Understanding and Optimizing LAT XML Models

Elizabeth Ferrara
Fermi Science Support Center

2013 Fermi Summer School
May 28 – June 7

XML Model Structure

```
<?xml version_info=”number”?>
<!-- Comment -->
<tag attribute=”value assigned”>
   <nested_tag attribute=”value assigned”/>
      <parameter attribute=”value”/>
   </nested_tag>
</tag>
```

NOTES FOR XML TAGS AND ATTRIBUTES

1) Start tags (<) must have end tags (/>)
2) Parameter tags usually start and end on the same line
3) Values for attributes are in double quotes (“ “)
4) NO SPACES within an attribute definition
5) If you can avoid touching the XML model, that’s best!
   (It’s EASY to make a typo that ruins your week.)

Example Point Source

```
<?xml version="1.0" ?>
<source_library title="source library">
    <source name="SwiftJ1644" type="PointSource">
    <!-- point source units are ph cm^-2 s^-1 MeV^-1 -->

        <spectrum type="PowerLaw2">
            <parameter free="1" max="10000.0" min="0.0001" name="Integral" scale="1e-07" value="1.0"/>
            <parameter free="1" max="5.0" min="0.0" name="Index" scale="-1.0" value="2.0"/>
            <parameter free="0" max="500000.0" min="20.0" name="LowerLimit" scale="1.0" value="100.0"/>
            <parameter free="0" max="500000.0" min="20.0" name="UpperLimit" scale="1.0" value="300000.0"/>
        </spectrum>

        <spatialModel type="SkyDirFunction">
            <parameter free="0" max="360.0" min="-360.0" name="RA" scale="1.0" value="251.2054"/>
            <parameter free="0" max="90.0" min="-90.0" name="DEC" scale="1.0" value="57.5808"/>
        </spatialModel>

    </source>

    *** ADD MORE SOURCES HERE ***

</source_library>
```

“source” Tag

```
<?xml version="1.0" ?>
<source_library title="source library">
    <source name="SwiftJ1644" type="PointSource">
    <!-- point source units are ph cm^-2 s^-1 MeV^-1 -->

        <spectrum type="PowerLaw2">
            <parameter free="1" max="10000.0" min="0.0001" name="Integral" scale="1e-07" value="1.0"/>
            <parameter free="1" max="5.0" min="0.0" name="Index" scale="-1.0" value="2.0"/>
            <parameter free="0" max="500000.0" min="20.0" name="LowerLimit" scale="1.0" value="100.0"/>
            <parameter free="0" max="500000.0" min="20.0" name="UpperLimit" scale="1.0" value="300000.0"/>
        </spectrum>

        <spatialModel type="SkyDirFunction">
            <parameter free="0" max="360.0" min="-360.0" name="RA" scale="1.0" value="251.2054"/>
            <parameter free="0" max="90.0" min="-90.0" name="DEC" scale="1.0" value="57.5808"/>
        </spatialModel>

    </source>

    *** ADD MORE SOURCES HERE ***

</source_library>
```

“source” Tag

```
<?xml version="1.0" ?>
<source_library title="source library">
  <source name="SwiftJ1644" type="PointSource">
```

name attribute can have any value. User-assigned.

type attribute must be one of the following:

- PointSource - For most sources
- DiffuseSource - For extended sources, galactic diffuse, isotropic diffuse

No **parameter** values associated with this tag.

```
  </source>

  *** ADD MORE SOURCES HERE ***

</source_library>
```

