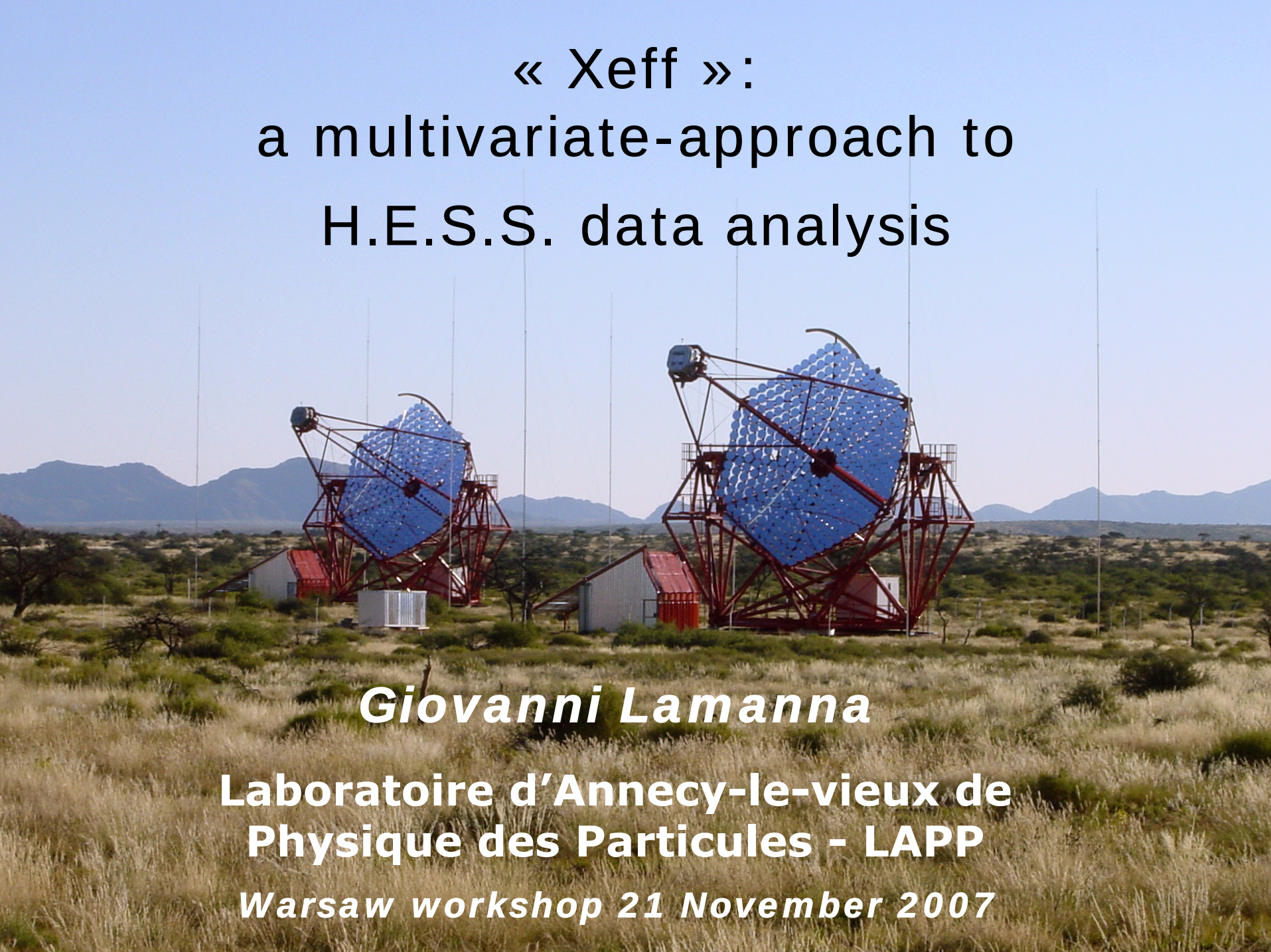
« Xeff »: a multivariate-approach to H.E.S.S. data analysis



Giovanni Lamanna

Laboratoire d'Annecy-le-vieux de Physique des Particules - LAPP

Warsaw workshop 21 November 2007

Introduction

Basic Idea: Need to go further a “simple” (Hillas) parameters cuts approach.

Status: New and successful attempts to a multivariate approach of IACT data analysis, Neural network, Random forest and any other possible classification and regression method suitable for the IACT application.

Purpose: Improving γ/**h** separation
(critical at lower energy threshold – HESS II, CTA)

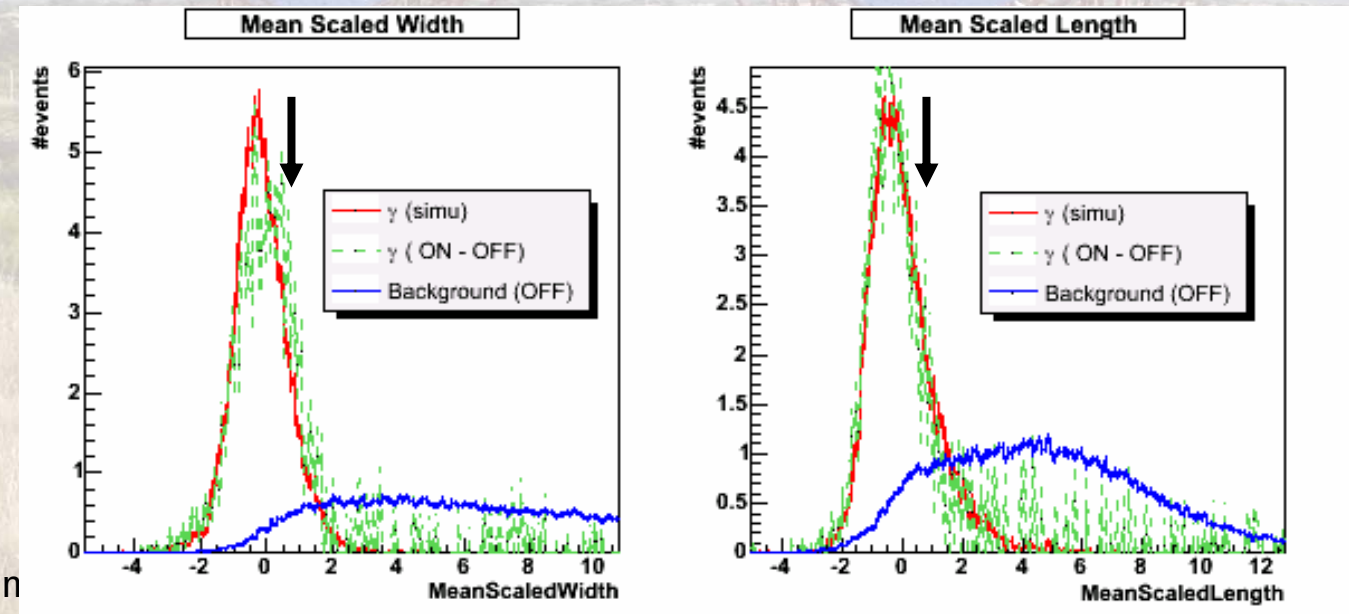
Intent: Looking for the form of a generalized predictor based on a set of uncorrelated “methods”.

The multi-dimensional space: Hillas (and some additional) parameters,
“Goodness” estimator of Model analysis,
“Rescaled-width” from 3D model analysis.

Training samples: MC-γ, ON and OFF data

Analysis methods

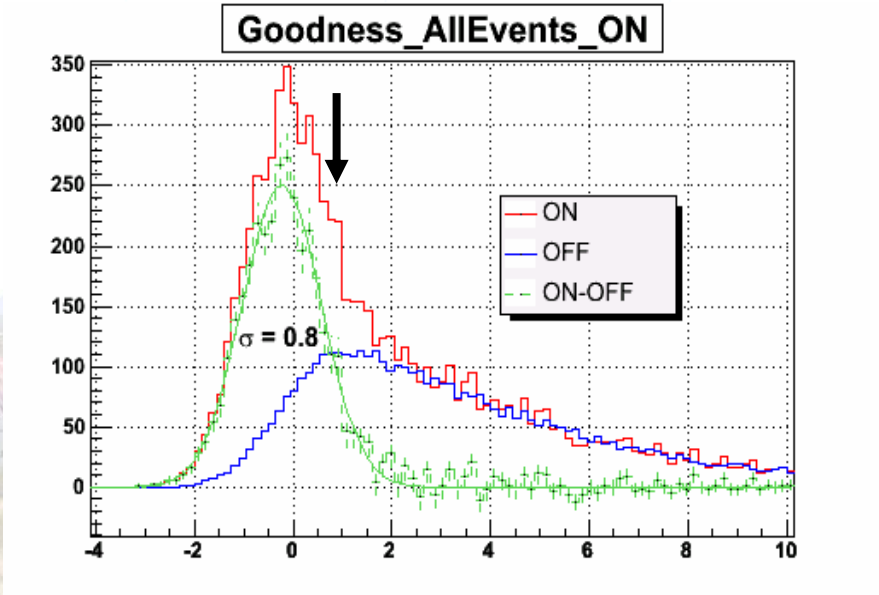
- Hillas :
  - 2D gaussian distributions of charge shared by camera pixels
  - Parameters :
    - Mean scaled width and length (variables independent on energy and direction)
    - Independent Cuts on them to minimize residual B tails.



