

PHYSTAT-Gamma 2022, 28-30 Sep 2022

Statistical methods for data analysis:

High-energy gamma-ray astronomy in a multiwavelength context

Welcome!

X-RAYS
T
A
T
RADIO
S
T
I
C
 γ -RAYS

LOC: Gerrit Spengler, Thomas Lohse, Ullrich Schwanke
(Humboldt University Berlin), Manuel Meyer (University of Hamburg)
PHYSTAT: Olaf Behnke (DESY), Louis Lyons (Imperial College)

<https://espace.cern.ch/phystat>



PHYSTAT-Gamma

Future PHYSTAT Workshops:

- * PHYSTAT- Gamma: High Energy Gamma Ray Astronomy in a Multi-Wavelength Context, <https://indico.cern.ch/event/1122011/> , 27-30 Sep 2022
- * PHYSTAT-Anomalies: Model-independent searches for New Physics, <https://indico.cern.ch/event/1138933/> , 24th and 25th May 2022

Links to past events: PHYSTAT workshops:

- PHYSTAT-Systematics workshop 1-3 Nov + 10 Nov 2021

- PHYSTAT-FLAVOUR 2020 virtual workshop 19-21 Oct 2020



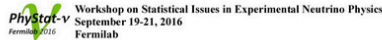
- PHYSTAT-DM 2019 (Stockholm University) Jul 31 - Aug 2, 2019 "Statistical Issues in direct-detection Dark Matter search experiments"



- PHYSTAT-nu 2019 (CERN) Jan 22-25



- PHYSTAT-nu 2016 (FNAL)



- PHYSTAT-nu 2016 (Kavli, Japan)



- PHYSTAT 2011 (CERN) Proceedings "Statistical issues related to discovery claims in search experiments, concentrating on those at the LHC, + Unfolding workshop"

- PHYSTAT 2007 (CERN) Link to proceedings "Statistical issues for LHC physics."

- PHYSTAT 2005 (Oxford) Proceedings  "Statistical Problems in Particle Physics, Astrophysics and Cosmology"

- PHYSTAT 2003 (SLAC)  "Statistical Problems in Particle Physics, Astrophysics and Cosmology"

- PHYSTAT 2002 (Durham) "Advanced statistical analysis techniques as used in measurements and searches in Particle Physics, including Astroparticle Physics"

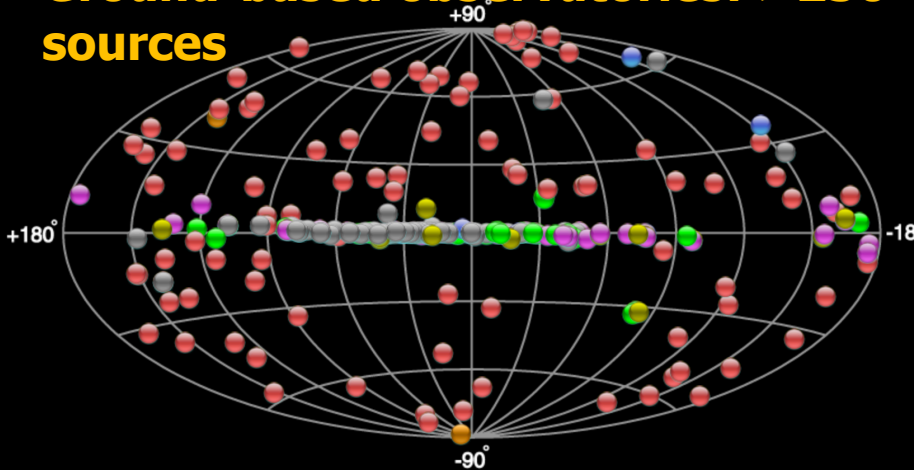
- Link to the agenda and materials

- Workshop on Confidence Limits 2000 (FNAL)

