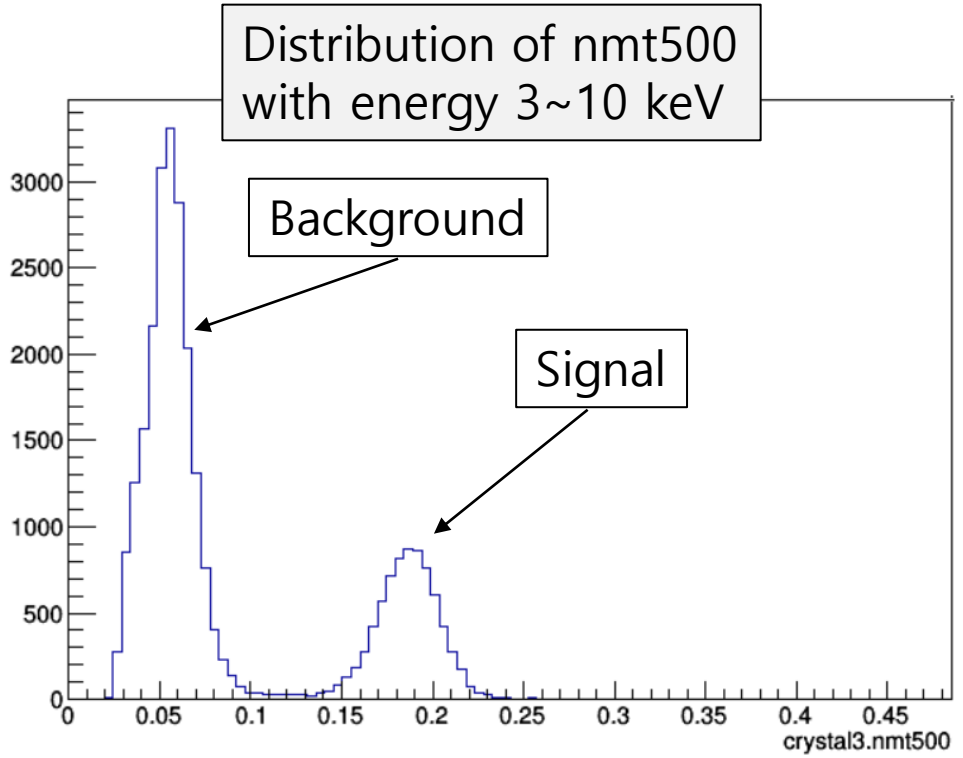
CNN For High Energy - Validation -

2020.03.24

Seungmok Lee

Training Data Selection

- Training data was selected using criteria below.
  - Signal tag
    - Runnum 1765, 1601 (Calibration run)
    - Have coincidence
    - Nmt500 > 0.12
  - Background tag
    - Runnum 1858, 1859 (Physics run)
    - No coincidence
    - Nmt500 < 0.12



Data Statistics

		Training data		Validation data		Test data	
		Background	Signal	Background	Signal	Background	Signal
Crystal 1	3~6 keV	146,089	414	18,264	47	18,273	38
	6~10 keV	62,381	452	7,805	48	7,806	47
Crystal 3	3~6 keV	142,487	1,367	17,815	164	17,792	187
	6~10 keV	50,477	1,909	6,322	226	6,326	222
Crystal 7	3~6 keV	146,834	1,892	18,365	224	18,371	218
	6~10 keV	63,592	2,761	7,959	333	7,980	312

CNN Structure

- 1dimensional CNN motivated by AlexNet was successful!
- Batch norm., ReLU activation applied.

Input shape	Layer	Output shape
(4080, 2)	Conv1D(filters=96, kernel_size=39, strides=3)	(1348, 96)
(1348, 96)	MaxPool1D(pool_size=4)	(337, 96)
(337, 96)	Conv1D(filters=256, kernel_size=23)	(315, 256)
(315, 256)	MaxPool1D(pool_size=3)	(105, 256)
(105, 256)	Conv1D(filters=384, kernel_size=9)	(97, 384)
(97, 384)	Conv1D(filters=384, kernel_size=9)	(89, 384)
(89, 384)	Conv1D(filters=256, kernel_size=9)	(81, 256)
(81, 256)	MaxPool1D(pool_size=3, strides=2)	(40, 256)
(40, 256)	Flatten()	10240
10240	Dense(512)	512
512	Dense(512)	512
512	Dense(512)	512
512	Dense(512)	512
512	Dense(2)	2

Compatibility

Trained Using		Tested on		Signal Eff. [%]	Background Eff. [%]	F1 Score [%]
Crystal	Energy [keV]	Crystal	Energy [keV]			
3	6~10	3	3~6	100	95.48	**97.69**
3	3~6	3	6~10	100	99.21	**99.60**
7	6~10	3	6~10	100	99.18	**99.59**
7	6~10	3	3~6	100	94.22	**97.03**
7	3~6	3	6~10	100	100	**100**
7	3~6	1	6~10	100	99.69	**99.84**
7	3~6	1	3~6	100	99.64	**99.82**
3	6~10	1	6~10	100	99.69	**99.84**
1	3~6	7	6~10	99.68	100	**99.84**

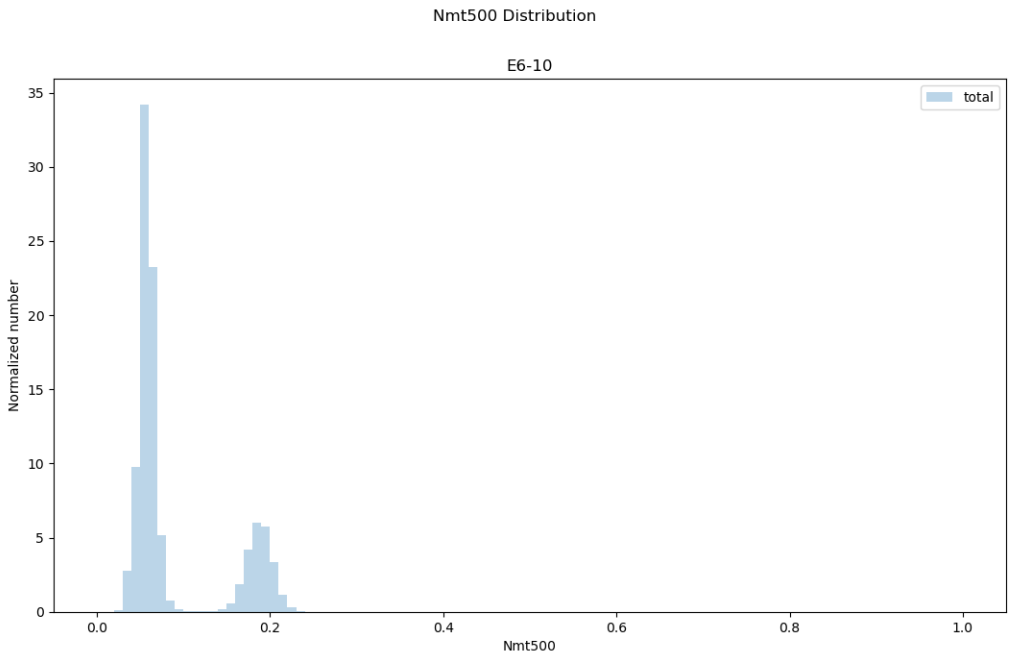
※ $\mathrm{F}_1\ \mathrm{score} = \dfrac{2}{\dfrac{1}{\text{Signal eff.}} + \dfrac{1}{\text{Background eff.}}}$

