GLAST Data Access and Analysis

Don Horner (RSIS/GSFC) and the GLAST Science Support Center Team

For more information please visit the GSSC Web Site: glast.gsfc.nasa.gov/ssc

Abstract

The scientific community will access the public GLAST data through the website of the GLAST Science Support Center (GSSC). For most data products the GSSC website will link to the NASA High Energy Astrophysics Science Archive Research Center's (HEASARC) Browse interface, which will actually serve the data. For example, data from the GLAST Burst Monitor (GBM) from a given burst will be packaged together and accessible through Browse. However, the photon data produced by the Large Area Telescope (LAT), GLAST's primary instrument, will require a customized GSSC interface. These photons will be detected over the LAT's large field-of-view, usually while the LAT is scanning the sky, and thus cannot be attributed to the observation of a particular object. Users will request all photons detected from a region on the sky over a specified time and energy range. This also means that analyzing data from the LAT will require sophisticated techniques. The GSSC will provide a suite of data analysis tools and libraries for use in analyzing the GLAST data. This software is being developed by the instrument teams with assistance from the GSSC to provide a solid and robust framework for interpreting and analyzing the GLAST data. Here we provide a short overview of the Standard Analysis Environment (SAE) software and the data products that will be served.

GLAST Mission Overview

GLAST is an international and multi-agency space mission that will study the cosmos in the 10 keV-300 GeV energy range. The main instrument, the Large Area Telescope (**LAT**), is a pair conversion telescope. It will have an effective area (>8000 cm^2), angular resolution (<3.5º @ 100 MeV, <0.15º @ >10 GeV), field of view (>2 sr), and deadtime (<100 µs) that will provide a factor of 30 or more advance in sensitivity compared to previous missions, as well the capability for studying transient phenomena. Data from the LAT are processed by the LAT Instrument Science Operations Center (**LISOC**) before being sent to the GSSC. The other instrument on GLAST, the GLAST Burst Monitor (**GBM**), consists of 14 detectors that will monitor >8 sr of the sky, including the LAT's field of view. Gamma ray bursts will be localized to 9° (1σ, brightest 40% of the bursts) by comparing the rates in different detectors. The GBM also extends GLAST's burst spectral sensitivity to the <10 keV to >25 MeV band. Data from the GBM are processed by the GBM Instrument Operations Center (**GIOC**) before being sent to the GSSC. Although pointed observations will be possible, the observatory will probably scan the sky continuously because of the LAT's large field of view; this survey mode is planned for at least GLAST's first year.

GLAST Data Policy and Access

All GBM data is public immediately from the beginning of the mission. During the first year of the mission, LAT event data are proprietary to the instrument team and the Interdisciplinary scientists, although lightcurves and spectra of detected transients and ~20 selected sources will be made public as soon as possible. A month after the end of the first year, the events data will become publicly available.

Starting the second year, all subsequent science data acquired by the spacecraft will be immediately in the public domain without a proprietary data period. Full details on the GLAST Data Policy can be found on the GSSC web site.

All public data from the GLAST mission will be available through the GSSC's website. Much of the data will be served through the HEASARC Browse (an interface to all of NASA's high energy astrophysics data from both current and previous missions). The GSSC website will link to this interface. Those data not available through Browse will be served directly from the GSSC's website.

Data Properties and Data Serving

It is impossible to define an "observation" for the LAT in the traditional sense. Given the large field of view of the LAT and the continuously scanning operating mode, any given object is constantly moving into and out of the instrument's field of view and is typically observed for ~30 minutes every 3 hours of observing. Also, the low angular resolution at low energies also results in photons from one source overlapping with photons from nearby sources and requires simultaneous fitting to disentangle closely spaced objects. Typically a user will need data from a part of the sky equal to a 20-30° diameter field of view to fully analyze an object at the center of that field. This corresponds to ~1-2% of the entire sky for a single analysis.

The LAT will trigger on a large number of events, only a small fraction of which will be actual photons. Filtering will reduce the ~2.5 kHz of triggers to ~30 Hz delivered to the GSSC from which ~2-5 Hz of photons will extracted. This corresponds to ~100 million photons and ~1 billion events a year. With the current data record specifications, this translates into ~10 GB of photon data and ~1 TB of event data a year with a typical photon query returning 100-200 Mbytes of data.

It is also expected that the LAT will detect thousands of sources. The large number of sources, combined with the moderately large data volume and high degree of source overlap make it unrealistic to prepackage the data on a source by source or even region by region basis. This would result in an unnecessary duplication of the same data multiple times in the data system.

All of these factors point toward the development of a dynamic data retrieval system that retrieves the data as needed according to the user's unique specifications instead of relying on static prepackaged files.

GBM Data

Data products for the GBM will be similar to those of the BATSE instrument on the Compton Gamma Ray Observatory. For bursts, catalogs with derived parameters, which include preliminary lightcurves and spectral fits, will be provided through Browse. The catalog entries will link to data packages which will include accumulated counts (CTIME and CSPEC files), time tagged events, response functions, backgrounds, and other files necessary to analyze the data. New bursts will be added to the catalog as soon as the data is received from the GIOC. Although burst alert notices will go out immediately, the data will not enter the archive until the GIOC is finished processing the data. The entire day of data is processed by the GIOC to account for background variability before any data is sent to the GSSC.

LAT Data

Photon and event data will be served through the GSSC's LAT data server. The GSSC populates the photon database based on event data received from the LISOC.