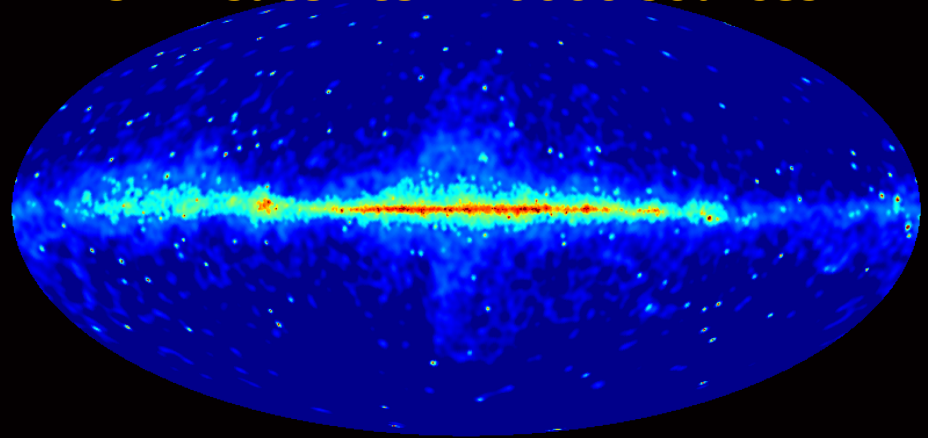
- 1st Workshop on Confidence Limits 2000 (CERN)

Why now a PHYSTAT-Gamma?

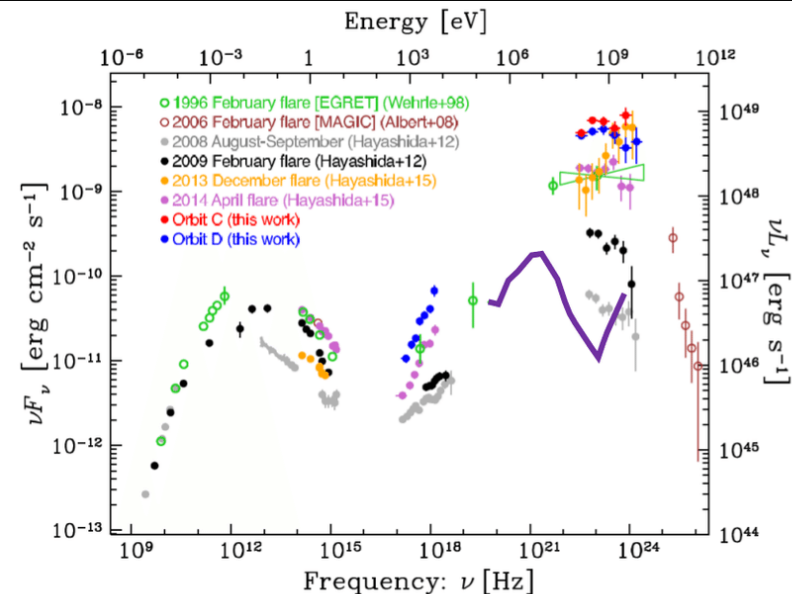
Ground-based observatories: >250 sources



Fermi satellite: > 6600 sources



- In the past 20 years, the current gamma-ray observatories (satellites and ground-based instruments) have detected numerous sources
- Their (astro-)physics can only be understood in a multiwavelength context (X-rays, radio, optical, ...)



Why now a PHYSTAT-Gamma?

- **In particular the ground-based Cherenkov instruments are in a transition phase from the current experiments (H.E.S.S., MAGIC, VERITAS) to the CTA observatory**
-

Cherenkov Telescope Array

Large-Sized
Telescope (**LST**)

Small-Sized
Telescope (**SST**)

Medium-Sized
Telescope (**MST**)

+ atmosphere monitoring

α configuration		LSTs	MSTs	SSTs
CTA North	La Palma (Spain)	4	9	
CTA South	Paranal (Chile)	0	14	37

Why now a PHYSTAT-Gamma?

