The status of pass2 cascade 11 year data diffuse analysis

Zelong Zhang zelong.zhang.1@stonybrook.edu

Background And Purpose

- GlobalFit with diffuse numu and cascade project is on-going.
  - https://events.icecube.wisc.edu/event/127/contributions/7505/attachments/5869/6940/GlobalFit_SnowStorm.pdf
  - Going to use pass2 data and pass2 snowstorm mc
- Cascade analysis based on Hans Niederhausen's work on pass1 (2012-2015) 4 year data.
  - https://wiki.icecube.wisc.edu/index.php/Multi_Year_Cascades
  - https://wiki.icecube.wisc.edu/index.php/Multi_Year_Cascades/Paper
- Goal: Request to unblind pass2 (2010-2020) 11 year cascade data for the GlobalFit numu+cascades
- https://wiki.icecube.wisc.edu/index.php/Pass2_Multi_Year_Cascade_Analysis

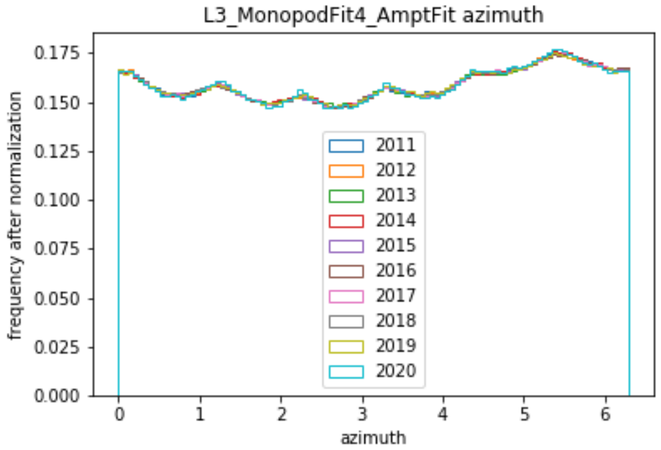
Cascade Final level Cuts

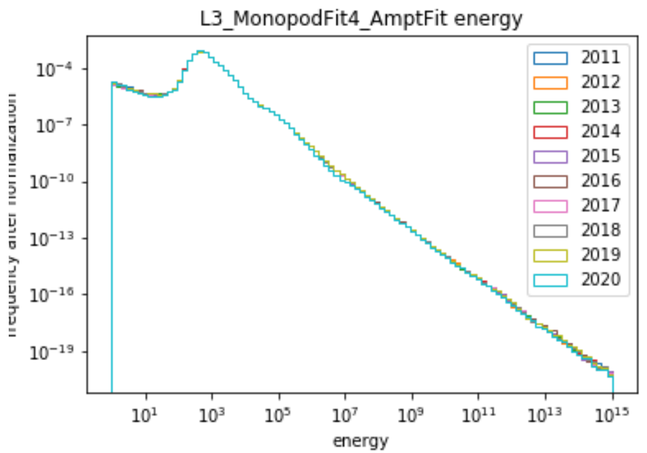
- Event selections we use are the same as for previous (pass1) cascade analysis
  - Low energy (H. Niederhausen): https://wiki.icecube.wisc.edu/index.php/Multi_Year_Cascades/Low_Energy
    - Some clean cuts:
      - anti-dust layer cuts
      - containment cuts
      - large delay time cuts (safetymargin)
    - boosted decision tree event classification
      - final_cascade
      - final_hybrid
      - final_muon
  - High energy (Y. Xu): https://wiki.icecube.wisc.edu/index.php/Multi_Year_Cascades/High_Energy

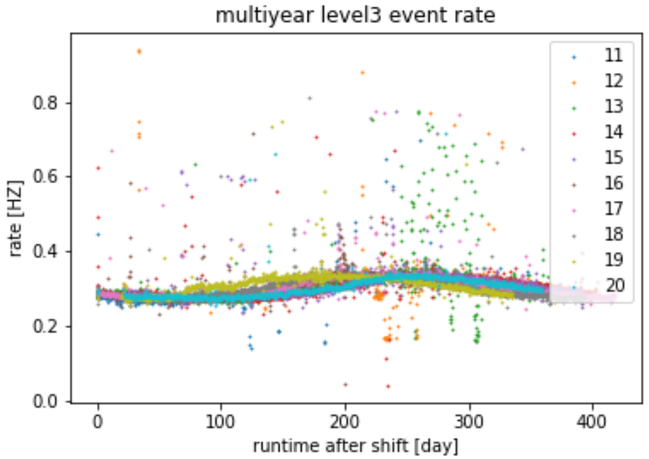
Pass2 comparison of variables across years