“spectrum” Tag

```
<?xml version="1.0" ?>
<source_library title="source library">
    <source name="SwiftJ1644" type="PointSource">
    <!-- point source units are ph cm^-2 s^-1 MeV^-1 -->

        <spectrum type="PowerLaw2">
            <parameter free="1" max="10000.0" min="0.0001" name="Integral" scale="1e-07" value="1.0"/>
            <parameter free="1" max="5.0" min="0.0" name="Index" scale="-1.0" value="2.0"/>
            <parameter free="0" max="500000.0" min="20.0" name="LowerLimit" scale="1.0" value="100.0"/>
            <parameter free="0" max="500000.0" min="20.0" name="UpperLimit" scale="1.0" value="300000.0"/>
        </spectrum>

        <spatialModel type="SkyDirFunction">
            <parameter free="0" max="360.0" min="-360.0" name="RA" scale="1.0" value="251.2054"/>
            <parameter free="0" max="90.0" min="-90.0" name="DEC" scale="1.0" value="57.5808"/>
        </spatialModel>

    </source>

    *** ADD MORE SOURCES HERE ***

</source_library>
```

“spectrum” Tag

```
<?xml version="1.0" ?>
<source_library title="source library">
    <source name="SwiftJ1644" type="PointSource">
    <!-- point source units are ph cm^-2 s^-1 MeV^-1 -->

        <spectrum type="PowerLaw2">
```

type attribute defines the spectral model to be fit:

- **PowerLaw, BrokenPowerLaw**
- **PowerLaw2, BrokenPowerLaw2 -** Use the integrated flux as a free parameter
- **SmoothBrokenPowerLaw**
- **LogParabola -** Simplest curved spectrum
- **ExpCutoff**
- **BPLExpCutoff**
- **PLSuperExpCutoff -** Most frequently used for pulsars
- **Gaussian**
- **BandFunction -** Use for GRBs
- **ConstantValue -** Use if spatial model includes spectral information (e.g. Galactic Diffuse)
- **FileFunction -** Allows user to define a custom spectral shape (e.g. Isotropic Diffuse)

“spectrum” Tag

type attribute defines the spectral model to be fit:
PowerLaw, BrokenPowerLaw, PowerLaw2, BrokenPowerLaw2, SmoothBrokenPowerLaw, LogParabola, ExpCutoff, BPLExpCutoff, PLSuperExpCutoff, Gaussian, BandFunction, ConstantValue, FileFunction

```
<spectrum type="PowerLaw2">
  <parameter free="1" max="10000.0" min="0.0001" name="Integral" scale="1e-07" value="1.0"/>
  <parameter free="1" max="5.0" min="0.0" name="Index" scale="-1.0" value="2.0"/>
  <parameter free="0" max="500000.0" min="20.0" name="LowerLimit" scale="1.0" value="100.0"/>
  <parameter free="0" max="500000.0" min="20.0" name="UpperLimit" scale="1.0" value="300000.0"/>
```

parameter attributes are used to refine how the spectral model is fit to the data

- **free = “0 or 1”** : Will the parameter be fit? 1 means yes, it will.
- **min = “number”** : Minimum of the fit range
- **max = “number”** : Maximum of the fit range
- **scale = “number”** : Factor to multiply the fitted value to find the actual value
- **value = “number”** : Starting guess for the parameter value
- **name = “Name”** : Name of the model parameter (depends on the model selected)
  Prefactor, Index(1/2), Scale, Integral, LowerLimit, UpperLimit, BreakValue, Cutoff, norm, alpha, beta, Value, Normalization, etc...

Available spectral and spatial models at:
http://fermi.gsfc.nasa.gov/ssc/data/analysis/scitools/source_models.html

“spatialModel” Tag

```
<?xml version="1.0" ?>
<source_library title="source library">
    <source name="SwiftJ1644" type="PointSource">
   <!-- point source units are ph cm^-2 s^-1 MeV^-1 -->

       <spectrum type="PowerLaw2">
          <parameter free="1" max="10000.0" min="0.0001" name="Integral" scale="1e-07" value="1.0"/>
          <parameter free="1" max="5.0" min="0.0" name="Index" scale="-1.0" value="2.0"/>
          <parameter free="0" max="500000.0" min="20.0" name="LowerLimit" scale="1.0" value="100.0"/>
          <parameter free="0" max="500000.0" min="20.0" name="UpperLimit" scale="1.0" value="300000.0"/>
       </spectrum>

       <spatialModel type="SkyDirFunction">
          <parameter free="0" max="360.0" min="-360.0" name="RA" scale="1.0" value="251.2054"/>
          <parameter free="0" max="90.0" min="-90.0" name="DEC" scale="1.0" value="57.5808"/>
       </spatialModel>

    </source>

    *** ADD MORE SOURCES HERE ***

</source_library>
```

