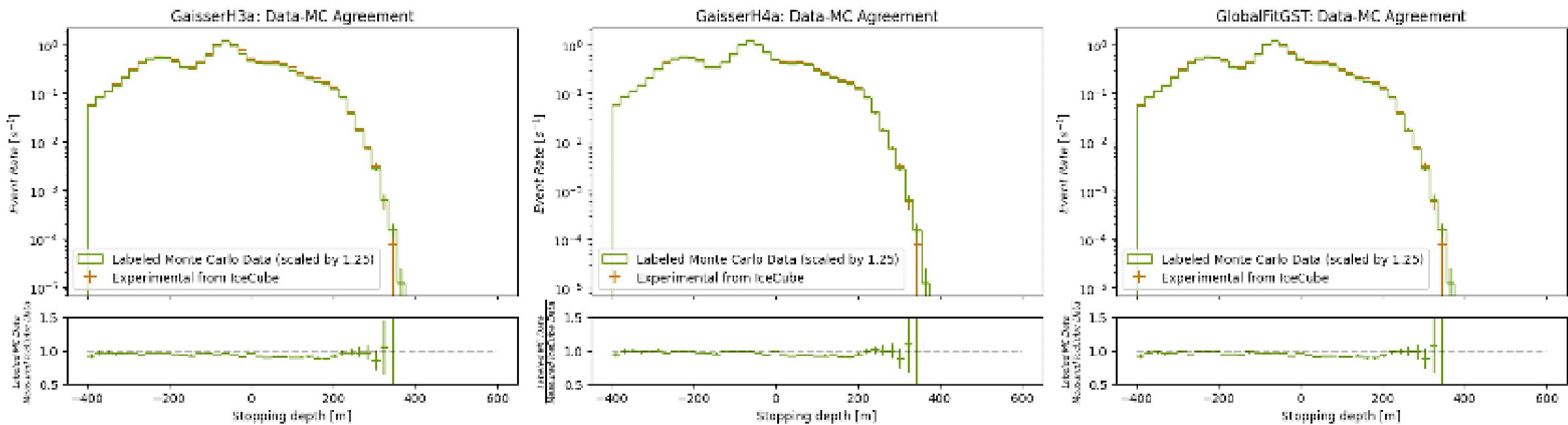


# Neutrino: z-vertex

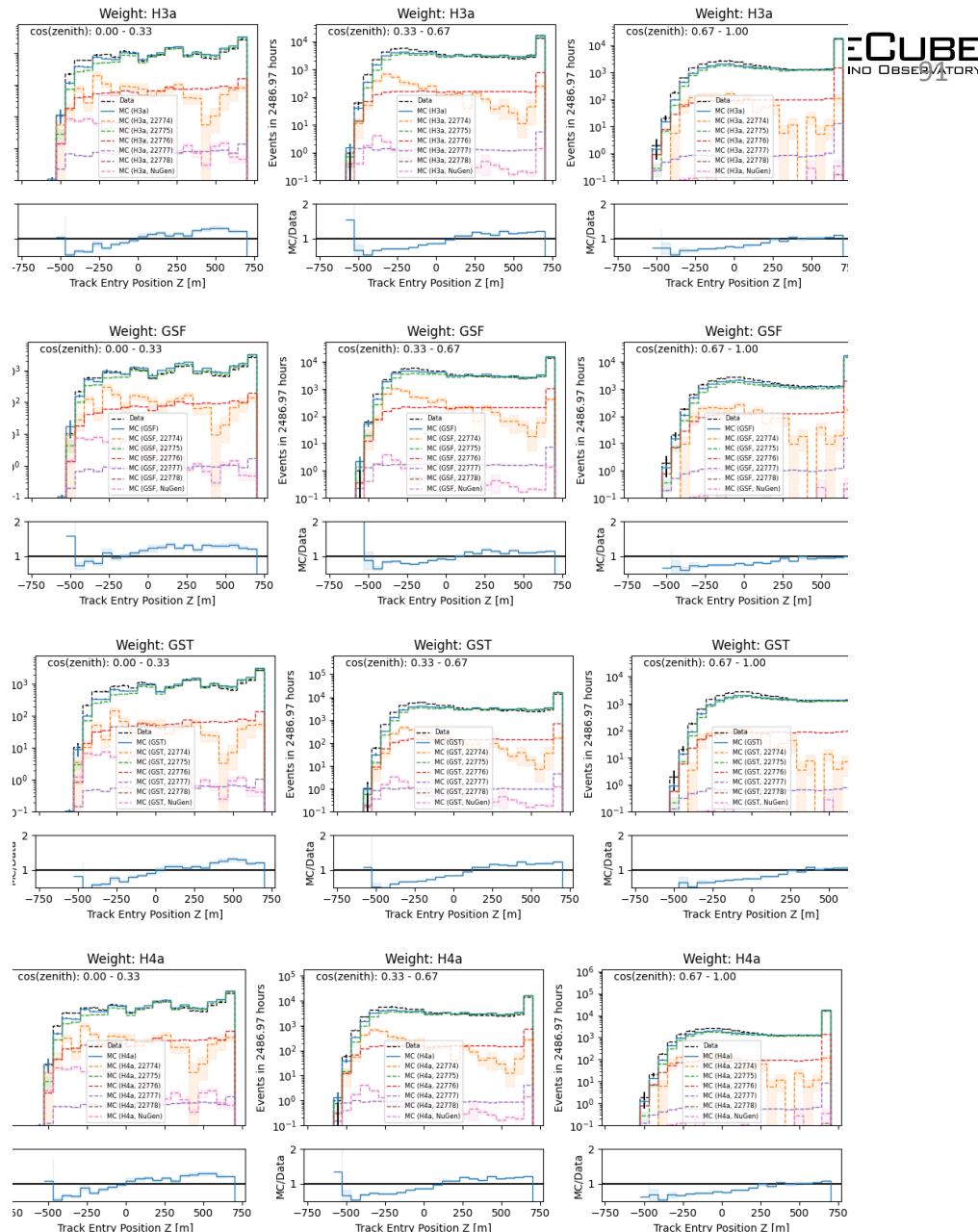
- Dissertation Mirco
- Same shape mismatch for neutrinos as for muons