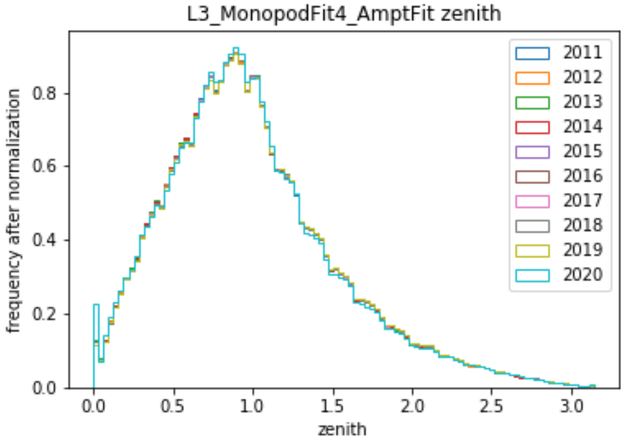
Distribution of all L3 events*

- Goal: 11 year data (2010-2020)
- Pass2 data is uniform across years (2010 is to be checked)
- Some reconstructed variables







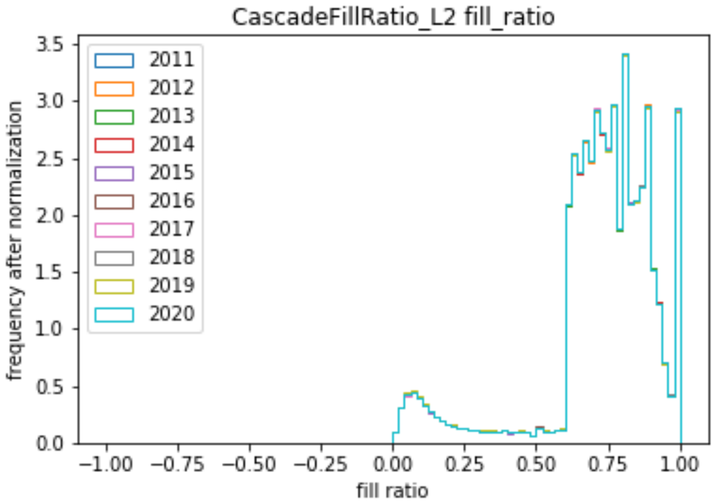


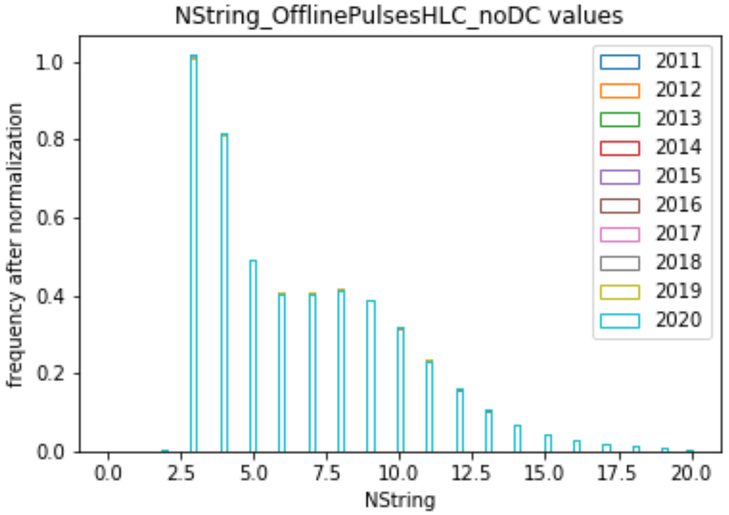
All L3 events includes single contained, single uncontained and coincidences contained events

Pass2 comparison of variables across years

Distribution of all L3 Events*

- Goal: 11 year data (2010-2020)
- Pass2 data is uniform across years (2010 is to be checked)
- Some low filter level variables



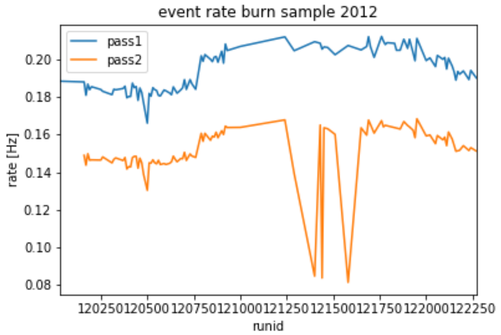


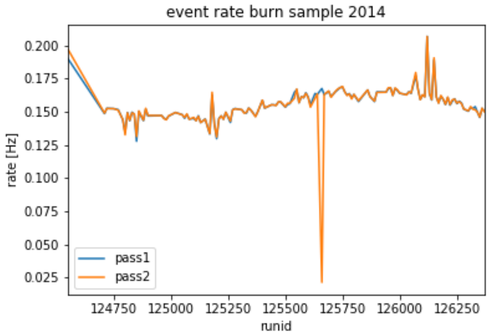
All L3 events includes single contained, single uncontained and coincidences contained events

