33RD INTERNATIONAL COSMIC RAY CONFERENCE, RIO DE JANEIRO 2013
THE ASTROPARTICLE PHYSICS CONFERENCE

ICRC 2013

The Fermi LAT Third Source Catalog

J. BALLET[1] AND T.H. BURNETT[2] FOR THE FERMI-LAT COLLABORATION.

[1] *Laboratoire AIM, CEA-IRFU/CNRS/Université Paris Diderot, Service d'Astrophysique, CEA Saclay, 91191 Gif sur Yvette, France*
[2] *Department of Physics, University of Washington, Seattle, WA 98195-1560, USA*

Jean.Ballet@cea.fr

Abstract: The Fermi Large Area Telescope has been routinely gathering science data since August 2008, surveying the full sky every three hours. The second source catalog (2FGL) was based on two years of data. We are preparing a third source catalog (3FGL) based on four years of reprocessed data. The reprocessing introduced a more accurate description of the instrument, which resulted in a better high-energy PSF. This improves both the detection threshold and the localization for hard sources. The new catalog also relies on a refined model of Galactic diffuse emission, particularly important for low-latitude soft sources. Finally the association process has also improved, thanks to dedicated multiwavelength follow-up, new surveys and better ways to select potential counterparts likely to be gamma-ray emitters. We describe the construction of this new catalog, its characteristics, and its remaining limitations.

Keywords: icrc2013, Fermi LAT, gamma rays, catalog

1 General context

The *Fermi* Gamma-Ray Space Telescope (*Fermi*) has been surveying the sky since August 2008. The Large Area Telescope (LAT) [7] is the main instrument on *Fermi*, covering the GeV energy range. The most recent LAT source catalog is 2FGL [18] which was based on two years of data and contained 1873 sources. That catalog is not an accurate representation of the γ-ray sky any more: fainter sources than could be reached with two years of data are becoming significant, the average fluxes of variable sources have changed, and newly bright AGN have appeared. For example PMN $1532-1319$ [12] and NVSS J141828+354250 [9] flared brightly in 2011 and 2012 respectively but were not detectable in 2FGL (before August 2010).

In parallel with accumulating new data, developments on the instrument side [10] led to reprocessing all *Fermi* data with new calibration constants, resulting in the Pass 7 reprocesssed data. The main advantage for the source catalog is that it improved the PSF above 10 GeV by 30% or so.

Together, those two reasons justified building an intermediate *Fermi* LAT catalog (3FGL) before the anticipated 5-year catalog. The 3FGL catalog will be based on four years of data. Note that the 5-year catalog, besides the additional time, will also benefit from improved data, since it will be based on an extensive refinement of the reconstruction and classification of LAT event data (Pass 8) [8]. Between 2FGL and 3FGL, we have worked on a dedicated catalog of high-energy (> 10 GeV) sources [19].

2 Gamma-ray sources

Like 2FGL, the 3FGL catalog is a catalog of point sources. We have not included in the process an automatic search for new extended sources. It is also a catalog of the sources detected in the integrated 4-year data set with the current methods (it does not necessarily include short-lived transients). Therefore it is not a superset of 2FGL.

2.1 Data set and diffuse emission

The 3FGL data set covers 2008 August 8 to 2012 August 2, excising intervals around bright GRBs and solar flares, times when the rocking angle of the spacecraft was larger than 52°, and events with zenith angles larger than 100°. The rocking angle remained set at 50° after September 2009, so the survey is slightly non uniform. The maximum exposure is reached at the North celestial pole. At 1 GeV it is 60% larger than the minimum exposure which is reached at the celestial equator.

The Pass 7 reprocessed data do not have event-by-event correspondence with the original Pass 7 data. The reprocessing involves more than changing the event characteristics. So it has a slightly different effective area as well. This means the Galactic diffuse model used for 2FGL (whose spectrum was fit to the data) could not be used directly. Since the additional depth also provides more constraints on the diffuse emission, a new Galactic diffuse model was built. It used the same Galactocentric ring framework as the previous one, fitting templates from CO, HI and IR surveys as well as an inverse-Compton model from Galprop and a preliminary list of point sources [11]. The main conceptual difference is that the large scale residuals were not addressed by sharp-edged patches, but by smoothly varying patches obtained from the data itself on large scales.

Additional templates are the residual Earth limb emission (significant only below 200 MeV around the celestial poles) and the projection on the entire sky of the moving celestial bodies (the Sun and the Moon) [14]. The latter is particularly important for the light curves to avoid false flares, but also reduces false detections along the ecliptic plane.

Events are extracted between 100 MeV and 300 GeV from the Clean class. After four years, the additional instrumental background very nearly offsets the gain in effective area offered by the Source class for sources at the detection limit outside the Galactic plane. We chose to use the Clean class because the Galactic diffuse model was derived using the Clean class, so the catalog is not sensitive to

residual differences at the percent level between the Source and Clean class.

