



Universidad del Desarrollo
Facultad de Ingeniería

GUÍA INTRODUCTORIA COMO CAPACITACIÓN EN EL ESTUDIO EN FUENTES DE RAYOS GAMMA

*Estudio y análisis de eventos espaciales de alta energía para promover el
interés y desarrollo de competencias necesarias para proyecto SWGO en
Chile*

POR: CARLOS BUSTAMANTE

JORGE MORGADO

VÍCTOR ZÚÑIGA

Proyecto de grado presentado a la Facultad de Ingeniería de la
Universidad del Desarrollo para optar al grado académico de
Magíster en Data Science

PROFESOR GUÍA:

Sr. GERMÁN GÓMEZ VARGAS

DICIEMBRE 2024

SANTIAGO, CHILE

AGRADECIMIENTOS

Como equipo este trabajo no habría sido logrado sin el apoyo y orientación de nuestro profesor guía Germán Gómez quien nos entregó feedback constantemente por medio de su gran expertis para definir bien los lineamientos de los objetivos principales y secundarios que teníamos considerados desde un principio, y que permitieron enmarcar nuestro trabajo hacia una necesidad industrial académica.

Por otro lado agradecer al trabajo colaborativo en cuanto al apoyo y esfuerzo que cada uno de nosotros entregó para poder concretar el desarrollo de cada tópico de este trabajo, que si bien nos sacó de nuestra zona de confort en cuanto al tipo de datos que acostumbramos a ver o trabajar, logramos sortear las dificultades inherentes del mismo gracias al trabajo en equipo.

TABLA DE CONTENIDO

Resumen.....	1
1. Introducción.....	2
2. Objetivo principal y secundarios.....	4
Objetivo principal.....	4
Objetivos secundarios.....	4
3. Datos y Metodología.....	6
3.1. Captura de los datos.....	6
3.2. Metodología.....	12
4. Resultados.....	15
4.1. Funciones de Respuesta del Instrumento IRF (Por Carlos Bustamante).....	15
OBJETIVO.....	15
RAYOS GAMMA.....	16
IRF.....	20
APLICACIÓN DE IRF EN RAYOS GAMMA.....	22
IRF EN CTA NORTH.....	24
ESTUDIO DEL ÁREA EFECTIVA.....	28
BACKGROUND.....	37
SOPORTE DE GAMMAPY PARA IRF ASIMÉTRICAS.....	43
MAPAS DE EXPOSICIÓN.....	45
CONCLUSIÓN.....	46
4.2. Predicción y Análisis Temporal de Eventos en Datos del Telescopio Fermi usando Machine Learning (Por Jorge Morgado).....	47
Contexto.....	47
Metodología.....	48
Resultados.....	51
Conclusión.....	63
4.3. Análisis con Gammapy (Por Víctor Zúñiga).....	65
Contexto.....	65
¿Por qué usar Gammapy?.....	66
Necesidad del campo educativo.....	67
Relación con el nested model.....	67
Objetivo del análisis.....	67
Manos a la obra (Step-by-Step).....	70
Conclusión.....	122

5. Conclusiones.....	123
6. Bibliografía.....	127
7. Anexo I: Encuesta Rebeca Burgos estudiante de pregrado en Astronomía.....	133
8. Anexo II: Encuesta Mauricio Baeza estudiante de pregrado en Astronomía.....	136
9. Anexo III: Descripción de atributos EventList.....	138
10. Anexo III: Descripción de atributos catálogo Fermi 3FHL.....	140
11. Autoevaluación/Coevaluación.....	145

Resumen

Este trabajo corresponde a una guía capacitadora e incluso motivacional para profesionales o interesados en el estudio del campo de la astronomía, precisamente en las fuentes espaciales de rayos gamma. Un campo de la ciencia que ha tomado gran relevancia en Chile especialmente por el observatorio terrestre SWGO que se encuentra en construcción y espera esté operativo en 2030. Este hito del observatorio SWGO en Chile ha procurado la necesidad de la industria académica e investigativa de formar más astrónomos o especialistas en este campo. Esto es nuestra principal motivación, razón por la cual nos aventuramos en crear esta guía especie starter pack (en base a nuestra expertise con proyectos de datos) para los eventuales incumbentes que quieren versarse y por qué no, desempeñarse laboralmente en el eventual observatorio de rayos gamma SWGO.

La idea es que esta guía sea un primer encuentro cercano hacia el estudio con datos de las fuentes de rayos gamma, de tal manera que sea amigable, fácil de leer, y lo más importante que permita al incumbente identificar patrones espaciales, energéticos en las emisiones de rayos gamma, validar teorías astrofísicas y desarrollar habilidades en el análisis de estos datos.

Para lo anterior abordaremos esta guía en 3 tópicos distintos pero que finalmente convergen al mismo objetivo. El primero será IRF's (Instrument Response Functions), luego análisis de fuentes de rayos gamma (con Gammapy), y finalmente Machine Learning aplicado con estos datos.

1. Introducción

El estudio de las fuentes energéticas de rayos gamma está tomando cada vez mayor auge en el mundo académico y de investigación en Chile. Esto se complementa con que en 2026 se pondrá en marcha la construcción del **observatorio de rayos gamma cósmicos SWGO** (Southern Wide field Gamma ray Observatory), la cual se espera estará operativo en 2030.

El investigador Walter Max-Moerbeck del Centro de Astrofísica y Tecnologías Afines (CATA), y académico de la Universidad de Chile sostiene que *“Los astrónomos que estudiamos rayos gamma aquí somos muy pocos. Una realidad muy similar a, por ejemplo, lo que ocurría antes de que instalaran ALMA en el norte, pero su implementación gatilló mucha especialización en las ondas milimétricas. Con este proyecto se abre una enorme posibilidad para que las futuras generaciones de astrónomos se focalicen en el área de los rayos gamma”*. (Centro de Astrofísica y Tecnologías Afines [CATA], n.d.) 1

A propósito de esta investigación, se encuestó a dos estudiantes de astronomía (anexos 1 y 2), donde se pudo diagnosticar que ambos están al tanto del proyecto, pero el primero demuestra un entendimiento más detallado sobre su diseño innovador (tanques de agua sellados para captar rayos gamma), ambos comprenden también que los rayos gamma son emisiones electromagnéticas altamente energéticas asociadas con eventos extremos en el universo, como quásares y supernovas.

