A Novel Maximum Likelihood Method For VERITAS Analysis

Tom Brantseg
for the VERITAS collaboration

August 10, 2017

IOWA STATE UNIVERSITY

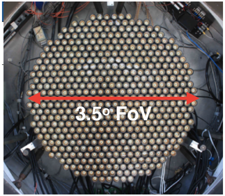
Outline

Extended gamma-ray sources

- Recent developments (T. Linden talk Tuesday, D. Hooper talk Wednesday, J. Hewitt talk Monday):
  - HAWC: Extended TeV emission from Geminga/Monogem
  - Also seen by Milagro, PAMELA, AMS02
- Important physics! (Positron excess)
- IACTs could fill in a nice part of the physical picture...
  - Good energy/spatial resolution, wide energy range
  - ...but it's a little tricky

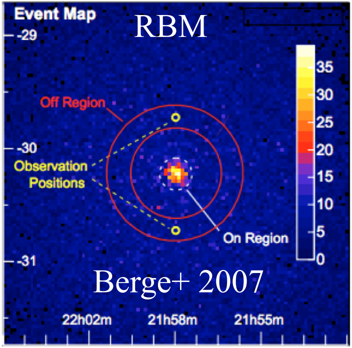
- 499 pixels/camera
- Energy range: 85 GeV to > 30 TeV
- Energy resolution: 20% @ 1 TeV
- Angular resolution (68% containment): 0.08° @ 1 TeV
- Point source sensitivity: 1% Crab in ~25h

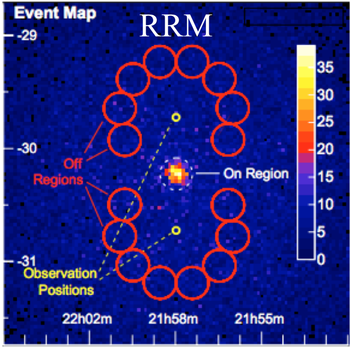
G. Hughes



IACTs and extended sources

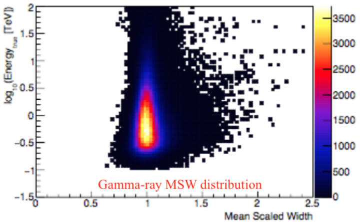
- Standard analysis methods have difficulty with extended sources
  - Large sources can "crowd out" background regions in field
  - Difficult to get enough background for statistically meaningful analysis of sources > about 0.3°
  - Hard to analyze sources like Geminga, Cygnus Cocoon, etc.

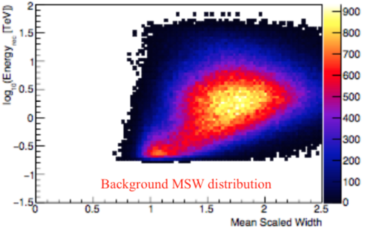




3D Maximum Likelihood

- Developed by ISU VERITAS group led by Amanda Weinstein
  - Most of the figures here are from Amanda's (recently defended) student Josh Cardenzana
- 3DMLM adds a third dimension to standard spatial ML fitting for gamma-ray/background separation
  - Mean scaled width (MSW)
  - Average Hillas width parameter for all images contributing to an event





Likelihood calculation

$$\log \mathcal{L}(\vec{s}) = N_{obs} \log(N_{exp}) - N_{exp} + \sum_i [S_{src}(\vec{r}_i|\vec{s}) W_{src}(w_i|\vec{s}) + S_{bkg}(\vec{r}_i) W_{bkg}(w_i)]$$

- Red: likelihood term measuring observed (obs) vs. expected (exp) counts
- Blue: likelihood term measuring correlation of the source and background spatial (S) and MSW (W) models with the data

Data binning

- Collect observed data with similar spatial/MSW background distributions (*field*) to improve stats
- Spatial and MSW background distributions depend on:
  - Detector configuration
  - Atmospheric conditions
  - Observational parameters
  - Telescope pointing
  - Energy
- Each field can be fit to its own set of spatial and MSW models