“spatialModel” Tag

```
<?xml version="1.0" ?>
<source_library title="source library">
    <source name="SwiftJ1644" type="PointSource">
   <!-- point source units are ph cm^-2 s^-1 MeV^-1 -->

       <spectrum type="PowerLaw2">
          <parameter free="1" max="10000.0" min="0.0001" name="Integral" scale="1e-07" value="1.0"/>
          <parameter free="1" max="5.0" min="0.0" name="Index" scale="-1.0" value="2.0"/>
          <parameter free="0" max="500000.0" min="20.0" name="LowerLimit" scale="1.0" value="100.0"/>
          <parameter free="0" max="500000.0" min="20.0" name="UpperLimit" scale="1.0" value="300000.0"/>
       </spectrum>

       <spatialModel type="SkyDirFunction">
```

type attribute defines the spatial model to be fit:

- **SkyDirFunction** - Gives RA, DEC of a point source in the sky
- **SpatialMap** - FITS template file with position in the header
  - Example: Extended sources
- **MapCubeFunction** - Spatial map with energy panes giving spectral information
  - Example: Galactic Diffuse model
  - Spectral component should not be fit (ConstantValue)
- **ConstantValue** - Component with only spectral information
  - Example: Isotropic diffuse model

“spatialModel” Tag

type attribute defines the spatial model to be fit:

- **SkyDirFunction** - Gives RA, DEC of a point source in the sky
- **SpatialMap** - FITS template file with position in the header
  - Example: Extended sources
- **MapCubeFunction** - Spatial map with energy panes giving spectral information
  - Example: Galactic Diffuse model
  - Spectral component should not be fit (ConstantValue)
- **ConstantValue** - Component with only spectral information
  - Example: Isotropic diffuse model

```
<spatialModel type="SkyDirFunction">
   <parameter free="0" max="360.0" min="-360.0" name="RA" scale="1.0" value="251.2054"/>
   <parameter free="0" max="90.0" min="-90.0" name="DEC" scale="1.0" value="57.5808"/>
```

parameter attributes are the same as for the spectral tag

- **free** = “True/False”
- **min** = “number” : For RA use -360, DEC use -90
- **max** = “number” : For RA use 360, DEC use 90
- **scale** = “number”
- **value** = “number” : For many parameters of a spatial model, this is set to 1
- **name** = “Name” : RA, DEC, Value, Normalization

Parameter Values

```
<parameter free="1" max="10000.0" min="0.0001" name="Integral" scale="1e-07" value="1.0"/>
<parameter free="1" max="5.0" min="0.0" name="Index" scale="-1.0" value="2.0"/>
<parameter free="0" max="500000.0" min="20.0" name="LowerLimit" scale="1.0" value="100.0"/>
<parameter free="0" max="500000.0" min="20.0" name="UpperLimit" scale="1.0" value="300000.0"/>

<parameter free="0" max="360.0" min="-360.0" name="RA" scale="1.0" value="251.2054"/>
<parameter free="0" max="90.0" min="-90.0" name="DEC" scale="1.0" value="57.5808"/>
```

NOTES FOR PARAMETER VALUES

1) If you try to set a value outside the min-max range, you will get an error

2) If you don't set scale so that value is a reasonable size, the optimizer may not converge (not true for fixed parameters)

3) If you try to integrate over an energy range that has no data, your fit will fail

