RUN 3 AmBe Simulation (with QF) + simple ER/NR discrimination

Simulation setup

1. Geant4 (**10^7 neutrons** from AmBe source, implemented QF)
2. Digitization (as in RUN 3)
3. Reconstructed as RUN 3

Only ~20000 of the events generate some kind of energy deposit in the gas (0.02%)

Not all of them are NR. More than half of them are ER due to neutrons interacting with PMMA and copper and produce photons/e-

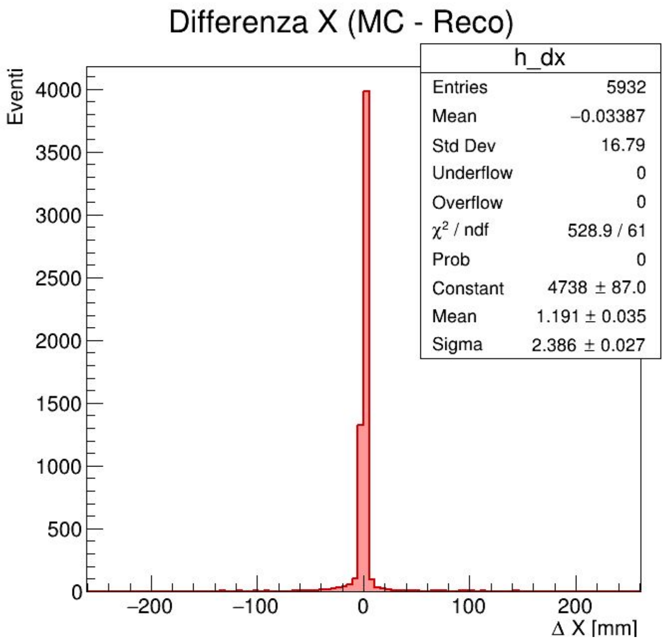
Matching MC truth with reconstructed tracks

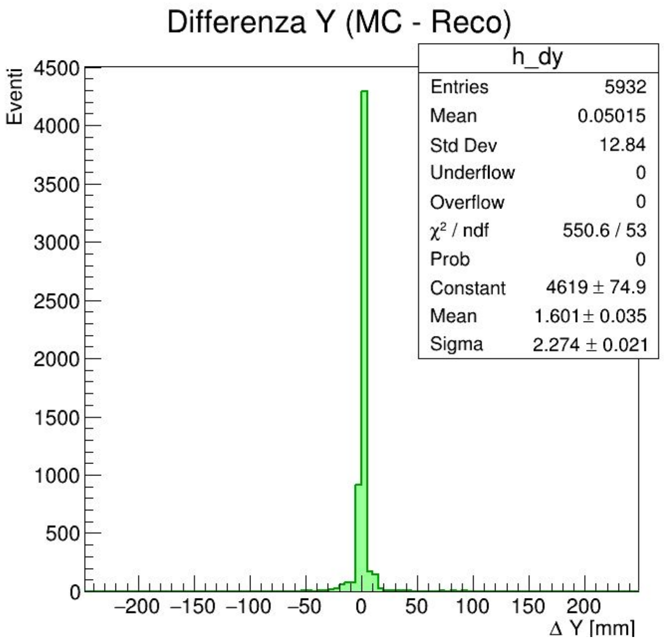
1. **Loop over reconstructed clusters**
   - For each cluster, compute the **reconstructed barycenter**.
2. **Read the corresponding MC truth (Geant4) event**
3. **Read Geant4 hits for that event**
   - Each G4 event can have multiple particles (we identifiy by **trackID_hits**).
   - For each unique **trackID**, select all hits belonging to that track.
4. **Compute barycenters in G4 (x,y)**
   - For each track (trackID), compute the **x-y barycenter** using its hits.
5. **Matching**
   - Compare each reconstructed barycenter with all MC barycenters in the event.
   - Assign the match to the **closest MC barycenter** (within a given threshold, es. 10 mm). See next slide for the **reasonable theshold**
6. **After match:**
   - Do all analysis you want

Differences between MC truth and reco

We select events with a single track, and reconstructed pictures with a single high-density track (density > 20). After the correct translation:

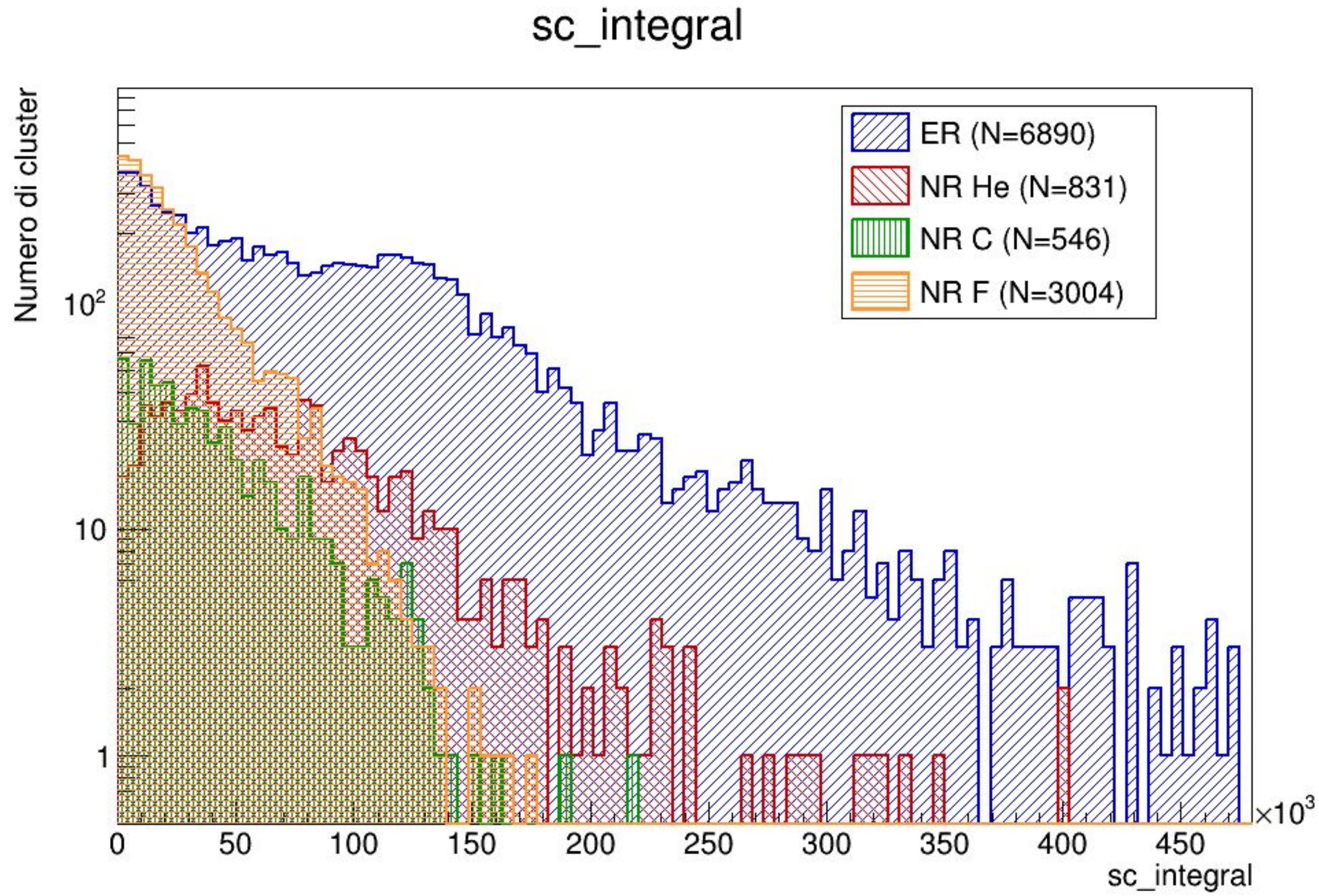
```python
def transform_mc_coordinates(barycenter_y_mm, barycenter_z_mm):
    """Trasforma coordinate MC"""
    y_offset = - barycenter_y_mm - 165
    z_offset = - barycenter_z_mm + 100
    return z_offset, y_offset

def transform_reco_coordinates(sc_xmean, sc_ymean):
    """Trasforma coordinate ricostruite"""
    reco_x = (sc_xmean - 1152) / 2304 * 346
    reco_y = (sc_ymean - 1152) / 2304 * 346
    return reco_x, reco_y
```