Respecto al ámbito educacional, los dos coinciden en que los programas académicos actuales tienen áreas de mejora, particularmente en la integración de teoría y práctica, y en el fortalecimiento de habilidades de

análisis de datos. Es por esto que en el contexto formativo y de especialización en el campo de la astronomía y de los rayos gamma, se logran identificar las siguientes necesidades y desafíos para los estudiantes:

- Dificultad de interpretación, análisis y visualización del dato.
- Uso avanzado de herramientas computacionales, de programación y de estadísticas para su manipulación.
- Falta de materiales pedagógicos simplificados y de alto nivel.

Dado lo anterior, en términos de la necesidad imperativa que existe en el mundo investigativo y profesional de formar y capacitar más especialistas en el estudio de rayos gamma (para el proyecto SWGO especialmente), así mismo el desafío que implica para los eventuales interesados en aprender y/o especializarse en este campo, es que motivó nuestro proyecto de tal manera de poder cubrir estas dos necesidades a través de una solución asequible, digital y fácil de interpretar.

Nuestra solución apunta a entregar una guía digital capacitadora alto nivel sobre el estudio de fuentes de rayos gamma, la cual hemos definido en 3 grandes tópicos: primero está la explicación y aplicación de los IRF's (instrument response functions) funciones matemáticas relevantes que permiten convertir las mediciones brutas obtenidas por el instrumento en estimaciones precisas de las propiedades de las fuentes de rayos gamma, como posición, espectro, y flujo.

Luego abordamos el tópico de análisis de fuentes de rayos gamma con una librería especializada en rayos gamma (Gammapy), entregando un tutorial o paso a paso de cómo abordar el análisis de las observaciones que detectan los instrumentos (en este caso aplicado a Fermi).

Finalmente esbozamos la aplicabilidad de herramientas de Machine Learning para mejorar la imagen capturada por el instrumento.

2. Objetivo principal y secundarios

Objetivo principal

Desarrollar una guía práctica y accesible para la capacitación en el análisis de datos de fuentes de rayos gamma, utilizando herramientas como las Funciones de Respuesta del Instrumento (IRFs), la librería Gammapy y técnicas de Machine Learning, con el fin de preparar a futuros profesionales para el estudio y la investigación en astrofísica de rayos gamma en el contexto del observatorio SWGO en Chile.

Objetivos secundarios

- Explicar y aplicar las Funciones de Respuesta del Instrumento (IRFs) para corregir y calibrar los datos obtenidos por telescopios de rayos gamma, mejorando la precisión en la estimación de propiedades como energía, flujo y localización de las fuentes.
- Diseñar y ejecutar un análisis de las fuentes de rayos gamma utilizando la librería Gammapy, facilitando la carga, el análisis y la visualización de los datos de rayos gamma, y promoviendo la identificación de patrones espaciales y energéticos en las emisiones.
- Implementar y evaluar técnicas de Machine Learning para el análisis predictivo y la detección de anomalías en los datos de rayos gamma, mejorando la capacidad de los modelos para identificar eventos extremos y clasificar fuentes en función de sus características temporales y espectrales.

3. Datos y Metodología

3.1. Captura de los datos

Los rayos gamma son un tipo de radiación electromagnética. y por tanto constituida por fotones, producida generalmente por elementos radiactivos o por proceso subatómicos como la aniquilación de un par positrón-electrón. También se genera en fenómenos astrofísicos de gran violencia.

La energía de esta naturaleza se mide en millones de electronvoltios (MeV), donde un MeV corresponde a fotones gamma de longitudes de onda inferiores a 10^{-11} m y a frecuencias superiores a 10^{19} Hz.

Los rayos gamma producidos en el espacio no llegan a la superficie terrestre, pues los absorbe la alta atmósfera. Para observar el universo en estas frecuencias es necesario usar globos de gran altitud u observatorios exo-espaciales (fenómenos astrofísicos de alta energía como explosiones de supernovas o núcleos de galaxias activas). (Donath et al., 2023)²

La captura de datos de rayos gamma, en comparación con otras longitudes de onda del espectro electromagnético (como radio, microondas, infrarrojo, ultravioleta, rayos X), utiliza tecnologías y métodos específicos adaptados a las propiedades físicas de los rayos gamma.

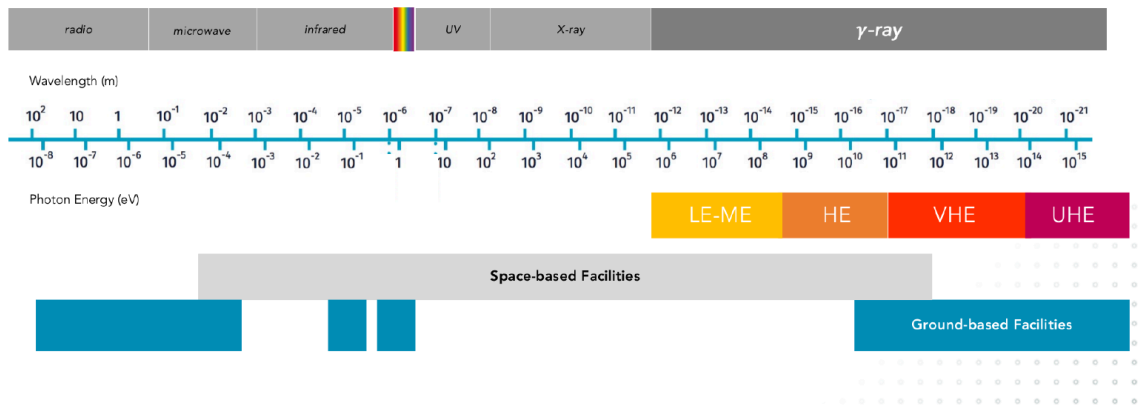


Imagen 1

Distintos instrumentos u observatorios capturan la energía (rayos gamma) en distintos rangos.

