DC2 Data Products Overview

S. W. Digel
Stanford Linear Accelerator Center
What’s a Data Product?

• 1’s and 0’s
• For this talk it will be the various kinds of simulated data available for analysis, or to support analyses, in DC2
What the science tools care about

- Event summary (FT1, or D-something)
  - High-level information – time, direction (absolute and in instrument coords), energy, event class for the “gamma rays”
  - Actual def used by science tools is http://www-glast.stanford.edu/cgi-bin/viewcvs/fitsGen/data/ft1.tpl?rev=1.10&view=markup

- Pointing/livetime/attitude/mode history (FT2)
  - Exactly as named - for calculating exposures
  - Also has information related to the location of the spacecraft, like geomagnetic coordinates, geographic coordinates
  - Definition is still somewhat idealized – for this data product we don’t need to keep track of each ANCILLARY and ATTITUDE packet, ~30 s updates are sufficient
  - Science tools definition is http://www-glast.stanford.edu/cgi-bin/viewcvs/fitsGen/data/ft2.tpl?rev=1.6&view=markup
What the science tools care about (2)

- Instrument response functions
  - As you heard from Jim Chiang, a collection of tables that the science tools read
  - They represent the angular resolution (PSF), effective area, and energy dispersion as functions of energy and angle (and front vs. back) for each event class
  - These are used for model fitting (likelihood analysis) and to build response matrices for spectral analysis in Xspec

- Pulsar ephemeris file
  - Pulsars are extremely stable ‘clocks’, except when they aren’t; they also are not bright enough in gamma rays (usually) for blind periodicity searches
  - As you will hear tomorrow, the ephemeris file collects contemporary timing information for pulsars that having timing information available from other wavelengths (typically radio)
    - Coordinates, period p, p,’p’’, range of validity, phase reference
    - Can also include timing parameters for binary pulsars (although for DC2 it does not)
What the science tools care about (3)

- **Live time cubes**
  - Will be mentioned by Gino Tosti tomorrow
  - Exposure calculation can be factored – although different event classes have different Aeffs, they all have the same pointing and livetime history.
  - The livetime (actually its distribution with inclination angle) can be tabulated for a grid of directions on the sky. This pre-computation greatly speeds up exposure calculation for diffuse sources (where you need to know how the exposure varies on the sky)
    - (No header template in CVS yet)
    - This has been implemented by Toby Burnett as table of HEALPix pixels (~1 deg²), with each entry containing an array of accumulated livetime for a grid of sqrt(1-cos(inclination angle))

- **Interstellar emission model**
  - As you will hear from Igor Moskalenko tomorrow, the model of interstellar diffuse emission is an important component of likelihood analyses
  - It is defined as an all sky map (in flat projection) giving differential gamma-ray intensity at 17 energies: 10, 20, 40, .... 655,360 MeV
  - It is in your science tools distribution in `externalData/galdiffuse/GP_gamma.fits`, and for good measure available from the GSSC’s server (see Tom’s talk)
What the science tools care about (4)

- Point source catalog (as input to source model definition)
  - The detailed specification has been discussed for some time; see https://confluence.slac.stanford.edu/display/ST/LAT+Source+Catalog+Contents
  - It has a few bells and whistles, including light curves
  - A very preliminary, limited version has been released today for DC2 analyses; Jean Ballet and David Landriu are producing it
  - Late data delivery and science tools-related software problems mean that a more complete preliminary catalog will not be ready until next week
What the science tools care about (some)

- GBM data products – More FITS files!
  - CSPEC
    - Counts accumulated by each detector over 8.192 s binned into 128 energy channels, ±4000 s around time of burst into one PHAII FITS file per detector per burst
    - ‘Daily’ version also provided
  - CTIME
    - Counts accumulated by each detector over 0.256 s binned into 8 energy channels, ±4000 s around time of burst One ‘PHAII’ FITS file per detector per burst; daily version, too
  - TTE – Time Tagged Event
    - Event list (arrival time and energy channel) for the counts in a burst. One file per detector
  - ASCII notice (summary info), RSP and BAK files - attend Valerie Connaughton’s talk tomorrow!
Formats of science tools data products

- All are FITS* format (http://fits.gsfc.nasa.gov), either tables or image arrays (cubes), or both
  - See the Web site maintained by Masa Hirayama for more information on the LAT data products: http://glast.gsfc.nasa.gov/ssc/dev/fits_def/
  - Tom Stephens will briefly cover some utilities for examining, modifying, visualizing contents of FITS files tomorrow morning

More fundamental LAT data products

- The FT1 files contain a (small) subset of the information in the Merit ntuple ROOT files from the Gleam runs that generated the DC2 event data
  - In particular variables used for event classification are in the Merit ntuple files only
  - (For DC2, an event in FT1 and its corresponding parent in the Merit files may be matched up via the Run ID and Event ID; in real life we’ll probably have a single, albeit very long, identifier)
  - See Tony Johnson’s talk for access information
  - ROOT->IsEasy to work with ([http://root.cern.ch](http://root.cern.ch))
- DC2 Background runs – Merit tuple ROOT files
- DC2 AllGamma - Same
Fundamental products (2)

- Digi and recon ROOT files are also available if needed
  - For look up by Run ID/Event ID
  - See Tony’s talk