Events are all triggers reconstructed by the LISOC and will include, e.g., electrons and protons in addition to photons. There are about 200 parameters associated with each event. Based on these parameters, a subset of events are classified as gamma-ray photons. We then use only the events classified as photons and extract about 20 of the 200 parameters for each photon that are most useful for scientific analysis to populate the photon database. Most users will only need the photon data, but the event data will be there for those who wish to rescreen the data themselves.

New data will be added to the sever as soon as they are received from the LISOC. This will occur within 72 hours of the data being taken but likely in much less time. A variety of source catalogs will be served through Browse. Burst and transient catalogs will be updated on a per event basis. Other catalogs will be updated infrequently. Also available will be an interstellar emission model. This will be refined as the mission progresses and updated infrequently.

LAT Data Server

The LAT Data Server will provide users with access to events, photons, and position and attitude history from the LAT. It will accessed by users through a customized web interface hosted by the GSSC.

The user enters the desired position with optional energy and time cuts and will then be served a page with the estimated query time and a link to a results page. When the data files are ready (FITS files), they will be made available for download on the results page.

The data server is implemented as a small cluster of Linux PCs which operate on event lists stored in FITS files. Based on the user's query, the photons matching the data cuts are extracted and presented to the user as a downloadable FITS file. The photon and event databases will be updated immediately when new data are received by the GSSC from the LAT instrument team.

Spacecraft Data

In addition to the science data, various data products dealing with the spacecraft will be available via the GSSC website to help proposers and observers. These will include proposed and observed science timelines for coordinating simultaneous multi-wavelength observations, spacecraft alerts, pointing and livetime history, and various other spacecraft related data.

Data Properties and Data Analysis

The LAT will detect photons that can be used in data analysis up to ~66° off-axis. The effective area decreases off-axis, but the solid angle increases, and therefore a large fraction of the data will be taken off-axis. The GBM will detect bursts down to the Earth's limb. While GLAST can point at individual sources, there will rarely be any advantage because of the LAT's large FOV. Usually GLAST will rock ~35° above and below the orbital plane once per orbit for uniform sky coverage.

The LAT's PSF will be ~3.5º at 100 MeV, <0.15º at 10 GeV (68% containment radius) with significant tails that will decrease as a power law with radius. LAT sources will be observed against a bright spatially varying Galactic and isotropic Extragalactic diffuse background. The average LAT count rate from astrophysical sources (including the background) will be 2-3 counts/s.

Most persistent sources will be observed by the LAT at a variety of detector orientations; each count must be analyzed using the response function appropriate for the detector orientation when the count was detected. Counts from different sources, including the diffuse background, will often overlap.

The Standard Analysis Environment

The GSSC will provide a suite of data analysis tools and libraries for use in analyzing the GLAST data called the **Standard Analysis Environment (SAE)** which is being developed by the instrument teams with assistance from the GSSC.

General Analysis: The SAE will consist of several general purpose tools to assist in analyzing data including a data sub-selection tool, tools to generate source models and extract source parameters from existing catalogs, and the workhorse of the GLAST data analysis, the Likelihood tool to perform maximum likelihood fits of the data with the specified models (see below). The suite also provides an event binning tool to create time, energy and spatially binned data sets and tools to compute exposure and response matrices.

GRBs: The SAE will provide several tools to assist in the study of gamma-ray bursts including tools for spectral and temporal data analysis and model fitting as well as tools for generating the necessary response functions and binning events for analyzing GLAST data with existing tools such as XSPEC. These tools can be used to analyze both LAT and GBM data, either individually or simultaneously.

Pulsars: The SAE will provide a number of tools to assist in pulsar analysis including a barycenter arrival time correction tool, period search and profiling tools, and a pulsar ephemeris extraction tool to retrieve pulsar ephemerides from a pulsar database.

Data Simulation: The SAE also provides an observation simulator that can simulate LAT data based on an input source model and optional spacecraft orbit profile.

FTOOLS—The SAE will be standard FTOOLS. A GUI interface will also be provided. Generic FTOOLS utilities can be used on the GLAST data files.

Documentation—The GSSC website will provide online and printable manuals include an Installation Guide, reference manual (description of all inputs to each tool), analysis threads (step-by-step examples), and a detailed manual (the methodology implemented in each tool).

Likelihood Tools

The LAT's large effective area, but low angular resolution, especially at low energies, means that many sources will be detected but that their PSFs will merge at low energies. Therefore, source analysis must include both spectral, spatial, and possibly temporal information. Since the LAT will usually survey the sky, a source will be also observed at different instrument orientations. Consequently, the source model will have many parameters.

For a typical analysis the source model must include: all point sources within a few PSF radii of the region of interest; diffuse sources (e.g., supernova remnants); diffuse spatially variable Galactic emission (which must be modeled); and diffuse isotropic extragalactic emission. A very large data space results. Even with 10^5 counts, this data space will be sparsely populated.

Therefore maximum likelihood tools are the foundation of our analyses (e.g., detecting sources, determining source intensities, fitting spectral parameters, setting upper limits). The likelihood is the probability of the data (the counts that were detected) given the model (the photon sources).

The data consist of both the counts that were detected, and the regions of parameter space where counts were not observed. A comparison of the likelihood for different models (e.g., with and without an additional point source) will indicate which model is preferred. The best-fit model parameters and their confidence regions will be calculated from the likelihood as a function of the model parameters.

GLAST Data Products