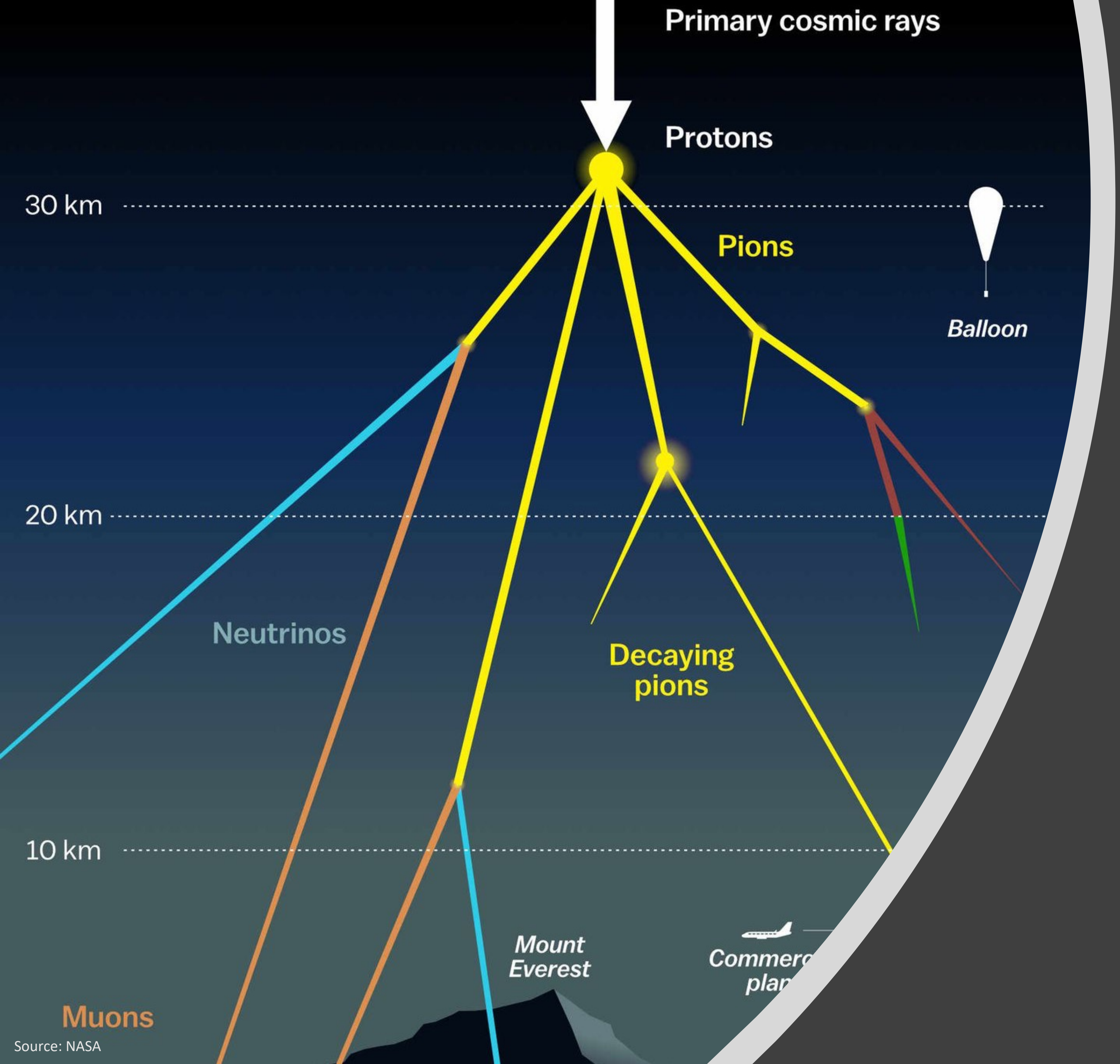


Source: NASA

# Prompt Muons Update - Coincident Events & Systematics


**Pascal Gutjahr**


August 15, 2025

WG reviewer: Dennis Soldin
Coll. reviewer: Anatoli Fedynitch
Technical reviewer: Karolin Hymon
Wiki: prompt wiki
Last update: Leadingness Cut

# Coincident Primaries on Final Level

Use dataset 22615

description: 5-component CORSIKA Snowstorm using icetray.v1.9.2 (up to Level2): spice_ftp-v1 Ice E^-2 spectrum {600GeV,1e8GeV}, SIBYLL 2.3d
jobs_submitted: 100000

```
/data/sim/IceCube/2023/filtered/level2/CORSIKA-in-ice/22615/
```

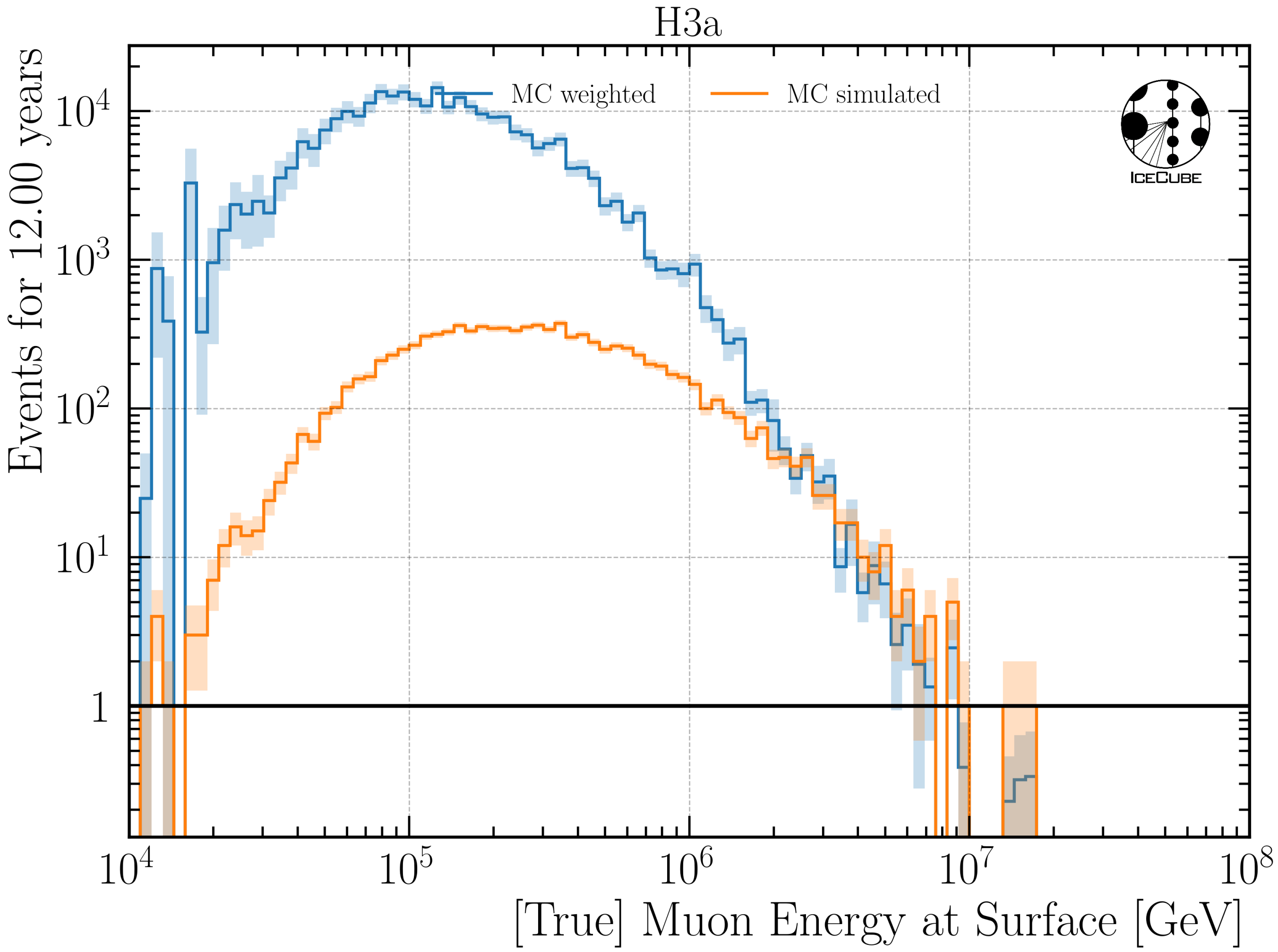
# I3MCTree

Signal

Background

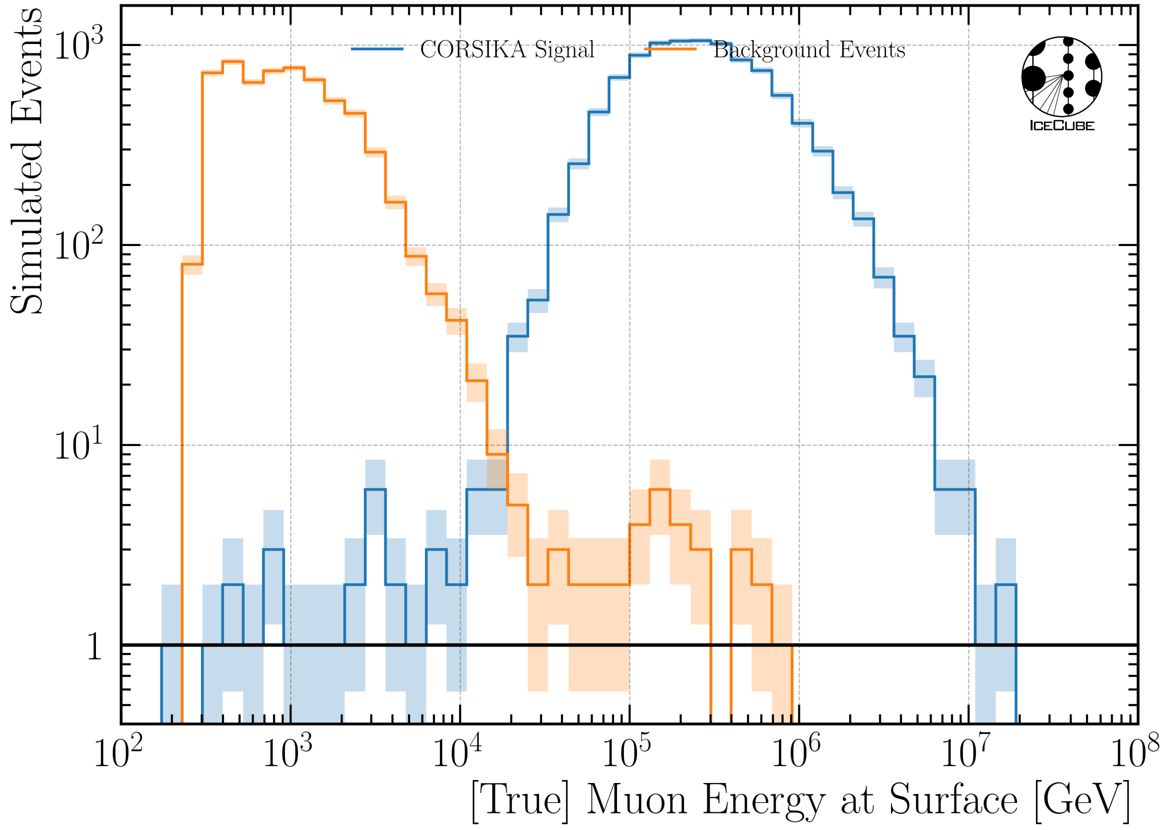
```
I3MCTree_preMuonProp [TreeBase::Tree<I3Particle, I3ParticleID, i3hash<I3ParticleID> >]:
[I3MCTree:
  2150 PPlus (-10016.7m, -15075m, 111889m) (9.26856deg, 236.318deg) -351822ns 15045.2GeV 95347.9m
    2151 NuMu (-65.0611m, -144.268m, 1950.08m) (9.27059deg, 236.355deg) 19752.8ns 6860.23GeV nanm
    2152 MuPlus (-64.7531m, -144.808m, 1950.08m) (9.2668deg, 236.307deg) 4734.62ns 1645.51GeV nanm
    2153 MuPlus (-66.0127m, -113.568m, 1950.08m) (9.52149deg, 237.478deg) 49671.8ns 195.36GeV nanm
    2154 MuMinus (-69.6836m, -164.987m, 1950.08m) (9.08885deg, 236.417deg) 17724.1ns 121.291GeV nanm
  965 PPlus (-173153m, -348327m, 100569m) (75.5732deg, 243.564deg) -1.32459e+06ns 3016.31GeV 356965m
    966 NuMuBar (-2078.95m, -5096.2m, 1950.08m) (75.5954deg, 243.482deg) -2912.83ns 506.092GeV nanm
  1116 PPlus (-2369.58m, -59870.8m, 111641m) (28.4463deg, 267.819deg) -402068ns 5277.49GeV 89989.7m
    1117 MuPlus (-63.9208m, -478.739m, 1950.08m) (28.4565deg, 267.656deg) 14091.1ns 765.373GeV nanm
]
```

# Check Signal Statistics for Muon Energy at Surface

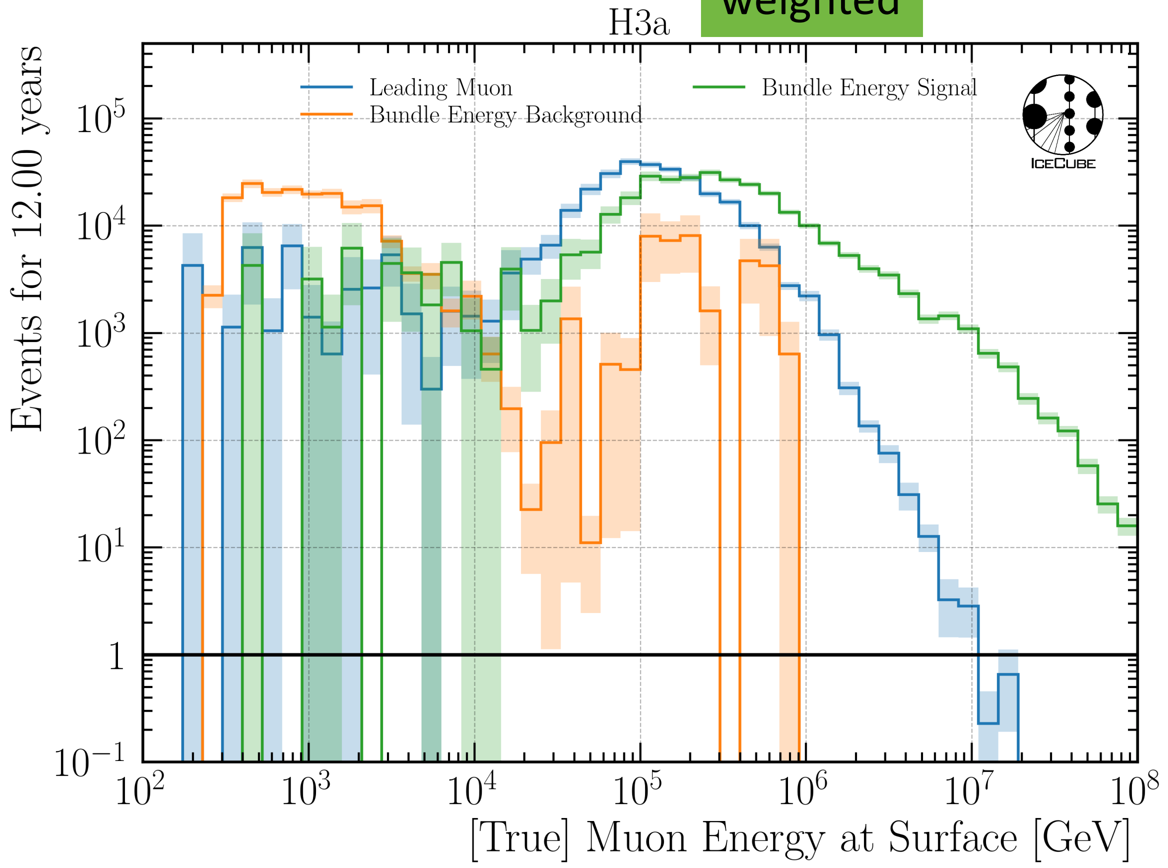


# Muon Energy at Surface

simulated

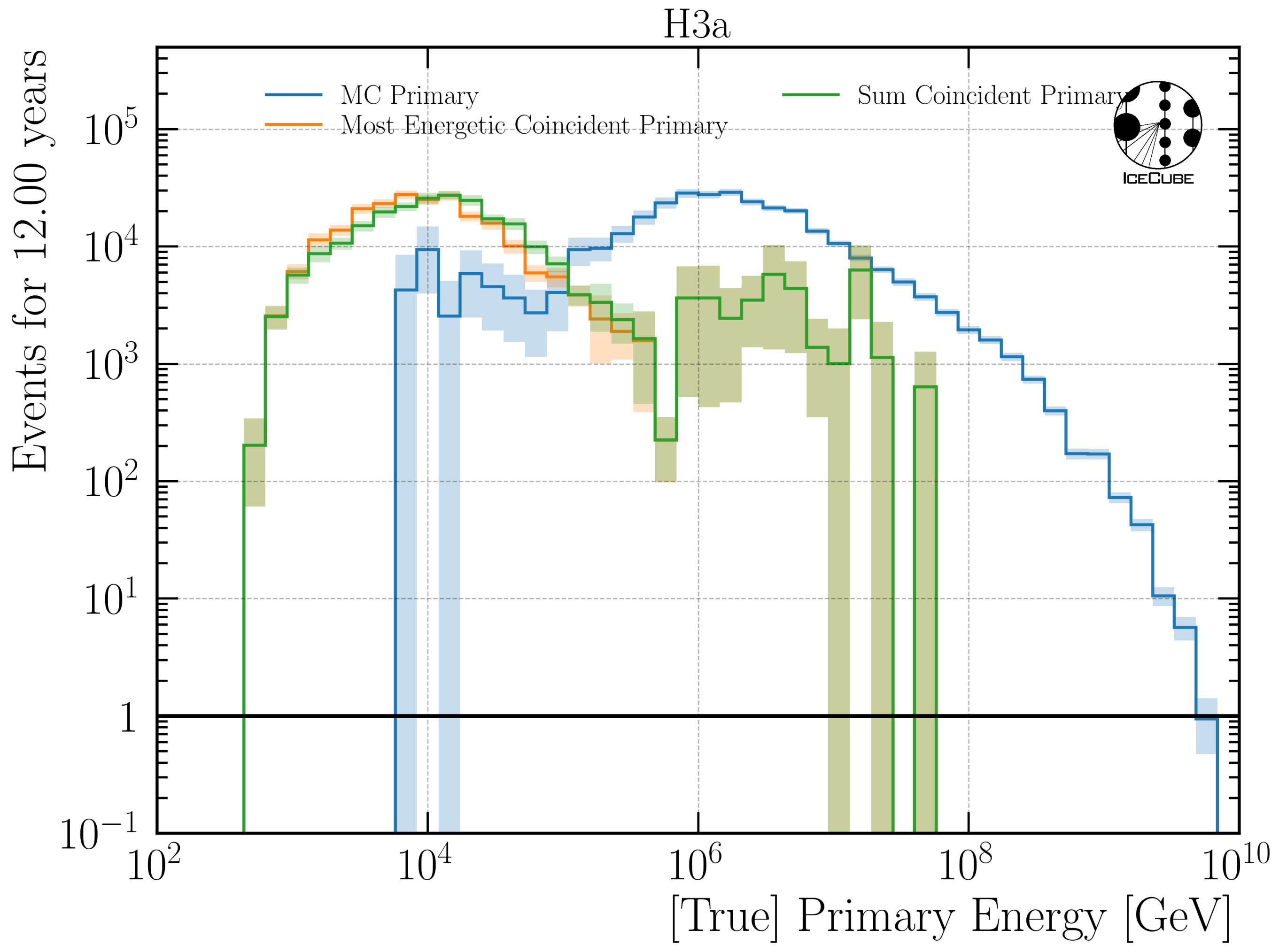



weighted



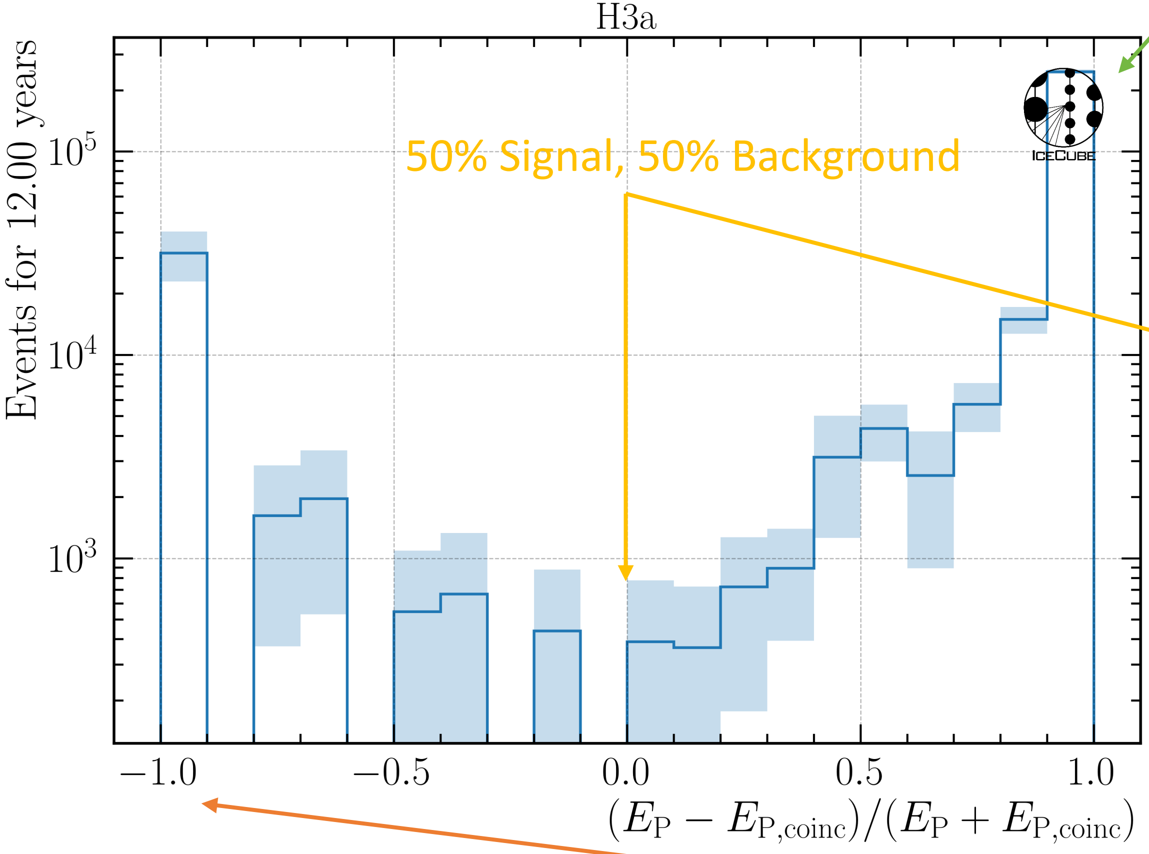

- Leading Muon: most energetic muon in tree
- Bundle Energy Signal: sum of all muons of ”signal” primary
- Bundle Energy Background: sum of all muons of coincident primaries

# Primary Particles

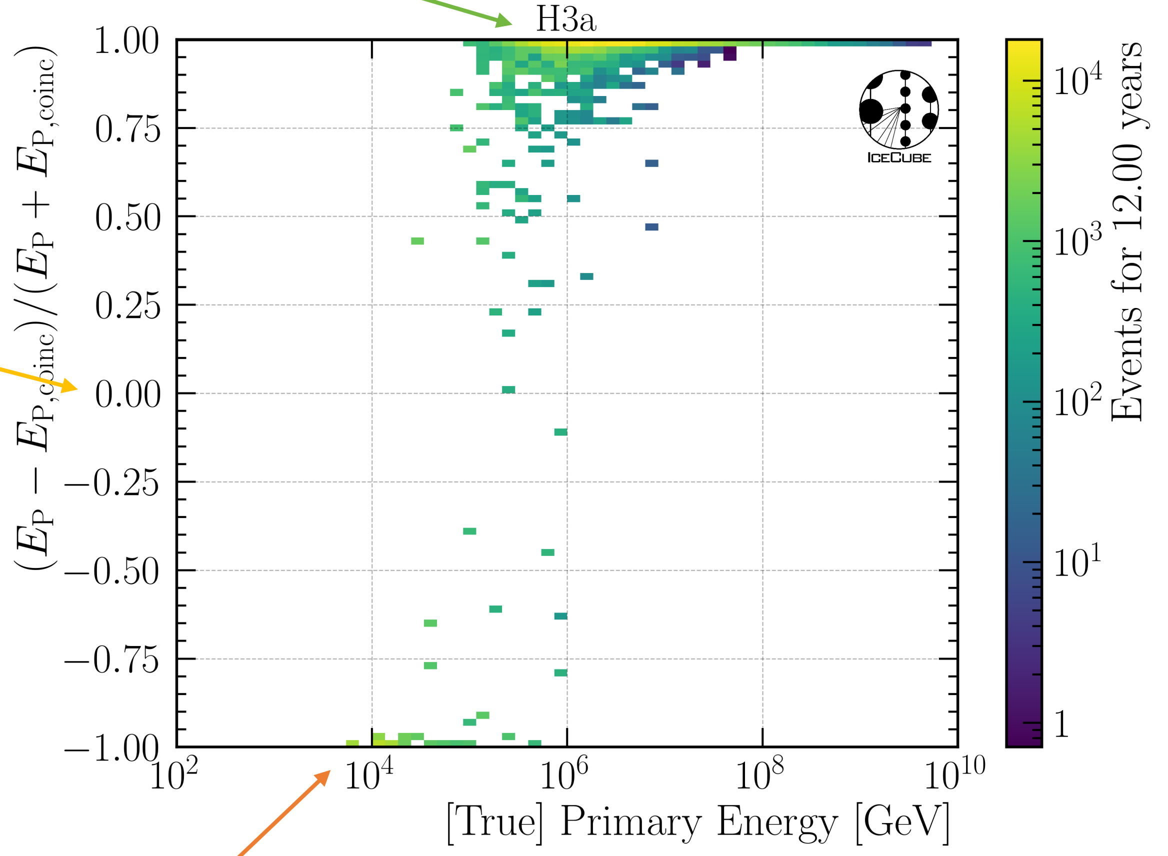


# Primary Particles on Event Level

100% Signal Primary



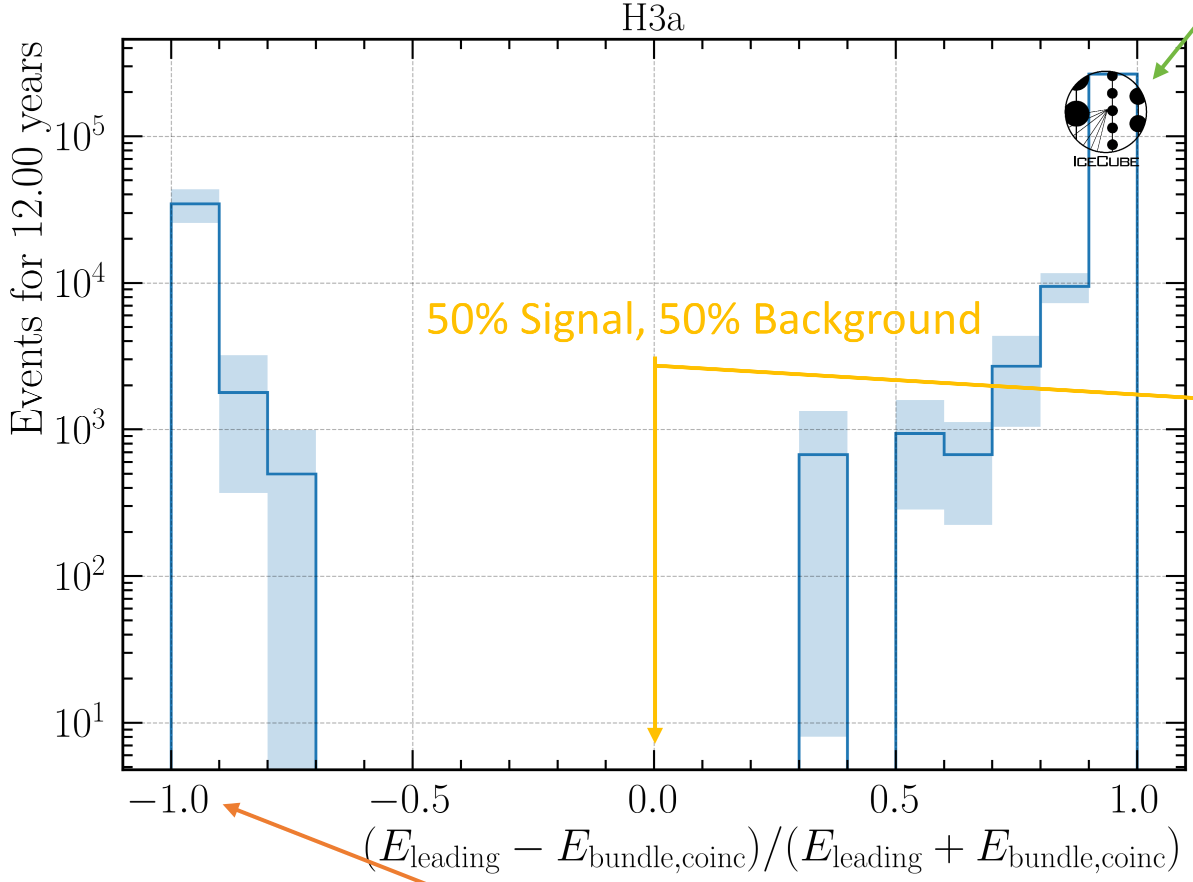

50% Signal, 50% Background



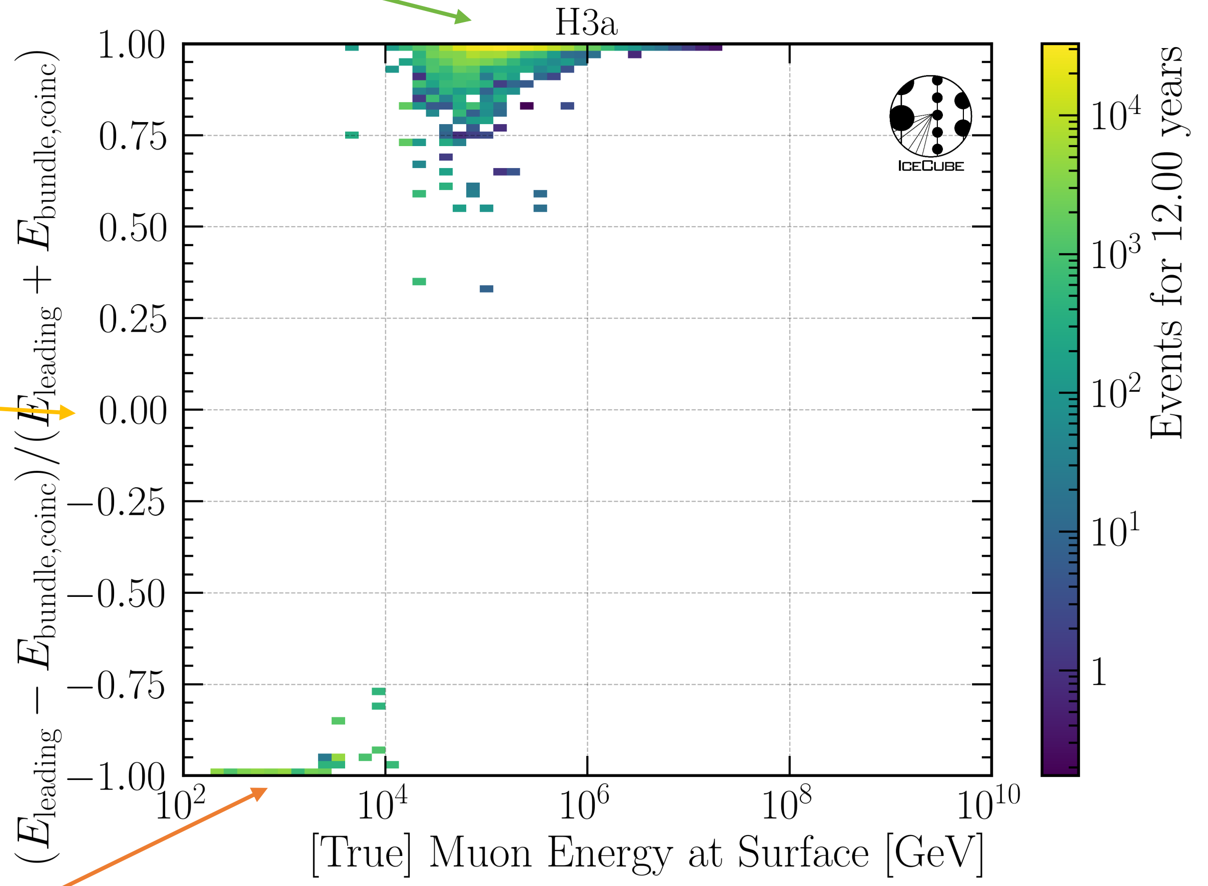

100% Background Primary

100% Signal Primary

# Muons on Event Level




50% Signal, 50% Background




100% Background Primary

- above 10 TeV, no background dominated events

# Estimate Rates with H3a

- Signal rate: 0.84 mHz (total event rate)
- Background rate: 0.68 mHz (event rate with at least 1 coincident primary)