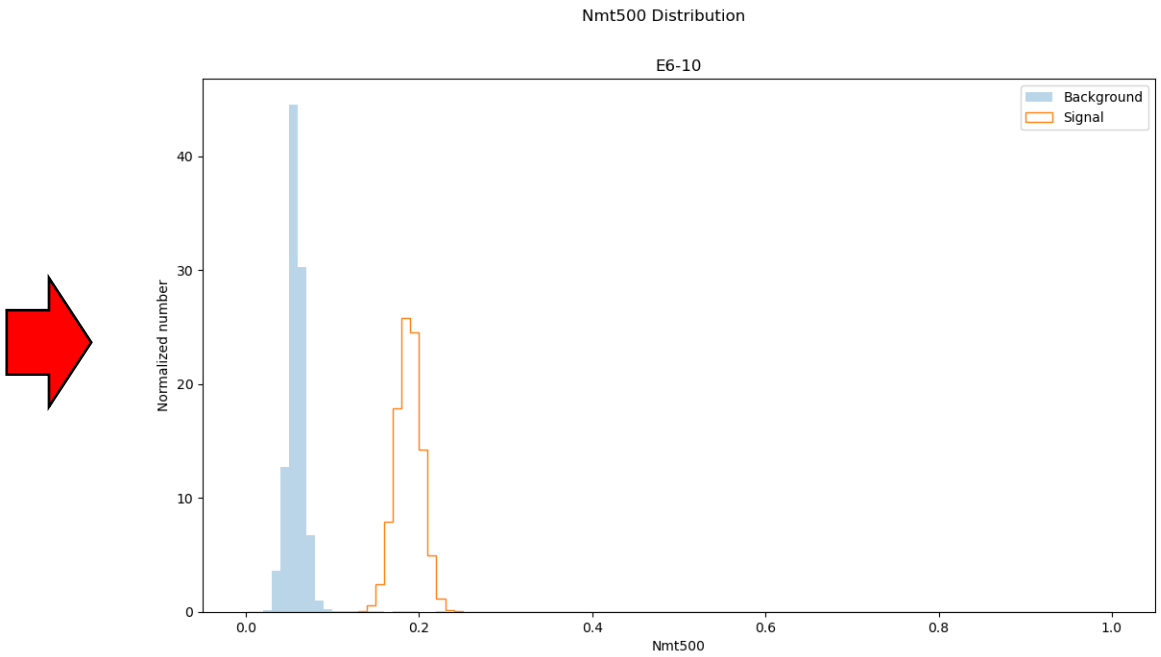
Testing Data Selection

- Testing data was selected without nmt cut.
  - Runnum 1765, 1601 (Calibration run)
    - Have coincidence
  - Runnum 1858, 1859 (Physics run)
    - No coincidence
  - Not tagged. Crystal3 was used.

Nmt500 Splitting Plot

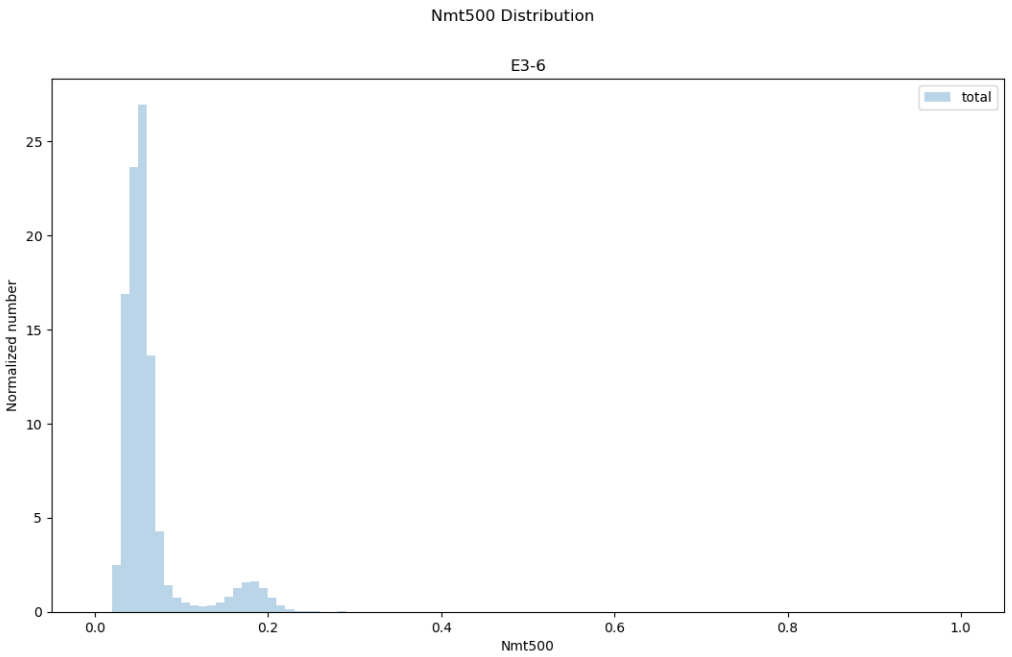
- Trained model could successfully split data without nmt cut.
  - Model was trained using crystal 7, energy 3~6 keV data.