4) If you try to fit a source that lies outside your ROI, your fit will fail

Extended Source

```
<source name="W44" type="DiffuseSource">
<!-- diffuse source units are cm^-2 s^-1 MeV^-1 sr^-1 -->

  <spectrum normPar="norm" type="LogParabola">
    <parameter free="1" max="100000" min="1e-05" name="norm" scale="1e-11" value="9.5"/>
    <parameter free="1" max="5" min="0" name="alpha" scale="1" value="2.3"/>
    <parameter free="1" max="5" min="-1" name="beta" scale="1" value="0.1"/>
    <parameter free="0" max="300000" min="20" name="Eb" scale="1" value="1000"/>
  </spectrum>

  <spatialModel file="$(LATEXTDIR)/Templates/W44.fits" type="SpatialMap" map_based_integral="true">
    <parameter free="0" max="1000" min="0.001" name="Prefactor" scale="1" value="1"/>
  </spatialModel>

</source>
```

- Extended source templates are on your USB stick and at the FSSC.
- Extract the templates into a specific directory and include the path in the XML/make symbolic link/set environment variable
- Remember that the diffuse response has already been calculated for Galactic and isotropic models. But you MUST run gtdiffrsp before fitting any other “DiffuseSource”.

Galactic Diffuse Model

```
<source name="gal_2yearp7v6_v0" type="DiffuseSource">
<!-- diffuse source units are cm^-2 s^-1 MeV^-1 sr^-1 -->

    <spectrum type="ConstantValue">
        <parameter free="1" max="10.0" min="0.0" name="Value" scale="1.0" value= "1.0"/>
    </spectrum>

    <spatialModel file="$(PATH_TO DIFFUSE)/gal_2yearp7v6_v0.fits" type="MapCubeFunction">
        <parameter free="0" max="1000.0" min="0.001" name="Normalization" scale= "1.0" value="1.0"/>
    </spatialModel>

</source>
```

- Add path to Galactic diffuse model FITS file (using whatever method you prefer).
- Galactic model contains spectral shape in the FITS file (map has 30 different energy planes). So only the model normalization is left free. Spectral shape is not fitted.

Isotropic Diffuse Model

```
<source name="iso_p7v6source" type="DiffuseSource">
<!-- diffuse source units are cm^-2 s^-1 MeV^-1 sr^-1 -->

    <spectrum file="$(PATH_TO_DIFFUSE)/iso_p7v6source.txt" type="FileFunction">
        <parameter free="1" max="1000" min="1e-05" name="Normalization" scale="1" value="1"/>
    </spectrum>

    <spatialModel type="ConstantValue">
        <parameter free="0" max="10.0" min="0.0" name="Value" scale="1.0" value="1.0"/>
    </spatialModel>

</source>
```

- Add path to isotropic diffuse model FITS file. (Easiest if both models are in the same location.)
- Again, the spectral shape is defined in the isotropic model text file. So only the model normalization is left free. The spectral shape is not fitted.

Earth Limb Template

```
<source name="EarthLimb" type="DiffuseSource">

   <spectrum file="$(PATH_TO_FILES)/limb_2year_P76_source_v0_smooth.txt" type="FileFunction">
      <parameter free="1" max="1000" min="1e-05" name="Normalization" scale="1" value="1"/>
   </spectrum>

   <spatialModel file="$(PATH_TO_FILES)/limb_2year_smooth.fits" type="SpatialMap">
      <parameter free="0" max="1000.0" min="0.001" name="Normalization" scale= "1.0" value="1.0"/>
   </spatialModel>
</source>
```

- Example of customized spatial AND spectral templates.
- Remember: This template was designed specifically for 2 years of all-sky data and would need to be adjusted for other data sets.
- Not an issue for smaller (~20 deg) ROIs. Limb emission gets incorporated into other diffuse components.
- For all-sky analysis, a tighter zenith cut can make this component less important, but at a cost to exposure.

Looking at XML can get mind-numbingly boring. Another good reason not to touch it....

