D1. Event summary (Level 1) database

Date: 7 May 2002 (draft v2, now includes sun and moon angles)
Contributors: S. Digel (SU-HEPL)

Function

This is the Level 1 (gamma-ray summary) database. It is not directly accessed by the analysis tools. It receives queries from, and passes data to, the Data Extractor (U2).

Inputs (for gamma rays)

See inputs for the Data Extractor (U2).

Databases required

NA

Outputs

The provisional contents of the database, defined in the report of the Data Products Working Group as LS-002 (and updated to include the angles of the sun and moon – see U3), are as follows:

	Column Name	Units
1	Energy of event	GeV
2	1-sigma uncertainty of energy	GeV
3	RA, J2000	deg
4	Dec, J2000	deg
5	Localization uncertainty, est. 1-sigma radius	deg
6	Time (Mission Elapsed Time)	s
7	Conversion layer	dimensionless
8	Number of SSD hits	dimensionless
9	Number of tracker hits NOT reconstructed	dimensionless
10	Conversion point, (x,y,z)	m
11	Reconstruction trajectory-event (dir. cosine)	dimensionless
12	Reconstruction trajectory-secondary 1 (dir. cosine)	dimensionless
13	Reconstruction trajectory secondary 2 (dir. cosine)	dimensionless
14	Secondary energies	GeV
15	ACD tiles hit (bit flags)	dimensionless
16	Quality_Parm--quality parameter for fitted trajectory	dimensionless
17	Data_Quality--Overall status of event	dimensionless
18	Deadtime accumulated since start of mission	s
19	Instrument mode (slew, diagnostic, ...)	dimensionless
20	TKR, CAL-only flag--2bit for CAL, TKR or Both	dimensionless
21	Zenith angle	deg
22	Earth azimuth angle (from north to east)	deg
23	S/C position from earth center, (x,y,z)*	km

24	Angle from sun	deg
25	Angle from moon	deg
26	Ground point--lat.	deg
27	Ground point--lon.	deg
28	Barycenter arrival time of event	s
29	Offset of Solar System Barycenter from S/C (x,y,z)	s (light travel time)
30	McIlwain B	gauss
31	McIlwain L	Earth radii
32	Geomagnetic Lat.	deg
33	Geomagnetic Long.	deg
34	Recon. version	dimensionless
35	Calib. table versions	dimensionless
36	Multiple event reconstruction	dimensionless
37	Reconstruction number within event	dimensionless
38	Original Event ID	dimensionless

* x-axis is direction RA,DEC = 0,0, z-axis is north, y-axis defined for earth-centered right-hand coordinate system, J2000

Also from the DPWG report:

- The reconstructed trajectory of the photon (item 13) is in instrument coordinates, and so specifies the inclination angle and azimuth of the photon.
- Instrument modes (item 19) should probably be defined in a second extension.
- Perhaps multi-gamma events should just have a 'primary' gamma ray defined here, plus a flag set to indicate that special processing should be done.
- Note that >1 photon can be claimed by Reconstruction (may translate into multiple Event Summary entries, with number_of_photons entry)
- Quality flags above are intended to be representative of the background rejection/PSF enhancement cut parameters

Performance requirements

Performance requirements will be considered in detail in the full requirements document. Requirements related to ingest rate, query service time, and database backup and restore have been identified.

Other modules required

None.

Host environment

Database server system

Existing counterparts

Nothing directly applicable. Spatial indexing methods (heirarchical tessellations of the sky) for facilitating access to data by region of the sky have been designed for other astronomical applications. Two widely-used indexing schemes (HTM and HEALpix) are well documented and some software is available for efficient coordinate-spatial index conversion is available.

Open issues for definition or implementation

1. Will there be a corresponding Event summary database that includes the events identified as cosmic rays?

2. Some input parameters for the Data Extractor (U2) might not be passed on to the Event summary database, depending on whether the Data Extractor does any ‘post processing’ on retrieved data. This is probably neither here nor there as far as the requirements summaries go, however.

D2. Pointing, livetime, and mode history

Date: 7 May 2002 (draft v2, added sun and moon positions)
Contributors: D. Band (GSFC-UMBC), S. Digel (SU-HEPL)

Function

This is the database of pointing and observation history that is needed to calculate exposures. It contains information about the orientation, location, and operation of the LAT for regular time intervals, ~30 s. It is not directly accessed by the analysis tools. Instead, it receives queries from, and passes data to, the Pointing/livetime history extractor (U8). The positions of the sun and moon are included here solely to facilitate cuts on their positions in the generation of exposure. They are both gamma-ray sources (the sun impulsively) and both of course shadow sources they pass in front of.

Inputs

See inputs for the Pointing/livetime history extractor (U8).

Databases required

NA

Outputs

The provisional contents of the database, defined in the report of the Data Products Working Group as LS-005 and augmented here to include the SAA flag and positions of the sun and moon, are as follows:

	Contents	Units
1	starting time of interval (Mission Elapsed Time)	s
2	ending time of interval (Mission Elapsed Time)	s

3	position of S/C at start of interval (x,y,z inertial coordinates)	km
4	viewing direction at start (LAT +z axis), 2 angles	dimensionless
5	orientation at start (LAT +x axis), 2 angles	dimensionless
6	zenith direction at start, 2 angles	dimensionless
7	LAT operation mode	dimensionless
8	livetime	s
9	SAA flag	dimensionless
10	S/C longitude	deg
11	S/C longitude	deg
12	S/C altitude	km
13	direction of the sun, 2 angles	Deg
14	direction of the moon, 2 angles	

Performance requirements

Performance requirements will be considered in detail in the full requirements document. None of the requirements is likely to be challenging in terms of implementation of the database.