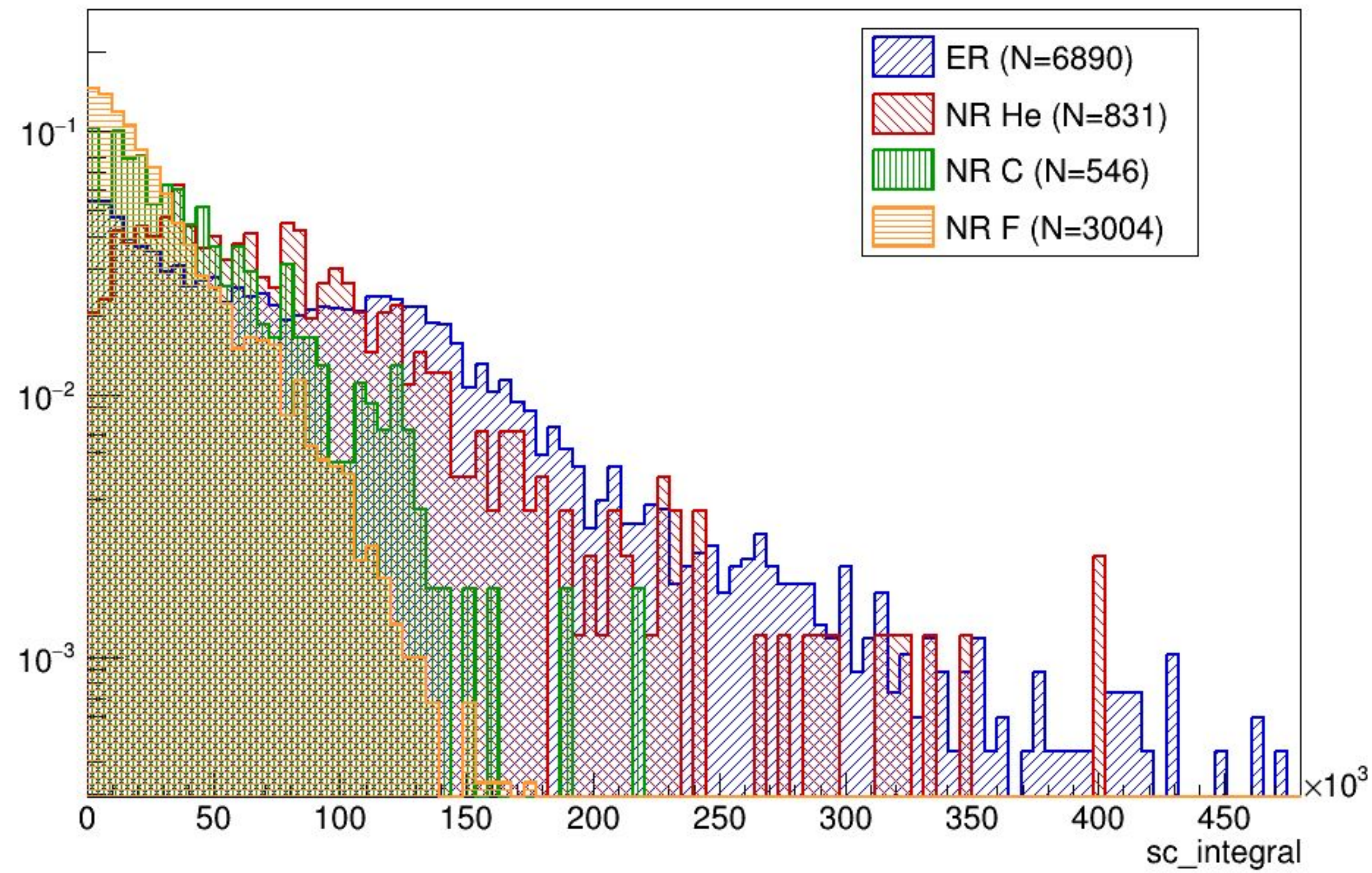


We use 10 mm as a threshold for a good match

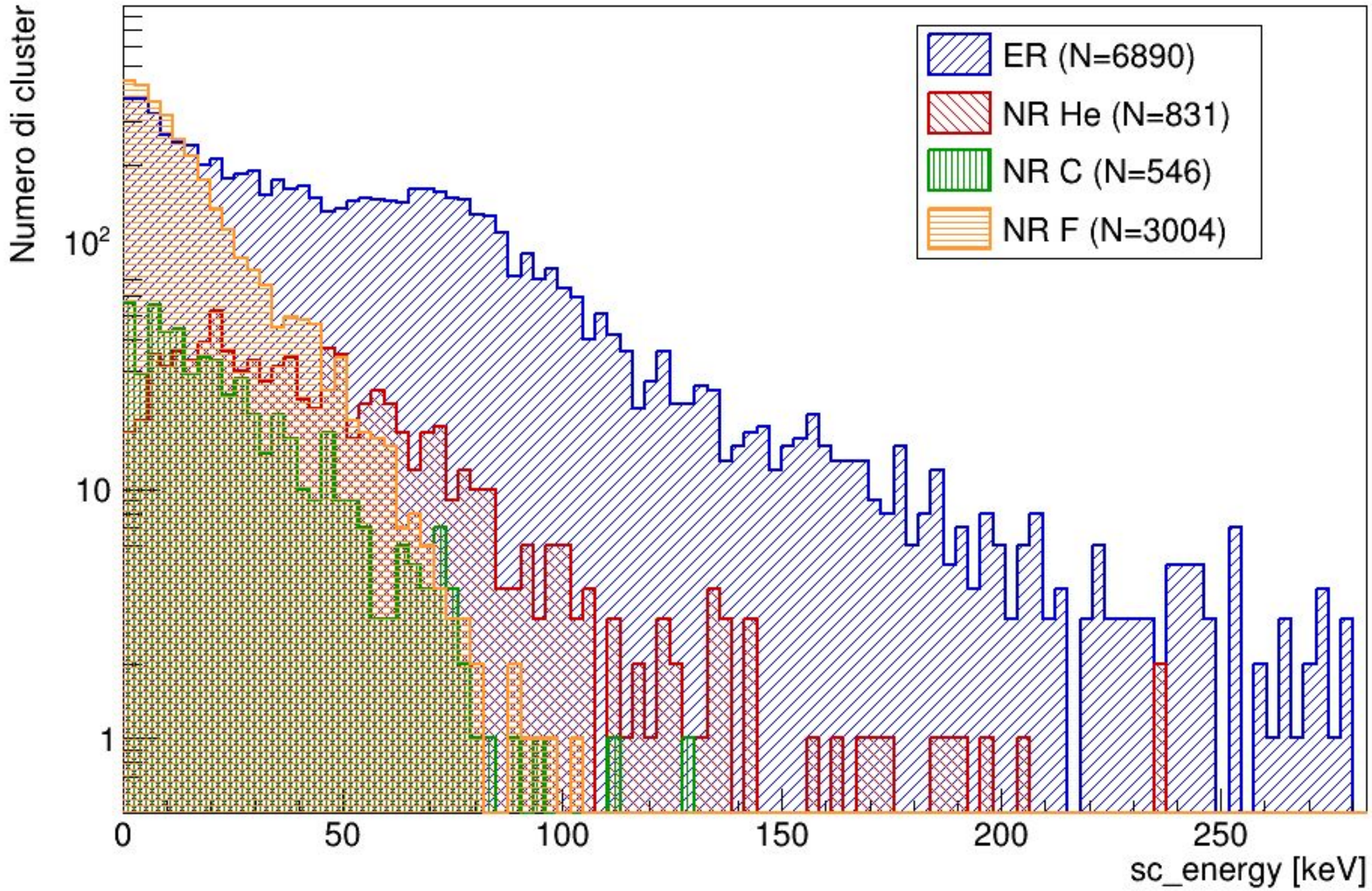
Reconstructed quantities
(sc_integral, sc_length, etc…)
by particle type


sc_integral
ER (N=6890)
NR He (N=831)
NR C (N=546)
NR F (N=3004)
Numero di cluster
1
10
10²
0
50
100
150
200
250
300
350
400
450
×10³
sc_integral

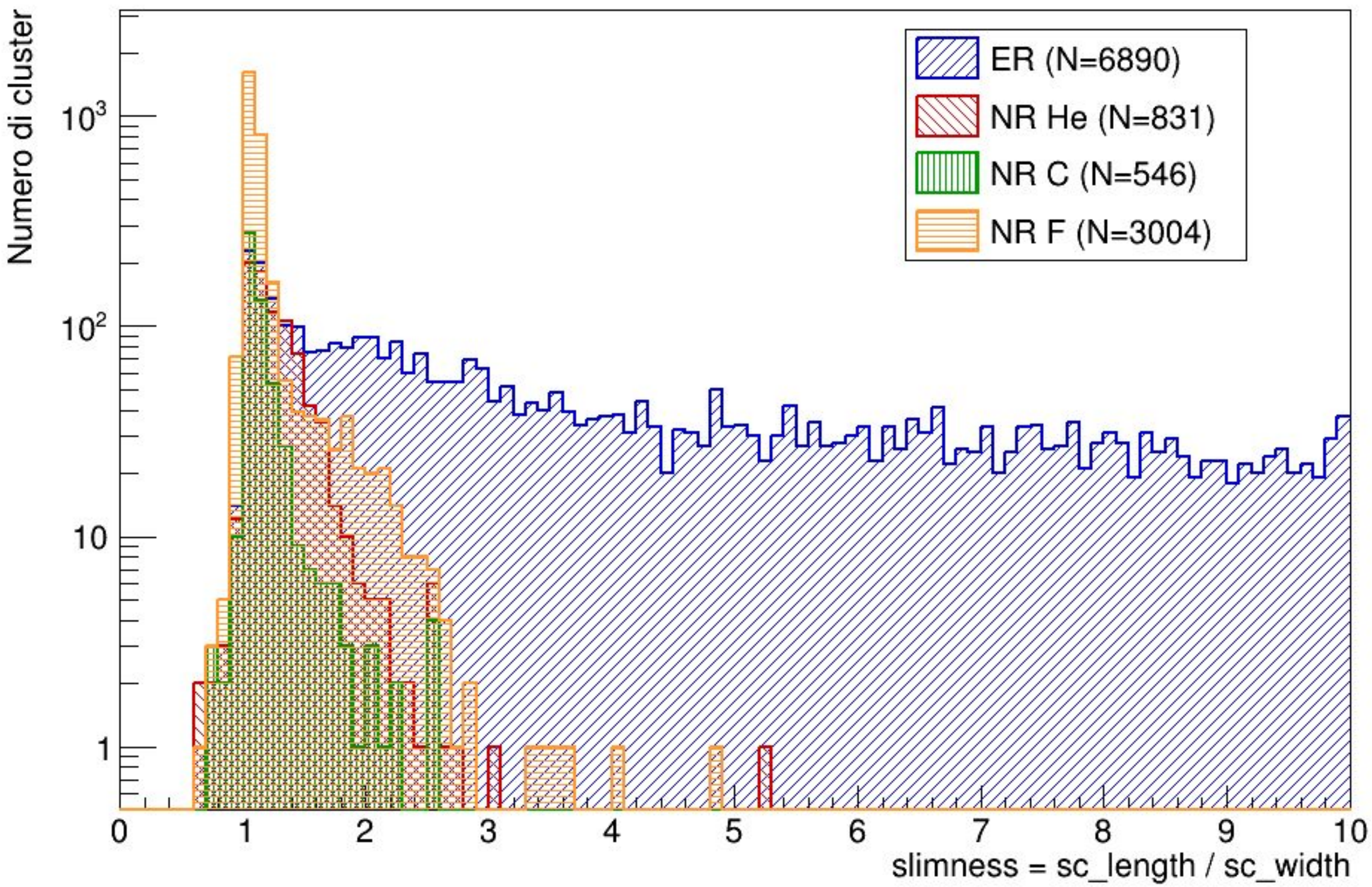
sc_integral (NORMALIZED)