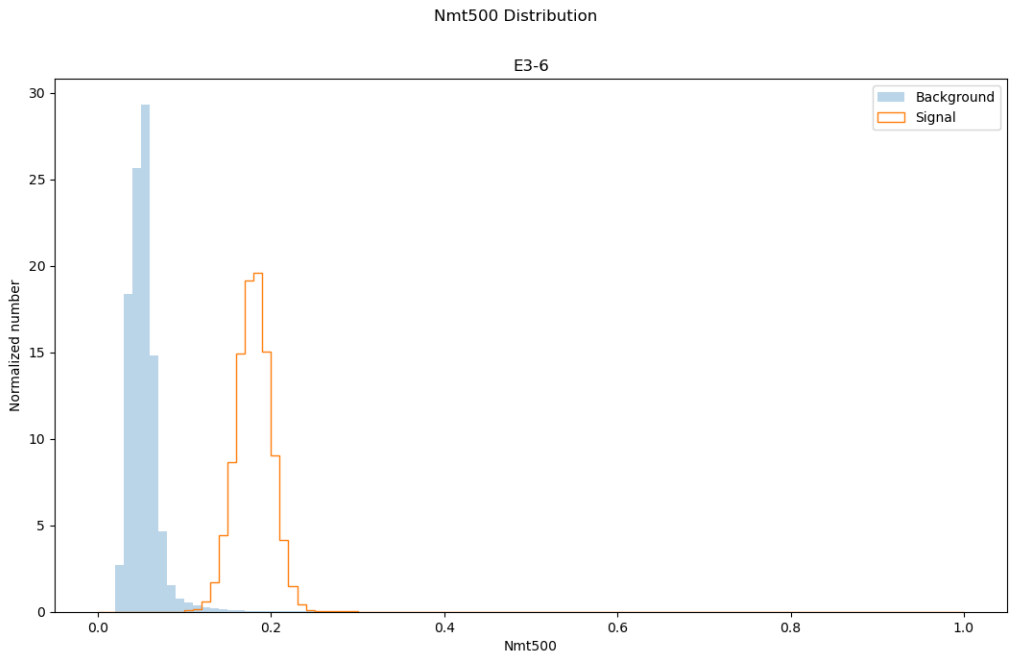




Nmt500 Splitting Plot

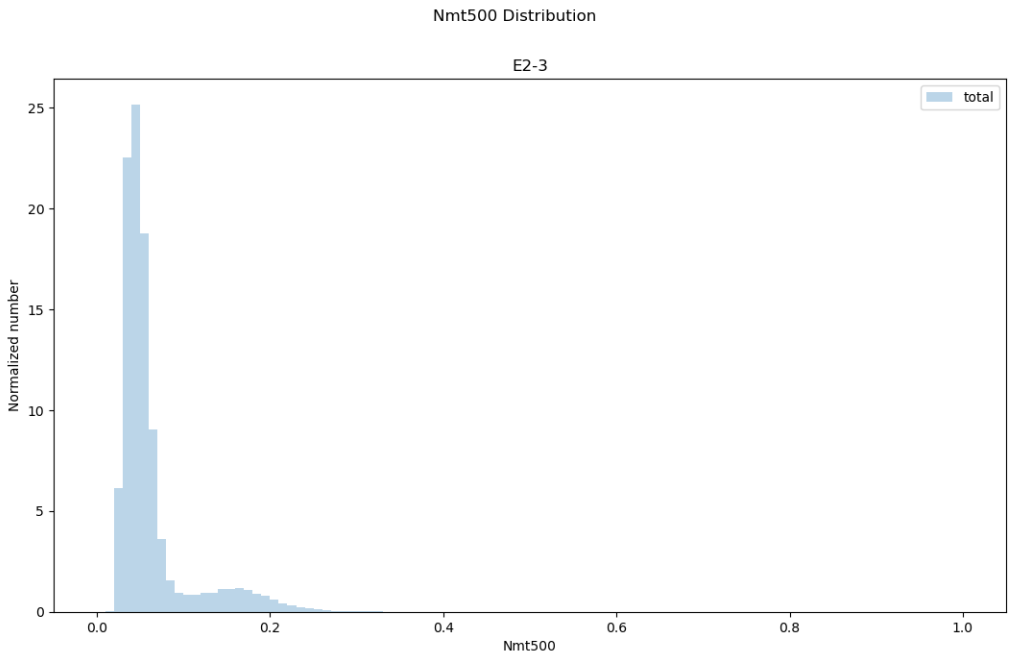
- 3~6 keV data showed very clear nmt distribution.
- It seems **not nmt-biased**.

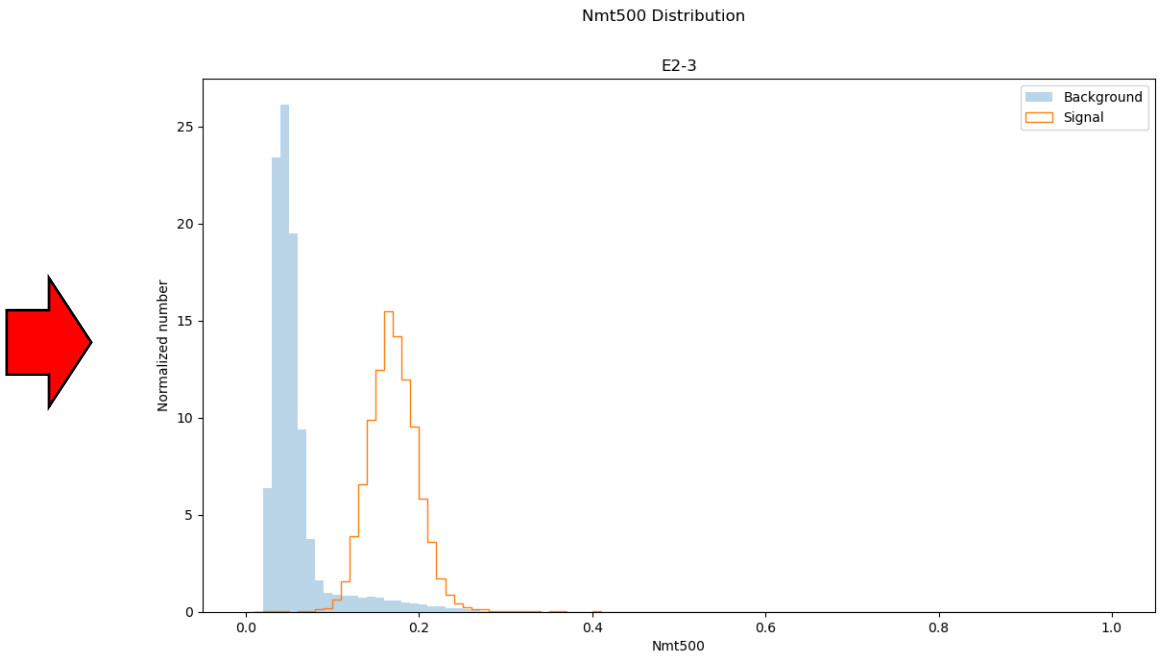




Nmt500 Splitting Plot

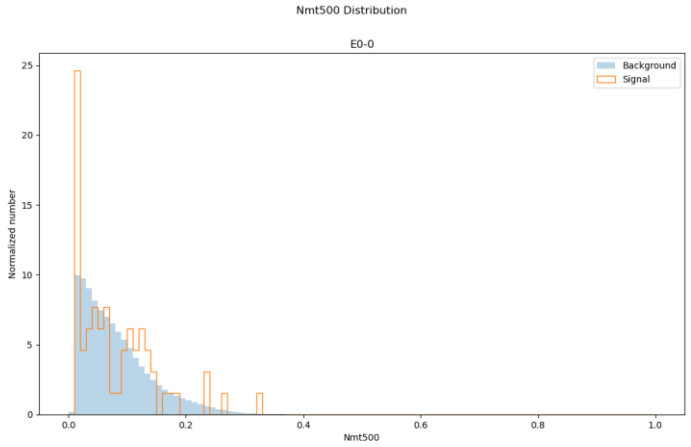
- 2~3 keV data showed clear nmt distribution, too.



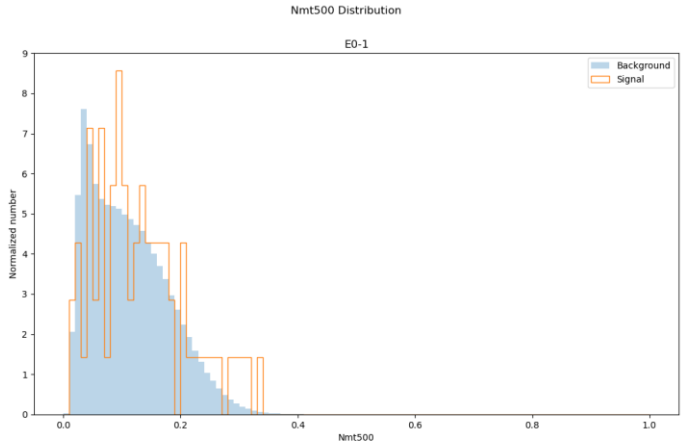


Nmt500 Splitting Plot

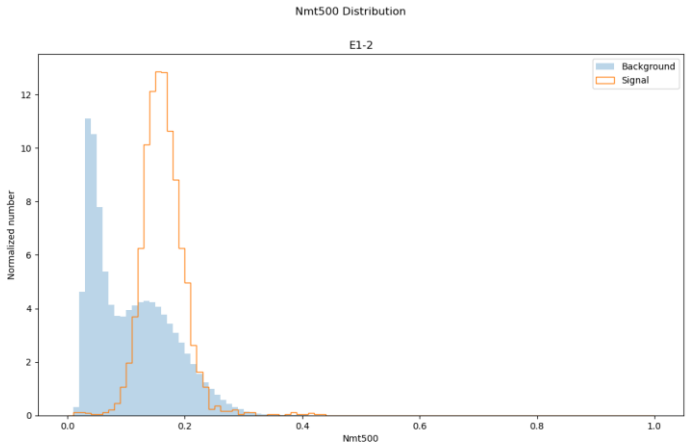
- Nmt distribution was not that clear in low energy.
  - Nmt is known to be unavailable in this low energy, so these are not intuitive.