Categoría	Rango de Energía	Significado
LE (Low Energy)	100 keV a 1 MeV	Baja energía.
ME (Medium Energy)	1 MeV a 100 MeV	Energía media.
HE (High Energy)	100 MeV a 100 GeV	Alta energía.
VHE (Very High Energy)	100 GeV a 100 TeV	Energía muy alta.
UHE (Ultra High Energy)	> 100 TeV	Energía ultra alta.

Tabla 1

Por ejemplo, en rangos LE-ME encontramos procesos térmicos de baja energía como interacciones de partículas cargadas en campos magnéticos, algunas fuentes típicas son estallidos de rayos gamma (GRB's), supernovas, y radiación difusa galáctica.

En los rangos HE y VHE encontramos eventos como púlsares, remanentes de supernova y núcleos galácticos activos.

Y en UHE procesos extremos como agujeros negros supermasivos y materia oscura, y hasta fenomenos aún no identificados.

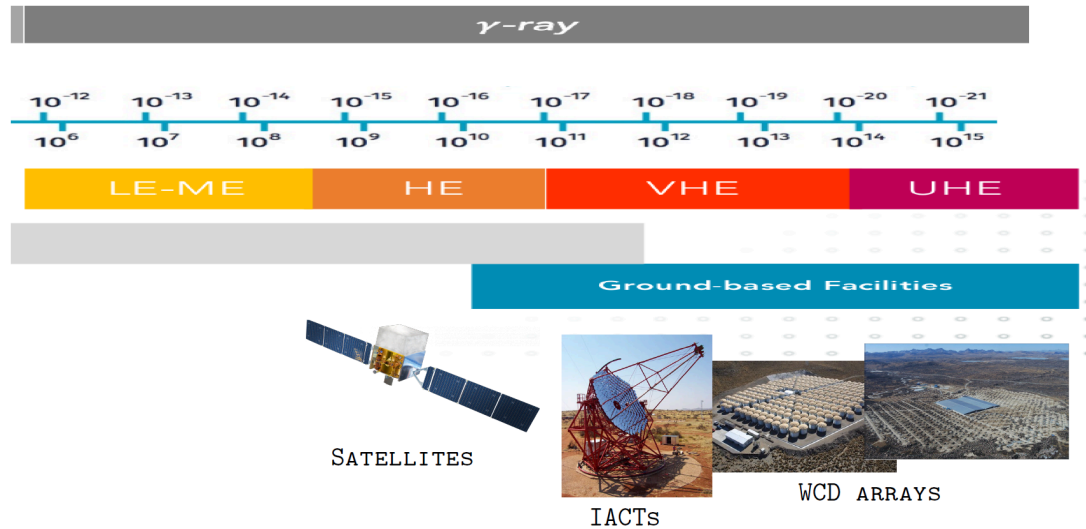


Imagen 2

En la gráfica anterior visualizamos los rangos en que operan distintos observatorios, ya sean espaciales como terrestres, en donde por ejemplo el Fermi idealmente detecta rayos gamma en energías medias y altas, mientras que los Imaging Atmospheric Cherenkov Telescopes (IACT's) y los Water Cherenkov Detectors arrays (WCD) exploran rangos de energía mucho más altas, entre los 50 GeV a 100 TeV y 100 GeV respectivamente.

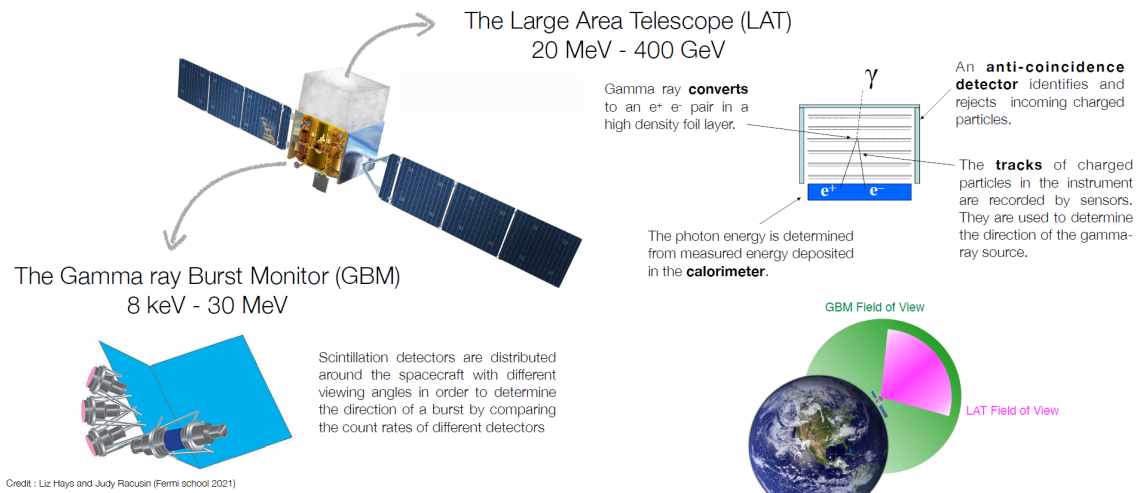


Imagen 3

El telescopio espacial Fermi Gamma-ray Space Telescope captura rayos gamma utilizando dos instrumentos principales: el LAT (Large Area Telescope) y el GBM (Gamma-ray Burst Monitor). El LAT detecta rayos gamma en el rango de 20 MeV a >300 GeV, convirtiendo los rayos gamma en pares electrón-positrón mediante láminas de tungsteno; sus trayectorias se rastrean con un detector de silicio y la energía se mide en un calorímetro. Este proceso permite reconstruir la dirección y energía del rayo gamma, mientras un sistema anticoincidencia filtra el ruido de fondo. El GBM, diseñado para detectar estallidos de rayos gamma (GRBs) en energías de 8 keV a 40 MeV, utiliza detectores de centelleo que convierten los rayos gamma en luz visible, la cual es procesada para determinar energía y tiempo de llegada. Ambos instrumentos transmiten datos a la Tierra, donde se procesan para identificar fuentes, espectros y eventos transitorios como GRBs, publicándose en catálogos como el 4FGL. Fermi escanea continuamente el cielo, ofreciendo un amplio rango energético y capacidad para detectar tanto fuentes persistentes como fenómenos transitorios, siendo una herramienta clave para explorar el universo extremo. (Wood et al., 2017)3