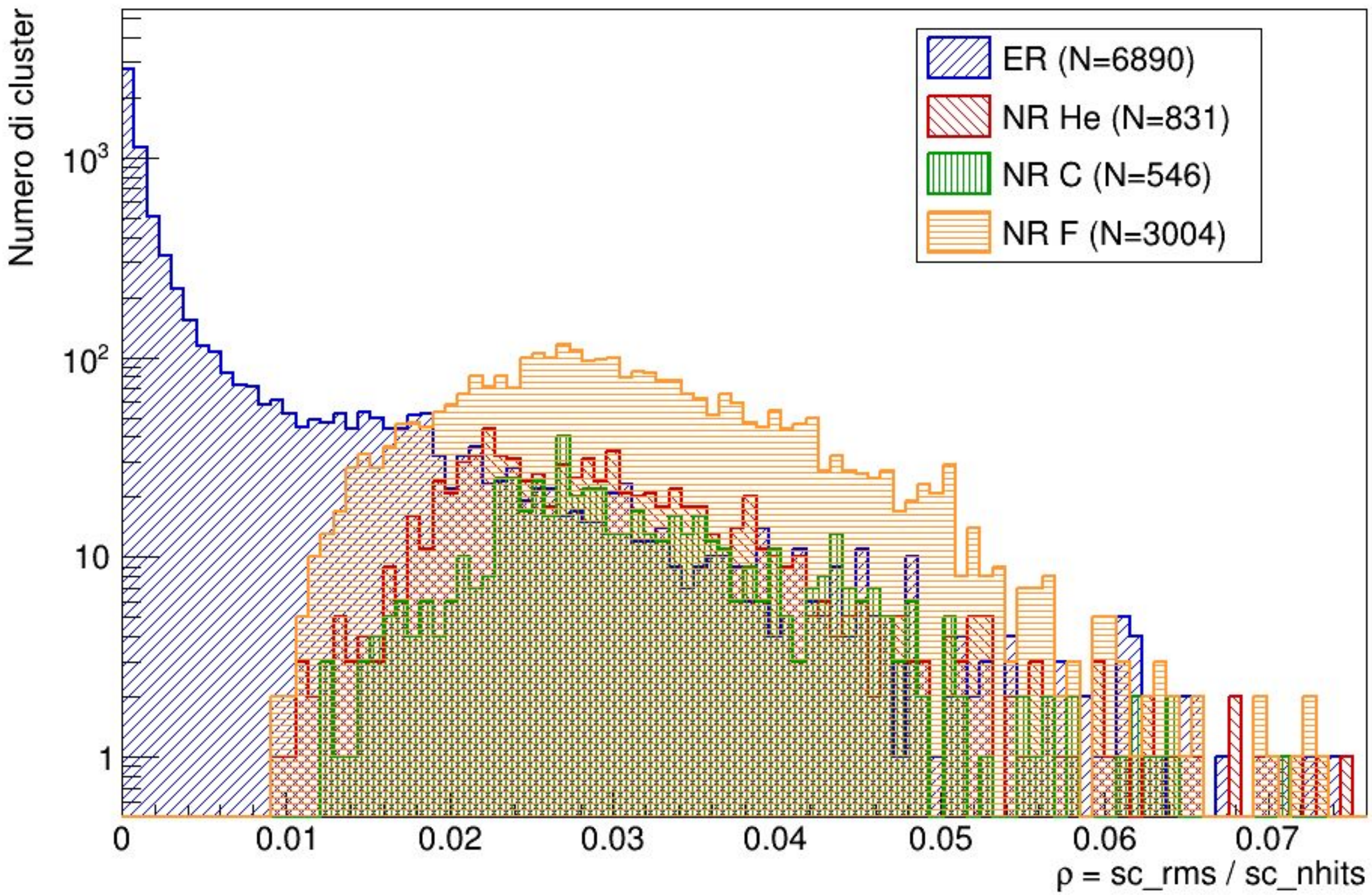
sc_energy calibrated with iron peak (5.9 keV) @ 10000 counts