E 0~0.5 keV



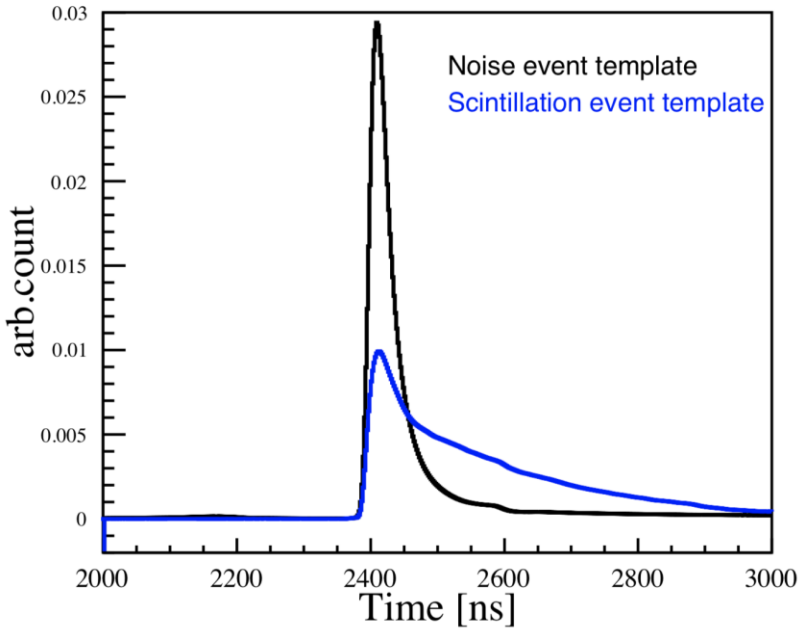
E 0.5~1 keV



E 1~2 keV

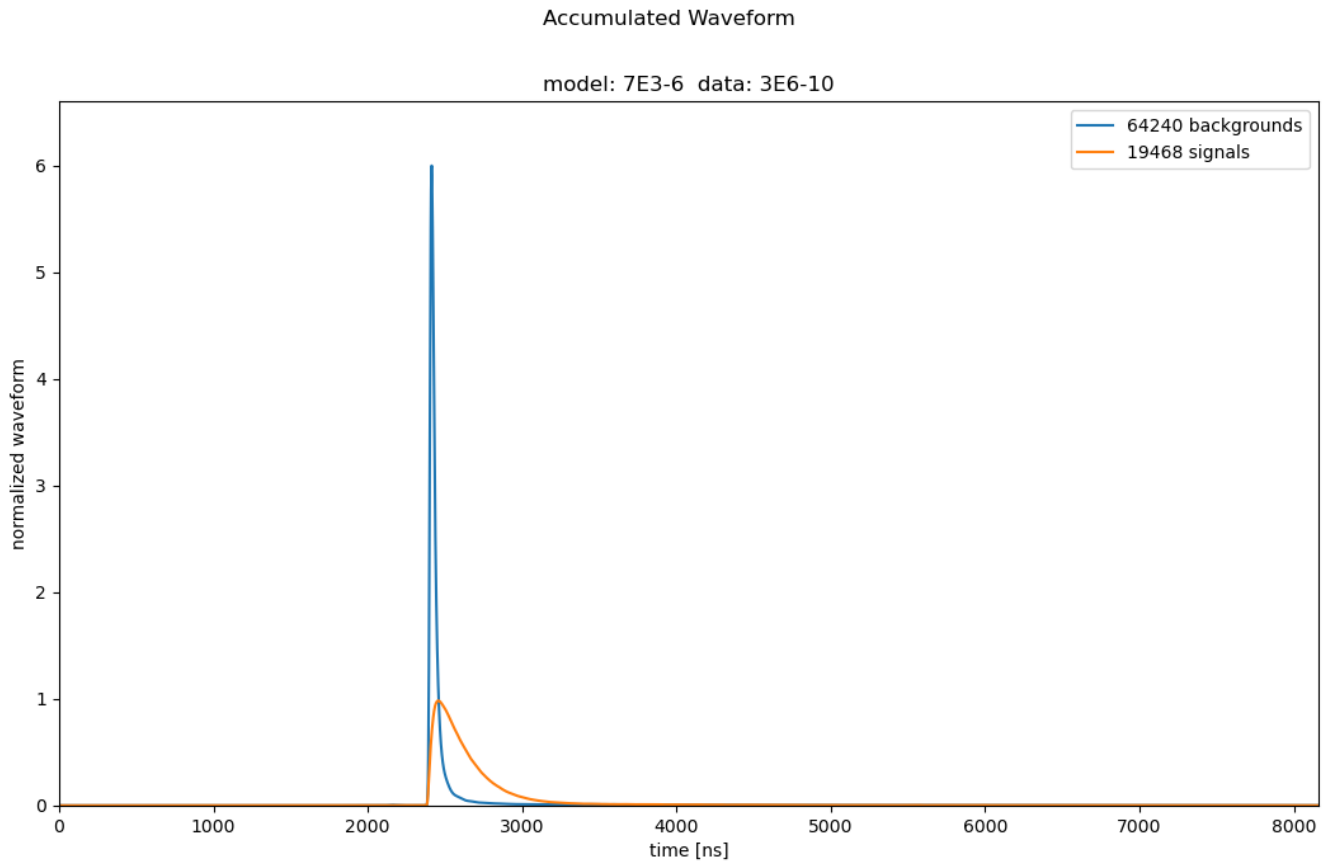
Validation – Accumulated Waveform

- Signal wave has typically longer waveform.
- Accumulating the waveform helps to visualize overall property.
- Figures after this slide were generated by model trained using crystal 7 energy 3-6 keV data.



