GammaLib - Action #1004

Make gammalib compatible with P7REP LAT data (add IRFs and diffuse models)

11/28/2013 10:46 AM - Schulz Anneli

Status:	Closed		**Start date:**	11/28/2013
Priority:	Normal		**Due date:**	
Assigned To:	Schulz Anneli		**% Done:**	100%
Category:			**Estimated time:**	0.00 hour
Target version:	1.0.0			

Description

Hi,
since Fermi officially switched to the reprocessed data, it would be good to implement the corresponding responses and diffuse models in gammalib, too.
General information about P&REP: http://fermi.gsfc.nasa.gov/ssc/data/analysis/documentation/Pass7REP_usage.html
the diffuse models are here: http://fermi.gsfc.nasa.gov/ssc/data/access/lat/BackgroundModels.html
and the IRFs are P7V15, coming with Science Tools 09-32-05.

Is there more to implement than to add

the response functions to $GAMMALIB/inst/lat/caldb...
and the diffuse to $GAMMALIB/inst/lat/test/data/ with a new folder?

Related issues:

Related to GammaLib - Feature # 603: Perform a scientific validation of the P...	**Closed**	
Related to GammaLib - Action # 1060: Investigate whether a more precise curva...	**Closed**	01/07/2014
Related to GammaLib - Bug # 1496: Fit errors in log parabola fit are too large	**Rejected**	07/01/2015

History

#1 - 11/28/2013 10:53 AM - Knödlseder Jürgen

I need to check if the response format has been changed.

#2 - 11/28/2013 12:51 PM - Schulz Anneli

I did a short check yesterday and it looks fine. I copied the files to my caldb folder and it seemed to work.

#3 - 11/29/2013 12:19 PM - Knödlseder Jürgen

Great!

The best would be to make a short check with some data, e.g. fitting the Crab or Vela with gtlike and then with ctlike to see whether the results are (basically) the same.

#4 - 01/07/2014 06:43 PM - Schulz Anneli

- % Done changed from 0 to 80

Hi,
I did a cross check with an analysis of the Crab region (following the example in the p7v6 folder), the results look the following:

science tools 09-32-05

Crab:
'Index': [-2.1354435334969293, 0.04594388482845476],
'Integral': [25.587001794287488, 1.5107021246857777], scale 1e-7
'TS value': 1432.7716170280473,

'gll_iem_v05':

'Index': [0.08682762462159843, 0.07162344769818912],
'Prefactor': [0.7287081943702254, 0.1440608280265844],
'TS value': 636.2220754477894},

'iso_source_v05':
'Normalization': [2.875818635678156, 0.6855552052125737],
'TS value': 53.92443256788101},
'logLike': 10808.505288286657

and ctools:

```
Name ......................: Crab
Integral .................: 1.99229e-06 +/- 1.19129e-07 [1e-12,0.0001] ph/cm2/s (free,scale=1e-07,gradient)
Index ....................: -2.05072 +/- 0.0491567 [-5,-0.1]  (free,scale=1,gradient)

Name ......................: gll_iem_v05
Prefactor ................: 0.774972 +/- 0.156895 [1e-05,100] ph/cm2/s/MeV (free,scale=1,gradient)
Index ....................: 0.0695083 +/- 0.0745843 [-1,1]  (free,scale=1,gradient)

Name ......................: iso_source_v05
Normalization ............: 3.48618 +/- 0.694252 [0,10]  (free,scale=1,gradient)
```

This looks reasonable and is similar to the results with p7v6.

I created a branch and put the files in the corresponding folders, as written in the description above. The log files for the two fits are in
$GAMMALIB/inst/lat/test/data/p7repv15
(including the file compare_gtlike_ctlike.txt, for more information on the fit.)

#5 - 01/07/2014 08:48 PM - Knödlseder Jürgen

Did you analyse the same dataset? I'm a little worried by the significantly different Crab flux. I would have expected much closer values ...