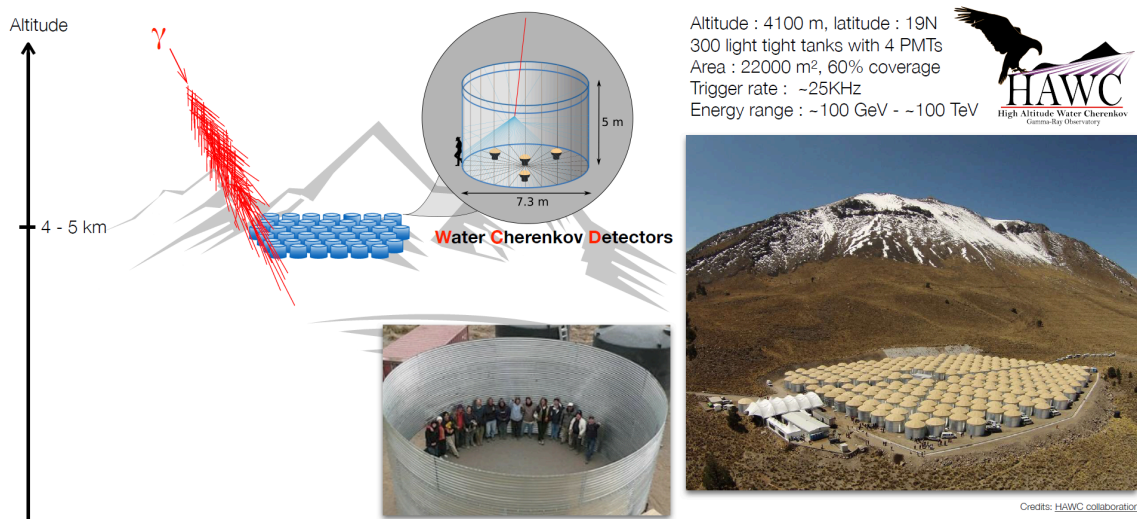


Imagen 4

Los Water Cherenkov Detectors (WCD) capturan rayos gamma de manera indirecta al observar la radiación Cherenkov generada por partículas secundarias creadas cuando un rayo gamma interactúa con la atmósfera terrestre, produciendo una cascada de partículas cargadas. Estas partículas atraviesan tanques de agua purificada, donde generan radiación Cherenkov si viajan más rápido que la luz en el agua, emitiendo un destello de luz azulada. Fotomultiplicadores (PMTs) dentro de los tanques detectan esta luz y la convierten en señales eléctricas, que contienen información sobre la energía, dirección y tiempo de llegada del evento. Con los datos de múltiples tanques, se reconstruyen las propiedades del rayo gamma original, como su dirección y energía, permitiendo identificar su fuente astrofísica. Este método, utilizado en observatorios como HAWC en México y LHAASO en China, ofrece monitoreo continuo del cielo, ideal para detectar rayos gamma de alta y ultra alta energía y estudiar fenómenos transitorios o fuentes puntuales.

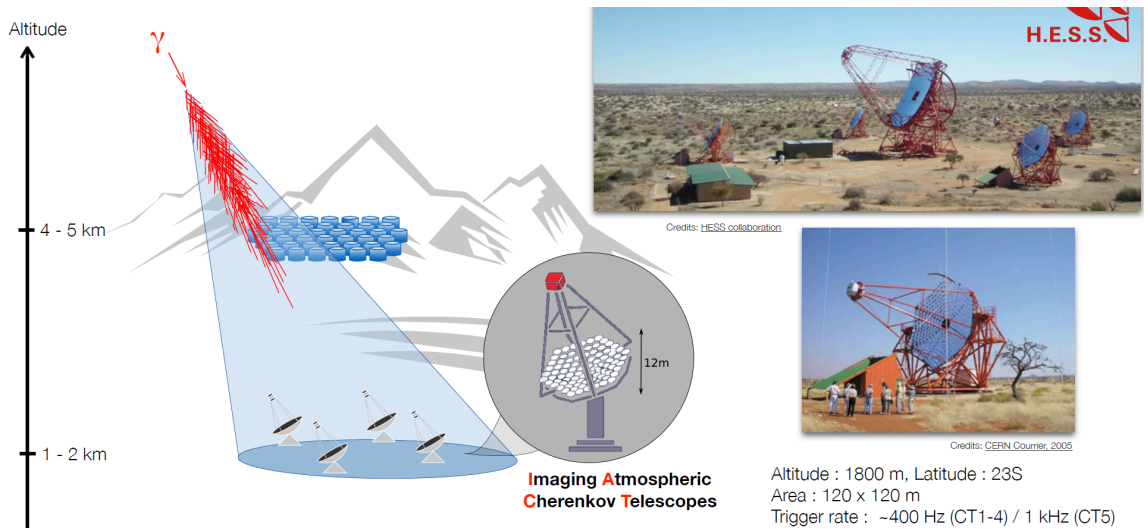


Imagen 5

Los Imaging Atmospheric Cherenkov Telescopes (IACTs) detectan rayos gamma de manera indirecta al observar la radiación Cherenkov generada en la atmósfera terrestre. Cuando un rayo gamma de alta energía interactúa con un núcleo en la atmósfera, produce una cascada de partículas cargadas (electrones, positrones y otros), que viajan a velocidades superiores a la velocidad de la luz en el aire, emitiendo un destello de luz azulada y breve, llamado radiación Cherenkov. Los IACTs, equipados con grandes espejos parabólicos, colectan esta luz y la dirigen hacia cámaras sensibles (fotomultiplicadores o detectores avanzados), que registran imágenes de las cascadas de partículas. Estas imágenes contienen información sobre la dirección, energía y características del rayo gamma original. Al combinar los datos de múltiples telescopios en una matriz (como HESS, MAGIC o VERITAS), se mejora la precisión en la reconstrucción de las propiedades del rayo gamma, permitiendo identificar su origen astrofísico y estudiar fuentes de alta energía, como púlsares, remanentes de supernova y núcleos galácticos activos.