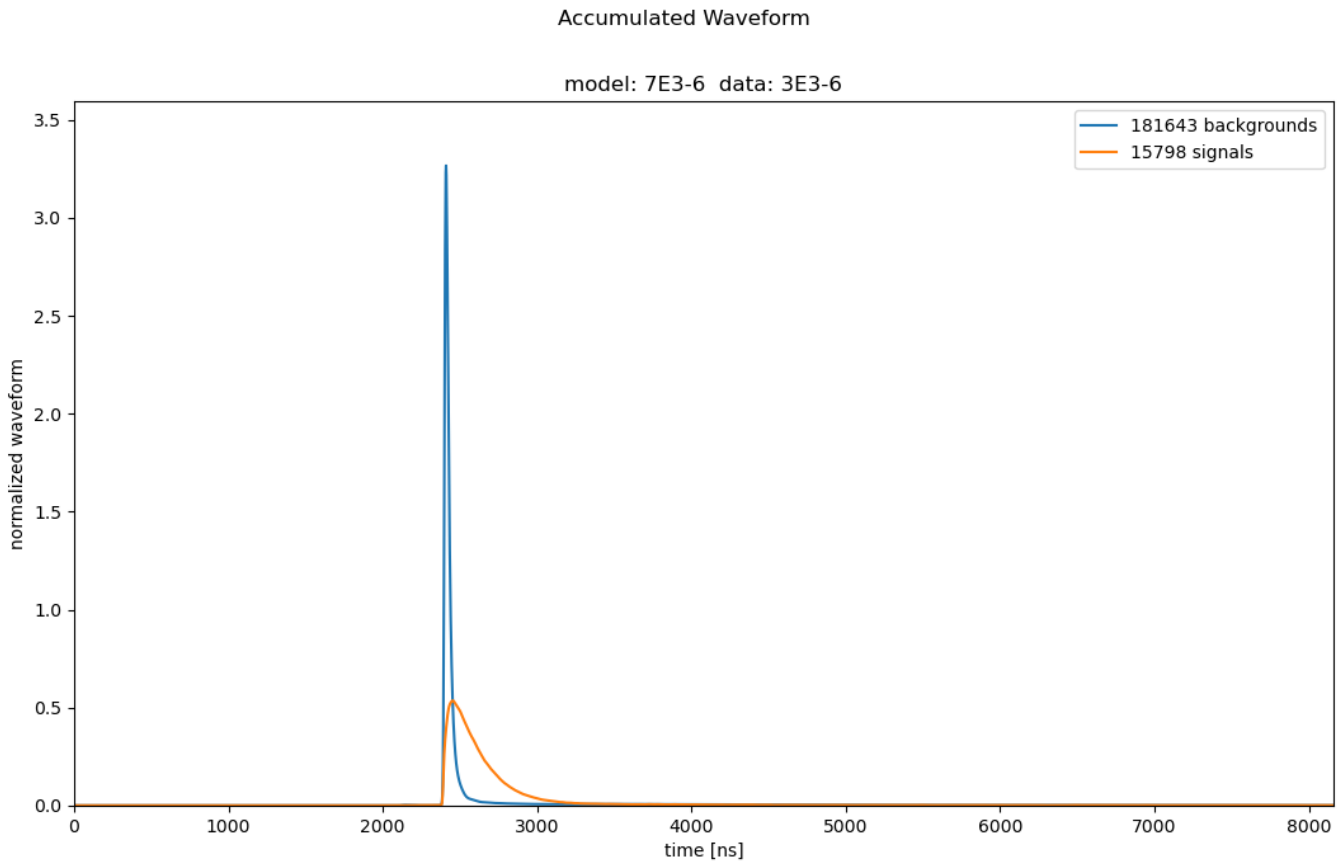
Validation – Accumulated Waveform

Crystal 3, energy 6-10 keV events were classified successfully.



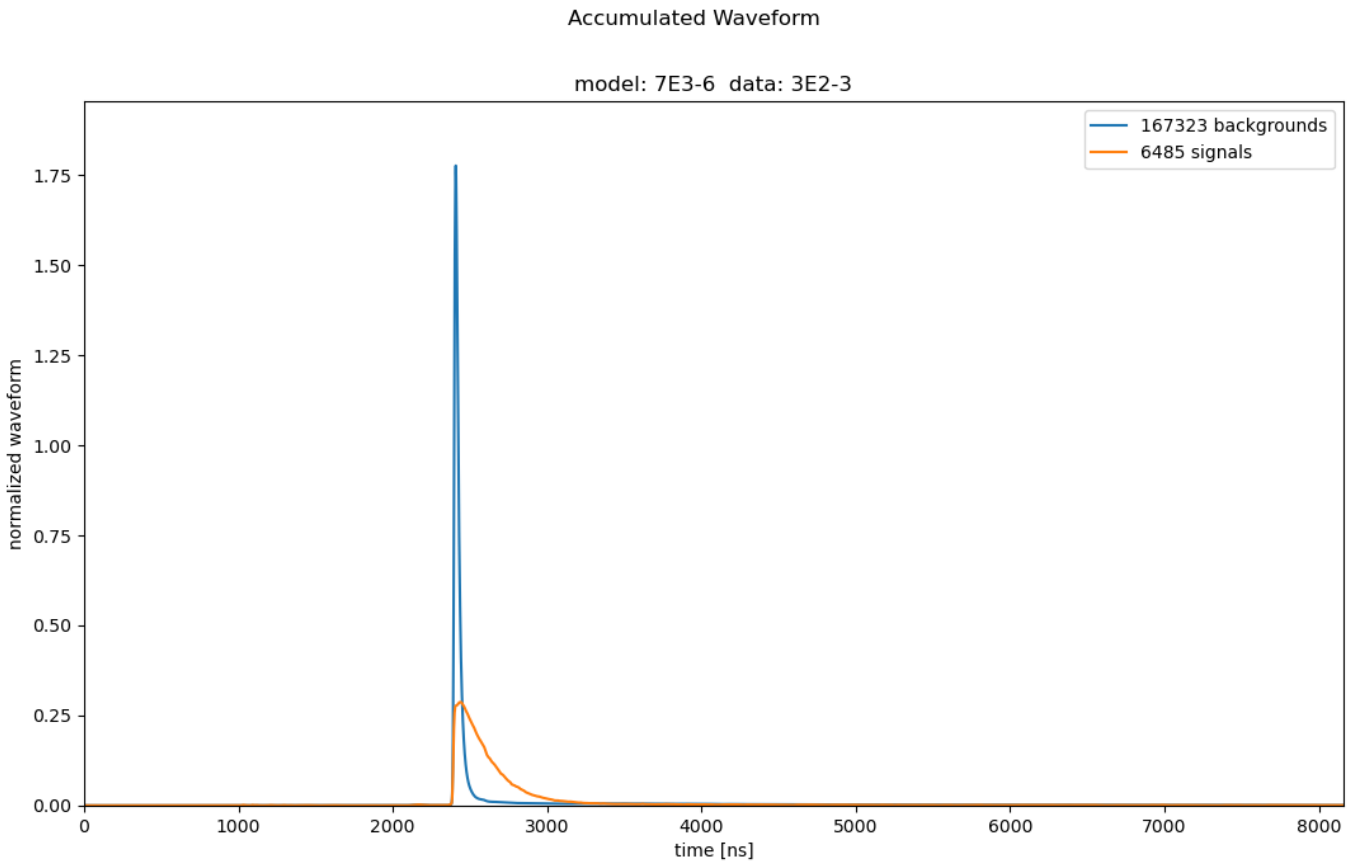
Validation – Accumulated Waveform

Crystal 3, energy 3-6 keV events were classified successfully, too.