Analysis methods

- Model :
  - Likelihood function depending on:
    - Shower energy
    - Shower direction
  - Discriminating variable Goodness-of-fit :
    - A mean scaled variable independent on energy and direction:

$$G = \frac{\langle L \rangle - L}{\sqrt{Ndof}}$$



The X_{EFF} estimator

Our original "Goal":

- Improve the morphological study of *extended sources* of limited statistics:
  By minimizing **B** population and maximizing the Quality factor *Q*
  (increasing the Significance σ).

(Implicitly we assume the method being already powerful for point-like sources.) [...]

Purpose:

- When the variables are less discriminating a combination (*Xeff estimator*) of n dimensions:
  gain in selection efficiency.
- Optimization of a unique cut on and "hadroness" (Xeff) estimator.
- *The hadroness estimator can be applied event by event for a weighting method approach*.

Xeff originally applied in the ALEPH-LEP experiment and described in
Buskulic et al. (Physics Letters B 384 449-460)

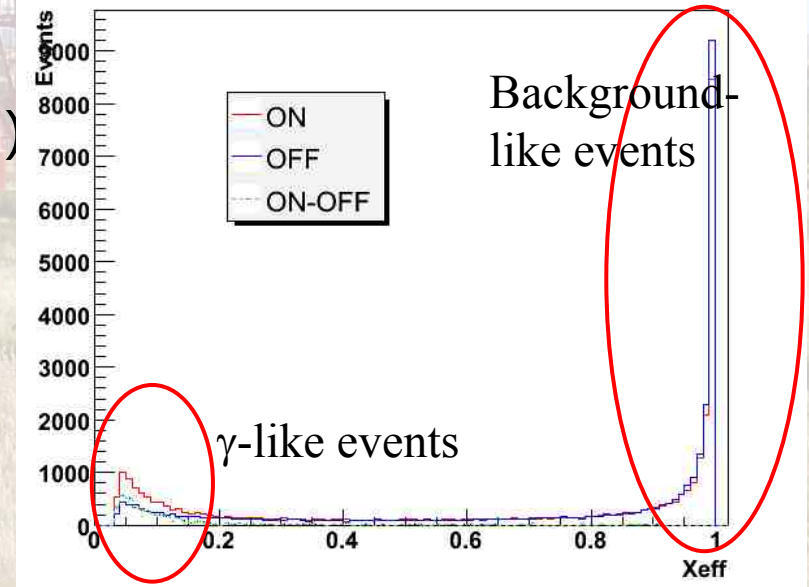
It is alternative to other iterative regression methods with final Hadroness estimator.

The X_{EFF} estimator

- Definition : $$X_{eff} = \frac{\eta.\Pi_B}{\eta.\Pi_B + (1-\eta).\Pi_S}$$

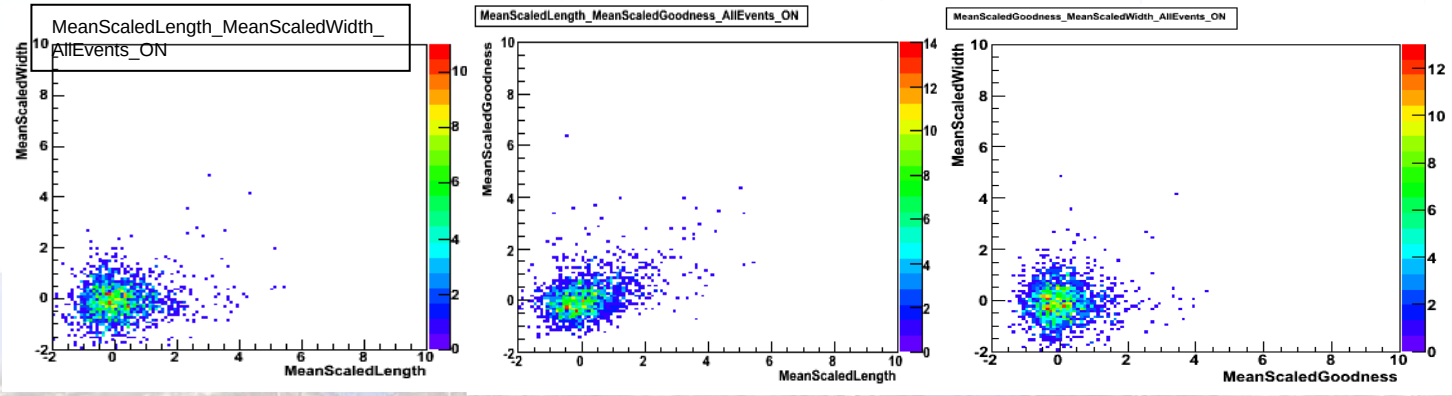
$$\Pi_{B,S} = \prod_{Variables} pdf_{B,S} \qquad \eta = \frac{n_B}{n_B + n_S}$$

- Combined variables :
  - MeanScaledGoodness (Model),
  - MeanScaledWidth (Hillas),
  - MeanScaledLength (Hillas).

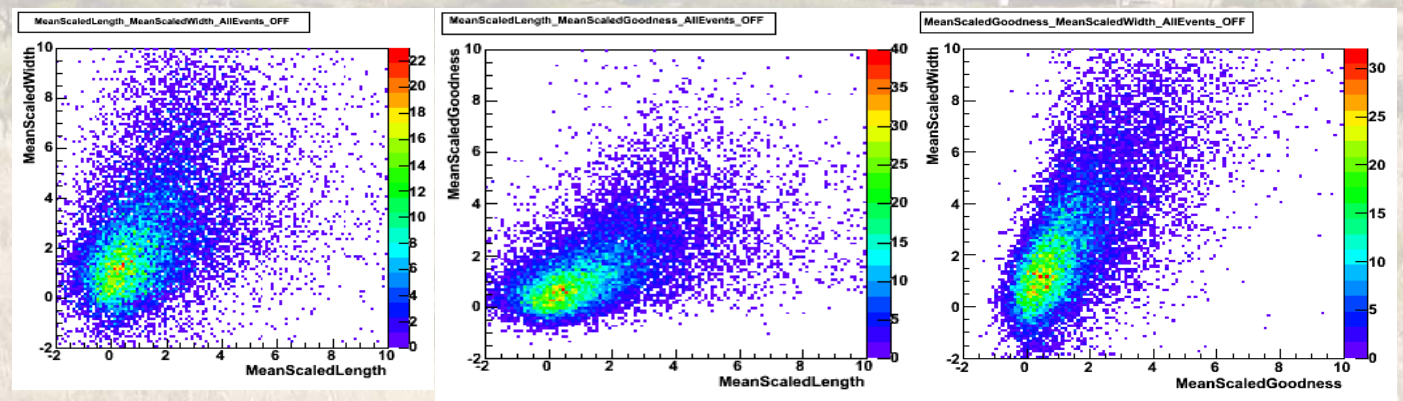


Variables' correlation

- Uncorrelated for signal:

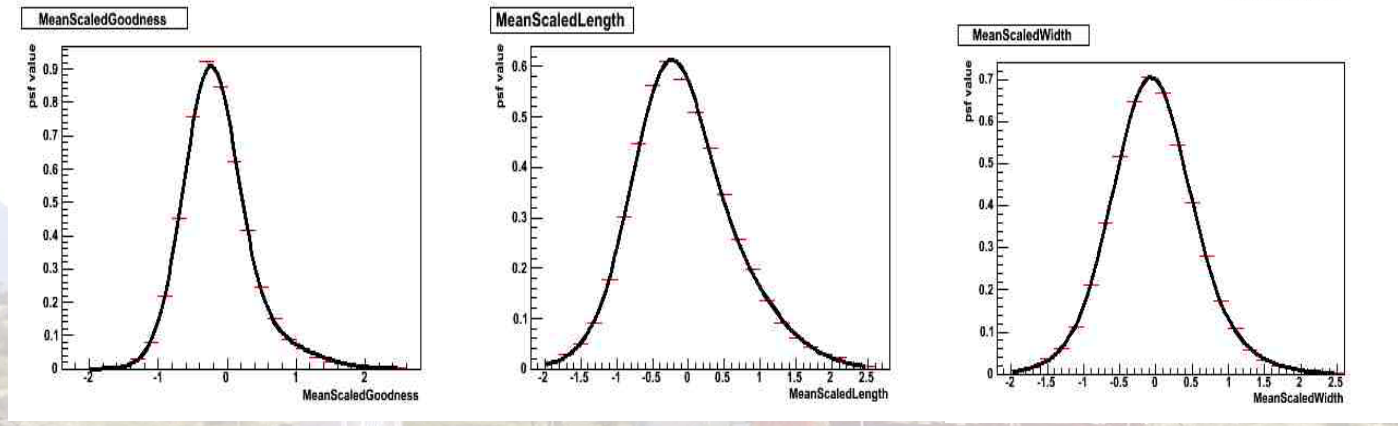


- Some correlation for the background case:

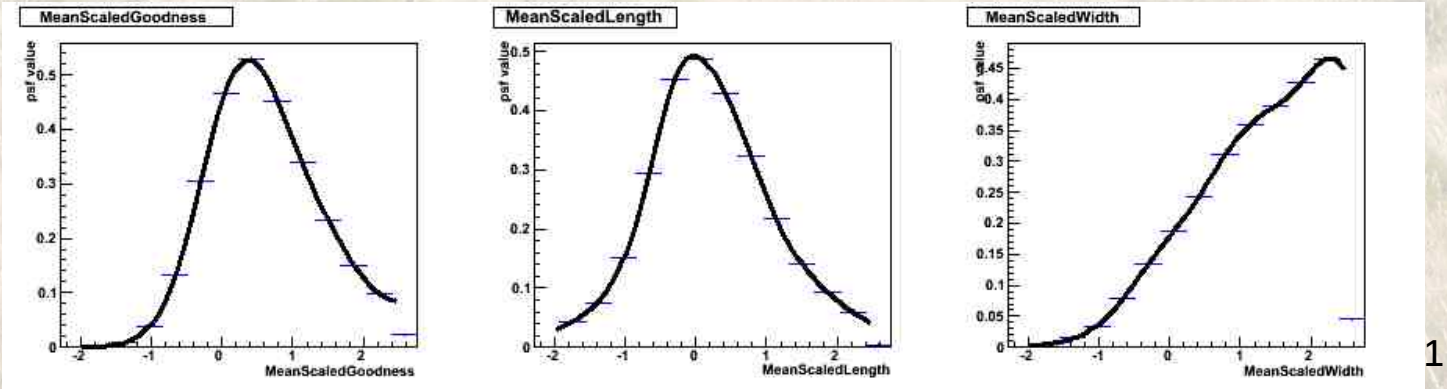


Combining analysis methods

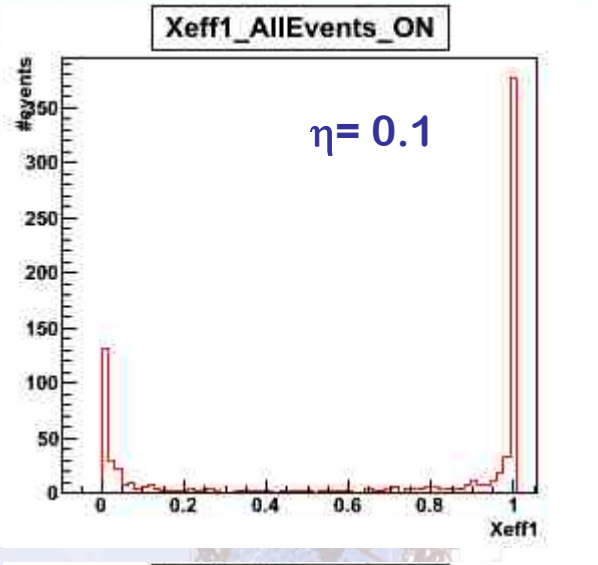
- Probability density finctions (Pdf) :
  - Monte Carlo for signal (γ),

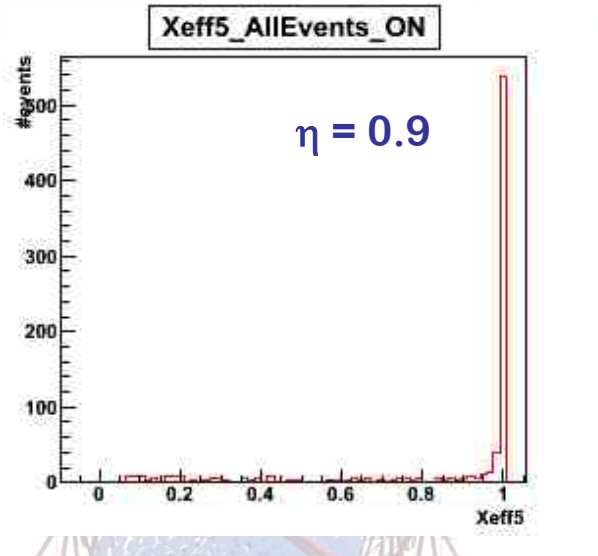


  - OFF-source data for a more robust Background estimation

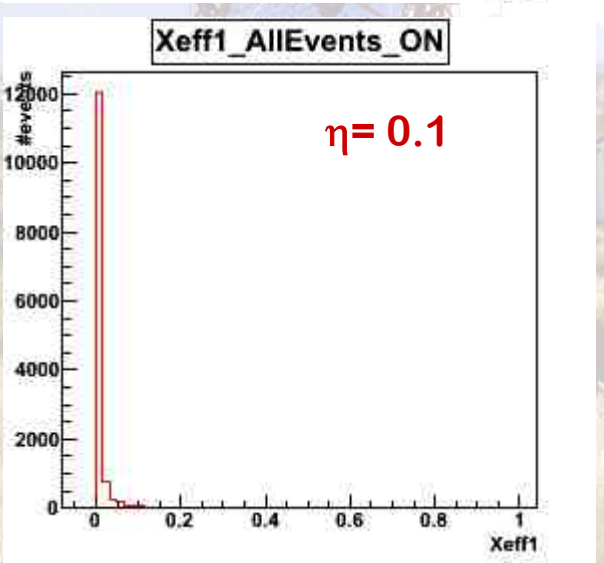


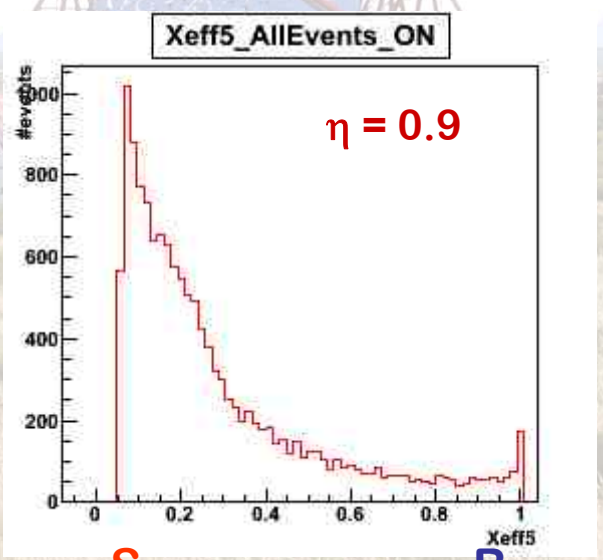
Combined estimator: MC crosscheck





Proton MC





Gamma MC

S **B**

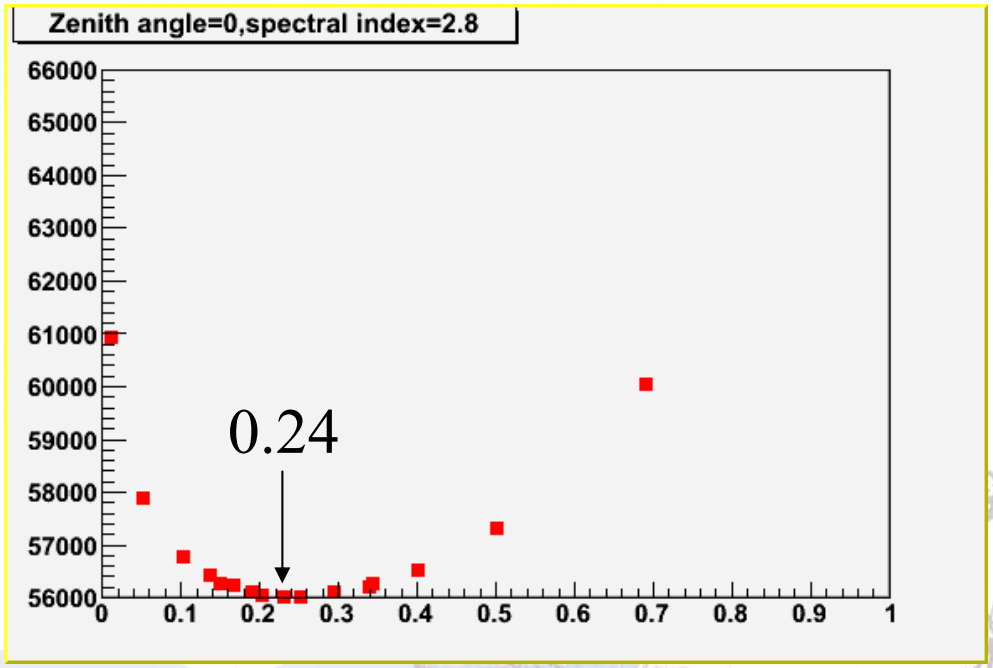
Finding ETA..