Other modules required

None.

Host environment

Database server system

Existing counterparts

Nothing directly applicable.

Open issues for definition or implementation

1. Should latitude, longitude, and altitude be translated into geomagnetic coordinates in anticipation of any need to generate exposure maps for different ranges of geomagnetic conditions? (This might be useful, e.g., if the background rejection is not extremely good.)

D3. LAT point source catalog

Date: 25 April 2002 (draft v1)
Contributors: S. Digel (SU-HEPL)

Function

This is the online form of the point source catalog. It is not directly accessed by the analysis tools, but instead receives queries from, and passes data to, the Catalog Access tool (U5).

Inputs

See inputs for the Catalog Access tool (U5).

Databases required

NA

Outputs

The provisional contents of the database, defined in the report of the Data Products Working Group as LS-008, are as follows:

	Contents	Units
1	source name ("telephone number")	dimensionless
2	RA	deg
3	Dec	deg
4	th68 semimajor, semiminor axis, and position angle	deg
5	th95 semimajor, semiminor axis, and position angle	deg
6	flux (>100 MeV, avg. for the time interval of the catalog)	cm-2 s-1
7	flux uncertainty, 1 sigma (as above)	cm-2 s-1
8	photon spectral index (avg)	dimensionless
9	variability index	dimensionless
10	significance (avg)	dimensionless
11	significance (peak)	dimensionless
12	peak flux (for time interval above?)	cm-2 s-1
13	peak flux uncertainty	cm-2 s-1
14	time of peak flux (wrt reference date)	s
15	interval of time	s
16	flux history	cm-2 s-1
17	flux uncertainty, 1 sigma (as above)	cm-2 s-1
18	start times of flux history entries	s
19	end times of flux history entries	s
20	candidate counterparts	dimensionless
21	degrees of confidence for the counterparts	dimensionless
22	flags (confusion, low latitude,…)	dimensionless

Also from the DPWG report:
For source naming conventions, see http://cdsweb.u-strasbg.fr/viz-bin/DicForm

Performance requirements

Performance requirements will be considered in detail in the full requirements document. This database is not expected to be particularly large or to have particularly demanding access requirements.

Other modules required

None.

Host environment

Database server system

Existing counterparts

Nothing directly applicable. Spatial indexing methods (heirarchical tessellations of the sky) like those developed for the Level 1 database (D1) might be carried over directly to this application, although the size of the database will be so much smaller that spatial indexing probably will not be required for adequate performance.

Open issues for definition or implementation

1. Are there any other quantities that should be included in the catalog entry for each point source? Note that typical sources, i.e., sources not much brighter than the detection limit, will have poorly-determined spectra, and so in general detailed spectral information probably does not belong in the catalog.
2. Should multiple versions of the catalog be defined? In the analysis process, a large catalog of all detections of sources (say, on a daily, weekly, or longer timescale) could be compiled, then that catalog cross-correlated with itself to determine which sources had (probably) been seen more than once. This process would also be useful for searching for transient sources.

D5. Pulsar ephemerides

Date: 26 April 2002 (draft v1)
Contributors: S. Digel (SU-HEPL)

Function

This is the radio pulsar timing information to be maintained during the LAT mission for assigning pulsar phases to gamma rays. It is not directly accessed by the user, but instead by the Pulsar Analysis tool (A3). (If the pulsar ephemerides are implemented in a true database system, then a front-end interface tool, equivalent to the Data Extractor for the gamma-ray data, will be what communicates directly with the database.) receives queries from, and passes data to, the Catalog Access tool (U5).

Inputs
See inputs for the Catalog Access tool (U5).

Databases required

NA

Outputs

The provisional contents of the database, defined based on the pulsar timing files used for EGRET are given below. Note that for generality and consistency with format provided by pulsar observers, times in this file should be specified in MJD rather than Mission

Elapsed Time. The second table below contains the additional information required for binary pulsars, but this two tables could be combined.

Parameters for any pulsar

	Contents	Units
1	Pulsar name	dimensionless
2	Right Ascension (J2000)	deg
3	Declination (J2000)	deg
4	Start of interval of validity for timing info (MJD)	days
5	End of interval of validity (MJD)	days
6	Infinite-frequency geocentric UTC arrival time of a pulse (MJD)	days
7	Pulsar rotation frequency	Hz
8	First derivative of pulsar frequency	Hz^2
9	Second derivative of pulsar frequency	Hz^3
10	Root-mean-square radio timing residual (periods)	dimensionless
11	Source of timing information	dimensionless
12	Flag for binary pulsars	dimensionless

Orbital parameters for binary pulsars

	Contents	Units
1	Pulsar name	dimensionless
2	Orbital period	s
3	Projected semi-major axis	s (light travel time)
4	Orbital eccentricity	dimensionless
5	Barycentric time (TDB scale) of periastron (MJD)	days
6	Longitude of periastron	deg
7	First derivative of longitude of periastron	deg per Julian year
8	Time-dilation and gravitational redshift parameter	s
9	First derivative of orbital period	dimensionless
10	Source of orbital parameters	dimensionless

Originator codes (items 11 and 10 in the two tables) need to be defined someplace, or perhaps the relevant information included directly as character strings.