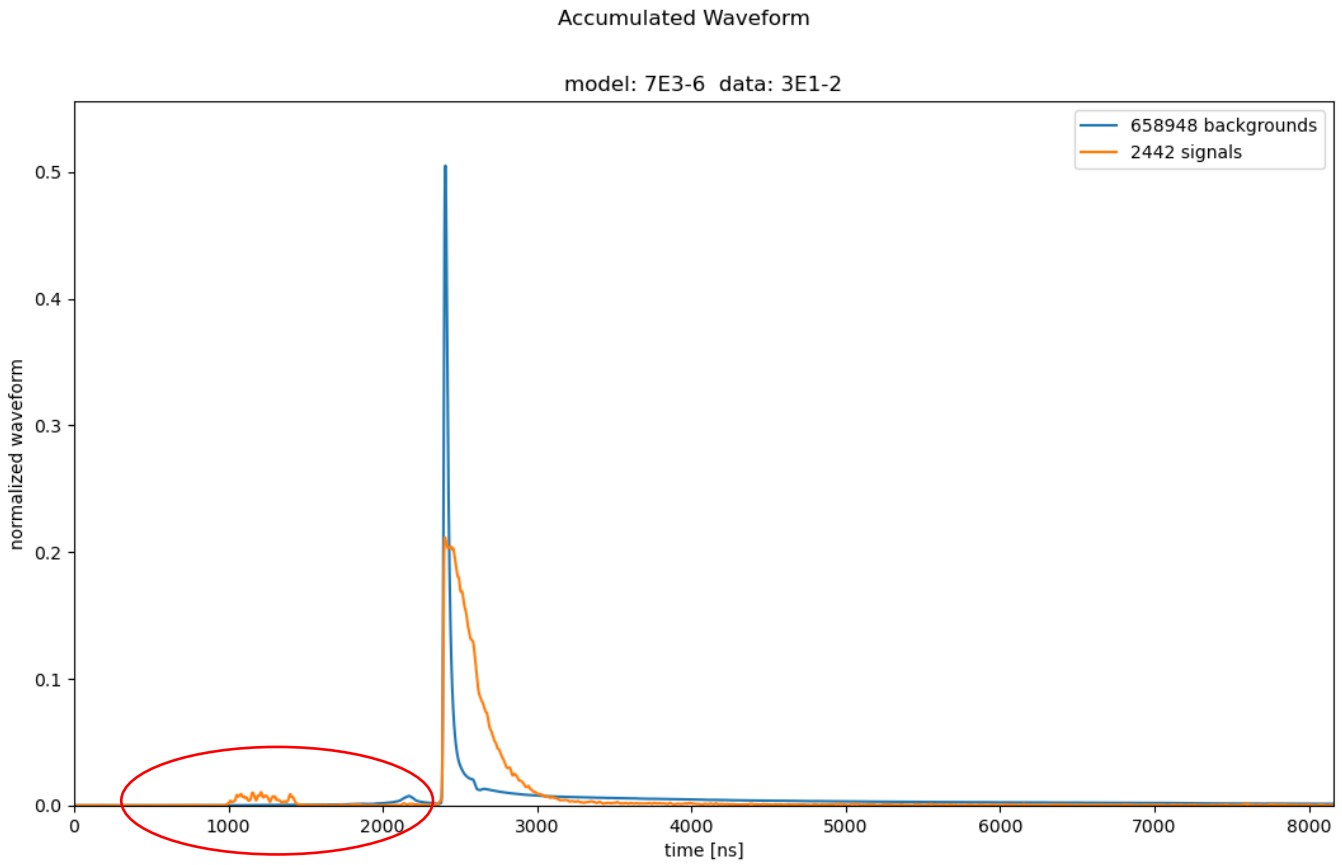
Validation – Accumulated Waveform

Crystal 3, energy 2-3 keV events were classified successfully, too.



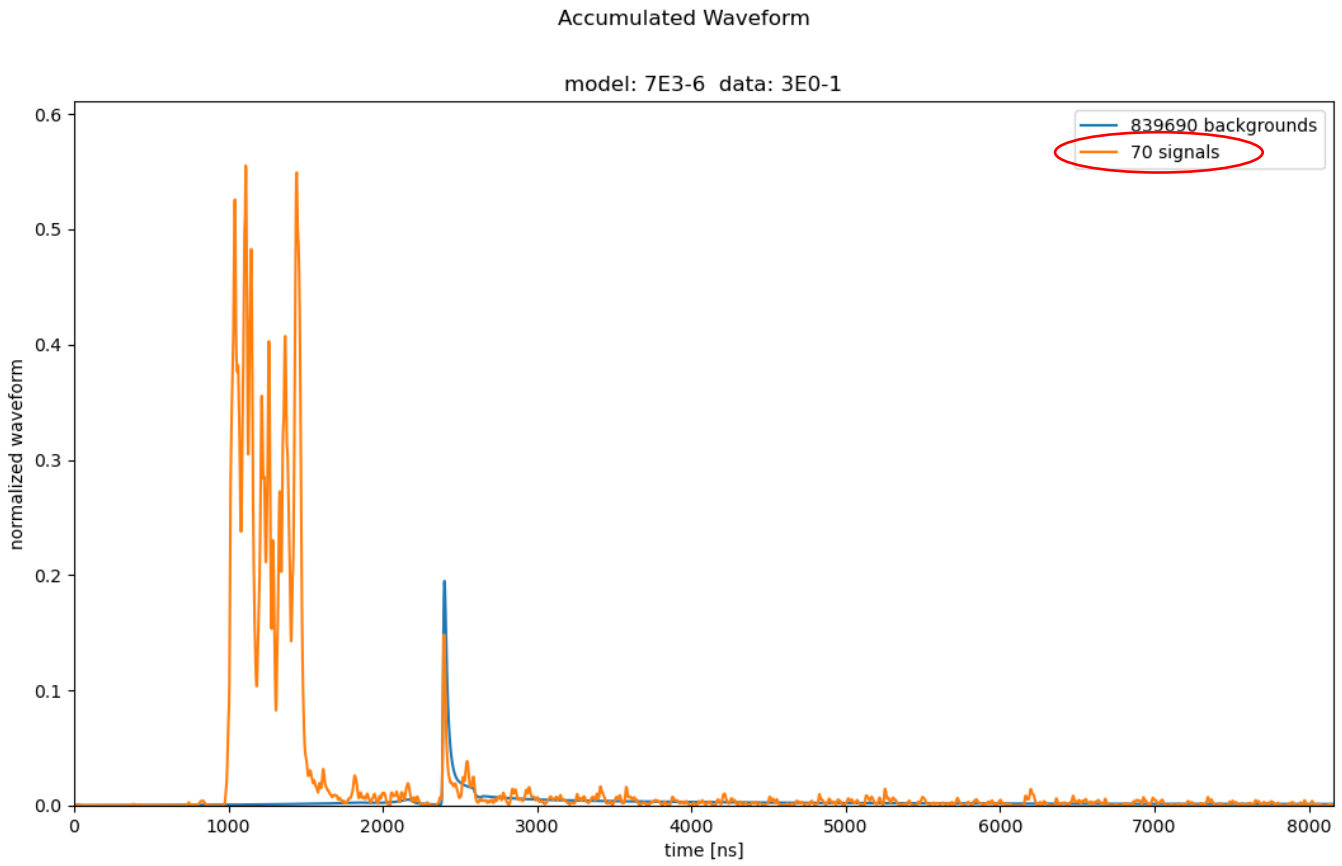
Validation – Accumulated Waveform

Crystal 3, energy 1-2 keV events classification result seems not bad, but 'something' arises at the front side.



