UNIVERSITÀ DEGLI STUDI DI UDINE

Searching for long faint astronomical high energy transients: a data driven approach

Riccardo Crupi, Giuseppe Dilillo, Kester Ward, Elisabetta Bissaldi, Fabrizio Fiore, Andrea Vacchi

October 2024 - Gravi-Gamma-Nu

SPOKE 3

ASTROPHYSICS & COSMOS OBSERVATIONS

ICSC
Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

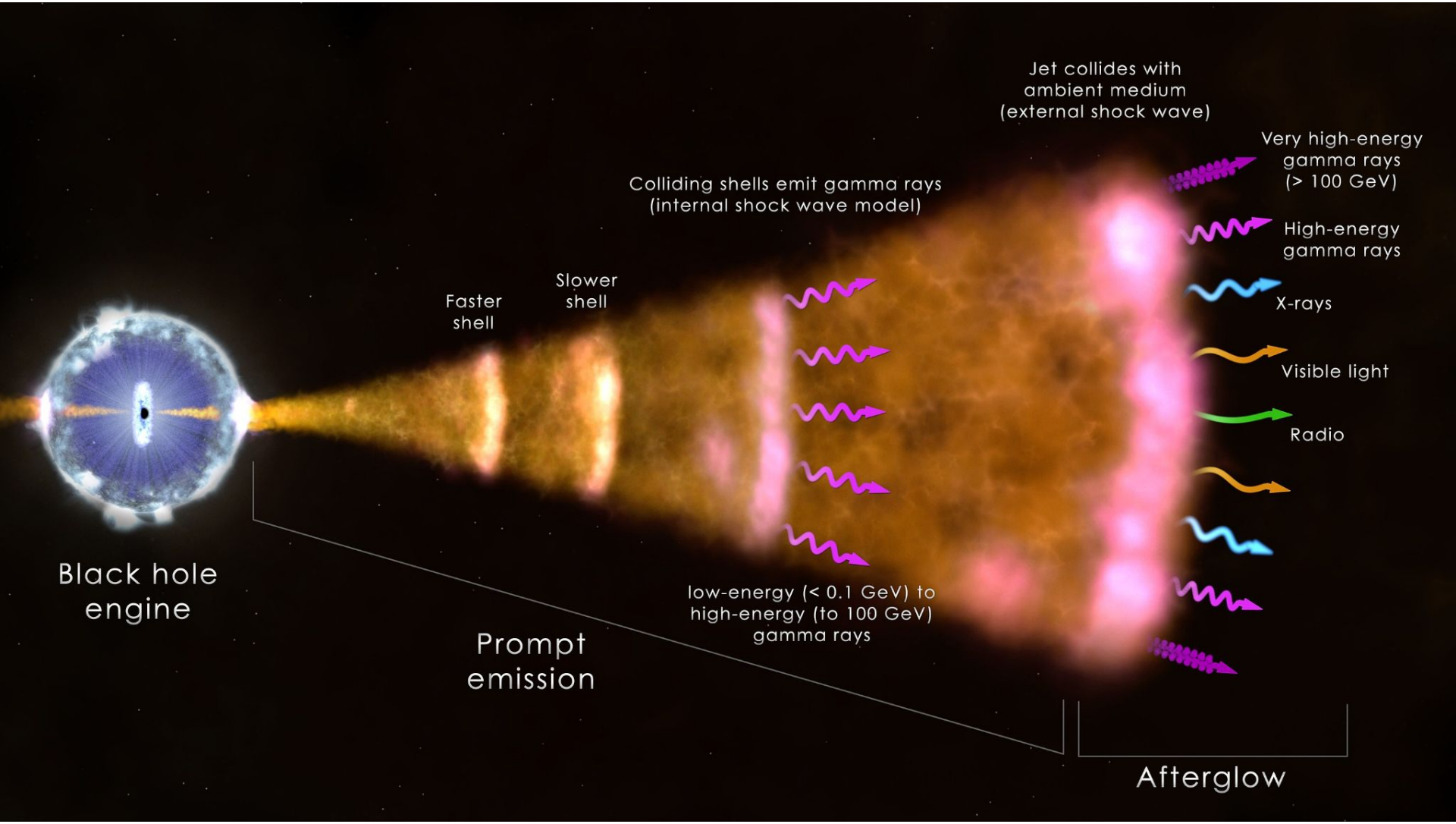
INTESA SANPAOLO

Time series use cases:

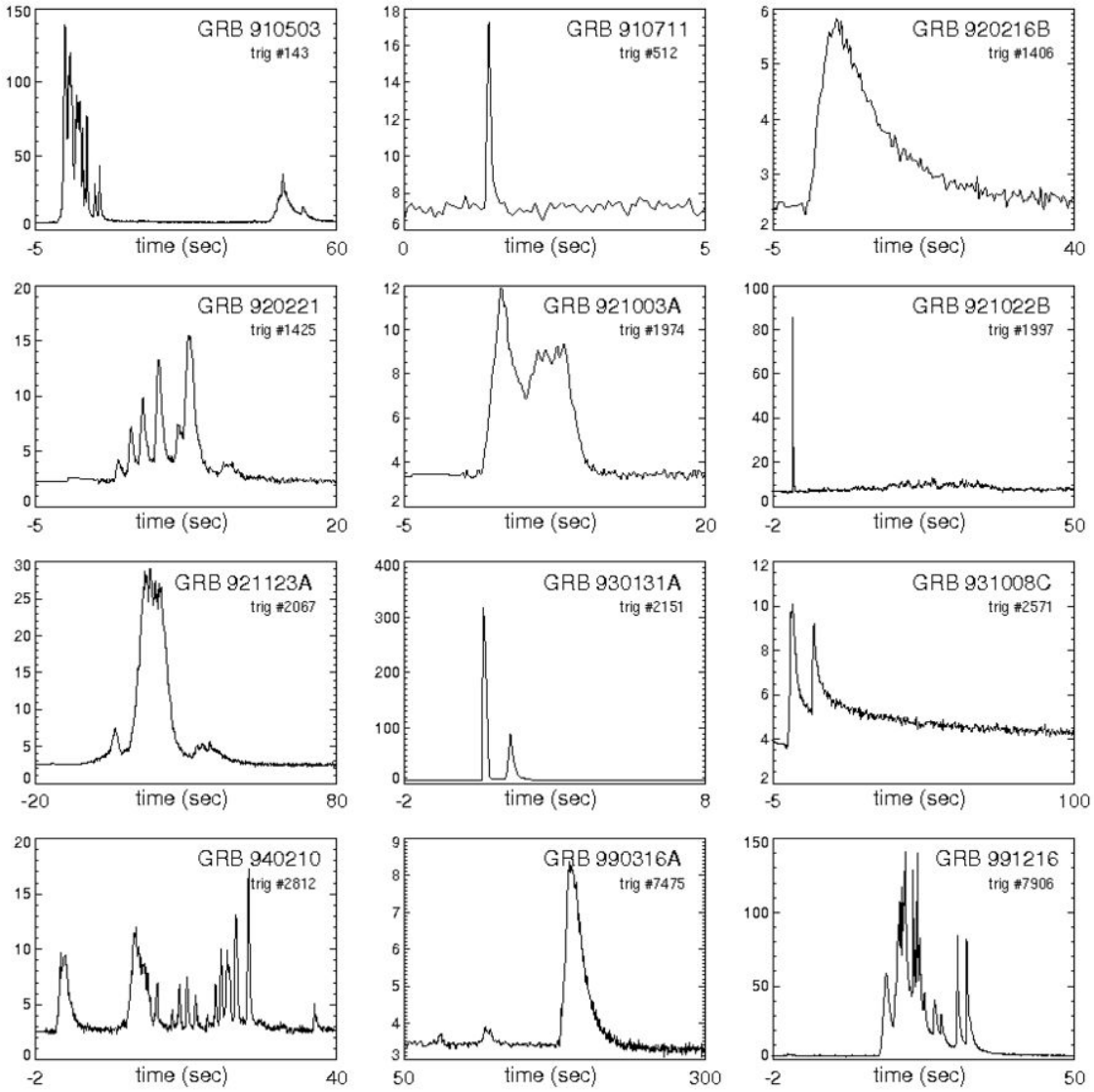
1. Data Quality control on banking accounts
2. Fraud Detection
3. Anomalies in Time Series

Summary

- **Sensitivity demonstration for HERMES**
- New method to estimate the background via a Neural Network
- Use Poisson-FOCuS to trigger events and present some results
- Propose an automatic classification of the events identified



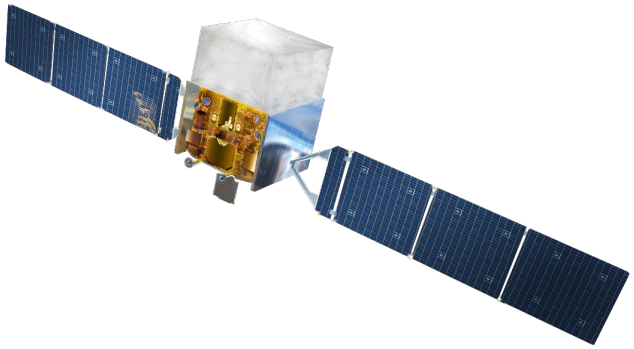
Credit: NASA/Goddard Space Flight Center/ICRAR