- **In particular the ground-based Cherenkov instruments are in a transition phase from the current experiments (H.E.S.S., MAGIC, VERITAS) to the CTA observatory**
 - **CTA will bring the source count to > 1000 and will serve a diverse scientific community**
 - **A good time to review the statistical methods for data analysis, to see what others (X-rays, radio) have done and to converge (e.g. on software and documentation/recommendations)**
 - **Benefit from the interchange with (astro-) statisticians and related fields (e.g. particle physics)**
-

Why now a PHYSTAT-Gamma?

- **In particular the ground-based Cherenkov instruments are in a transition phase from the current experiments (H.E.S.S., MAGIC, VERITAS) to the CTA observatory**
- **CTA will bring the source count to > 1000 and will serve a diverse scientific community**
- **A good time to review the statistical methods for data analysis, to see what others (X-rays, radio) have done and to converge (e.g. on software and documentation/recommendations)**
- **Benefit from the interchange with (astro-) statisticians and related fields (e.g. particle physics)**

- **Before I forget: The mastermind behind all this is**

Gerrit Spengler



Today

14:00

Workshop Introduction

Speaker: Ullrich Schwanke

14:10

Introduction: Making the sum greater than its parts

Without a proper accounting of known and unknown systematics and uncertainties, combining information across multiple surveys, wavelengths, and detectors may be risky. Realizing the true potential of multi-messenger and panchromatic astrophysics requires getting data integration right.

Speaker: Joshua Speagle (University of Toronto)

14:30

Discussion

⌚ 5m

14:35

Statistical methods for VHE gamma-ray astronomy: Gammapy

Speakers: Axel Donath, Regis Terrier

⌚ 30m

15:05

Discussion

⌚ 10m

15:15

Coffee break

⌚ 20m

15:35

Statistical methods for X-ray astronomy

Speaker: Johannes Buchner

⌚ 30m

16:05

Discussion

⌚ 15m

16:20

Statistical methods for the combination of data: ThreeML

As the multi-messenger era is now fully active, it is crucial that the community has a framework within which to analyze data from multiple messengers, wavelengths, and instruments in a statistically robust, common way. 3ML (<https://threeml.readthedocs.io>) provides an abstract, plugin-based data interface for instruments to combine analysis through each instrument's own unique likelihood. Users and instrument teams can create or use existing plugins to interface their data to a plethora of Bayesian and optimization packages in a uniform way. Analysis results are reported and stored in portable file formats that allow for the sharing and replication of results in a way that provides observers to produce robust scientific results that the community can interpret. 3ML currently supports, via standard plugins many ground and space-based observatories as well as being the analysis tool for some collaborations (HAWC, XIPE, Fermi-LAT, POLAR, GECAM).

Speaker: J Michael Burgees

⌚ 30m

16:50

Discussion

⌚ 10m

17:00

Selected problems from HAWC analyses

Speaker: James Linnemann

⌚ 30m

Thursday

14:00

Session introduction

Speaker: Ullrich Schwanke

14:05

Astrostatistics: Overview and highlights

Speaker: Eric Feigelson

⌚ 40m 

14:45

Chi-square, K-S, and bootstrap: Fitting astrophysical models to data

Complicated models from astrophysical theory are often fit to observational data. There are several issues with the classical procedures used in astronomy literature. First, 'chi-square minimization' is commonly used for fitting functions often disregard mathematical assumptions. Second, the Kolmogorov-Smirnov (K-S) test for goodness-of-fit is misused in astronomy when the model parameters are estimated from the dataset under study. Third, the KS is inefficient at detecting deviations between the data and model at the tails of the distribution. Fourth, the K-S test cannot justifiably be applied to multivariate data as KS is no longer distribution-free. Recent developments of bootstrap resampling method, a simple Monte Carlo procedure on data, will be described, to address these issues.