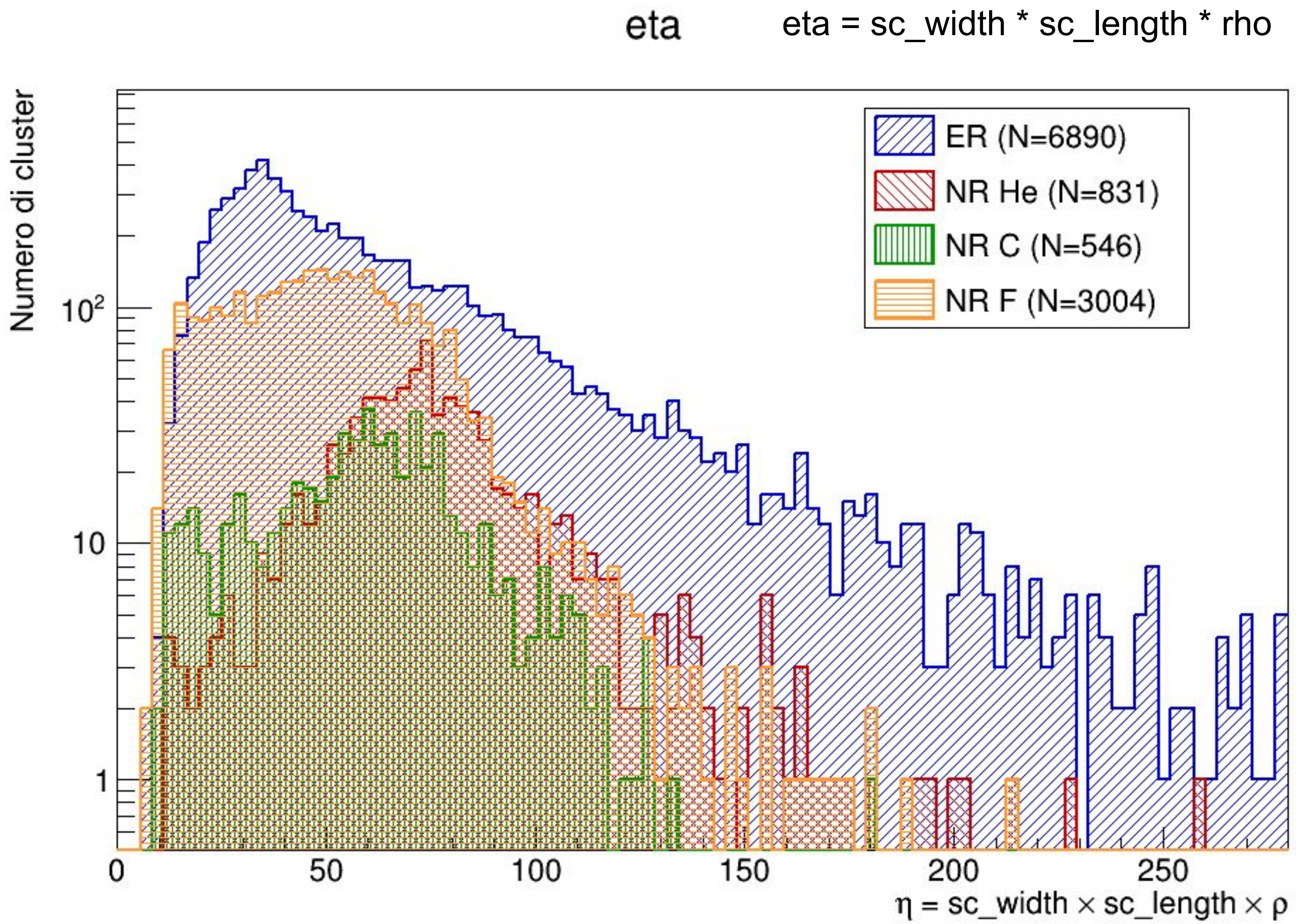


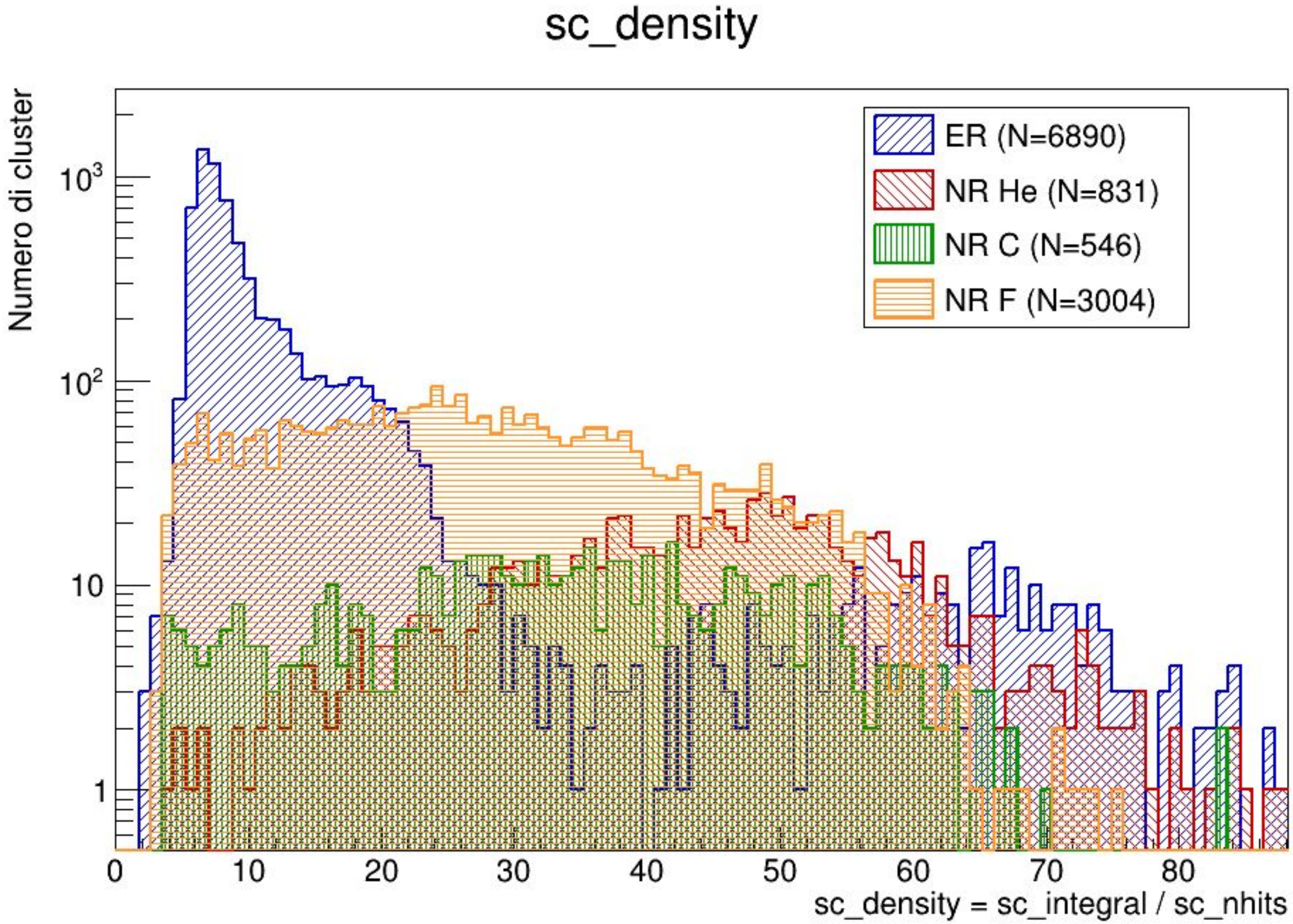
slimness sc_length / sc_width

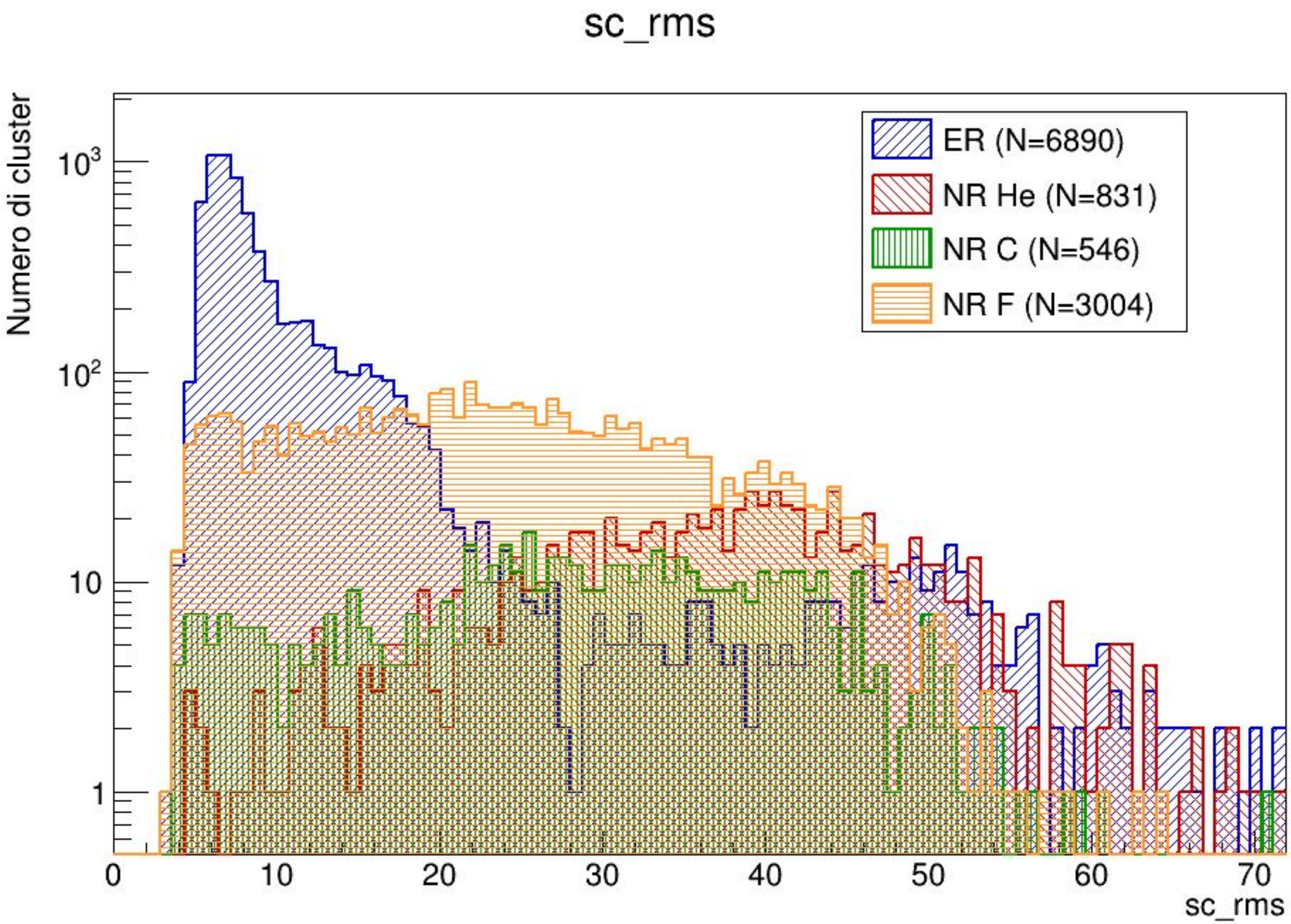


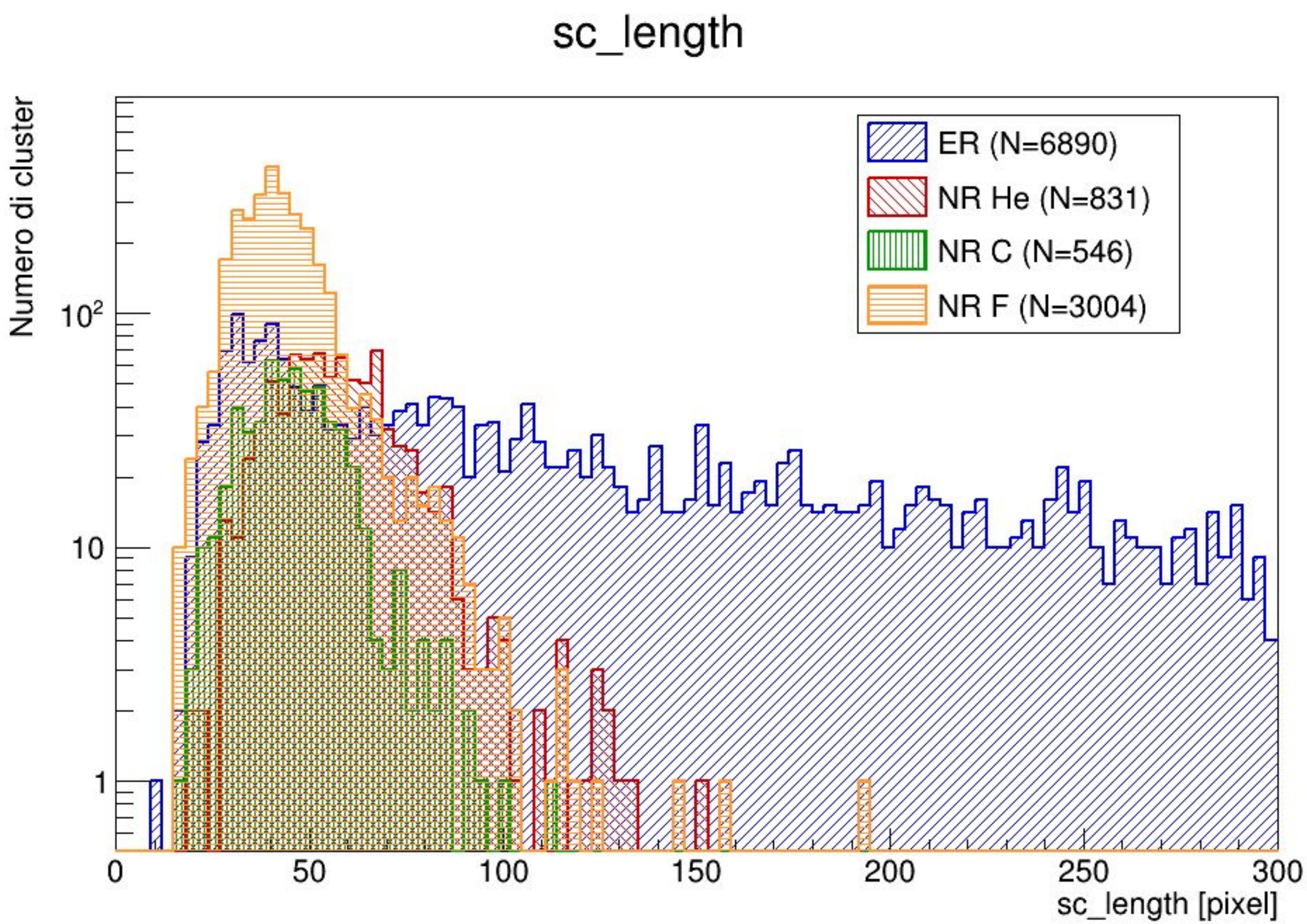
rho rho = sc_rms / sc_nhits

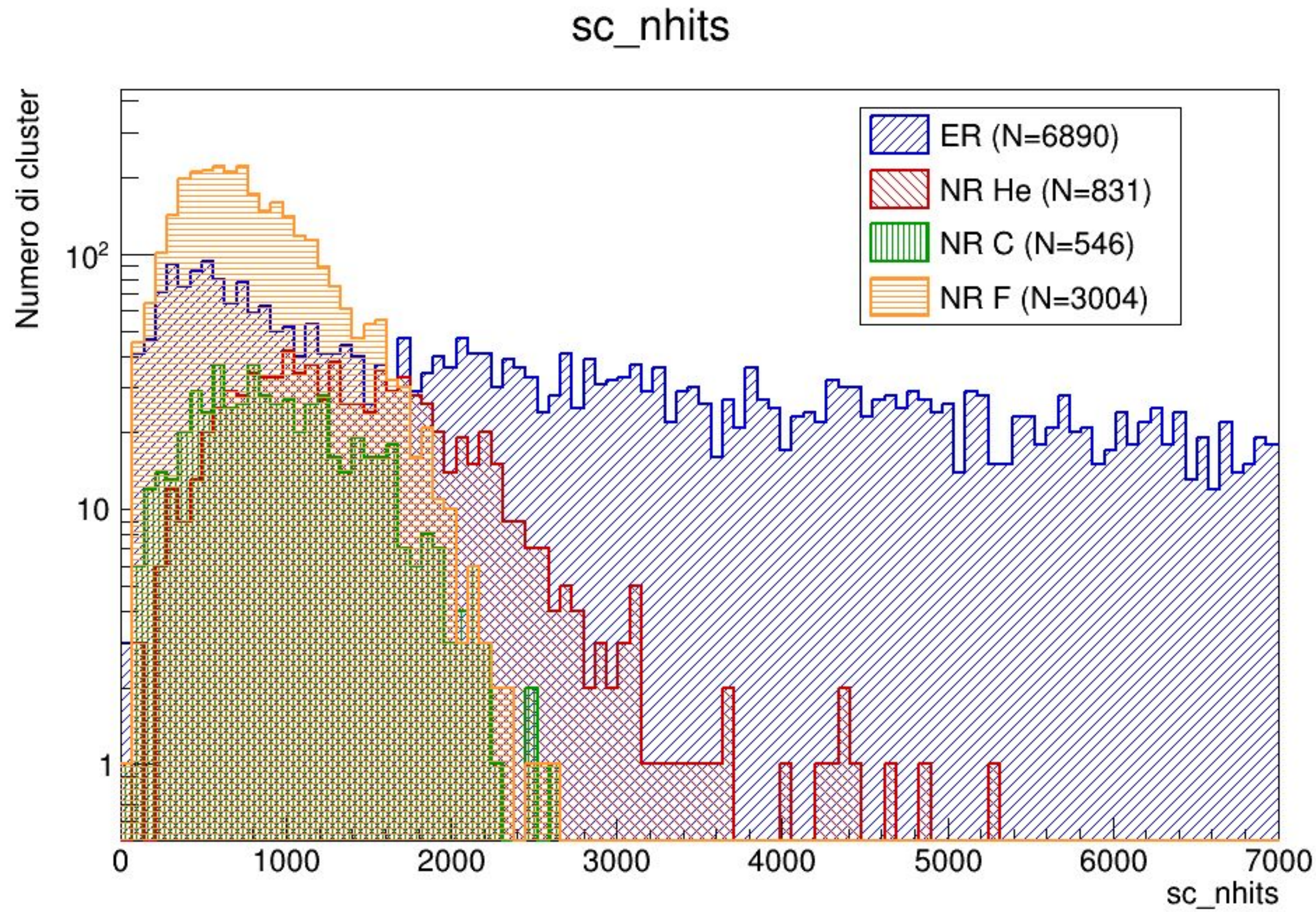



eta
eta = sc_width * sc_length * rho
Numero di cluster
ER (N=6890)
NR He (N=831)
NR C (N=546)
NR F (N=3004)
10^2
10
1
0
50
100
150
200
250
η = sc_width × sc_length × ρ


sc_density