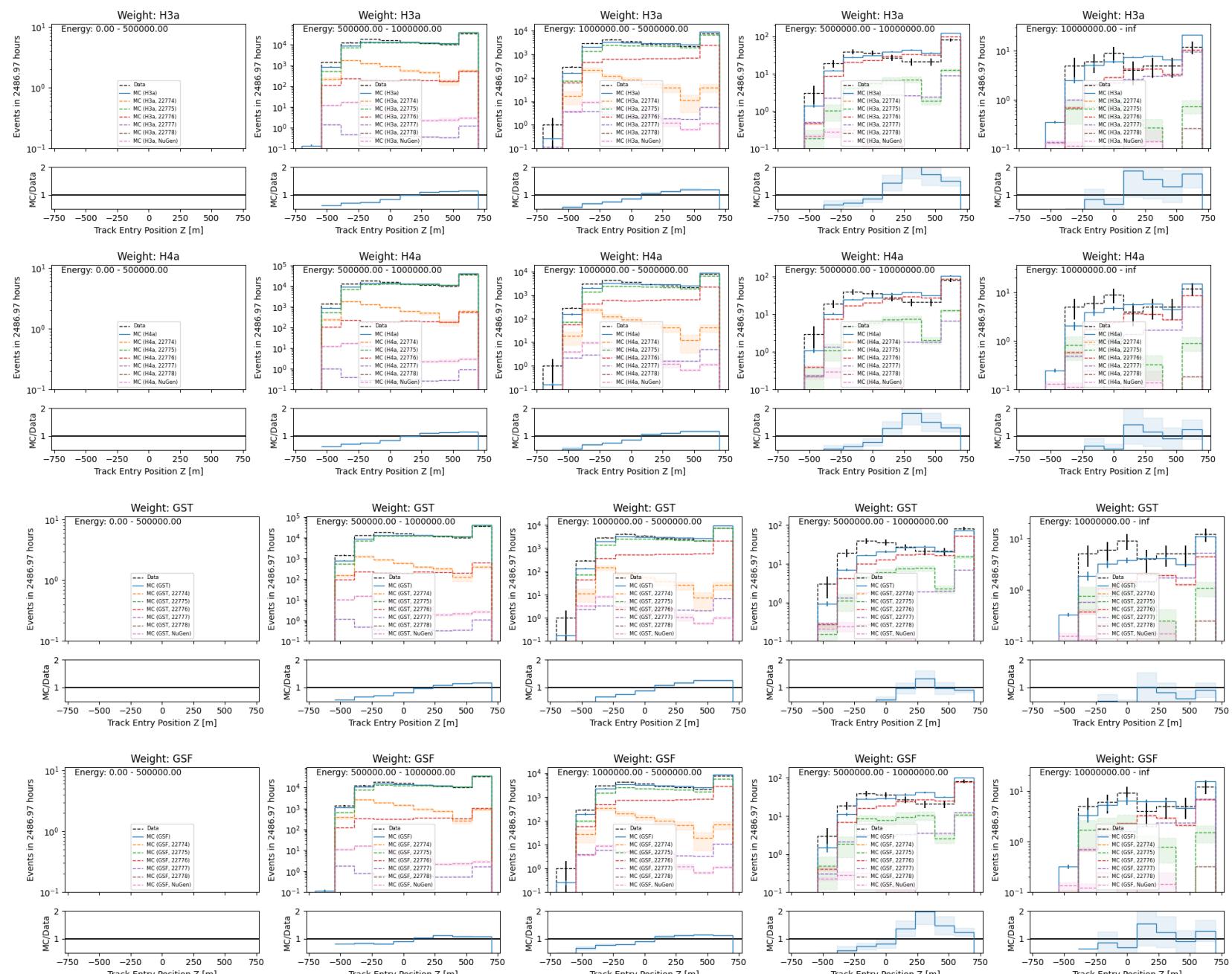
# Stopping muons



# Entry position z - different cos(zenith)



Entry position z  