Cabe mencionar que proyecto en curso SWGO operará de forma similar a los observatorios como HAWC en México y LHAASO en China, donde se emplean estanques de agua equipados con sensores para detectar la radiación Cherenkov producida por partículas secundarias generadas cuando los rayos gamma interactúan con la atmósfera terrestre. Por lo tanto, SWGO se asemeja a los WCD arrays (Water Cherenkov Detector arrays) en su diseño y método de detección. (Conceição et al., 2023)⁴

3.2. Metodología

Como marco de trabajo utilizamos el Modelo anidado de Munzner ya que ofrece un marco sistemático para diseñar, evaluar y justificar visualizaciones efectivas especialmente cuando se trata de datos complejos como los astronómicos.

Con esta metodología buscamos una mayor empatía en los intereses de los usuarios finales y entender sus necesidades.

Este modelo consta de cuatro niveles y tres preguntas, las cuales facilitan el análisis:

- Caracterización del problema de dominio y los datos

- ¿Quiénes son los usuarios objetivo?

Estudiantes de posgrado y astrónomos interesados en el análisis de datos de rayos gamma que busquen abordar necesidades de investigación emergentes en Chile como el proyecto SWGO.

- Abstracción en operaciones y tipos de datos

- Traducir de los elementos específicos del dominio al vocabulario de visualización

- ¿Qué se está mostrando? data abstraction

Un artículo digital que sirve como guía capacitadora y motivacional para que eventuales profesionales se sumergan en la manipulación, entendimiento y análisis de los datos a los que podrían enfrentarse trabajando en proyectos como el SWGO en Chile. La idea es esbozar cómo se manipulan y muestran datos de eventos de alta energía capturados por telescopios de rayos gamma, incluyendo ubicaciones espaciales, espectros de energía y series temporales, así como también visualizar la distribución de las fuentes de rayos gamma, observar la intensidad y variación de las emisiones en distintos rangos y momentos. Todo a través del mejor proxy que tenemos, como son los datos del Fermi-Lat.

- ¿Por qué el usuario está mirando esto? task abstraction

El usuario a través de este artículo tendrá un primer encuentro cercano y amigable en el entendimiento de datos de alta energía que lo van a preparar para desenvolverse en la industria investigativa de fuentes de alta energía como la de los rayos gamma. Aprenderá a identificar patrones espaciales y energéticos en las emisiones de rayos gamma, validar teorías astrofísicas y desarrollar habilidades en el análisis de datos complejos.

- Diseño de codificación visual e interacción
 - ¿Cómo se está mostrando?

Se eligen representaciones visuales que faciliten la interpretación de datos complejos para estudiantes y

profesionales, como gráficos de series temporales para analizar cambios en la emisión a lo largo del tiempo. Además, se consideran representaciones 3D y modelos animados que faciliten el entendimiento de la disciplina. Existen herramientas como Gammapy y DS9 para una exploración interactiva y facilitar el aprendizaje.

- **Diseño de algoritmos**
 - **Computación eficiente**

Se implementan herramientas y bibliotecas como Gammapy, AstroML y Astropy, pero principalmente Gammapy para el análisis de datos de rayos gamma, permitiendo el procesamiento y visualización interactiva. Se busca que el usuario pueda observar ajustes de parámetros, aplicar filtros y experimentar con los datos, facilitando una comprensión práctica y global de la reducción de distintos procesos dentro del estudio de los rayos Gamma, como la reducción de ruido, la clasificación de fuentes y el análisis de fenómenos de alta energía.

4. Resultados

4.1. Funciones de Respuesta del Instrumento IRF (Por Carlos Bustamante)

OBJETIVO

El objetivo principal de esta sección es dar solución al problema detectado tanto desde la investigación como desde las encuestas, donde se reveló que actualmente hay una necesidad de información en cuanto a los profesionales de la astronomía y el estudio de los rayos gamma.

Es por esto que en el presente documento, “se cuenta una historia”, la cual parte desde las IRF, es por eso que el primer objetivo específico de esta sección es entregar una visión general de los rayos gamma, para luego proporcionar una explicación detallada sobre qué son las Instrument Response Functions (IRFs), posteriormente comentar su relevancia en el análisis de datos de rayos gamma y finalmente aplicar esto en estudios astrofísicos avanzados reales, en este caso el del observatorio CTA North.

Las funciones IRF desempeñan un papel fundamental en la caracterización del rendimiento de los instrumentos de detección, permitiendo interpretar de manera precisa los datos observacionales.

En este documento, se explora el uso práctico de las IRFs mediante la librería Gammapy, una herramienta de software abierta y ampliamente utilizada en la comunidad científica para el análisis de astronomía de rayos gamma. Se presentarán ejemplos aplicados que abarcan desde la carga y visualización de IRFs hasta su implementación en modelos analíticos, brindando así un enfoque práctico para entender su impacto en la reconstrucción de eventos astrofísicos.

RAYOS GAMMA

Solemos pensar en la luz como la radiación que nos permite ver, pero esta luz, conocida como visible, es sólo una pequeña parte de un fenómeno mucho más amplio. La luz está compuesta por ondas electromagnéticas cuya energía depende de su longitud de onda: a menor longitud, mayor energía.

La luz visible abarca longitudes de onda entre 0.4 y 0.7 micras, donde los distintos colores corresponden a rangos específicos dentro de este intervalo. Más allá de 0.7 micras y hasta 0.1 milímetros, encontramos la luz infrarroja, imperceptible al ojo humano. Por encima de 1 milímetro se ubican las ondas de radio, mientras que en el extremo opuesto, con longitudes menores a 0.4 micras, están la luz ultravioleta, los rayos X y, con las longitudes de onda más cortas y energéticas, los rayos gamma. (Pasaje a la Ciencia, 2021)⁵

Los rayos gamma son la forma más energética de radiación electromagnética, generada en eventos cósmicos violentos como explosiones de supernova y formación de estrellas. Al viajar sin desviarse por campos magnéticos, pueden aportar información sobre sus fuentes de origen, aunque son absorbidos por la atmósfera terrestre, que actúa como escudo protector.