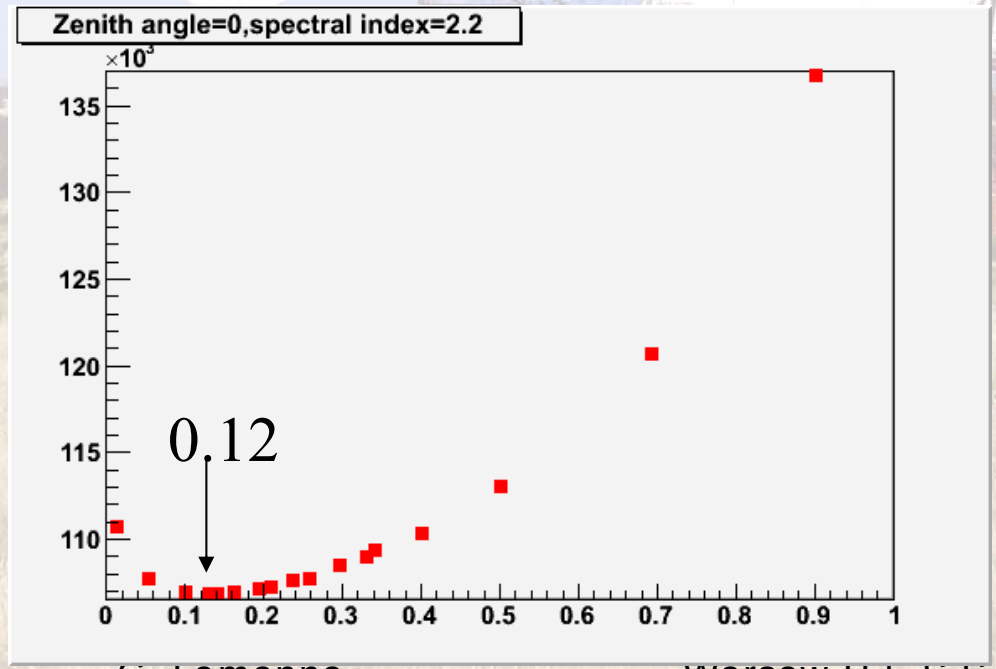
η = b/(s+b) estimated by means of the maximum likelihood method

$$X_{Like} = \eta \Pi_B + (1-\eta)\Pi_S$$

$$\text{Minimum} \left[\sum_{Events} -\ln(X_{Like}(\eta, Events)) \right] \longrightarrow \eta$$

– Validation through MC , knowing “a priori” the «b» and «s» statistics:





DST MC γ and protons :

n_S=8983

n_B=2911

Expected :

η=0.2447

One different γ sample:

n_S=22215

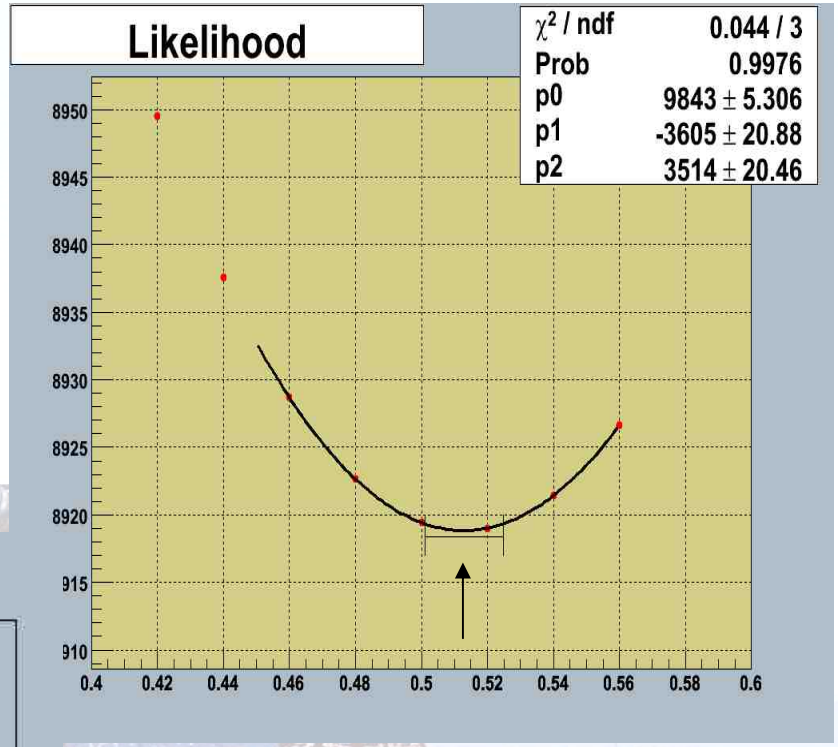
n_B=2911

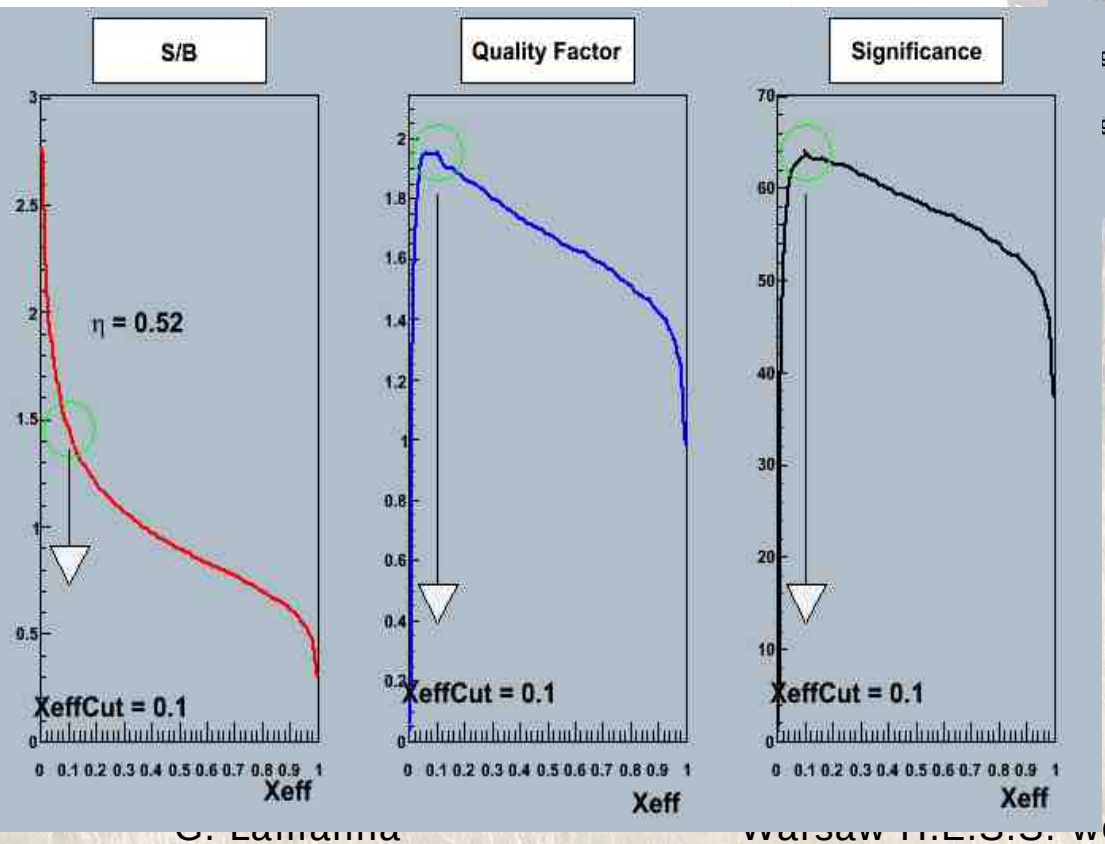
Expected :

η=0.1159

Galactic Centre

Signal-Pdfs: Gamma-MC
Background-Pdfs: OFF data

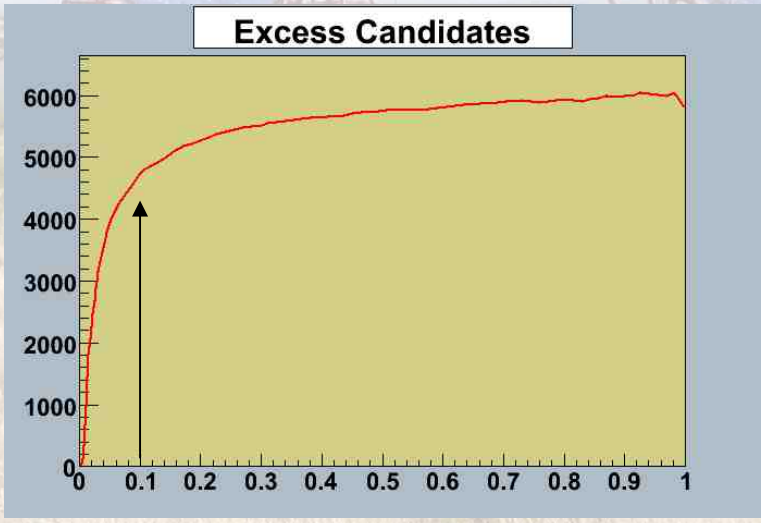




Best Fit results:
Min(-log(L_max) +/- ½) →
Eta = 0.513 + 0.009 - 0.012

COMPARING RESULTS

GC (2004+2005)	$\Delta\theta^2$	OFF	γ	S/B	σ
CombinedCut2	0.02	36640	5140	0.9	56
Multivariables η =0.52 Xeffmax=0.1	0.02	**18231**	**4142**	**1.6**	**63**