Speaker: Jogesh Babu

⌚ 40m 

15:25

Discussion⌚ 15m 

15:40

Coffee break

⌚ 20m

16:00

Bayesian methods

Speaker: Tom Loredo

⌚ 40m 

16:40

Discussion ¶⌚ 15m 

16:55

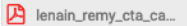
Time Series Analysis In the Dynamic Universe

Speaker: Jeff Scargle

 phystat_talk_repaire...⌚ 40m 

17:35

Discussion⌚ 20m 

14:30	Session introduction Speaker: Ullrich Schwanke	5m
14:35	Counterpart identification: Overview Speaker: Tamas Budavari	30m
15:05	Discussion	15m
15:20	VHE gamma-ray surveys with CTA The Cherenkov Telescope Array (CTA) will be the first astronomical observatory fully covering the gamma-ray sky in an energy range from 20 GeV up to 300 TeV. The observatory will be composed of two arrays of tens of telescopes located in La Palma, Spain, and Paranal, Chile. Among the Key Science Projects proposed by the CTA Consortium, Galactic and extragalactic surveys will be conducted during the first years of operation. With an unprecedented sensitivity and improved angular resolution, CTA surveys promise the discovery of several hundred of new gamma-ray sources, but the challenges coming along with the analyses of these data will also scale up. We will focus on the challenges of source variability, extended sources modeling, source confusion, source association with multi-wavelength catalogues, classification in source populations, and sources contamination due to the systematic errors in the modeling of instrumental and astrophysical backgrounds. Speakers: Jean-Philippe Lenain, Quentin Remy 	30m
15:50	Discussion	15m
16:05	Coffee break 15m	
16:20	Identifying correct counterparts to high-energy sources by "multiwavelength educated guesses" imbibed in a Bayesian statistic environment The identification of the counterparts to sources detected by instruments with large positional uncertainties can not be done using match in coordinates, due to the very high number density of the ancillary source catalogs. In addition, given that now the entire sky is literally covered by a plethora of multiwavelength surveys, the search for the counterparts by using a single band at a time is outdated. Instead, the entire SED for every single source in the sky can be created and used for discriminating the actual emitter from the field population. Finally, at least with respect to X-ray observations, we have more than 20 years of XMM and Chandra detection with a secure counterpart that can be used for creating a training sample to educate our guess. This is the basis of NWAY, a cross-matching code based Bayesian statistics that works with arbitrarily many catalogs, can handle varying positional errors, can incorporate additional prior information (the educated guesses and works accurately and fast in small areas and all-sky catalogues. In my talk, I will present how NWAY is now routinely used in the determination of the counterparts to Xray sources detected by e.g. ROSAT, XMMslew, NUSTAR, and eROSITA. In particular, I will show how the prior (based on photometry, colors, parallax, and SNR of the detection) was built for eROSITA using Random Forest and tested on a validation sample providing 96% completeness and purity. The final goal is to discuss with the audience how a similar approach could be built for CTA. Speaker: Mara Salvato	30m
16:50	Discussion	15m
17:05	Radio surveys Speaker: Beatriz Mingo	30m
17:35	Discussion	15m

Friday

Technicalities

- **Session are scheduled from 2pm-6pm CEST**
 - **Please make sure you are muted most of the time**
 - **We will collect questions in the chat – but might be forced to make a selection**
 - **The workshop is recorded and videos will be made available on the phystat web page if speakers agree**
-

PHYSTAT-Gamma 2022, 28-30 Sep 2022

Statistical methods for data analysis:

High-energy gamma-ray astronomy in a multiwavelength context

Enjoy the
Workshop!

X-RAYS
T
A
T
RADIO
S
T
I
C
 γ -RAYS

LOC: Gerrit Spengler, Thomas Lohse, Ullrich Schwanke
(Humboldt University Berlin), Manuel Meyer (University of Hamburg)
PHYSTAT: Olaf Behnke (DESY), Louis Lyons (Imperial College)

<https://espace.cern.ch/phystat>