Performance requirements

Performance requirements will be considered in detail in the full requirements document. This database is not expected to be particularly large or to have particularly demanding access requirements.

Other modules required

None.

Host environment

Database server system?

Existing counterparts

Nothing directly applicable. The corresponding files for the EGRET analysis system are plain ASCII, and probably could be just a static file for the LAT as well.

Open issues for definition or implementation

1. As written, the tables assume that each pulsar could be represented by multiple entries, one for each interval for which a given set of timing parameters is valid. In principle, a different organziation, indexed only by the pulsar names, could be defined. In terms of implementation, the database is likely to be small enough that the organization will not matter.

U1. Event data extractor

Date: 25 April 2002 (draft v6)
Contributors: S. Digel (SU-HEPL)

Function

This is the front end to the Level 1 databases – the gamma-ray summary and the event summary databases. It can be used as a standalone tool to produce FITS files for export outside of the analysis system, or it can be run by standard high-level analysis tools. The latter will generally run on the data requestor's computer while the Data Extractor will run on the SSC or LAT team-based server computers hosting the databases.

The Data Extractor utility constructs queries to the Level 1 database (D1) and writes FITS files containing the retrieved events.

Inputs (for gamma rays)
A complete set of selection criteria describing the desired data set, including the following:

- Time range
- Region of sky (specified as center and radius or coordinate ranges)
- Energy range
- Zenith angle cuts (energy dependent, and possibly also on inclination and azimuth or even plane of conversion[1])

[1] Cuts on zenith angle will be required to eliminate albedo gamma rays produced in the upper atmosphere. Crude cuts on zenith angle will be applied onboard (to limit the impact of albedo gamma rays on the average data rate). However, more detailed cuts will have to be applied in ground processing. Owing to

- Inclination angle range
- Azimuth range
- Gamma-ray type (assigned in Level 1 processing based on sets of background rejection/PSF enhancement cuts that have corresponding IRFs tabulated in CALDB)
- Solar system object flag (for indicating whether a suitable radius around solar system objects – the moon and sun in particular - should be excluded from the selection)

The coordinate system for the search should be selectable, among celestial, Galactic, instrument-centered, earth-centered (and possibly sun and moon-centered).

Note that the selection criteria specified here are identical to the criteria supplied to the Exposure Calculator to calculate the corresponding exposure.

Databases required

Level 1 photon summary D1

Outputs

A table containing the selected events. (Draft FITS headers are available in the report of the DPWG.) The output must include as part of the header the complete specification of selection criteria.

Performance requirements

The requirements for the Data Extractor are distinct from the requirements for the Level 1 database itself. With the query processing being handled by the data server, the principal processing work done by the Data Extractor should amount to reformatting the data that have been extracted from the database. Requiring only that the overhead for this work be much less (<10%?) of the time required by the database system for a typical query is probably sufficient.

Other modules required

GUI front end (if we define such a thing as a general utility for providing user interfaces to server-run tools)

Host environment

Database server system

Existing counterparts

Nothing directly applicable.

---

the strong dependence of the PSF size on energy, to a lesser extent on inclination and azimuth, and other parameters like plane of conversion), more conservative cuts must be applied at lower energies (and larger inclinations) for the same rejection efficiency. These same cuts must also be incorporated into the exposure calculation. These cuts may be difficult to specify as selection criteria for extracting events from the databases, a more practical approach may be to assign an 'albedo-ness' value to each event during Level 1 processing. The albedo value (basically yes or no) could be specified as an input to the Data Extractor. If this implementation is adopted, care must be taken to ensure that the Exposure Calculator has the same algorithm as the Level 1 processing pipeline for the albedo cuts.

Open issues for definition or implementation

1. Should all of the coordinate systems listed above be available, and which should be the primary system? (Database performance is likely to be significantly greater for the primary system if data are sorted by the corresponding spatial index before they are ingested.)

2. As currently conceived, the event summary database contains summary information (output of reconstruction and classification) for all telemetered events and the gamma-ray summary database contains only those events that are classified as gamma rays by at least one of the standard sets of classification cuts. Conceptually, it would be cleaner to keep all events together in one database, but at least two good reasons indicate that a separate gamma-ray database is desirable for enhanced performance: gamma rays will represent only ~10% of the total events, and they will be most usefully accessed by region of the sky (as opposed to arrival direction in instrument coordinates for cosmic rays).

3. Pulsar phase (involving specification of a pulsar with known timing information) would also be a useful selection criterion, but implementation is TBD. A more practical approach may be to make a standalone tool for subsetting selections and scaling exposures.

4. Selection criteria related to, e.g., geomagnetic cutoff might also be useful if it turns out that residual cosmic rays are a problem. What geomagnetic parameters will be the most useful in this regard?

5. The number of events returned by a search could easily be in the millions. Is this a problem in terms of staging data for transfer and transferring it over the network?

U2. Pointing/livetime history extractor

Date: 1 May 2002 (draft v1)
Contributors: S. Digel (SU-HEPL)

Function

This is the front end to the Pointing, livetime, and mode history database D2. Most commonly, this will be run by the Exposure Calculator U3 to extract the relevant portion of the pointing/livetime history for a particular exposure calculation. It could also be used as a standalone tool to produce FITS files for export outside of the analysis system, but the information is likely to be useful only for exposure calculations.