- different cuts on  
bundle energy at  
surface

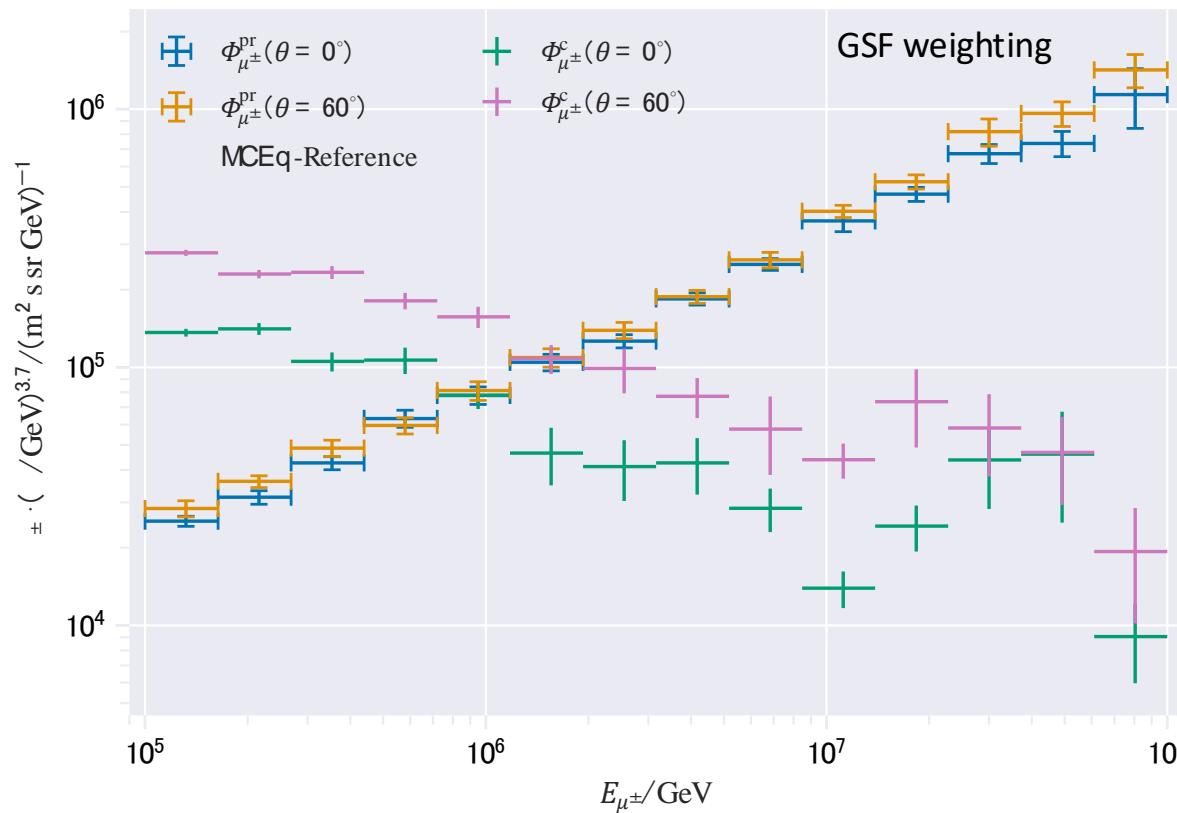


# New CORSIKA simulation