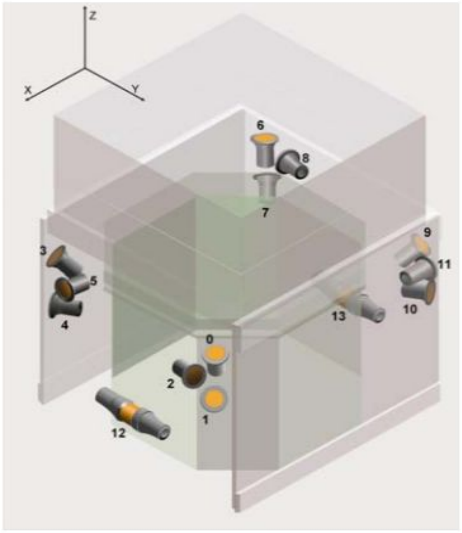


Credit: Daniel Perley http://gammaray.msfc.nasa.gov/batse/grb/catalog/

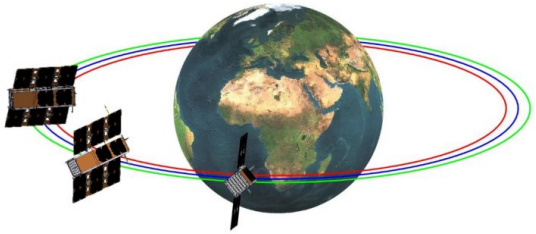
Spacecraft for X/ɣ observations

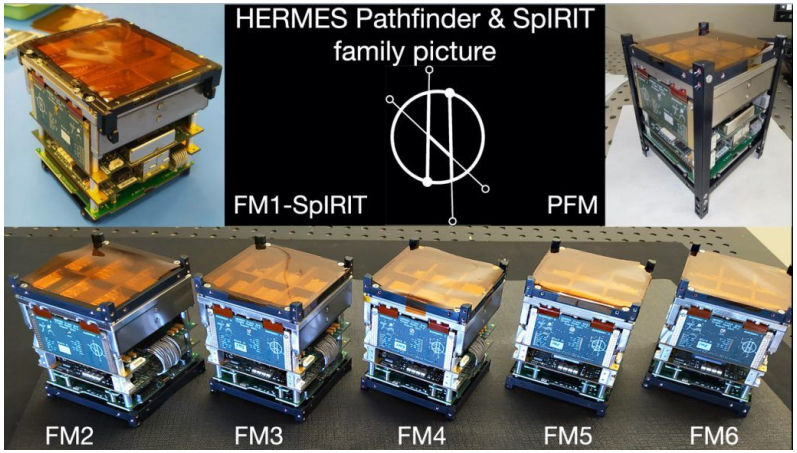
Fermi GBM





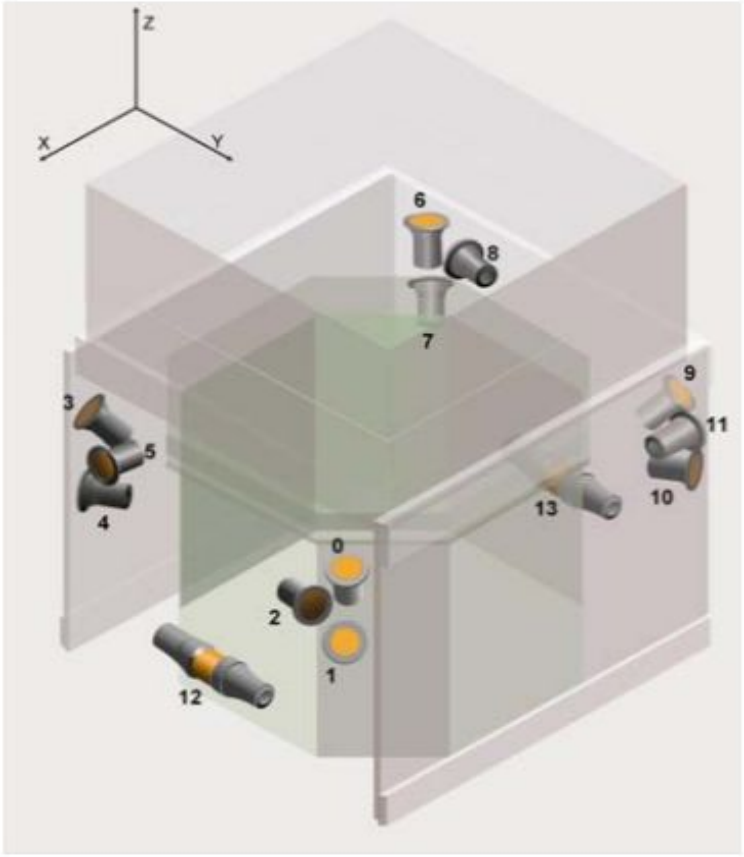
HERMES

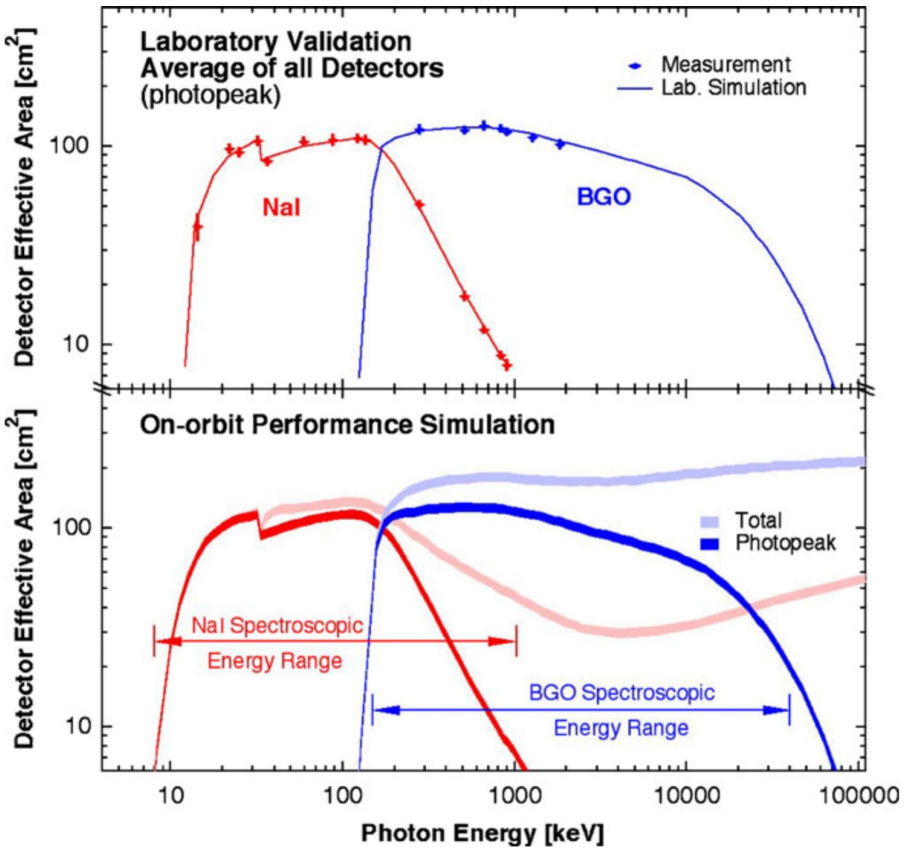




https://www.hermes-sp.eu/

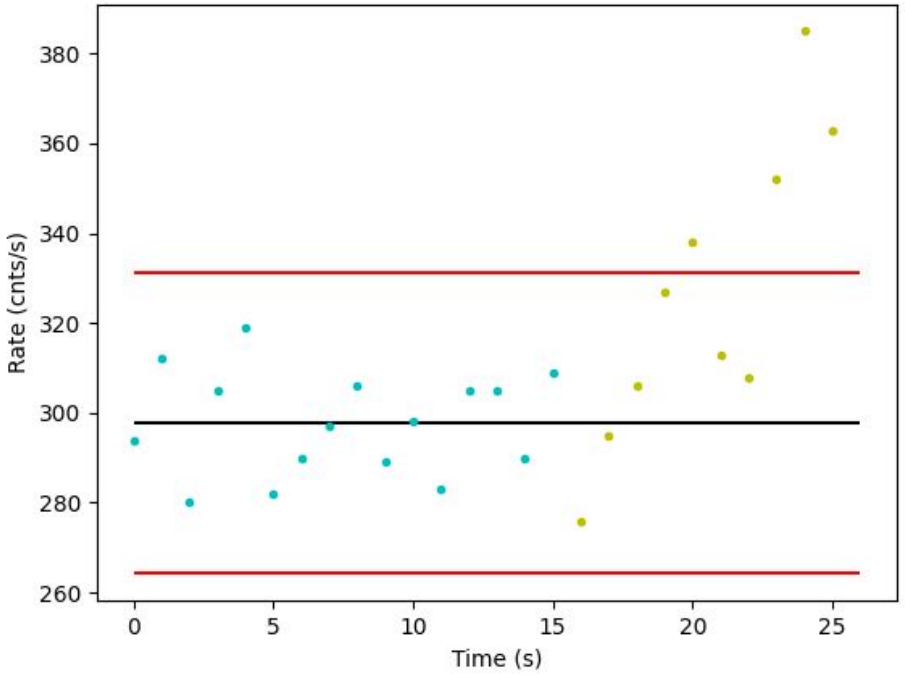
Spacecraft for X/γ observations

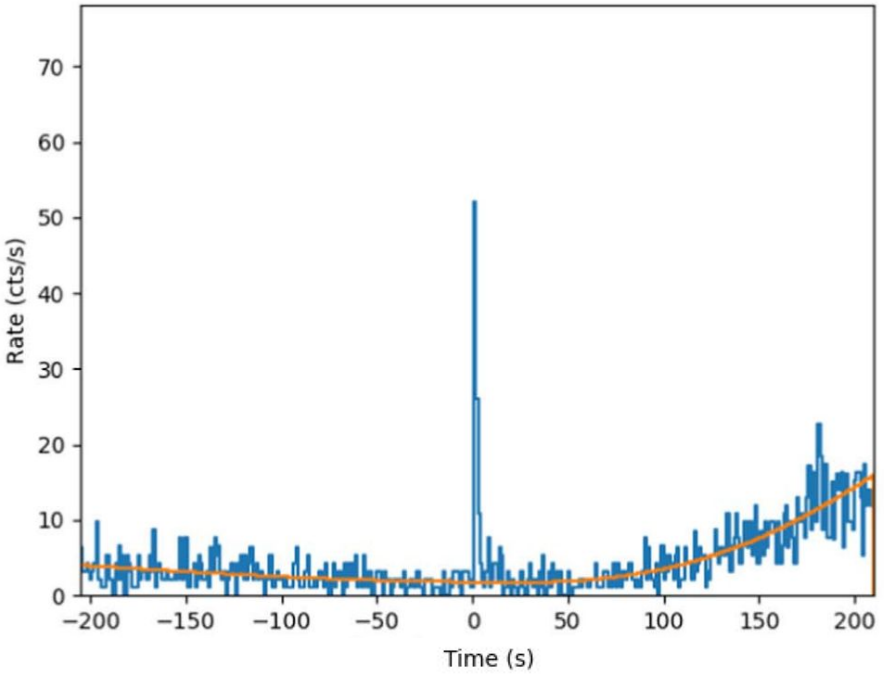




Charles Meegan, et al. “The fermi gamma-ray burst monitor.” The Astrophysical Journal, 702(1):791, 2009.
Elisabetta Bissaldi, et al. “Ground-based calibration and characterization of the fermi gamma-ray burst monitor detectors.” Experimental Astronomy, 24(1-3):47–88, 2009.

Background estimation





Ajello, M., et al. "A decade of gamma-ray bursts observed by Fermi-LAT: the second GRB catalog." The Astrophysical Journal 878.1 (2019): 52.

Background estimation

Background physical model*:

- Detector Response Matrices (DRMs)
- Earth albedo
- South Atlantic Anomaly (SAA)
- point sources (e.g., the Sun)
- extended sources (e.g., cosmic gamma-ray background)
- McIlwain L-parameter (magnetic intensity)
- Other factors…

*Bjorn Biltzinger, Felix Kunzweiler, Jochen Greiner, Kilian Toelge, and J Michael Burgess. A physical background model for the fermi gamma-ray burst monitor. Astronomy & Astrophysics, 640:A8, 2020.

ML Background estimation

Input variables:

- Latitude satellite
- Longitude satellite
- Altitude satellite
- Detectors pointing
- Sun position
- McIlwain L-parameter
- …

ML Background estimation

Input variables:

- Latitude GBM
- Longitude GBM
- Altitude GBM
- Detectors pointing
- Sun position
- McIlwain L-parameter
- …

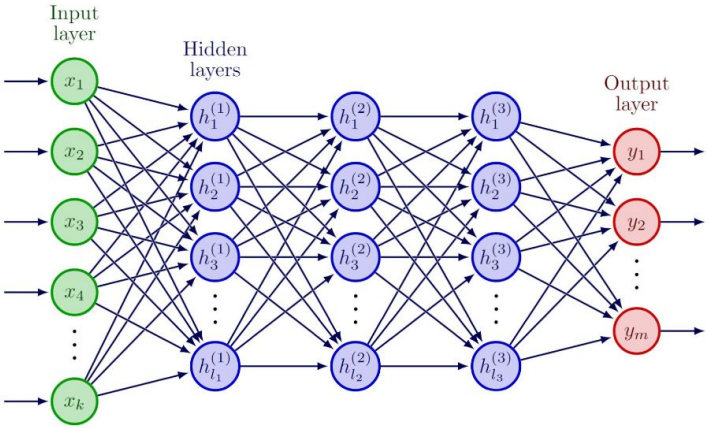
output variables:

- Counts det n0
- Counts det n1
- …
- Counts det n9
- Counts det na
- Counts det nb

ML Background estimation

Input variables:

- Latitude GBM
- Longitude GBM
- Altitude GBM
- Detectors pointing
- Sun position
- McIlwain L-parameter
- …



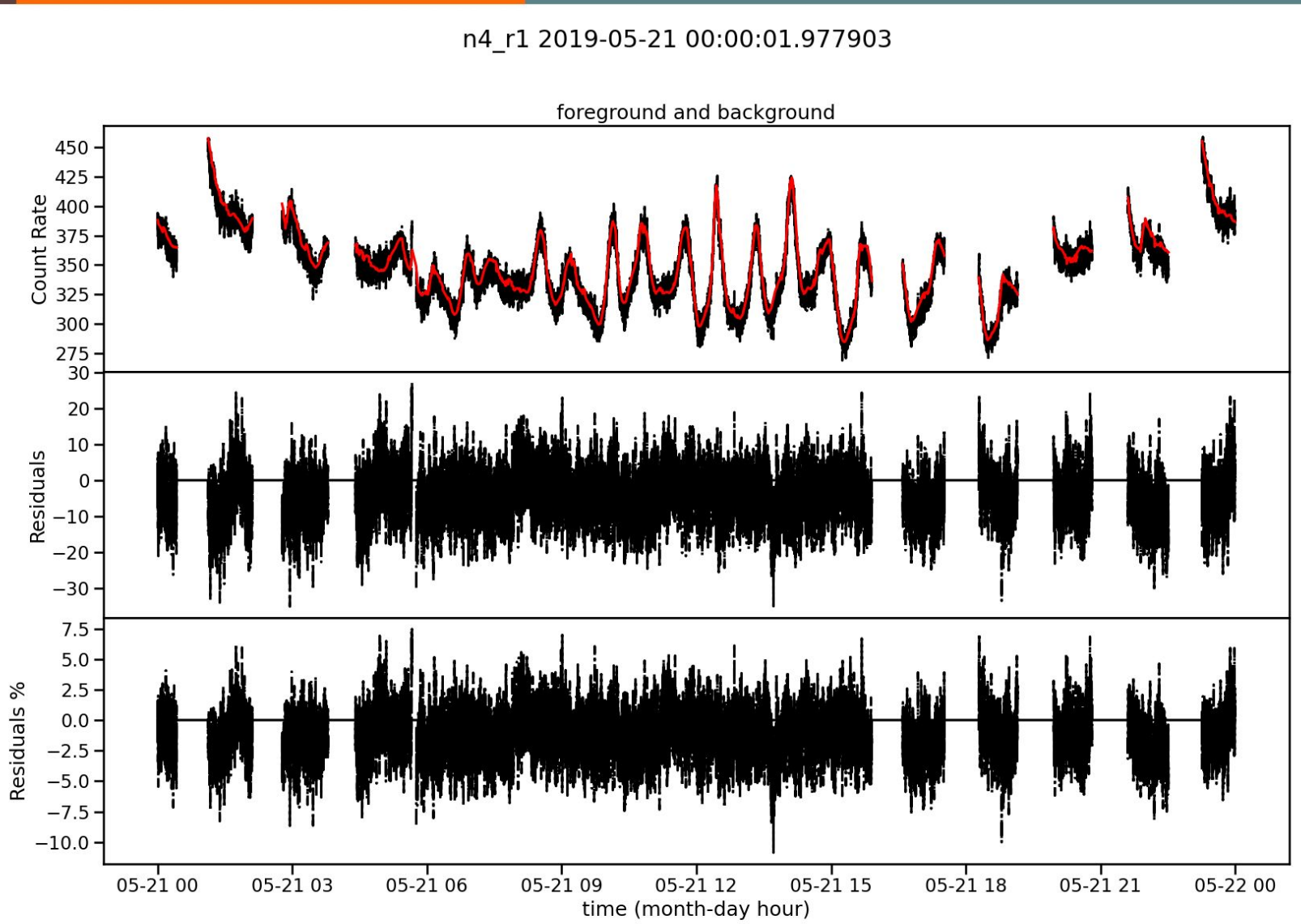
output variables:

- Counts det n0
- Counts det n1
- …
- Counts det n9
- Counts det na
- Counts det nb