Inputs

- Time range

Databases required

Pointing, livetime, and mode history D2

Outputs

A table containing the entries in the Pointing/livetime history database for the selected time interval. (Draft FITS headers are available in the report of the DPWG.) The output must include as part of the header the complete specification of selection criteria.

Performance requirements

With the query processing (minimal as it is) being handled by the database server (for D2), the principal processing work done by the Pointing/livetime extractor should amount to reformatting the data that have been extracted from the database. Requiring only that the overhead for this work be much less (<10%?) of the time required by the Exposure calculator U3) for a typical query is probably sufficient.

Other modules required

GUI front end (if we define such a thing as a general utility for providing user interfaces to server-run tools)

Host environment

Database server system

Existing counterparts

Nothing directly applicable.

Open issues for definition or implementation

1. Should the operating mode of the LAT be an input? The idea is to incorporate flexibility for different operating modes (e.g., in the event of a tower failure) that correspond to completely different sets of instrument response functions.

2. The high-level analysis in general, and the exposure calculation in particular, may be based on a tessellation scheme for subdividing the sky. (The tessellation also provides a spatial index for the Level 1 database D1.) Exposure calculations in particular will be much faster for having precalculated values of the angles between each cell and every other cell in the tessellation. Should the pointing/livetime history extractor include the cell values for the relevant directions (z-axis, y-axis, zenith) in the portions of the timeline that it extracts? The alternative would be to have these directions translated into cell index values by the Exposure calculating tool U3.

3. Pre-extracted pointing/livetime histories for some standard time intervals (e.g., monthly or yearly), or whatever intervals appear to be commonly used for analyses could be made available.

U3. Exposure calculator

Date: 7 May 2002, v3 (22 Apr 2002 v1; 1 May 2002 in v2 clarified the function of the tool, its inputs, and a possible implementation; 7 May 2002 in v3 further clarified inputs & outputs)
Contributors: D. Band (GSFC-UMBC), S. Digel (SU-HEPL)

Function

This module calculates the exposure map for a given sky region over a given time range. The exposure can be provided at different levels of detail, i.e., as a function of different detector parameters. Exposures are necessary inputs to most high-level analysis tools and are also required for Observation simulator O2.

Inputs

- Pointing/livetime history extracted by U2 from Pointing/livetime database D2 or generated by pointing/livetime simulator O1
- Desired detail of exposure map; user may want the total exposure or the exposure as function of detector parameters such as inclination angle, so a set of parameters to be 'marginalized' can be specified (see Open Issues below)
- Specification of time range, energy range, inclination angle range, zenith angle cuts (and their energy dependence), photon type or quality, possibly even cuts on geomagnetic parameters and cuts on the directions of the sun and moon, etc. These specifications must be *exactly* the same as the selection criteria given to the Event data extractor (U1) and might even be extracted for the user from the output of U1.

Databases required

D6 - CALDB for the effective area component of the IRFs and for possible application to the zenith angle cuts (but see Open Issues below).

Outputs

An exposure array or matrix (i.e., a map with multiple dimensions) at the desired level of detail. The specification of parameter ranges on input and the versions of the instrument response functions used must be recorded with the output, e.g., as part of the header.

The exposure itself could be thought of as a large table of values (cm^2 s) as a function of spatial cell in the region of interest, and a grid of inclination, azimuth, energy, event class, and any detector parameters that are included in the definitions of the instrument response functions (perhaps like plane of conversion). For some of the dimensions, parameterization of the variation of the exposure may be possible. In this case, the exposure array would contain coefficients of the parameterization, and the table would be a lot smaller.

Performance Requirements

TBD, but at a minimum time required for execution should be comparable to the Data Extractor's retreival time for the corresponding gamma-ray data.

Other Modules Required

Pointing/livetime history extractor U2

Host Environment

Run on central server or client computer.

Existing Counterparts

None. The calculation of exposures for EGRET (by the program INTMAP) would require extensive revision to accommodate the exposure calculations for the LAT, which are more complicated in some respects (e.g., number of parameters in the instrument response functions) and simpler in others (e.g., only one basic operating mode for the LAT vs. dozens of combinations of EGRET's trigger telescopes).

Open Issues for Definition or Implementation

1. A user may want a time history of the exposure. Should this be implemented as an option within the exposure calculator, or should the calculator be run multiple times for different (adjacent) ranges of time?

2. Pointing/livetime information also can be used to derive the temporal window function (i.e., visibility profile) for a specified direction and time range. Should the extraction of window functions be an option in the Exposure calculator tool?

3. Spatial access to the gamma-ray data will likely be facilitated using a tessellation of the sky. Each tile in the tessellation is assigned an index. Each tile will be small enough (less than degree sized) that the exposure does not vary significantly across a tile. The analysis tools, e.g., for point-source detection may be implemented to use the same indexing scheme for the exposure, in which case the standard exposure matrix would not be a map but rather a tabulation (of multidimensional quantities).