Have you pushed the new branch already in the repo (I could not find anything)?

#6 - 01/08/2014 09:29 AM - Schulz Anneli

Hi,
I forgot to push yesterday, sorry. But it should be there now. There was a message that the file size for the galactic diffuse
(inst/lat/test/data/p7repv15/gll_iem_v05.fit) is too large, so I'm not sure if it worked out.

About the fit results: I was surprised about the big difference in integral flux. But the values for both fits were quite similar to the ones for p7v6:
in the README there:

gtlike results:
Crab:
Integral: 24.2369 +/- 1.49572
Index: -2.12467 +/- 0.0476363
LowerLimit: 100
UpperLimit: 500000
TS value: 1300.37
Flux: 2.41076e-06 +/- 1.50551e-07 photons/cm^2/s

Extragal_diffuse:
Normalization: 2.28565 +/- 0.462091
Flux: 0.000472695 +/- 9.55313e-05 photons/cm^2/s

Galactic_diffuse:
Value: 0.861504 +/- 0.0652024
Flux: 0.00041505 +/- 3.14108e-05 photons/cm^2/s

WARNING: Fit may be bad in range [100, 1258.93] (MeV)
WARNING: Fit may be bad in range [1995.26, 5011.87] (MeV)

Total number of observed counts: 2211
Total number of model events: 2211

-log(Likelihood): 10794.50987

GammaLib Results without efficiency correction:
=== GOptimizerLM ===
 Optimized function value ..: 10814.4
 Absolute precision ....: 1e-06
 Optimization status ....: converged
 Number of parameters ......: 13
 Number of free parameters .: 4
 Number of iterations ......: 9
 Lambda ....: 1e-12
=== GModels ===
 Number of models ....: 3
 Number of parameters ......: 13
=== GModelDiffuseSource ===
 Name ....: Extragal_diffuse
 Instruments ....: all
 Model type ....: "ConstantValue" * "FileFunction" * "Constant"
 Number of parameters ......: 3
 Number of spatial par's ...: 1
  Value ....: 1 [0,10] (fixed,scale=1,gradient)
 Number of spectral par's ..: 1
  Normalization ....: 2.89709 +/- 0.441824 [0,1000] (free,scale=1,gradient)
 Number of temporal par's ..: 1
  Constant ....: 1 (relative value) (fixed,scale=1,gradient)
=== GModelDiffuseSource ===
 Name ....: Galactic_diffuse
 Instruments ....: all
 Model type ....: "MapCubeFunction" * "ConstantValue" * "Constant"
 Number of parameters ......: 3
 Number of spatial par's ...: 1
  Normalization ....: 1 [0.001,1000] (fixed,scale=1,gradient)
 Number of spectral par's ..: 1
  Value ....: 0.862843 +/- 0.0637907 [0,1000] (free,scale=1,gradient)
 Number of temporal par's ..: 1
  Constant ....: 1 (relative value) (fixed,scale=1,gradient)
=== GModelPointSource ===
 Name ....: Crab
 Instruments ....: all
 Model type ....: "SkyDirFunction" * "PowerLaw2" * "Constant"
 Number of parameters ......: 7
 Number of spatial par's ...: 2
  RA ....: 83.6331 [-360,360] deg (fixed,scale=1)

```
   DEC .....................: 22.0145 [-90,90] deg (fixed,scale=1)
 Number of spectral par's ..: 4
  Integral ................: 1.95868e-06 +/- 1.23463e-07 [1e-14,0.0001] ph/cm2/s (free,scale=1e-07,gradient)
  Index ...................: -2.0426 +/- 0.0509271 [-5,5]  (free,scale=1,gradient)
  LowerLimit ...............: 100 [10,1e+06] MeV (fixed,scale=1)
  UpperLimit ...............: 500000 [10,1e+06] MeV (fixed,scale=1)
 Number of temporal par's ..: 1
  Constant ................: 1 (relative value) (fixed,scale=1,gradient)
```

#7 - 01/08/2014 04:35 PM - Knödlseder Jürgen

Schulz Anneli wrote:

> Hi,
> I forgot to push yesterday, sorry. But it should be there now. There was a message that the file size for the galactic diffuse
> (inst/lat/test/data/p7repv15/gll_iem_v05.fit) is too large, so I'm not sure if it worked out.