Spatial Models

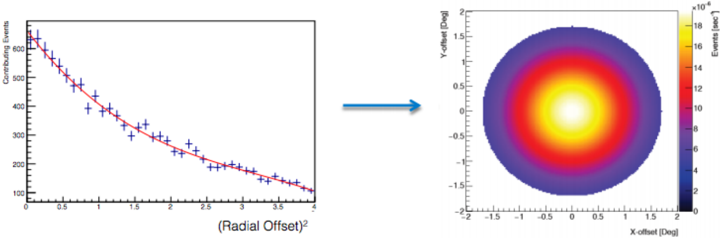
- Calculated from:
    - Effective area (A)
    - Energy dispersion (R)
    - PSF (P)
    - Source spatial morphology (B)
    - Spectral parameters (S)
- Spatial model M calculated as:

$$M(\vec{\mathbf{r}}_{i,j}|\vec{s}) = \sum_{k,m,n} B_{m,n} P_{m,n}(\vec{r}, E'_k) A_{m,n}(E'_k) \int_{E_{min}}^{E_{max}} R_{m,n}(E, E'_k) dE \times \int_{E'_{k_{lo}}}^{E'_{k_{up}}} S(E'|\vec{s}) dE'$$

- Red term can be pre-computed based on simulations and templates for each field
- Blue term must be recalculated at each fit iteration

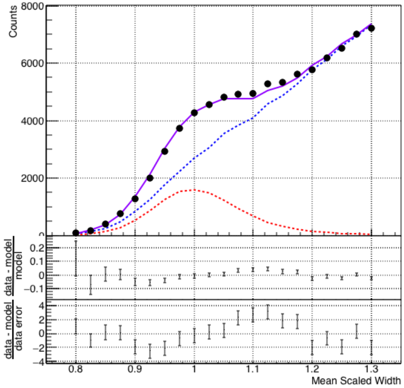
Spatial model morphology

- Source models generated from pre-existing templates
- Background models taken from data on weak sources and blank fields
- Bright stars and potential sources excluded
- Correction for zenith angle dependence of Cherenkov light
- Approximated by polynomial fit to generate spatially symmetric background model



MSW models

- Source models calculated from standard VERITAS simulations
- Background models calculated from observational blank field data



- Red: source MSW model from simulations
- Blue: background MSW model
- Purple: total MSW model

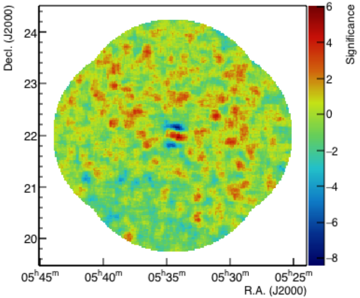
MLM vs. RBM check (Crab)

- Want to check quality of both spectral and spatial fits
- Spectral fit:
  - Log parabola model

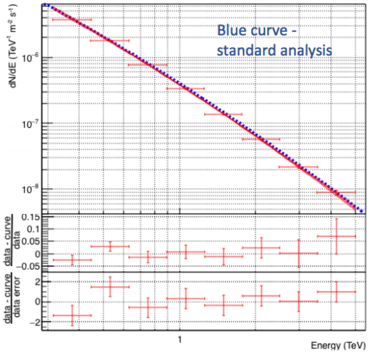
$$\frac{dN}{dE} = A\left(\frac{E}{E_0}\right)^{\alpha+\beta\log(E/E_0)}$$

- Spatial fit:
  - Measure quality of spatial models with residual map
  - Sky map - (source spatial model + background spatial model)
  - Should be basically blank if we've gotten it right

MLM vs. RBM check (Crab)