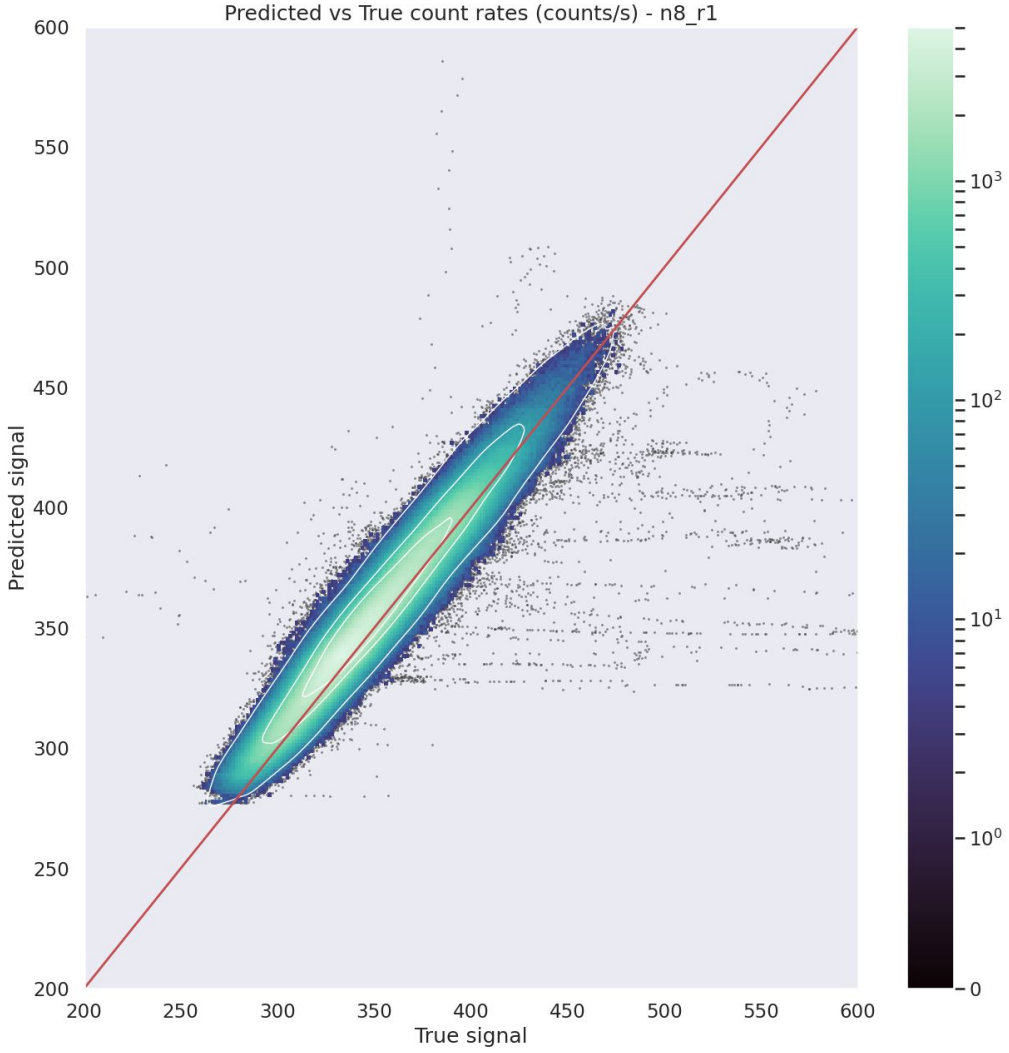
4 seconds bin

Energy range (KeV): [28, 50], [50, 300], [300, 500]

Background estimation



Predicted vs Detected

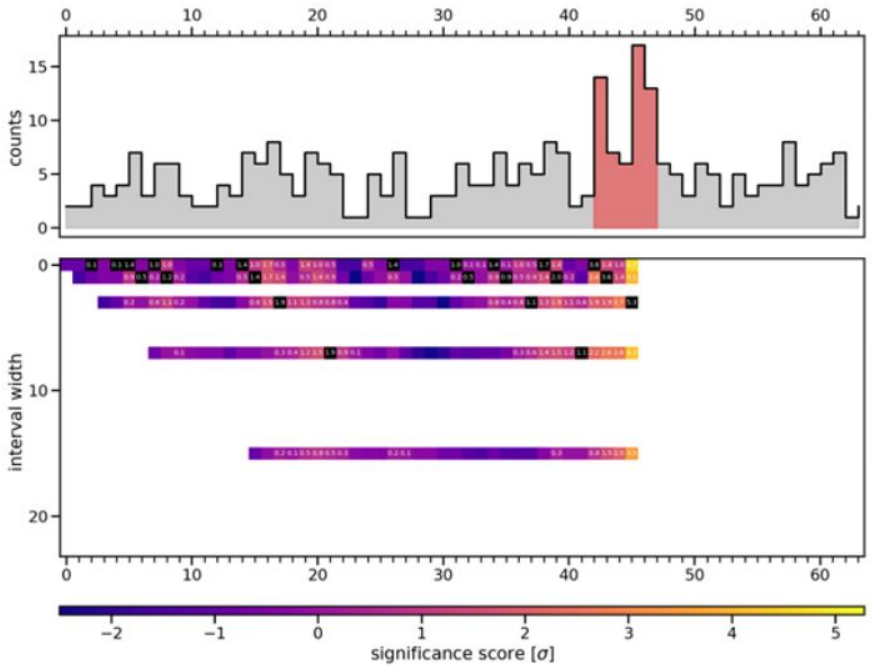


det range	MAE train	MAE test
r0	4.942 ± 0.331	4.953 ± 0.328
r1	6.088 ± 0.167	6.098 ± 0.163
r2	1.790 ± 0.044	1.792 ± 0.045
average	4.273	4.281

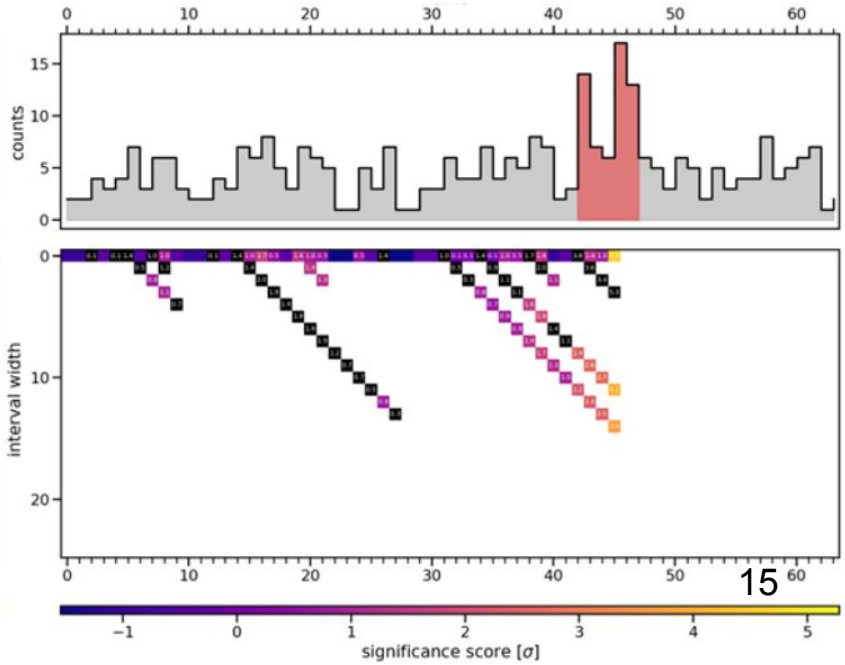
Trigger algorithm: Poisson-FOCuS

- Conventional algorithms checks for counts excess over pre-defined timescales.
- FOCuS tests for count excess over timescales which are optimal and dynamically assessed.
- Faster and more accurate (less false negatives).