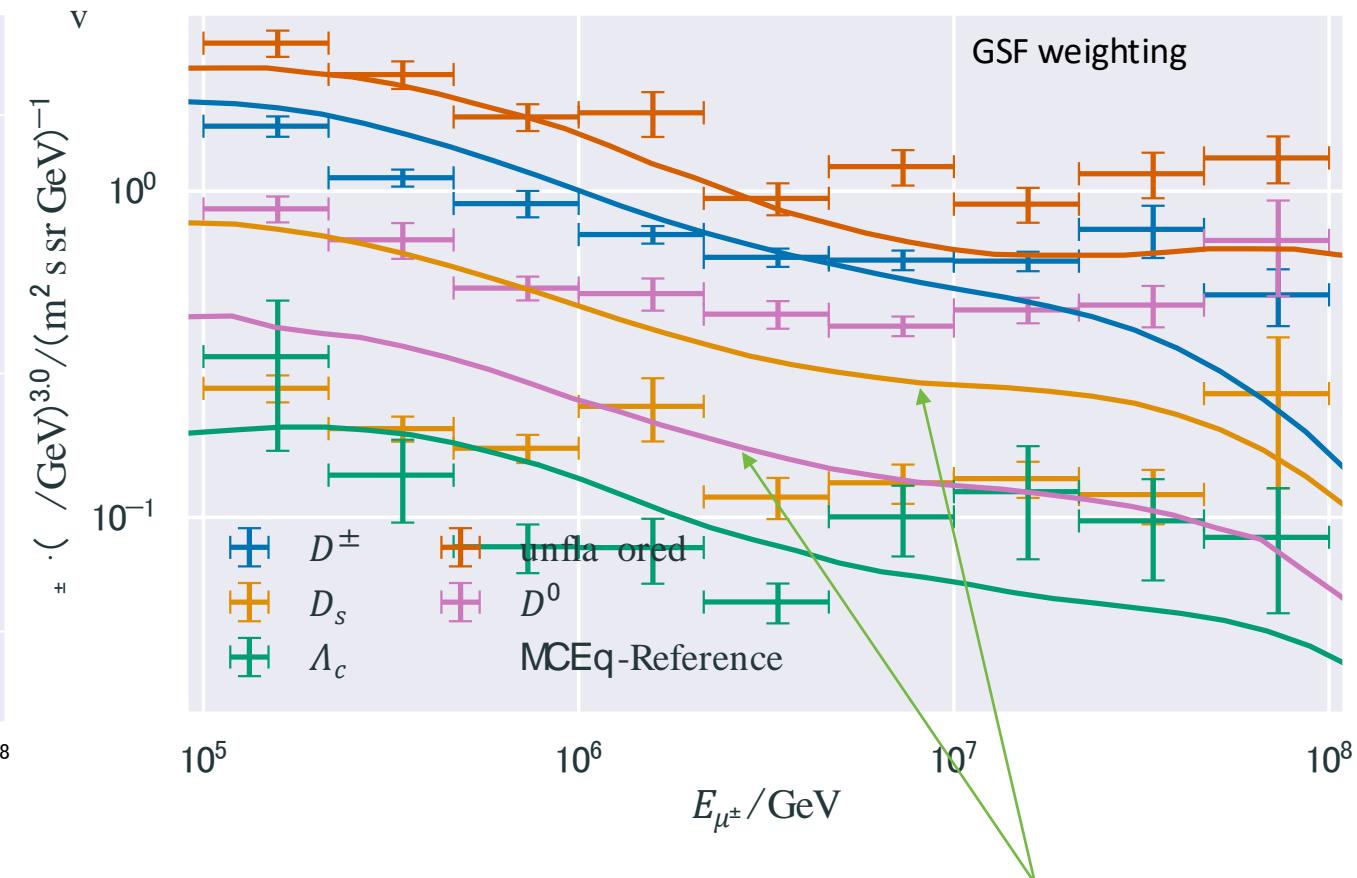
with extended history option for information about the parent particles