(b) Crab Source Subtracted Significance Map

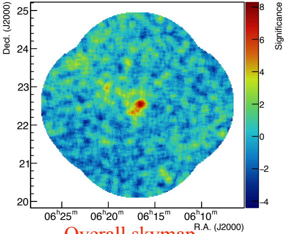


	MLM	Standard
Norm at 1 TeV	3.49 ± 0.09	4.05 ± 0.07
α	-2.49 ± 0.02	-2.44 ± 0.02
β	-0.12 ± 0.03	-0.15 ± 0.03

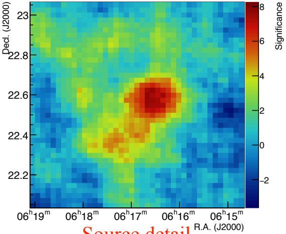
Known (small) extended source: IC443

- Field for this is larger than for the Crab case
- NOTE Not a realistic analysis, more of a software check
  - Only uses a portion of the IC443 data set
  - Higher than normal lower energy threshold (modeling issues at low energy)
- Spatial fit and spectral fit checks as before
  - Power law spectral model
  - Models computed using disk templates
  - 0.15° - brightest part of the emission
  - 0.35° - entire shell

IC443 sky maps

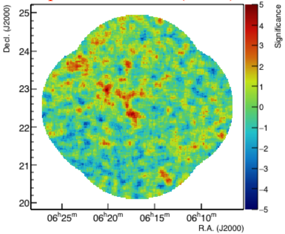


Overall skymap

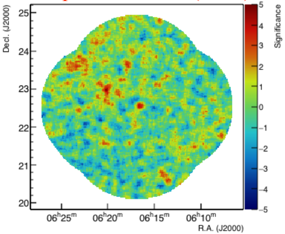


Source detail

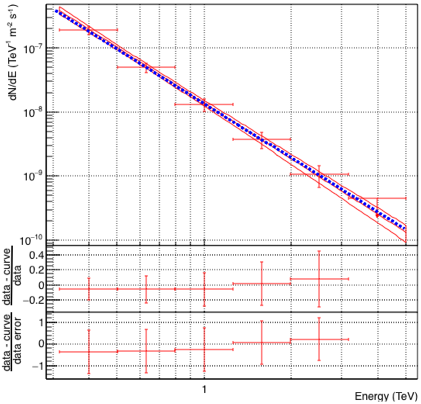
Spatial residuals (0.15°)



Spatial residuals (0.35°)



IC443 spectral fits



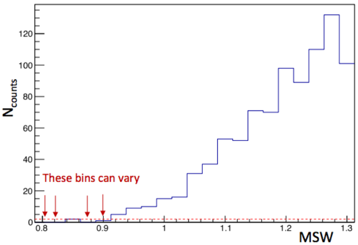
- Blue line: standard VERITAS analysis fit
- Red: 3D MLM spectral fit

Overall

- General agreement with standard analysis results on point sources (Crab, sample of blazars)
- Fitting extended source (IC443) with extended templates gives promising results!
- Background statistics in low-count areas subject of ongoing work

Blank-Field Issue

- "Phantom sources" detected fitting blank fields to extended source templates
- Issue partially tracked to low-count bins in MSW distributions
- Clear that a significant portion of the problem is due to handling background statistics
- Examining ways to modify the likelihood statistic (Barlow & Beeston 1993) to resolve the issue



Future course of work

- Resolve statistical and implementation issues with Barlow-Beeston test statistic to resolve blank field issue
- Re-check previous validation studies
- Further studies on other extended sources/3ML plugin

Possible fix

- Treat MSW distribution with modified likelihood stat from Barlow & Beeston (1993)

$$\log \mathcal{L} = \sum_i d_i \log f_i - f_i - \sum_{i,j} a_{ji} \log A_{ji} - A_{ji}$$

- Term in red is a 'penalty' term that measures contribution of individual model components
- Each component in each bin varies independently if that improves overall $\mathcal{L}$
- Currently resolving with statistical/implementation issues for this method
  - Toy MC studies of MSW distributions + blank field samples

How well can you do?

- Flux UL for a sample of 50 runs