4. The exposure 'map' may have so many dimensions to be unwieldy to tabulate for any particular analysis. The calculation of exposure can be factored somewhat into tabulation of total livetime for each spatial cell as a function of inclination, azimuth, zenith angle, and instrument mode (different modes corresponding to different sets of instrument response functions) which would be the output of this tool. (If it is determined that we will need to include geomagnetic parameters in the cuts for the analysis, then the tabulation must also include them. The tabulation would be even more unwieldy.) Then any other tool desiring exposures could calculate them from this tabulation plus a specification of the response functions (energy and angle dependence of the effective area) along with the prescription for the zenith angle cutoff. This approach adds flexibility, too, because for the same gamma-ray data set, different response functions might want to be considered, e.g., for gamma rays that convert in the front vs. the back of the TKR. The same tabulation of livetime distribution can be applied to each.

5. Exposure calculations for time intervals short relative to the ~30–60 s interval between entries in the Pointing/livetime database (D2) will require special treatment. The orientation and position of the LAT will not vary enough over such short time

intervals to make a significant difference in the exposure calculation, but the accumulated livetime can be determined only by examining the livetime counter values recorded with each event in that time interval. To a good approximation, these accumulated livetimes can be used to scale the overall exposure for the time interval between entries in the Pointing/livetime database. An exposure cannot be associated with a particular gamma ray; the shortest time intervals for which exposure can be calculated correspond to the shortest intervals between two events that are telemetered from the LAT. A special tool for calculating exposures on short time intervals, or perhaps for constructing profiles of livetime accumulation during short intervals, might need to be defined.

U4. Map generator

Date: 23 April 2002 (draft v1)
Contributors: S. Digel (SU-HEPL)

Function

This module constructs binned maps of gamma rays ('counts'), exposure, and intensity from the output of the Event data extractor (U1) and the Exposure Calculator (U3). Calculations of exposure maps will likely involve interpolation of exposures calculated for the standard tessellation grid.

Inputs

- Coordinate gridding of the maps and coordinate projection.
- Parameters to be 'marginalized' (integrated over)
- Gridding of the remaining parameters (e.g., energy or inclination angle)
- Data set produced by the Event data extractor U1
- Exposure matrix produced by the Exposure calculator U3

Databases required

None

Outputs

Counts and/or exposure maps (perhaps multidimensional). The intensity map, i.e., the ratio of counts and exposure, can be calculated as well. All output must include (in a header) the specifications of selection criteria used (for the input gamma-ray and exposure datasets) as well as of the parameters marginalized as part of the map generation.

Performance Requirements

TBD. The map generation is not likely to be computationally intensive.

Other Modules Required

None.

Host Environment:

Run on central server or client computer.

Existing Counterparts

None. The EGRET analogs, MAPGEN (counts maps) and INTMAP (exposure and intensity maps), are based on single viewing periods,

Open Issues for Definition or Implementation

1. What coordinate systems should be supported? Systems other than in the photon dataset?

2. What coordinate projections should be supported?

3. Will any analysis tools require as input any of the maps generated here as input? Or will the maps primarily be for visualizations?

U5. Catalog access

Date: 24 April 2002 (draft v1)
Contributors: S. Digel (SU-HEPL)

Function

This module is a catch-all front end for the catalogs of astronomical sources to be made available as part of the LAT analysis environment. The kinds of catalogs will include the LAT point source catalog, the EGRET point source catalog, flat spectrum radio sources, radio pulsars, SNRs, OB associations, blazars, etc. The idea would be to make available the catalogs that would be of interest for gamma-ray studies. It is likely that the catalogs would be stored on a database server, although in principle (especially for any large or dynamic catalogs) access requests for remote servers could be managed. This module should have both an interactive and scriptable (or API?) interface. The search criteria for the catalogs will at a minimum include region of the sky, but could also be specific to each particular catalog.

Inputs

- Region of the sky
- Catalogs of interest
- Subselection parameters for the individual catalogs; this may include specifications of what information from each catalog is to be returned.

Databases required

Collection of catalogs that are available

Outputs

Tables of sources extracted from the catalog(s).

Performance Requirements

No obvious requirements; queries are not likely to take long to process, i.e., responses should be almost at interactive speeds for regions of the sky of 10's of square degrees.

Other Modules Required

None.

Host Environment:

Run on central server or client computer. The catalogs themselves will be stored on the server if not at a remote site.

Existing Counterparts

Should be investigated carefully. NED?

Open Issues for Definition or Implementation

1. The output should in principle be readable by the model-definition module. This is probably not difficult, just a matter of using consistent naming schemes for the columns so the model-definition module can recognize coordinates and fluxes. (The model-definition module is currently defined under O2, the observation simulator, but is also relevant for the likelihood analysis tool A1.)

U7. Source model definition tool

Date: 7 May 2002 (draft v2; clarified outputs)
Contributors: D. Band (GSFC-UMBC), S. Digel (SU-HEPL)

Function

This module interactively constructs a source model for simulations and model fitting. It can be the input to the Observation simulator O2 or to the Likelihood analysis tool A1. The model per se is just a collection of parameter values; it is not an image or a multidimensional data set. In particular it does not contain the interstellar emission model for the region of interest, but instead just values for any adjustable parameters of the model and a specification of the version of the model that is being referenced. It does contain a specification of its boundaries on the sky. The Likelihood analysis tool recognizes the boundaries as well as the parameters of the model and their values, and it can use the Source model definition tool to write an updated model when the components or parameter values are changed.