Numero di cluster
10^3
10^2
10
1
ER (N=6890)
NR He (N=831)
NR C (N=546)
NR F (N=3004)
0
10
20
30
40
50
60
70
80
sc_density = sc_integral / sc_nhits

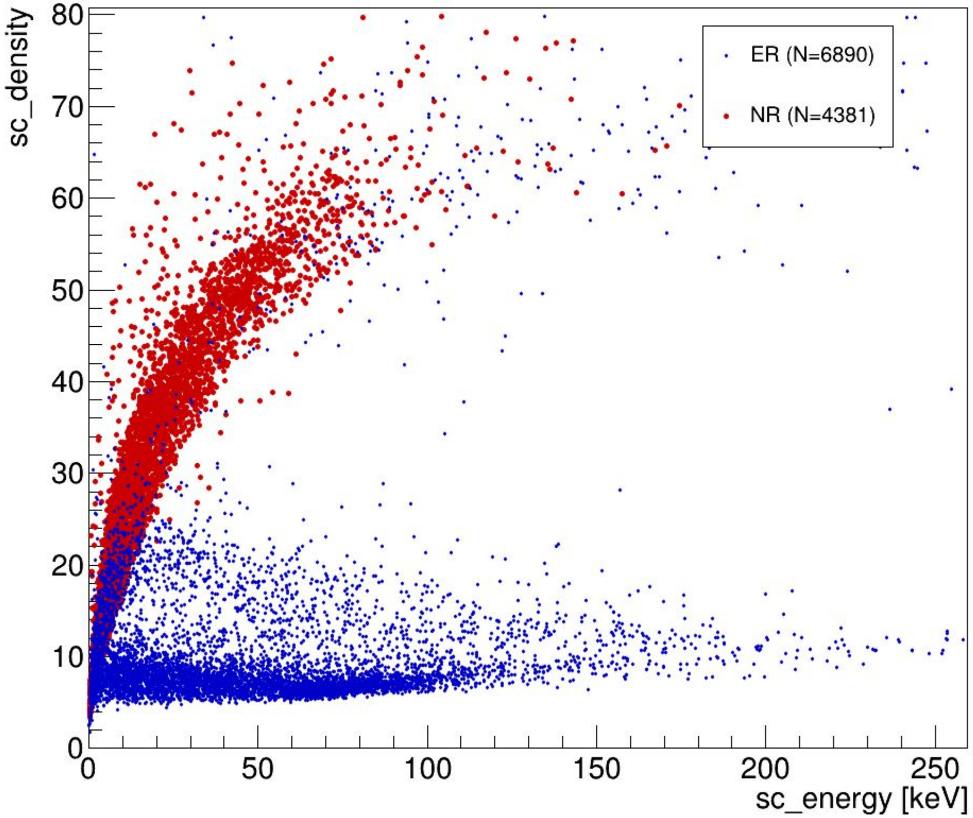

sc_rms
Numero di cluster
ER (N=6890)
NR He (N=831)
NR C (N=546)
NR F (N=3004)
10^3
10^2
10
1
0
10
20
30
40
50
60
70
sc_rms


sc_length
Numero di cluster
ER (N=6890)
NR He (N=831)
NR C (N=546)
NR F (N=3004)
10^2
10
1
0
50
100
150
200
250
300
sc_length [pixel]