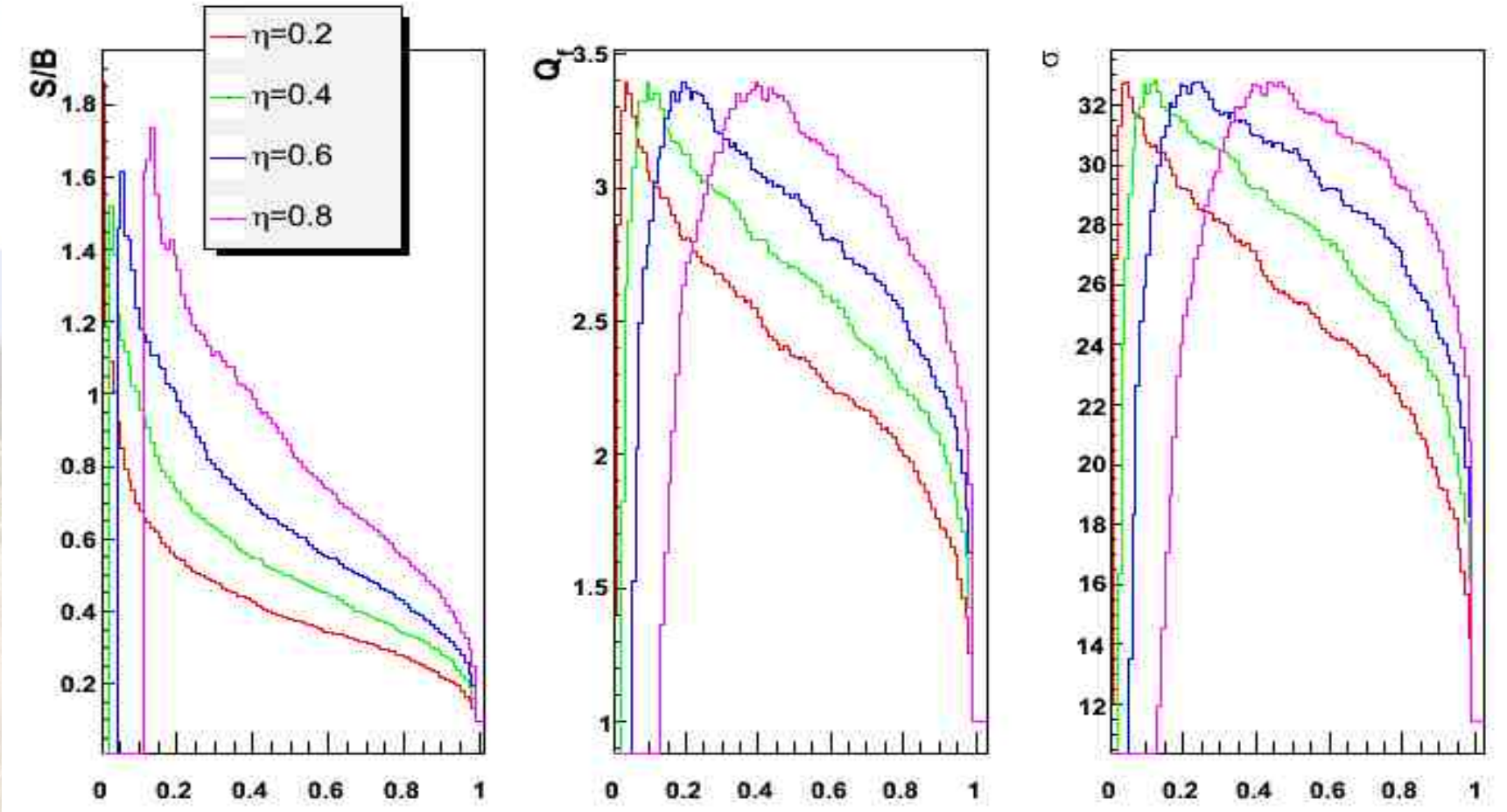


Privileging S/B, Q and σ

η DEPENDANCE

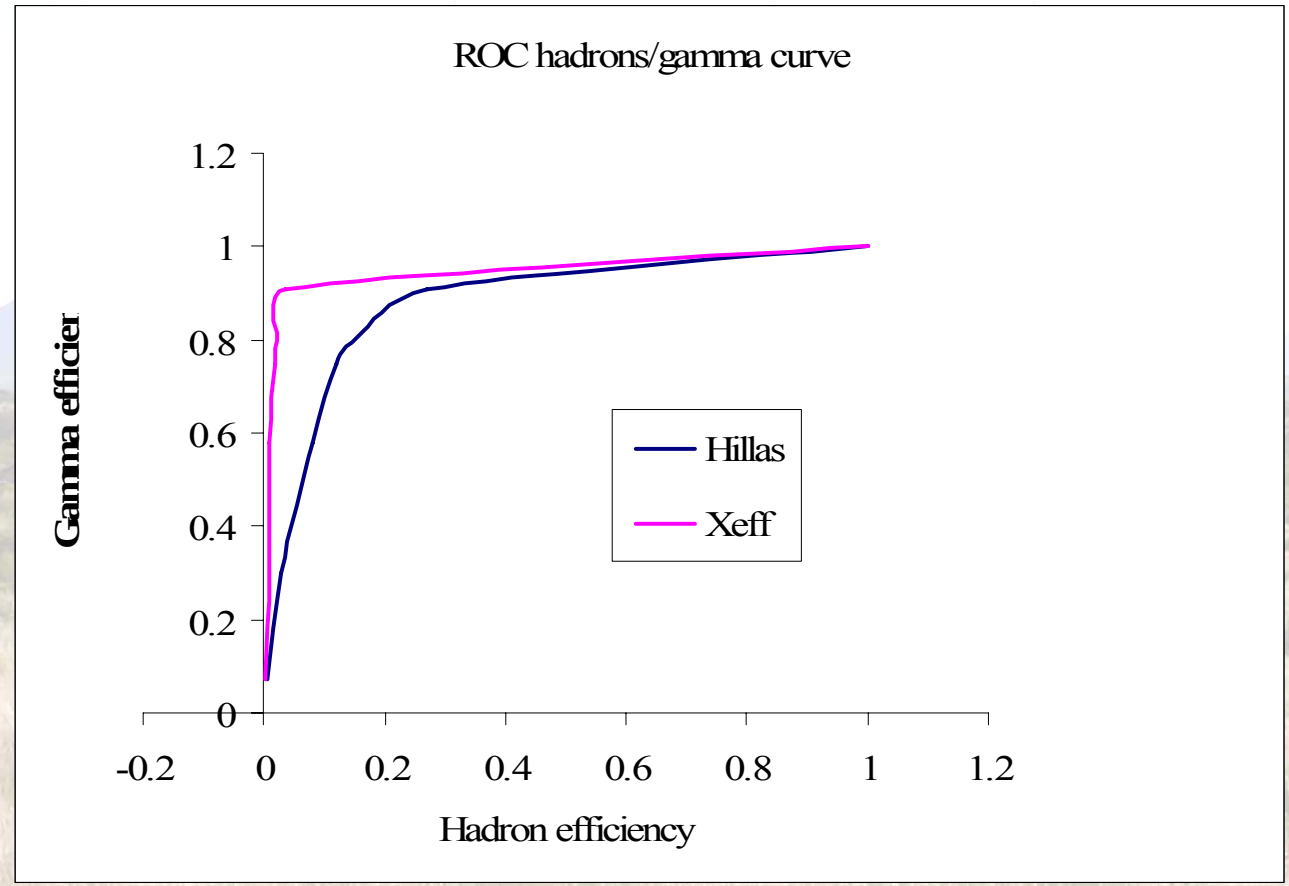
Correlation between η value and $X_{eff\text{-}CUT}$

Chosen η value : 0.5



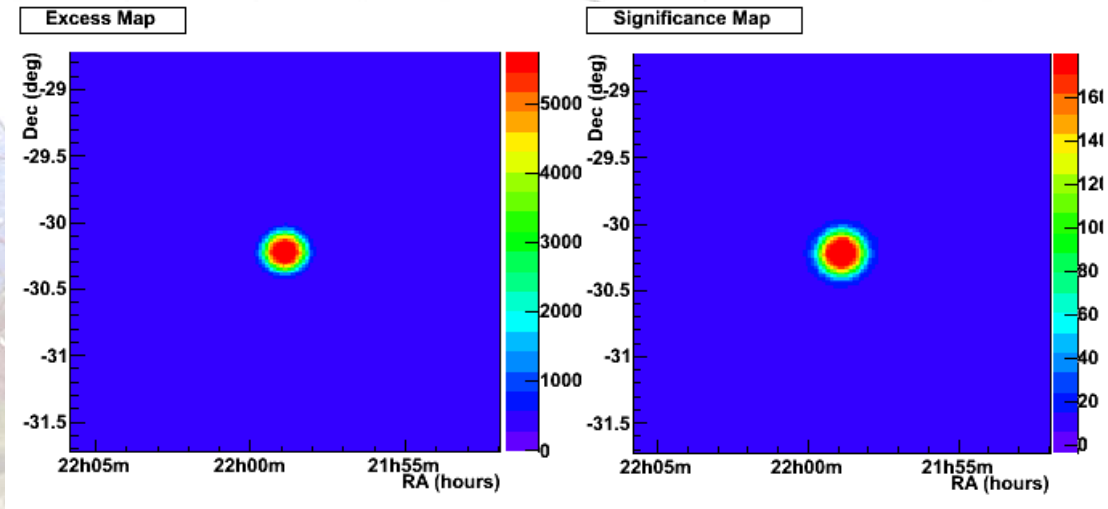
Receiver Operator Characteristic (ROC) diagrams