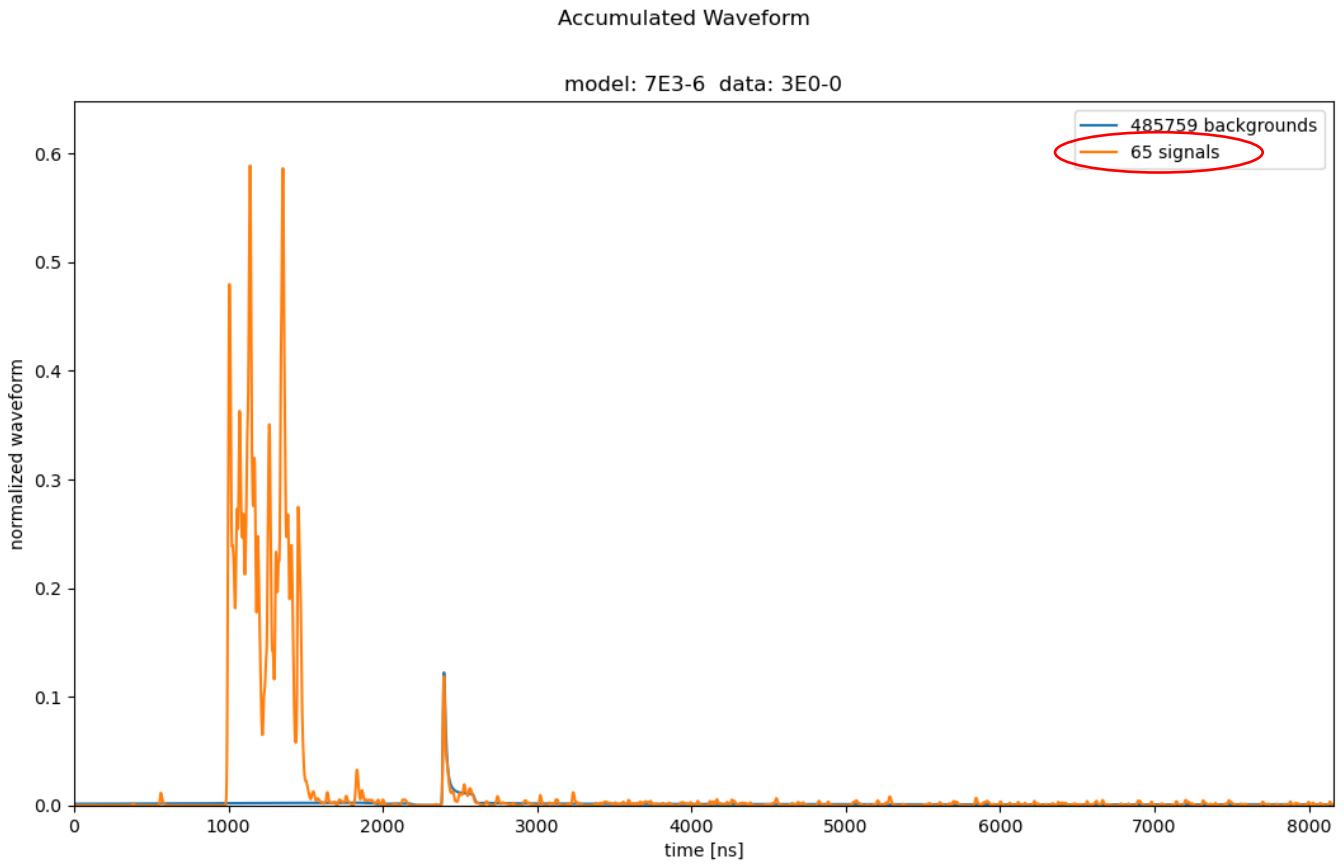
Validation – Accumulated Waveform

Crystal 3, energy 0.5-1 keV events classification seems not that good. However, signal-classified event number is too small.



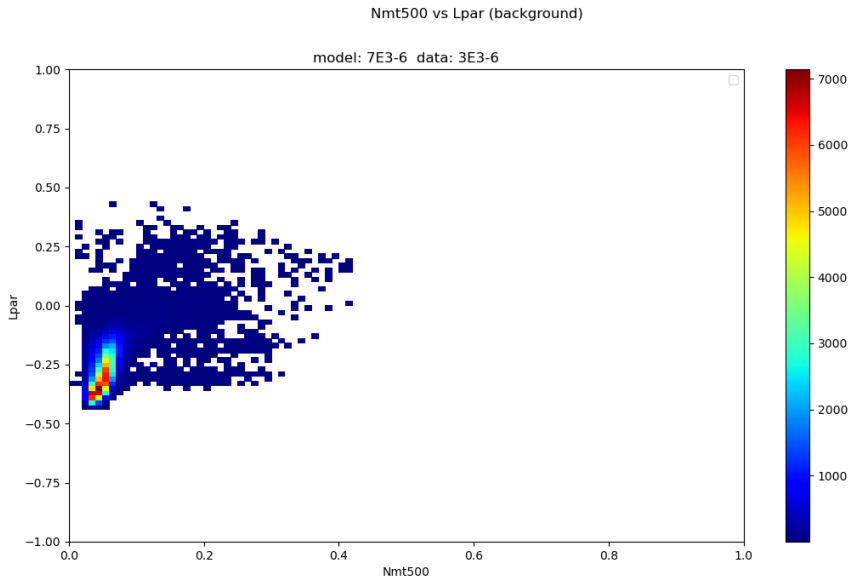
Validation – Accumulated Waveform

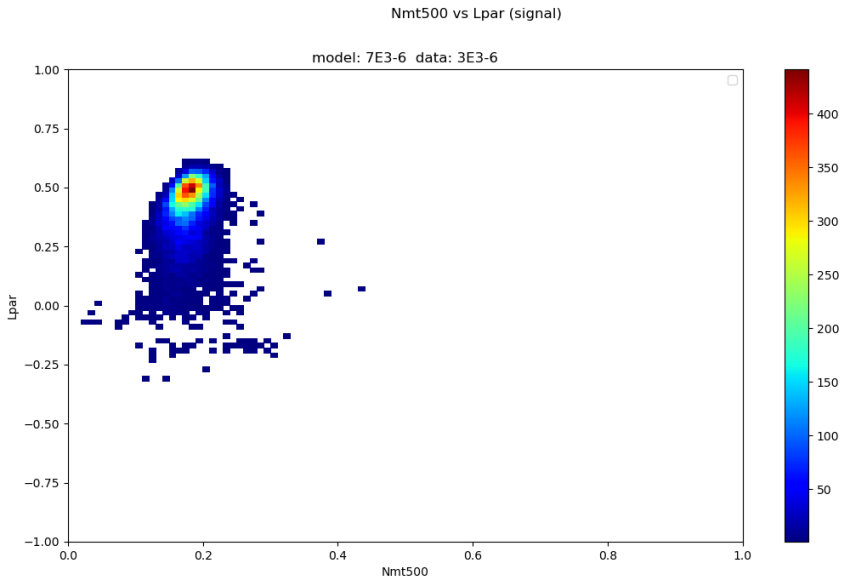
Crystal 3, energy 0-0.5 keV events classification seems also not that good. However, signal-classified event number is too small.



Validation – lpar:nmt500 Plot

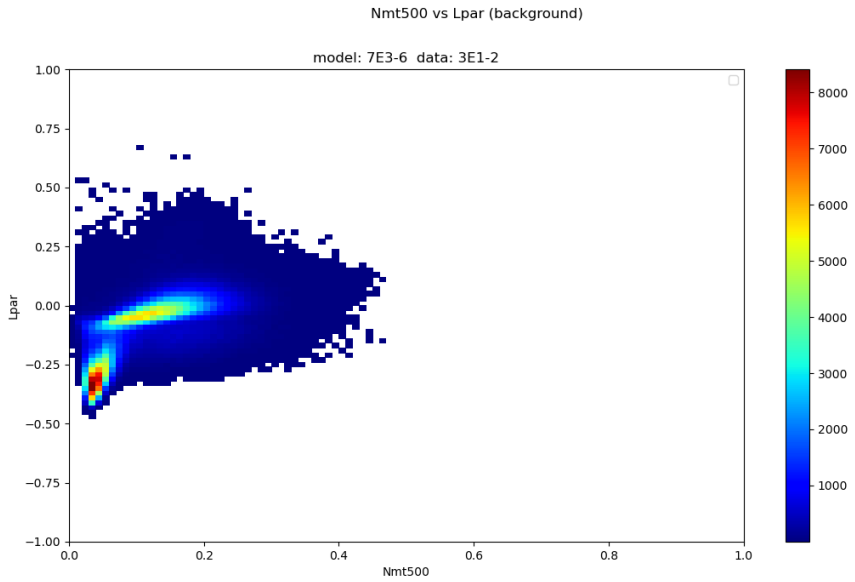
Energy 3~6 keV data is split clearly.