XML Avoidance

- **Very useful user-contributed tool: make2FGLxml.py**
  - **Download from FSSC User Contributions page:**
    - ***http://fermi.gsfc.nasa.gov/ssc/data/analysis/user/***
  - **Place somewhere your python can find it**
  - **Put 2FGL catalog file in your working directory**

```
>>> from make2FGLxml import *
This is make2FGLxml version 03.
NOTE: You must have run gtselect on the event file you use as input.
>>> mymodel = srcList('gll_psc_v07.fit','filtered_gti.fits','model_name.xml')
>>> mymodel.makeModel('gal_2yearp7v6_v0.fits','gal_2yearp7v6_v0','iso_p7v6source.txt',
'iso_p7v6source',extDir='Templates/')
Creating file and adding sources for 2FGL
Added 55 point sources and 0 extended sources
```

- **Defaults will give the following output**
  - **All 2FGL sources within your ROI plus 5 degrees**
  - **Sources outside ROI have all parameters fixed**
  - **Comments to indicate distance from ROI center**
  - **Initial guesses are set to 2FGL catalog values!**
- **Contract region of free parameters with 'radLim=X' option (X in degrees)**
- **Fix low-significance sources with 'signif=Z' option (Z in sigmas)**
- **Replace extended sources with point sources with 'psForce=True" option**

XML Avoidance

- **Very useful user-contributed tool: make2FGLxml.py**
  - **Download from FSSC User Contributions page:**
    - ***http://fermi.gsfc.nasa.gov/ssc/data/analysis/user/***
  - **Place somewhere your python can find it**
  - **Put 2FGL catalog file in your working directory**

```
>>> from make2FGLxml import *
This is make2FGLxml version 03.
NOTE: You must have run gtselect on the event file you use as input.
>>> mymodel = srcList('gll_psc_v07.fit','filtered_gti.fits','model_name.xml')
>>> mymodel.makeModel('gal_2yearp7v6_v0.fits','gal_2yearp7v6_v0','iso_p7v6source.txt',
'iso_p7v6source',extDir='Templates/')
Creating file and adding sources for 2FGL
Added 55 point sources and 0 extended sources
```

- **Defaults will give the following output**
  - **All 2FGL sources within your ROI plus 5 degrees**
  - **Sources outside ROI have all parameters fixed**
  - **Comments to indicate distance from ROI center**
  - **Initial guesses are set to 2FGL catalog values!**
- **Contract region of free parameters with 'radLim=X' option (X in degrees)**
- **Fix low-significance sources with 'signif=Z' option (Z in sigmas)**
- **Replace extended sources with point sources with 'psForce=True" option**

Model Performance

- **You now have:**
  - **XML model**
    - ***with ALL the sources (55 instead of 4)***
    - ***detected in the first 2 years of data***
    - ***that lie within your ROI + 5 deg***
    - ***plus Galactic and isotropic diffuse models***

- **Great! If you're analyzing the 2 years of data used for the catalog analysis**
  - **With less data, faint sources will less significant**
    - ***Fit is likely to be degraded for your source of interest***
  - **With more data, there may be additional sources that were not in the 2FGL catalog**
    - ***You will need to add these sources to get a proper fit***

  - **In both cases the fit may converge, but is STILL WRONG**

Evaluating the Model

- **When you finish fitting, look at fit parameters to evaluate how well your model fit the data**
  - **Too many sources will have many low-significance sources**
    - ***Too many degrees of freedom -> poorly constrained fit***
  - **Too few sources will affect the diffuse normalization**
    - ***Not an accurate representation of the data***