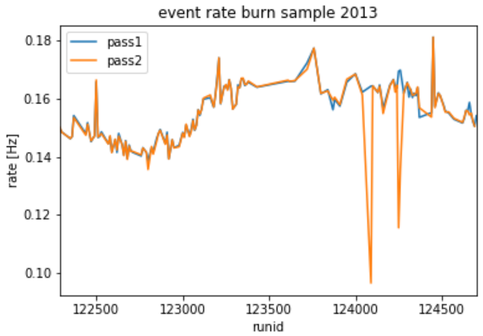
Pass2 vs Pass1 comparison

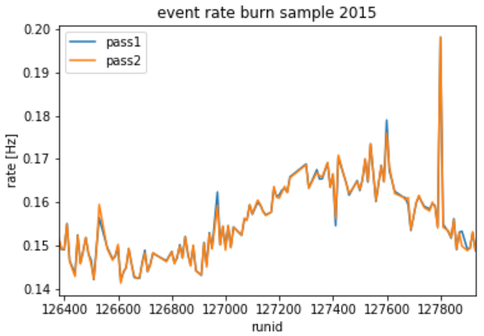
Level3 single contained event rate

- 2012 is special, 2012 pass1 data used ‘CascadeFilter_12’ online filter while other years used ‘CascadeFilter_13’ online filter. And all pass2 data used ‘CascadeFilter_13’ online filter
- For other years pass1 and pass2 data match with each other, except certain runs, which will be removed by good run list if they are outliers.