Inputs

- Region of the sky, energy range, and time range of validity of the model. Note that the model needs to be defined for a larger region of the sky than will be used for analysis, owing to the breadth of the PSF.

- Parameter values for the Interstellar emission model (A7) [These parameters may include an isotropic emission spectrum that accounts for the unresolved component of the extragalactic intensity.]
- Names, coordinates, fluxes, and spectra of point sources
- Same for extended sources (not included in the interstellar emission model), including extent and shape parameters (see Open Issues below)
- Optional for output of intensity map: coordinate gridding and energy range

Databases required

A7 - Interstellar emission model of the Milky Way
D3 - LAT point source catalog
D4 - Other high-level databases (i.e., other catalogs of point sources)
D6 - CALDB (if model intensity maps are to be constructed; see Open Issues below)

Outputs

- Region of the sky, energy range, and time range of validity of the model
- A file containing the definition of the model:
  1. A list of the point sources. For each source the following are provided:
  A. Source coordinates
  B. Spectral model assumed (e.g., power law)
  C. Parameters of this spectral model (e.g., normalization, spectral index)
  D. Optionally, a name
  2. Same for any extended sources; see Open Issues below
  3. Parameters for the underlying diffuse emission model
- Optionally, an intensity map corresponding to the model for the specified coordinate gridding and energy range can be generated.

Performance Requirements

None worth specifying; this tool is not likely to be computationally intensive.

Other Modules Required

None

Host Environment

Central server or client computer

Existing Counterparts

None directly applicable. The Sherpa modelling tool (and S-Lang underlying it) for Chandra data analysis should be investigated. The FluxService in GLEAM (the LAT instrument simulation) already has a method for specifying parameters of sources that should be investigated.

Open Issues for Definition or Implementation

1. Specifying time dependences of source fluxes may be tricky to implement in a flexible, convenient way, especially if detailed time profiles are desirable. If time dependence is

included, and quite likely it shouldn't be, then in addition to specifying a region of interest on the sky, the user must also specify a range of time.
2. Including pulsars in the flux model will likely require careful thought to correctly assign the arrival times of the gamma rays. In principle, all of the information needed to 'undo' the barycenter correction is available in the livetime/pointing history, although interpolation between entries will be required.
3. Is there a useful way that the input to this tool could be implemented graphically, e.g., for placement of sources or specification of their fluxes and spectra?
4. Should the specification of parameter values include confidence ranges? There's no need for this in terms of the Observation simulator, but as output of the Likelihood analysis, confidence ranges might be worth keeping track of.
5. Do we need to define a library of template spectral shapes along with their adjustable parameters? What if someone wants to define a new spectral shape? [This is not likely to be a great concern owing to the limited numbers of gamma rays in a typical spectral analysis, at least for point sources.]
6. What about extended sources? How do we define them? Should we have a template set of shapes – circular, square, ellipsoidal, etc.? How about an arbitrary distribution of intensity?

U8. Image and plot display

Date: 22 April 2002 (draft v1)
Contributors: S. Digel (SU-HEPL)

Function

This is a general-purpose plotting and image display package. It should provide platform-independent operation (among the operating systems supported), and be capable of generating PostScript output. It should accept input directly from other tools and have provision for returning graphical inputs, e.g., cursor positions. At a minimum, x-y plotting with linear and combinations of log scales should be supported, along with 2-dimensional contour maps and image displays with color bars. A selection of color tables must be provided, as well as means for specifying and saving user-defined color tables. Image scaling must also be adjustable. Overlaying of images and contours and images and plotted points should also be supported. For images, axis ranges must be specifiable and must be properly indicated in the displays. The package might also have an interactive version, in which case FITS images and tables and ASCII tables must be accepted for data input and the interface must be scriptable and optionally log-able.

Desirable features: 3-dimensional plots, surface plots, histogramming

Inputs (for both interactive and non-interactive versions)

- Data values
- Plots: axis ranges, axis scaling, plotting symbol
- Images: axis ranges, contour levels or color table and image scaling

Databases required

None

Outputs

A display window or a PostScript file

Performance requirements

The display generation must be quick enough for useful interactive use of the analysis tools that generate displays, but this is not likely to be a problem.

Other modules required

None

Host environment

Client computer.

Existing counterparts

Perhaps ChIPS (Chandra Image Processing System, http://cxc.harvard.edu), Paul Kunz's plotting package (descended from Hippodraw), and likely others

Open issues for definition or implementation

None.

A1. Likelihood analysis

Date: 7 May 2002 (draft v2; added to Open Issues section)
Contributors: S. Digel (SU-HEPL), D. Band (UMBC-GSFC)

Function

This is the general-purpose analysis tool for source detection and characterization. It uses the likelihood function to fit models to observed gamma rays and the likelihood ratio test to define confidence ranges for parameters of the models. Specific applications include point-source detection, flux estimation, spectrum characterization, and source location determination. Analyses are for fixed time intervals and the models are static in time; flux histories of sources may be obtained by likelihood analyses of data for successive time intervals with the same source model.

Some applications of the Likelihood analysis tool are necessarily interactive, such as the search for point sources in confused regions of the sky, and others can be run automatically, like an initial search for point sources or generation of source location maps for specified sources.