2.2 Data analysis

After the 2FGL paper we realized that there was an inconsistency in the way we handled the *Front* and *Back* events. We computed their contributions to the likelihood function separately using their respective IRFs, but neglected to account for the different instrumental backgrounds. The Earth limb contamination is also different for *Front* and *Back* events. These are now corrected.

After that we noticed differences remained between *Front* and *Back* results in the Galactic plane. We traced that to a slight inconsistency between the *Front* and *Back* effective areas [10], up to 10% at 100 MeV. This resulted in a systematic overestimate (respectively underestimate) of the Galactic diffuse level with *Front* (respectively *Back*) events at low energy, affecting the sources because the Galactic diffuse is about ten times brighter than the sources in the Galactic plane. The ratio of *Front* to *Back* effective areas has been corrected manually (while preserving the total effective area) in the P7REP_V15 IRFs that we used for 3FGL.

Even though the background and IRFs accounted correctly for the sum of *Front* and *Back* events in 2FGL, there remained a net effect on sources because the log(likelihood) function is not symmetric with respect to source flux around its minimum.

The 2FGL analysis resorted to binned likelihood because we had found biases in the way unbinned likelihood worked on simulations. This has been corrected since then and we checked that the binned and unbinned versions of *gtlike* return consistent results. For 3FGL we used binned likelihood below 3 GeV and unbinned likelihood above 3 GeV, in order to limit computing time and memory use without any loss of information. We used Science Tools v9r31p2.

Another improvement since 2FGL is that we put in place an automatic iteration procedure checking that the all-sky result is stable (2FGL used a fixed number of five iterations). This is necessary because a single call to *gtlike* cannot cover the whole sky so at step n sources covering the outer parts of a region of interest (ROI) are taken from nearby ROIs at step $n-1$ and not refitted (unless they are within 2° of sources in the center of the ROI). In practice this changes nothing at high latitude, but improves convergence in the Galactic plane.

Finally we entered more spatially extended sources into the model. Thirteen have been (or will soon be) published since 2FGL and are listed in Table1 (mostly Galactic SNRs and PWNe). The total number of extended sources in the model was 25.

The Galactic plane remains by far the most difficult region of the sky, due to the very strong and structured Galactic diffuse emission and the large source confusion. Separate efforts [4, 5, 13] are ongoing to address specifically pulsars, PWNe and SNRs. Pulsars are handled via timing, and PWNe and SNRs are often viewed as extended sources even at the modest spatial resolution of *Fermi*. Many point sources found in the Galactic plane remain unassociated, and in general their reliability is much less that that of high-latitude sources. Quantitatively, systematic uncertainties due to the imperfections of the Galactic diffuse model are larger than statistical uncertainties in the plane.

Extended Source	Spatial	Spectral	Ref
S 147	Map	PowerLaw	[15]
Puppis A	Disk	PowerLaw	[16]
Vela Junior	Disk	PowerLaw	[21]
HESS J1303-631	2D Gauss	PowerLaw	[5]
HESS J1614-518	Disk	PowerLaw	[16]
HESS J1616-508	Disk	PowerLaw	[16]
HESS J1632-478	Disk	PowerLaw	[16]
RX J1713.7−3946	Map	PowerLaw	[3]
HESS J1837−069	Disk	PowerLaw	[16]
HESS J1841-055	2D Gauss	PowerLaw	[5]
gamma Cygni	Disk	PowerLaw	[16]
Cygnus Cocoon	2D Gauss	PowerLaw	[6]
HB 21	Disk	LogParabola	[20]

Table 1: List of all sources that are modeled as extended sources in 3FGL but were not modeled as extended in 2FGL.

2.3 Source detection and localization

This step was very similar to that in 2FGL. We started with the 2FGL sources, relocalized them and looked for new seeds in the residual TS map. The main difference is that the localization was based on the log(likelihood) from the spectral model (TS_{model} in [18]) rather than that from the sum of 14 bands (TS_{band} in [18]). The former is more sensitive (it does not have as many degrees of freedom) but it requires that the spectral model be a good representation of the data. Particular care was taken to avoid spectral residuals at this step.

The resulting source localization is significantly better than 2FGL, not only for a given source (because of the improved depth) but also at a given significance. This is illustrated in Figure 1 in which we selected sources close to the detection limit ($25 < TS < 100$) outside the Galactic plane. The 3FGL error radii were multiplied by the same 1.1 factor as the 2FGL ones for the comparison. The average (computed in log space) 95% error radius for that sample in 2FGL was 0.129°; it is down to 0.109° in 3FGL. This is the result of two factors: mostly the improved high-energy PSF of the reprocessed data but also the trend that, for a given spectrum, a fainter source in a deeper exposure needs more high-energy photons (resulting in a smaller error radius) to reach the same significance as a brighter source in a shallower exposure because at low energy is the data is dominated by the diffuse background. Quantitatively, the average high-latitude source close to the detection limit is background dominated in the core of the PSF (i.e. the g factor in Eq. A1 of [1] is everywhere less than 1) up to 1 GeV.