La detección de rayos gamma comenzó en 1961 con satélites, y avances posteriores como el Compton Gamma-Ray Observatory y el telescopio Fermi/LAT permitieron mapear el Universo en rayos gamma, identificando numerosas fuentes conocidas y desconocidas.

Para observar fotones de mayor energía, se desarrollaron técnicas terrestres basadas en la radiación Cherenkov, producida cuando partículas generadas por rayos gamma interactúan en la atmósfera. Esta luz azulada permite detectar eventos energéticos aprovechando la atmósfera como parte del

detector, superando las limitaciones de tamaño de los telescopios satelitales. (Orellana, s.f.)⁶

Los rayos gamma son clave para estudiar los procesos más energéticos y extremos del universo, como agujeros negros, supernovas y el centro de galaxias. Su observación también permite investigar fenómenos fundamentales como la materia oscura, proporcionando pistas sobre su naturaleza.

En este contexto, el Southern Wide-field Gamma-ray Observatory (SWGGO), que se construirá en el Parque Astronómico Atacama, posicionará a Chile como un centro global para la astronomía gamma. Equipado con más de 3.000 estanques de agua a gran altitud, el observatorio podrá detectar partículas generadas por rayos gamma, ampliando nuestra capacidad para explorar el universo.

Este proyecto, liderado en Chile por Claudio Dib, representa un avance científico crucial, combinando el estudio del cosmos con el desarrollo de tecnologías avanzadas y formando nuevas generaciones de expertos en astropartículas, electrónica e inteligencia artificial. Programado para estar operativo en 2030, el SWGGO consolidará a América Latina como protagonista en la investigación de alta energía. (Universidad Técnica Federico Santa María, 2024) ⁷

La astronomía de rayos gamma utilizando telescopios Cherenkov tuvo sus inicios en 1968 con la construcción del primer telescopio de este tipo, denominado WHIPPLE. Sin embargo, no fue hasta 1989 cuando se logró detectar la primera fuente astronómica de rayos gamma, la Nebulosa del Cangrejo. Desde entonces, esta área de estudio ha experimentado un avance significativo.

Actualmente, existen cuatro observatorios operativos en el mundo equipados con telescopios Cherenkov de segunda generación, y ya se están desarrollando instrumentos de tercera generación. El número de fuentes

detectadas supera las 80, y el catálogo continúa expandiéndose de manera constante.

El estudio de los rayos gamma hoy en día permite explorar fenómenos cósmicos extremos y abordar preguntas fundamentales sobre el origen de los rayos cósmicos, la naturaleza de la materia oscura y las propiedades del universo primitivo.

- Supernovas: Estas explosiones estelares liberan enormes cantidades de energía, generando rayos gamma. Analizar su origen puede revelar si los rayos cósmicos provienen de colisiones de protones, un misterio sin resolver desde hace un siglo. Telescopios como MAGIC ya han detectado rayos gamma en remanentes de supernovas como la Nebulosa del Cangrejo.

- Púlsares: Restos colapsados de supernovas, estas estrellas de neutrones giran rápidamente, emitiendo radiación en todo el espectro, incluidos los rayos gamma. Su observación proporciona datos únicos sobre condiciones físicas extremas imposibles de replicar en la Tierra.

- Núcleos activos de galaxias: Los agujeros negros supermasivos en el centro de las galaxias generan chorros de partículas a velocidades cercanas a la luz, produciendo rayos gamma. Estos permiten medir la luz de fondo extragaláctica y reconstruir la historia del universo. Observaciones de galaxias lejanas, como 3C279, sugieren que la cantidad de esta luz de fondo es menor de lo esperado, reescribiendo modelos cosmológicos.

- Materia Oscura: Aunque se desconoce su composición, se teoriza que la materia oscura podría producir rayos gamma al colisionar entre sí. Los telescopios Cherenkov buscan estas señales en regiones ricas en materia oscura, como el centro de la Vía Láctea, aunque aún no se han encontrado pruebas concluyentes. (Pasaje a la Ciencia, 2021)⁵

Las IRF Instrument Response Function (IRF) o Función de Respuesta del Instrumento son un conjunto de características que definen el comportamiento de un instrumento, como un telescopio Cherenkov, y que

deben considerarse durante el análisis de datos (una de ellas es, por ejemplo, el área de colección del instrumento).

Estas funciones son fundamentales porque ningún instrumento es perfecto; su precisión depende de diversas limitaciones, ya sean físicas, tecnológicas o ambientales. Las IRFs se encargan de reflejar esas limitaciones, proporcionando modelos que permiten corregir, calibrar y analizar los datos con mayor precisión. No sólo detallan la sensibilidad y resolución del instrumento, sino que también establecen cómo los datos captados se conectan con los eventos reales que se están observando.

IRF

Cuando se recopilan datos experimentales, idealmente se espera que estos reflejen fielmente la realidad sin alteraciones. Sin embargo, en la práctica, todos los instrumentos introducen algún tipo de distorsión en las mediciones. Esta distorsión es lo que se conoce como la respuesta del instrumento. A continuación un ejemplo intuitivo:

Imagine que intenta leer un cartel en la carretera, pero su visión es borrosa.

Lo que sus ojos y cerebro perciben es una imagen distorsionada (borrosa), aunque la imagen real sea nítida.

En este caso, sus ojos y cerebro actúan como el instrumento y la distorsión (imagen borrosa) es el resultado de la respuesta del instrumento.

El objetivo en un experimento es corregir o al menos entender estas distorsiones para interpretar correctamente los datos obtenidos.