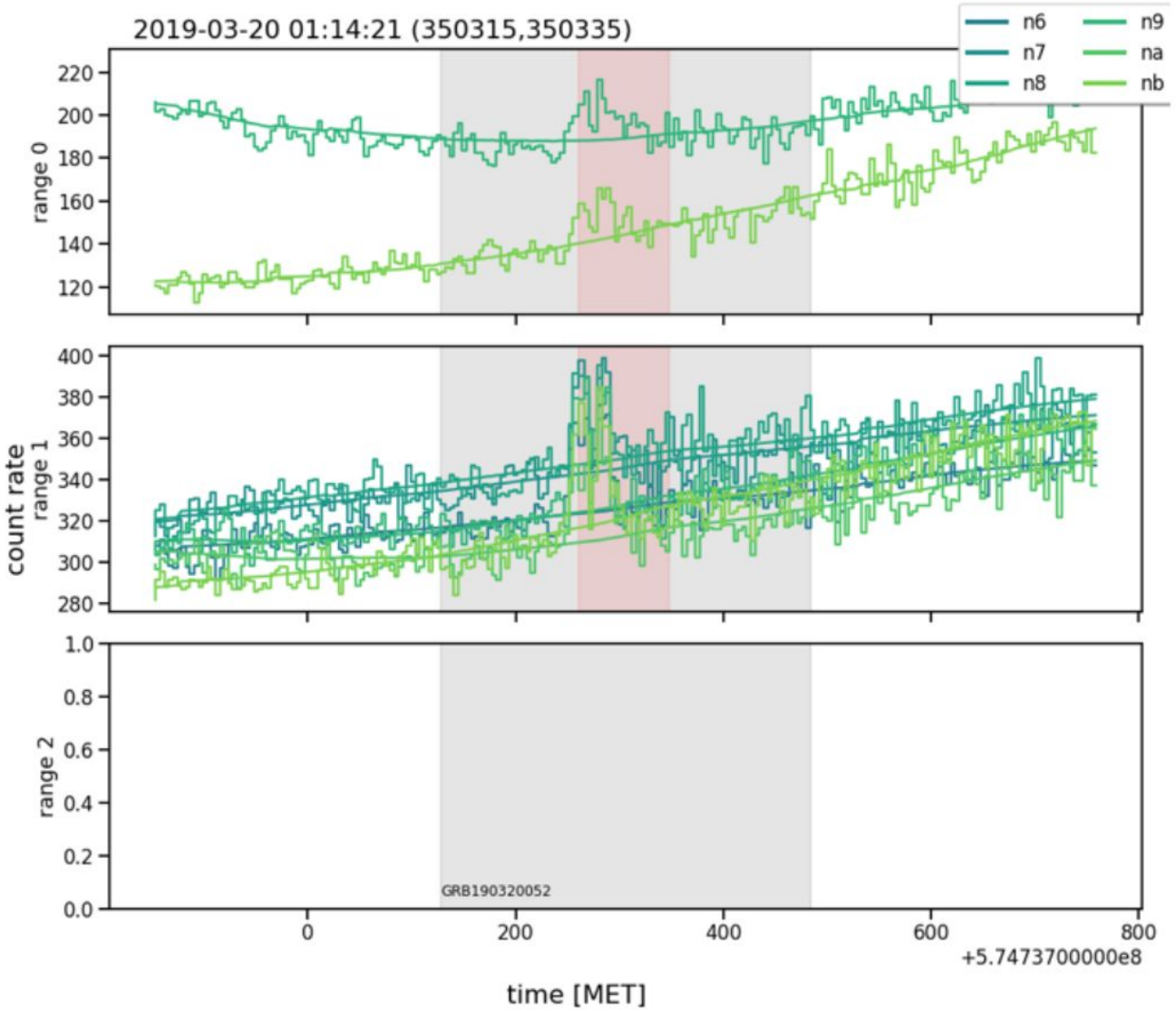
Conventional algorithm trace



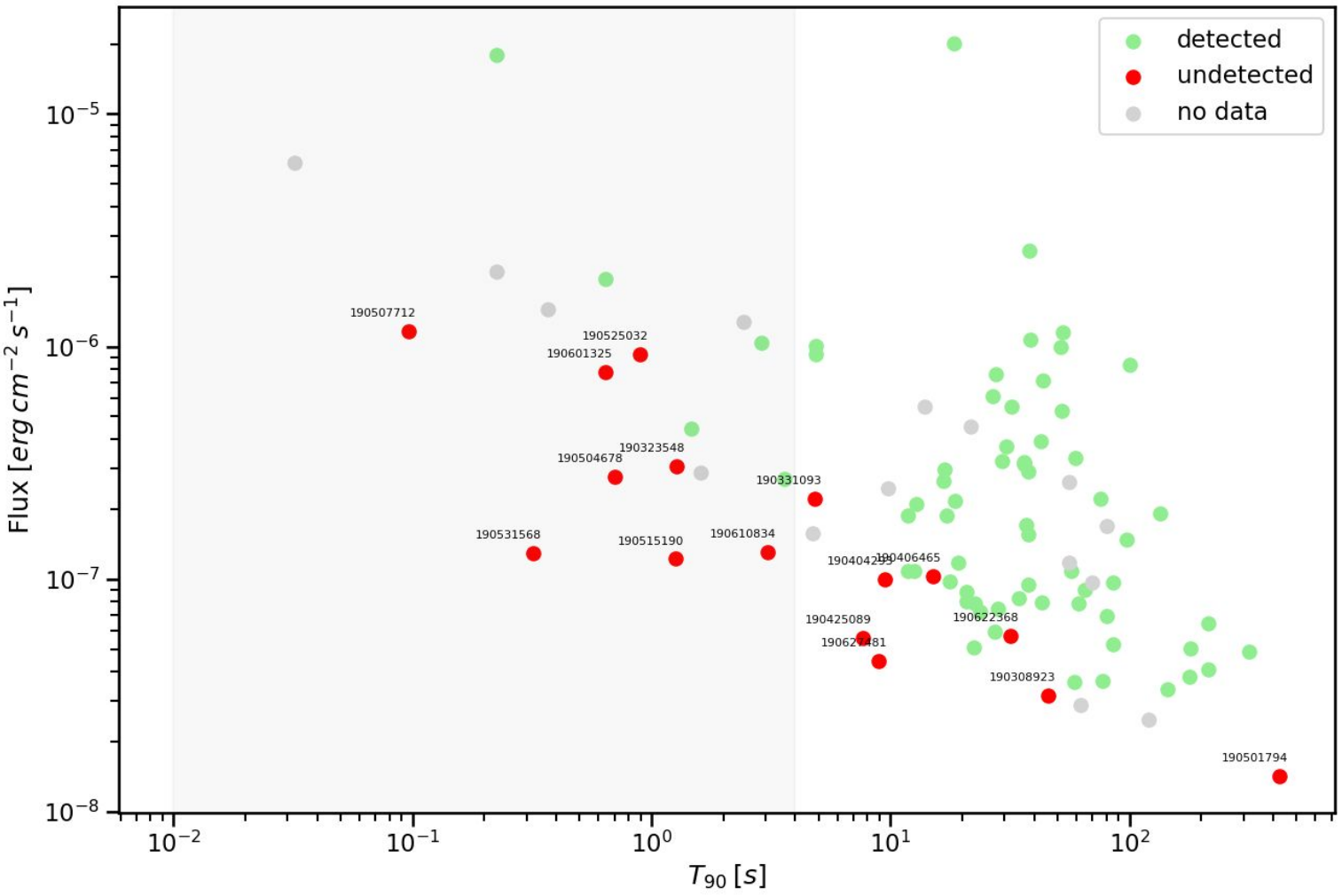
FOCuS algorithm trace



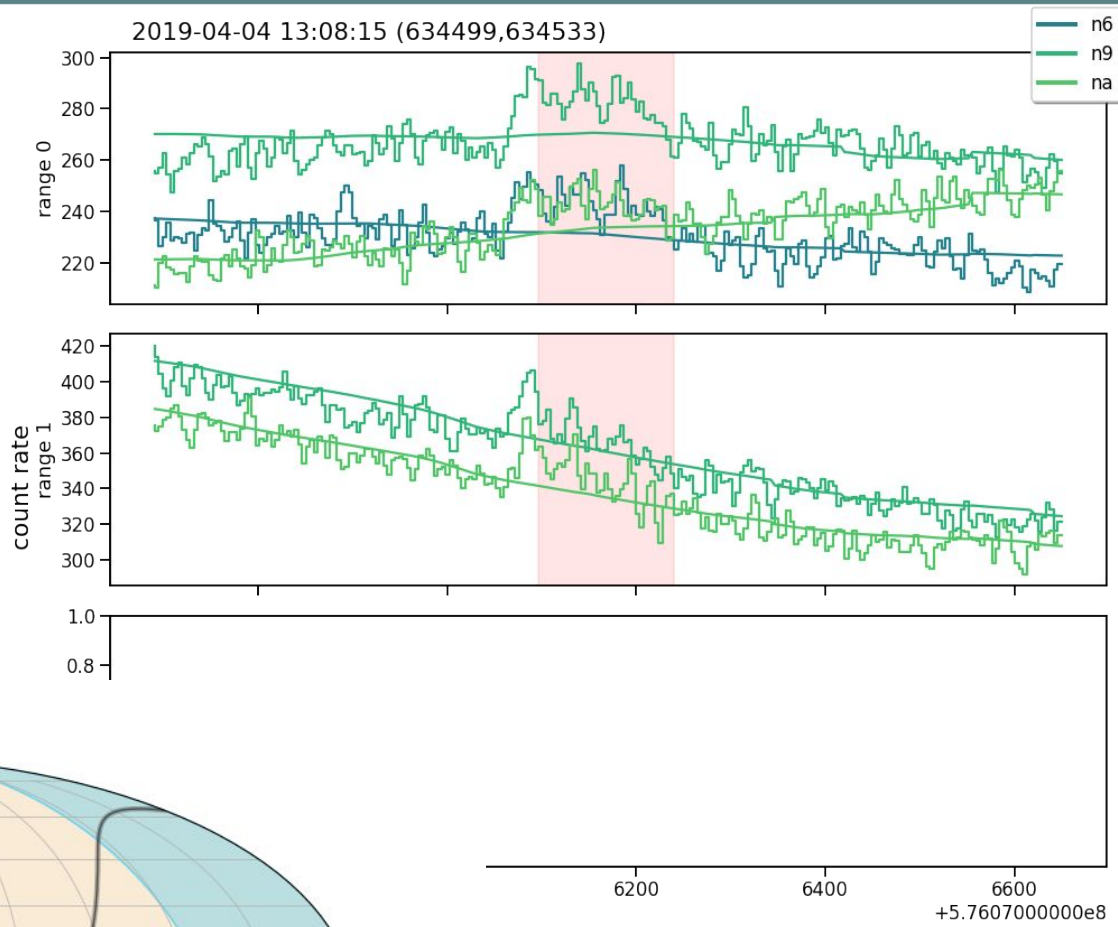
GRB 190320

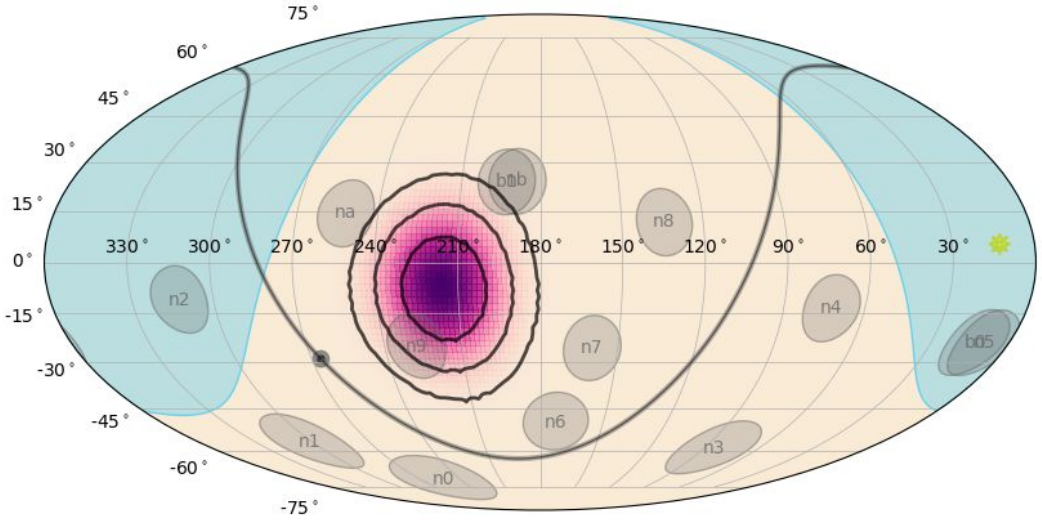

2019-03-20 01:14:21 (350315,350335)
n6
n7
n8
n9
na
nb
range 0
220
200
180
160
140
120
count rate
range 1
400
380
360
340
320
300
280
range 2
1.0
0.8
0.6
0.4
0.2
0.0
GRB190320052
0
200
400
600
800
+5.7473700000e8
time [MET]

Results


detected
undetected
no data
Flux [erg cm^{-2} s^{-1}]
T_{90} [s]
10^{-5}
10^{-6}
10^{-7}
10^{-8}
10^{-2}
10^{-1}
10^{0}
10^{1}
10^{2}
190507712
190525032
190601325
190323548
190504678
190331093
190531568
190515190
190610834
190404295
190406465
190425089
190622368
190621481
190308923
190501794

Event 190404





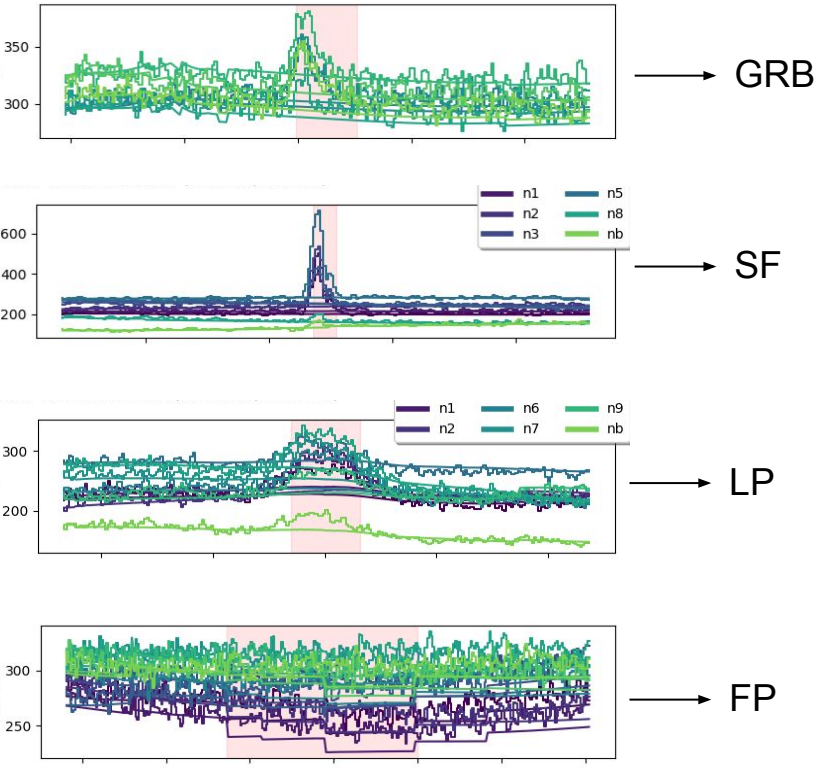
Attempting an automatic classification

type of transient	num. known	num. unidentified	total
GRB	127	30	157
SF	78	24	102
UNC(LP)	3	57	60
GF	0	10	10
TGF	1	8	9
UNC	1	3	4
FP	0	52	52

Data periods:

- November 2010 to 19 February 2011
- January 2014 to 28 February 2014
- March 2019 to 9 July 2019

Attempting an automatic classification



Class	Metric	DT	RF	FuleFit	Manual rule
GRB	Precision	81%	**93%**	87%	87%
	Recall	84%	**90%**	87%	84%
	F1-score	83%	**92%**	87%	85%
SF	Precision	80%	80%	**87%**	83%
	Recall	**95%**	**95%**	**95%**	**95%**
	F1-score	87%	87%	**91%**	89%
LP	Precision	43%	73%	**100%**	73%
	Recall	83%	67%	58%	**92%**
	F1-score	57%	70%	74%	**81%**
FP	Precision	50%	82%	**89%**	71%
	Recall	**100%**	90%	80%	**100%**
	F1-score	67%	**86%**	84%	83%

Feature extracted ~ 200: statistical, wavelet transform, catalog info.

Thank you,
I hope you have been triggered

crupi.riccardo@spes.uniud.it
riccardo.crupi@intesasanpaolo.com

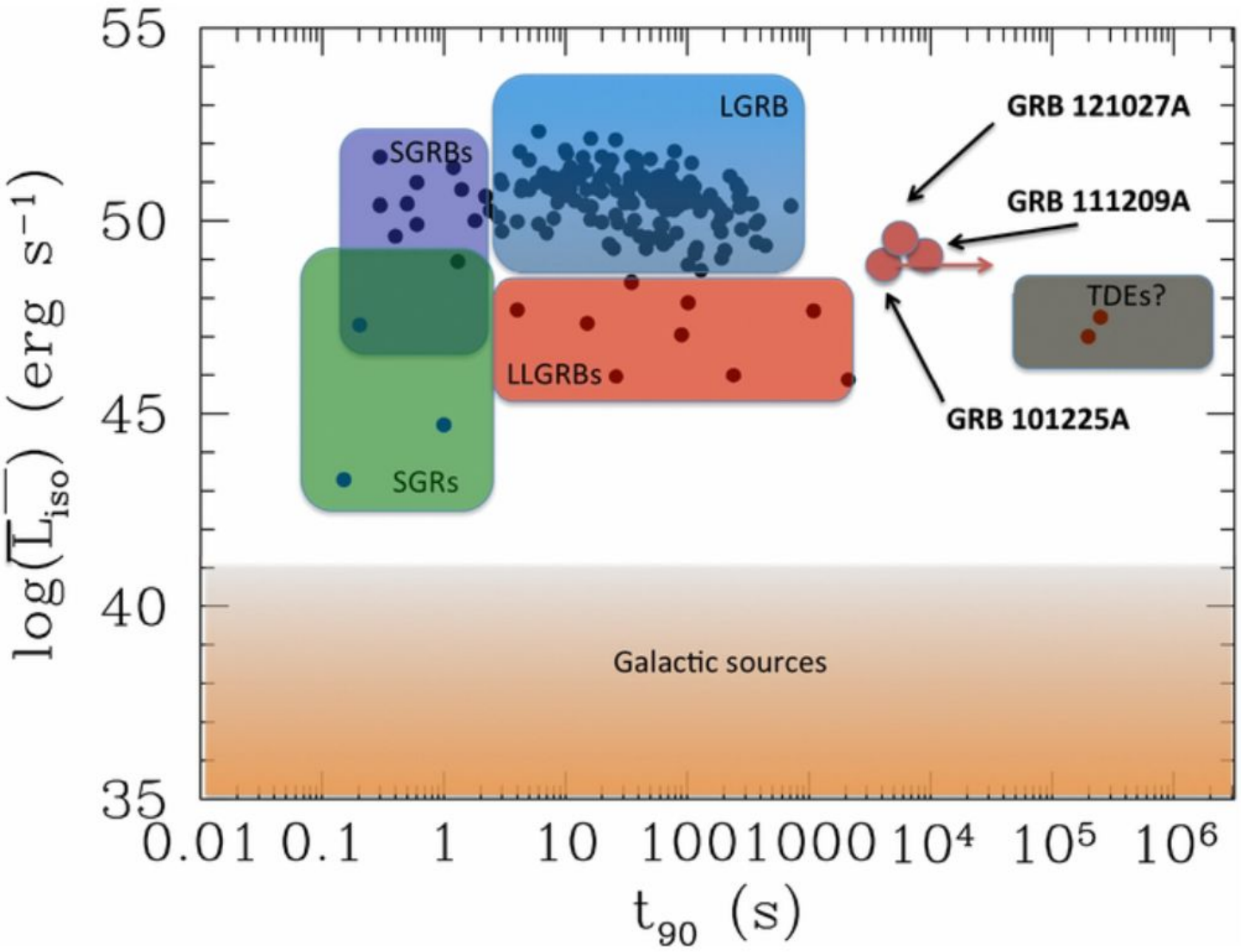
GitHub repository
https://github.com/rcrupi/DeepGRB

References

- Riccardo Crupi, Giuseppe Dilillo, Kester Ward, Elisabetta Bissaldi, Fabrizio Fiore, Andrea Vacchi: **Searching for long faint astronomical high energy transients: a data driven approach**. Experimental Astronomy 56.2 (2023): 421-476.