They show gamma acceptance as function of hadron acceptance,
to compare γ/h separation in a test sample (Vela X) by Xeff and by Hillas scaled



Point-like source

- PKS2155
  - July 2006-flare

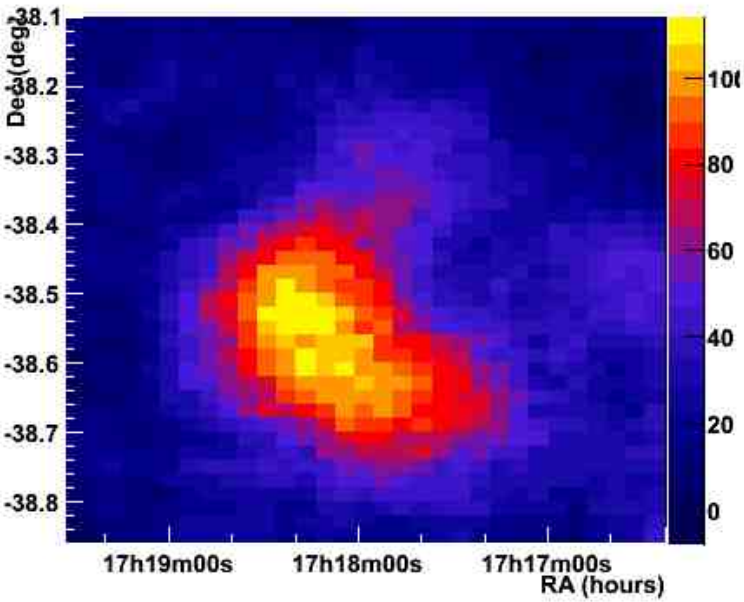


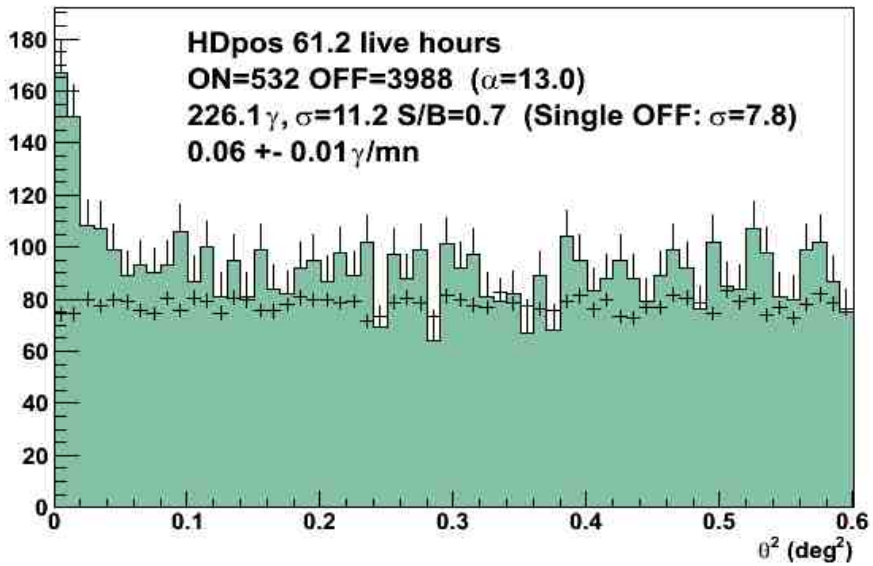
	N_{ON}	N_{OFF}	N_{γ}	S/B	σ
Analyse X_{eff}	8993	446(5.0)	8904	99.8	169.7
Analyse publié	12480	3296(4.7)	11771	16.8	168

Example 2: J1718-385

- **Same preselection as published but Xeff applied**

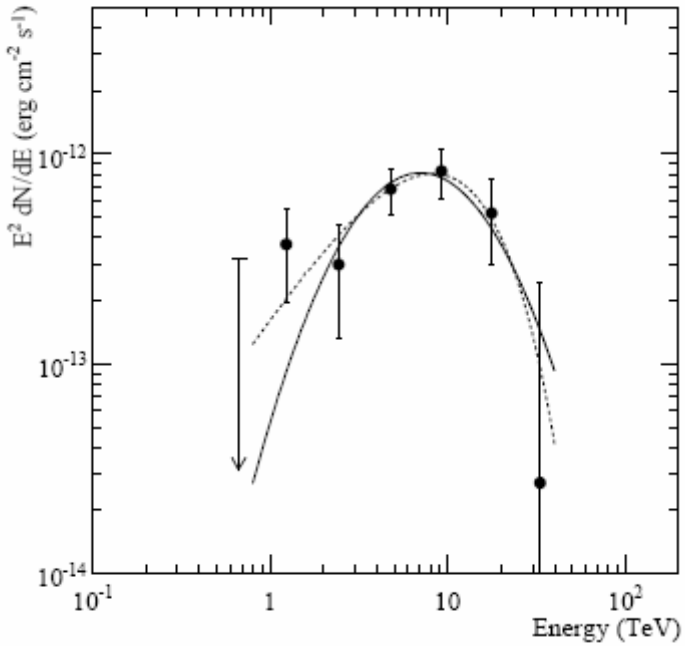
	γ	OFF	σ	S/B
X_{eff}	226	3988	11.2	0.7
Hillas (published)	343		7.6	





Published best fit

$$\frac{dN}{dE} = N_0 \left(\frac{E_{\text{peak}}}{1\ \text{TeV}} \right)^{-2} \left(\frac{E}{E_{\text{peak}}} \right)^{(\beta \cdot \ln E/E_{\text{peak}}) - 2.0}$$



The peak energy E_{peak} is $(7 \pm 1_{\text{stat}} \pm 1_{\text{sys}})$ TeV, the differential flux normalisation $N_0 = (1.3 \pm 0.3_{\text{stat}} \pm 0.5_{\text{sys}}) \times 10^{-12}$ TeV^{-1} cm^{-2} s^{-1} and $\beta = -0.7 \pm 0.3_{\text{stat}} \pm 0.4_{\text{sys}}$. This fit has a $\chi^2/d.o.f. = 3.2/3$. The integral flux between 1 – 10 TeV is about 2 % of the flux of the Crab nebula in the same energy range (Aharonian et al. 2006b). This spectral fit was used to derive the energy flux used later in Table 2.

J1718-385 comparison

(Curved Power Law)

$$\frac{dN}{dE} = \Phi_0 \left(\frac{E}{E_0}\right)^{\alpha - \beta \log(E/E_0)}$$

	α	β	χ^2/ndf
X_{eff}	0.99±0.40	0.23±0.13	37.8/20
Hillas (Not published)	**-0.1±3.4**	**1.1±0.2**	**0.92/3**

?

J1718-385 comparison

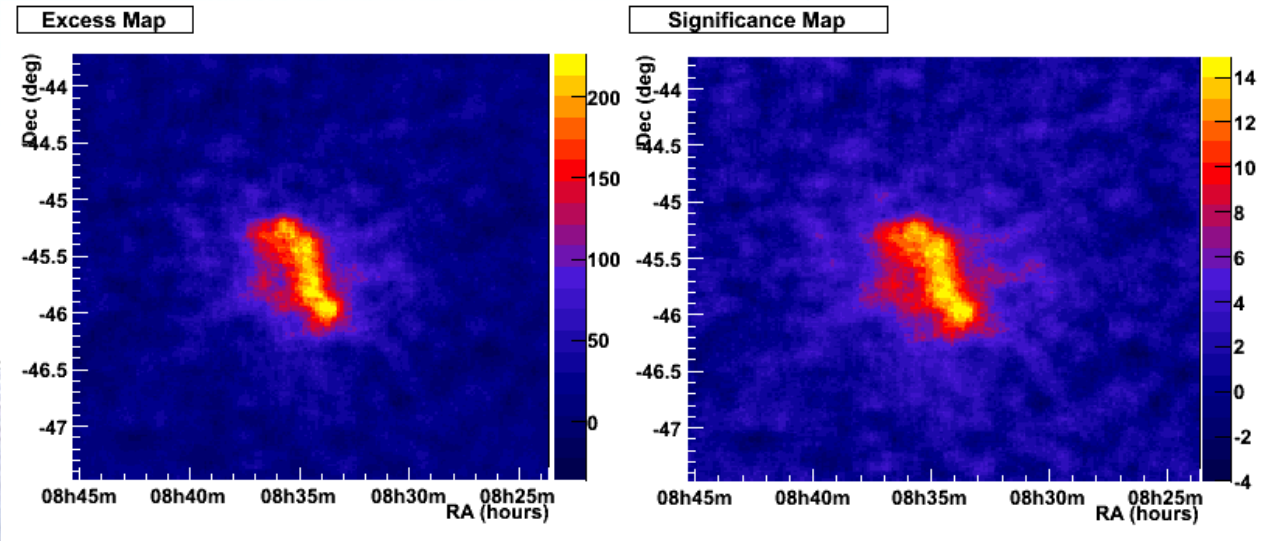
(Exponentially cut-off Power Law)

$$dN/dE = N_0 E^{-\Gamma} e^{-E/E_{cut}}$$

	Γ	E_{Cut} (TeV)	χ^2/ndf
X_{eff}	$1.0 \pm 0.3_{Stat}$	$9.7 \pm 4.7_{Stat}$	34.3/20
Hillas (published)	**$0.7 \pm 0.6_{Stat}$**	**$6 \pm 3_{Stat}$**	**1.6/3**