2.4 Spectral modeling

The majority of sources are modeled with the same spectral shapes as in 2FGL, namely power-law or LogParabola for ordinary sources, and PLExpCutoff for pulsars. We have more than 130 pulsars in the 3FGL catalog. The main improvements are the following:

- Very bright sources are modeled with a PLSuperExpCutoff model

$$\frac{\mathrm{d}N}{\mathrm{d}E} = K\left(\frac{E}{E_0}\right)^{-\Gamma} \exp\left(-\frac{E^b - E_0^b}{E_c^b}\right) \qquad (1)$$

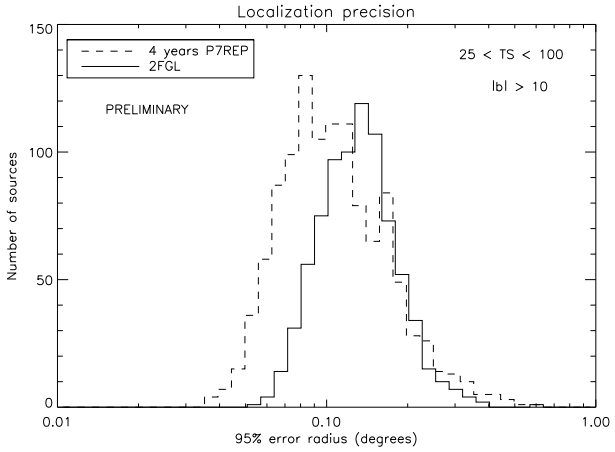

Figure 1: Distribution of the localization precision for faint high-latitude sources in 2FGL (solid line) and the future 3FGL (dashed line). The histograms are offset horizontally by a small amount to avoid overlap.

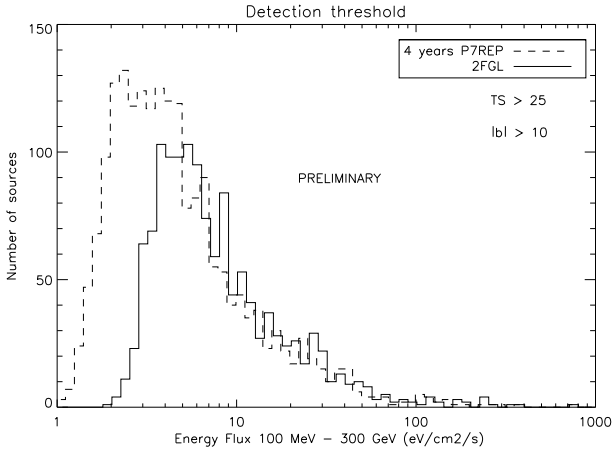

Figure 2: Distribution of the energy flux for high-latitude sources in 2FGL (solid line) and the future 3FGL (dashed line). The histograms are offset horizontally by a small amount to avoid overlays.

which has one more parameter than PLExpCutoff: b defining the shape of the cutoff. In practice b is always less than 1 (so it is a sub-exponential cutoff). In the present version this applies to six sources: PSR J0835 − 4510 (Vela), PSR J0633+1746 (Geminga), 3C 454.4, PSR J1836+5925, PSR J1709 − 4429 and PSR J0007+7303 (CTA 1).

- The Crab is modeled with three components: a PLExpCutoff for the pulsar, a hard power law ($\Gamma = 1.64$ [2]) for the Inverse Compton component of the nebula and a soft power-law for the synchrotron component of the nebula. The Inverse Compton component was left fixed. The other two were fitted.
- The bright sources are accounted for further from the RoI boundary than the faint sources. For 2FGL sources further than 7° from the boundary were never entered into the model.

These considerations are important to improve the low-energy modeling of the vicinity of the very bright sources. The large dynamic range of the *Fermi* sources (a factor 5000 in energy flux between the Vela pulsar and the fainter extragalactic sources) justifies this level of detail.

The detection threshold in 3FGL is significantly better than in 2FGL. Figure 2 shows that the log$N$ - log$S$ distributions are indeed the same at high fluxes, but 3FGL extends further toward low fluxes. The detection threshold in 3FGL is around 3×10^{-12} erg cm^{-2} s^{-1} (it was 5×10^{-12} erg cm^{-2} s^{-1} in 2FGL). Part of the gain, however, is because 3FGL is less uniform than 2FGL because the rocking angle for survey observations was increased in September 2009 resulting in a deeper survey toward the North celestial pole. This, together with the spectral index distribution, contributes to the broad "flat top" in the histogram. The better PSF with Pass 7 reprocessed data also reduces the detection threshold, particularly for hard sources. In terms of numbers, 3FGL contains more than 2500 sources vs. 1873 in 2FGL.