- Signal rate: 0.74 mHz (leading muon energy at surface > 10 TeV)
- Background rate: 0.58 mHz (event rate with at least 1 coincident primary & lead. muon E. > 10 TeV)

- Signal rate: 0.74 mHz (leading muon energy at surface > 10 TeV)
- Background rate: 0.02 mHz (event rate with at least 1 coincident primary & lead. muon E. > 10 TeV & coincident muon bundle energy at surface hast at least 10% of leading muon energy)

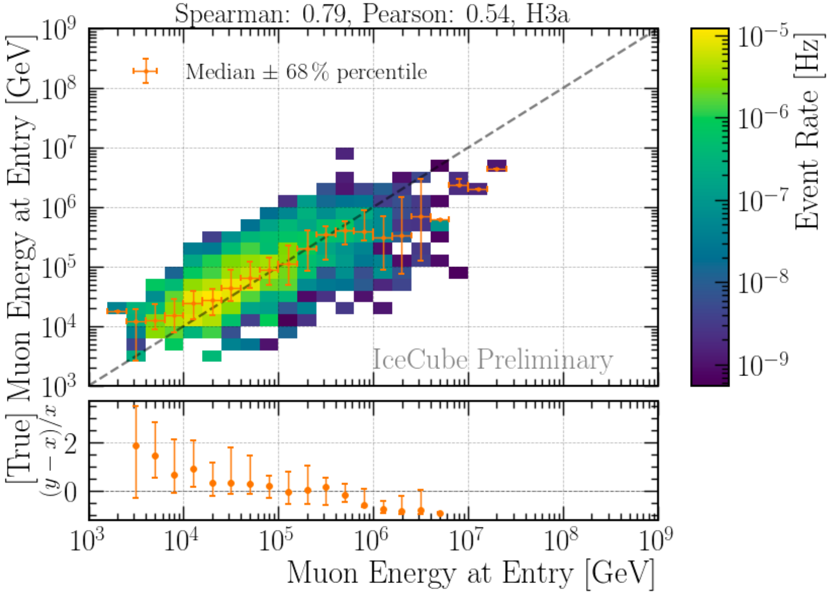
➢ The light of the background events overlaps with the signal → no chance to separate

➢ A 10% bundle energy contribution would shift the measured light up by roughly 10% → within the uncertainties of the energy prediction

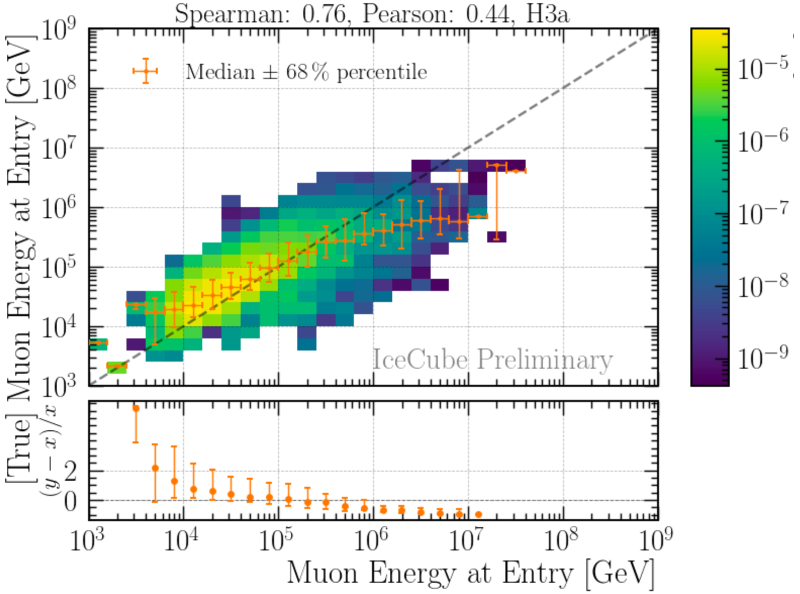
➢ Networks have been trained on MC with coincident events → they are able to subtract a little light in case they assume there is a background event, however, this is not quantified

# Reconstructions

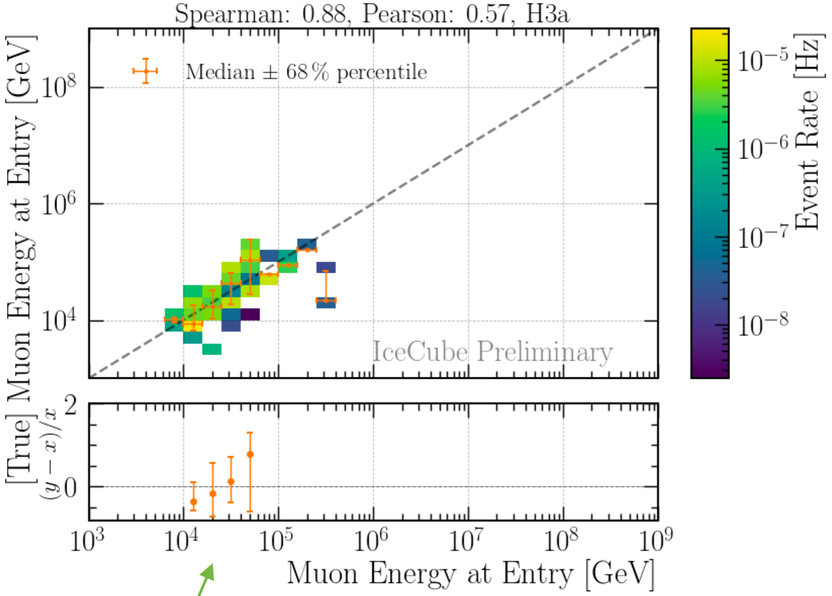
no background event




$$\frac{\text{bundle energy at surface}}{\text{leading energy at surface}} < 0.1$$




$$\frac{\text{bundle energy at surface}}{\text{leading energy at surface}} > 0.1$$



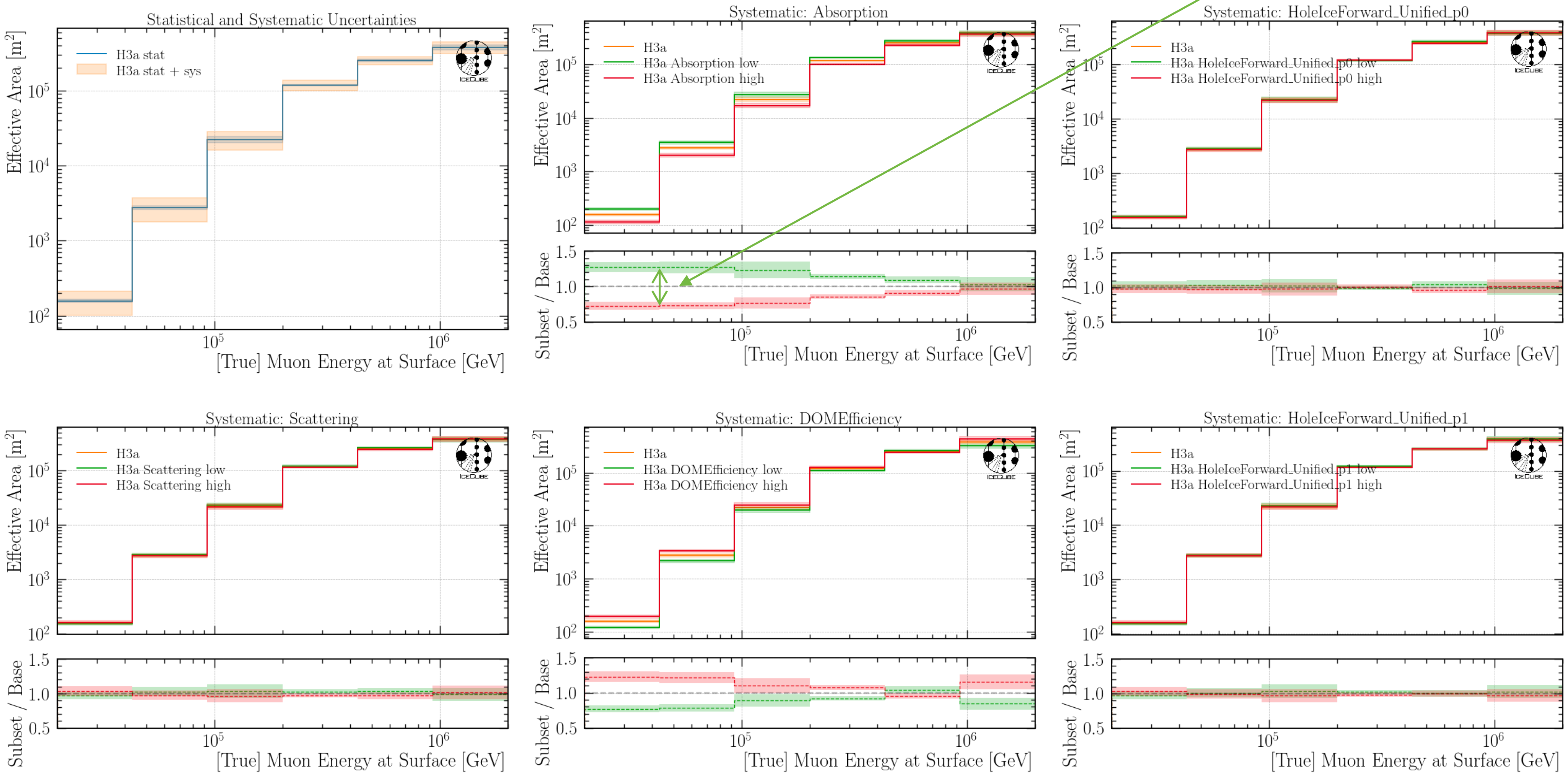

- sufficient energy reconstruction for events with background

# Systematic Uncertainties

# Effective Area → Systematics

- Baseline: entire set
- Subset: above/below center

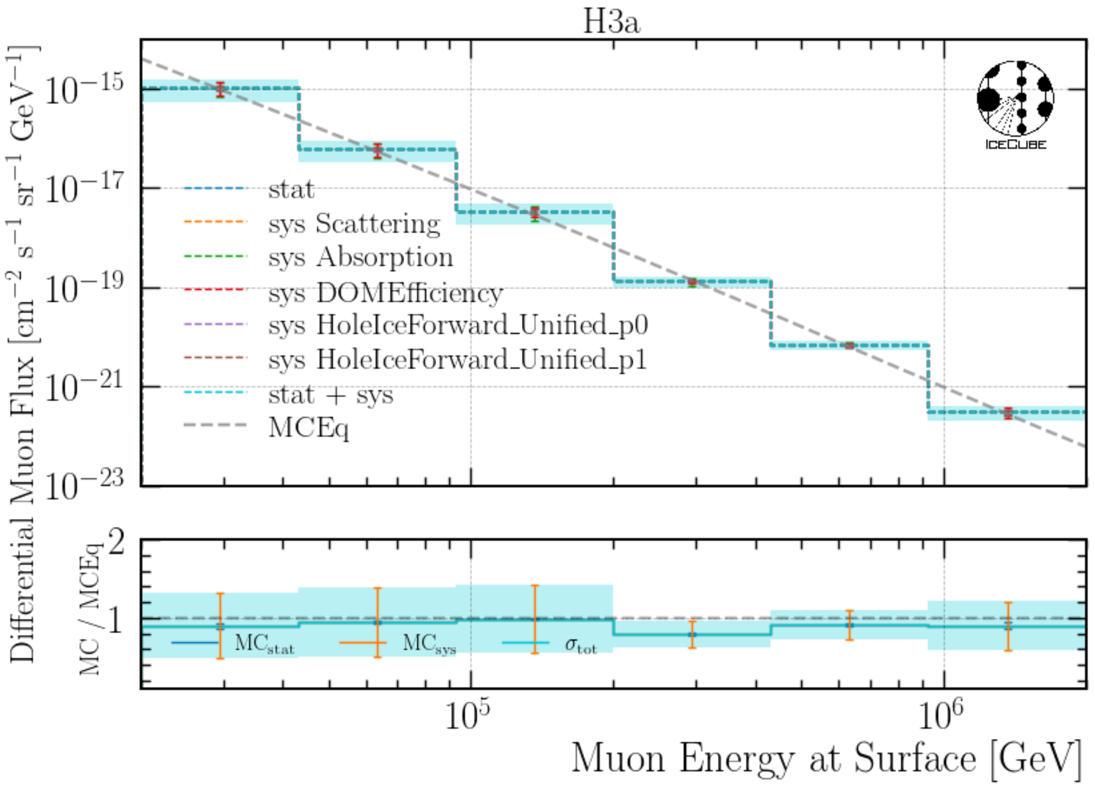
$$\sigma_{sys} = \max(A_{low} - \sigma, A_{high} + \sigma) - A_{base}$$



# Effective Area with Systematics

Systematics: above/below center
Statistical: sum(weights**2)

$$\sigma_{\mathrm{tot}} = \sqrt{\sigma_{\mathrm{stat}}^2 + \sum \sigma_i^2}\,, i\text{: scat, abs, DOME, holeice p0, p1}$$




- “blind” uncertainty estimation by up- and down-scaling of the simulated systematics
- provides a conservative uncertainty estimation

➢ systematics are included in the fit → further knowledge available → shrinks the uncertainty dortmund university

BERKELEY LAB

ICECUBE NEUTRINO OBSERVATORY

# Unfolding: LLH minimization

| | x0 | x1 | x2 | x3 | x4 | x5 | x6 | x7 | x8 | x9 | x10 | x11 | x12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| x0 | 2.45e+04 | -0.015e6 **(-0.578)** | -0.011e6 **(-0.474)** | -0.001e6 **(-0.059)** | 1e3 **(0.100)** | 1.1e3 **(0.180)** | 0.33e3 **(0.146)** | 0.037 **(0.002)** | -3.7283 **(-0.454)** | -982.08e-3 **(-0.579)** | 452.13e-3 **(0.240)** | 0.00 | -0.0000 |
| x1 | -0.015e6 **(-0.578)** | 2.6e+04 | 0 **(0.001)** | -0.010e6 **(-0.440)** | -1e3 **(-0.043)** | 1.1e3 **(0.174)** | 0.33e3 **(0.142)** | -0.121 **(-0.007)** | 0.5301 **(0.063)** | -19.05e-3 **(-0.011)** | 239.28e-3 **(0.124)** | -0.00 | 0.0000 |
| x2 | -0.011e6 **(-0.474)** | 0 **(0.001)** | 2.13e+04 | -0.001e6 **(-0.066)** | -4e3 **(-0.352)** | -0.7e3 **(-0.124)** | -0.05e3 **(-0.024)** | -0.030 **(-0.002)** | 2.7110 **(0.354)** | 488.48e-3 **(0.309)** | -445.07e-3 **(-0.254)** | 0.00 | -0.0000 |
| x3 | -0.001e6 **(-0.059)** | -0.010e6 **(-0.440)** | -0.001e6 **(-0.066)** | 2.17e+04 | -1e3 **(-0.087)** | -2.9e3 **(-0.493)** | -0.78e3 **(-0.367)** | 0.215 **(0.013)** | 1.6780 **(0.217)** | 662.43e-3 **(0.415)** | -504.42e-3 **(-0.285)** | 0.00 | -0.0000 |
| x4 | 1e3 **(0.100)** | -1e3 **(-0.043)** | -4e3 **(-0.352)** | -1e3 **(-0.087)** | 7.46e+03 | -0.1e3 **(-0.021)** | -0.36e3 **(-0.284)** | 0.046 **(0.005)** | -0.5199 **(-0.115)** | 55.37e-3 **(0.059)** | 74.65e-3 **(0.072)** | 0.00 | -0.0000 |
| x5 | 1.1e3 **(0.180)** | 1.1e3 **(0.174)** | -0.7e3 **(-0.124)** | -2.9e3 **(-0.493)** | -0.1e3 **(-0.021)** | 1.57e+03 | 0.40e3 **(0.699)** | -0.161 **(-0.037)** | -0.5235 **(-0.252)** | -151.77e-3 **(-0.354)** | 141.65e-3 **(0.297)** | -0.00 | 0.0000 |
| x6 | 0.33e3 **(0.146)** | 0.33e3 **(0.142)** | -0.05e3 **(-0.024)** | -0.78e3 **(-0.367)** | -0.36e3 **(-0.284)** | 0.40e3 **(0.699)** | 210 | -0.071 **(-0.045)** | -0.1471 **(-0.193)** | -53.26e-3 **(-0.339)** | 41.73e-3 **(0.239)** | -0.00 | 0.0000 |
| x7 | 0.037 **(0.002)** | -0.121 **(-0.007)** | -0.030 **(-0.002)** | 0.215 **(0.013)** | 0.046 **(0.005)** | -0.161 **(-0.037)** | -0.071 **(-0.045)** | 0.0121 | 0.0000 **(0.002)** | 0 | -0 **(-0.003)** | 0.000 | -0.0000 |
| x8 | -3.7283 **(-0.454)** | 0.5301 **(0.063)** | 2.7110 **(0.354)** | 1.6780 **(0.217)** | -0.5199 **(-0.115)** | -0.5235 **(-0.252)** | -0.1471 **(-0.193)** | 0.0000 **(0.002)** | 0.00275 | 0.25e-3 **(0.441)** | -0.24e-3 **(-0.376)** | 0.0000 | -0.0000 |
| x9 | -982.08e-3 **(-0.579)** | -19.05e-3 **(-0.011)** | 488.48e-3 **(0.309)** | 662.43e-3 **(0.415)** | 55.37e-3 **(0.059)** | -151.77e-3 **(-0.354)** | -53.26e-3 **(-0.339)** | 0 | 0.25e-3 **(0.441)** | 0.000117 | -0.03e-3 **(-0.255)** | 0 | -0 |
| x10 | 452.13e-3 **(0.240)** | 239.28e-3 **(0.124)** | -445.07e-3 **(-0.254)** | -504.42e-3 **(-0.285)** | 74.65e-3 **(0.072)** | 141.65e-3 **(0.297)** | 41.73e-3 **(0.239)** | -0 **(-0.003)** | -0.24e-3 **(-0.376)** | -0.03e-3 **(-0.255)** | 0.000145 | -0 | 0 |
| x11 | 0.00 | -0.00 | 0.00 | 0.00 | 0.00 | -0.00 | -0.00 | 0.000 | 0.0000 | 0 | -0 | 0.09 | -0.0000 |
| x12 | -0.0000 | 0.0000 | -0.0000 | -0.0000 | -0.0000 | 0.0000 | 0.0000 | -0.0000 | -0.0000 | -0 | 0 | -0.0000 | 0.0025 |

- Likelihood with 13 fit parameter
  2: over-/underflow bin
- 6: unfolding bins
- 5: ice systematic parameter

➢ minimize using Minuit

➢ provides full covariance matrix (best fit + uncertainty)

Idea:

- use best fit value of each systematic → scale up and down by uncertainty → fix systematics in unfolding
- 5 systematics → 10 unfoldings in total → re-run unfolding 10 times ($f_{\mathrm{sys},i}$)
- $\sigma_{\mathrm{sys}} = \sqrt{\sum (f_{\mathrm{sys},i} - f_{\mathrm{base}})^2}$, $i$: 10 sys. unfoldings

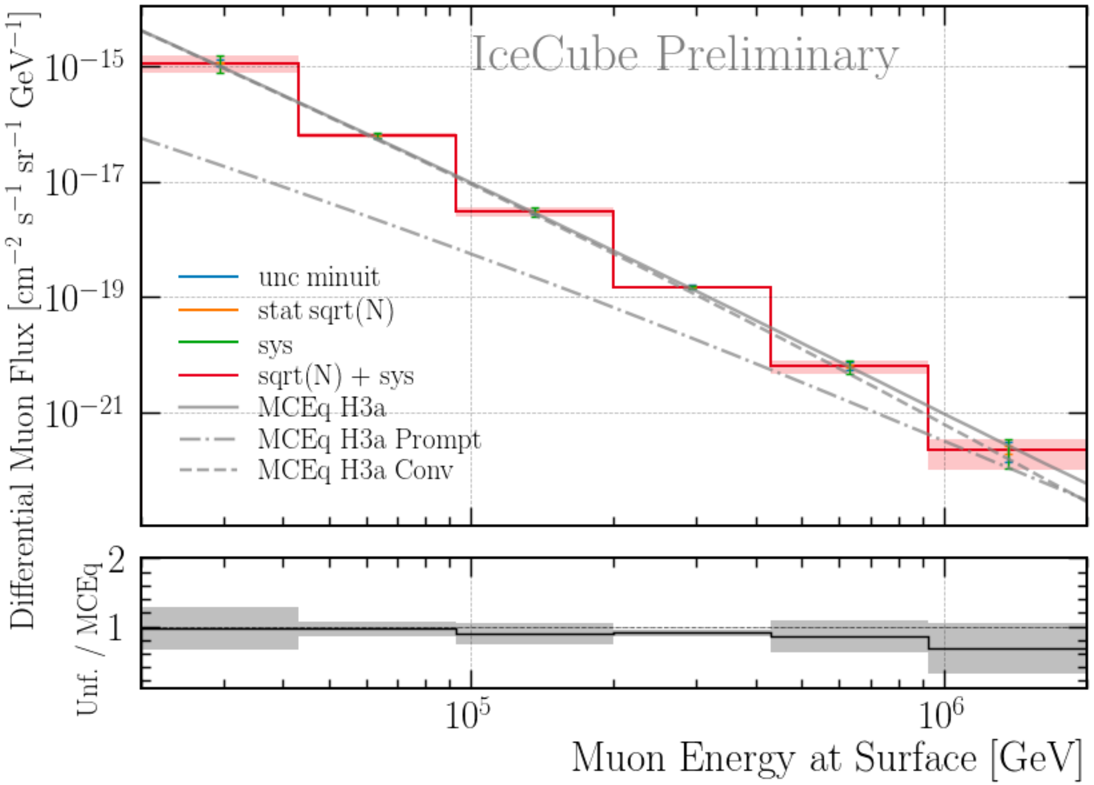
$$\sigma_{\mathrm{tot}} = \sqrt{\sigma_{\mathrm{stat}}^2 + \sum \sigma_i^2}\,, i\text{: scat, abs, DOME, holeice p0, p1}$$

pascal.gutjahr@tu-dortmund.de

# Test Method on MC

- run unfolding on MC
- re-run 10 times
  - fix all systematics in fit
  - take center value of simulation range → scale one systematic parameter up/down by 50% of the simulation range

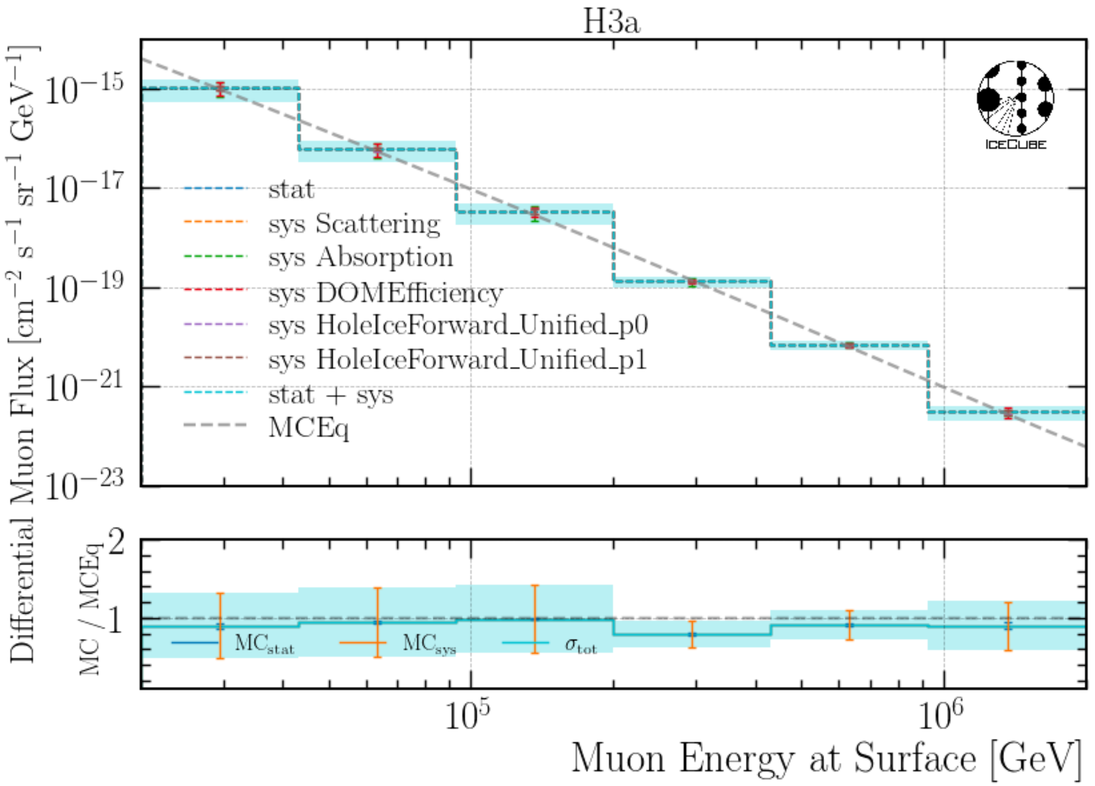
Systematics: center ± 50% simulation range
Statistical: sqrt(N)



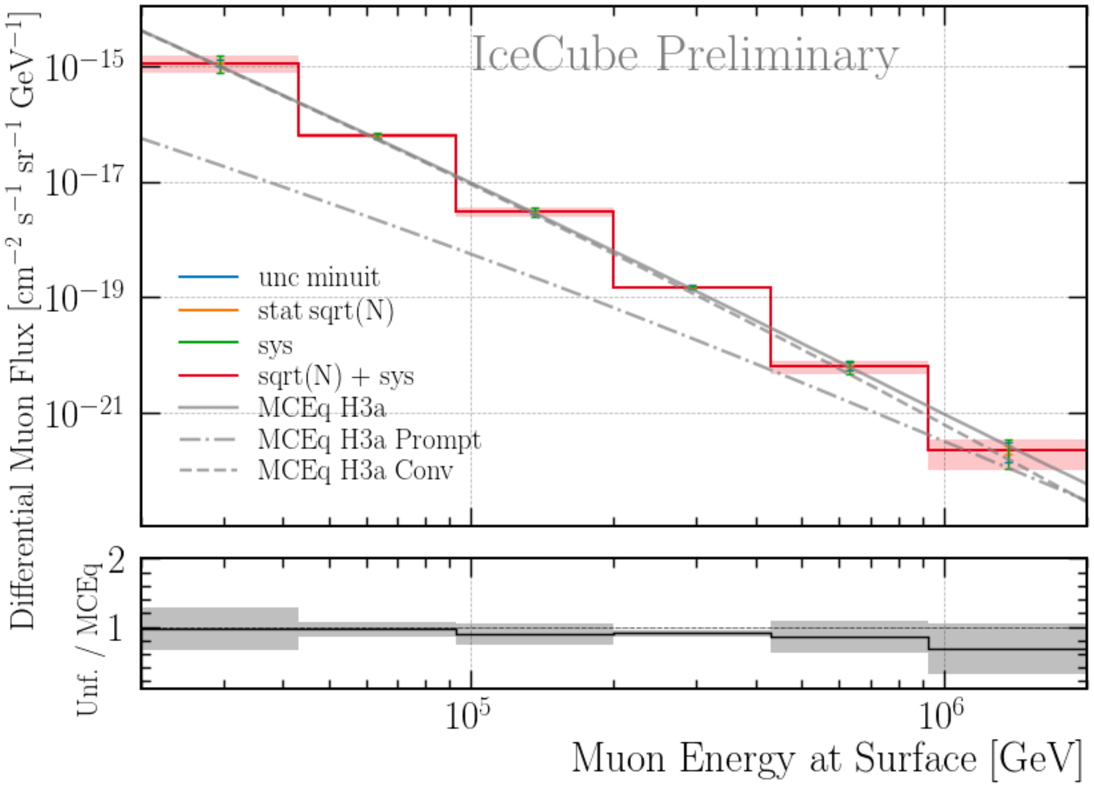
tu dortmund university BERKELEY LAB ICECUBE NEUTRINO OBSERVATORY

# "Round trip"

Systematics: above/below center
Statistical: sum(weights**2)



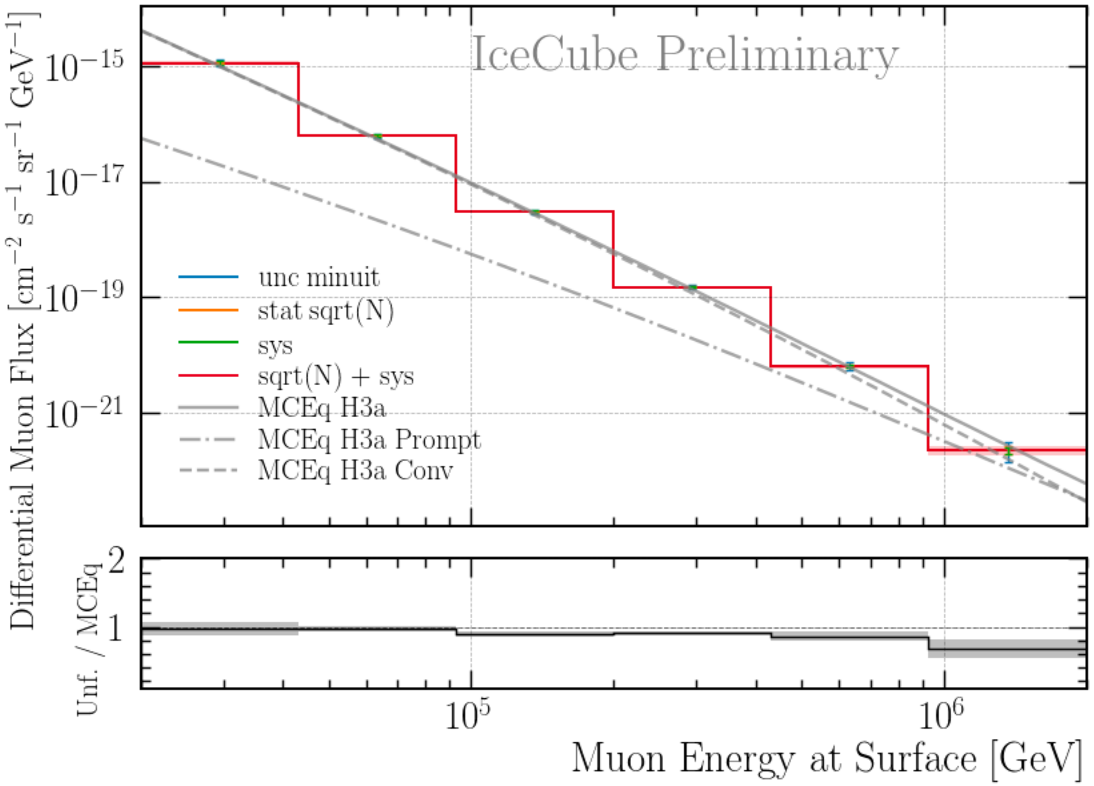

Systematics: center ± 50% simulation range
Statistical: sqrt(N)