Las Funciones de Respuesta del Instrumento son un conjunto de características que definen el comportamiento de un instrumento, como un telescopio Cherenkov, y que deben considerarse durante el análisis de datos (una de ellas es, por ejemplo, el área de colección del instrumento). (Cherenkov Telescope Array Observatory, s.f.)⁸

Siendo más específicos, el uso de las IRF en el estudio de rayos gamma se basa, en grandes rasgos, en describir matemáticamente cómo un instrumento de detección transforma las propiedades reales de un fotón en valores medidos o reconstruidos durante un experimento. (ctools, s.f.)⁹

En general, para los proyectos de Fermi o CTA (North/South), las IRF se conforman en base a tres importantes tópicos:

- Área efectiva (AEFF): describe cuán eficiente es el instrumento para detectar rayos gamma según su energía y dirección. Sirve para indicar qué proporción de los rayos gamma que llegan al telescopio son efectivamente captados.
- Dispersión de energía (EDISP): evalúa la diferencia entre la energía registrada por el telescopio y la energía real del rayo gamma. Este término es clave para entender qué tan bien el instrumento puede medir los espectros de las fuentes astrofísicas.
- Función de dispersión puntual (PSF): refleja la capacidad del telescopio para determinar con precisión la dirección de un rayo gamma. Describe cómo se dispersa o difumina la luz desde una fuente puntual ideal debido a las limitaciones del sistema de detección (instrumentos, óptica, atmósfera, etc.) (Gammapy Development Team, s.f.)¹⁰

De igual forma es posible encontrar IRF's relacionadas a otras materias.

APLICACIÓN DE IRF EN RAYOS GAMMA

Este caso de estudio está basado en los datos del primer Data Challenge realizado por CTA North. Este es un observatorio perteneciente a la organización CTAO, la cual tiene tres grandes objetivos:

1. Entender el origen y el rol de las partículas relativísticas:

Los rayos cósmicos son partículas atómicas comunes, compuestas en un 99% por protones y núcleos de helio, y en un 1% por núcleos más pesados, electrones, muones y otras partículas. Aunque son partículas ordinarias, se destacan por estar aceleradas a energías extremadamente altas, viajando casi a la velocidad de la luz. Su origen sigue siendo un misterio debido a que, al ser partículas cargadas, sus trayectorias se ven desviadas por los campos magnéticos del universo, impidiendo rastrear su fuente. Sin embargo, los rayos gamma, que se generan en las interacciones de los rayos cósmicos, ofrecen una forma indirecta pero precisa de estudiarlos, ya que no están cargados eléctricamente y viajan en línea recta. El CTAO (Cherenkov Telescope Array Observatory), con su amplio rango energético y alta resolución angular, permitirá identificar fuentes potenciales de rayos cósmicos y analizar su papel en la formación estelar y la evolución de galaxias.

2. Explorar entornos extremos:

El CTAO detecta rayos gamma con energías mucho mayores que los rayos X o los gamma observados por instrumentos espaciales, proporcionando información sobre los procesos físicos en los entornos más energéticos del universo. Estos incluyen objetos compactos como agujeros negros, estrellas de neutrones y enanas blancas, que surgen de la explosión de estrellas masivas al final de sus vidas. Sus intensos campos magnéticos, discos de acreción, chorros de partículas relativistas y vientos poderosos los convierten en objetivos clave para la astronomía de rayos gamma. Las

capacidades del CTAO permitirán estudiar con una precisión sin precedentes los mecanismos físicos que operan en estos sistemas extremos.

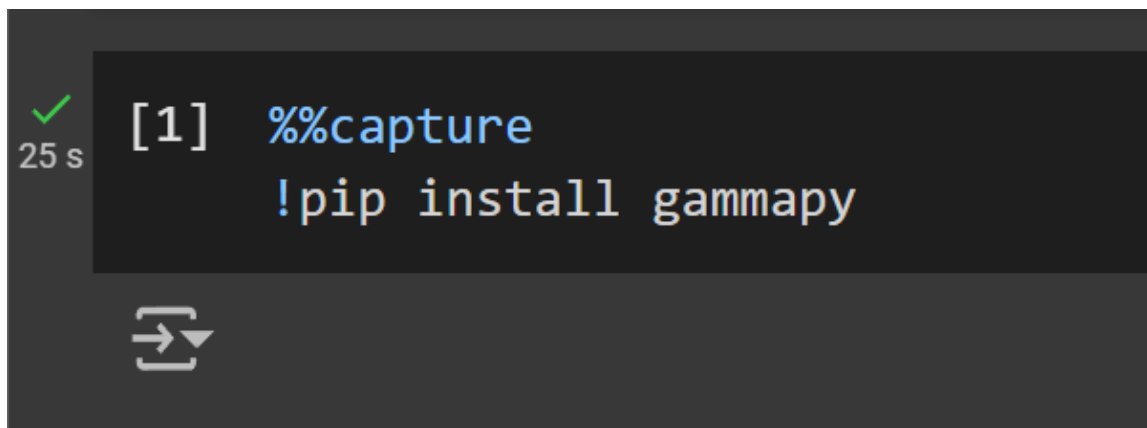
3. Explorar las fronteras de la física:

El CTAO representa un gran avance en sensibilidad y cobertura energética, abriendo la puerta a descubrimientos en física fundamental sobre el comportamiento básico del universo. Su misión incluye investigar la naturaleza y propiedades de la materia oscura, explorar la posible existencia de partículas tipo axión y poner a prueba desviaciones en la Teoría de la Relatividad Especial de Einstein. Cualquiera de estos hallazgos revolucionaría la física de partículas y la cosmología. (Cherenkov Telescope Array Observatory, s.f.)⁸

IRF EN CTA NORTH

En la siguiente sección, se instalan e importan todas las librerías necesarias para ejecutar el análisis y procesamiento de datos relacionados con la astronomía de rayos gamma, puntualmente registros obtenidos del observatorio CTA North (telescopios Cherenkov). Esta guía está basada en el ejemplo desarrollado en la página web de Gammapy "first data challenge". (Gammapy Development Team, s.f.)¹¹

Primero, se asegura la instalación de la librería Gammapy, que es esencial para trabajar con datos de rayos gamma. Esta biblioteca proporciona funciones y módulos específicos para analizar IRFs, como áreas efectivas, funciones de dispersión energética y funciones de dispersión espacial.



```
[1] %%capture
!pip install gammapy
```