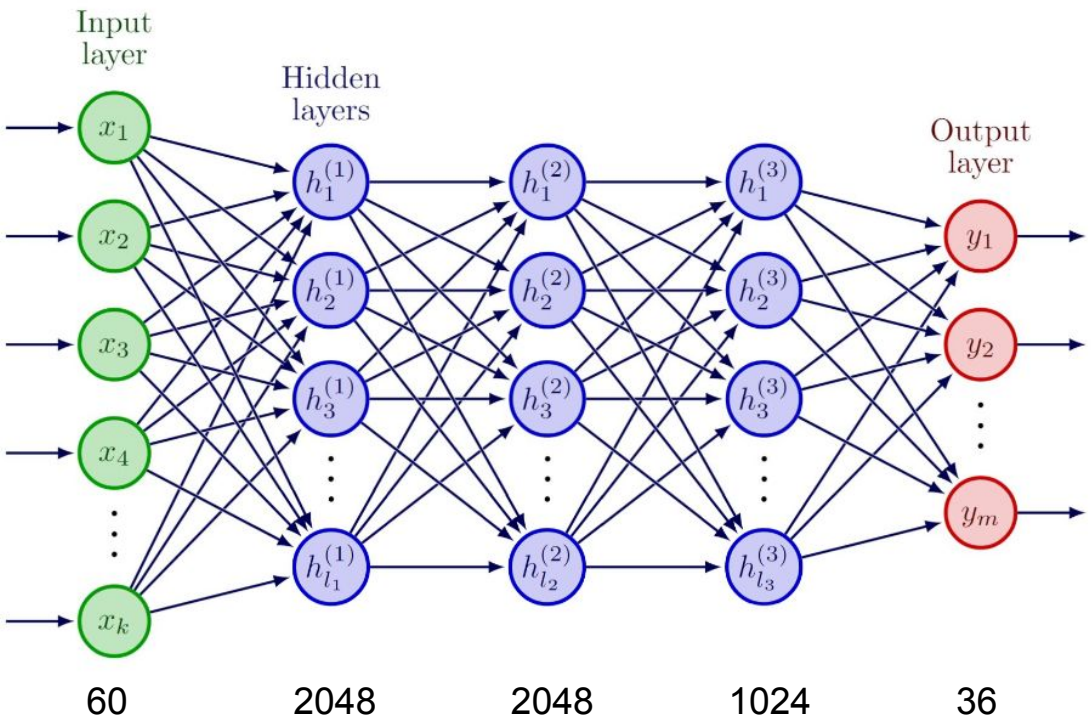
- Kester Ward, Giuseppe Dilillo, Idris Eckley, Paul Fearnhead. Poisson-FOCuS: An efficient online method for detecting count bursts with application to gamma ray burst detection. Journal of the American Statistical Association (2023): 1-13.

Backup slides



Levan, Andrew J., et al. "A new population of ultra-long duration gamma-ray bursts." The Astrophysical Journal 781.1 (2013): 13.

NN hypermarameters



- Optimizer: Nadam
- Batch normalization and dropout 0.02 each hidden layer
- Decay learning rate 0.01 first 4 epochs, 0.0016 up to 12 epochs and then 0.0004
- Batch size: 2000
- Max epochs: 64
- L1 loss (Mean Absolute Error)

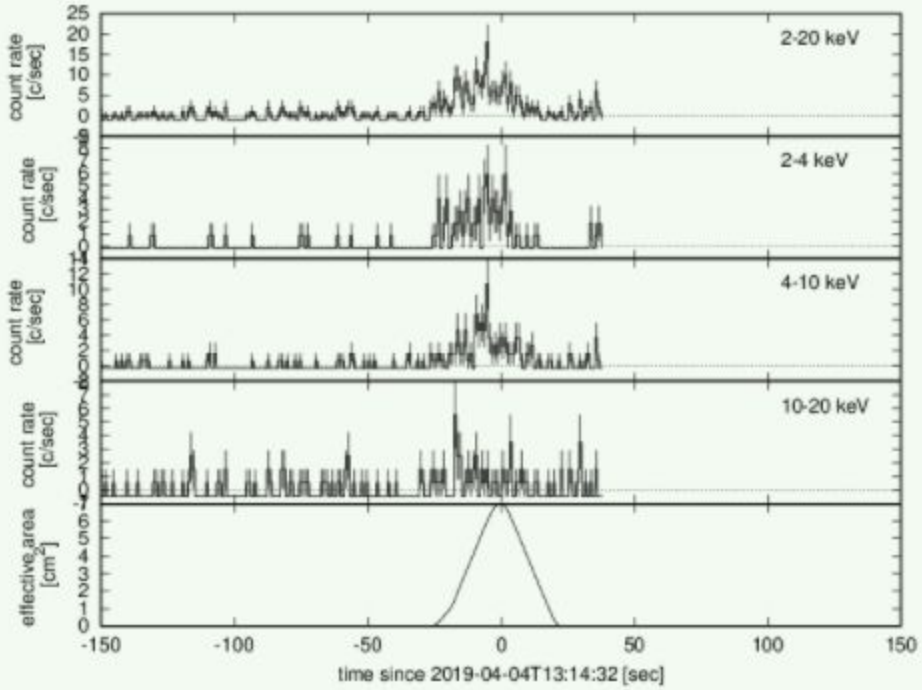
GRB 190404B

Data	med
SN	32.9
Flux ($\times 10^{-8}$ erg cm^{-2} s^{-1})	1.09 ± 0.08
Hardness	0.27 ± 0.06
Time bin	1 sec

http://maxi.riken.jp/grbs/190404b/

light curves

no effective area correction

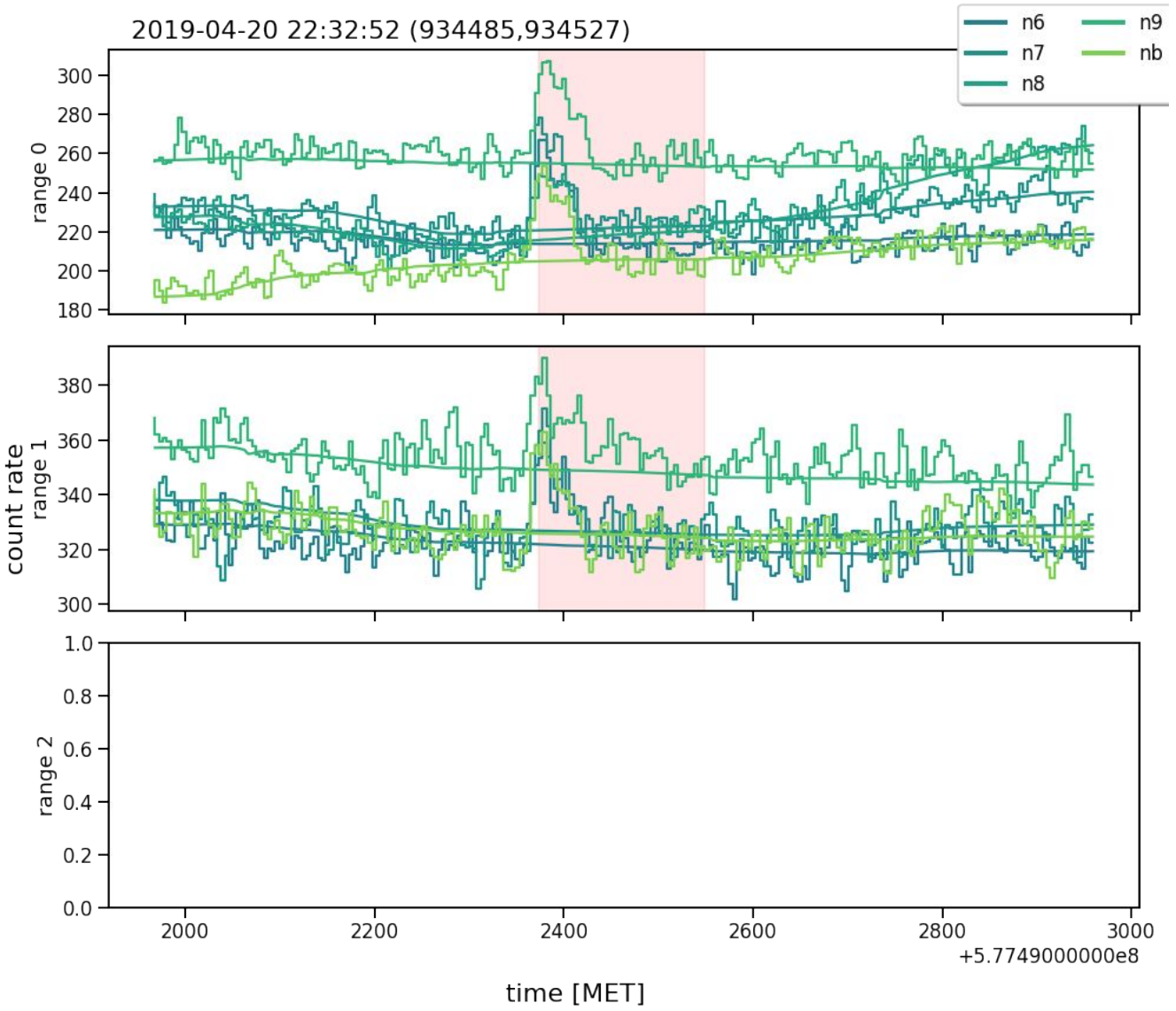


The MAXI/GSC nova alert system triggered on a bright uncatalogued X-ray transient source at 13:14:08 UT on 2019 April 04.
Assuming that the source flux was constant over the transit,
we obtain the source position at
(R.A., Dec) = (220.987 deg, -22.632 deg) =
(14 43 56, -22 37 55) (J2000)
with a statistical 90% C.L. elliptical error region
with long and short radii of 0.13 deg and 0.11 deg, respectively.

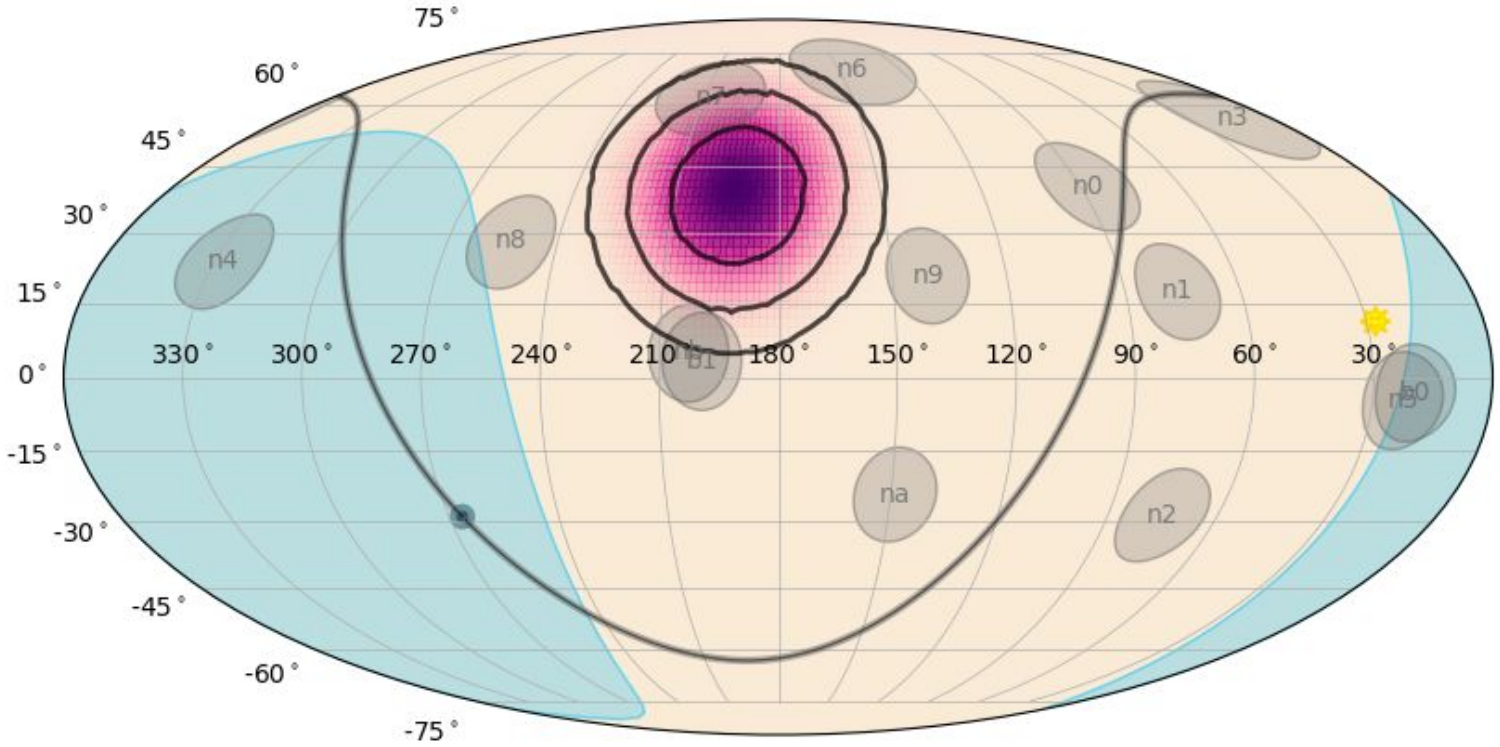
GCN:
https://gcn.gsfc.nasa.gov/gcn3/24049.gcn3
https://gcn.gsfc.nasa.gov/other/190404.gcn3

