
Data Challenge 2: Aims and Status

Julie McEnery

What is DC2?

- DC2 is the second in a sequence of 3 mock data challenges to prepare for LAT data after launch.
- ~ 1 month of simulated LAT data + burst data from GBM.
- Exercises the simulation/analysis chain from low level detector simulations to top level science analysis and data servers.
- Drives the development of the next round of IRFs/instrument performance metrics.
- Kickoff in January 2006, where the simulated data are released.

What is new?

- ~ 1 month of LAT observations will produce a significantly more detailed view of the GeV gamma-ray sky than that provided by EGRET.
 - Sky model needs to be (and will be) much more detailed than the model used for DC1.
 - Greater range of source classes.
 - More detailed models of source behaviour (variable AGN, periodic pulsars...).
 - Astrophysically motivated luminosity distributions and source locations.
 - Updated diffuse Galactic emission model.
 - Possible hardware failures/glitches (i.e. you cannot assume a perfect detector).
 - More detailed orbit/attitude profile and include effect of SAA.
 - With a richer sky model, and more mature science tools we anticipate that there will be many detailed and innovative analyses of the DC2 data – there will be lots to discover.
 - Some of this is explicitly listed as a DC2 goal – eg producing a catalog.
 - Analyses could include variability analysis, periodicity analyses, studies of source localisation and studies of diffuse sources and more.

Current Status - 1

- Toward the next iteration of instrument response.
 - Update of background model mostly complete (cosmic-rays, albedo etc).
 - New orbit/attitude model has been adopted.
 - Improvements to the event reconstruction code are ongoing.
 - Plan to produce moderately (10-20M) sized datasets of AllGamma and background simulations by the end of June. Both to exercise the MC generation system and to provide a starting point for people wishing to get involved in background rejection, event classification etc.
 - Main (i.e. large) AllGamma and background datasets will be generated later in July.

Current Status - 2

- Sky Model

- Defined an initial sky model, mostly to check the technical details.

- Population of variable AGN (Chiang, McEnery, Tosti).
 - Population of periodic pulsars (.).
 - Updated model of Galactic diffuse emission.

These have been partially tested in the second science tools checkout.

- GRB with both LAT and GBM simulated data.
 - Currently verifying that these sources can be correctly simulated in a full detector simulation (Gleam).

- Next Steps

- Add diffuse sources (SNR etc).
 - Increase the level of detail for pulsar/AGN.
 - Add more source classes.
 - Study the details of how to interleave/sample residual background events (after all filters and cuts) into one month of simulated gamma-ray sky.

Current Status - 3

- Science Tools
 - Continue to improve and be refined.
 - Tested and exercised in a series of “checkouts”.

Upcoming Events

- DC2 prep workshop at GSFC on June 27,28 and 29.
 - f2f meeting for people working on DC2 preparation. Essentially a software meeting.
- Extended/Open science tools checkout starting at the beginning of September.
 - Not technically part of DC2.
 - Opportunity for people to be come acquainted with the science tools, help test the sky-model and provide feedback on the data servers and documentation (more later)
- DC2: kickoff meeting in January followed by a closeout meeting 2-3 months later.
 - Large(ish) meeting to encourage maximum collaboration participation in DC2.
 - Unveil final sky simulation data.
 - Describe instrument performance
 - Describe science tools and how to use them.
 - (at closeout) discuss results and lessons learned.

Extended Science Tools Checkout

- The science tools have been tested for usability and performance via a series of “checkouts”
 - Small group of people recruited/assigned to try out various aspects of the science tools and provide feedback to the developers.
- The next checkout will be open to all who wish to participate.
- The checkout is not strictly part of DC2. It is not necessary to participate in the checkout to participate fully in DC2.
 - Useful to test and get feedback on science tools, data servers, documentation, installers and sky/source models (all critical for the success of DC2).
 - Opportunity to get involved/familiar with using the science tools, will provide a dataset for developing further analysis methods.

Extended Science Tools Checkout

- Generate at least a month of data using the observation simulator (gammas only, no need for updated response functions).
- Use a sky model to the highest level of detail that we know.
 - Except that the source locations might not be realistic (this will be a lot of work to implement well, so it is best to avoid duplicating it).
- The answers (i.e. MC truth) will be provided at the start so that users can verify that the source properties are simulated correctly and the analysis working correctly.
- The data will be served from the GSSC data server (and possibly the SLAC server).
- May kickoff this checkout with a half day tutorial on how to use the science tools sometime during the week of the August collaboration meeting.

Catalog and diffuse group contributions

- Two models of Galactic diffuse emission will be needed. One widely distributed and used as the background model. The other known in detail only to a few people and used in the simulation. What about analysis? There can be only one model in the simulation, but many different analyses. Will this group develop analysis for diffuse sources. What are the goals for ESTCO and DC2?
- Helpful to have a basic catalog analysis complete before the DC2 kickoff. The results could be provided at the kickoff meeting to provide a baseline/guidance for people wishing to find sources to look at in greater detail or to compare with their own catalog analysis. More detailed analyses can follow after the kickoff. May also be interesting to use the ESTCO dataset to try out the catalog analysis.
- DC2 provides an opportunity to try out the prototype catalog production tools on a relatively detailed simulated dataset.