Sources étendues

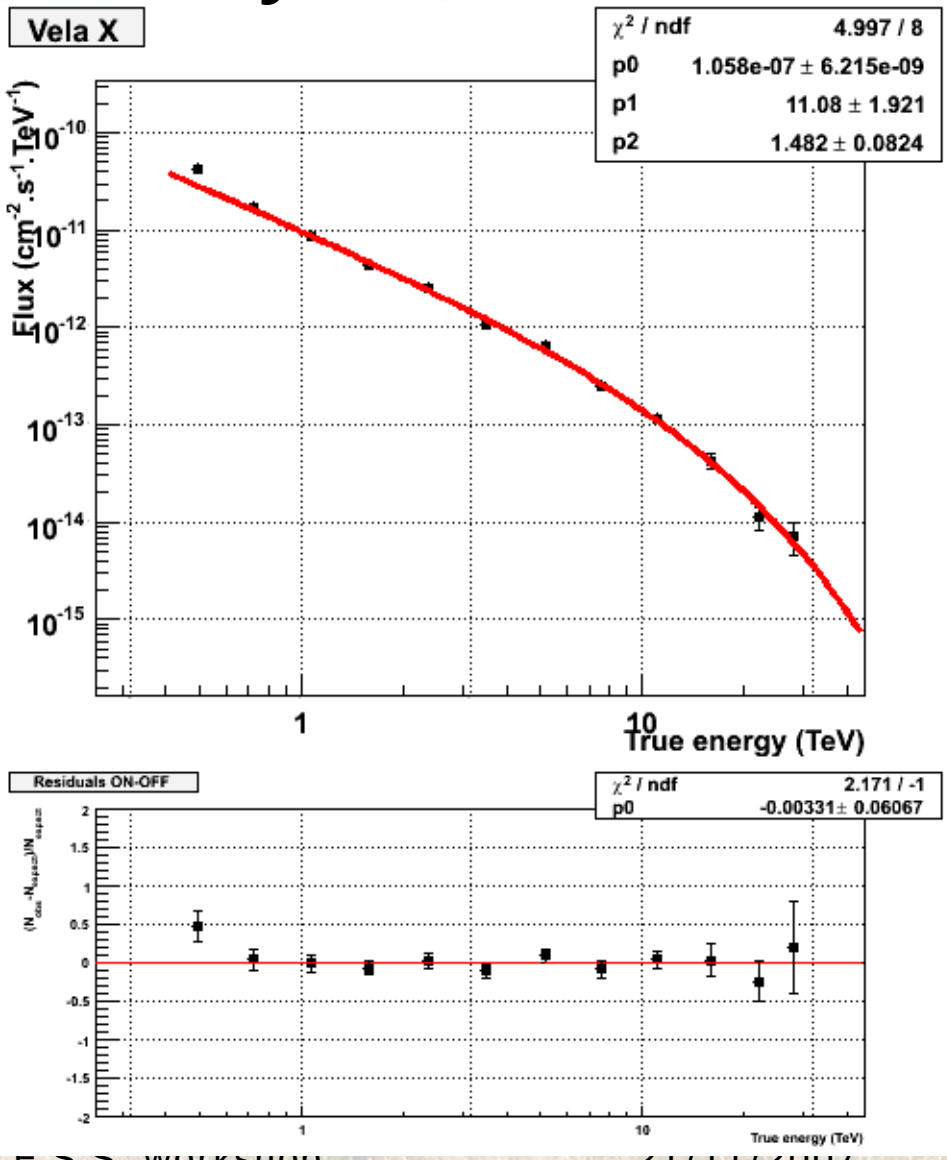
- Vela Pulsar :



	N_{ON}	N_{OFF}	N_{γ}	S/B	σ
Analyse X_{eff} (200 pe)	6668	7478	2539	0.61	28.2
Analyse Hillas (200 pe)	19861	19926	4210	0.27	23.7
Analyse X_{eff} (80 pe)	33506	33806	5418	0.19	22.9

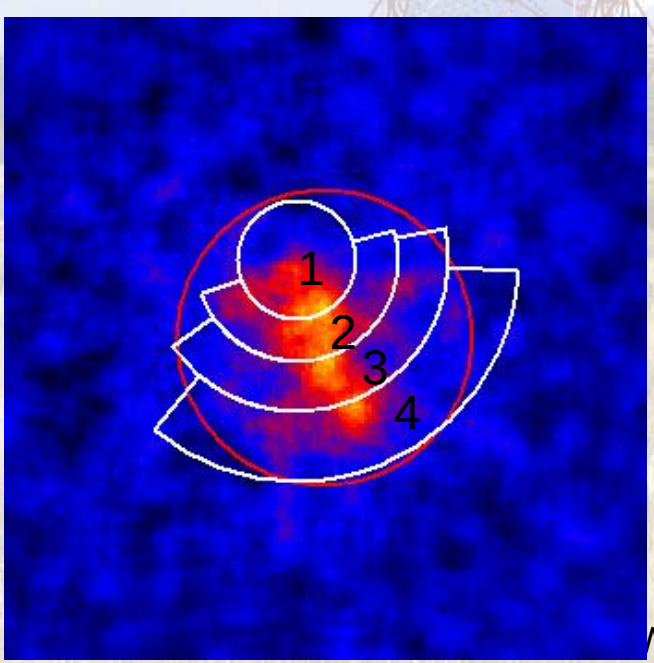
Spectral analysis

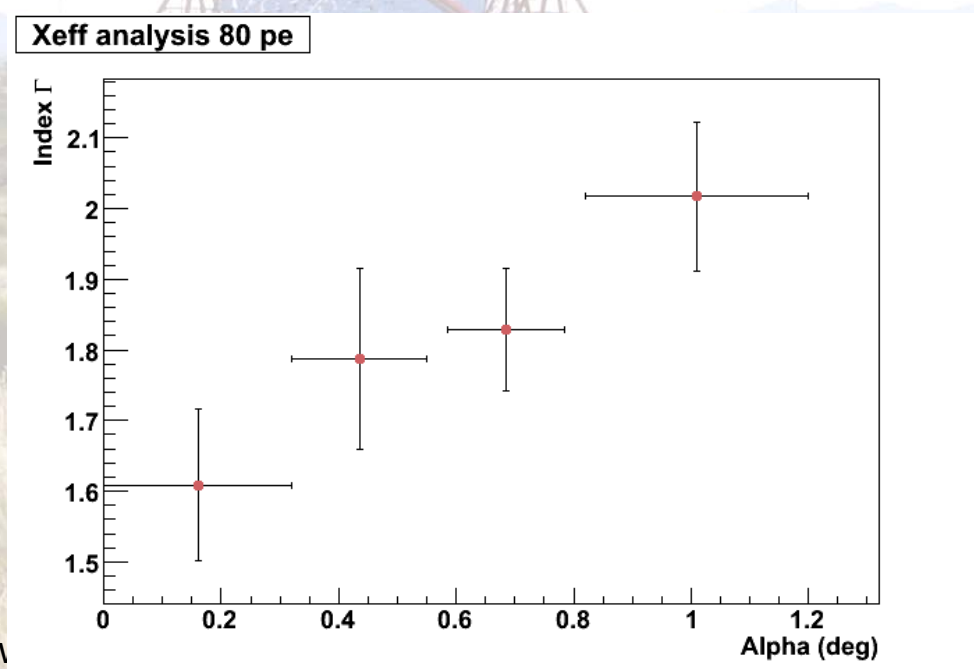
	N_0 ($10^{-8}TeV^{-1}m^{-2}.s^{-1}$)	Γ
Published results		1.45±0.(
X_{eff} analysis 80 pe	**10.6±0.6**	**1.48±0.(



X_{eff}-Morphology

- 4 wedges :
  - 0° to 0.32° from the pulsar,
  - 0.32° to 0.57°,
  - 0.57° to 0.8°,
  - 0.8° to 1.2°.
- Spectrum fitted by Power law from 0.4 TeV to 20 TeV