Swift afterglow:
https://www.swift.ac.uk/xrt_live_cat/20887

Event 190420



Event 190420

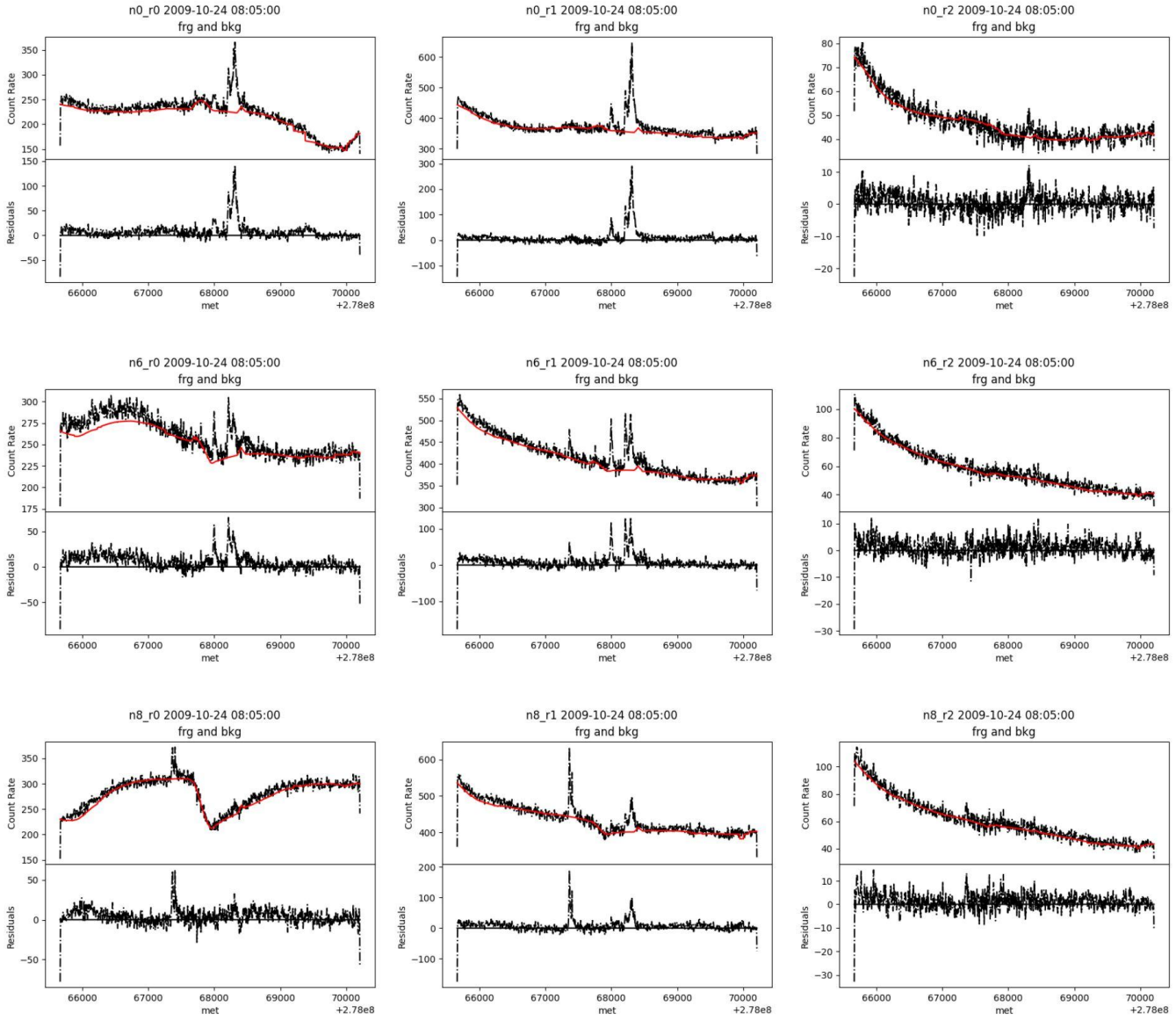


det range	MAE train	MAE test
r0	4.942 ± 0.331	4.953 ± 0.328
r1	6.088 ± 0.167	6.098 ± 0.163
r2	1.790 ± 0.044	1.792 ± 0.045
average	4.273	4.281

Table 1: The NN MAE loss function ($\pm$ standard deviation) per energy range, over the training and the testing datasets, averaged on detectors. The complete table can be found in table 7).

