Analyzing Gamma Sources

Using the Databases and the Likelihood Tool (A1)

Patrick Nolan, Stanford University

GLAST collaboration meeting, Rome

16 September 2003
Analyzing Gamma Sources: the Steps

Once the pipeline processing is done, the data analyst will want to select a set of photons and find information about astrophysical sources. Here I describe the steps to look for one or more point sources.

There are two main operations:
• Acquire the photon and spacecraft data from the D1 (photon list) and D2 (spacecraft pointing & deadtime) databases using the web-based U1/U3 tools.
• Analyze the data using the A1 Likelihood tool.

In this type of analysis we assume that all sources emit steadily. Variation can be studied by repeating this analysis for several time intervals. Undoubtedly there will be more advanced tools to handle this problem later.

The tools described here will be ready for DC1 in November. They are currently in alpha or beta stages of development. The GRB detection and analysis path (via XSPEC) won't be described here.
U1/U3 Web Interface Page Layout

- Database selection (D1/D2/Both)
- Position search
- and/or Time search
- and/or Energy search

GLAST Event

Which database do you want to query?

- Database
- File
- spacecraft data

Do you want to search around a position?

- Coordinates
- Area
- Circle

For the circle, enter the radius in degrees.
For a rectangle, enter the length, width, and rotation.
For an ellipse, enter the semi-major and semi-minor axes, and rotation.
Rotation is defined astronomically. Zero degrees is north, positive angles are to the east.

Search area shape:
- circle (default)
- box
- ellipse

and/or search by time?

- Time

Start and end times for search.

- MJD
- Gregorian (UTC)

and/or search by energy?

- Energy

Minimum and maximum energy to search.
U1/U3 Web Results Page

GLAST Event

Your query was successful.

GLAST Event Data
local/data/fafnr3/jstephen/data/repository/D1/U1WEBB101281831615542321610_results.fits

GLAST Spacecraft Data
local/data/fafnr3/jstephen/data/repository/D2/U1WEBB101281831615542321610_02_results.fits

These file(s) will remain on the public FTP server for approximately 30 days. Please retrieve them as soon as possible. For future reference, your query ID is U1WEBB101281831615542321610

Thank you!
The GLAST SSC Team

Glast SSC D1 Web Interface
A1

• The function of the A1 (likelihood) tool is to produce a maximum-likelihood fit of photon data with a model of gamma-ray emitters. The model includes adjustable parameters for multiple point sources and diffuse emitters.
• The same tool can do several types of analysis: detecting sources, measuring spectra and positions, measuring variability, and characterizing extended sources.
• We analyze the data by model-fitting because of low fluxes, a bright foreground from the Milky Way, and overlap of nearby point sources.
• The major parts of the tool exist in CVS: packages Likelihood and optimizers. There’s no user interface yet.
The required inputs are

- FITS files for the photon list (D1) and spacecraft data (D2).
- Instrument response functions (IRFs) from CALDB: effective area, point spread function, energy dispersion.
- A precomputed FITS exposure map based on D2 and IRFs.
- XML files containing spectrum models, diffuse source models, and cuts.
- An XML file containing a list of point sources with initial guesses for positions and spectral indices.

This list shows the need for the U4 and U7 tools to compute the exposure file and to make the initial list of point sources.

Currently-available components for spectrum models are power laws, gaussian lines, and absorption edges.

Sources can be points, isotropic, or arbitrary shapes defined by FITS files.
Simulation

• The development process relies heavily on realistic simulated data. Jim Chiang developed observationSim, which uses the “glast25” instrument parameters to create photons. Any residual cosmic-ray background is ignored. It uses the EGRET catalog and diffuse models, augmented with faint point sources chosen randomly from plausible distributions.
30-Day Counts Maps of the Cygnus Region

30-65 MeV

65-140 MeV

140-300 MeV

300-650 MeV

650-1400 MeV

1.4-3 GeV

Galactic Longitude

Galactic Latitude

Counts per pixel

EGRET sources
Exposure

• It is important to keep track of where the instrument points and how much instrumental dead time there is. As it scans across a point on the sky, the effective area and point spread function change continuously. Each photon will be detected by an instrument with different characteristics. Model-fitting procedures need to take this into account.

• Rocking the spacecraft will partly even out the exposure, but not completely.

• Photons from the Earth’s horizon must be eliminated. This cut will complicate all the convolutions of response functions.

• If sources are assumed to emit steadily, it is possible to accumulate maps of exposure by convolving the response functions with the pointing and dead time histories. An average point spread function can be created for each source. This is the current level of development for A1.
30-Day Exposure Maps of the Cygnus Region

Based on “glast25” IRFs
Finding point sources is an iterative process. Some method is needed to produce an initial list of sources, with rough estimates of their positions and perhaps fluxes. The U7 source definition tool will support one or more methods.