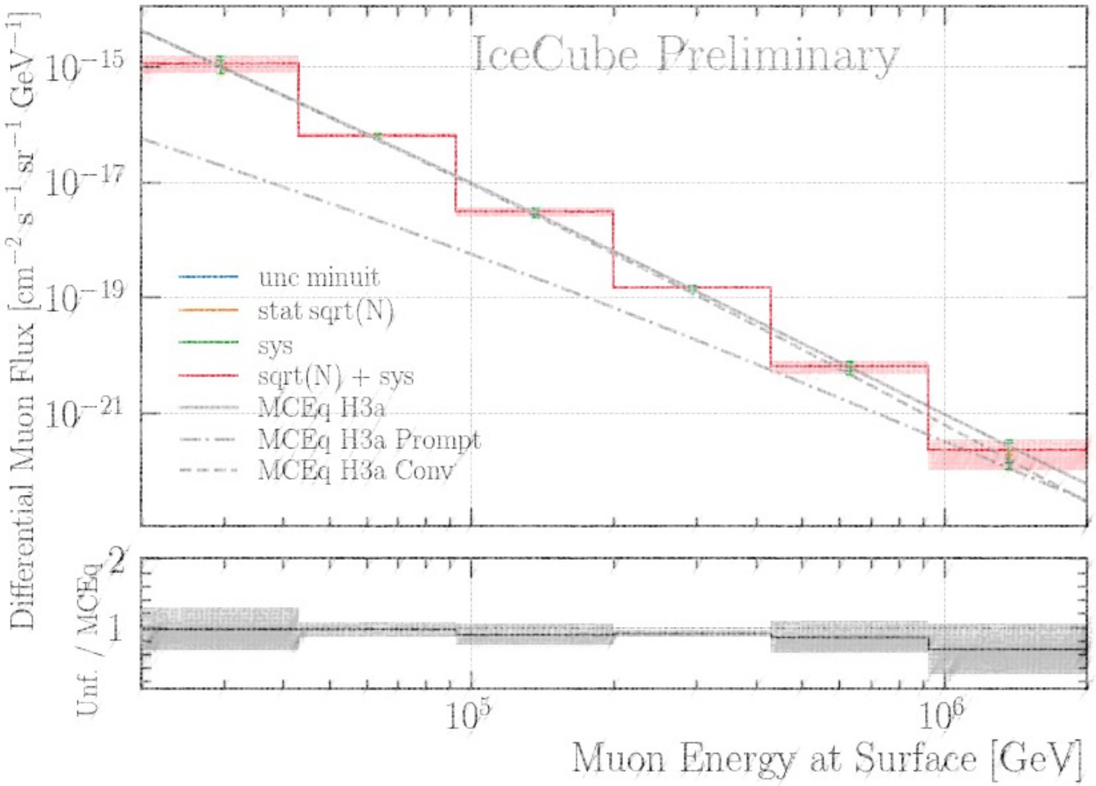

- both methods provide similar uncertainties
- it is NOT expected that the uncertainties are the same
(left method also includes limited MC statistics from the weights)

# Burnsample MC

Systematics: best fit ± sigma
Statistical: sqrt(N)



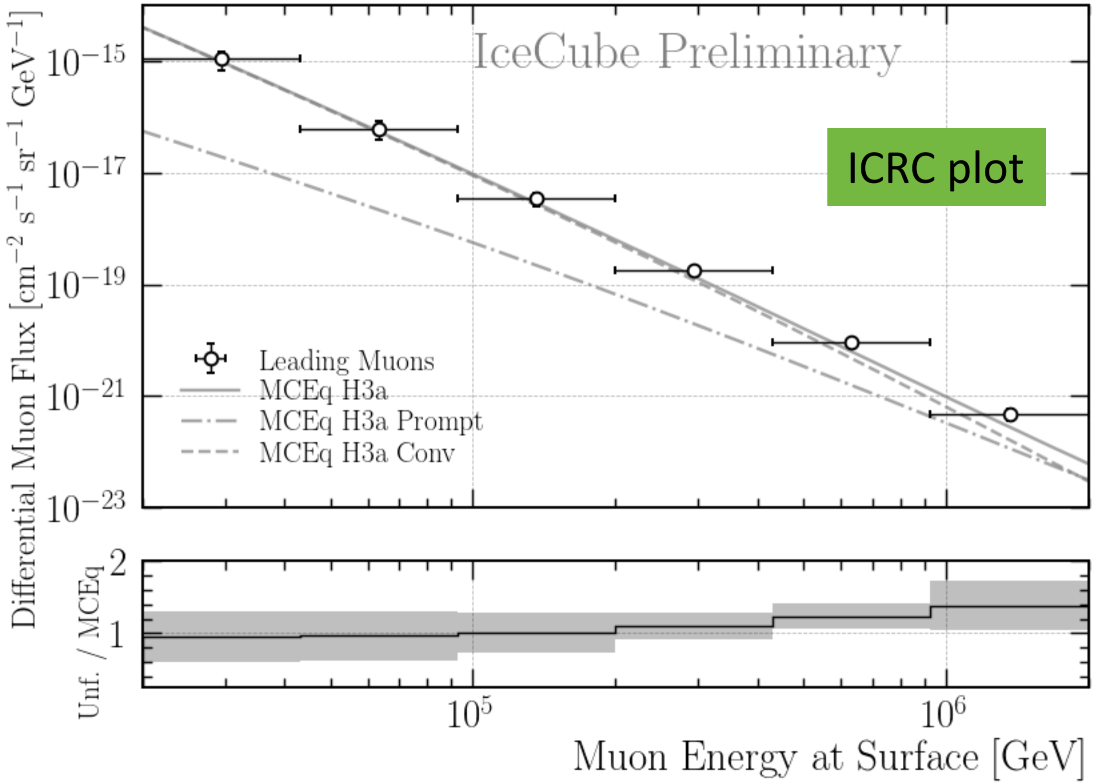

Systematics: center ± 50% simulation range
Statistical: sqrt(N)

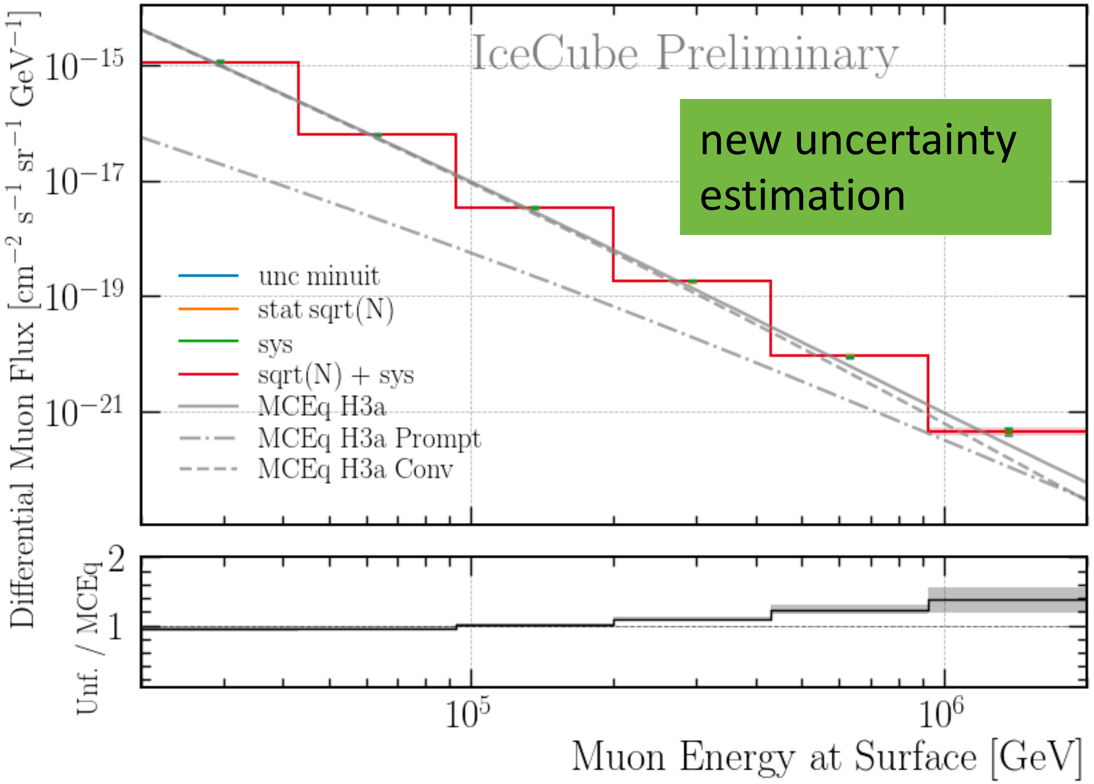


# Burnsample Data

Systematics: ± 50% on MC effective area
Statistical: minuit uncertainty

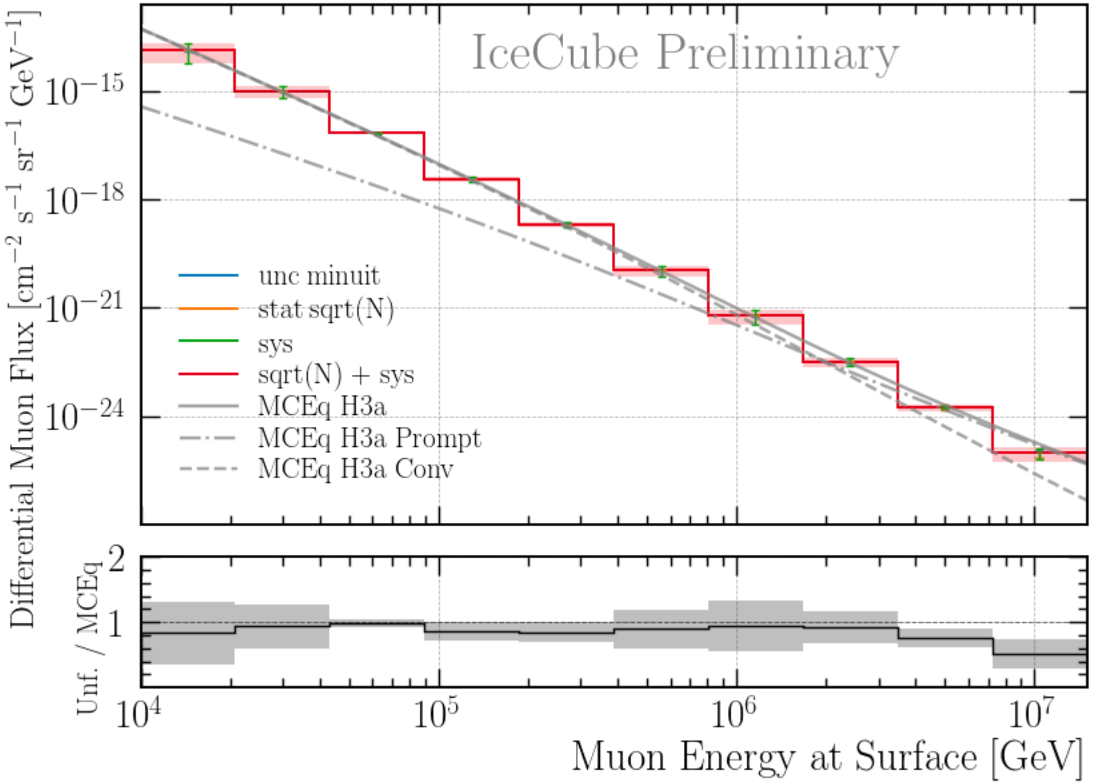
Systematics: best fit ± sigma
Statistical: sqrt(N)



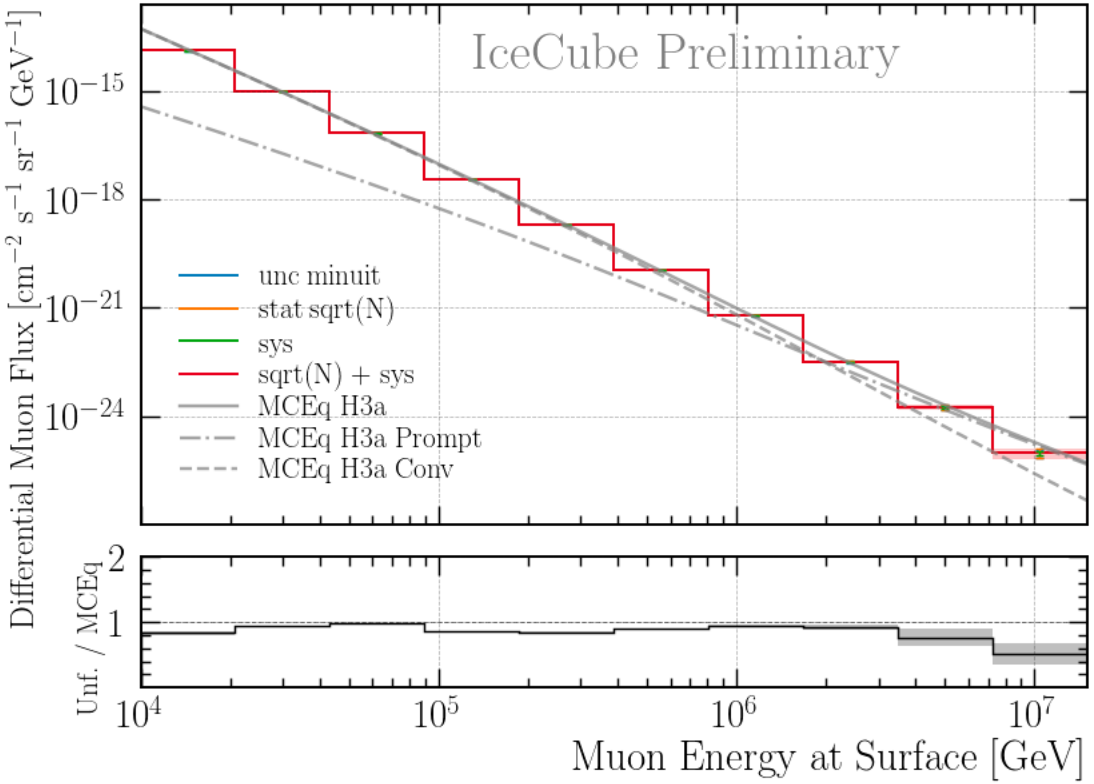



# Full Statistics MC

Systematics: center ± 50% simulation range
Statistical: sqrt(N)




Systematics: best fit ± sigma
Statistical: sqrt(N)

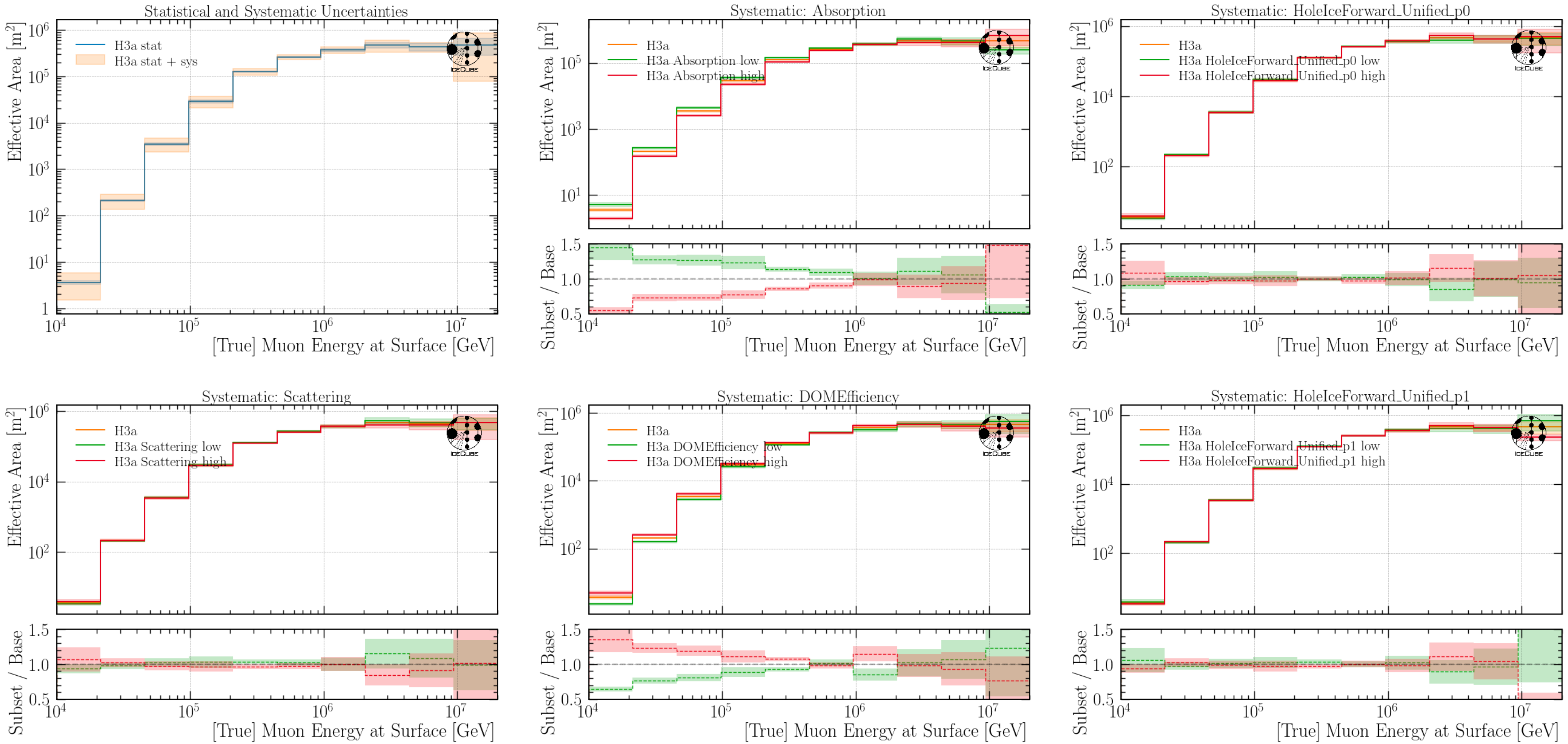


# Conclusion

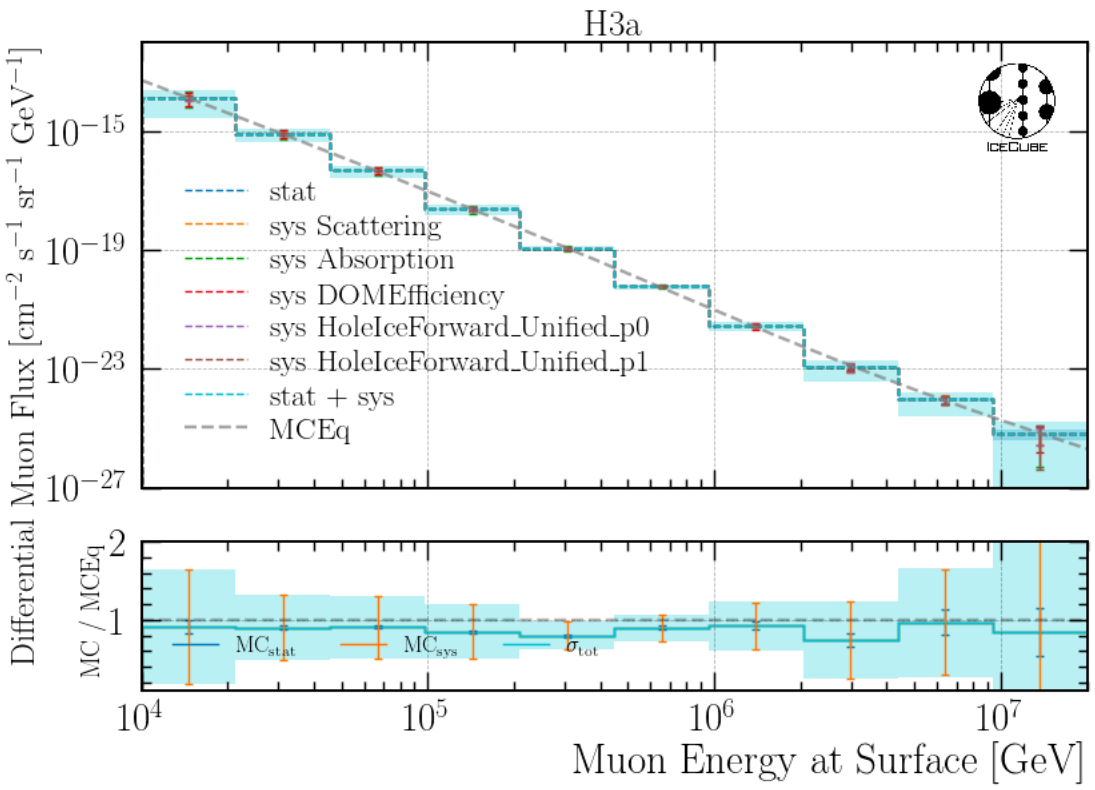
- Coincident primaries estimated → negligible towards higher energies
  - primarily low-energy contribution
- Estimate systematic uncertainties via scaling of systematics and re-run unfolding
  - more precise uncertainty estimation
  - “tested” against up/down-scaling of effective area calculation
    - similar results (same results are not expected)

Questions?

# Effective Area Systematics

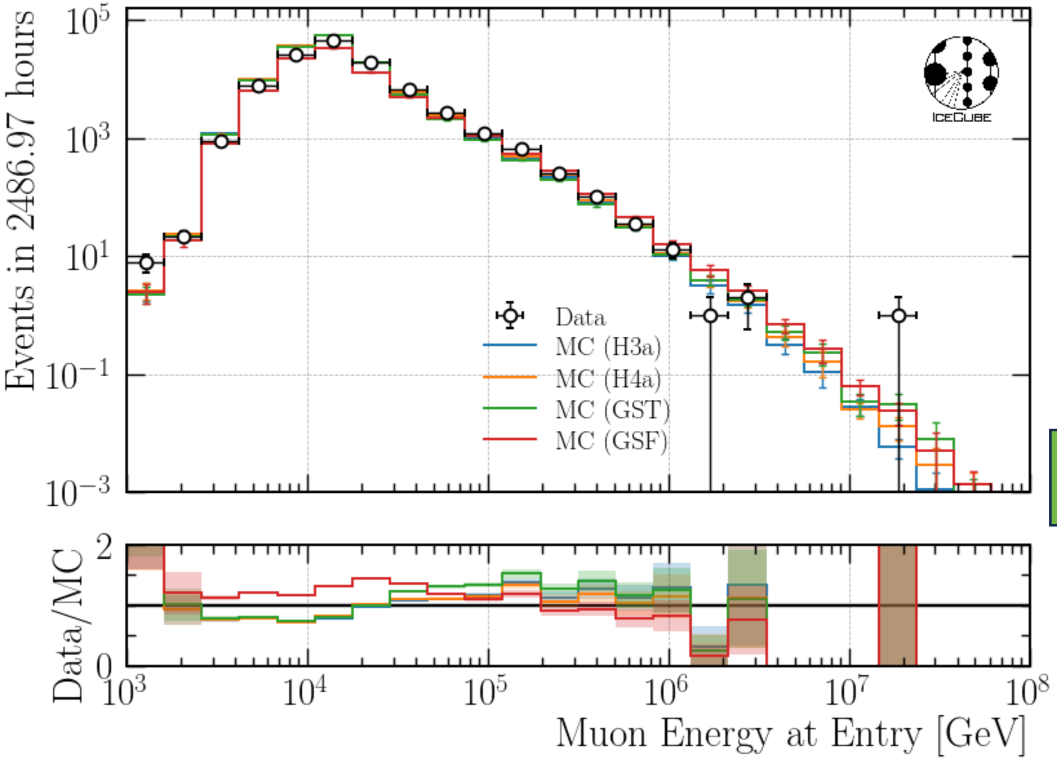


# Full Statistics MC Systematics

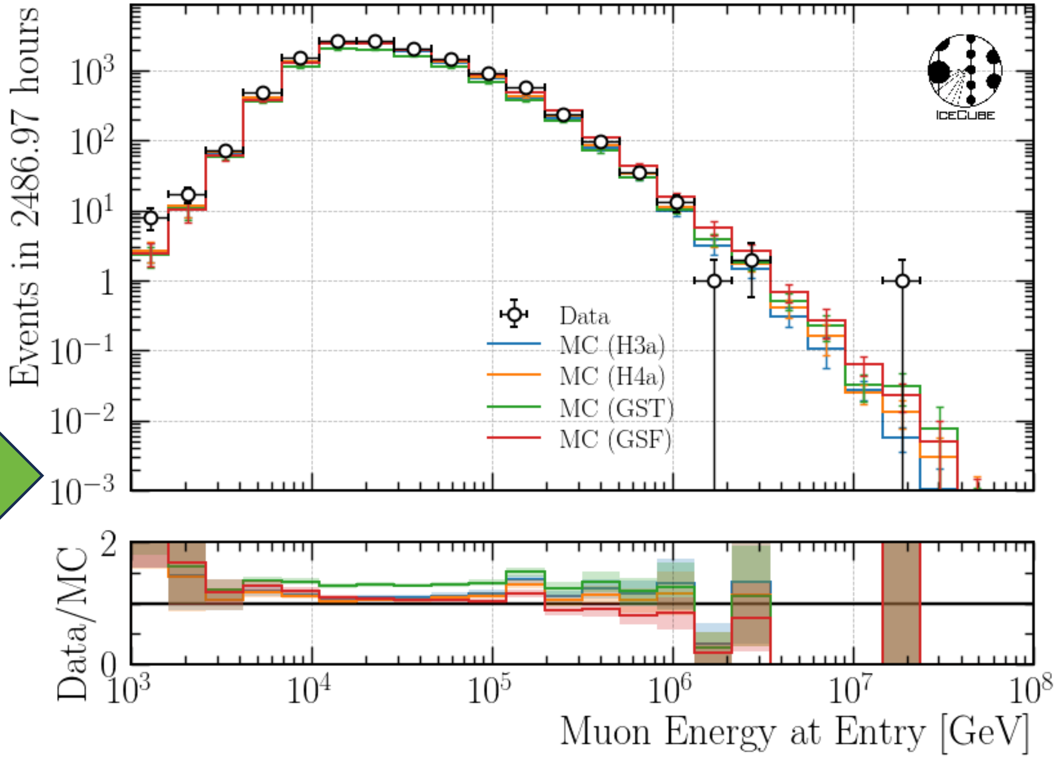


# Data—MC Proxy Variable

$$\text{Leadingness} = \frac{\text{Leading Energy}}{\text{Bundle Energy}}$$




Leadingness > 40%

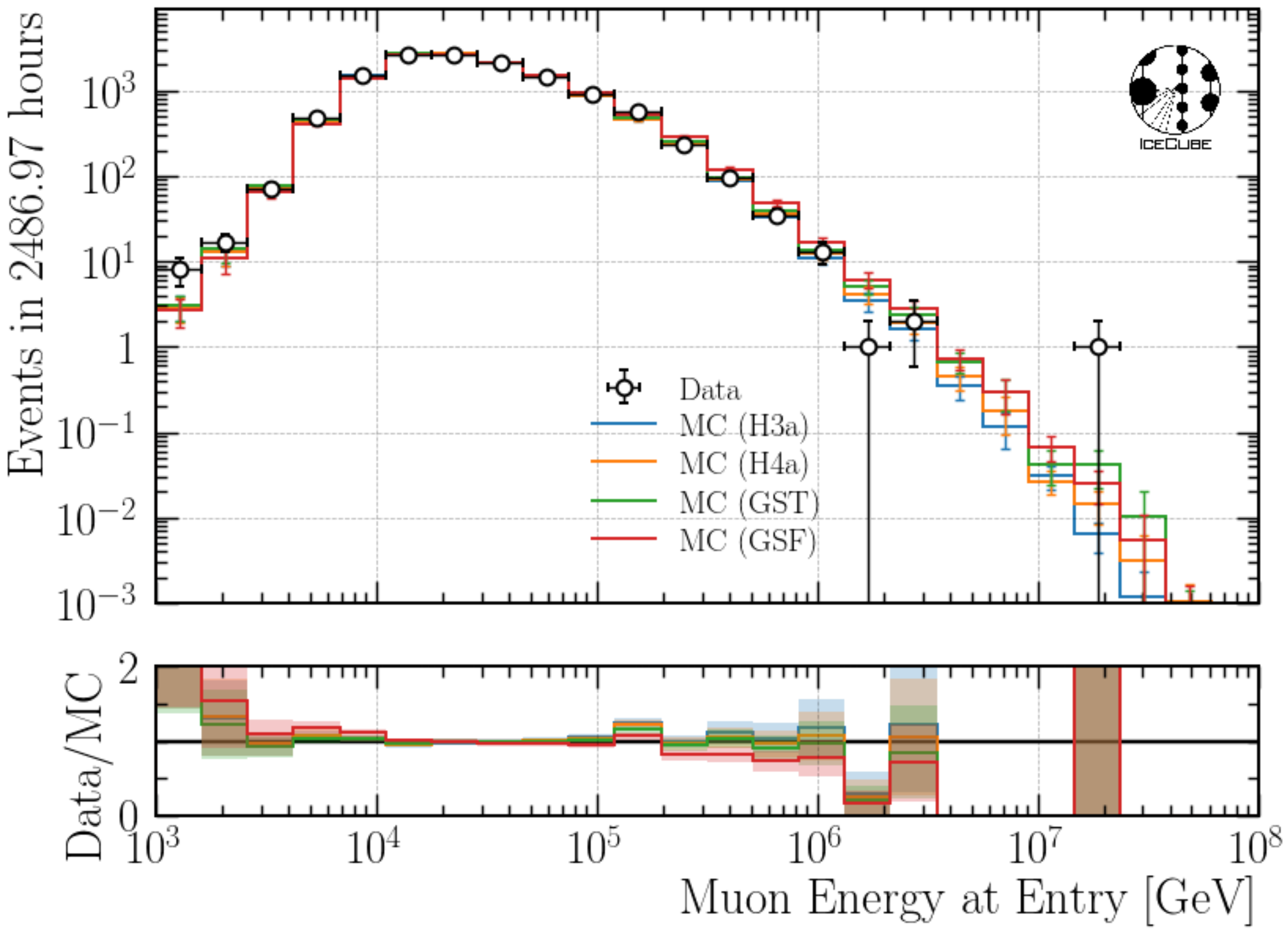


dortmund university BERKELEY LAB ICECUBE NEUTRINO OBSERVATORY

# Data—MC Proxy Variable: Leadingness > 40 %

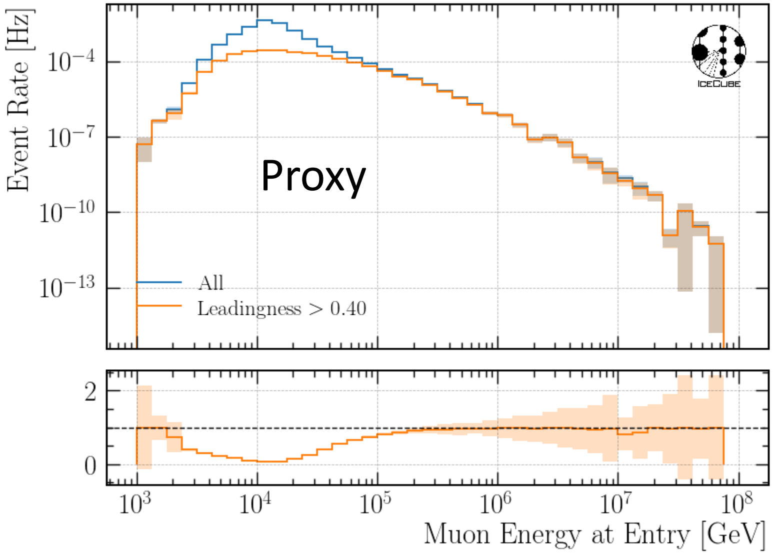
Nomalizations:

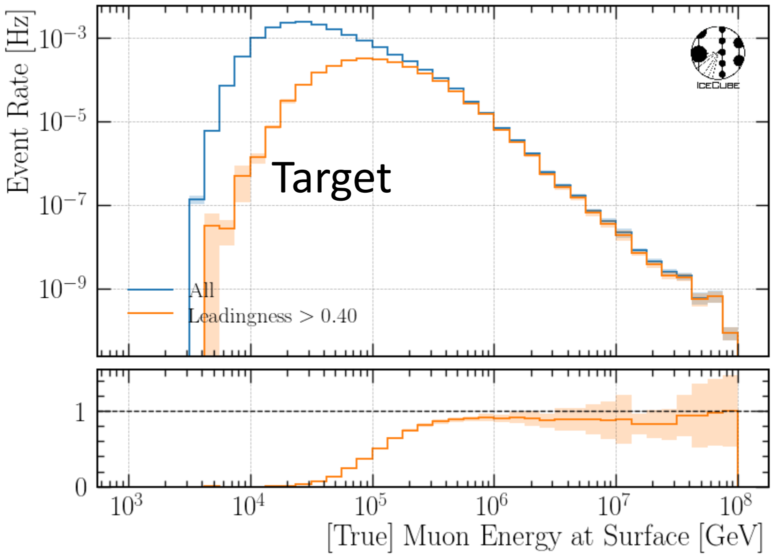
- H3a: + 12 %
- H4a: + 9%
- GST: + 32%
- GSF: + 9%