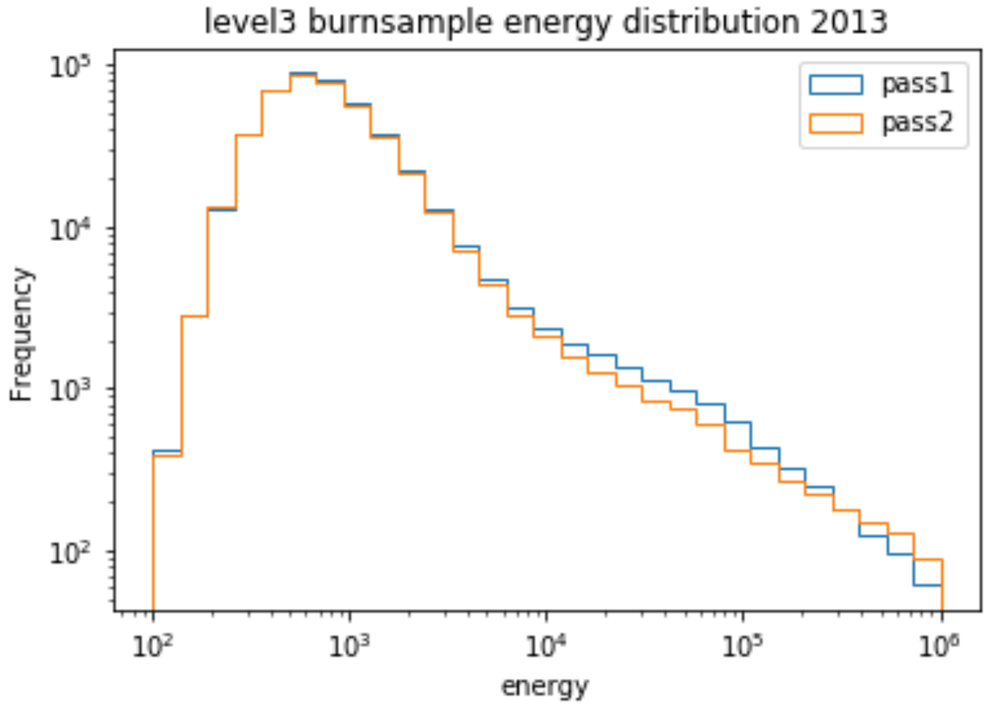




Pass2 vs Pass1 comparison

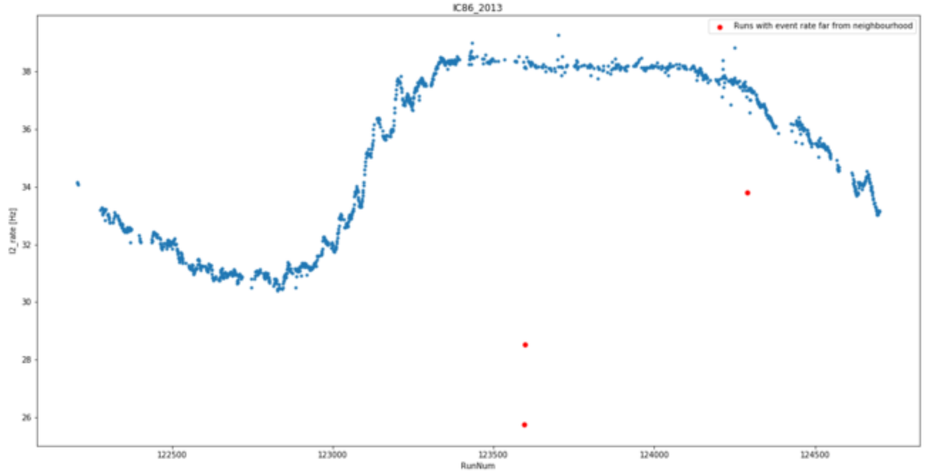
Level3 single contained event energy distribution

- Low energy part matches well
- High energy part: the difference is caused by background, has been verified that it will disappear at final cut level

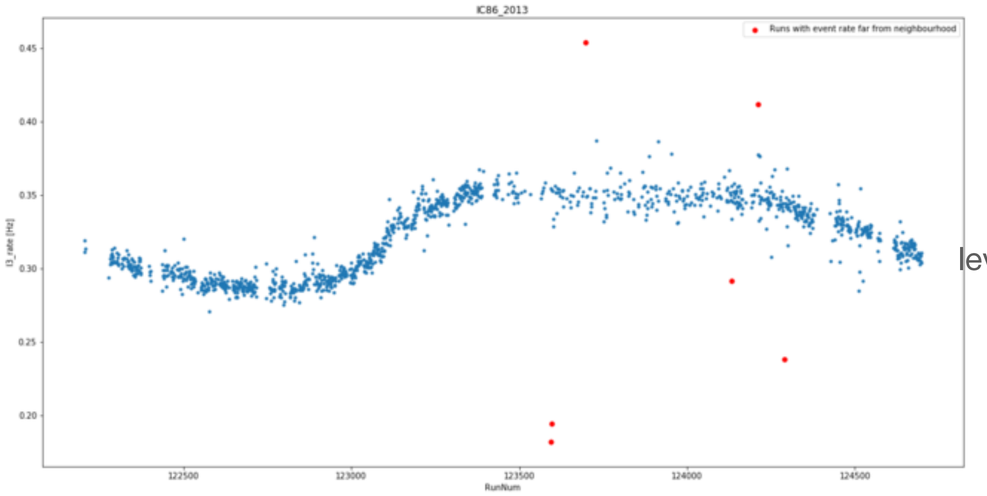


Pass2 Good Run List

- Basing on official Good Run List for pass2
- Select runs with 86-string configuration
- Remove outliers (find by eyes, larger than 20% deviation) according to **Cascade L3 rate**
- Final Good Run Lists are uploaded to my wiki page https://wiki.icecube.wisc.edu/index.php/Pass2_Multi_Year_Cascade_Analysis