Inputs

- Gamma-ray data extracted by the Data Extractor U1 and the corresponding exposure matrix calculated by the Exposure calculator U3. These tools are the front ends to the Event summary database D1 and the Pointing/livetime history database D2 for all of the analysis tools.
- Any subselections desired on the gamma-ray and exposure data, e.g., restriction on inclination angle range or harsher cut on zenith angle.
- A model defined by the source model definition module U7. The source modelling module defines the parameters of the model, the region of interest on the sky and the energy range of validity.
- Initial guesses for parameter values for the model. (Note that the source model will include values for all of the parameters, but within the Likelihood analysis tool these can be adjusted.)
- Constraints on parameter values (especially specification of those to be held constant). For TS maps, the ranges and increments of the variable parameters.

Databases required

NA

Outputs

These include the model with updated parameter values, and confidence limits, along with a specification of the data (at least file names) sufficient to be read in by the Likelihood analysis tool for subsequent analyses.

For TS maps, a FITS file can be written. Via the image and plot display tool U8, the map (or 1-dim cuts) can also be displayed. Coordinate limits for subsequent 'fine' maps can also be selected graphically.

Via the source modelling module U7, the model being used for the analysis can be updated, e.g., to add or remove point sources, within the Likelihood analysis tool and additional likelihood analyses performed.

The Likelihood analysis tool should have a provision for command logging to facilitate re-running or scripting analyses

Performance requirements

Performance requirements will be considered in detail in the full requirements document. The most pressing requirement on performance will probably relate to the time available in the Level 2 pipeline for transient source detection.

Other modules required

Source model definition module U7

Host environment

Client computer

Existing counterparts

The EGRET likelihood analysis tool LIKE is not likely to carry over to LAT data analysis. The user interface is awkward and the interface to the instrument response functions is completely different. Also, the likelihood analysis for EGRET was designed for pointed observations, for which the instrument response functions (in particular the PSF) that apply to all gamma rays in a given region of the sky are the same. This will not be the case for LAT observations obtained in scanning mode.

Some time ago, Pat Nolan prototyped a sensible interactive interface to the LAT likelihood analysis for point sources and this should be revisited for contents and/or implementation. See http://www-glast.slac.stanford.edu/software/Workshops/January01Workshop/talks/glike.pdf

Open issues for definition or implementation

1. Should the gamma-ray data be binned before the likelihood function is evaluated? If the data are to be binned, what binning is optimal, or what is the tradeoff in sensitivity for various levels of binning? Binning might be performed on any of the parameters of the gamma rays, e.g., region of the sky of arrival, range of inclination angle (region of the sky in instrument coordinates), energy, or photon class. Unbinned analysis is the most sensitive in principle, using all of the information from each gamma-ray event, but implementation is difficult (owing to multidimensional integrations that must be performed accurately), and the loss of sensitivity if carefully chosen bins are used is likely to be small.
2. What are the implications of the Protassov et al. paper (http://arxiv.org/abs/astro-ph/0201547) about the validity of the likelihood ratio test for determining, e.g., whether or not a point source is present at a given location? At the very least, distributions of likelihood ratio values in the null hypothesis should be investigated thoroughly using simulations.
3. Should there be a facility within the likelihood analysis tool for automated analyses of the same region for multiple time ranges? This would require the exposure calculator U3 to generate multiple versions of the exposure (see Open Issues in U3), and at least a good educated guess about the proper time interval for the analyses. This kind of study will undoubtedly be done frequently. The question is whether the capability explicitly should be part of A1 and U3 or whether the scriptability of both can adequately accommodate the need.
4. How will analyses of moving sources, in particular solar system bodies, proceed? The moon is a fairly bright EGRET source. The sun is not a steady source, but is impulsively very bright. During those times, it probably could be analyzed as any point source, but more generally the question needs to be addressed of how to accommodate sources that are not fixed on the sky (but have calculable trajectories).

A7. Interstellar emission model

Date: 26 April 2002 (draft v1)
Contributors: S. Digel (SU-HEPL)

Function

This is the model of the interstellar gamma-ray emission of the Milky Way used both for analysis of gamma-ray data and generation of simulated data. For both applications, the distribution of interstellar intensity on the sky and in energy is one component of a more detailed model of the gamma-ray intensity that includes point sources and small extended sources. The interface to the interstellar emission model will be the source modelling module [which must have an option to write out a model].

The model of interstellar emission will be defined with TBD adjustable parameters, most likely related to the spatial and energy distributions of cosmic rays. In principle, the optimum values for these parameters will be determined from global analysis of LAT data and for subsequent studies these parameters could be assigned fixed values. However, owing to the need (at least initially) for adjustability of the parameters, the Interstellar emission model cannot be planned as a static database as far as the analysis environment is concerned.

Keep in mind that the requirements specified here apply only to the module that provides the interstellar intensities to the source modelling module. The actual development of the interstellar emission model (via models of the interstellar matter and radiation and cosmic rays) is not the subject of these requirements.

Inputs

[Subject to review] The inputs are the region of the sky and energy range, along with the values of the TBD adjustable parameters. Depending on the implementation of the source modelling module, the coordinate system, gridding, and energy gridding of the model might also be inputs.

Databases required

NA

Outputs

The provisional outputs, defined in the report of the Data Products Working Group as LS-010, are shown in the table below. The DPWG approach assumes that the coordinate grid on the sky will be the same as used for generating the exposure and indexing the gamma rays for spatial access. For completeness, the output may also include the celestial coordinates of the spatial grid and the specific energies for the model. The output should also include specification of the version of the underlying model of the interstellar emission and the values of the adjustable parameters.