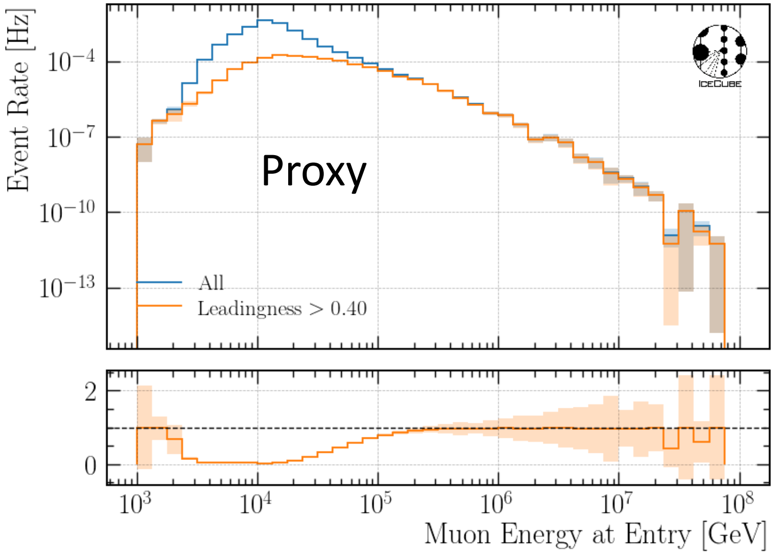

- ➢ MC shapes agree with data
- ➢ Total normalization has NO impact on unfolding

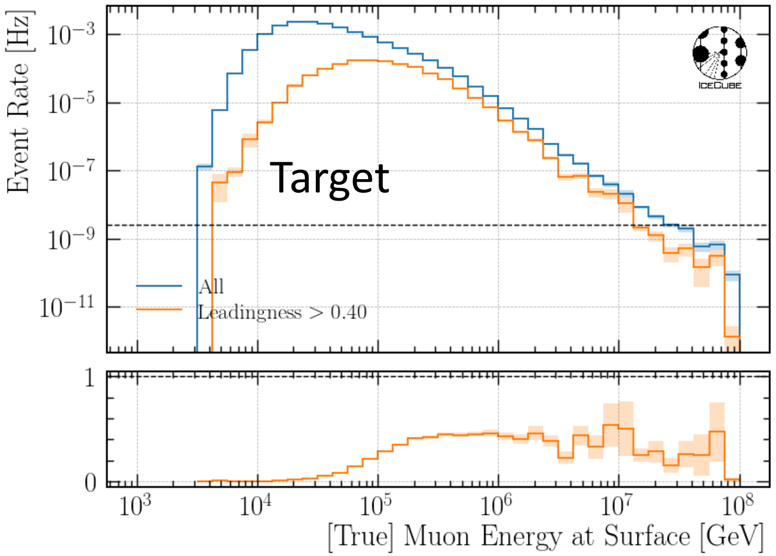
# Proxy & Target Distributions Before/After cut (True & Reco)



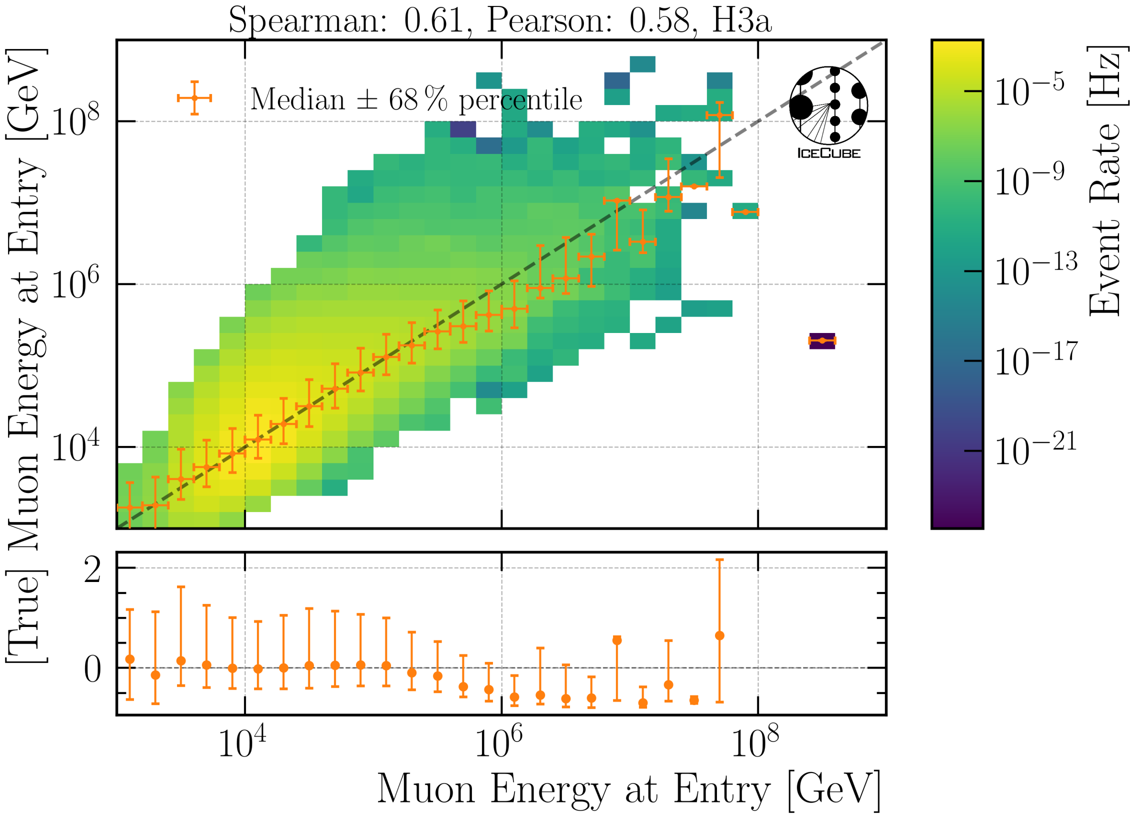

Cut on TRUE leadingness

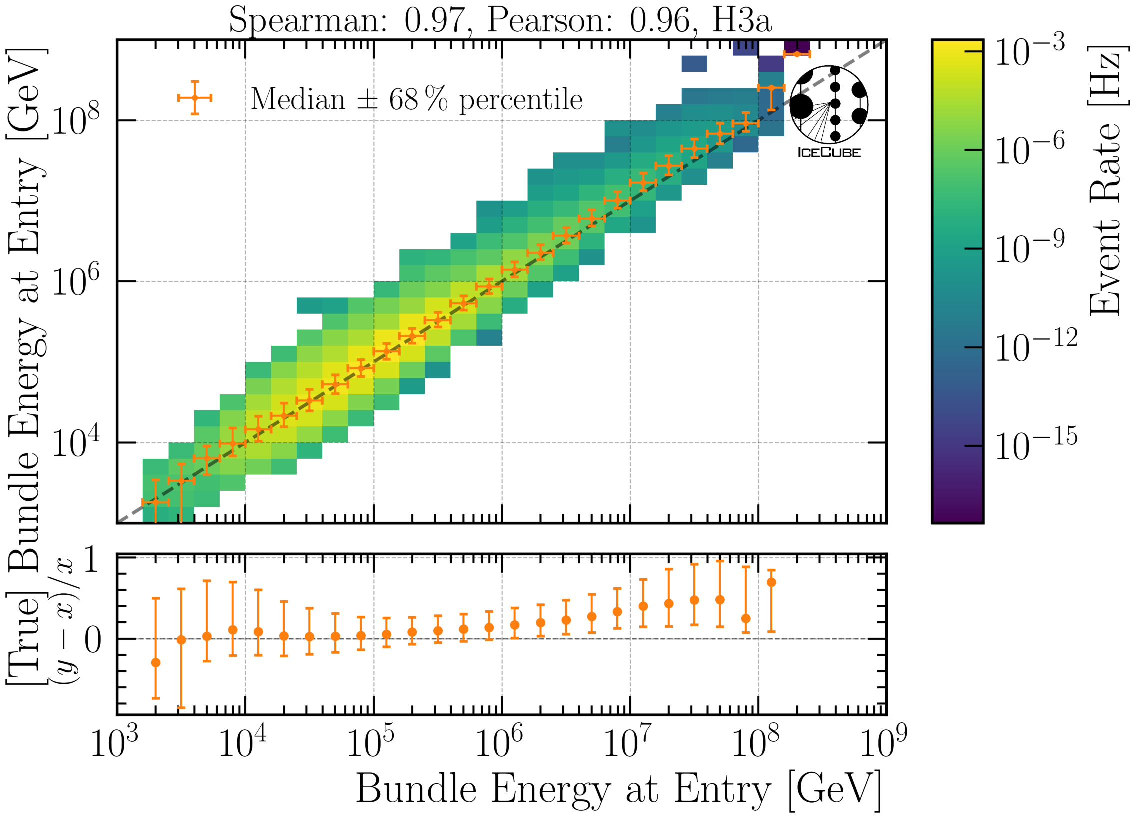


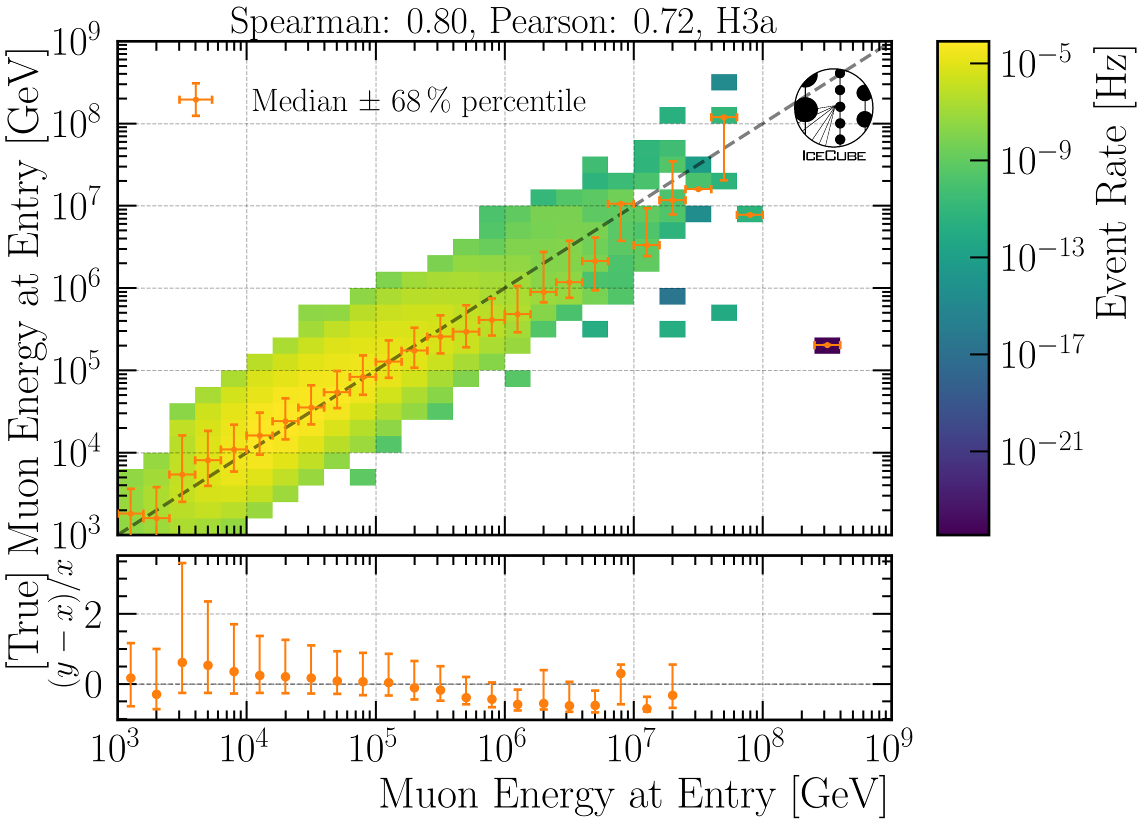

Cut on PRED leadingness

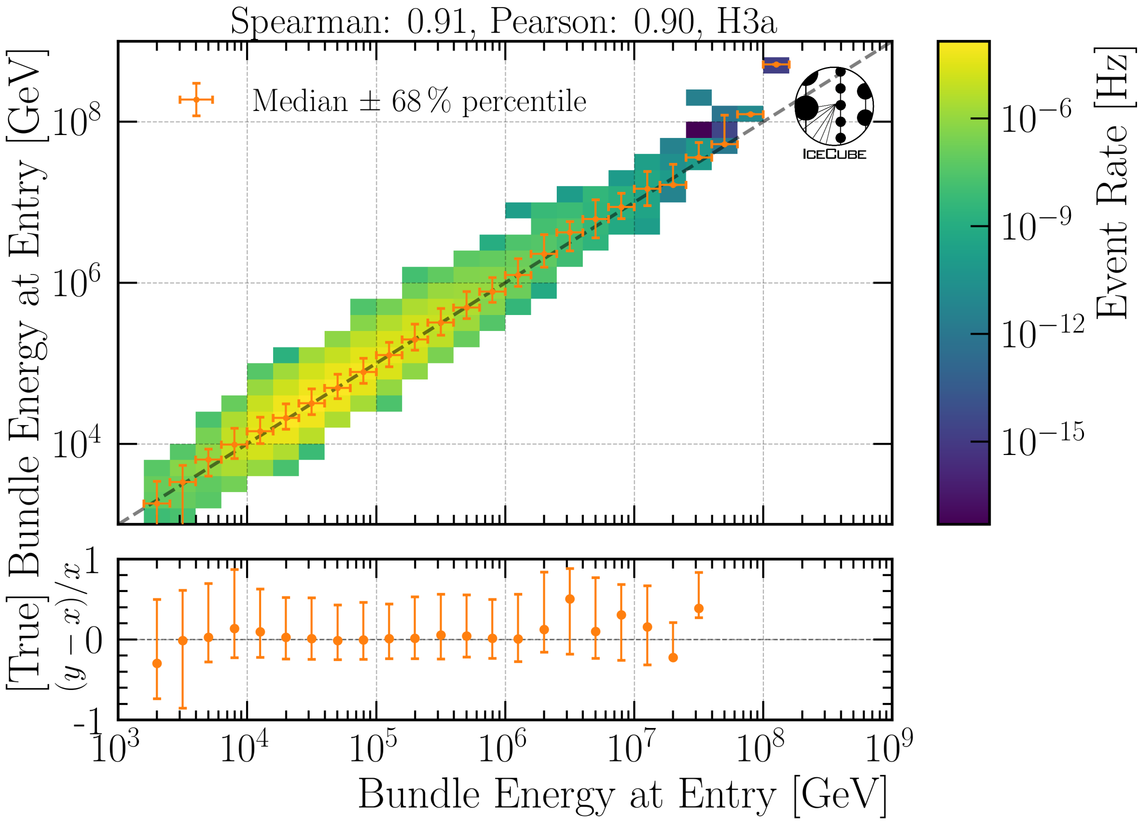
# Reconstructions for Leadingness: Level 5



# Reconstructions after Leadingness cut: Final Level

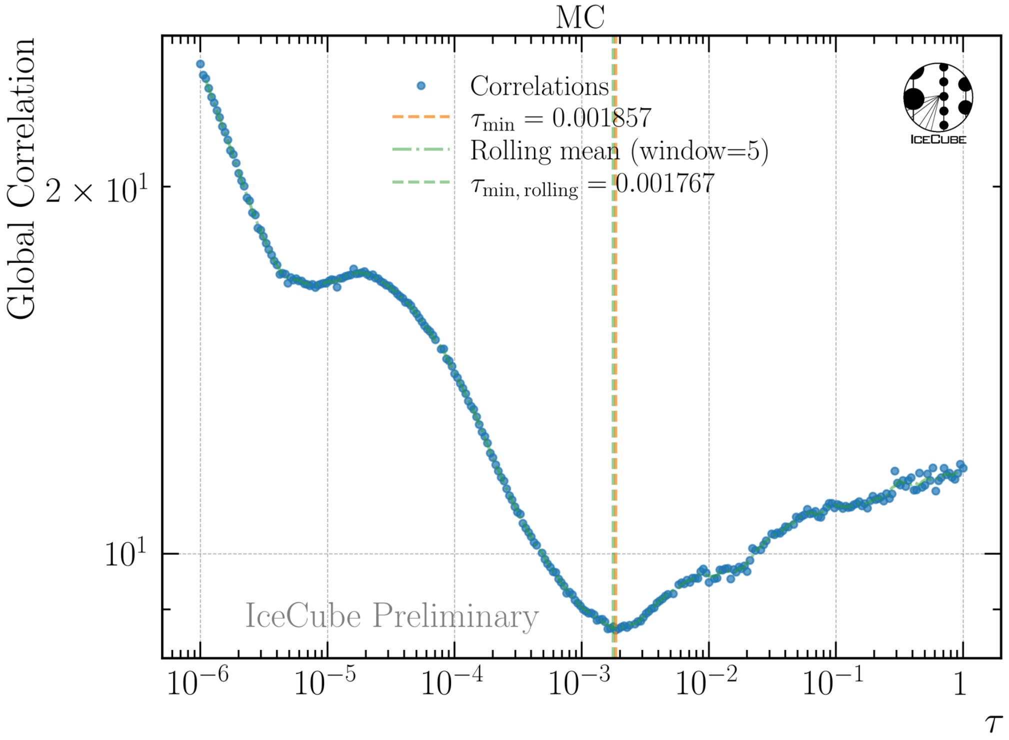


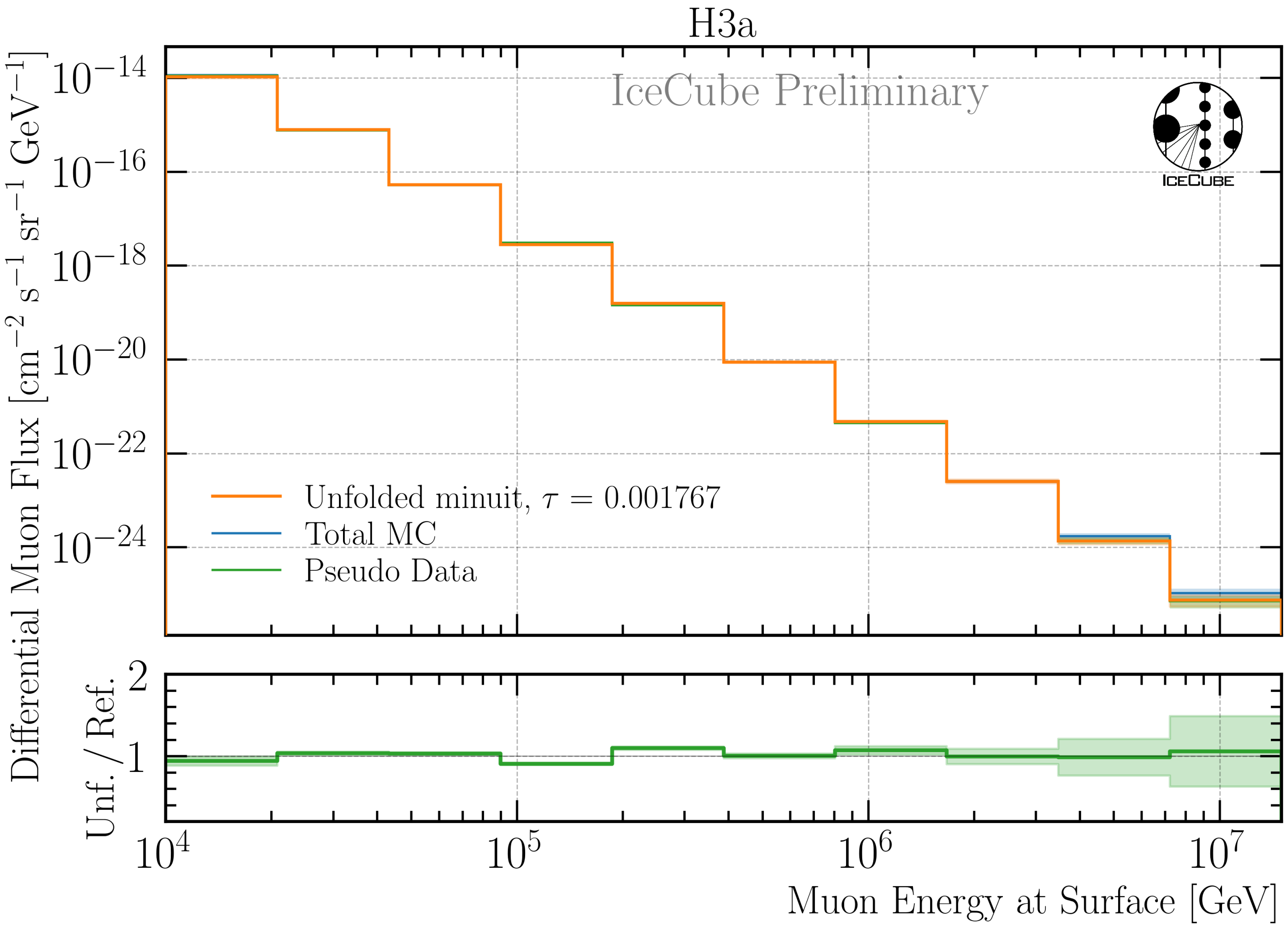




- ➢ Improved reconstructions

# Unfolding on 12.12 years MC Pseudo data

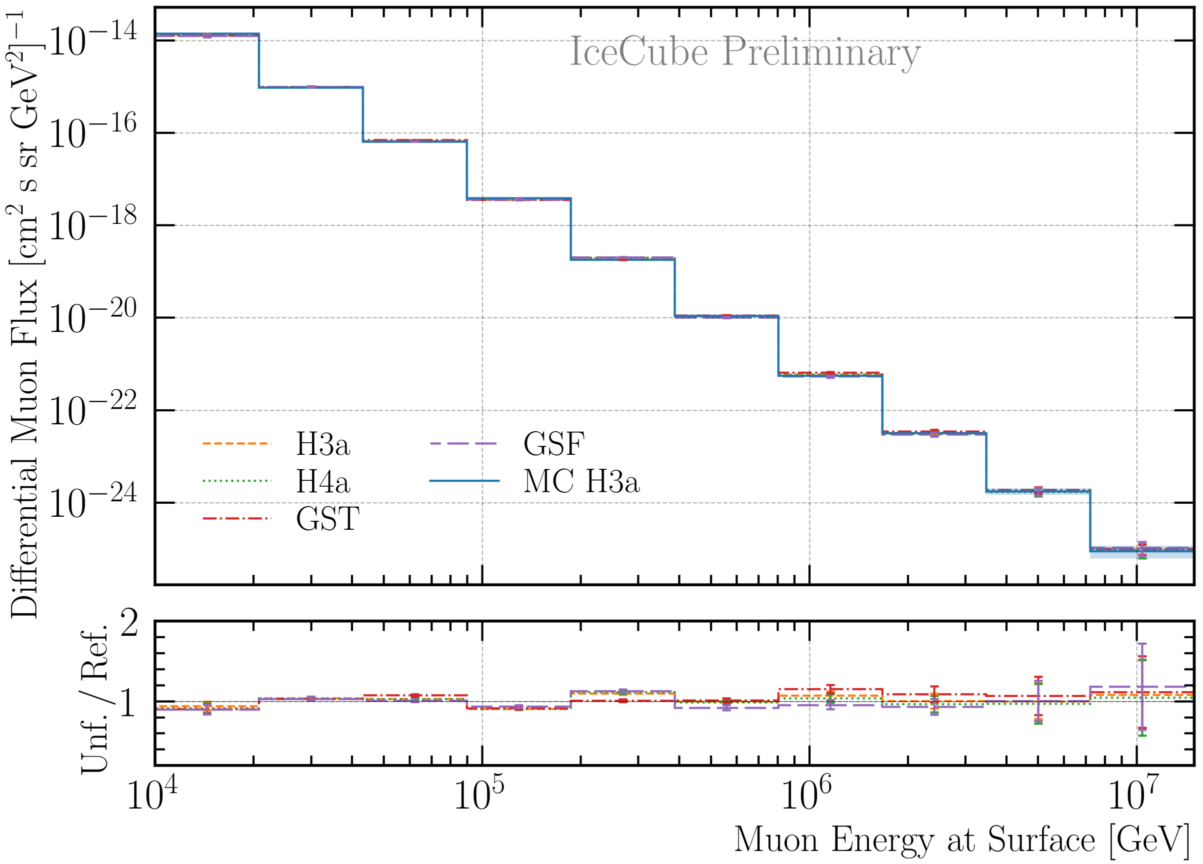
- Find reguluraization with minimal correlation: $\rho = \sum_i \sqrt{1 - \left(\mathbf{V}_{ii} \cdot \mathbf{V}_{ii}^{-1}\right)}$
  (V: Covariance Matrix)



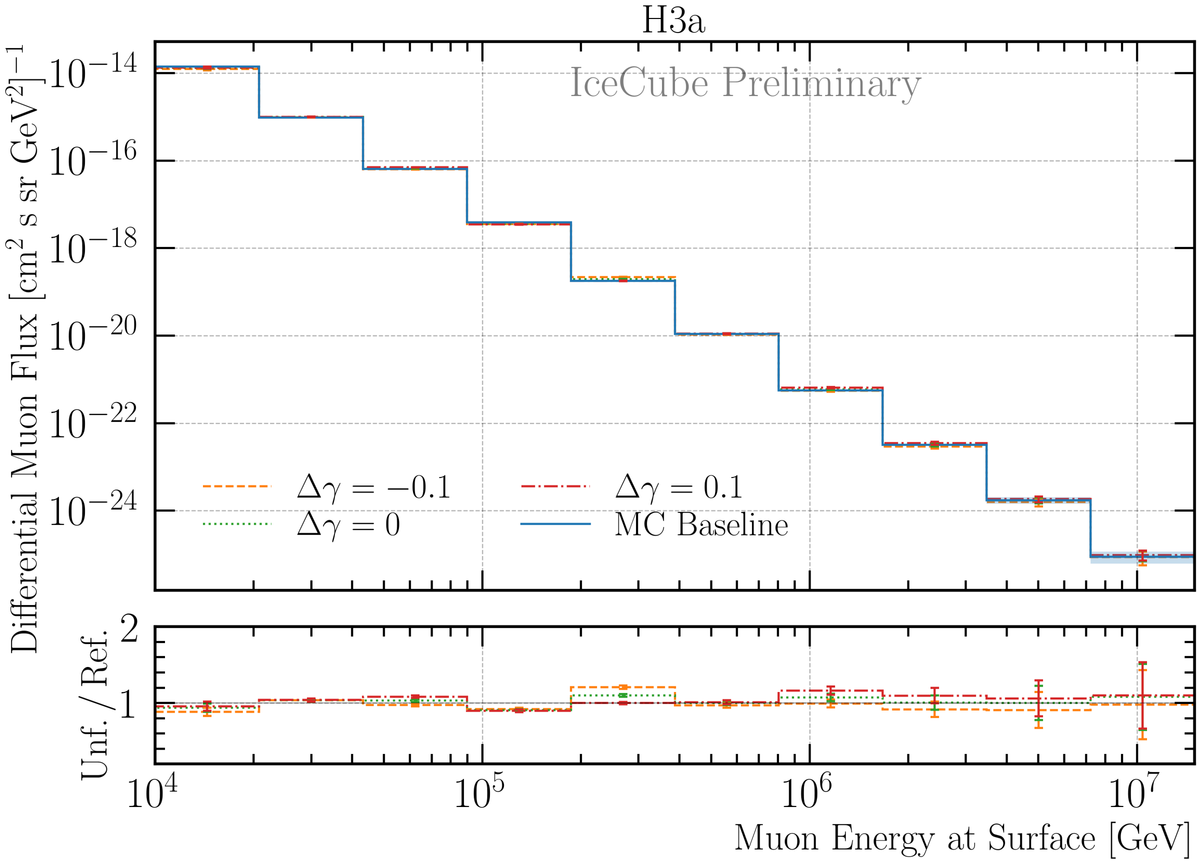



# Robustness Tests

- Build unfolding matrix on H3a, H4a, GST and GSF
- Unfold H3a as “test data”

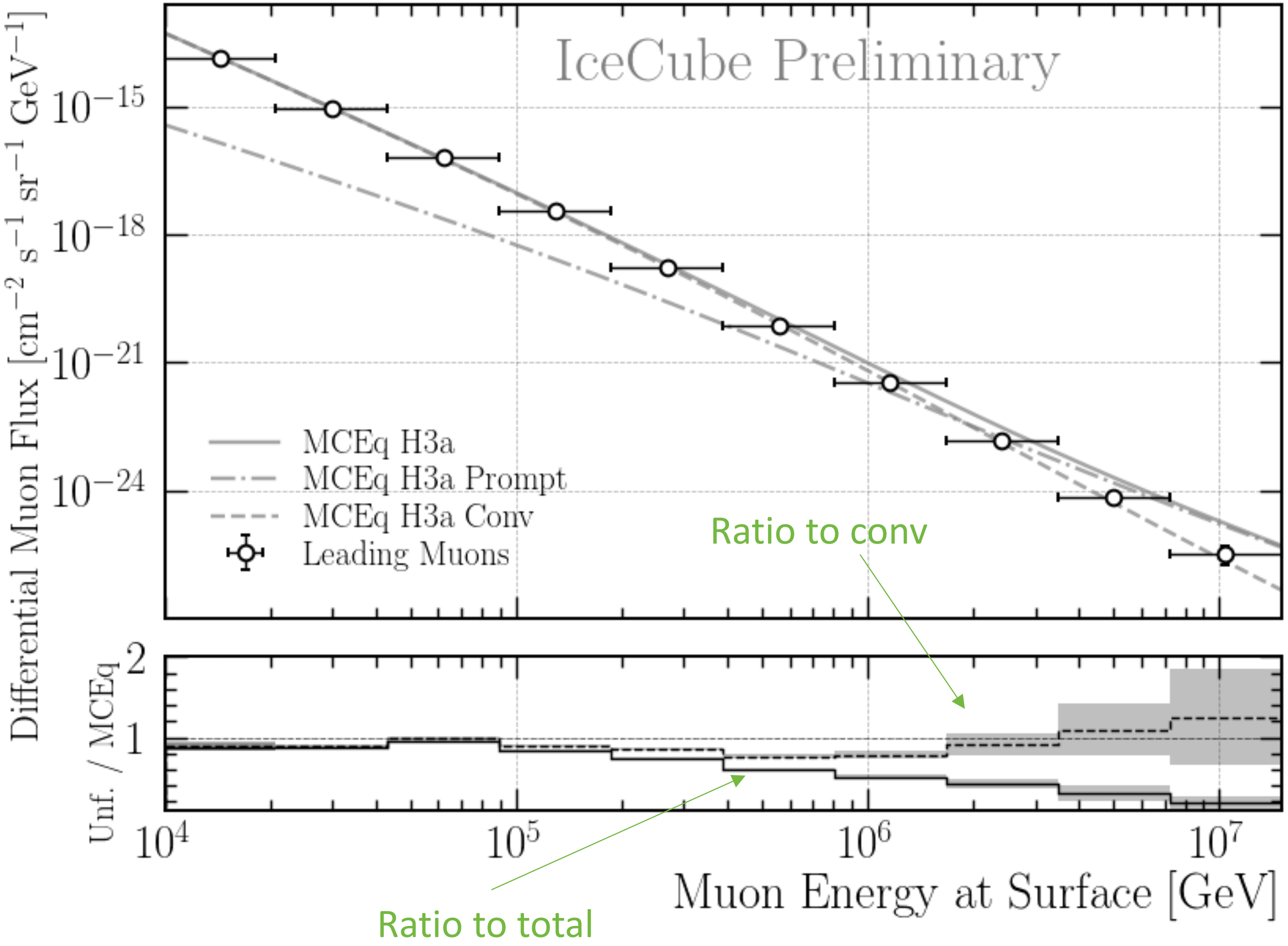



- Build unfolding matrix on H3a with $\gamma \pm 0.1$
- Unfold H3a as “test data”