It came across, but we never should add such huge files to GammaLib as otherwise the package becomes very large.

The diffuse model is in fact not really used in the Fermi/LAT analysis with GammaLib and only pre-convolved source models are supported. So the srcmap.fits files are ultimately needed. But also here one should pay attention that the files are not too big.

> About the fit results: I was surprised about the big difference in integral flux. But the values for both fits were quite similar to the ones for p7v6:
> in the README there:
>
> [...]

I have to look this up, but maybe I used a very limited number of energy bins to get a small srcmap.fits file. The purpose here is of course to get a unit test, not a scientific test.

I did some Fermi/LAT Science Tools comparison some time ago: https://cta-redmine.irap.omp.eu/projects/gammalib/wiki/Validation_of_Pass_7_IRFs. Indeed, the differences were similar.

Maybe you can add an equivalent page with your results.

#8 - 01/08/2014 04:55 PM - Knödlseder Jürgen

- Status changed from New to In Progress

- Assigned To set to Schulz Anneli

I created a new branch 1004_fermi_p7rep_no_diffuse without the diffuse map, as I still have problems with the old branch hanging around, and trying to push around a 400 MB file.

#9 - 01/08/2014 05:46 PM - Knödlseder Jürgen

I still was struggling with the big file problem and just copied all the text files except of the diffuse model into the actual devel branch and deleted the 1004_fermi_p7rep_no_diffuse branch. Now everything seems to be okay.

#10 - 01/08/2014 06:15 PM - Schulz Anneli

Hi Jürgen,
sorry for causing chaos with the too large file, it won't happen again smile.png

The data files (srcmap, exposure cube and ltcube) are the same, so the difference for the integral flux is strange.

I started a wiki page: https://cta-redmine.irap.omp.eu/projects/gammalib/wiki/Validation_of_Pass_7REP

I will do some more checks and put them on the wiki page.

#11 - 01/08/2014 09:02 PM - Knödlseder Jürgen

Schulz Anneli wrote:

> Hi Jürgen,
> sorry for causing chaos with the too large file, it won't happen again smile.png

I screwed up the synchronization with github, as github limits files to 100 MB. As Christoph uses github there was a risk that he could not synchronize with the latest code updates ...

> The data files (srcmap, exposure cube and ltcube) are the same, so the difference for the integral flux is strange.

Indeed. But the maps are used slightly differently (I use the bin centres while the Fermi/LAT Science Tools use the bin edges). And the fitting factors are not the same.

There is in fact feature #603 that asks for a scientific validation of the P6V7 response. I propose we use this to follow-up on this issue. I'll start to write down some ideas about test that should be done.

> I started a wiki page: https://cta-redmine.irap.omp.eu/projects/gammalib/wiki/Validation_of_Pass_7REP

> I will do some more checks and put them on the wiki page.

Great.

#12 - 02/12/2014 05:19 PM - Knödlseder Jürgen

- Target version set to 00-08-02

- % Done changed from 80 to 90

I discovered a bug in the energy scaling of the back PSF. The scaling of the front PSF has in fact been applied, which explains the offsets between ScienceTools and gammalib (see https://cta-redmine.irap.omp.eu/projects/gammalib/wiki/Validation_of_Pass_7REP). We probably still want to make more tests, but we made a big step forward in validating the P7REP response implementation.

#13 - 03/12/2014 10:49 PM - Knödlseder Jürgen

Anneli reported further problems on https://cta-redmine.irap.omp.eu/projects/gammalib/wiki/Validation_of_Pass_7REP. This problems occurred when dealing with more complex problems where several sources are in the field of view with different spectral laws. Specifically, the number of predicted events by ctlike was considerably in excess of the number of predicted events by the ScienceTools.