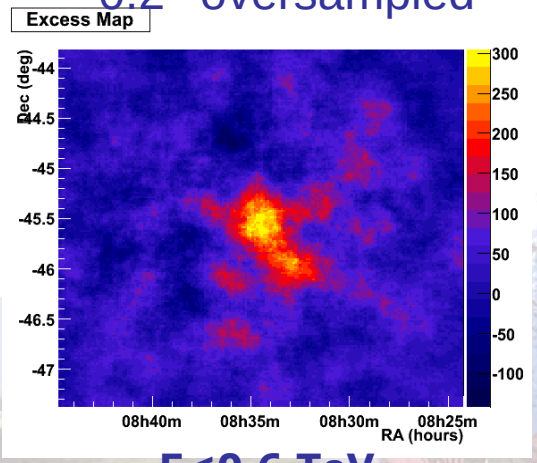




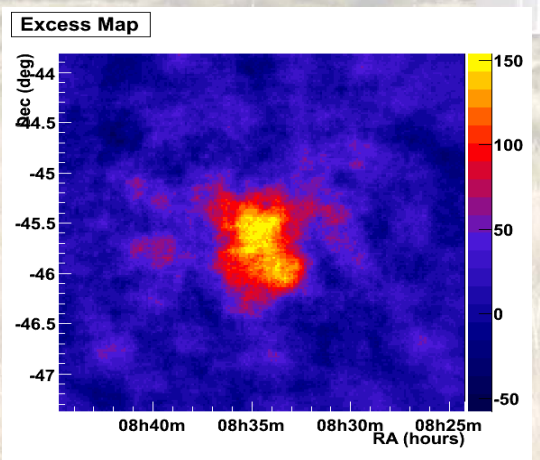
SkyMap vs Energy

Excess SkyMaps vs Energy

0.2°-oversampled



E<0.6 TeV



0.6<E<1 TeV

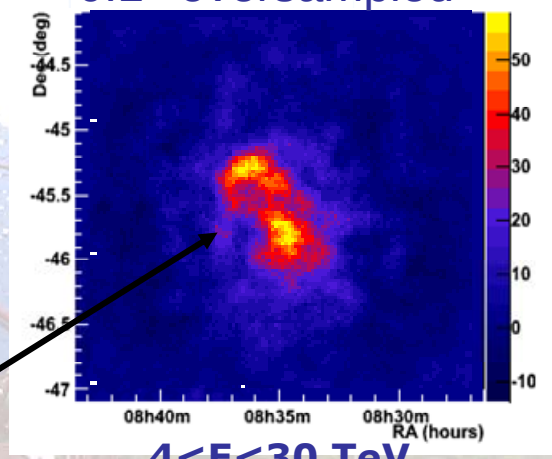
4 energy bands (TeV) : [0;0.6], [0.6;1], [1;4], [4;30]

- **Diffuse emission at lower energy**
- **Vela X extension above 0.8°**

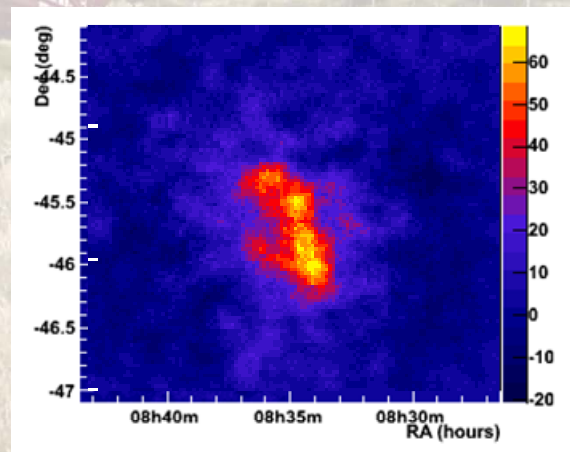
Two hot spots along major axis

Correlated to X-ray ?

0.1°-oversampled



4<E<30 TeV



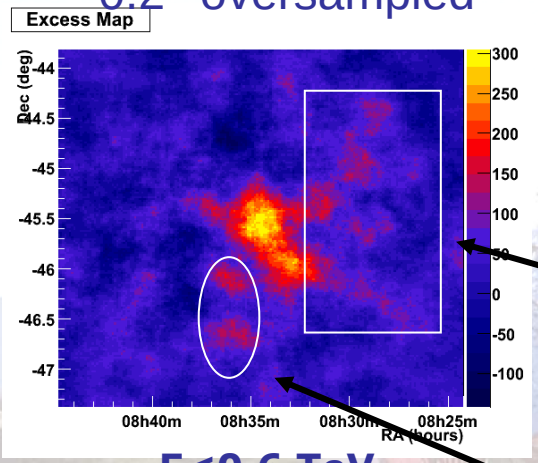
1<E<4 TeV

SkyMap vs Energy

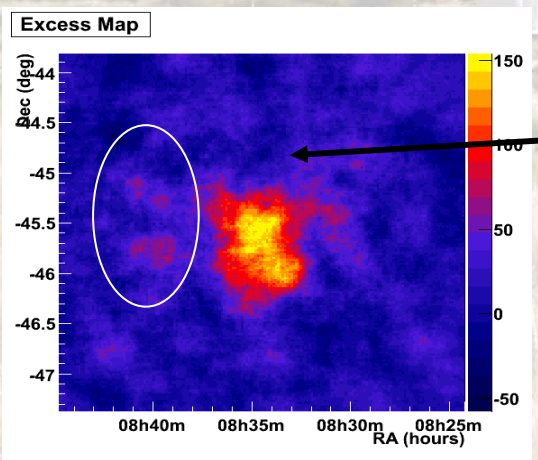
Excess SkyMaps vs Energy

0.2°-oversampled

4 energy bands (TeV) : [0;0.6], [0.6;1], [1;4], [4;30]

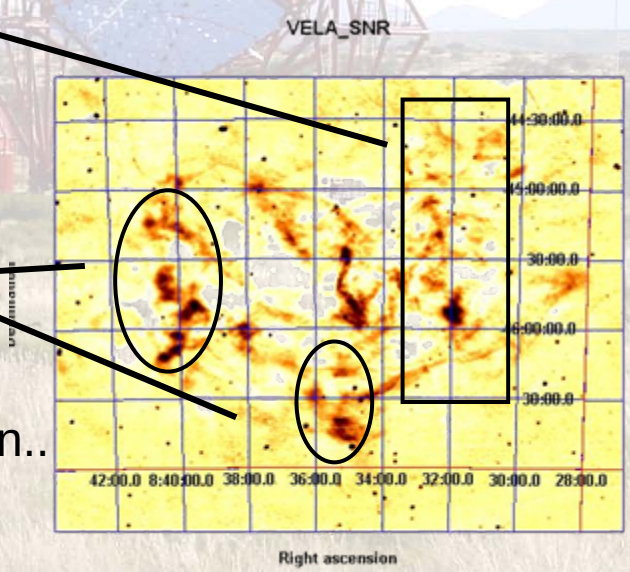


E<0.6 TeV



0.6<E<1 TeV

- **Diffuse emission at lower energy**
- **Vela X extension above 0.8°**

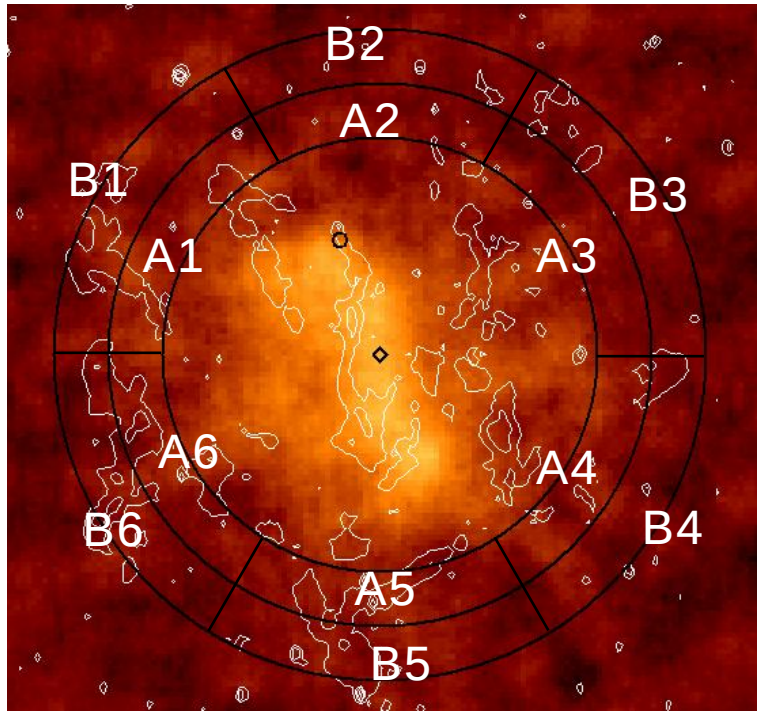


Some structures outside the cocoon..

Correlated with Radio filaments ?

Vela X extension

	Nγ	σ		Nγ	σ
A1	119	3.1	B1	60	1.5
A2	56	2.2	B2	77	2.9
A3	37	0.9	B3	45	1.1
A4	81	2.0	B4	57	1.4
A5	34	1.3	B5	53	2.0
A6	108	2.9	B6	40	1.0



- Excess events found at the East of the source between 0.8° and 1°, and at the North (not correlated with radio emission), but not really significant.
- Need more observation to have a better estimation.

Conclusions

- Multivariate approach provide simultaneous application of different complementary, uncorrelated analysis methods
- Xeff is validated, successful and promising in S/B especially at low energy (e.g HESSII, CTA)
- Morphology studies improved already in HESSI data analysis.
- To be done: get rid of residual correlation among variables for bg. (Correlation function instead Π of Pdfs)