# Test: Pseudo Data set includes ONLY conv

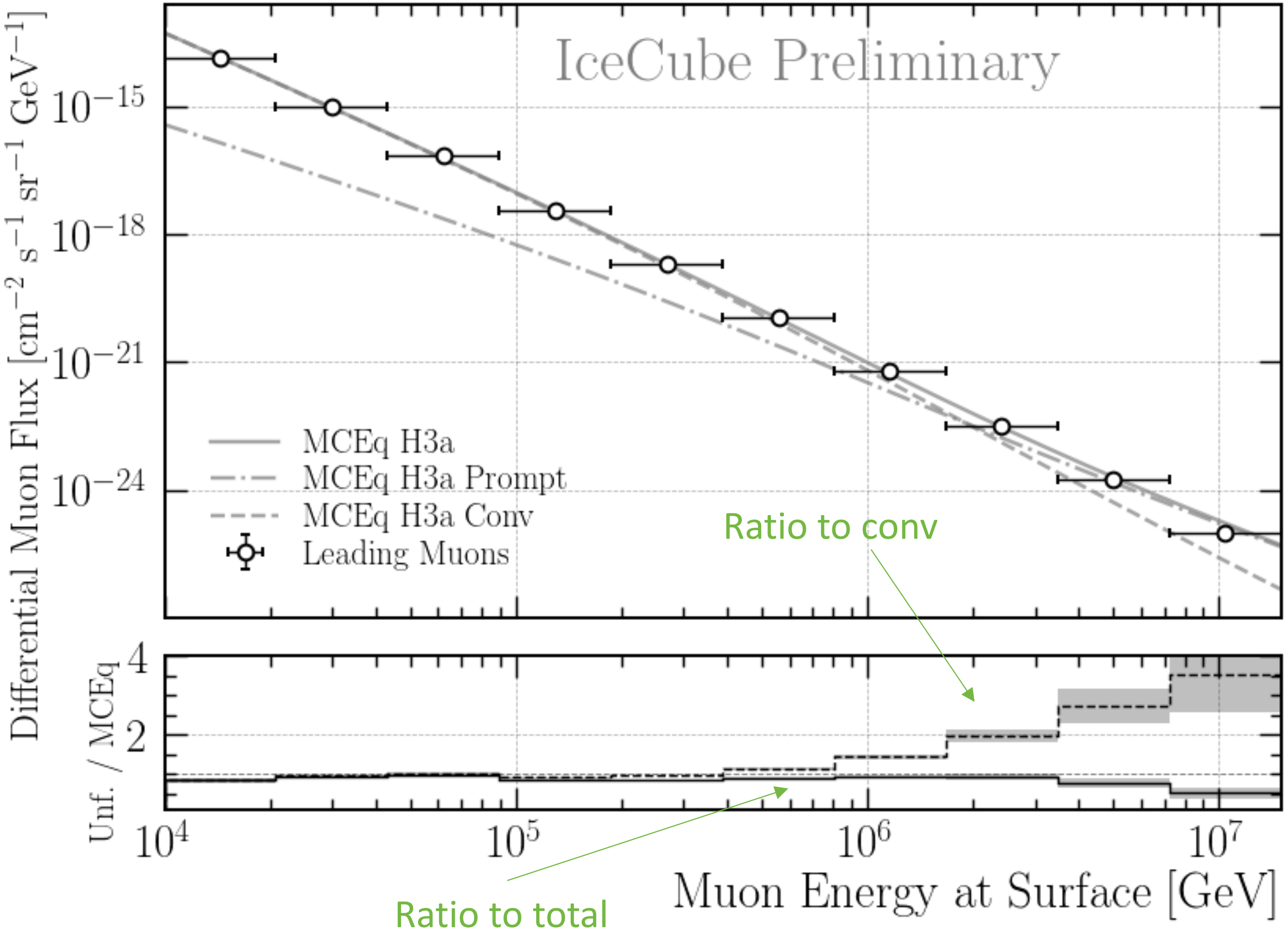
- Expectation:
  Unfolding agrees with conventional
- Works as expected




No effective area are uncertainties shown

# Test: Pseudo data includes prompt & conv

- Expectation:
  Unfolding agrees with total
- Works as expected




No effective area are uncertainties shown

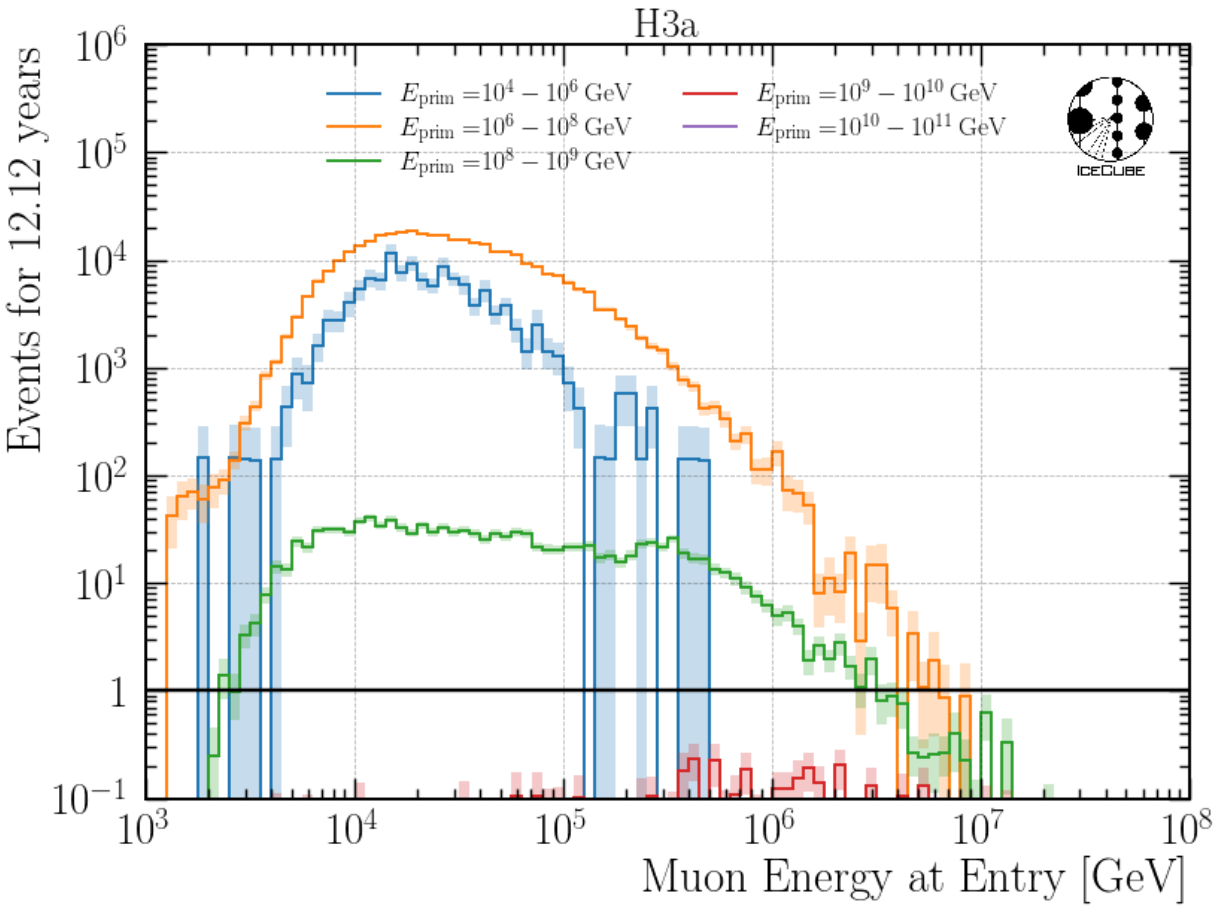
# Conclusion

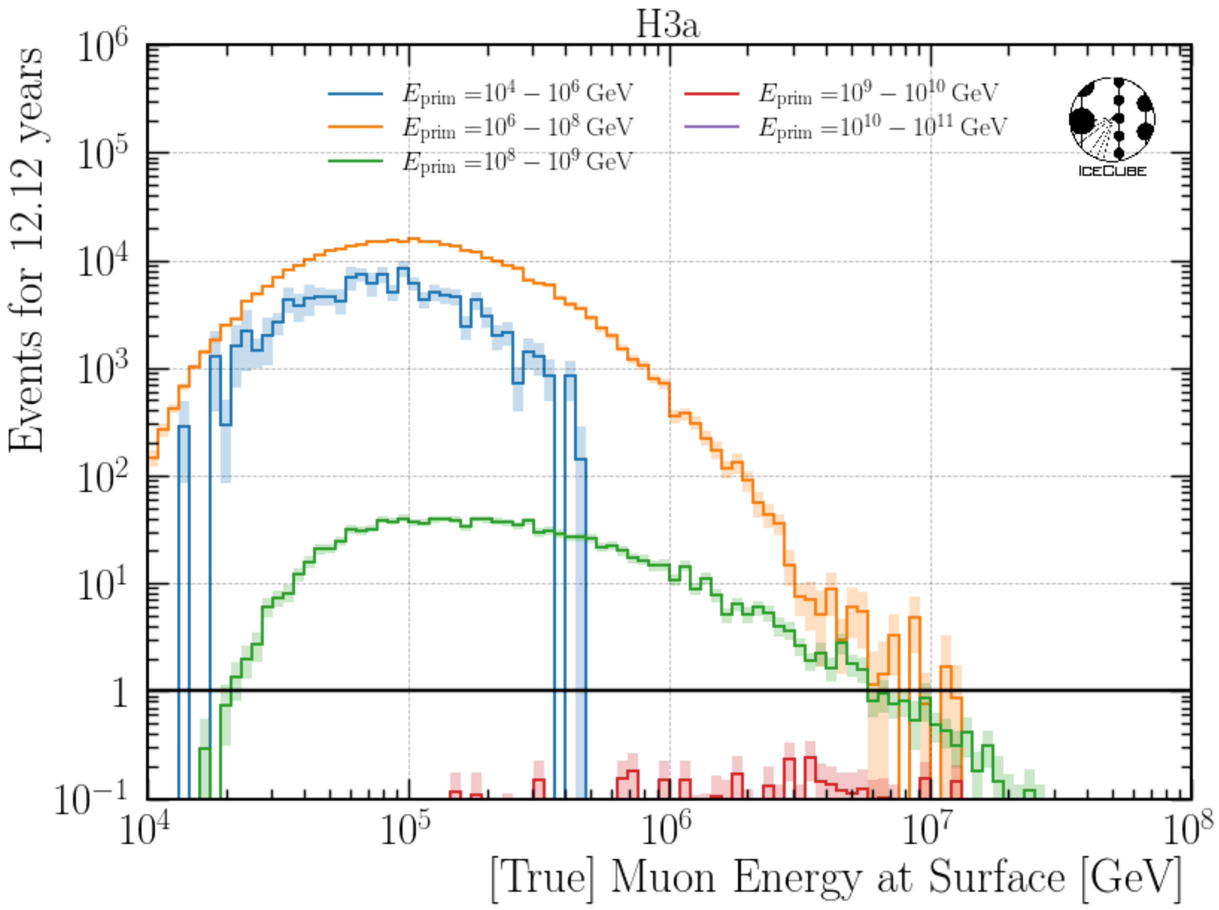
- Leadingness > 40 % improves data—MC of proxy variable
  - Additional cut → Final Level
  - Better reconstructions
  - Better correlation
- Unfolding works on MC for 12.12 years of data
- Robustness for different primary mass compositions and $\gamma \pm 0.1$ presented
- Algorithm is sensitive to prompt
  - When injecting conv only → conv is revovered
  - When injecting conv and prompt → total is recovered

# Binning – info for myself

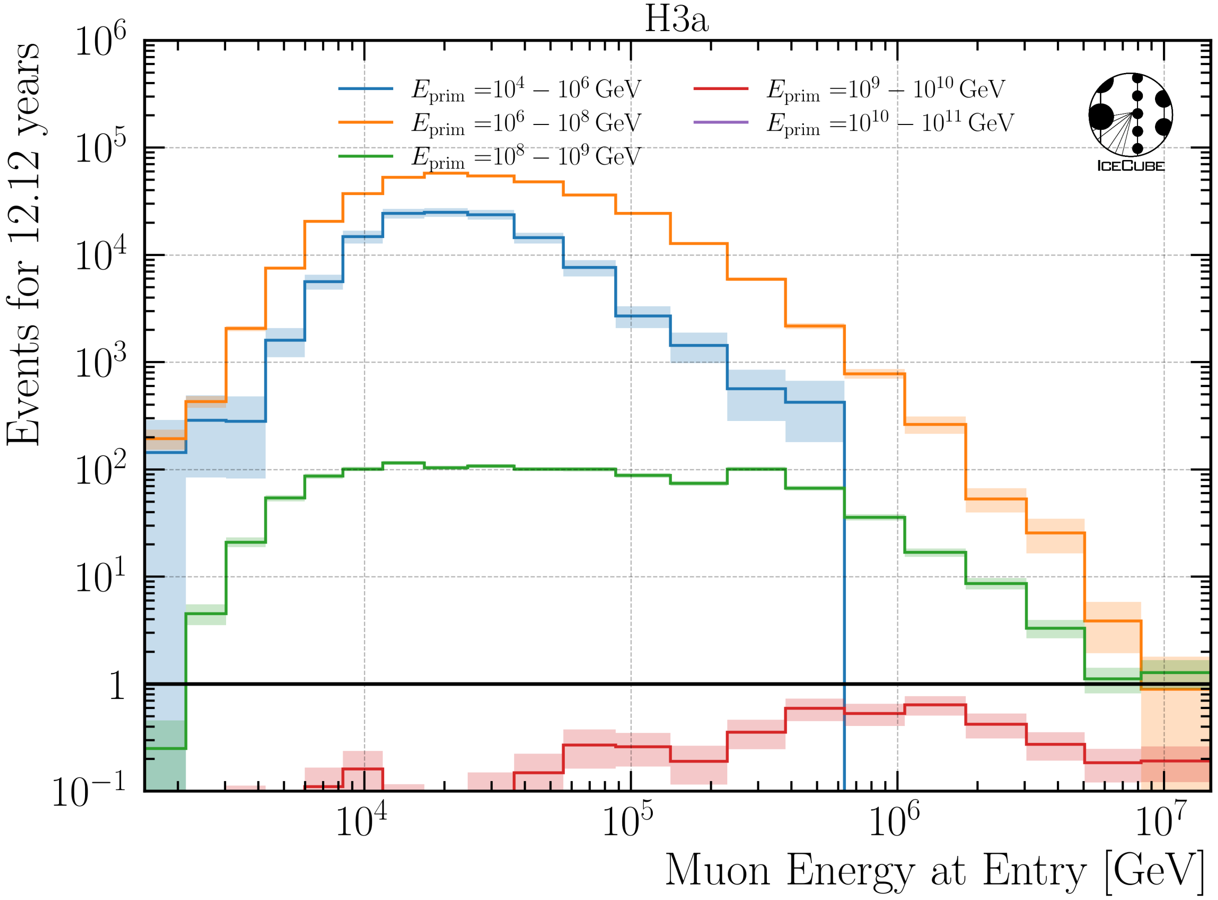
- Plots in these slides have been created with the proxy resolution binning k = 1
- Proxy: bins = np.geomspace(1.5e3, 1.5e7, 11)
- Unf: unf_bins = np.geomspace(1e4, 1.5e7, 11)
- Test set: H3a, seed 42

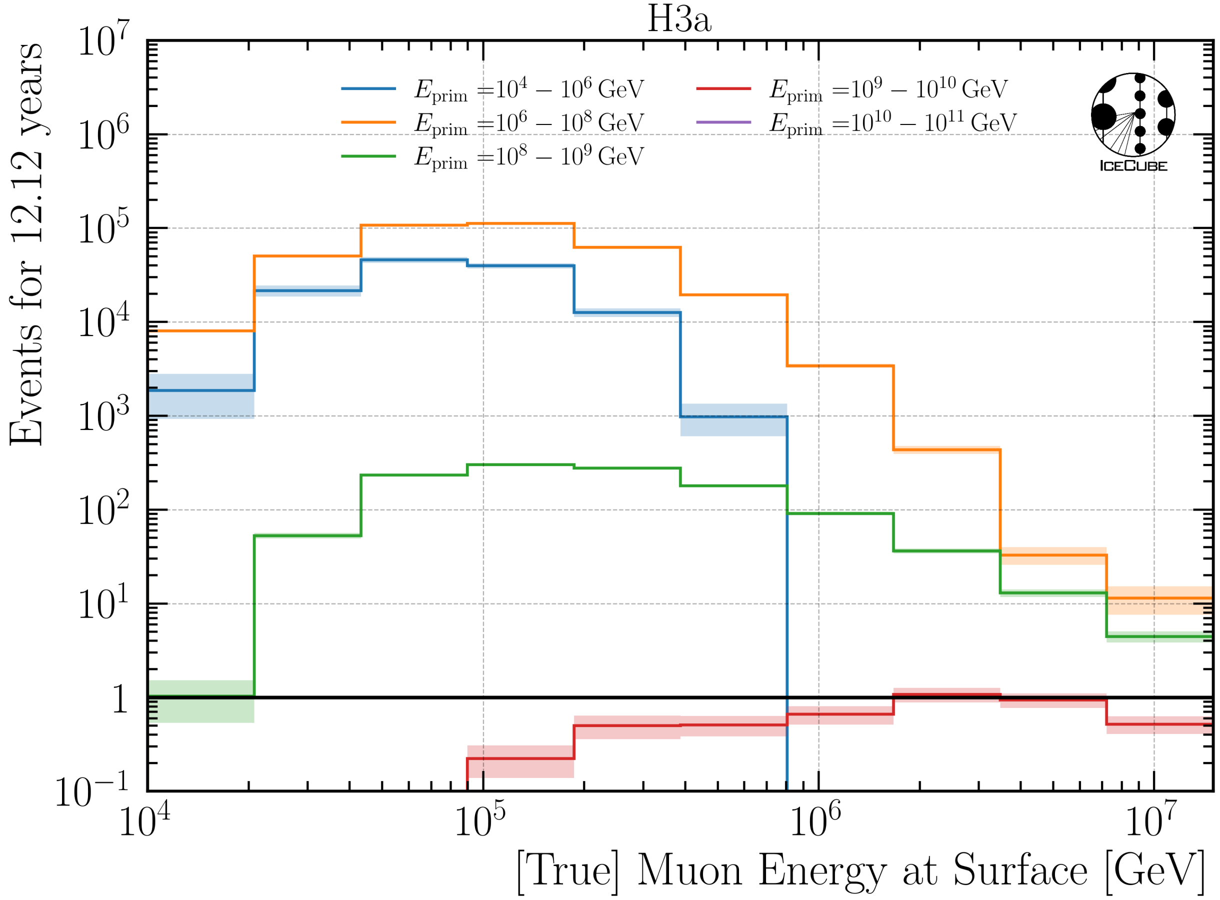
# Final Level Statistics



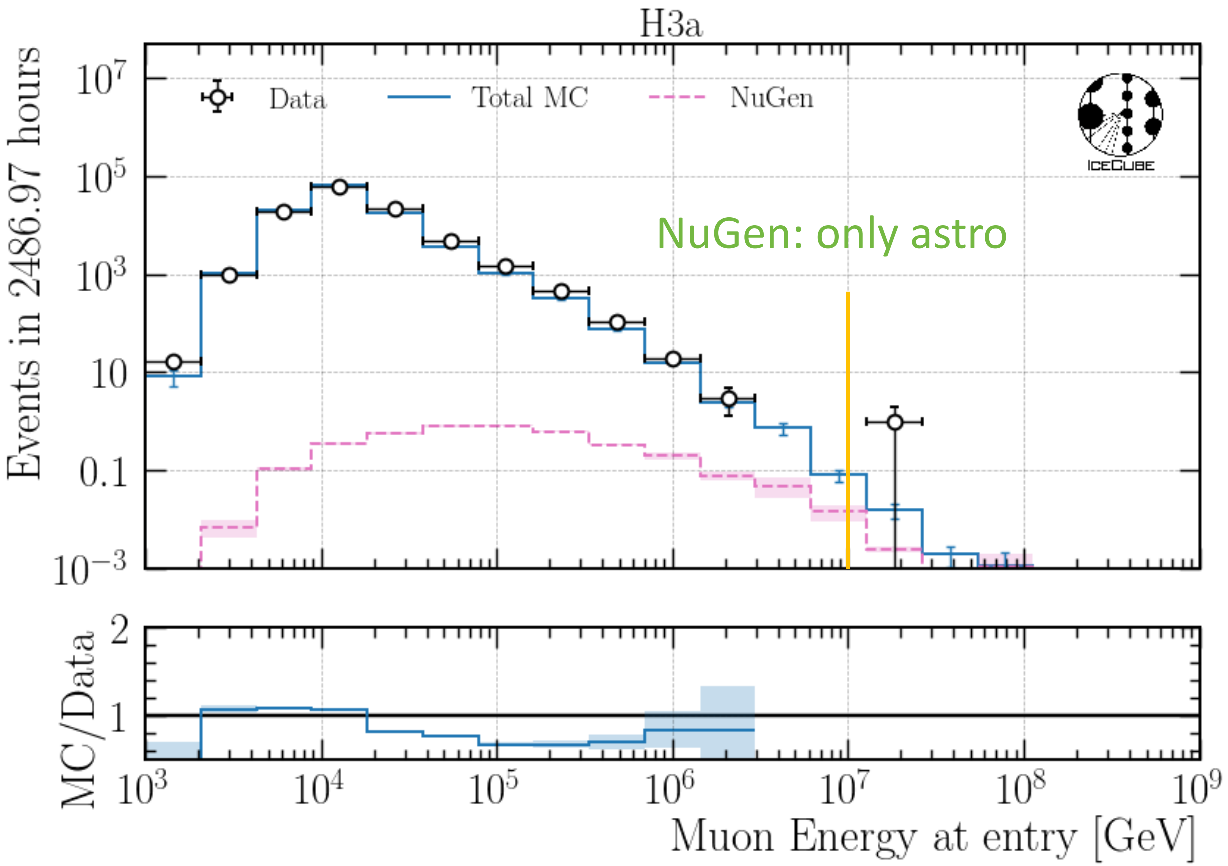
# Final Level Statistics: Proxy & Target Binning

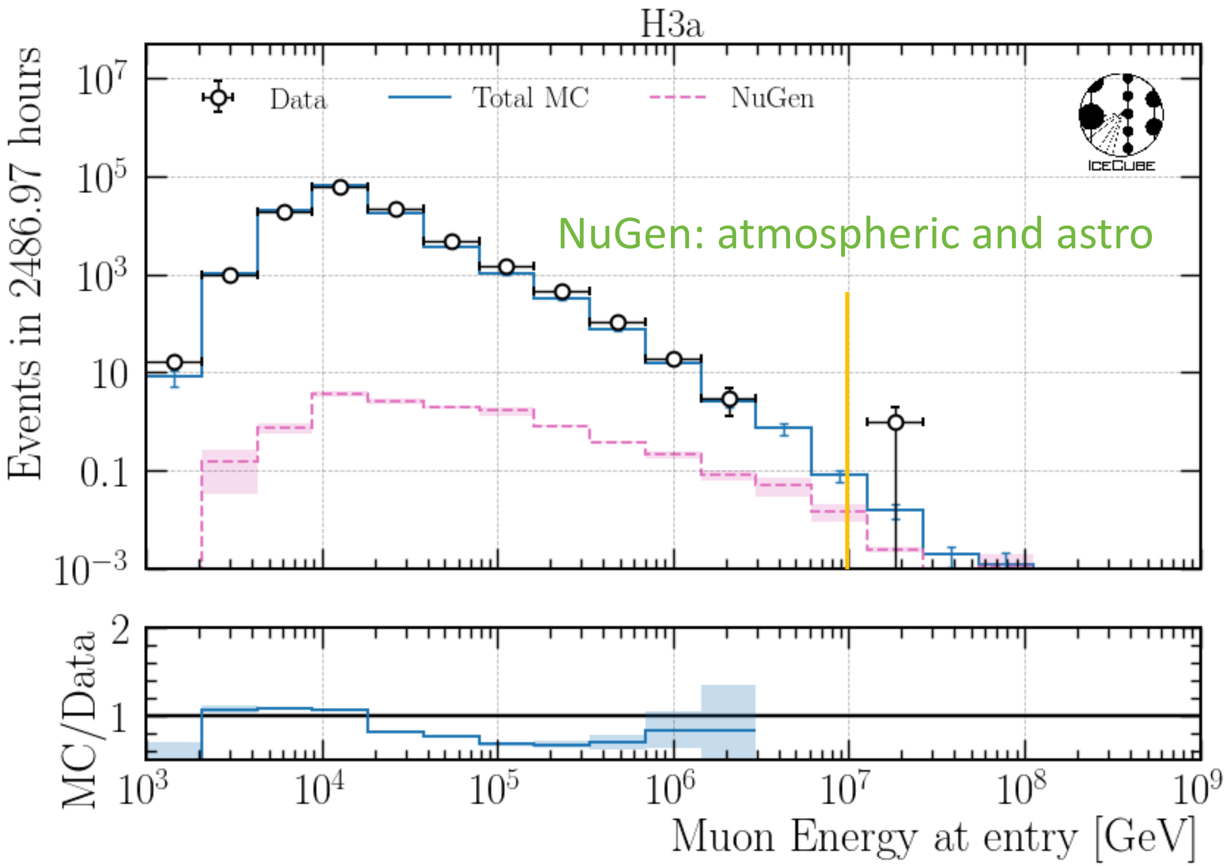


# Impact of Astrophysical Neutrinos – Muon Energy at Entry

$n = 1.5, \gamma = 2.6$




- Neutrinos contribute ~10% to total flux at high energies > 5 PeV

# Assume Latest Astrophysical Diffuse Results by IceCube

## Measurement of the astrophysical diffuse neutrino flux in a combined fit of IceCube's high energy neutrino data

**The IceCube Collaboration**
(a complete list of authors can be found at the end of the proceedings)

*E-mail:* rnaab@icecube.wisc.edu, erik.ganster@icecube.wisc.edu, zelong.zhang@icecube.wisc.edu

The IceCube Neutrino Observatory has discovered a diffuse neutrino flux of astrophysical origin and measures its properties in various detection channels. With more than 10 years of data, we use multiple data samples from different detection channels for a combined fit of the diffuse astrophysical neutrino spectrum. This leverages the complementary information of different neutrino event signatures. For the first time, we use a coherent modelling of the signal and background, as well as the detector response and corresponding systematic uncertainties. The detector response is continuously varied during the simulation in order to generate a general purpose Monte Carlo set, which is central to our approach. We present a combined fit yielding a measurement of the diffuse astrophysical neutrino flux properties with unprecedented precision.

ArXiv: 2308.00191

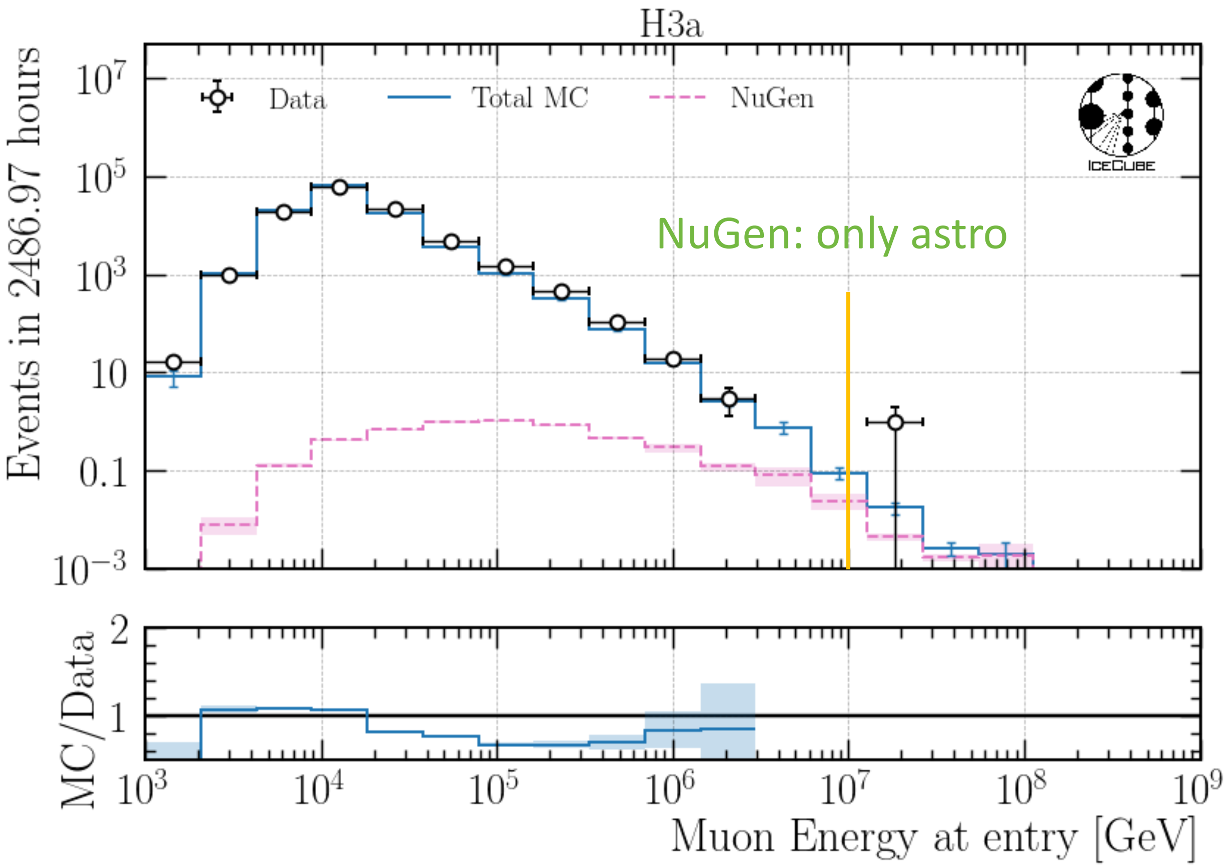
| Astrophysical Model | Result | Energy Range (90% CL) | $-2\Delta\log\mathcal{L}$ over SPL |
|---|---|---|---|
| SPL | $\Phi^{\nu+\bar{\nu}}_{@100\mathrm{TeV}} / C = 1.80^{+0.13}_{-0.16}$<br>$\gamma = 2.52^{+0.04}_{-0.04}$ | 2.5 TeV to 6.3 PeV | - |
| LogP | $\Phi^{\nu+\bar{\nu}}_{@100\mathrm{TeV}} / C = 2.13^{+0.16}_{-0.19}$<br>$\alpha_{LP} = 2.57^{+0.06}_{-0.05}$<br>$\beta_{LP} = 0.23^{+0.10}_{-0.07}$ | 8.0 TeV to 2.2 PeV | 16.4 |
| BPL | $\Phi^{\nu+\bar{\nu}}_{@100\mathrm{TeV}} / C = 1.77^{+0.15}_{-0.11}$<br>$\log_{10}(E_{\mathrm{break}}/\mathrm{GeV}) = 4.39^{+0.09}_{-0.08}$<br>$\gamma_1 = 1.31^{+0.50}_{-1.21}$<br>$\gamma_2 = 2.74^{+0.06}_{-0.07}$ | 13.7 TeV to 4.7 PeV | 24.7 |