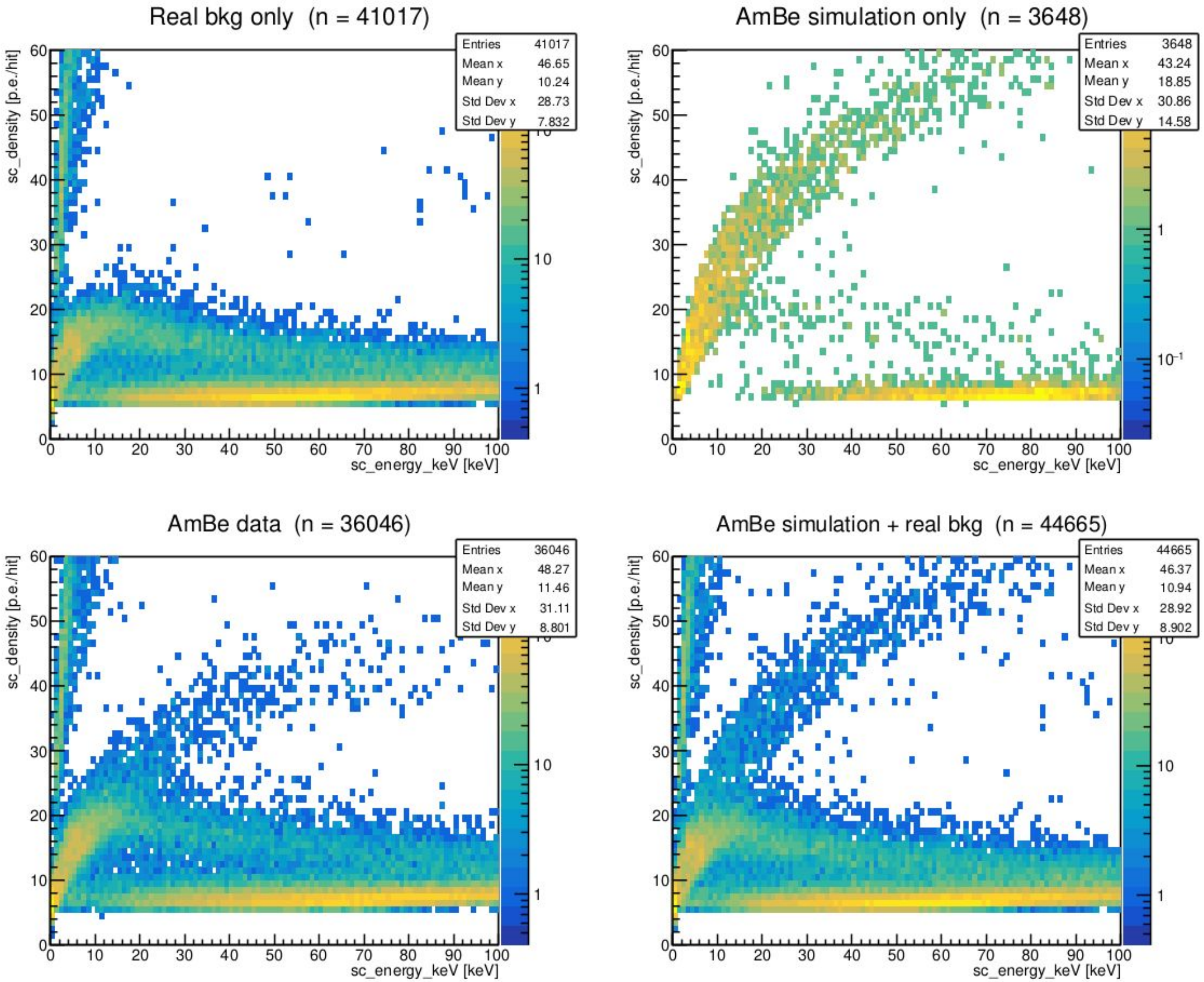


sc_nhits
Numero di cluster
ER (N=6890)
NR He (N=831)
NR C (N=546)
NR F (N=3004)
10^2
10
1
0
1000
2000
3000
4000
5000
6000
7000
sc_nhits