Some problem has been unveiled in the computation of the isotropic model. This seems to be some bug in the Fermi/LAT Science Tools, and should be followed up with Jim Chiang. At the end, it was not the isotropic model that led to the overestimation in Npred.

It has turned out that the PSF was not normalized correctly for sources falling outside the event cube. This has now been corrected, and results look more reasonable. Here the fitted parameters for W49B, the galactic diffuse and the isotropic diffuse emission obtained from the Science Tools and ctlike. Results look pretty similar:

Parameter	Science Tools	ctlike
Iso Normalization	0.3567 +/- 0.0362	0.4318 +/- 0.0393
Gal Prefactor	1.0626 +/- 0.0026	1.0396 +/- 0.0042
Gal Index	- 0.0051 +/- 0.0009	- 0.0066 +/- 0.0015
W49B Prefactor	7.4432 +/- 0.3704	7.9027 +/- 0.5843
W49B Index	- 2.3976 +/- 0.0164	- 2.4112 +/- 0.0250

#14 - 03/14/2014 05:45 PM - Schulz Anneli

Jürgen's bugfix solved the issue:
I checked with several sources free (in one field of view) and get the following results: (always ctlike compared to gtlike):

2FGL J1906.5+0720
parameter: norm with value 0.422176 +/- 78634.9 compared to 0.4092975931 +/- 0.03962327838
parameter: alpha with value 2.55612 +/- 8238.58 compared to 2.504185881 +/- 0.05900088563
parameter: beta with value 0.0565304 +/- 0.0400038 compared to 0.0682810382 +/- 0.0255311659
parameter: Eb with value 0.782425 +/- 57014.3 compared to 0.7740598176 +/- 0.02977020576

2FGL J1914.4+0951c
parameter: norm with value 1.71845 +/- 93958.5 compared to 1.683374541 +/- 0.8176208382
parameter: alpha with value 3.32881 +/- 18785 compared to 3.303447011 +/- 0.2521842347
parameter: beta with value 0.571837 +/- 0.277524 compared to 0.5790399427 +/- 0.2131126854
parameter: Eb with value 1.72019 +/- 28254.4 compared to 1.708008588 +/- 0.2663808932
2FGL J1916.1+1106
parameter: norm with value 63.7346 +/- 99999.9 compared to 57.28774491 +/- 56.51042741
parameter: alpha with value 3.04577 +/- 30.2178 compared to 2.937955134 +/- 0.1494343719

parameter: beta with value -0.0293269 +/- 0.102382 compared to -0.01658470348 +/- 0.03581787606
parameter: Eb with value 0.303468 +/- 156.329 compared to 0.291565867 +/- 0.1084661751

W49B
parameter: Prefactor with value 7.73985 +/- 0.601714 compared to 7.458333573 +/- 0.3711425726
parameter: Index with value -2.40573 +/- 0.0258459 compared to -2.398122557 +/- 0.01644587742
gll_iem_v05
parameter: Prefactor with value 1.03762 +/- 0.00430308 compared to 1.062561236 +/- 0.00261299233
parameter: Index with value -0.00614701 +/- 0.00156984 compared to -0.005113371265 +/- 0.0009463204093
iso_source_v05
parameter: Normalization with value 0.441461 +/- 0.0396752 compared to 0.3574049615 +/- 0.03619618379

the errors for log parabola get huge with the correlated parameters. Fixing E_b leads to:

2FGL J1906.5+0720
parameter: norm with value 0.433633 +/- 0.0185852 compared to 0.4090445333 +/- 0.03961582417
parameter: alpha with value 2.55493 +/- 0.0886896 compared to 2.503967361 +/- 0.05899313288
parameter: beta with value 0.0565298 +/- 0.0400038 compared to 0.06835095727 +/- 0.0255339505