The DWPG assumed that in addition to the intensity, the Interstellar emission model would also provide a specification of the uncertainty in the model. This has not been explored in terms of the overall analysis system. Whether the uncertainty represents a statistical uncertainty or a systematic uncertainty is not yet specified, nor is how these uncertainties would be quantitified. Neither has the possible incorporation of this information in model fitting has not been investigated.

Keep in mind that the output of this module will not have been convolved with any instrument response functions. Also, the intensities are differential. (For EGRET analysis, integral intensities were calculated for standard energy ranges.) This allows for more straightforward interpolation in energy and will also facilitate spectral analysis within the general Likelihood analysis tool A1. (If fluxes are modelled on integral energy ranges, as they were for EGRET, then spectra must be derived after the fact by 'forward folding' models to find agreement with the integral fluxes found from the likelihood analysis.)

	Column Name	Units
1	pixel number	dimensionless
2	intensity spectrum	photon cm-2 s-1 sr-1 GeV-1
3	intensity uncertainty (1-sigma)	photon cm-2 s-1 sr-1 GeV-1

Performance requirements

Performance requirements will be considered in detail in the full requirements document. No particularly demanding requirements are apparent.

Other modules required

None

Host environment

Database server system

Existing counterparts

Nothing directly applicable. For EGRET analyses, the models were entirely precomputed

Open issues for definition or implementation

1. In terms of the Likelihood analysis tool A1, the model may need to be recomputed many times during a given analysis with different values of the adjustable parameters. If the model depends non-linearly on the parameters, there is not a lot that we can do to speed up the process. However, if the dependence on any of the parameters is linear, and if performance is a concern, then separate model components for each of the linear terms should be returned as output.
2. Other open issues, related to the assumptions made by the DPWG in defining the output of the Interstellar emission module, are included in the Output section above

O1. Livetime/pointing simulator

Date: 24 April 2002 (draft v1)
Contributors: D. Band (UMBC), S. Digel (SU-HEPL)

Function

This module generates the equivalent of the output of U2, the pointing, livetime, and mode history extractor, for a simulated observation. It understands the observing strategies available and constraints on the observations, such as limitations on slewing rates, requirements for Sun-normal orientation of the solar panels, and no data taking during passages through the SAA. It can be used to predict the spacecraft's pointing for tools required for supporting the mission timeline, or planning GI observations.

Inputs

- Orbit parameters – initial time, inclination, altitude, longitude of ascending node (likely to be standard)
- Observation type (survey, pointed, pointed scan)
- Observing strategy, e.g., for avoiding occultation by the earth (likely to have a set standard options available)

Databases required

None LAT-specific, although the position of the sun and possibly also the moon must be calculated from the appropriate ephemeris.

Outputs

Same as a pointing/live time history extracted from D2 by U2. The parameters of the simulation must be recorded in the header.

Performance Requirements

The spacecraft attitude and the position of the sun, earth and moon should be accurate to 1° (TBR) for a 1 year integration. These simulations are not expected to be computationally intensive, and although the simulator will be available as a tool, sets of precomputed pointing/livetime histories for various observing strategies will likely be sufficient for most users interested in generating gamma rays for simulated observations.

Other Modules Required

None.

Host Environment

Run on central server. [SD: Run on client computer]

Existing Counterparts

None. A good orbit/attitude simulator, such as written by Eric Stoneking (GSFC), could be used as the heart of this tool.

Open Issues for Definition or Implementation

1. Is it worth interfacing this tool with the image and plot display utility, U8?

O2. High-level observation simulator

Date: 24 April 2002 (draft v1)
Contributors: D. Band (GSFC-UMBC), S. Digel (SU-HEPL)

Function

This module generates a list of gamma rays with their simulated as-observed parameters for a simulated observation. The output will be in the same for as the output of the Event summary data extractor (U1) and therefore suitable for analysis with any of the standard tools. This will be useful for establishing the correctness of the analysis algorithms as well as the scientific performance of the LAT and the feasibility of any proposed study.

Inputs

- Flux model (see below)
- Region of the sky of interest (potentially a subset of the coverage of the flux model)
- Exposure map, from either an actual or simulated exposure history. The exposure may be at various levels of detail (i.e., a function of different detector parameter sets). [SD: Actually, what is needed more fundamentally is the pointing/livetime/observing mode history, because the instrument response functions depend on observing mode.]

Databases required

D6 (CALDB), the instrument response functions

Outputs

Photon list with observed parameters. The photon list must be trimmed according to the parameter specifications in the exposure map. The output must also include a specification of the input model, e.g., the source fluxes, positions, spectra, or at least the name of a file containing a specification of the model.

Performance Requirements

No particular requirements. Generation of simulated gamma rays at this high level is not expected to be computationally intensive.

Other Modules Required

Source model definition tool U7 and Interstellar emission model A7

Host Environment

Run on central server. [SD: Run on client computer]

Existing Counterparts

None.

Open Issues for Definition or Implementation

1. How should/could time dependence of source fluxes, or even periodicity of sources, be incorporated into the simulation? Time dependence of fluxes will ultimately be very important to simulate; what is the best way to specify the parameters?