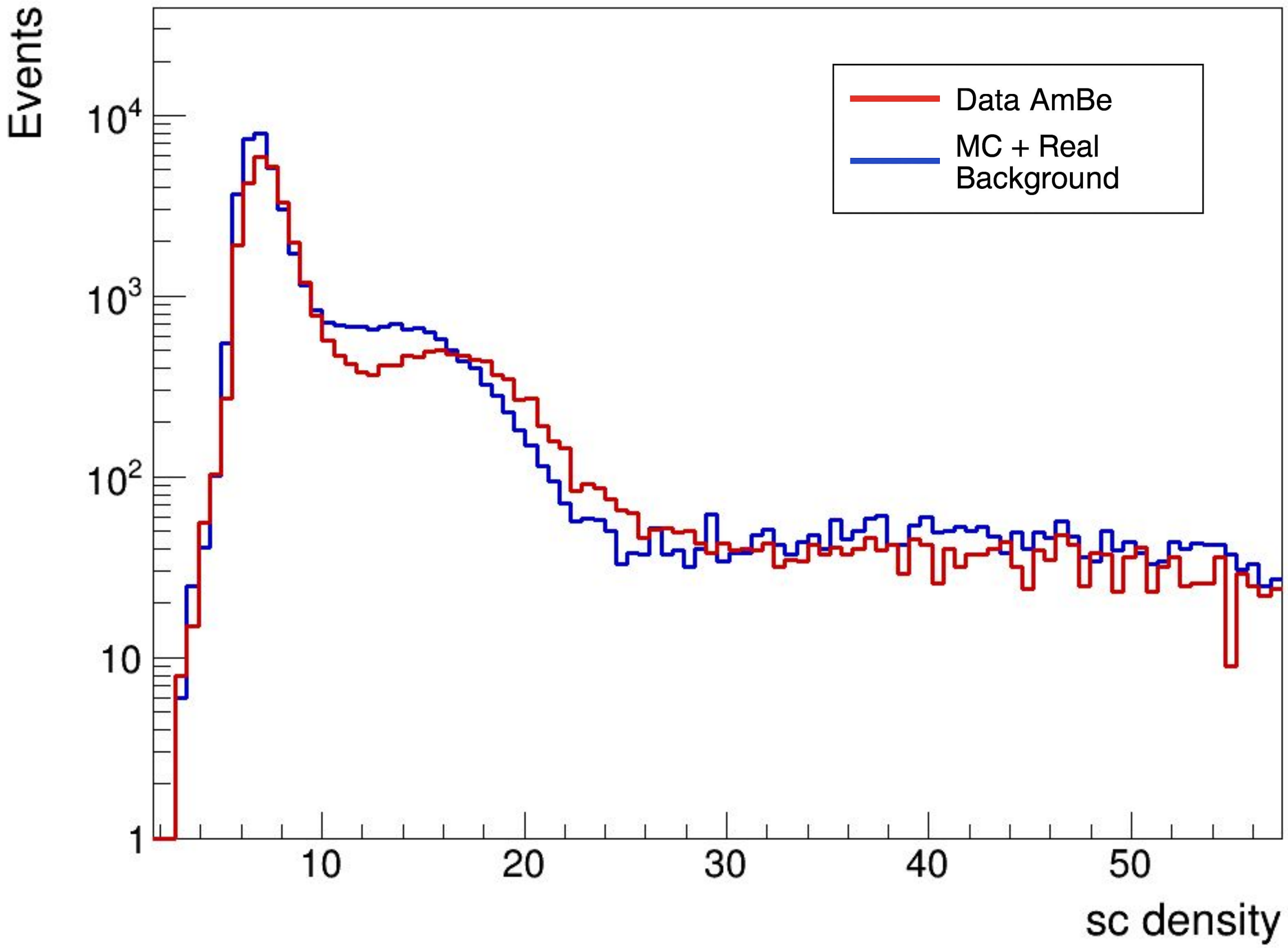
Geant4+digitization+reconstruction



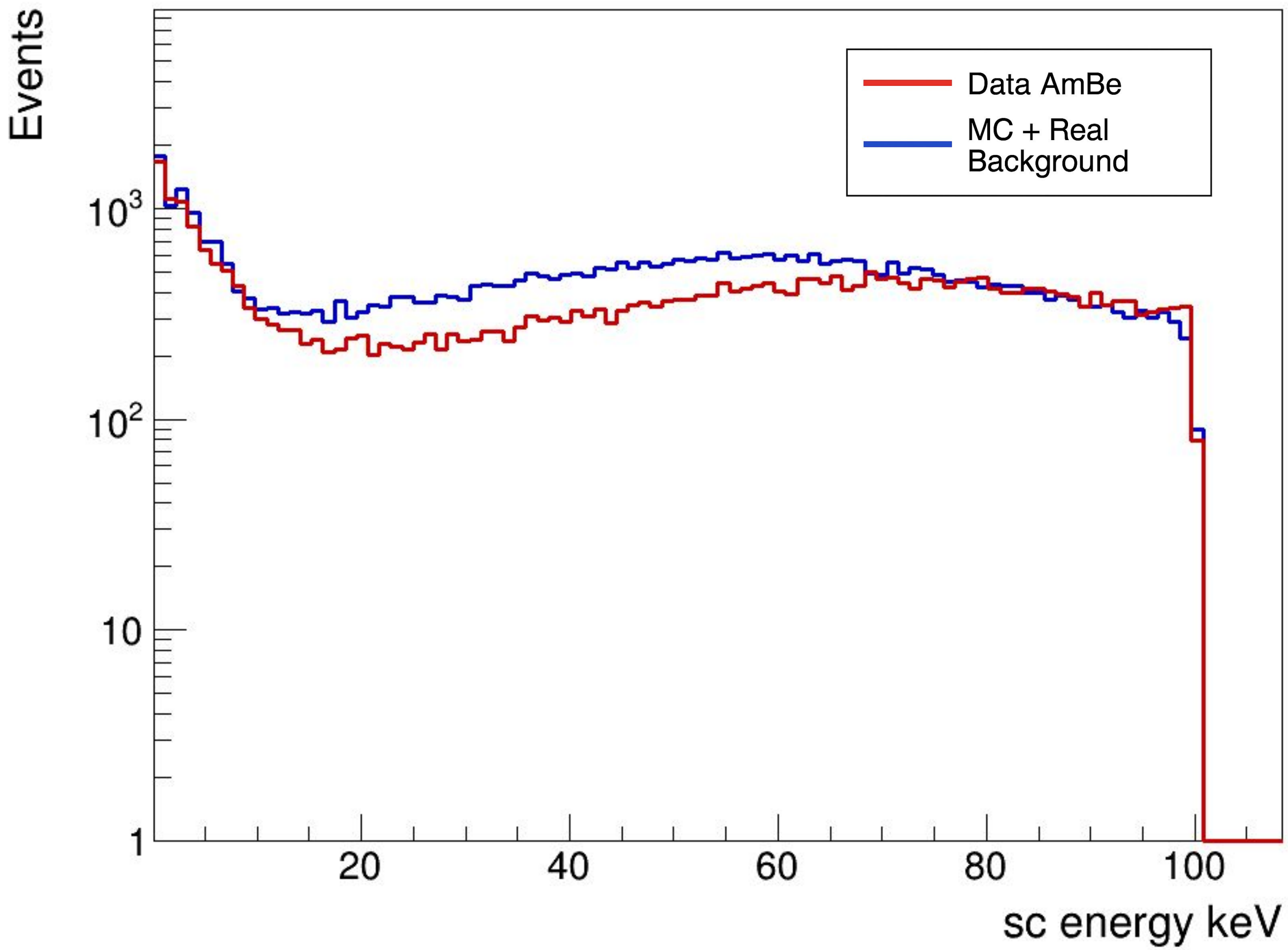
Data/MC comparison RUN 3



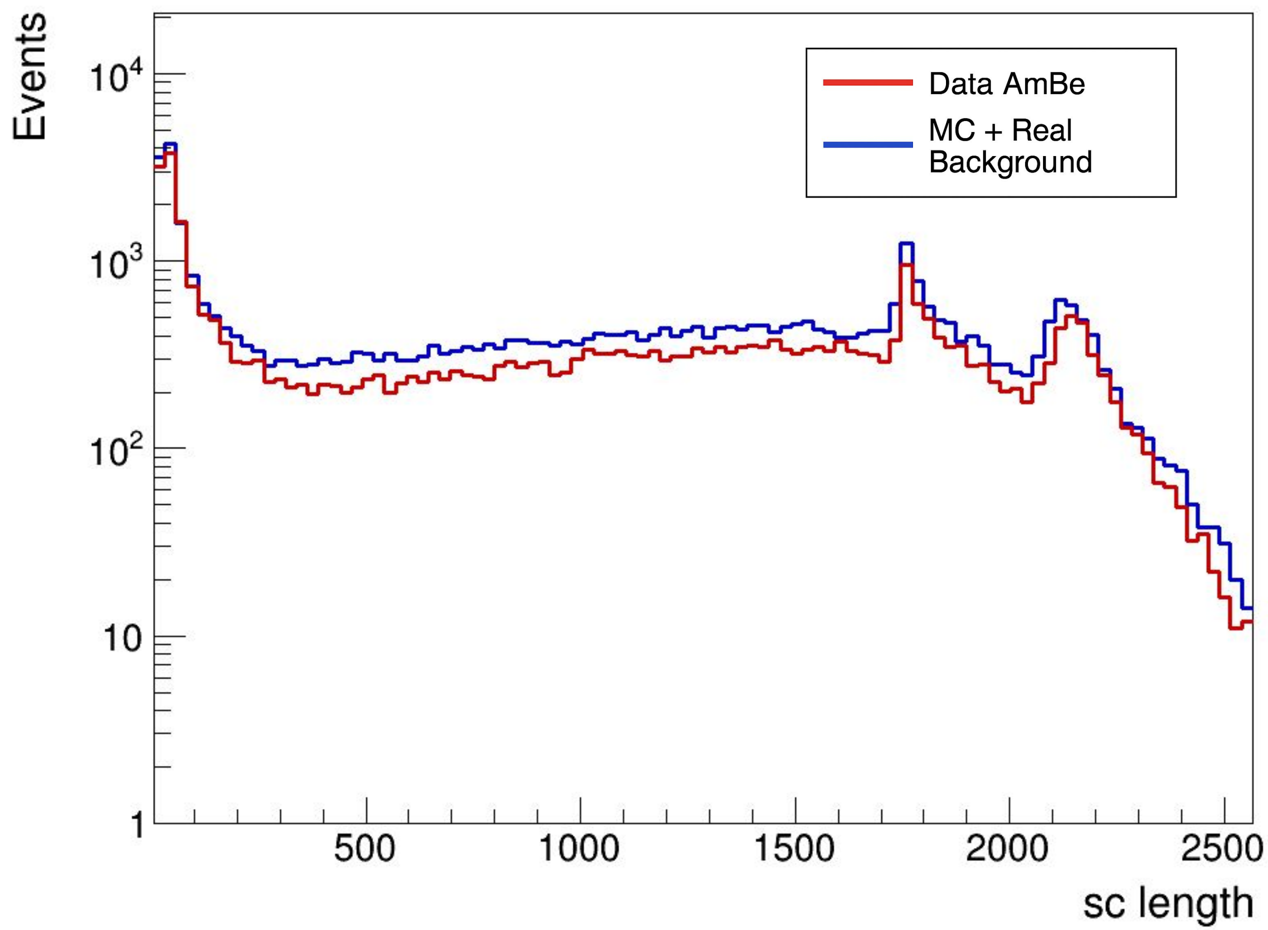
In RUN 3, starting from a trigger rate of 1.6 Hz without the source and 1.8 Hz with the AmBe source, a net excess of 0.2 Hz (i.e., two events every 10 seconds) is observed, which can be attributed to neutrons since the gamma contribution from AmBe is negligible; given that the source emits 200 neutrons per second, this means only 1 in 1000 is detected—corresponding to the product of the geometrical acceptance and the DAQ efficiency—where GEANT4 simulations show that 3 out of 1000 generated neutrons interact in the gas, implying a DAQ efficiency of about 33%, and thus $0.003 \times 0.33 \approx 0.001$, so to simulate one hour of AmBe exposure one should generate (3600 s) × (200 n/s) × (0.33) = 237,600 neutrons, without multiplying by the geometrical acceptance again since GEANT4 already includes the full geometry.



Events
10^4
10^3
10^2
10
1
10
20
30
40
50
sc density
Data AmBe
MC + Real
Background



Events
Data AmBe
MC + Real Background
10^3
10^2
10
1
20
40
60
80
100
sc energy keV

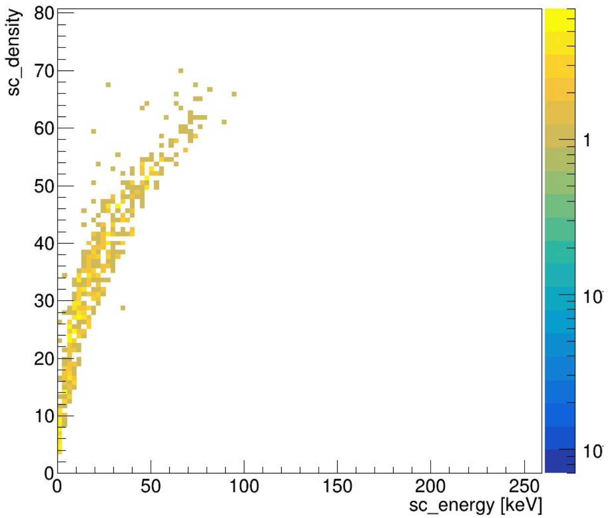


Events
10^4
10^3
10^2
10
1
500
1000
1500
2000
2500
sc length
Data AmBe
MC + Real Background

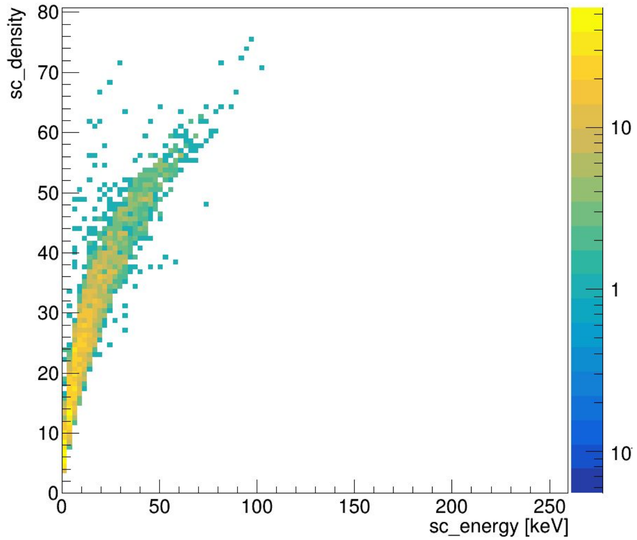
How He, C, and F differ from each other?

Color map is "<u>number of clusters</u>"

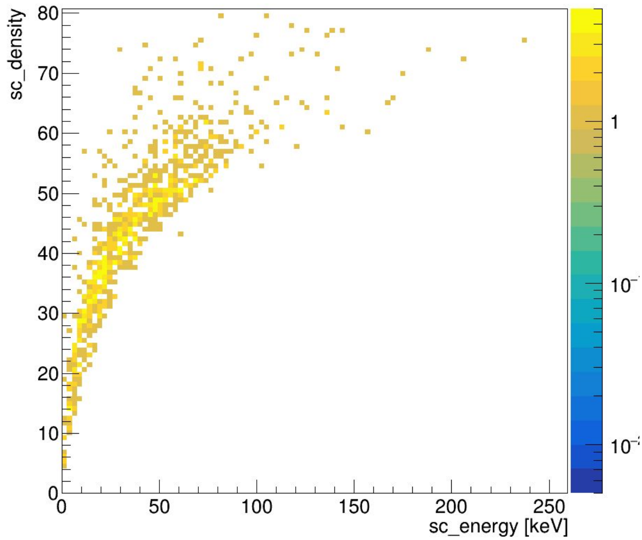
Carbon



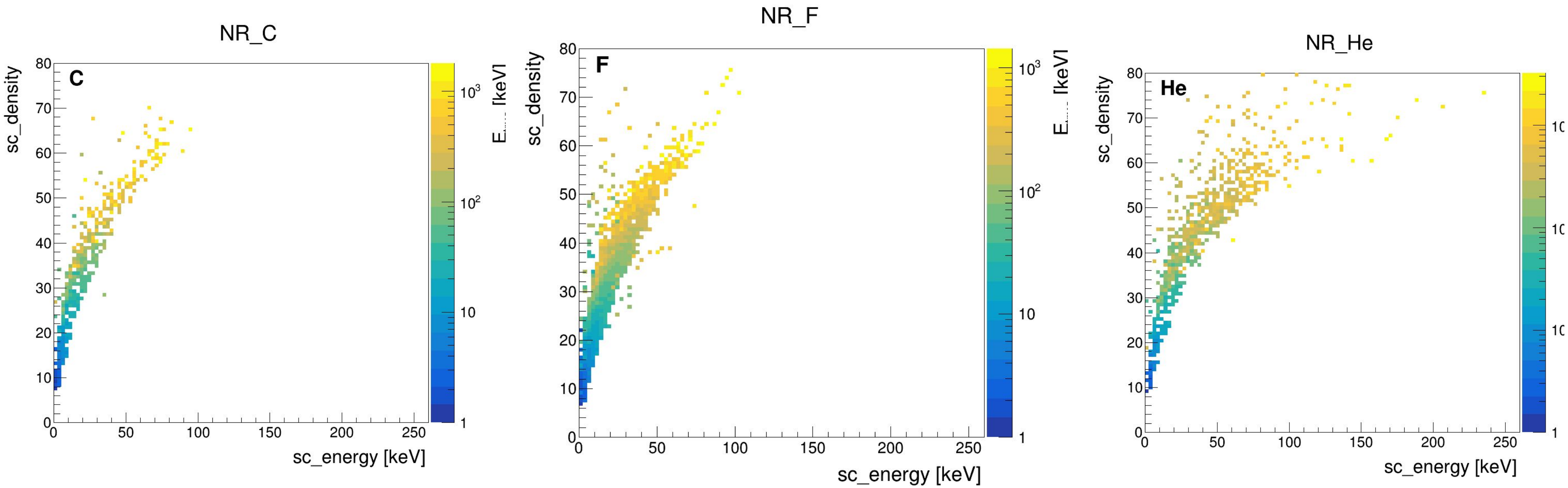
Fluorine



Helium



Color map is "real energy in MC truth"