2FGL J1914.4+0951c
parameter: norm with value 1.77126 +/- 0.234328 compared to 1.695391293 +/- 0.8232425651
parameter: alpha with value 3.3184 +/- 0.253384 compared to 3.308084063 +/- 0.2538721804
parameter: beta with value 0.571821 +/- 0.277522 compared to 0.5854030934 +/- 0.2139677668

2FGL J1916.1+1106
parameter: norm with value 71.637 +/- 24.7915 compared to 56.88787447 +/- 56.07531643
parameter: alpha with value 3.04775 +/- 0.417986 compared to 2.934170604 +/- 0.1494810659
parameter: beta with value -0.0292527 +/- 0.10239 compared to -0.01558510458 +/- 0.03587226734

W49B
parameter: Prefactor with value 7.73985 +/- 0.601714 compared to 7.443152099 +/- 0.3704410753
parameter: Index with value -2.40573 +/- 0.0258459 compared to -2.397552635 +/- 0.01644322714
gll_iem_v05
parameter: Prefactor with value 1.03762 +/- 0.00430308 compared to 1.062612981 +/- 0.002613131318
parameter: Index with value -0.00614704 +/- 0.00156984 compared to -0.005119404545 +/- 0.0009463250926
iso_source_v05
parameter: Normalization with value 0.441462 +/- 0.0396752 compared to 0.3567742776 +/- 0.03619719838

#15 - 03/14/2014 09:41 PM - Knödlseder Jürgen

Looks good.

I added some information to issue #1060 which should tackle the error problem. As the Science Tools seem to have good errors, one solution is to base the GammaLib code on the Science Tools code, which itself takes the minuit2 code. This is probably quite some work (to understand what exactly minuit2 is doing, then implement and test the code), but it's probably worth the money.

#16 - 03/18/2014 02:28 PM - Schulz Anneli

I did an example plot which shows the nicely matching results on the wiki page:
https://cta-redmine.irap.omp.eu/projects/gammalib/wiki/Validation_of_Pass_7REP

#17 - 03/18/2014 02:34 PM - Knödlseder Jürgen

Great!

I also made progress on the conceptional side for improving the error computation. I think I now know how to implement this, I'm still thinking about the most efficient way (as this step can be quite time consuming). And then I need to find some time for coding and testing, hence I won't promise you now a fixed date for this, but I try'll work on this as soon as I have some spare time.

#18 - 07/19/2014 02:08 AM - Knödlseder Jürgen

- Target version changed from 00-08-02 to 00-09-00

Skip 0.8.2 release.

#19 - 11/11/2014 11:27 PM - Knödlseder Jürgen

- Target version changed from 00-09-00 to 1.0.0

#20 - 06/30/2015 12:58 PM - Knödlseder Jürgen

I reanalysed the problem:

```
$ ctlike debug=yes
Event list, counts cube or observation definition file [srcmap_W49B_2fgl_pl_CEL_0.1dpb_AIT_500-300000MeV_P7REP_SOURCE_V15.fits]
 observation_fermi_2fgl_pl.xml
Source model [crab_1_10.xml] fit_W49B_gal.xml
Source model output file [crab_results.xml] results.xml
```

The optimizer gave the following result:

```
2015-06-30T11:11:56: === GOptimizerLM ===
2015-06-30T11:11:56:  Optimized function value ..: -403202.585
2015-06-30T11:11:56:  Absolute precision ........: 0.005
2015-06-30T11:11:56:  Acceptable value decrease .: 2
2015-06-30T11:11:56:  Optimization status .......: converged
2015-06-30T11:11:56:  Number of parameters ......: 467
2015-06-30T11:11:56:  Number of free parameters .: 4
2015-06-30T11:11:56:  Number of iterations ......: 3
2015-06-30T11:11:56:  Lambda ....................: 1e-06
2015-06-30T11:11:56:  Maximum log likelihood ....: 403202.585
2015-06-30T11:11:56:  Observed events  (Nobs) ...: 1057409.000
2015-06-30T11:11:56:  Predicted events (Npred) ..: 1060092.776 (Nobs - Npred = -2683.78)
```

and here the results for some of the sources:

Name	Prefactor	-Index
W49B	7.0503e-09 +/- 5.39854e-10	2.38619 +/- 0.0256042

gll_iem_v05	1.06361 +/- 0.00409294	0.00867046 +/- 0.0014983

#21 - 06/30/2015 02:12 PM - Schulz Anneli

It would be great to include also tests with Pass 8, since it's public now: http://fermi.gsfc.nasa.gov/ssc/data/access/

#22 - 06/30/2015 03:47 PM - Knödlseder Jürgen

Another run:

$ ctlike debug=yes
Event list, counts cube or observation definition file [events.fits] observation_fermi_2fgl_pl.xml
Source model [$CTOOLS/share/models/crab.xml] binned_fitted_W49B_2fgl_pl.xml
Source model output file [crab_results.xml] results.xml

Here the result of the optimizer:

2015-06-30T14:00:58: === GOptimizerLM ===
2015-06-30T14:00:58: Optimized function value ..: -403245.514
2015-06-30T14:00:58: Absolute precision ....: 0.005
2015-06-30T14:00:58: Acceptable value decrease .: 2
2015-06-30T14:00:58: Optimization status ....: errors are inaccurate
2015-06-30T14:00:58: Number of parameters ......: 470
2015-06-30T14:00:58: Number of free parameters .: 17
2015-06-30T14:00:58: Number of iterations ......: 3
2015-06-30T14:00:58: Lambda ....: 1e-06
2015-06-30T14:00:58: Maximum log likelihood ....: 403245.514
2015-06-30T14:00:58: Observed events (Nobs) ...: 1057409.000
2015-06-30T14:00:58: Predicted events (Npred) ..: 1059799.479 (Nobs - Npred = -2390.48)

and here the result of the source fits:

Name	Prefactor	-Index	Curvature	Pivot energy
2FGL J1906.5+0720	4.06993e-11 +/- 7.42991e-06	2.49713 +/- 9950.35	0.0680539 +/- 0.0410746	772.916 +/- 5.65051e+07
2FGL J1914.4+0951c	1.68762e-12 +/- 9.38756e-08	3.29031 +/- 19115.9	0.565359 +/- 0.28774	1703.55 +/- 2.88003e+07
2FGL J1916.1+1106	2.31198e-10 +/- 4e-07	2.96791 +/- 26.6823	0.0228827 +/- 0.106874	293.75 +/- 171240

So it looks like that the problem when fitting all parameters for a logparabola model is still there.

#23 - 06/30/2015 03:48 PM - Knödlseder Jürgen

Schulz Anneli wrote:

> It would be great to include also tests with Pass 8, since it's public now: http://fermi.gsfc.nasa.gov/ssc/data/access/

Would be great, but I have no idea what handling of Pass 8 implies in terms of software. I'd prefer to defer this after release 1.0 to get not stuck with that.

#24 - 06/30/2015 04:04 PM - Schulz Anneli

Knödlseder Jürgen wrote:

> Schulz Anneli wrote:
>
> > It would be great to include also tests with Pass 8, since it's public now: http://fermi.gsfc.nasa.gov/ssc/data/access/
>
> Would be great, but I have no idea what handling of Pass 8 implies in terms of software. I'd prefer to defer this after release 1.0 to get not stuck with that.

Sure. I'm not sure either what Pass 8 implies. I completely agree: It's more important to have a stable release 1.0 out and then check pass 8 later and it's just out for a week or so.

#25 - 07/01/2015 02:13 PM - Knödlseder Jürgen

- Status changed from In Progress to Closed

- % Done changed from 90 to 100

- Remaining (hours) set to 0.0

Not clear how Fermi can fit all parameters of a log parabola, as the pivot energy is degenerate.

Close this now.