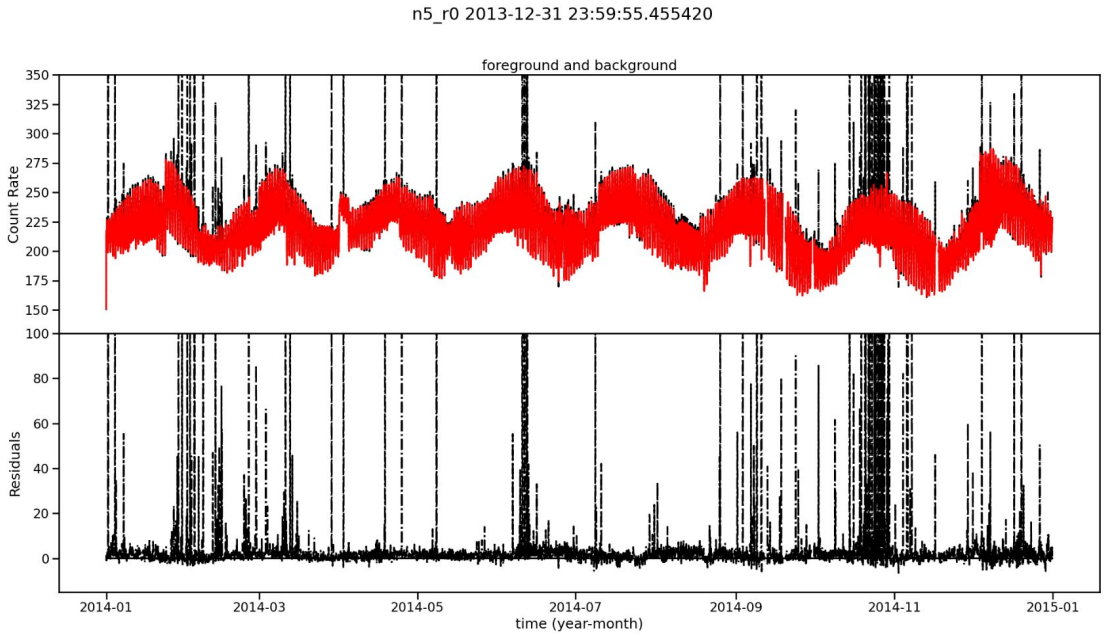
Background estimation

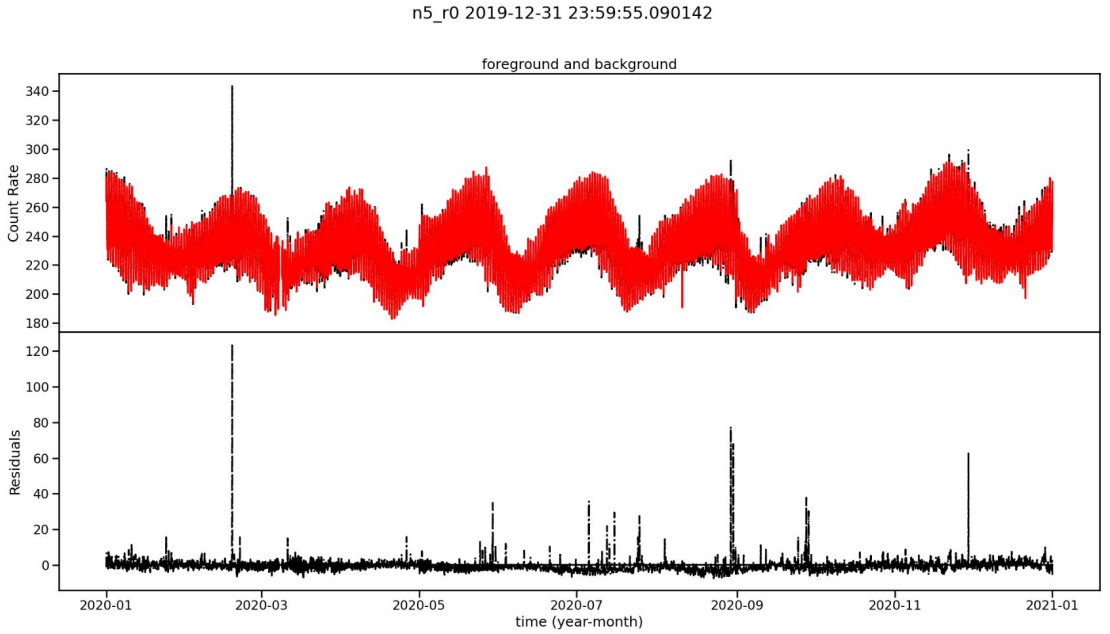
GRB 091024



Background estimation

Solar
Maxima
Minima



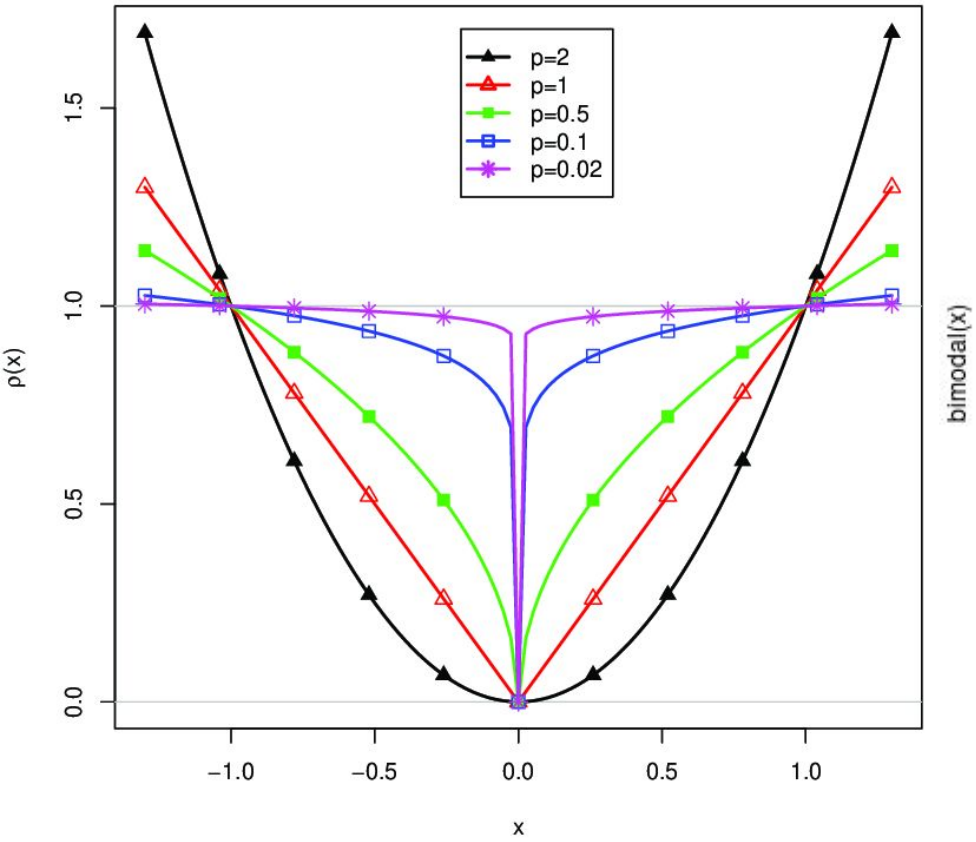


Partial input - output

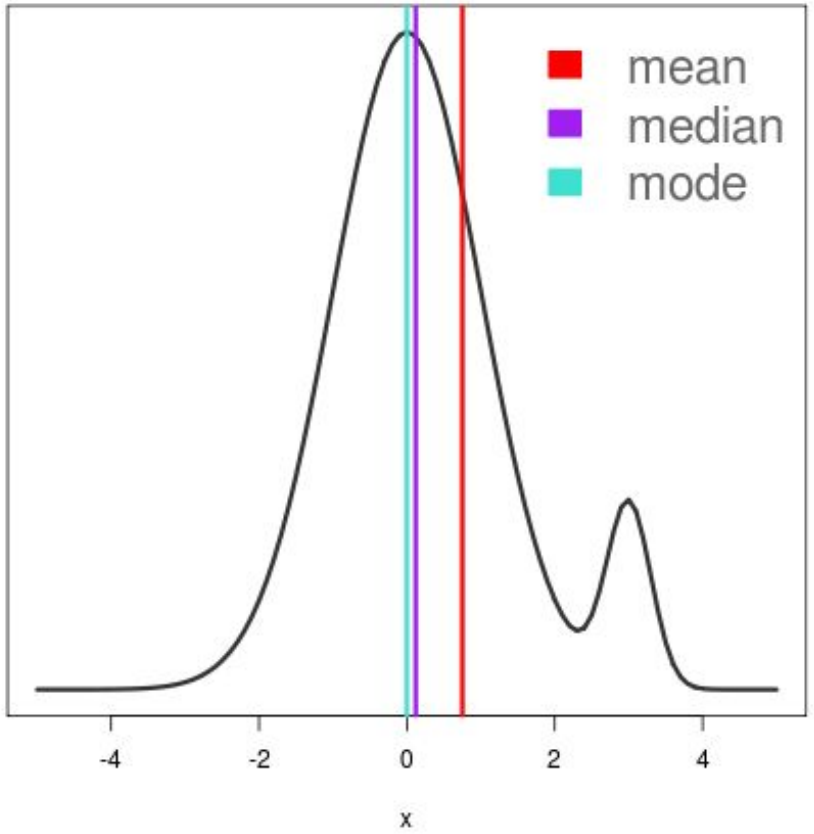
lat	lon	alt	n0_dec	n0_ra
-21.74720	16.13370	536687.34686	-60.61029	193.13023
-21.68540	16.38606	536703.44551	-60.45834	193.21557
-21.62316	16.63820	536719.48134	-60.30667	193.30146
-21.56049	16.89012	536734.64282	-60.15521	193.38776
-21.49737	17.14180	536750.18760	-60.00350	193.47457
-21.43384	17.39325	536765.63889	-59.85207	193.56186
-21.36987	17.64446	536780.25793	-59.70084	193.64969
-21.30547	17.89543	536795.28120	-59.54942	193.73788
-21.24065	18.14616	536809.89110	-59.39822	193.82666
-21.17541	18.39666	536825.25390	-59.24726	193.91585
-21.10975	18.64692	536840.85849	-59.09606	194.00553
-21.04367	18.89693	536855.23207	-58.94514	194.09561
-20.97716	19.14671	536869.53731	-58.79449	194.18599
-20.91025	19.39625	536884.27429	-58.64356	194.27703
-20.84292	19.64555	536898.11610	-58.49301	194.36832
-20.77518	19.89461	536911.85559	-58.34254	194.46017
-20.70703	20.14341	536925.07311	-58.19195	194.55239
-20.63847	20.39198	536939.09638	-58.04163	194.64502
-20.56950	20.64031	536952.73061	-57.89151	194.73807
-20.50013	20.88840	536965.88525	-57.74122	194.83153

n0_r0	n0_r1	n0_r2
131.84344	275.56137	46.64100
274.10376	548.45300	91.94076
280.94550	541.39386	92.78934
280.45737	535.99884	95.24504
280.56808	537.93964	94.87277
266.68027	535.44684	97.44323
264.70062	534.06450	100.62796
265.55270	535.52313	97.31207
262.96994	521.39954	92.15605
266.53240	521.16160	93.38451
265.80035	529.88270	95.71757
263.83970	521.91174	91.17808
264.19520	509.61572	90.19204
267.47410	504.76535	93.12515
270.79575	513.12540	91.04234
257.90674	513.11414	89.93607
249.80258	503.28592	89.44306
257.17477	505.39258	91.40993
253.60439	498.98840	87.97018
252.25041	490.51416	83.91988

Loss function



Some distribution



https://www.r-bloggers.com/2012/03/how-not-to-draw-a-probability-distribution/

Robitzsch, Alexander. "Lp loss functions in invariance alignment and Haberman linking with few or many groups." Stats 3.3 (2020): 246-283.

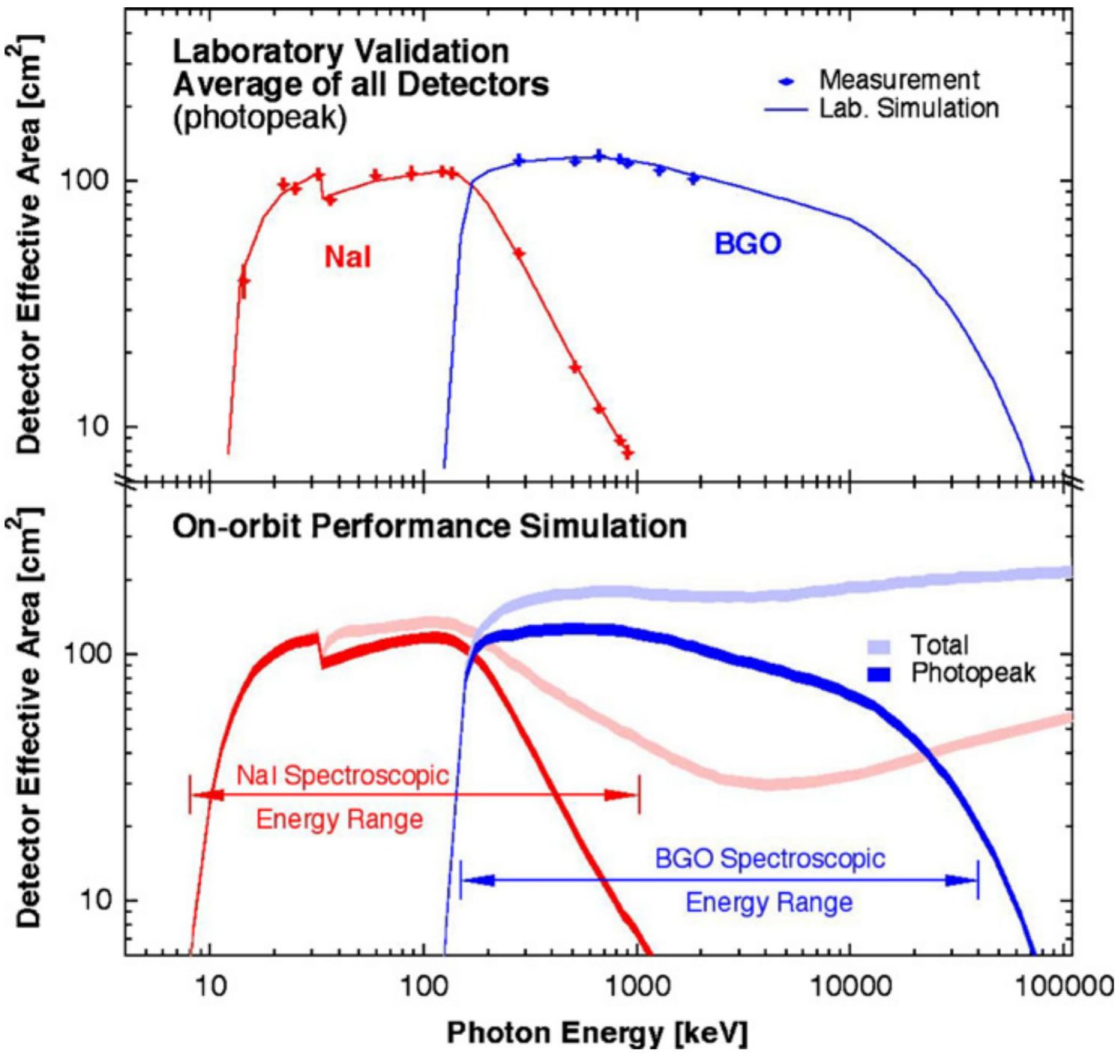
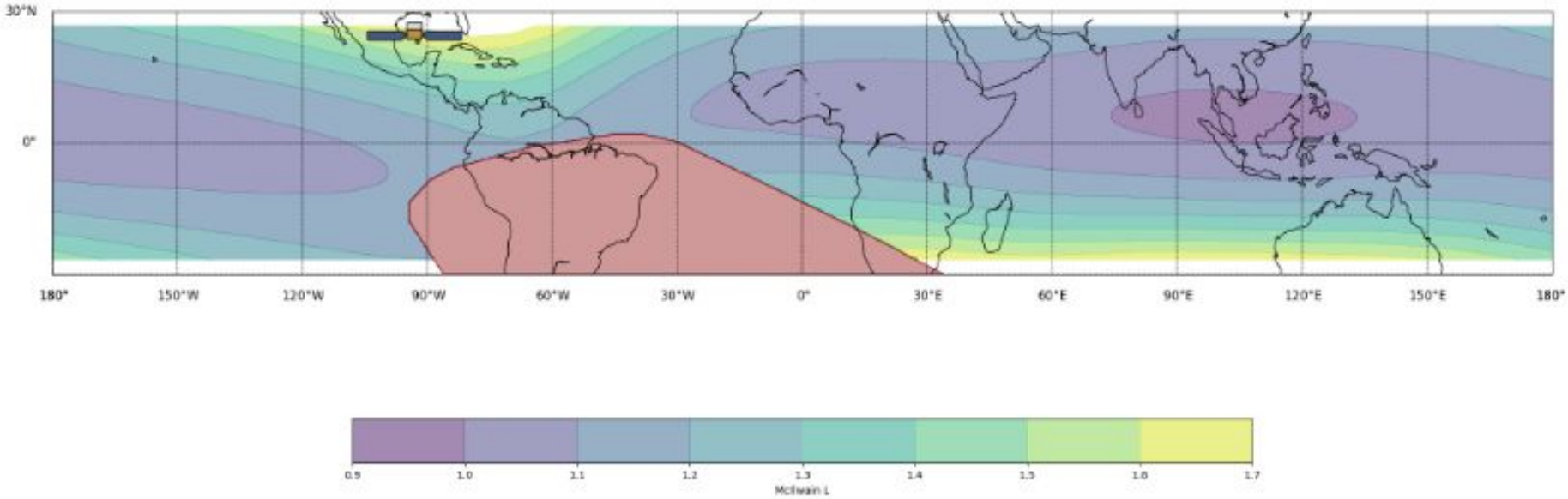
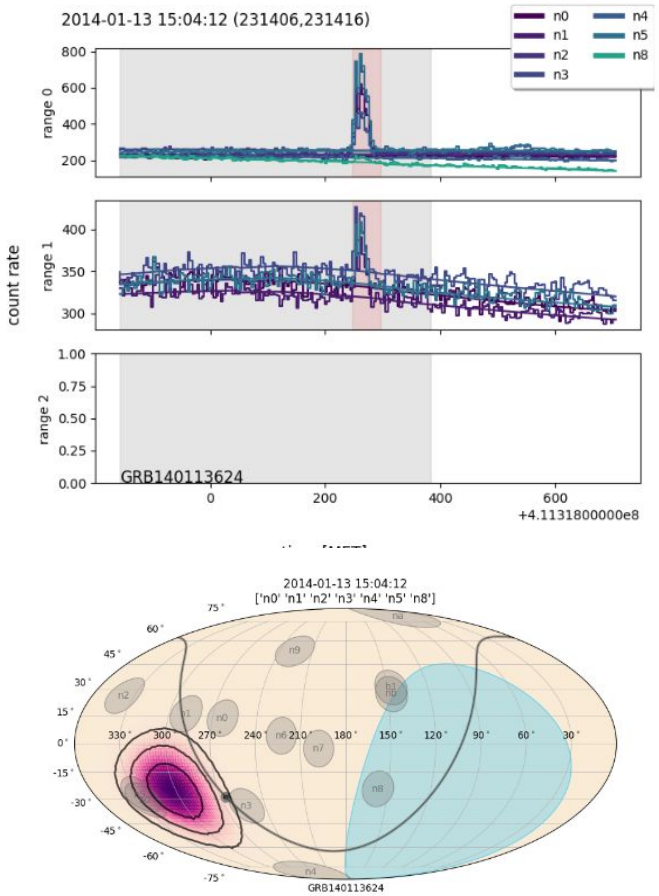


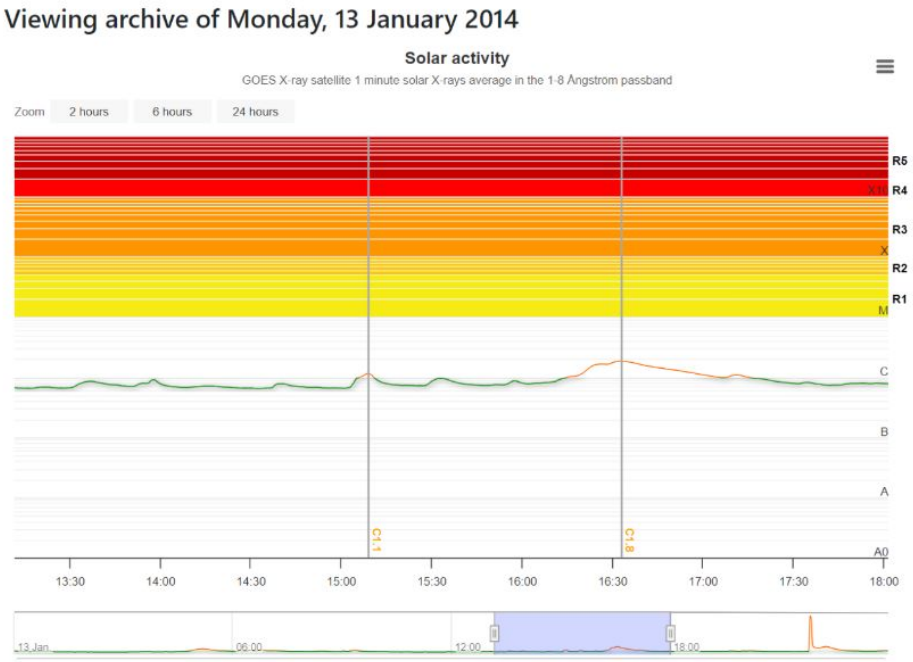
Figure 11. Energy dependence of the effective area at normal incidence, for both detector types. The lower panel includes the simulated effects of the spacecraft for a representative detector.