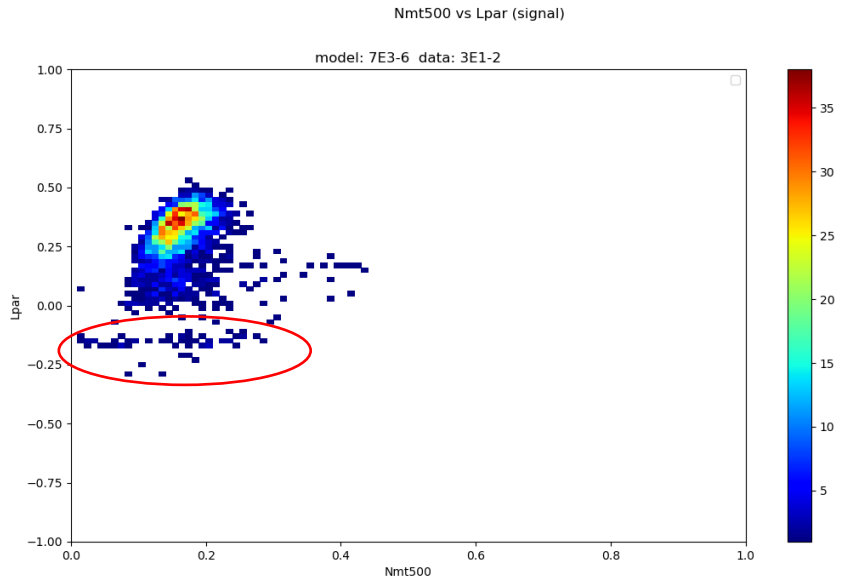




Validation – lpar:nmt500 Plot

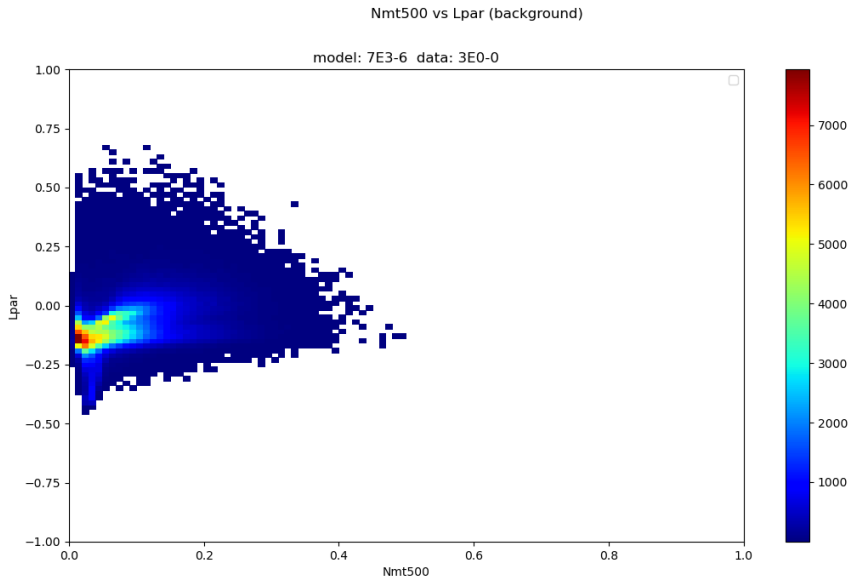
Energy 1~2 keV data is split not clearly,
and low lpar part appears.

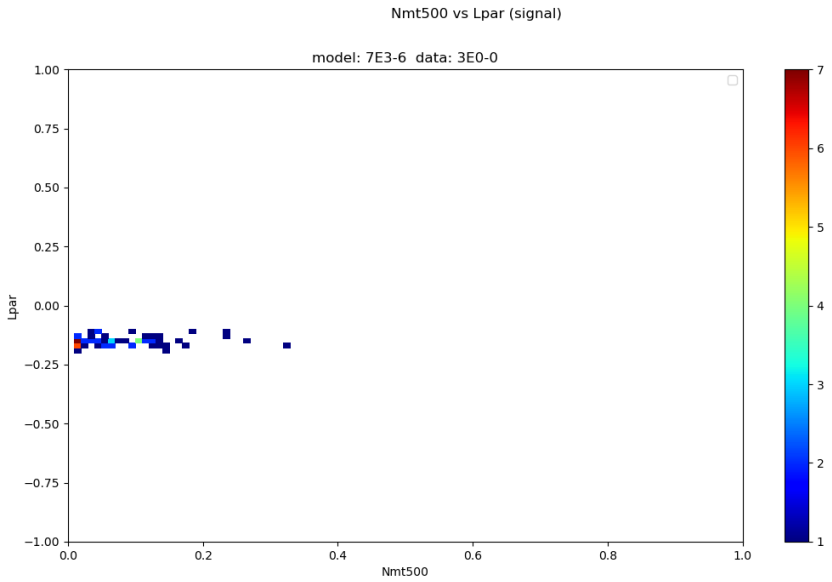




Validation – lpar:nmt500 Plot

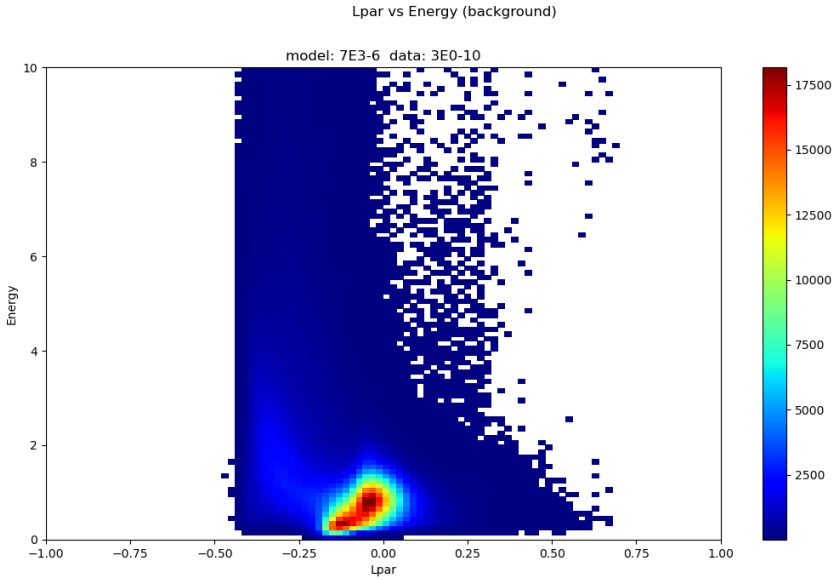
Energy 0~0.5 keV data is not split. But signal data is too rare.

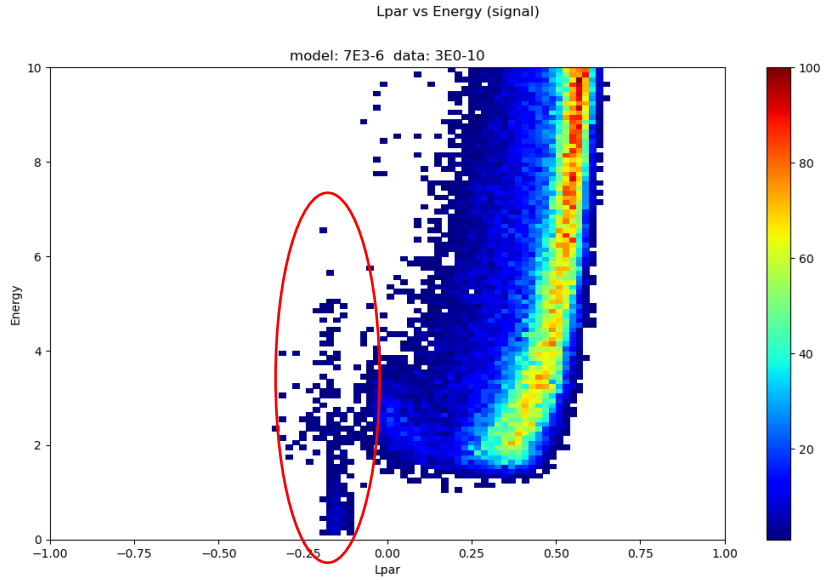




Validation – energy:lpar Plot

Some low lpar samples were classified as signal.
These samples should be studied more.





Validation

- I need some more validation ways.
- Maybe training dataset in lower energy range is useful.