Candidates include:
- EGRET-style TS map
  - $\ln(L)$ for a test source
  - Known to be slow
- Jeff Scargle’s Bayesian Blocks
- Wavelets
- Independent Component Analysis?
- Catalog of sources in other bands
- Eyeball
- ?
XML

XML is being used routinely for specifying information used by programs, such as cuts.

```xml
<?xml version='1.0' standalone='no'?>
<!DOCTYPE Region-of-Interest SYSTEM "RoiCuts.dtd">
<Region-of-Interest title="Cygnus Region">
  <timeInterval start="0" stop="4" unit="days"/>
  <energies emin="30." emax="3.1623e5" unit="MeV"/>
  <acceptanceCone longitude="75." latitude="0." radius="30" coordsys="Galactic"/>
</Region-of-Interest>
```

and source lists for simulations

```xml
<source_library title="Simulated sources in Cygnus">
  <!-- Center of field (deg): l,b = 75.0000, 0.000000     Radius of field (deg): 50.0000     Flux limit (>100 MeV): 1.00000e-008     Number of sources: 369     Mean and dispersion of spectral index: 2.00000, 0.400000     Written by: gen_xml.pro, Wed Sep 03 07:50:02 2003 UTC -->
  <source name="cat_001" flux = "0.00616736">
    <spectrum escale="MeV">
      <particle name="gamma">
        <power_law emin="30." emax="100000." gamma="1.31319"/>
      </particle>
      <galactic_dir l="119.920" b="10.5400"/>
    </spectrum>
  </source>
  etc....
</source_library>
```
and spatial models for fitting

```xml
<?xml version='1.0' standalone='no'?>
<!DOCTYPE source_library SYSTEM "A1_Sources.dtd" >
<source_library title="prototype sources" function_library="A1_Functions.xml">
  <source name="Bright Point Source" type="PointSource">
    <spectrum type="Bright Source Power-law"></spectrum>
    <spatialModel type="SkyDirFunction">
      <parameter name="RA" value="0" unit="degrees" min="-180" max="180" free="false" scale="1" />
      <parameter name="DEC" value="0" unit="degrees" min="-90" max="90" free="false" scale="1" />
    </spatialModel>
  </source>
  <source name="Galactic Diffuse Emission" type="DiffuseSource">
    <spectrum type="Galactic Diffuse Power-law"></spectrum>
    <spatialModel type="SpatialMap" file="gas.cel"></spatialModel>
  </source>
  <source name="Extragalactic Diffuse Emission" type="DiffuseSource">
    <spectrum type="Extragalactic Diffuse Powerlaw"></spectrum>
    <spatialModel type="ConstantValue">
      <parameter name="Value" value="1" unit="none" min="0" max="10" free="false" scale="1" />
    </spatialModel>
  </source>
</source_library>
```

and spectrum models with constraints

```xml
<?xml version='1.0' standalone='no'?>
<!DOCTYPE function_library SYSTEM "A1_Functions.dtd" >
<function_library title="prototype Functions">
  <function name="Galactic Diffuse Power-law" type="PowerLaw">
    <!-- Use this prototype with the EGRET Galactic diffuse model map for E > 100 MeV. -->
    <parameter name="Prefactor" value="6.31" unit="cm^-2 s^-1 MeV^-1" min="1e-3" max="1e3" free="true" scale="1e-3" />
    <parameter name="Index" value="-2.1" unit="none" min="-3.5" max="-1" free="true" scale="1" />
    <parameter name="Scale" value="100" unit="none" min="50" max="200" free="false" scale="1" />
  </function>
</function_library>
```
The basic Likelihood and optimizer packages in C++ are fairly well developed. They need glue to put the parts together.

SWIG is being used to wrap C++ classes for use with python. This makes it easy to write python scripts to put the pieces together. We are currently studying how to integrate SWIG with CMT. So far it only works on Linux. It’s sensitive to the versions of SWIG and python.

The current optimizers are Minuit and quasi-Newton methods. These take no account of the (possibly) sparse nature of the problem, nor of the structure of the likelihood function.
Performance

- It’s really slow. A recent test run used 4.7 days of data from the Cygnus region with no diffuse emission (~10K events). The model included 19 of the brightest point sources, held at fixed positions. The fluxes and spectral indices were allowed to vary, for a total of 38 adjustable parameters. Real problems are likely to be considerably larger.

- On a 1.8 GHz computer, the optimizer converged to a satisfactory answer in about 1000 seconds. This is not suitable for interactive data exploration.
To Be Done

- Clean up the development environment
- Speed up the optimizer:
  - Find a specialized optimizer for sparse problems?
  - Use the EM algorithm for optimizing?
  - Use shortcuts in convolving the response functions?
- Settle on a method to find the initial set of sources.
- Allow point source positions to move.
- Perhaps make a GUI, if there’s time.