https://fermi.gsfc.nasa.gov/ssc/data/analysis/gbm/gbm_data_tools/gdt-docs/

XAI - Anchor rule



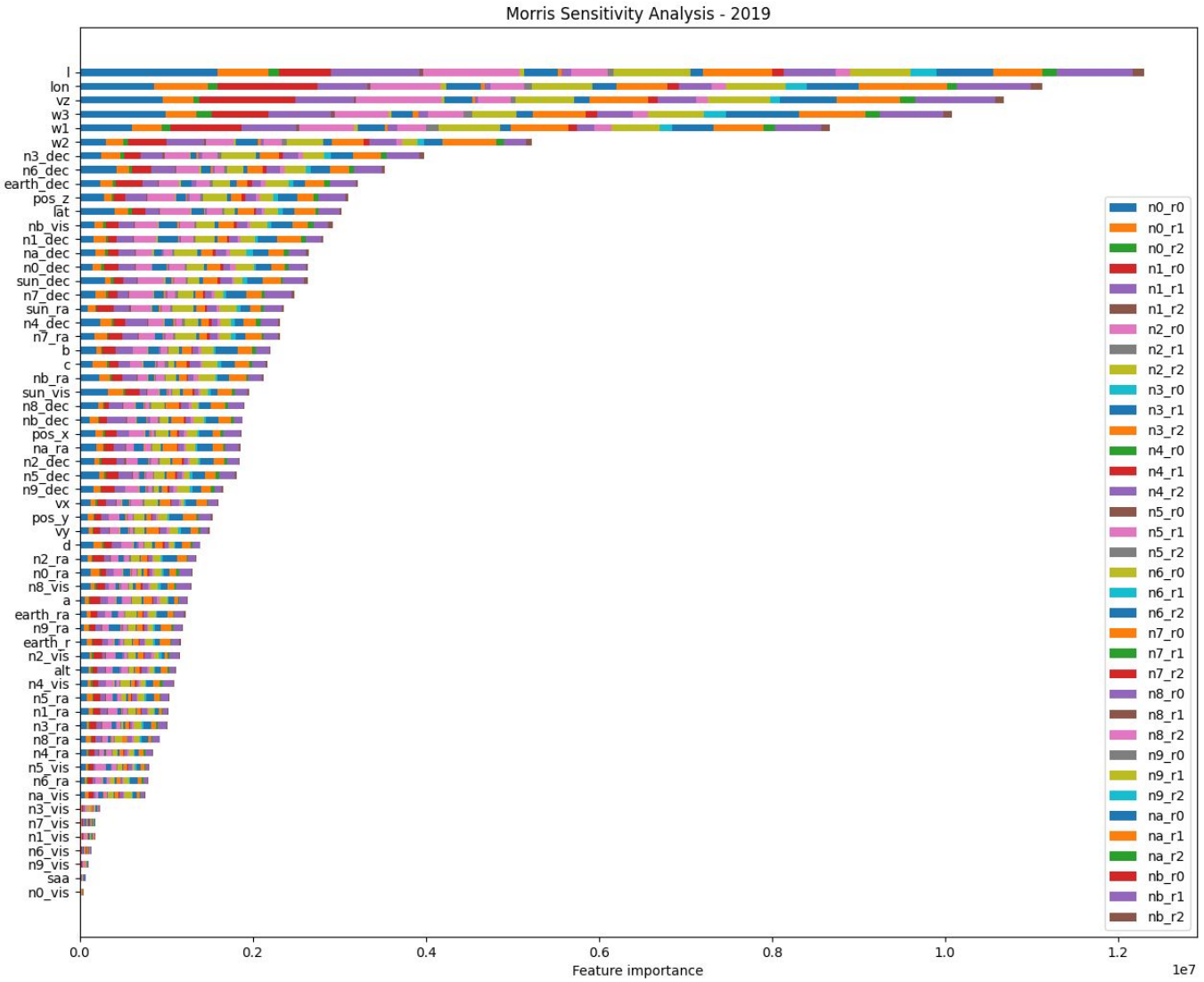


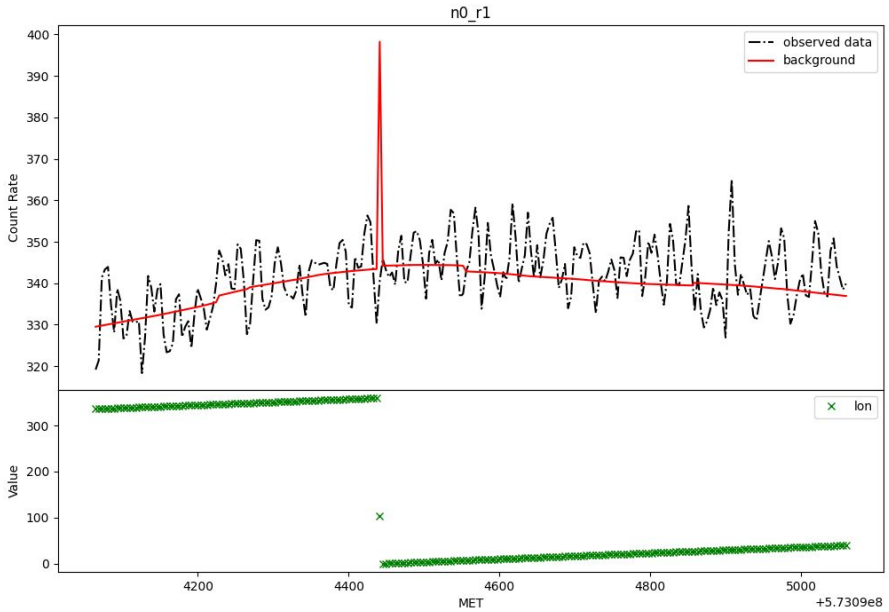
Anchor rule on RF for SF:

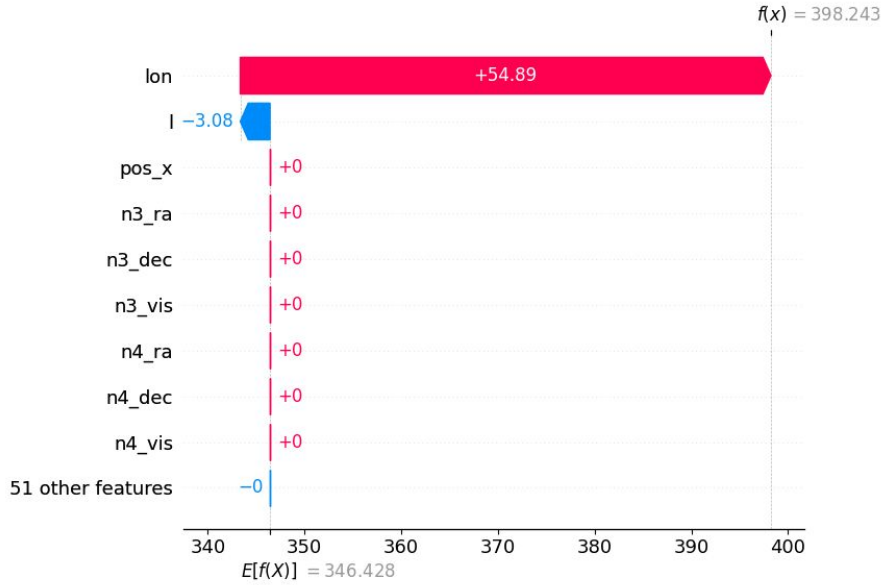
(diff_sun ≤ 32.58 AND HR10 ≤ 0.27)

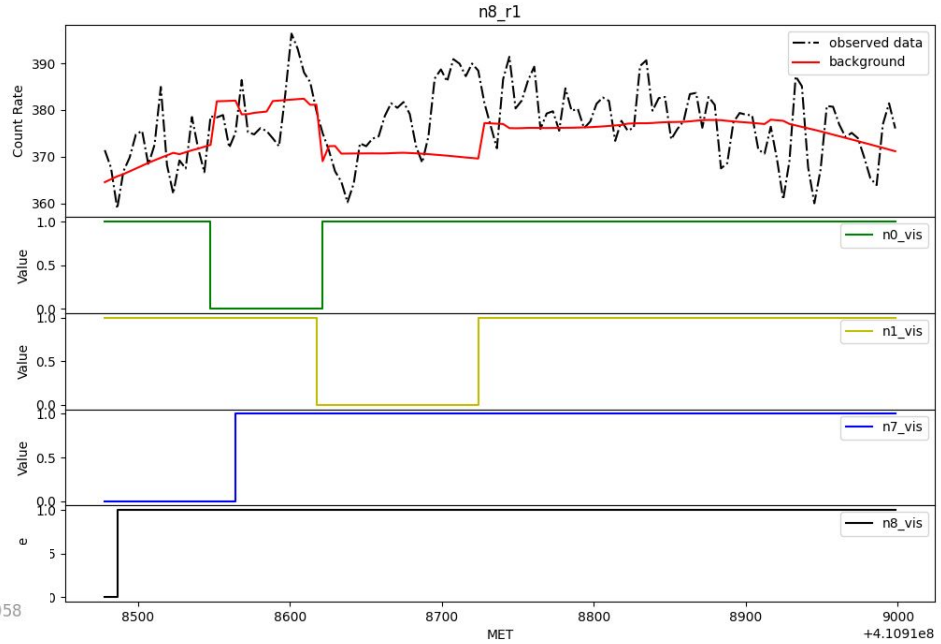
https://github.com/marcotcr/anchor

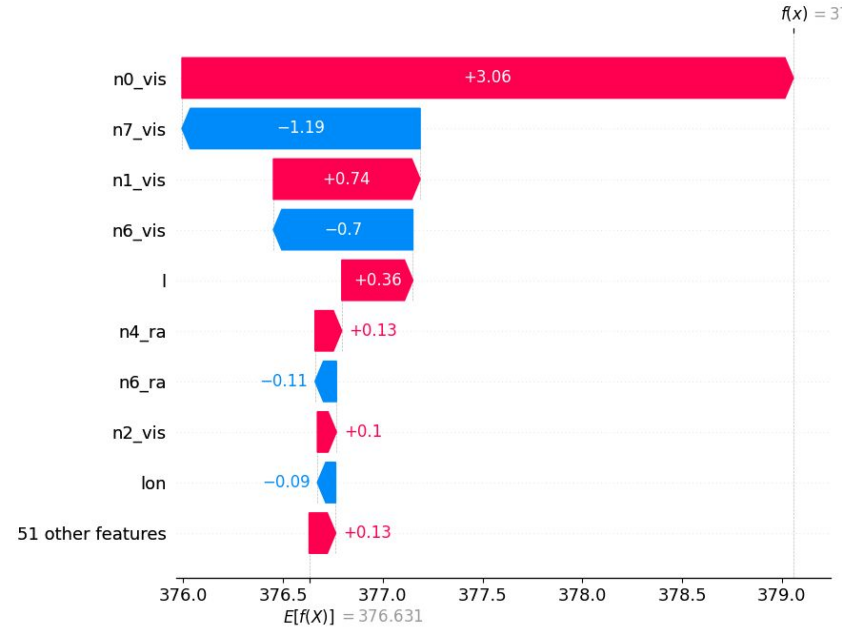
https://www.swpc.noaa.gov/products/goes-x-ray-flux


Morris Sensitivity Analysis - 2019
l
lon
vz
w3
w1
w2
n3_dec
n6_dec
earth_dec
pos_z
lat
nb_vis
n1_dec
na_dec
n0_dec
sun_dec
n7_dec
sun_ra
n4_dec
n7_ra
b
c
nb_ra
sun_vis
n8_dec
nb_dec
pos_x
na_ra
n2_dec
n5_dec
n9_dec
vx
pos_y
vy
d
n2_ra
n0_ra
n8_vis
a
earth_ra
n9_ra
earth_r
n2_vis
alt
n4_vis
n5_ra
n1_ra
n3_ra
n8_ra
n4_ra
n5_vis
n6_ra
na_vis
n3_vis
n7_vis
n1_vis
n6_vis
n9_vis
saa
n0_vis
n0_r0
n0_r1
n0_r2
n1_r0
n1_r1
n1_r2
n2_r0
n2_r1
n2_r2
n3_r0
n3_r1
n3_r2
n4_r0
n4_r1
n4_r2
n5_r0
n5_r1
n5_r2
n6_r0
n6_r1
n6_r2
n7_r0
n7_r1
n7_r2
n8_r0
n8_r1
n8_r2
n9_r0
n9_r1
n9_r2
na_r0
na_r1
na_r2
nb_r0
nb_r1
nb_r2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Feature importance
1e7


n0_r1
observed data
background
Count Rate
400
390
380
370
360
350
340
330
320
lon
Value
300
200
100
0
4200
4400
4600
4800
5000
MET
+5.7309e8


f(x) = 398.243
lon
+54.89
l
−3.08
pos_x
+0
n3_ra
+0
n3_dec
+0
n3_vis
+0
n4_ra
+0
n4_dec
+0
n4_vis
+0
51 other features
−0
340
350
360
370
380
390
400
E[f(X)] = 346.428


n8_r1
observed data
background
Count Rate
390
380
370
360
n0_vis
n1_vis
n7_vis
n8_vis
Value
1.0
0.5
0.0
8500
8600
8700
8800
8900
9000
MET
+4.1091e8


f(x) = 379.058
n0_vis +3.06
n7_vis −1.19
n1_vis +0.74
n6_vis −0.7
l +0.36
n4_ra +0.13
n6_ra −0.11
n2_vis +0.1
lon −0.09
51 other features +0.13
376.0
376.5
377.0
377.5
378.0
378.5
379.0
E[f(X)] = 376.631

DATA CLEANING
BUILD A NEURAL NETWORK, BUILD ON TOP EXPLAINABILITY, FIND THE ERRORS AND CORRECT THE DATASET
imgflip.com