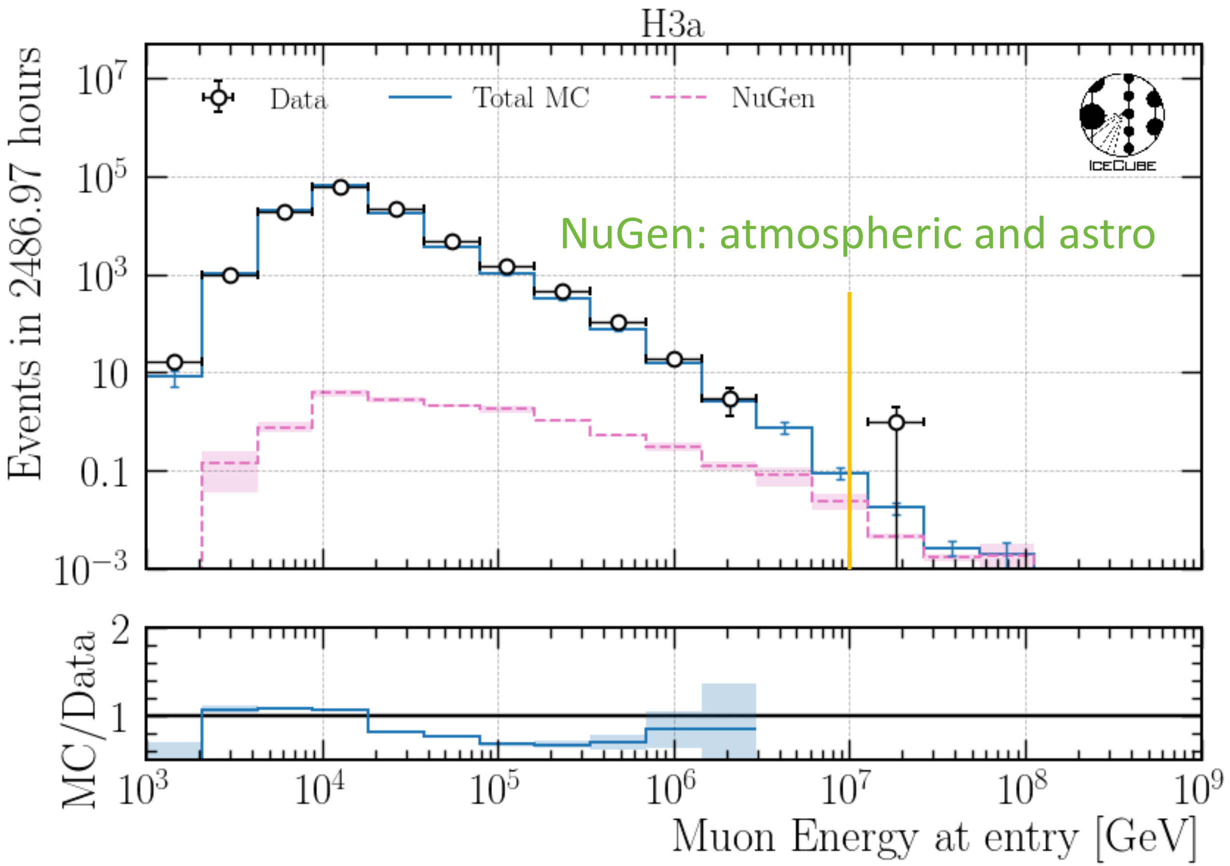
➢ Test SPL and BPL

# Impact of Astrophysical Neutrinos – Muon Energy at Entry

$n = 1.8, \gamma = 2.52$

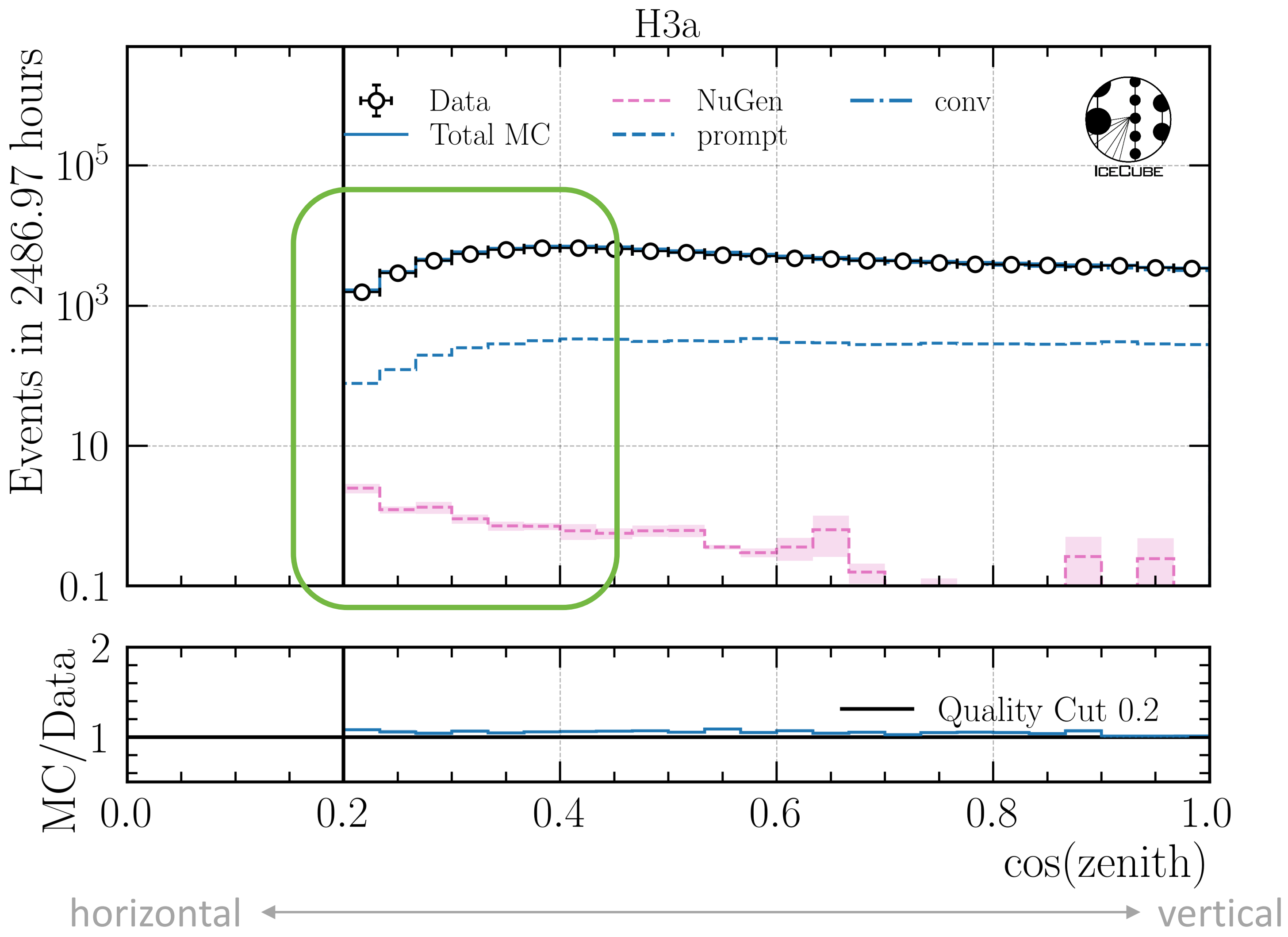
2308.00191



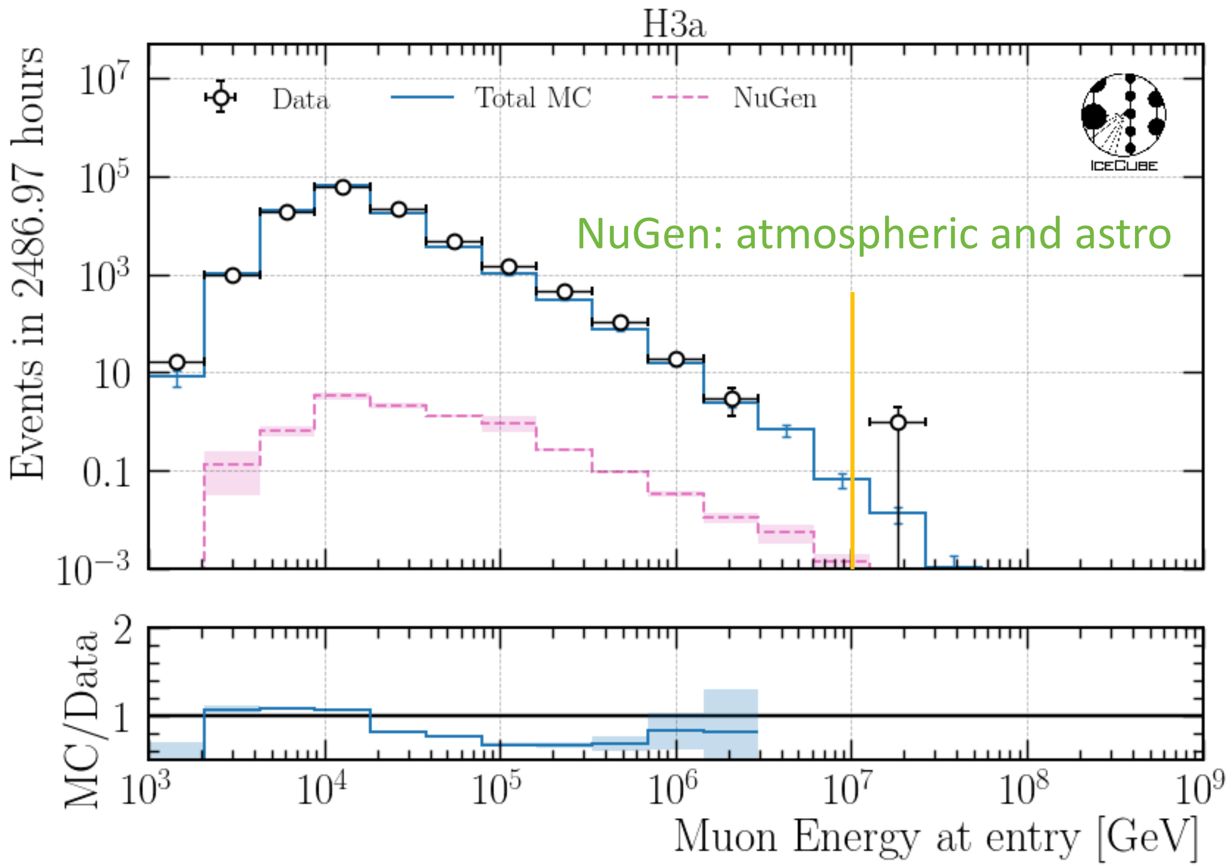

- Neutrinos contribute ~10% to total flux at high energies > 5 PeV

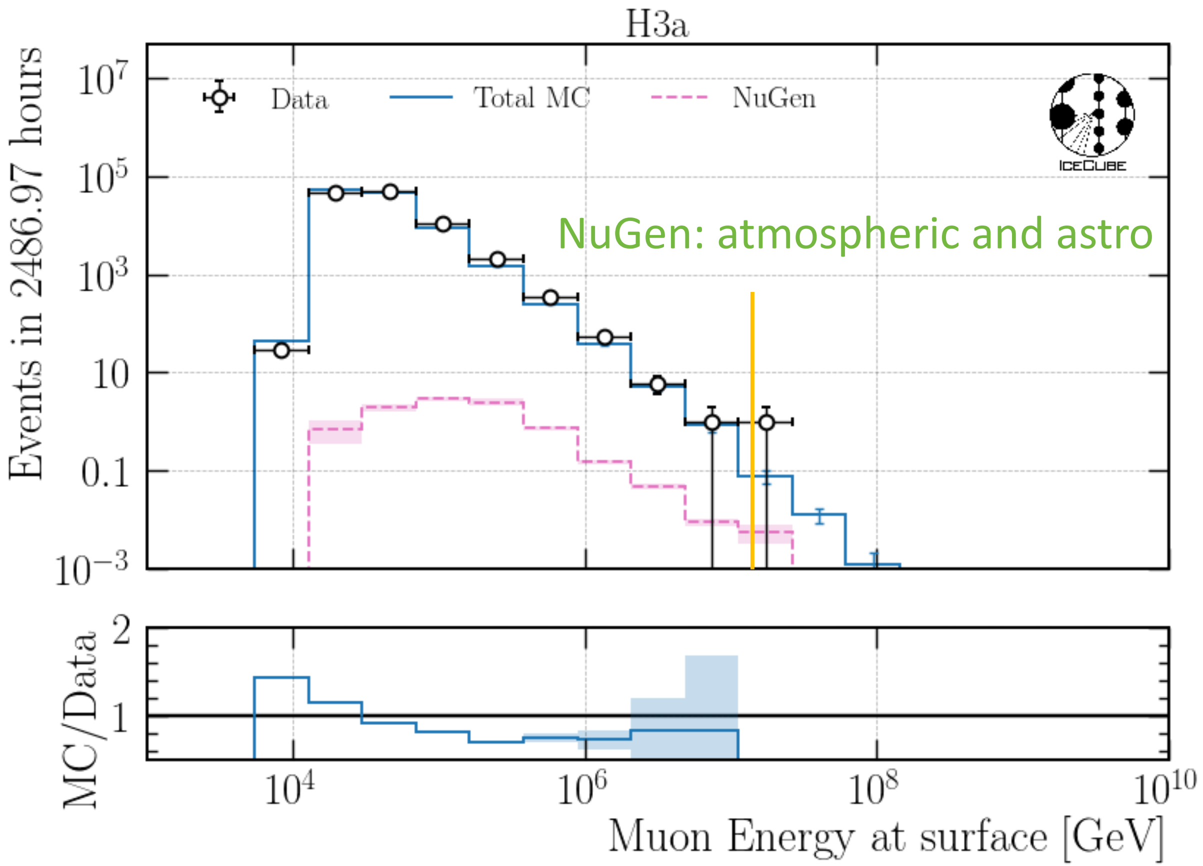
# Cos(zenith) Distribution




- Expect more neutrinos from the horizon
- ➢ Apply stronger zenith cut to remove neutrinos

# Impact of Astrophysical Neutrinos – Muon Energy at Surface/Entry




$$n = 1.77, \gamma_1 = 1.31, \gamma_2 = 2.74$$
$$\log10(E_{\text{break}}) = 4.39,$$

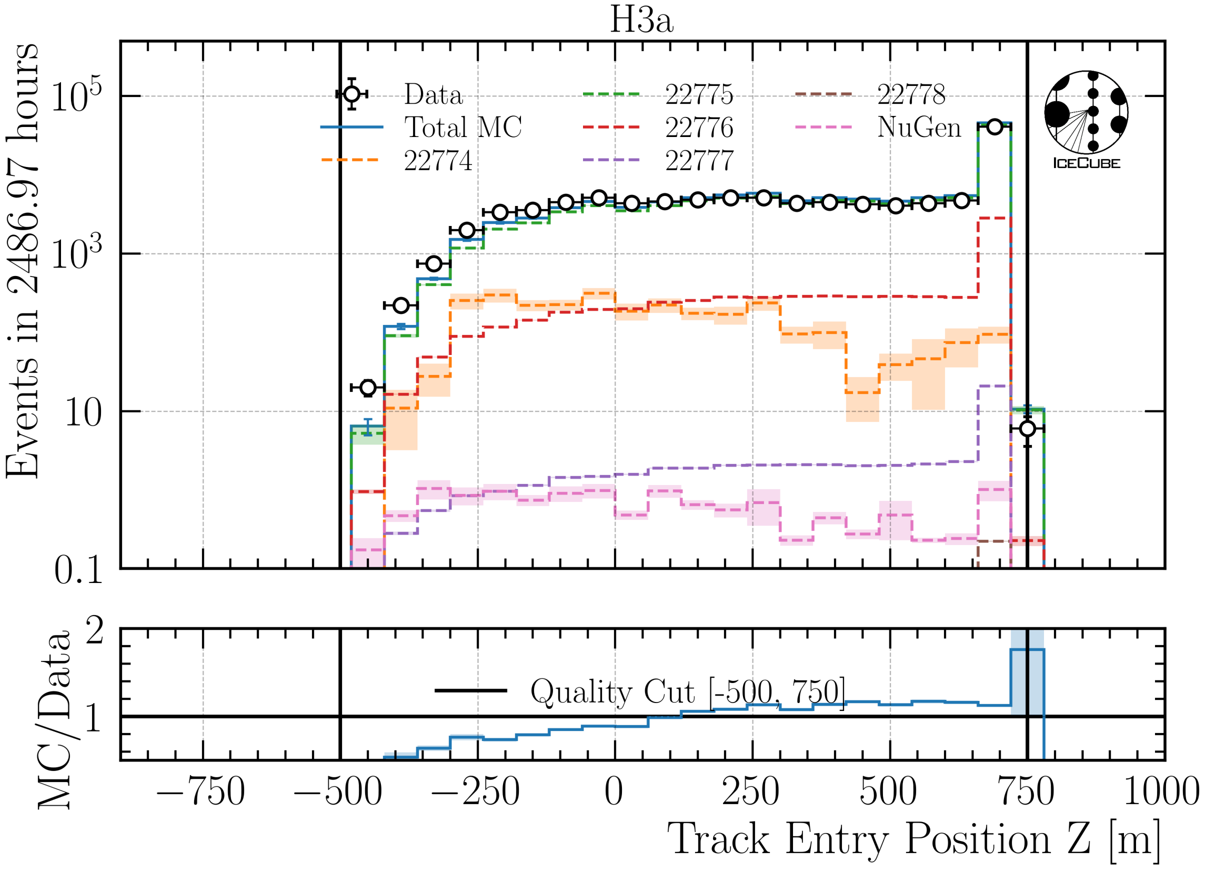
2308.00191

- For BPL: Neutrinos contribute less than 10%

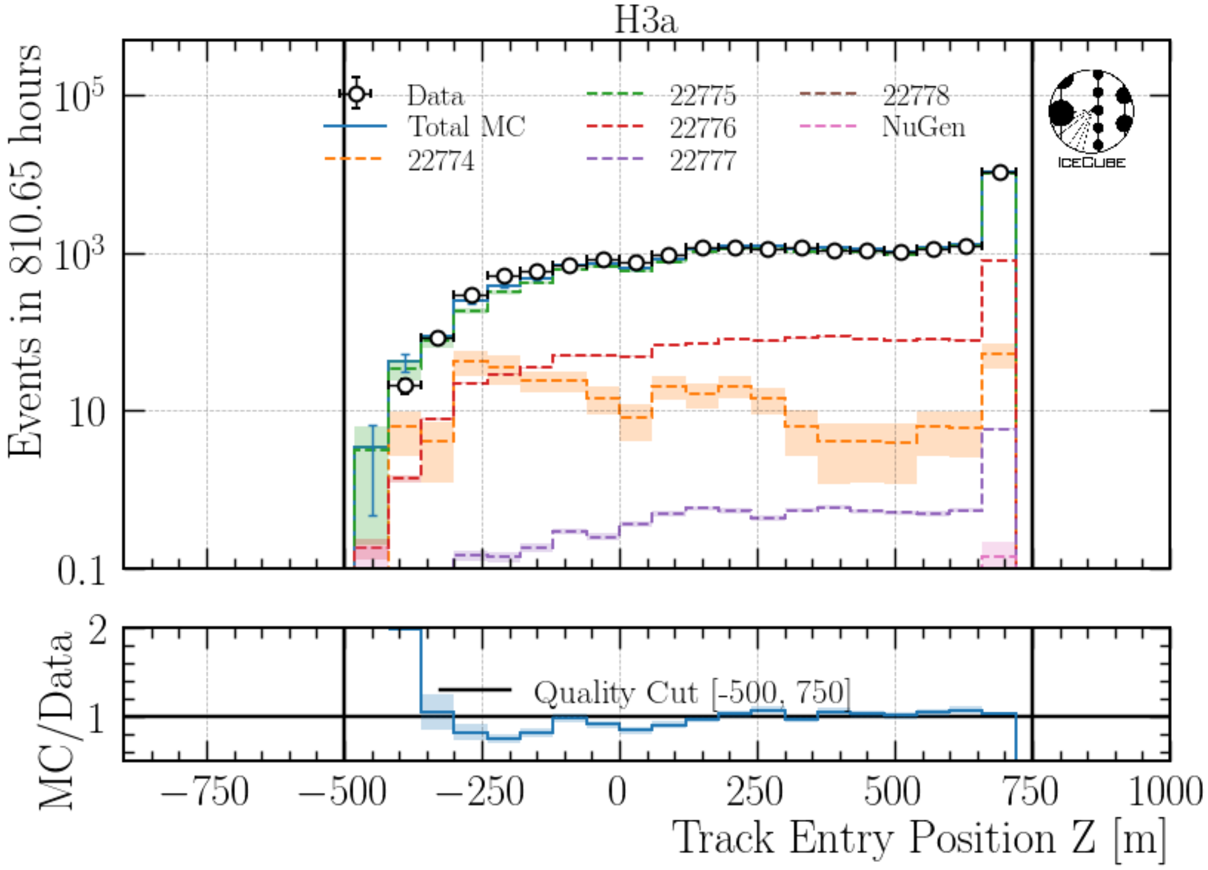
# Z—Vertex Data-MC

- Networks were trained on several old CORSIKA simulation datasets to make them robust
- Re-train on new ice model spice_ftp-v3m