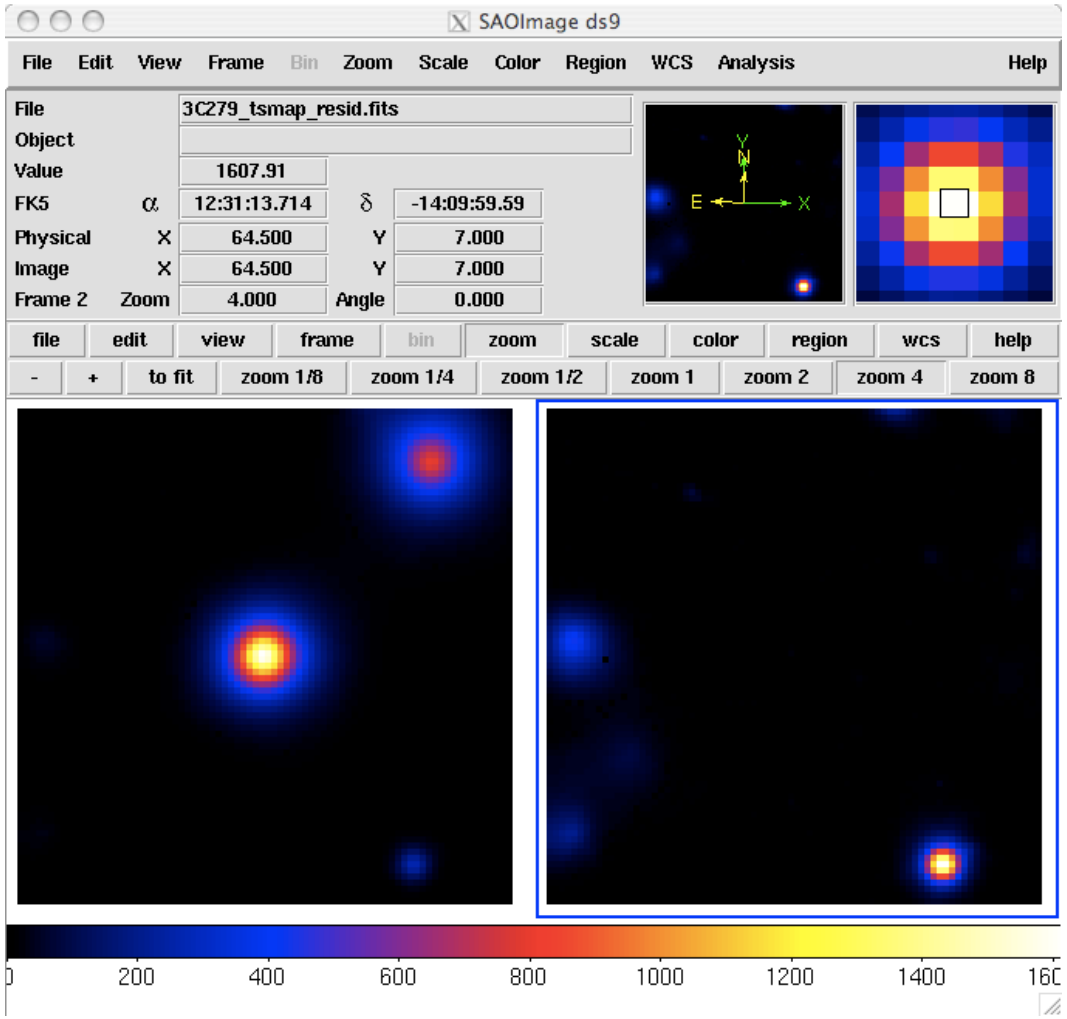
- **To correct for too few sources, remove low-significance sources and refit**
  - **Two ways: Npred and Test Statistic**
  - **If source has ~1 predicted count, it is unlikely your data can describe it well**
  - **If source is not significant (low TS) then....**

- **To correct for too few sources, add new ones in!**

Finding new sources - Unbinned

- **For Unbinned Likelihood, you can generate a TS map with all sources modeled**
  - **gttsmap creates a grid within your data**
  - **adds a putative source to your model at each position**
  - **performs likelihood fit**
  - **records the TS of a new source at that position**