# CORSIKA 7 tagging and MCEq comparison

MCEq: tool to numerically solve the cascade equations that describes the evolution of particle densities as they propagate through a gaseous, dense medium  
<https://github.com/mceq-project/MCEq>

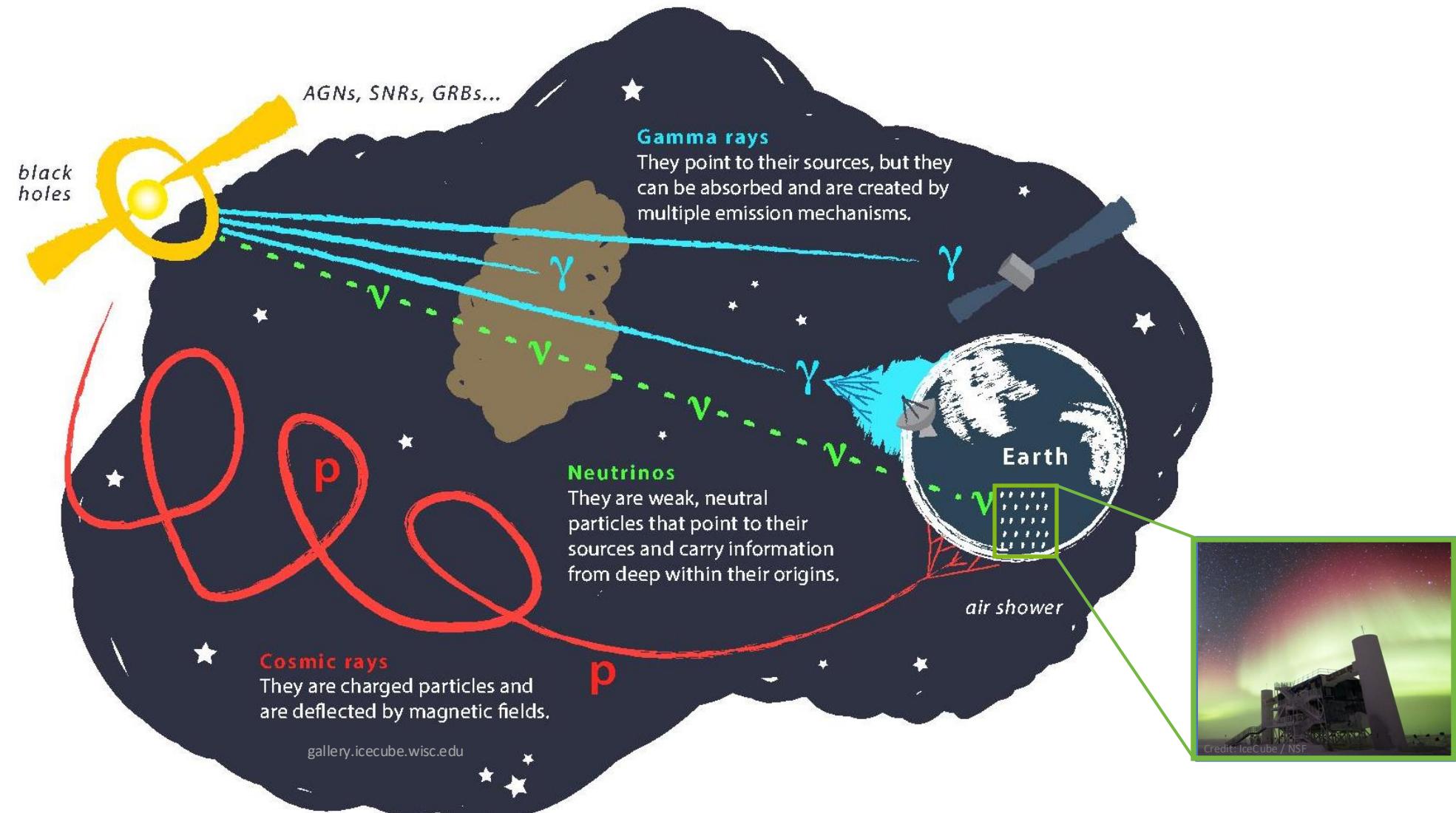


➤ Good agreement in total prompt and conv muon flux



➤  $D^0$  and  $D_s$  are swapped here but this is fixed in MCEq

# Astroparticle physics



## Cosmic ray flux