# Z—Vertex Data-MC: Entry Position

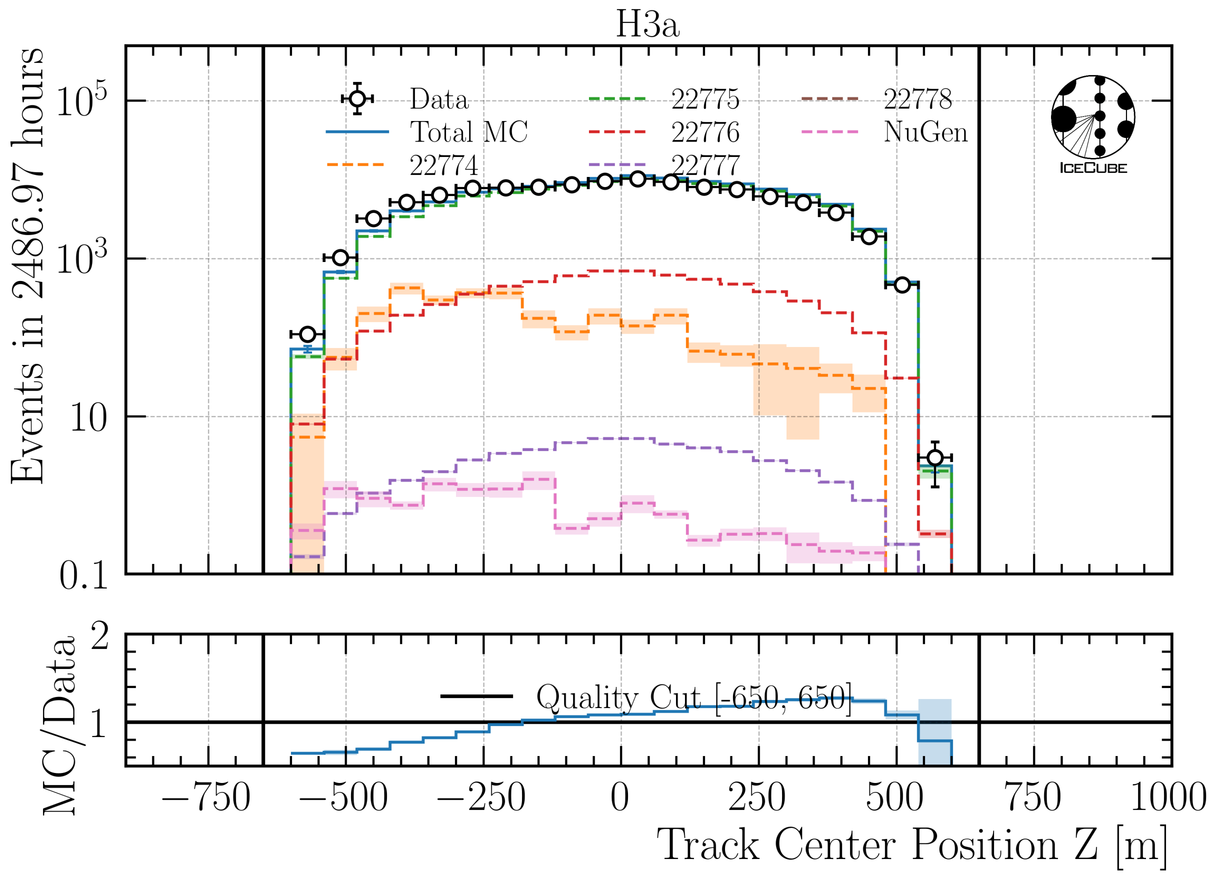



Old ice model

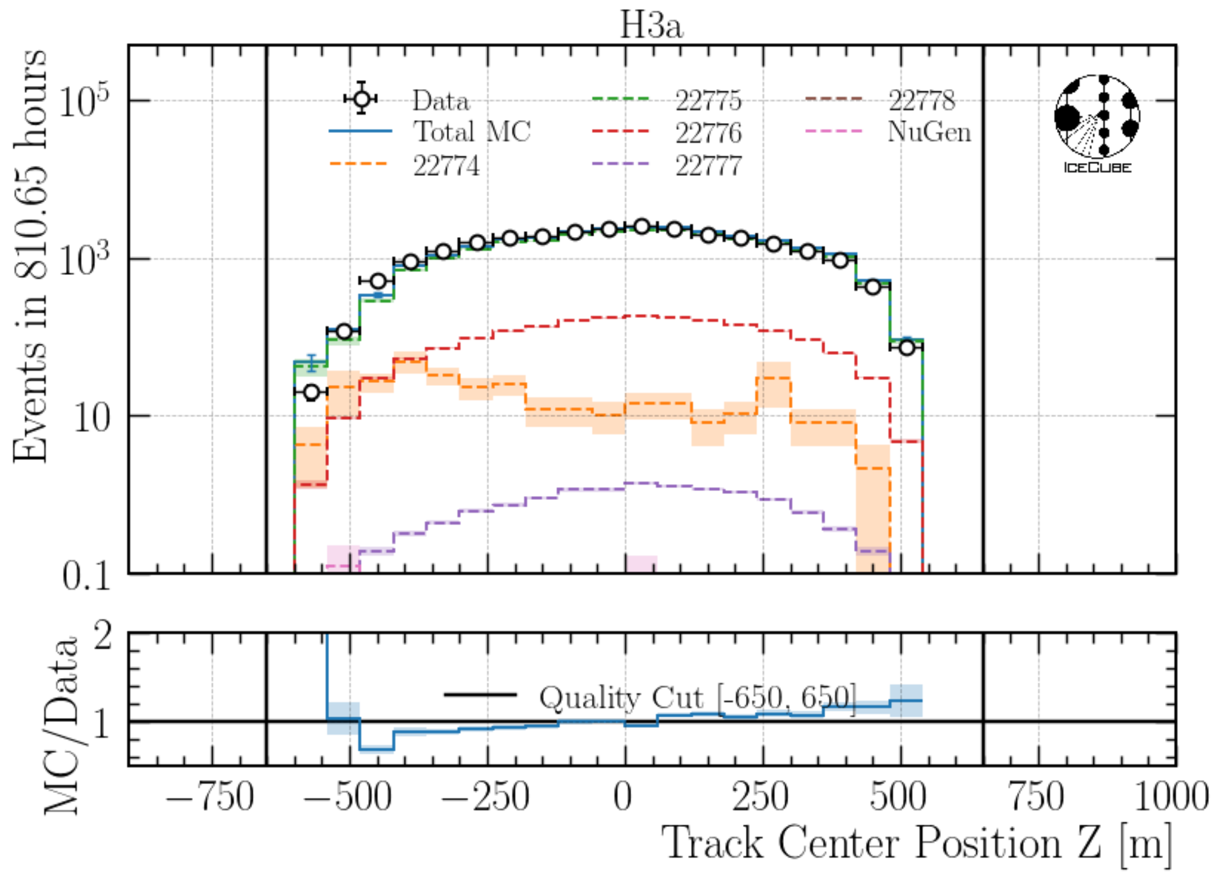



New ice model

# Z—Vertex Data-MC: Center Position

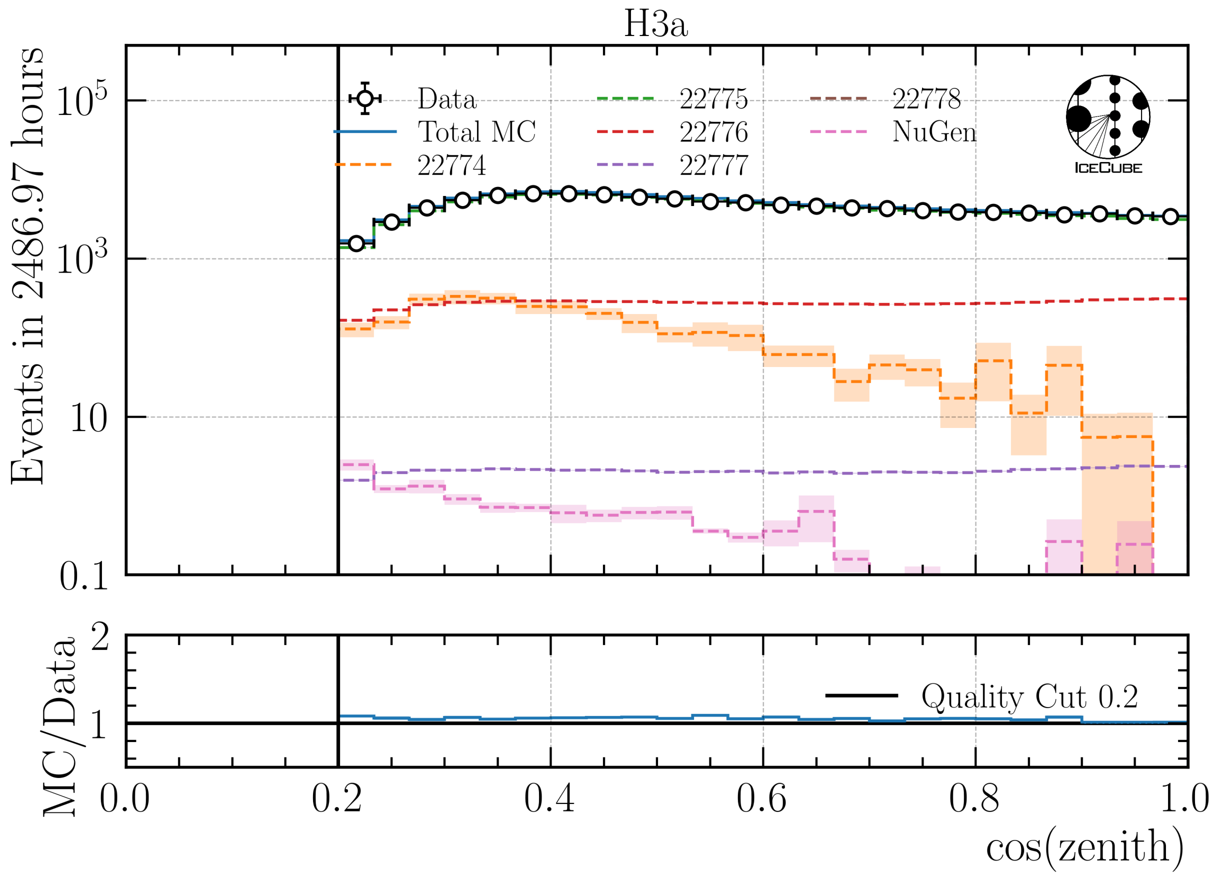



Old ice model

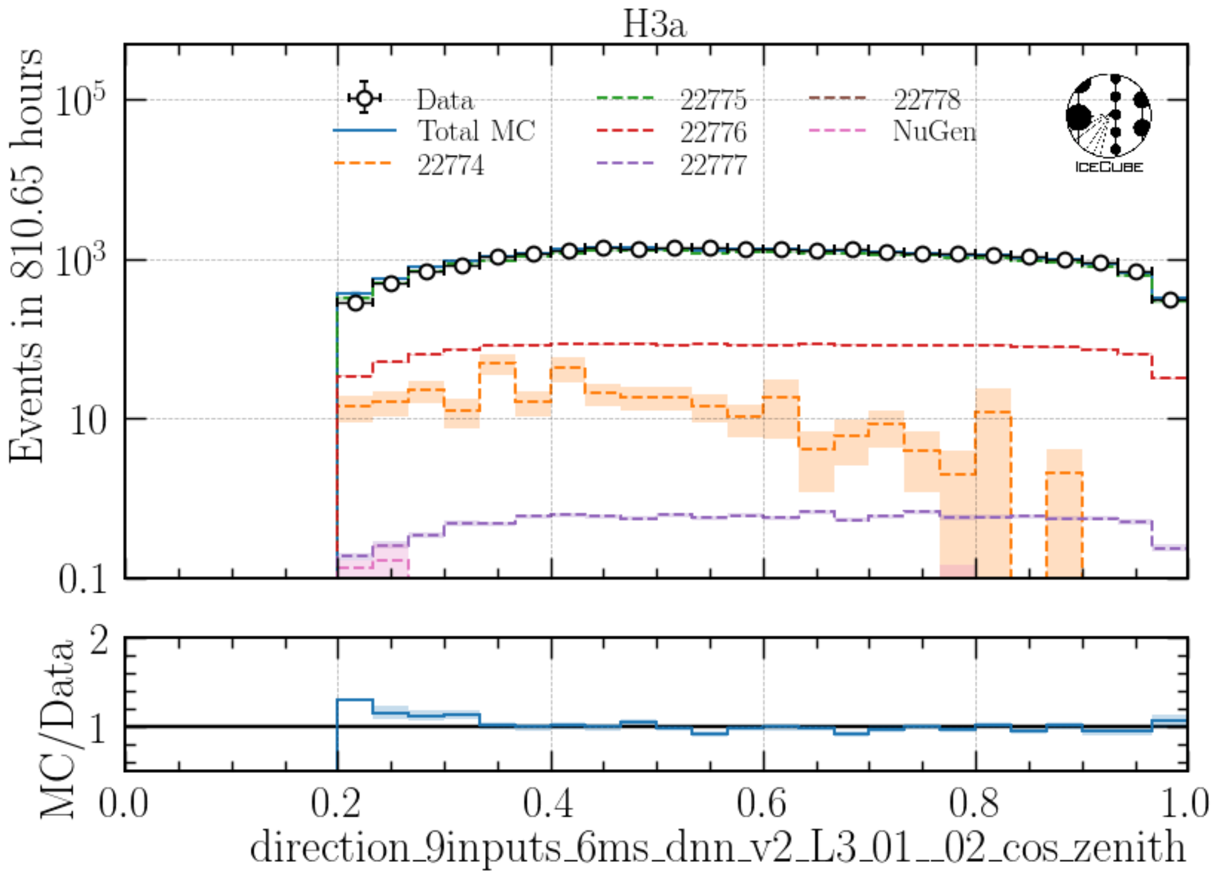



New ice model

# Cos(zenith)




Old ice model



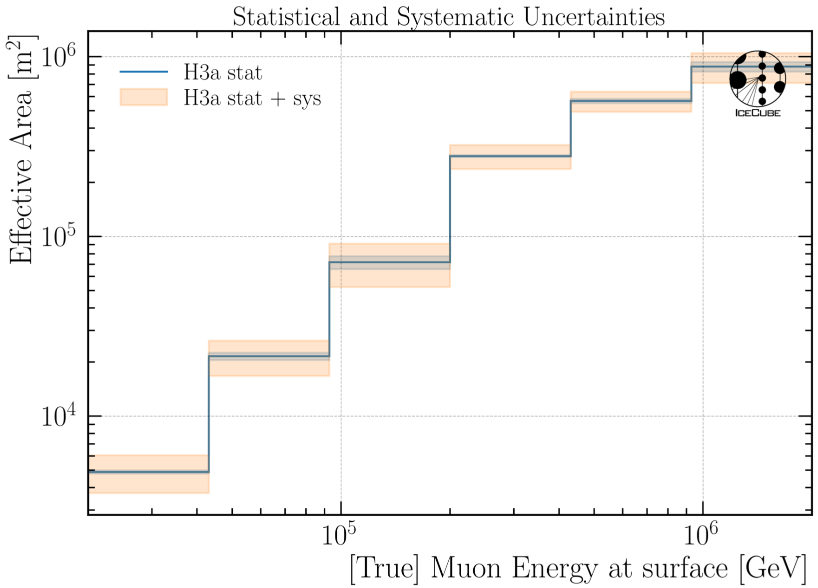

New ice model

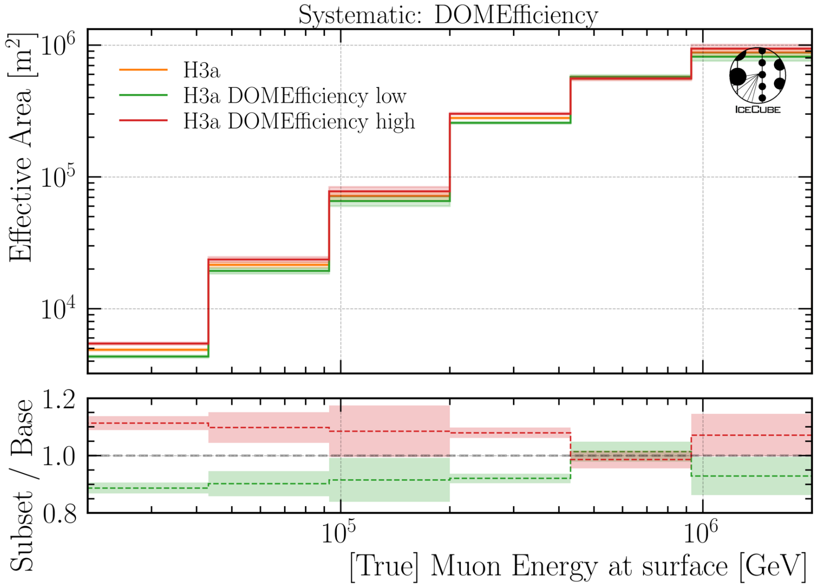
- Artefacts observed in zenith reconstruction
  - ➢ Re-train zenith reco network

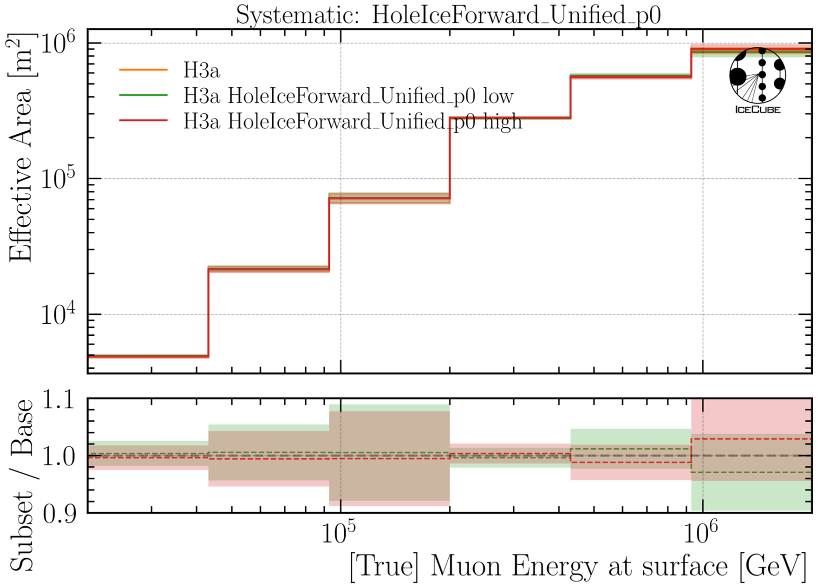
# Include Systematic Uncertainties on Effective Area

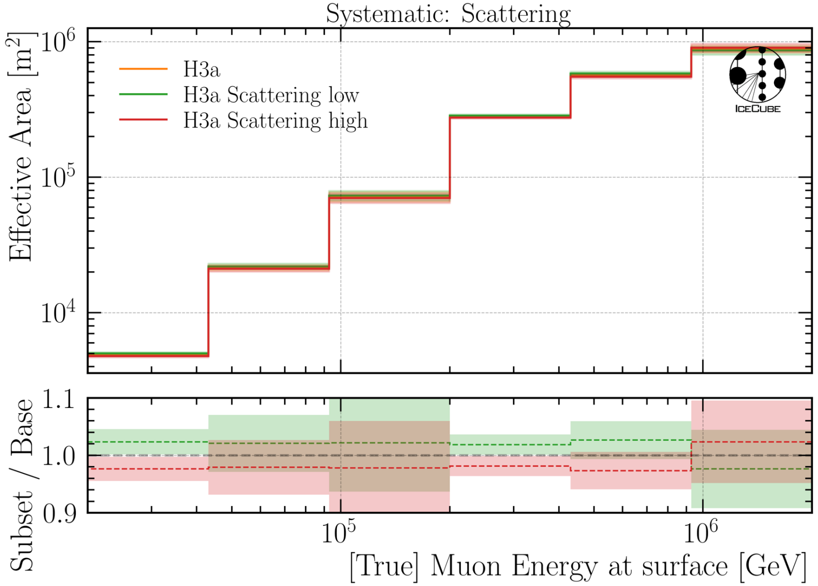
# Effective Area

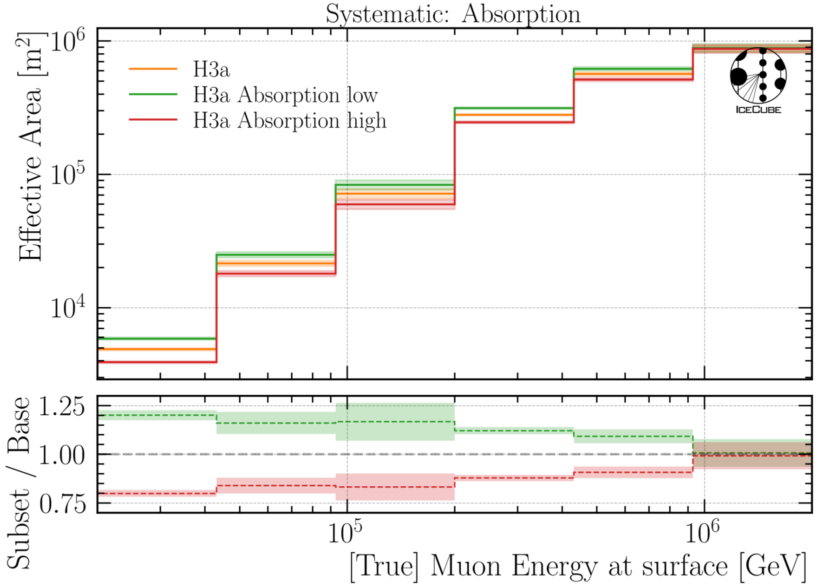
- Baseline: entire set
- Subset: above/below median

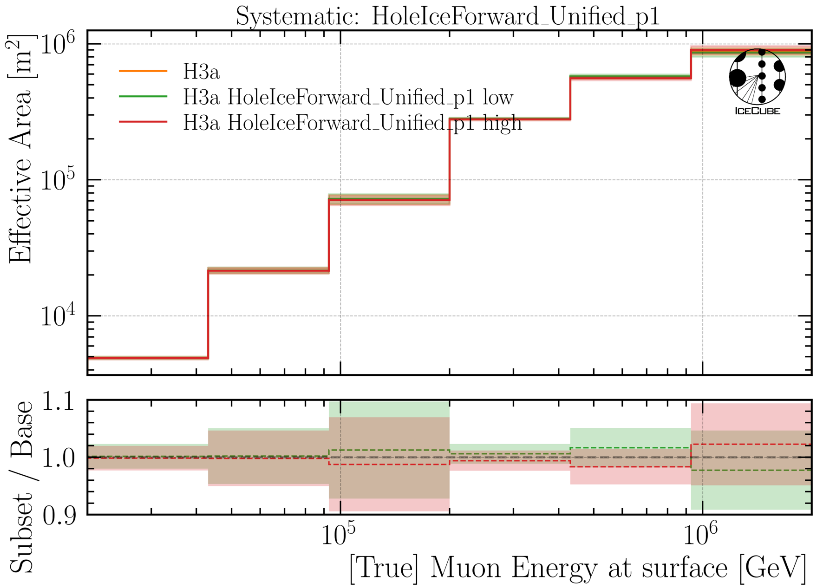


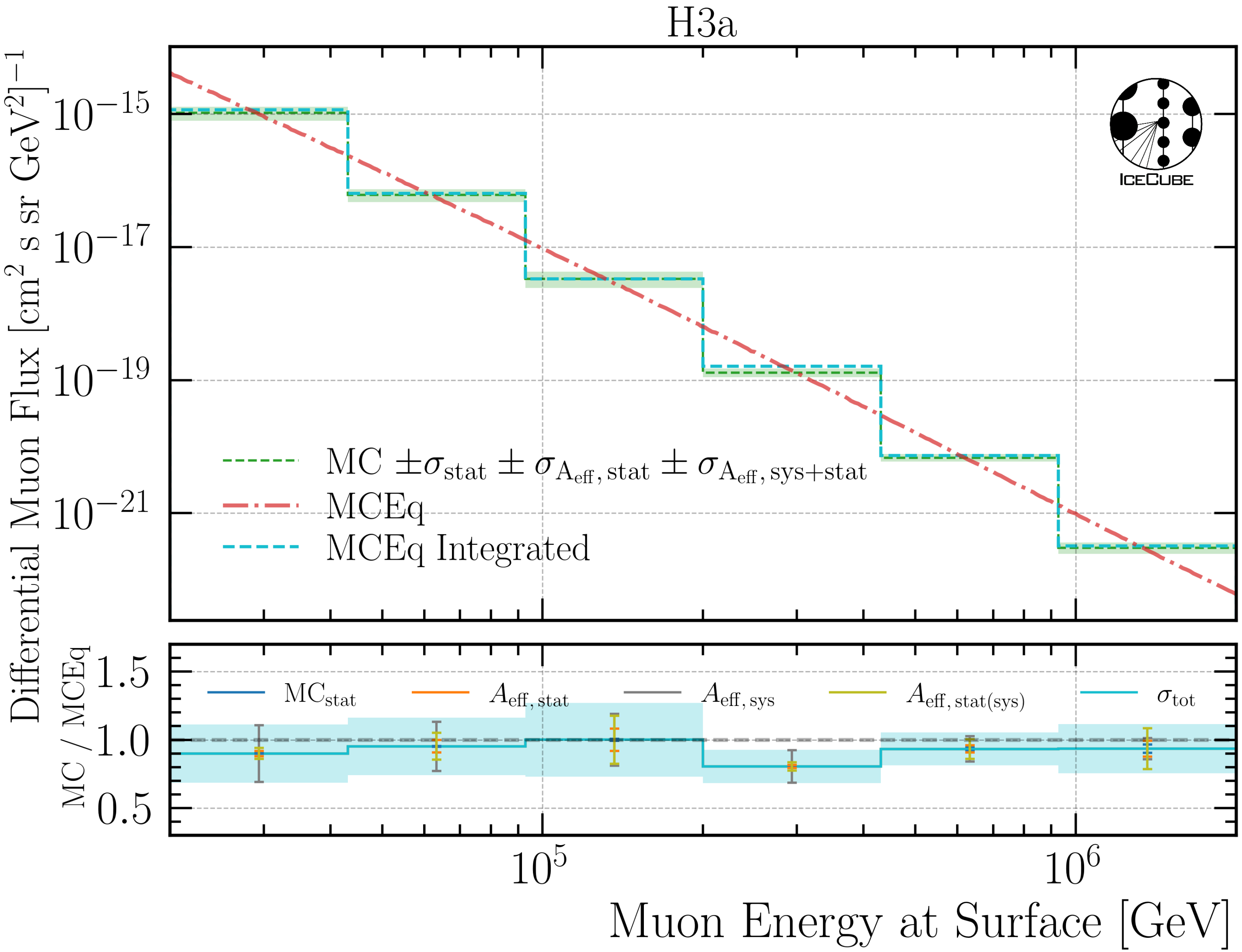
# MCEq vs CORSIKA

Uncertainties on Flux

- CORSIKA stat.
- Eff. area stat.
- Eff. area sys.
- Eff. area stat. on sys. subset

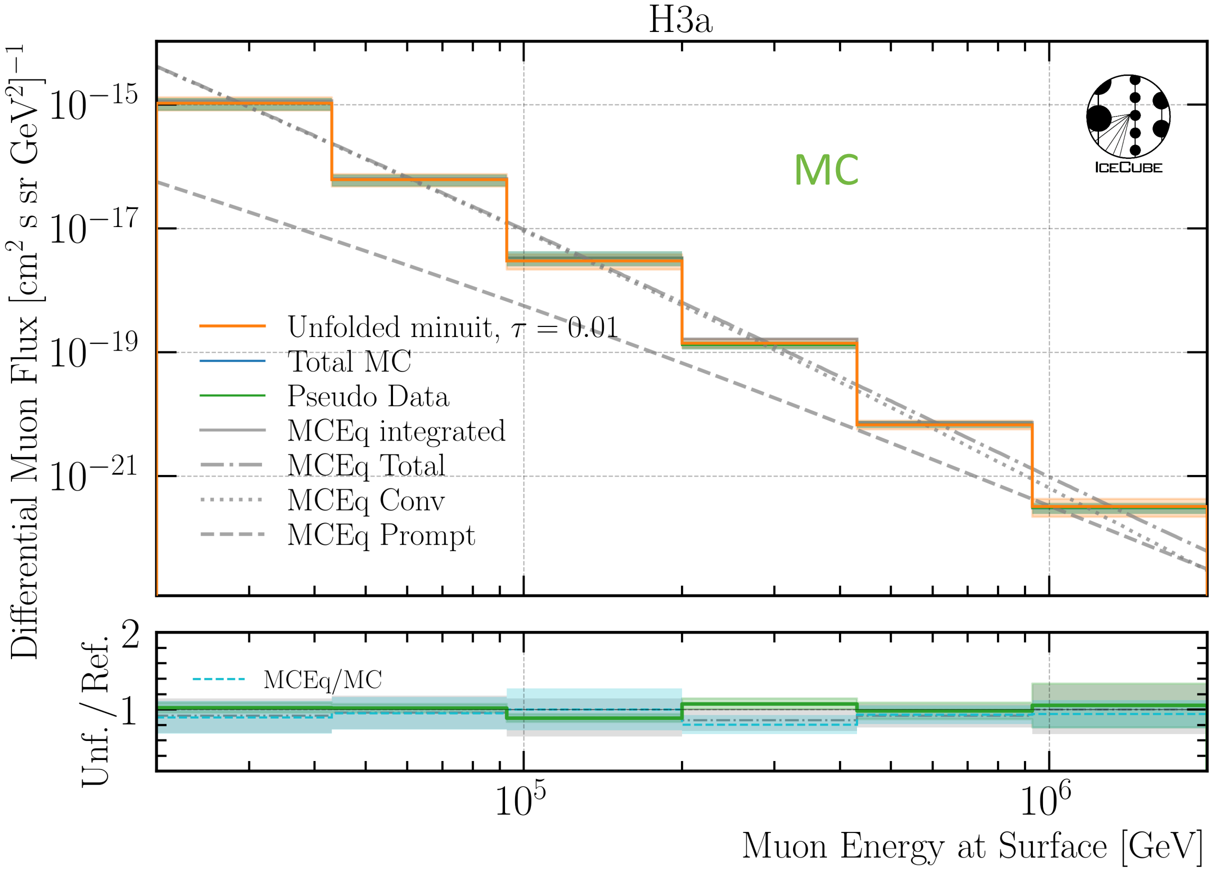
$\chi^2$– Test (H3a):
p—value: 0.683

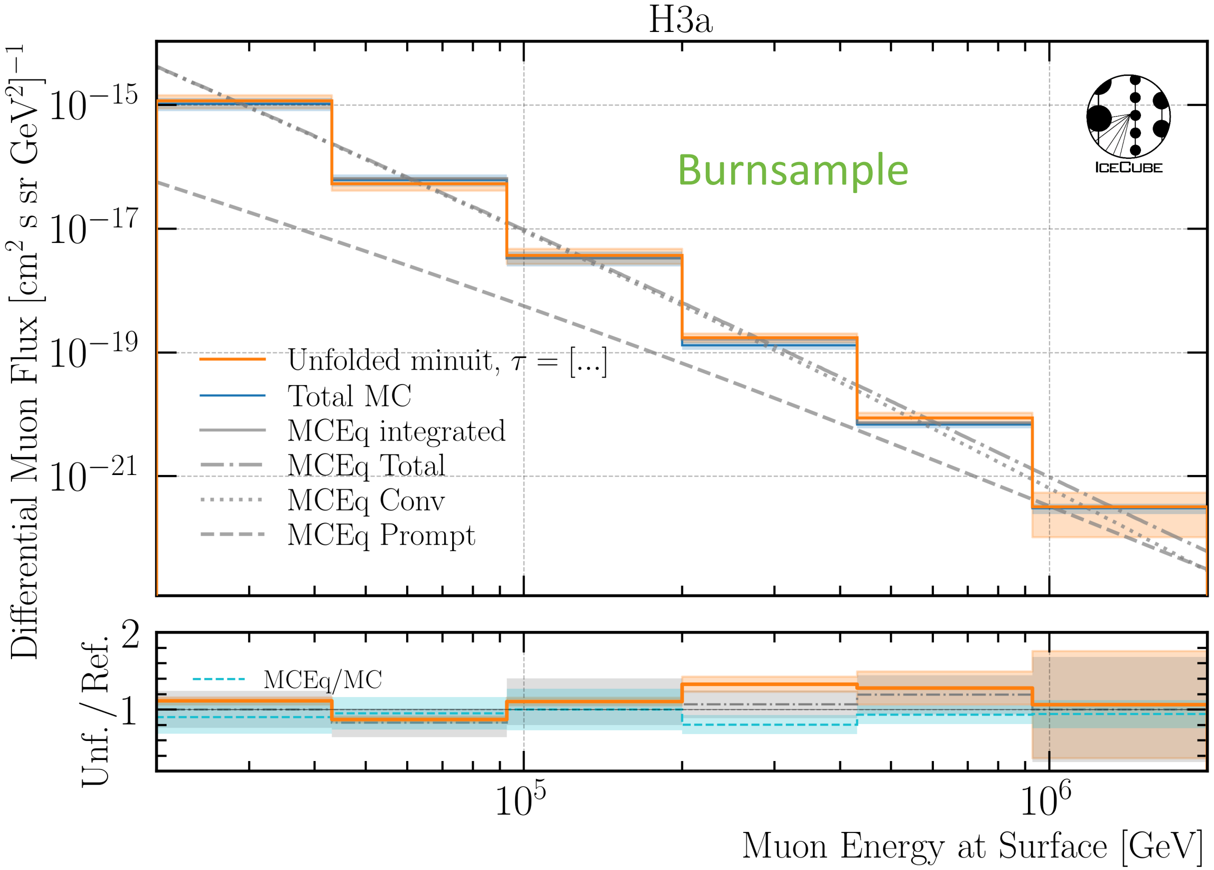
➢ Good agreement between MCEq & CORSIKA



# Unfold Burnsample

# Unfold Burnsample




$\chi^2$– Test (H3a) Conv only:
p—value: 0.587

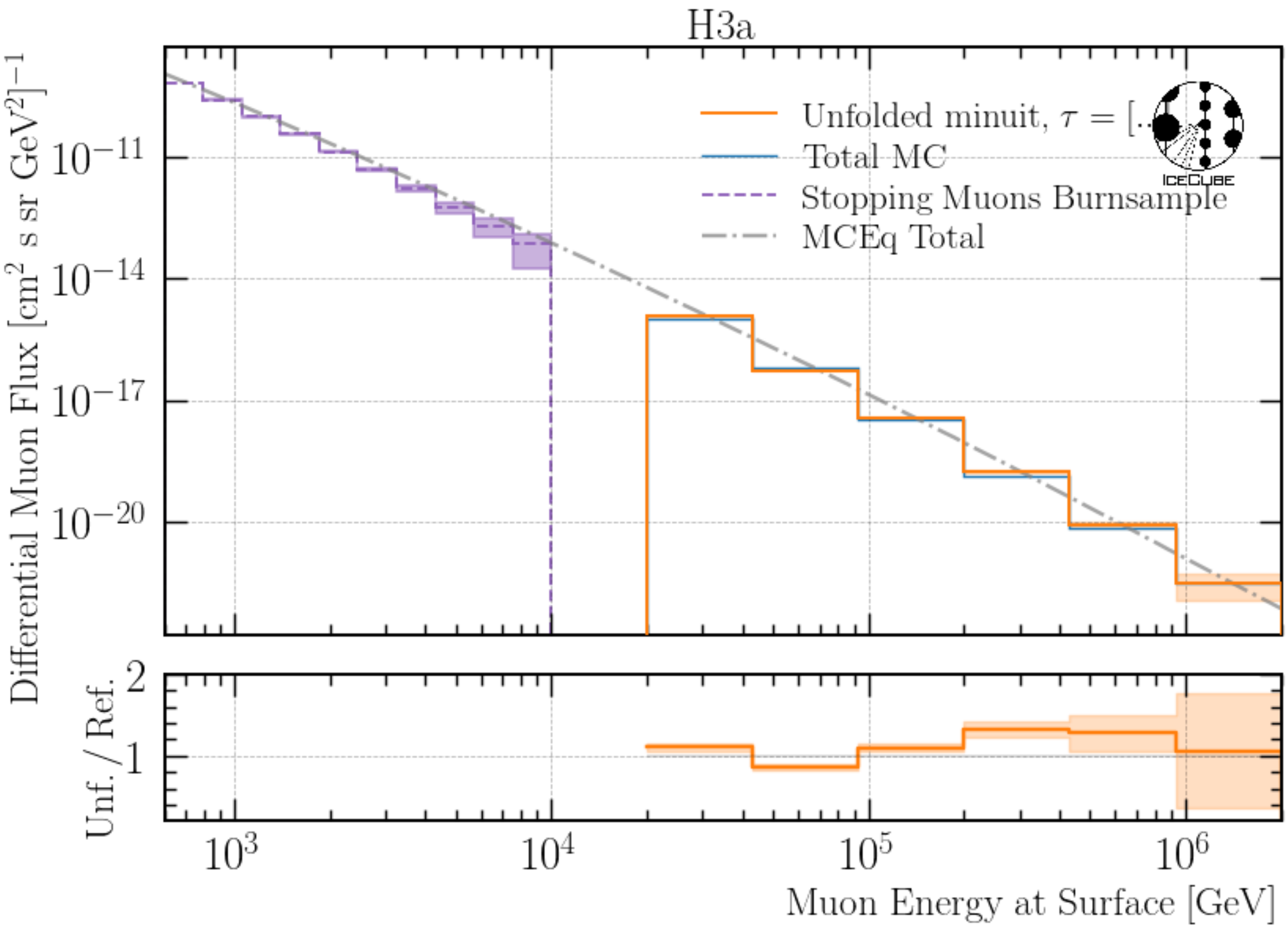
➢ Unfolding works on MC & burnsample

# Stopping Muons and Leading Muons

Work in progress

- Use stopping muons to unfold low energy muon spectrum
- Use leading muons to unfold high energy muon spectrum

➢ Atmospheric muon flux unfolding is promising



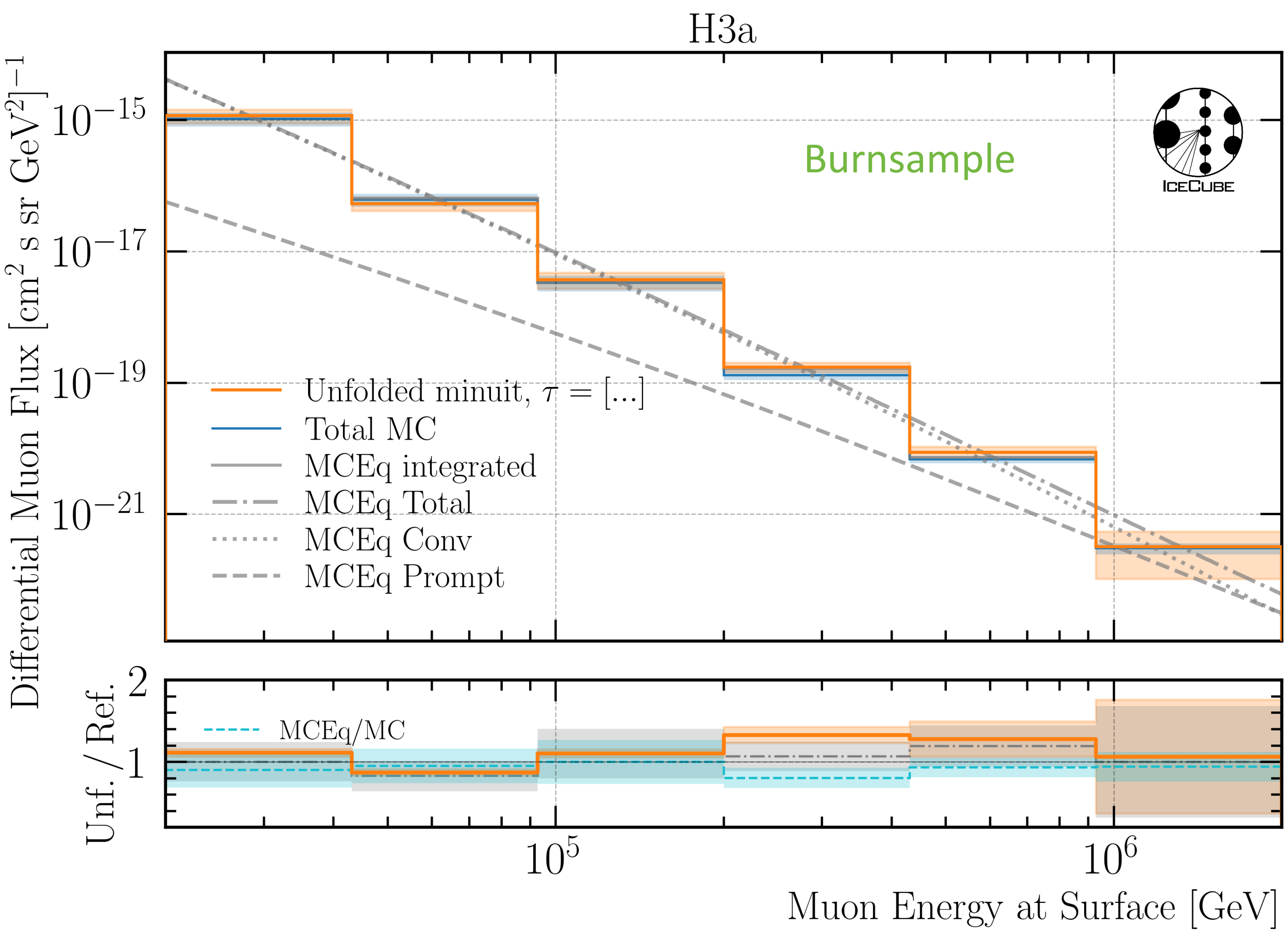
# Conclusion & Outlook

- Neutrino background investigated
  - BPL smaller impact than SPL
- Include uncertainties on effective area
  - statistics & systematics
- Agreement between MCEq & CORSIKA
- Unfolded burnsample up to 2 PeV
- Re-trained DNNs on new ice models
  - z—vertex Data-MC improved