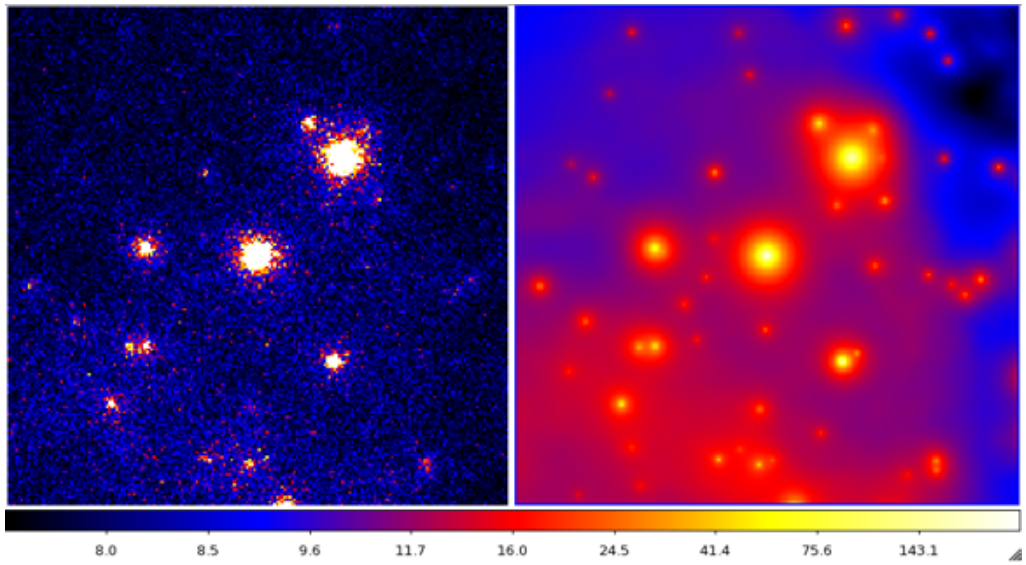
- **Very computationally expensive!**

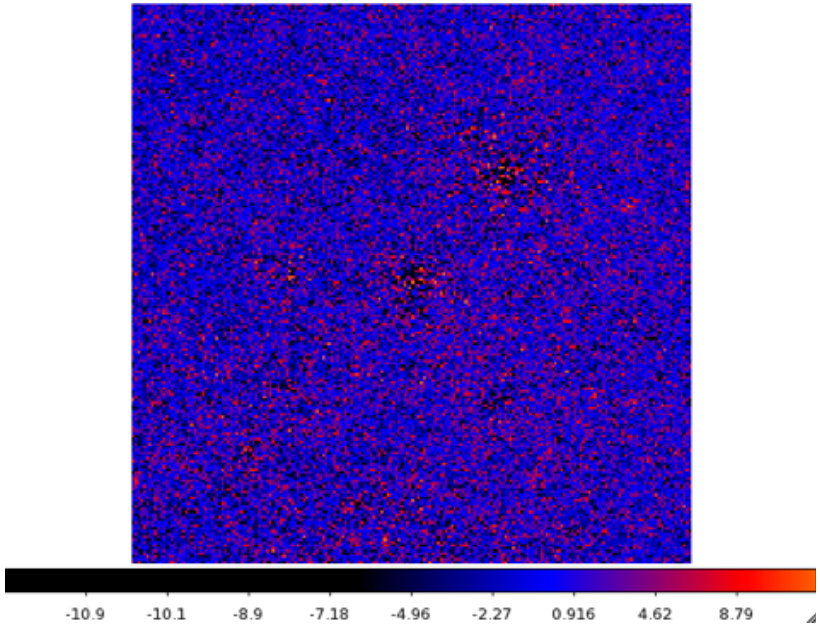


Finding new sources - Binned

- For Binned Likelihood, create a residual map
  - use gtmodel to create an image of the expected counts
  - use gtbin to create map of the observed counts (use identical binning!)
  - use farith to divide the two maps
- Quick! And good enough!

- Can also be used with Unbinned Likelihood
  - Must run gtsrcmaps before gtmodel





Iterate!

- **Add/subtract sources to/from your model then refit**

- **To add sources there are two methods**
  - **ModelEditor**
    - ***prompt> modeleditor &***
    - ***Very finicky, but should write XML properly***
  - **Open up the XML in a text editor and add sources manually**
    - ***Ick! Be careful!***

- **There are several methods of removing sources that don't require opening an XML file**
  - **ModelEditor again!**
  - **Text editor again!**
    - ***Ick again!***
  - **python tools!**
    - ***Jeremy's turn!!111!!!1one!!1!!***