level2

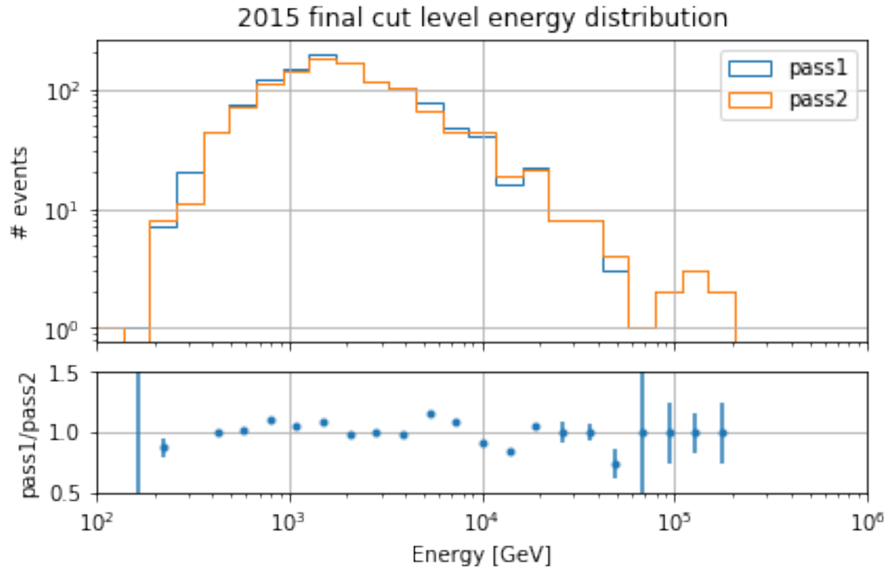


level3

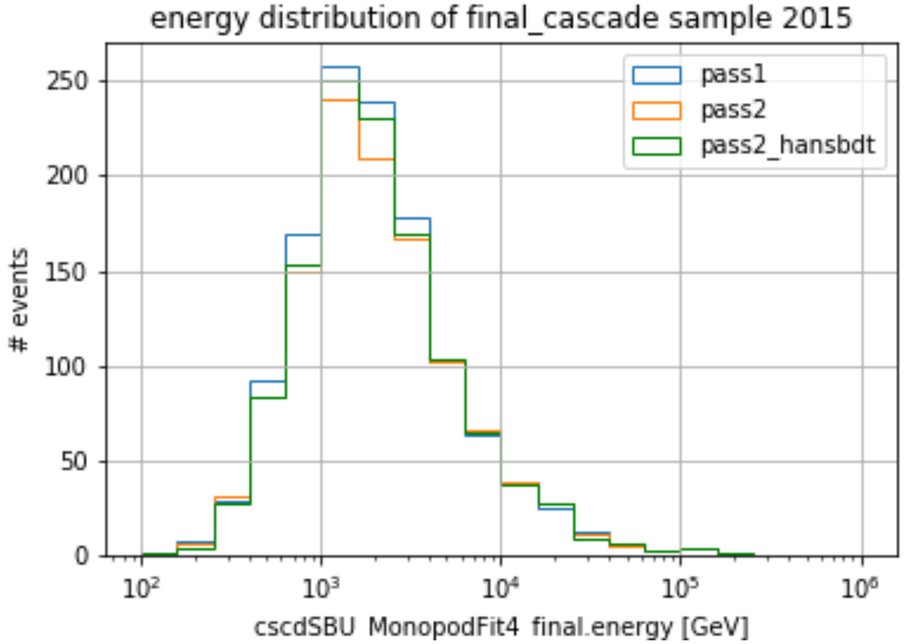
Final cut level:

Pass2 and Pass1 comparison

- Good match for high energy events (>30 TeV)
  - Only four events are lost and no event introduced. The reason of lost events are known (one due to early hit pulse, two due to good run list and one located at the boundary of LE and HE selection)
- Pass1 has overall higher rate at low energy
  - Part of the reason (version of xgboost) have been found and fixed.

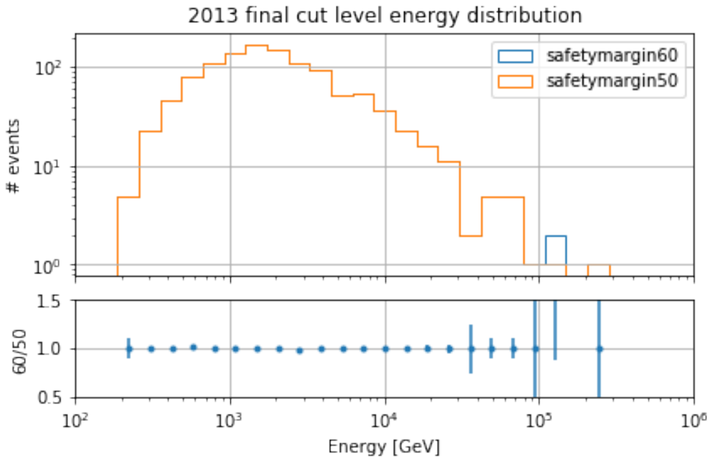


After the xgboost version fix

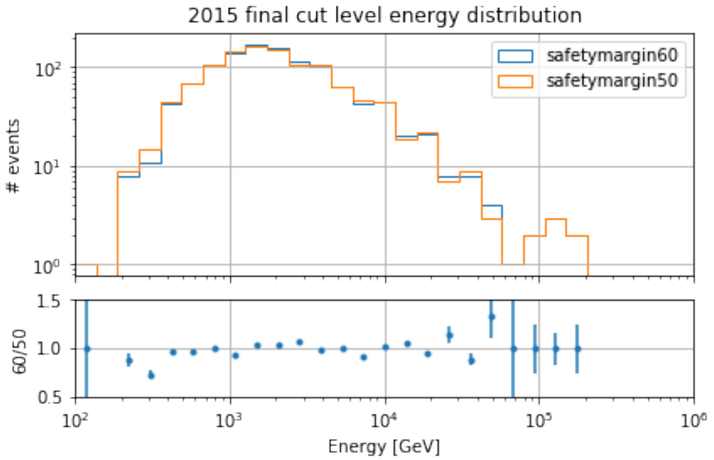


Safetymargin Issue

- Safetymargin is a parameter in function ACausalHitSelector '/data/user/zzhang1/combo/stable/src/CascadeL3_IC79/python/level4/veto.py'. It is used to count early hit pulses.
- Changed safetymargin from 50 to 60 compared to previous pass1 cascade analysis
- Have studied the effect of change.