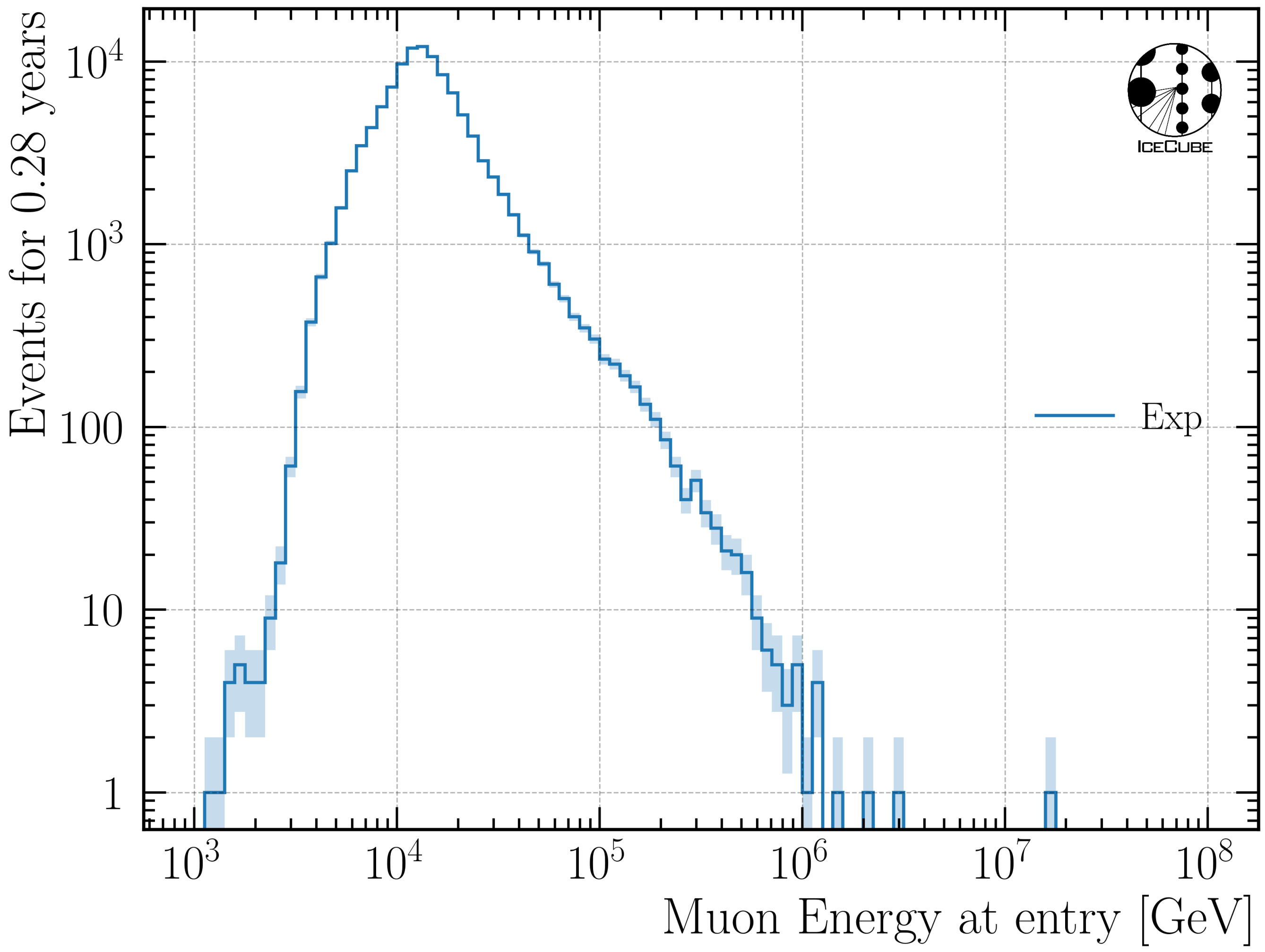
## Outlook

- Unfold for different neutrino background assumptions
- Evaluate new reconstructions in detail
- Unfold on selection with new networks

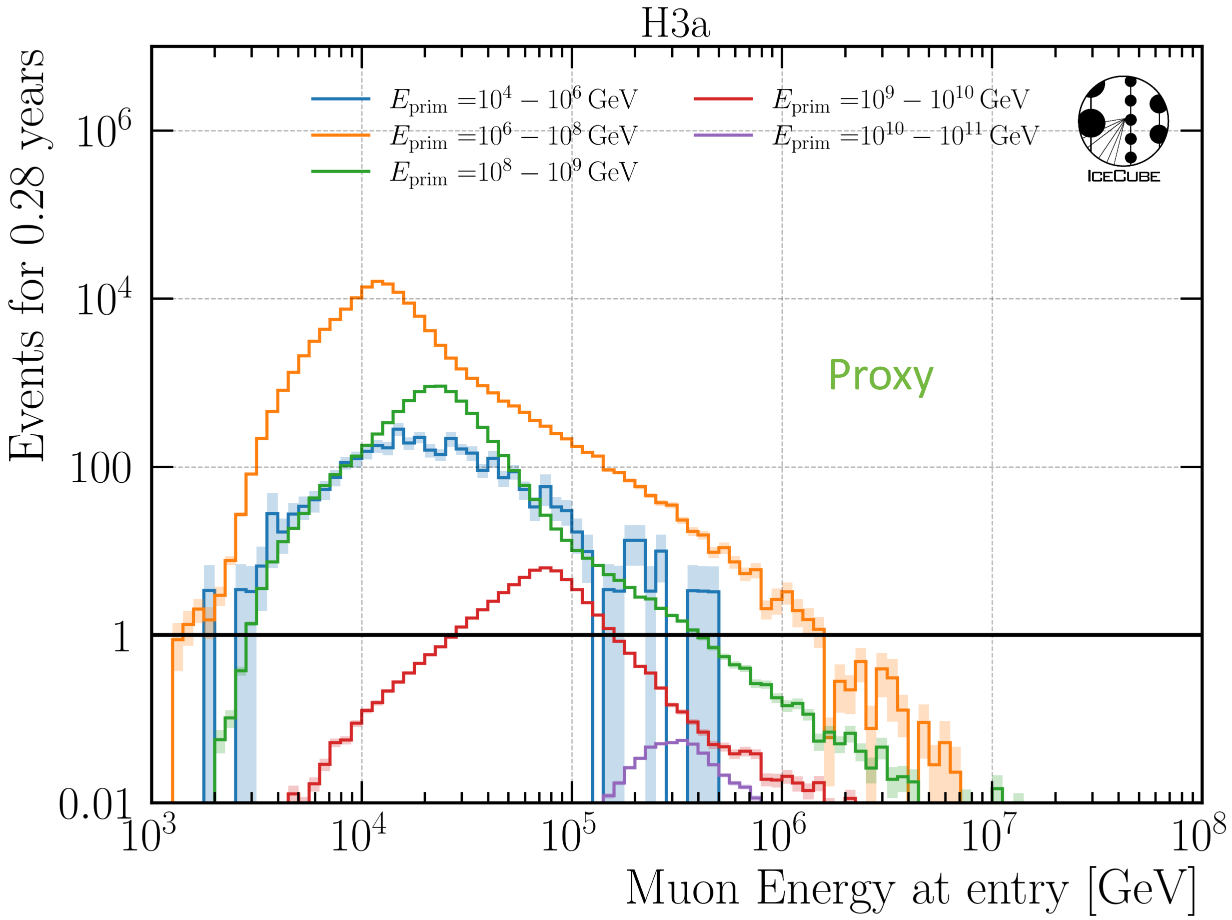
➢ Unfold on 12.12 years of data



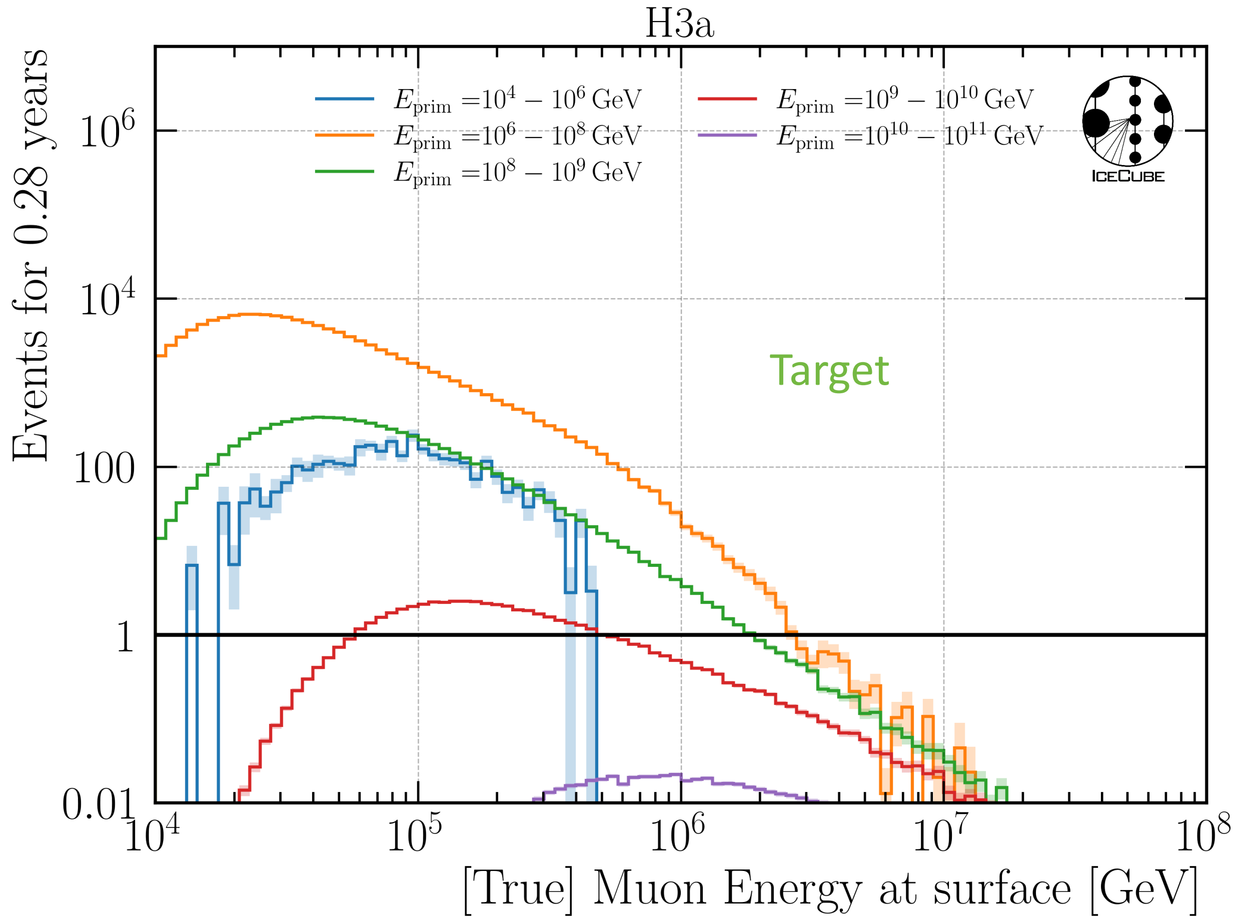
# Burnsample Proxy



# Statistics

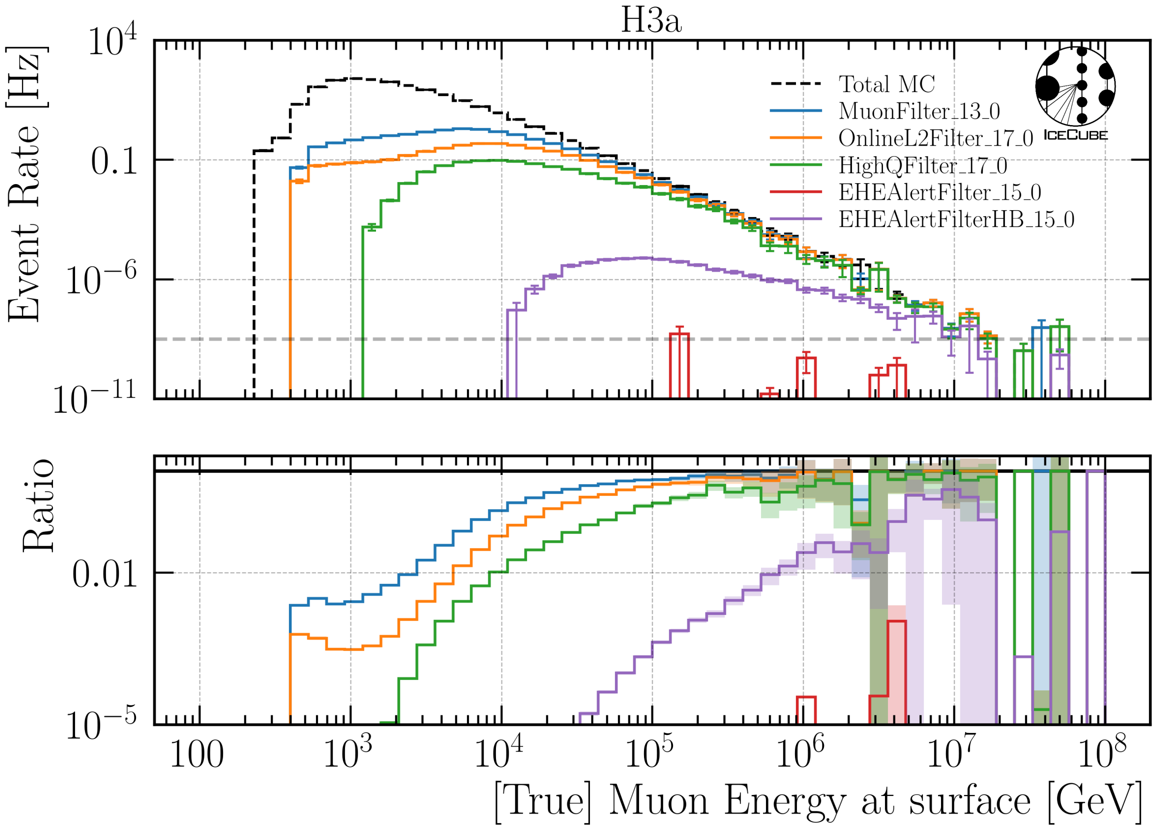


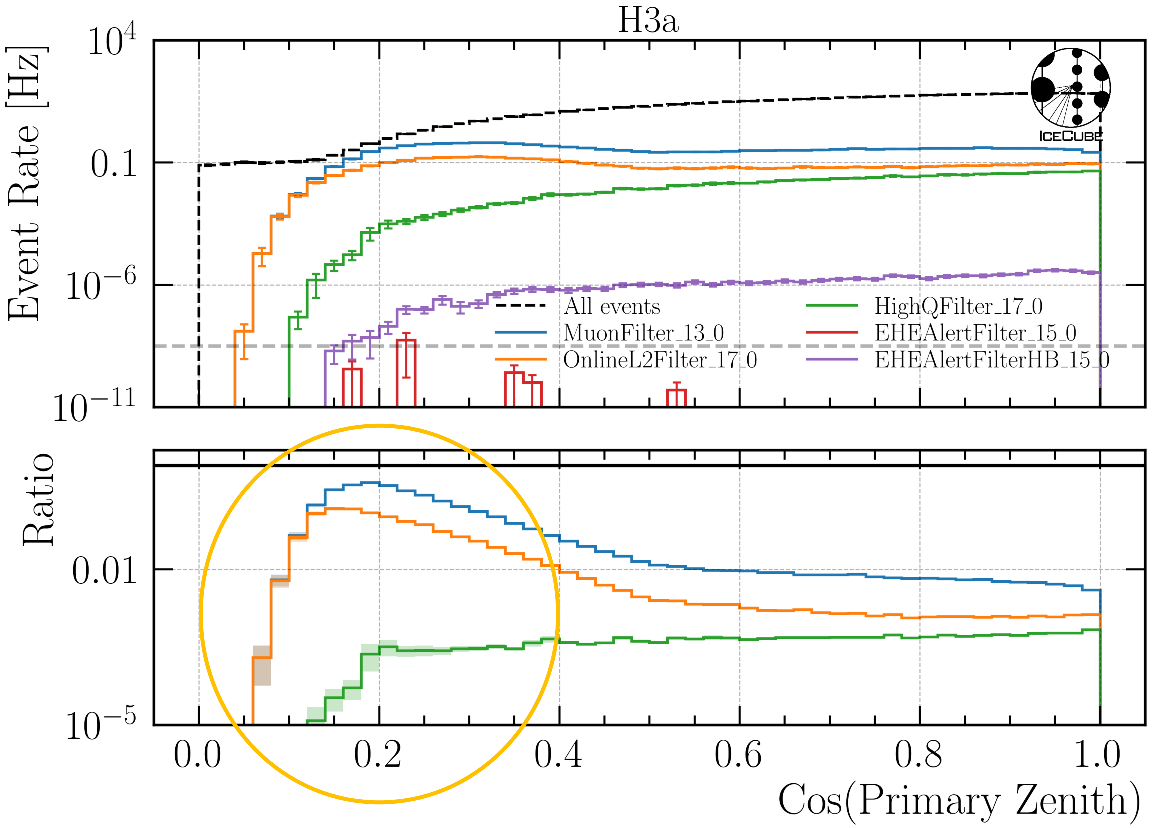
H3a
Events for 0.28 years
Muon Energy at entry [GeV]
$E_{prim} = 10^4 - 10^6$ GeV
$E_{prim} = 10^6 - 10^8$ GeV
$E_{prim} = 10^8 - 10^9$ GeV
$E_{prim} = 10^9 - 10^{10}$ GeV
$E_{prim} = 10^{10} - 10^{11}$ GeV
IceCube
Proxy

H3a
Events for 0.28 years
[True] Muon Energy at surface [GeV]
$E_{prim} = 10^4 - 10^6$ GeV
$E_{prim} = 10^6 - 10^8$ GeV
$E_{prim} = 10^8 - 10^9$ GeV
$E_{prim} = 10^9 - 10^{10}$ GeV
$E_{prim} = 10^{10} - 10^{11}$ GeV
IceCube
Target

# Q: Why do you choose the muon filter?

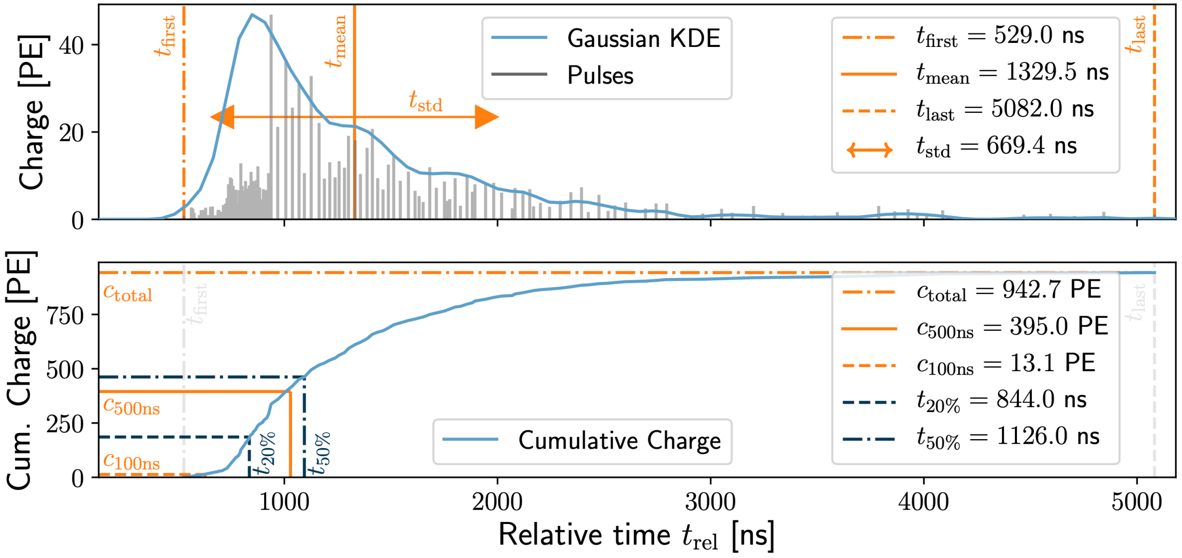



- HighQFilter: removes more horizontal events → muons travel large distances, not enough energy left to pass high-charge filter
- Goal: keep AMAP high-energy events at surface → muon filter

# Q: Which quantities do you reconstruct and how do you do this?

10.1088/1748-0221/16/07/P07041




## Energy

- Bundle energy at entry / surface
- Leading muon energy at entry / surface

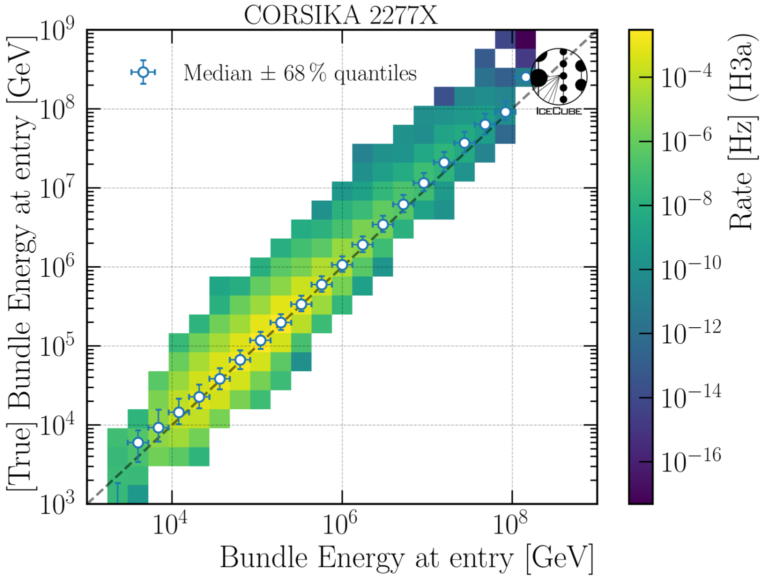
## Direction

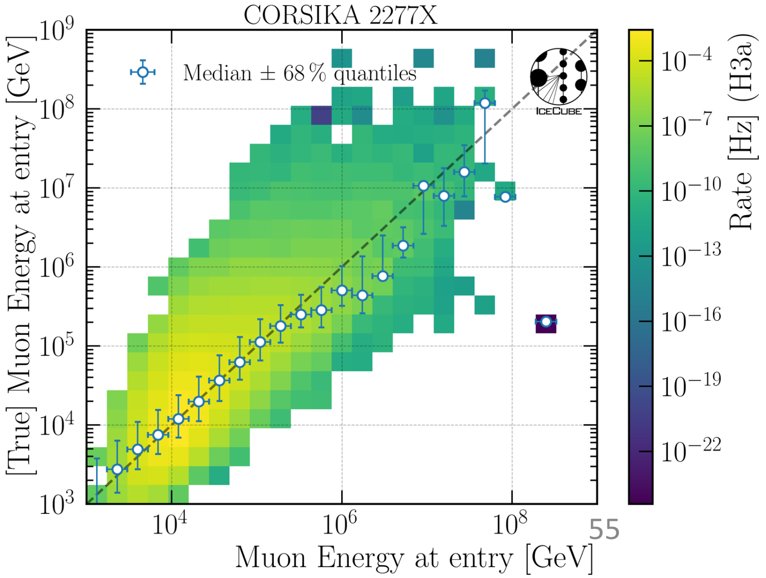
- Zenith / azimuth of leading muon

## Geometry

- Propagation length
- Entry point

- dnn_reco framework (Mirco Hünnefeld)
- Low-level features based on charge and timer per DOM



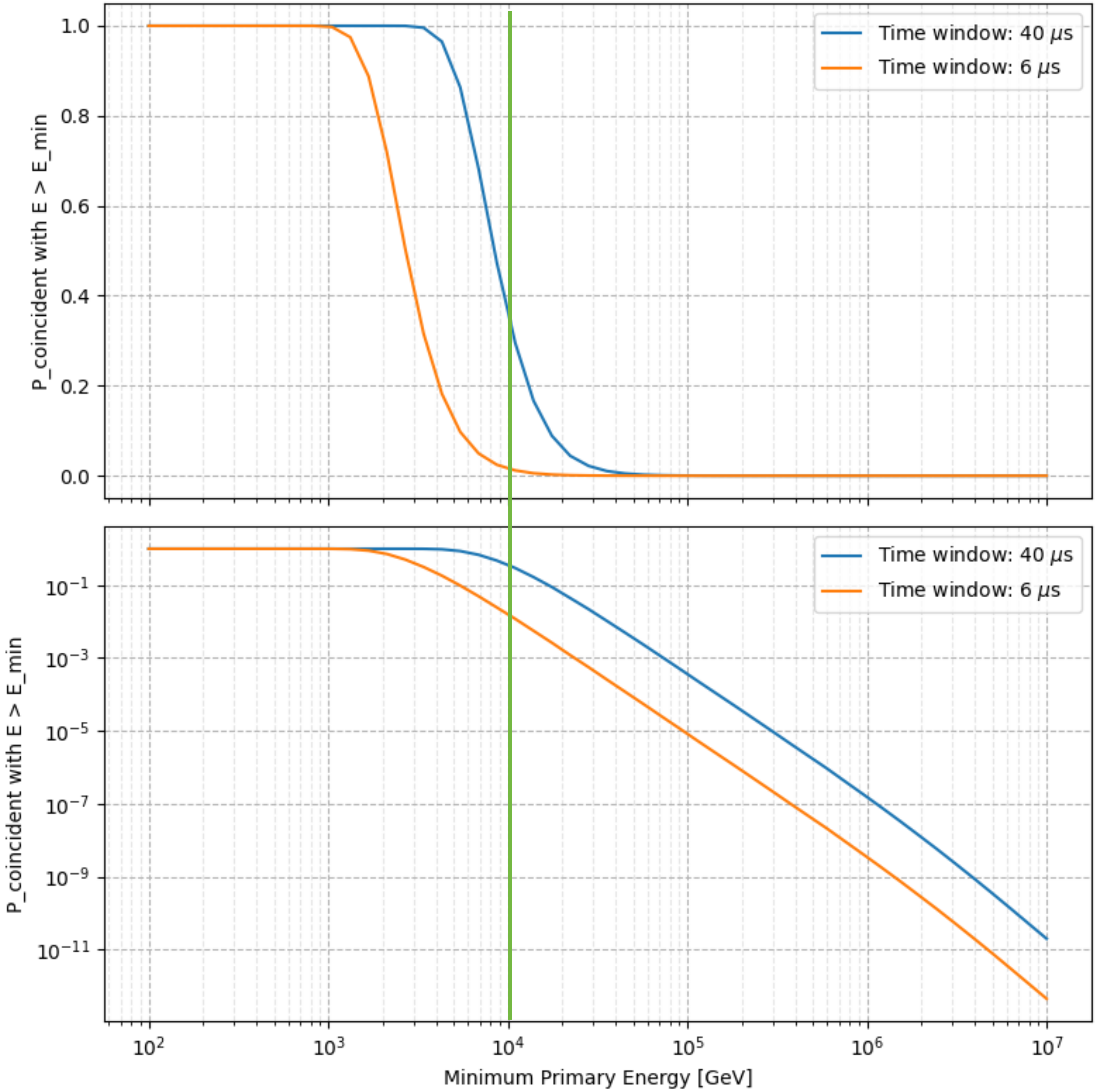




pascal.gutjahr@tu-dortmund.de

# Q: Coincident events are not simulated. Can you show that this does not cause any problems?

- For primary energies above 10 TeV, using a time window of 6 µs, the chance for a coincident event is ≤ 1 %

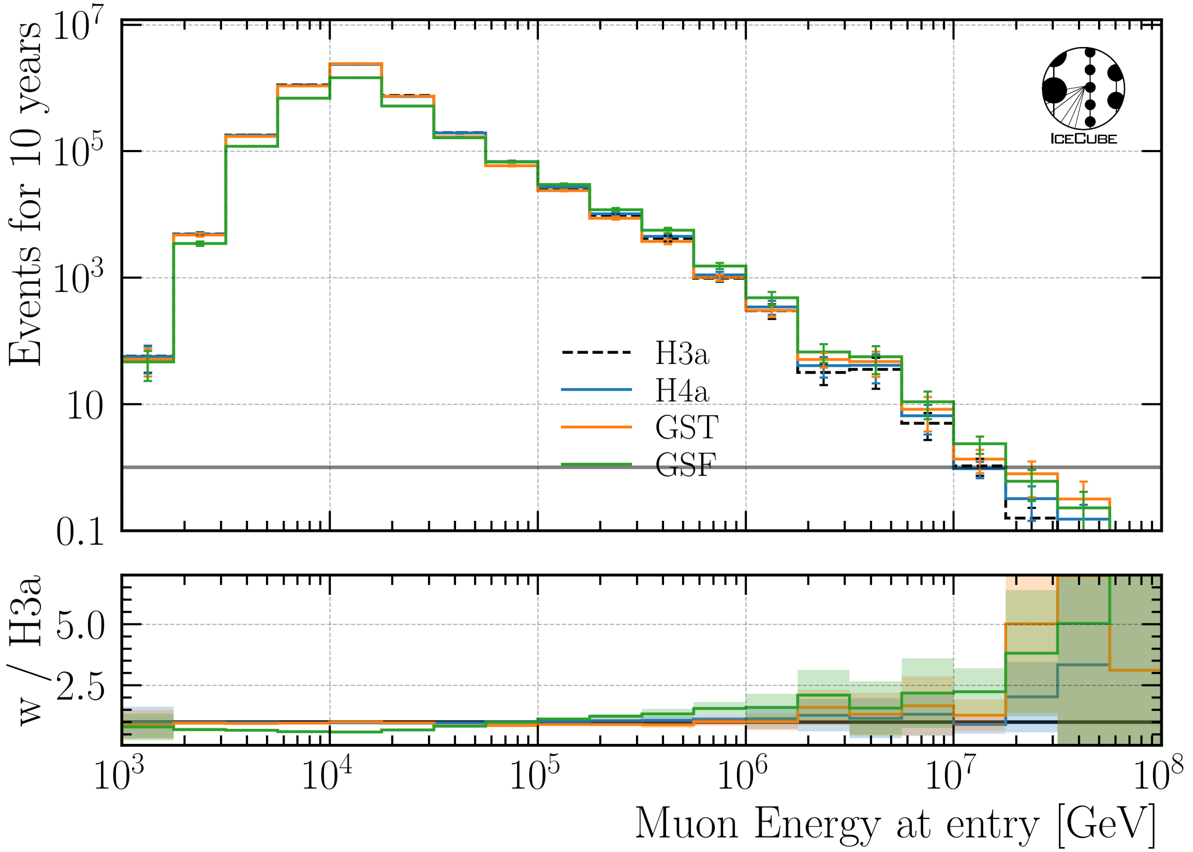
➢ Apply time-window cleaning of 6 µs to pulses

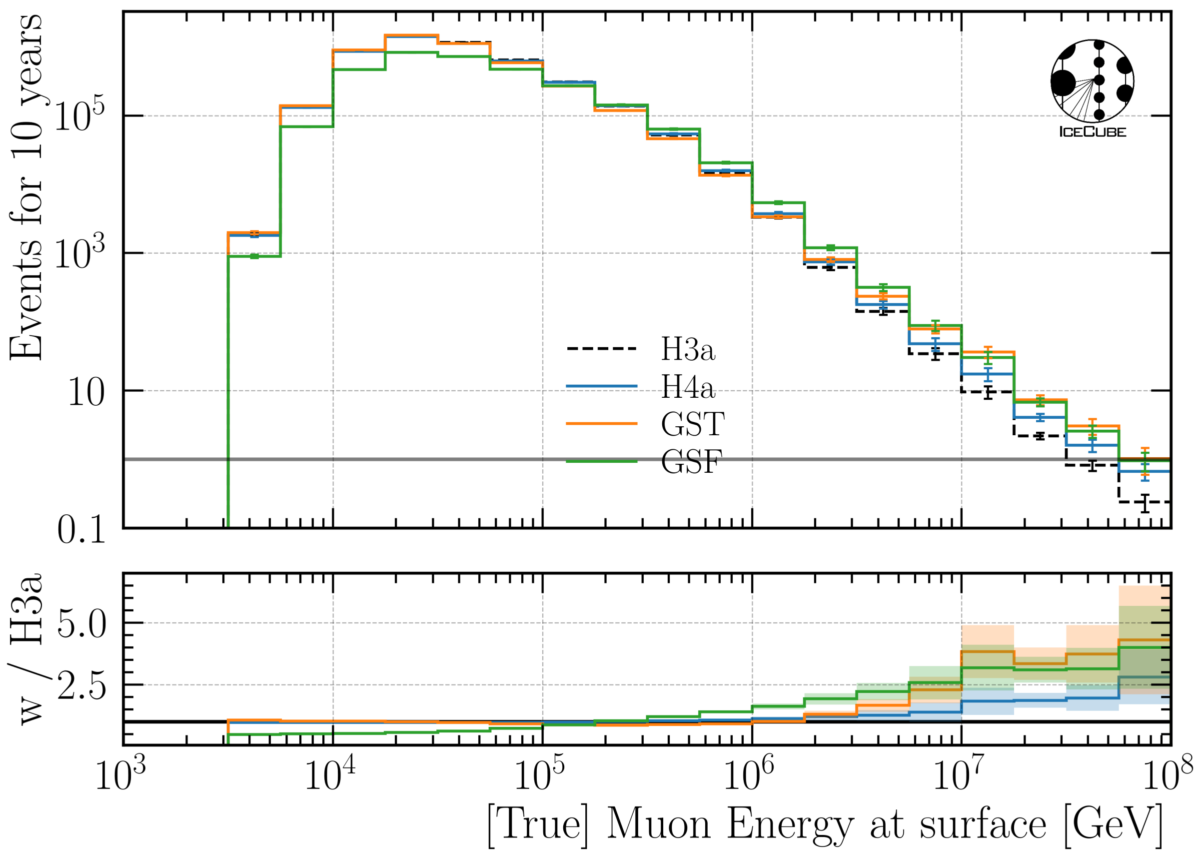



Linear scale

Log scale

# Q: What is the impact of the primary flux on your (most important) distributions?




➢ Primary models diverge towards higher energies