The catalog format will evolve a little from 2FGL. It will contain additional columns `Exp_Index` and `Unc_Exp_Index` to reflect parameter b in Eq. 1. In order to make SED extraction easier, it will also contain for each band a column containing explicitly the energy flux density e.g., `nuFnu100_300` for the first band (100 to 300 MeV). Since those fluxes are obtained using a fixed (power-law) spectral representation, the relative error on νF_ν is the same as that on the photon flux, so it does not require a separate uncertainty. A robust estimate of νF_ν averaged over an energy band $[E_1, E_2]$ is obtained by dividing the energy flux in that band by the width of the band in natural log, $\ln(E_2/E_1)$. The energy flux is obtained from the photon flux (given in the catalog) using the same power-law shape as in the fitting step. For curved spectra the spectral index is set, as in 2FGL, to the local spectral slope at the logarithmic mid-point of the band $\sqrt{E_1 E_2}$, restricted to be in the interval [0,5].

3 Associations

The improved localization of 3FGL sources (Figure 1) facilitates the association procedure (by reducing the fraction of random associations). On the other hand the lower detection threshold (Figure 2) makes it more difficult, as the counterparts are expected to be fainter as well, possibly reaching the detection threshold of the counterpart catalogs.

At this point the association process is only beginning. What can be done now is only compare with the 2FGL process at the same point (rather than the end result). Concentrating on point sources at high latitude ($|b| > 10°$), the early fraction of associated sources in 2FGL was 68% (the final fraction was 79%). Using the same counterpart catalogs on 3FGL results in 66% of associated sources. Adding the recent catalog of WISE blazars [17] increases that fraction to 68%, similar to 2FGL. So early prospects are that the fraction of associated sources may remain similar to 2FGL. This implies that more than 400 additional high-latitude sources would get an association in 3FGL. Among the new sources (not obviously associated to a 2FGL source), more than half can be associated to a counterpart from external catalogs.

4 Summary

The *Fermi* LAT collaboration is preparing the 3FGL catalog, based on four years of Pass 7 reprocessed data. This catalog will contain more than 2500 γ-ray sources, reaching down to 3×10^{-12} erg cm^{-2} s^{-1}. The average 95% error radius of faint sources is 0.11°, allowing to associate a large fraction of the new sources to known counterparts. Systematic uncertainties due to the Galactic diffuse emission remain large in the Galactic plane.

Acknowledgments: The *Fermi* LAT Collaboration acknowledges support from a number of agencies and institutes for both development and the operation of the LAT as well as scientific data analysis. These include NASA and DOE in the United States, CEA/Irfu and IN2P3/CNRS in France, ASI and INFN in Italy, MEXT, KEK, and JAXA in Japan, and the K. A. Wallenberg Foundation, the Swedish Research Council and the National Space Board in Sweden. Additional support from INAF in Italy and CNES in France for science analysis during the operations phase is also gratefully acknowledged.

References

[1] A.A. Abdo et al., ApJS 188 (2010), 405
[2] A.A. Abdo et al., ApJ 708 (2010), 1254
[3] A.A. Abdo et al., ApJ 734 (2011), 28
[4] A.A. Abdo et al., arXiv:1305.4385 (2013), accepted in ApJS
[5] F. Acero et al., arXiv:1306.5735 (2013), accepted in ApJ
[6] M. Ackermann et al., Science 334 (2011), 1103
[7] W.B. Atwood et al., ApJ 697 (2009), 1071
[8] W. Atwood et al., arXiv:1303.3514 (2013)
[9] E. Bernieri, R. Campana, E. Massaro, A. Paggi and A. Tramacere, A&A 551 (2013), L5
[10] J. Bregeon, E. Charles and M. Wood, arXiv:1304.5456 (2013)
[11] J.M. Casandjian, 33rd ICRC (2013), 0966
[12] D. Gasparrini and S. Cutini, ATel 3579 (2011)
[13] J.W. Hewitt, 33rd ICRC (2013), 0785
[14] G. Johannesson and E. Orlando, 33rd ICRC (2013), 0957
[15] J. Katsuta et al., ApJ 752 (2012), 135
[16] J. Lande et al., ApJ 756 (2012), 5
[17] F. Massaro et al., ApJ 750 (2012), 138
[18] P.L. Nolan et al., ApJS 199 (2012), 31
[19] D. Paneque et al., arXiv:1306.6772 (2013), submitted to ApJS
[20] I. Reichardt, E. de Ona Wilhelmi, J. Rico and R. Yang, A&A 546 (2012), A21
[21] T. Tanaka et al., ApJ 740 (2011), L51