- 2012-2014: only several events are different,
- pass1 and pass2 spectrum almost perfectly match with each other

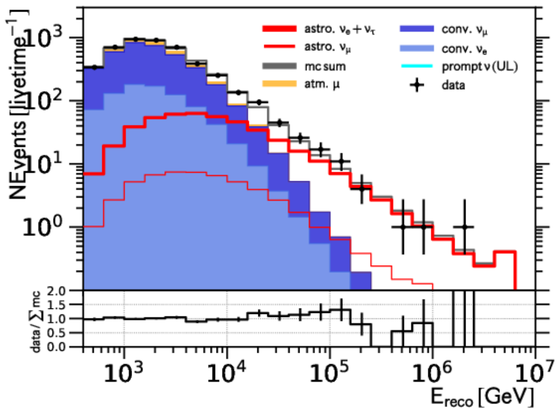


- 2015: the difference is relatively larger, total number of events is 1144(safetymargin50)/ 1145(safetymargin60), and the number of overlap events are 1095
- pass1 and pass2 spectrum basically match with each other
- Doing the double check

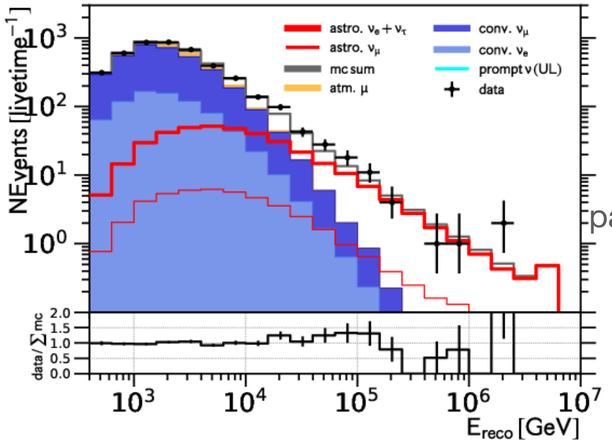
Final level

Fit result - Pass2 vs Pass1 data cross check

- I reproduced Hans' fit result with Pass1 data and Pass1 mc as a proof of I use his code correctly
  - 4 years (2012-2015) Pass1 data Final_cascade (66 bins), Final_hybrid (11 bins), Final_muon (1bin) samples,
  - Plus 2 years (2010-2011) data cascade_mlb (45 bins) samples.
- Fitted Pass2 data with Pass1 mc. I get a reasonable fit result and goodness of fit (115, given the number of degrees of freedom of chi-square distribution which is 112, so the Cumulative distribution function is 0.596)
  - 4 years (2012-2015) pass2 data Final_cascade (66 bins), Final_hybrid (11 bins), Final_muon (1bin) samples,
  - Plus 2 years (2010-2011) pass1 data cascade_mlb (45 bins) samples.
  - Different data point at highest energy bin dues to treatment of dust layer. Energy reconstruction of dust layer events will be done with GBDT or/and DirectFit for pass2 analysis, as it was done for pass1 cascade analysis.



pass1 data vs pass1 mc

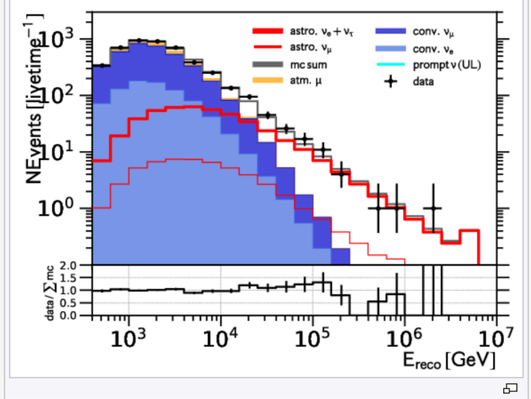


pass2 data vs pass1 mc

Final level

Fit result - pass2 vs pass1 cross check

Pass1 data pass1 mc fit result:

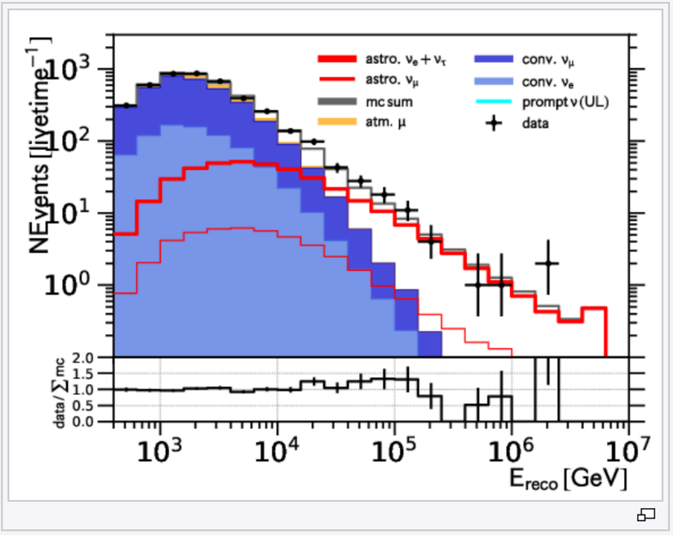


RESULTS (fitstatus: 0, edm: 5.00117e-05)
muon_norm = 1.44998 ± 0.04
conv_norm = $1.0689^{+0.13}_{-0.12} \times \Phi_{HKKMS06}$
prompt_norm = 0.00929372
muon_norm_mlb = 0.864457
delta_cr = 0.0168859 ± 0.03
absorption = $1.03359^{+0.05}_{-0.04}$
dom_efficiency = $1.025^{+0.08}_{-0.07}$
scattering = 1.02091 ± 0.03
zholeice_scattering = 1.72221 ± 0.19
astro_norm = $1.65672^{+0.25}_{-0.27}$
astro_index = 2.5277 ± 0.07
likelihood value (gof): 84.37
likelihood value (abs): 437.836
time to find minimium and evaluate exact likelihood: 1063.82s
number of llh evaluations: 765

The fit result of pass1 data and pass1 mc. Energy distribution of final cascade sample.

From Hans' paper
https://wiki.icecube.wisc.edu/index.php/Multi_Year_Cascades/Paper

Pass2 data pass1 mc fit result:



RESULTS (fitstatus: 0, edm: 0.000653766)
muon_norm = 1.21388
conv_norm = 1.16293
prompt_norm = 0.930632
muon_norm_mlb = 0.838293
delta_cr = -0.0543748
absorption = 1.03829
dom_efficiency = 0.963024
scattering = 1.06913
zholeice_scattering = 1.58903
astro_norm = 1.74359
astro_index = 2.46977
likelihood value (gof): 115.751
likelihood value (abs): 467.965
time to find minimium and evaluate exact likelihood: 1430.31s
number of llh evaluations: 990

Finallevel

pass2 data vs pass2 snowstorm mc

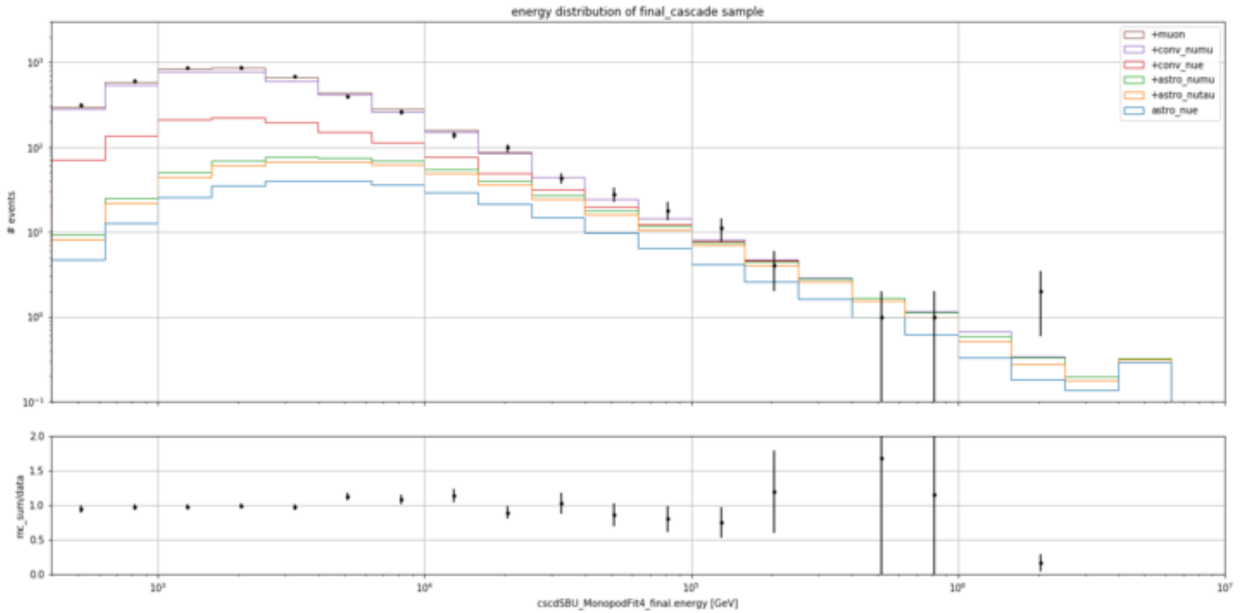
- No systematic fit, the parameters are only conv_norm, muon_norm, astro_norm, astro_index
- Final_cascade, final_hybrid, final_muon sample (78 bins in total)
- Reasonable fit result:

  muon_norm = 1.265

  conv_norm = 1.053

  astro_norm = 1.527

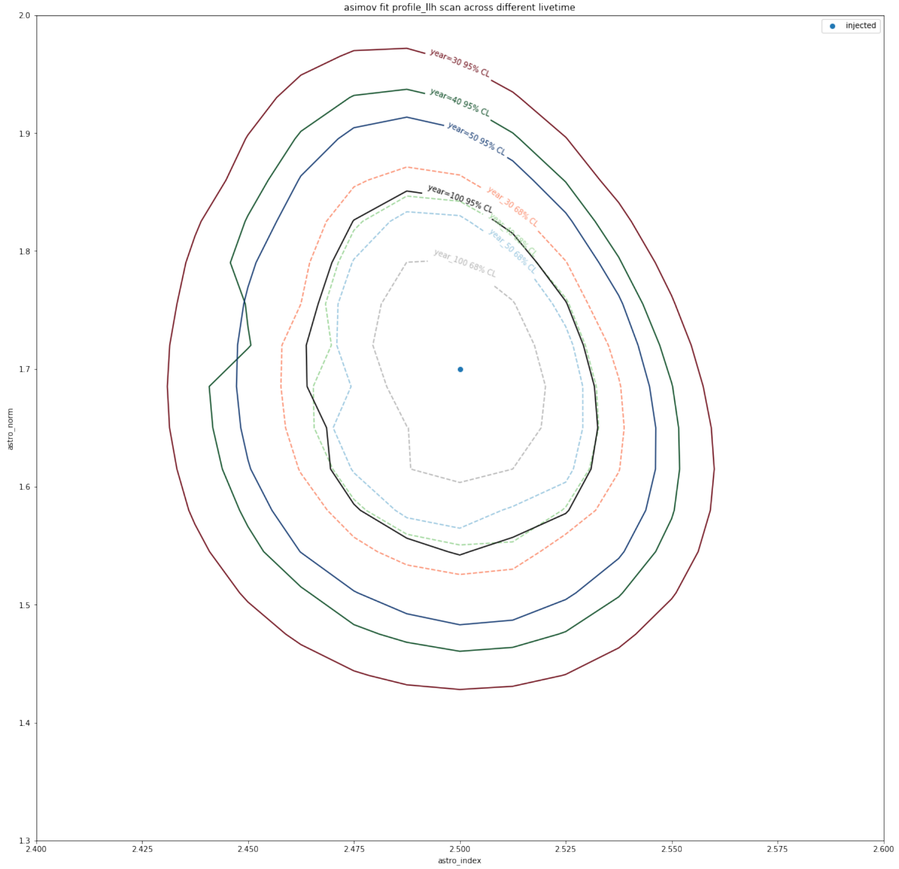
  astro_index = 2.608
- Goodness of fit is poor (135, given dof of chi2 distribution is 74).
- Study of pass2 data vs pass2 snowstorm mc is on going.



Outlook

asimov profile llh scan over different livetime

- Asimov fit injected point:
  - astro_norm: 1.7
  - astro_index: 2.5
- profile llh scan across 30, 40, 50, 100 years livetime
- Give us a feeling of the statistic and systematic uncertainty of Gen2



Conclusions

- Goal: Request to unblind 11-year (2010-2020) pass2 cascade filter data for GlobalFit (numu+cascades)
- The pass2 cascade selections were slightly modified compared to pass1 analysis (Good Run List and safetymargin)
- Pass2 data is uniform and match with pass1 data for overlapping years (2012-2015)
- Pass2 2010 and 2011 data are being processed.
- Pass2 data (2012-2015) vs pass1 mc fit are made as a cross check of pass1 data vs pass2 data comparison.
- See https://wiki.icecube.wisc.edu/index.php/Pass2_Multi_Year_Cascade_Analysis for more detailed information

Back up

Pass2 comparison across years

Event rate (includes 2010) - all L3 events

- Goal: 11 year data (2010-2020)
- Pass2 data is uniform across years (2010 is to be checked)
- All L3 events (includes single contained, single uncontained and coincidences contained events)