Simple ER/NR discrimination

Simple cuts
(based on previous distributions)

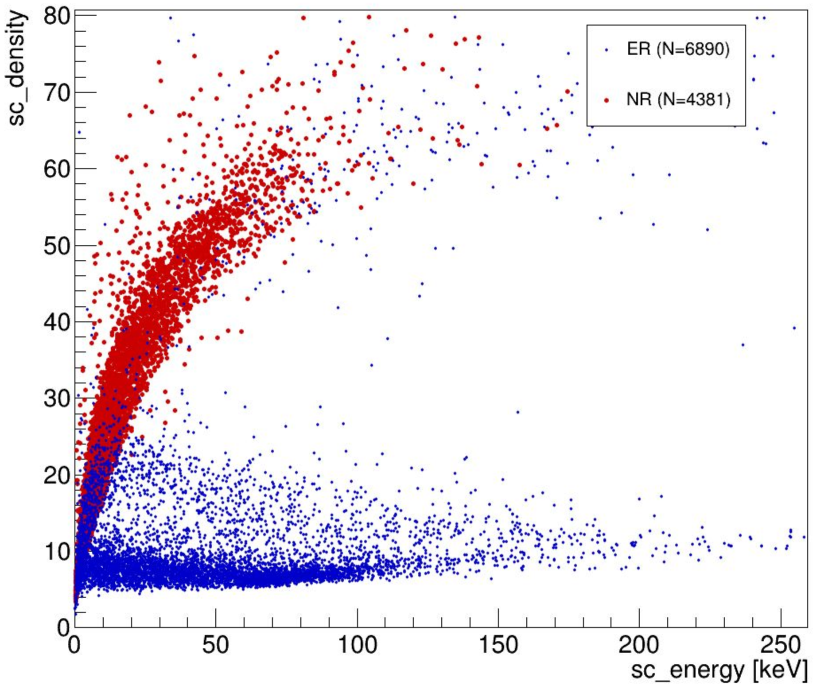
```
eta < 150
rho > 0.01
sc_integral < 300
sc_length < 200
sc_nhits < 5000
sc_tgausssigma < 10
sc_width < 100
slimness < 6
```

Note: ER and NR are all from AmBe source. There is no LIME background

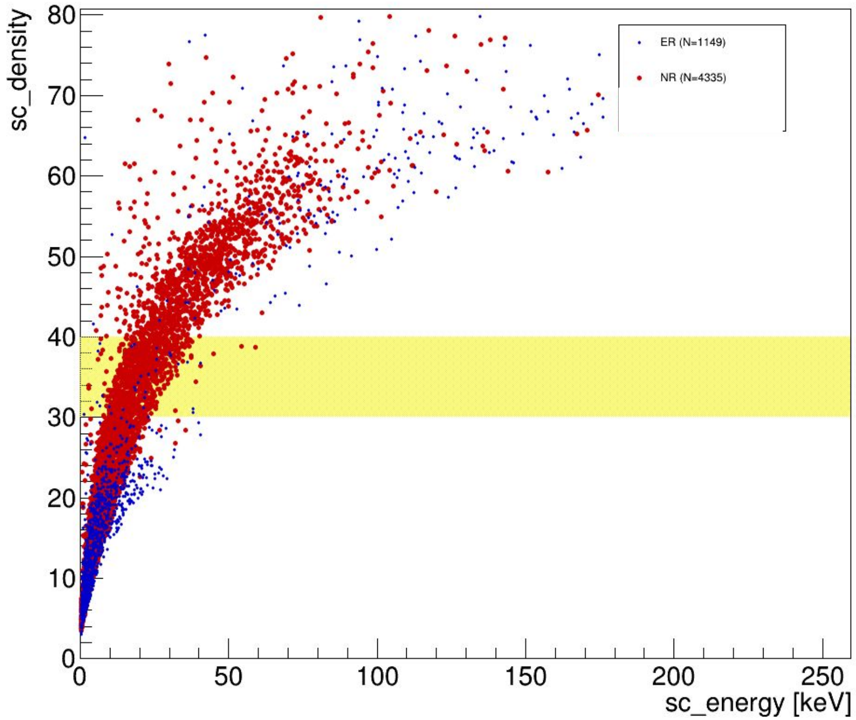
Next slide visual result

```
======================================================================
STATISTICHE SUI TAGLI
======================================================================

PRIMA DEI TAGLI:
  Background (ER): 6890
  Segnale (NR):    4381
  Totale:          11271

DOPO I TAGLI:
  Background (ER): 1149
  Segnale (NR):    4335
  Totale:          5484

EVENTI PERSI:
  Background (ER): 5741 (83.32%)
  Segnale (NR):    46 (1.05%)

📊 EFFICIENZA SEGNALE: 98.95% (4335/4381)
📊 REIEZIONE BACKGROUND: 83.32% (5741/6890)
📊 FATTORE DI REIEZIONE: 6.00x

DETTAGLIO PER TIPO:
  ER     :  6890 →  1149  (sopravvivenza: 16.68%)
  NR_He  :   831 →   802  (sopravvivenza: 96.51%)
  NR_C   :   546 →   545  (sopravvivenza: 99.82%)
  NR_F   :  3004 →  2988  (sopravvivenza: 99.47%)

======================================================================
EVENTI RIGETTATI DA OGNI SINGOLO TAGLIO
======================================================================

Taglio               ER rigettati    NR rigettati
--------------------------------------------------
eta                  399     (5.8%)  39      (0.9%)
rho                  5677    (82.4%) 3       (0.1%)
sc_integral          216     (3.1%)  8       (0.2%)
sc_length            5022    (72.9%) 0       (0.0%)
sc_nhits             4147    (60.2%) 1       (0.0%)
sc_tgausssigma       2643    (38.4%) 1       (0.0%)
sc_width             1904    (27.6%) 0       (0.0%)
slimness             3978    (57.7%) 0       (0.0%)
```

Before cuts



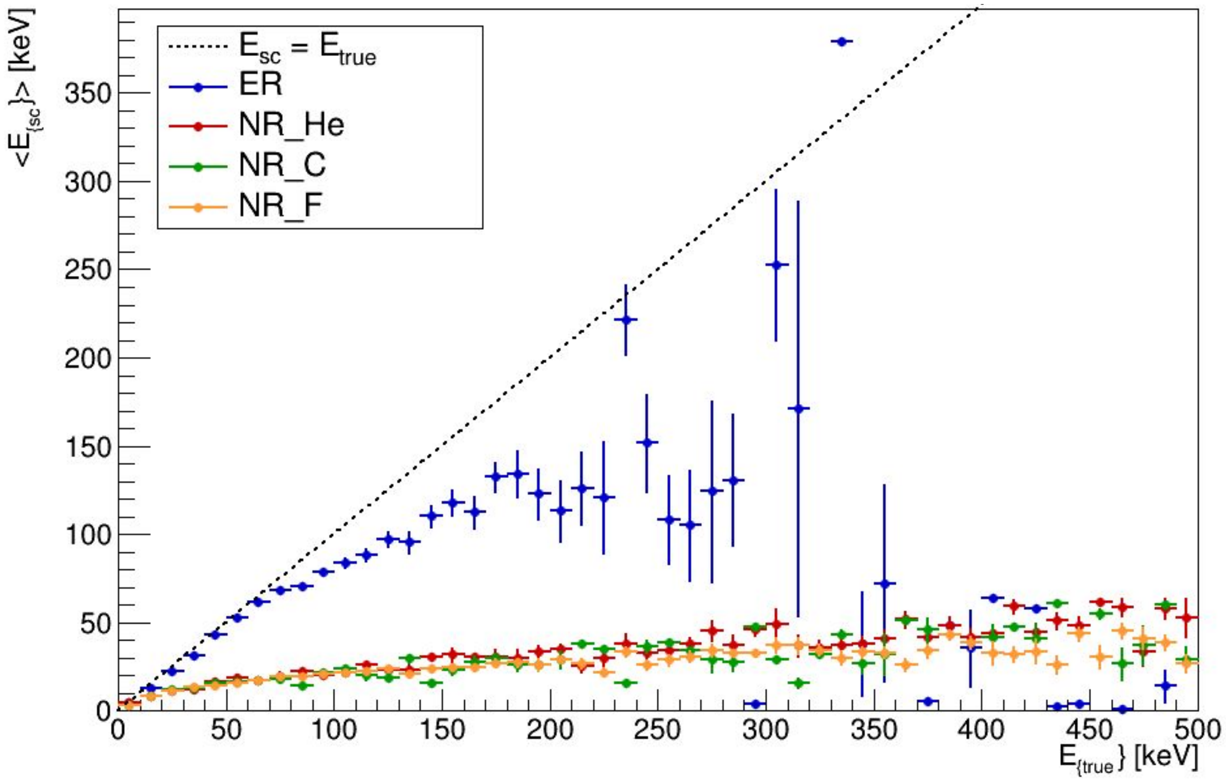
After cuts



Yellow region has no meaning. To fix

How does the QF affect NR energy?

Energy in MC truth vs reconstructed energy (10000 counts = 6 keV)



You can see how ER start to not be linear when they are too long (cut tracks), while NR due to QF

Next steps

- improve normalization for MC/data comparison
- find better cuts for ER/NR discrimination (looking at Atul's work)
- compare 3D reconstructed length (proportional to energy) for NRs