Imagen 6

La línea %%capture evita que se impriman mensajes en pantalla al instalar paquetes, manteniendo el entorno de trabajo limpio. Posteriormente se importan todas las librerías necesarias para realizar el análisis.

```
✓ 11 s #Librerías numéricas y de unidades astronómicas
import numpy as np
import astropy.units as u

#Coordenadas y visualización
from astropy.coordinates import SkyCoord
from astropy.visualization import quantity_support

#Gammapy IRF y funciones de soporte
from gammapy.irf import (
    IRF,
    Background3D,
    EffectiveAreaTable2D,
    EnergyDependentMultiGaussPSF,
    EnergyDispersion2D,
)
from gammapy.irf.io import COMMON_IRF_HEADERS, IRF_DL3_HDU_SPECIFICATION
from gammapy.makers.utils import (
    make_edisp_kernel_map,
    make_map_exposure_true_energy,
    make_psf_map,
)

#Mapas y geometría
from gammapy.maps import MapAxis, WcsGeom

#Visualización gráfica
%matplotlib inline
import matplotlib.pyplot as plt

#Bases de datos
import pandas as pd
```

Imagen 7

- NumPy: Proporciona herramientas para cálculos numéricos eficientes.
- Astropy Units: Maneja unidades físicas y astronómicas, asegurando consistencia en los cálculos y conversiones.

- SkyCoord: Facilita la gestión y transformación de coordenadas astronómicas.
 - quantity_support: Mejora la visualización de datos que incluyen unidades físicas en gráficos.
 - Background3D: Modela el fondo tridimensional esperado en función de la energía y posición.
 - EffectiveAreaTable2D: Representa el área efectiva del detector para diferentes energías.
 - EnergyDependentMultiGaussPSF: Modela la función de dispersión puntual (PSF) con múltiples componentes gaussianas.
 - EnergyDispersion2D: Gestiona la dispersión energética en dos dimensiones.
 - COMMON_IRF_HEADERS: Establece encabezados estándar para los archivos IRF.
 - IRF_DL3_HDU_SPECIFICATION: Define especificaciones para el formato de datos DL3 usados en análisis de rayos gamma.
 - make_edisp_kernel_map: Crea un mapa de dispersión energética para corregir mediciones.
 - make_map_exposure_true_energy: Calcula mapas de exposición en función de la energía verdadera.
- make_psf_map: Genera mapas de la función de dispersión puntual (PSF).
- MapAxis: Define ejes para mapas en función de coordenadas espaciales, energías o tiempos.
 - WcsGeom: Configura la geometría de mapas con proyecciones WCS (World Coordinate System).

- Matplotlib: Proporciona herramientas para generar gráficos y visualizaciones de los datos procesados.
- `%matplotlib inline`: Permite mostrar gráficos directamente en el entorno de ejecución (por ejemplo, en Jupyter Notebook).
- Pandas: bases de datos.

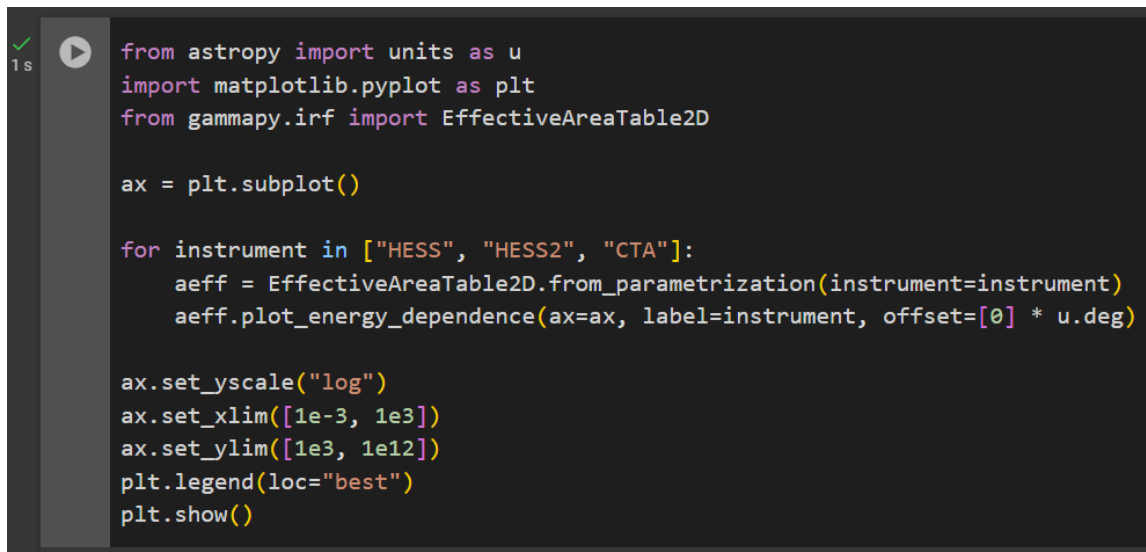
ESTUDIO DEL ÁREA EFECTIVA

El área efectiva se define como la relación entre energía verdadera (true energy) y ángulo de desajuste (ángulo de desvío contra el centro de la cámara). (Donath et al., 2023, p. 6)¹³

La librería Gammapy tiene una tabla con datos para CTA como para otros instrumentos. Estas tablas pueden interpolarse y graficarse. Se puede apreciar a raíz del análisis de estos datos que por ejemplo, los instrumentos seleccionados luego de 10^8 cm² de área efectiva, tienen a converger y no muestran mejoras significativas en su capacidad de percibir energía verdadera.

Los instrumentos HESS y HESS 2 "High Energy Stereoscopic System" son también observatorios de instrumentos Cherenkov, el primero inaugurado en 2003 y el segundo en 2012, ambos en Namibia.

(H.E.S.S. Collaboration, s.f.)¹⁴



```
from astropy import units as u
import matplotlib.pyplot as plt
from gammapy.irf import EffectiveAreaTable2D

ax = plt.subplot()

for instrument in ["HESS", "HESS2", "CTA"]:
    aeff = EffectiveAreaTable2D.from_parametrization(instrument=instrument)
    aeff.plot_energy_dependence(ax=ax, label=instrument, offset=[0] * u.deg)

ax.set_yscale("log")
ax.set_xlim([1e-3, 1e3])
ax.set_ylim([1e3, 1e12])
plt.legend(loc="best")
plt.show()
```

Imagen 8

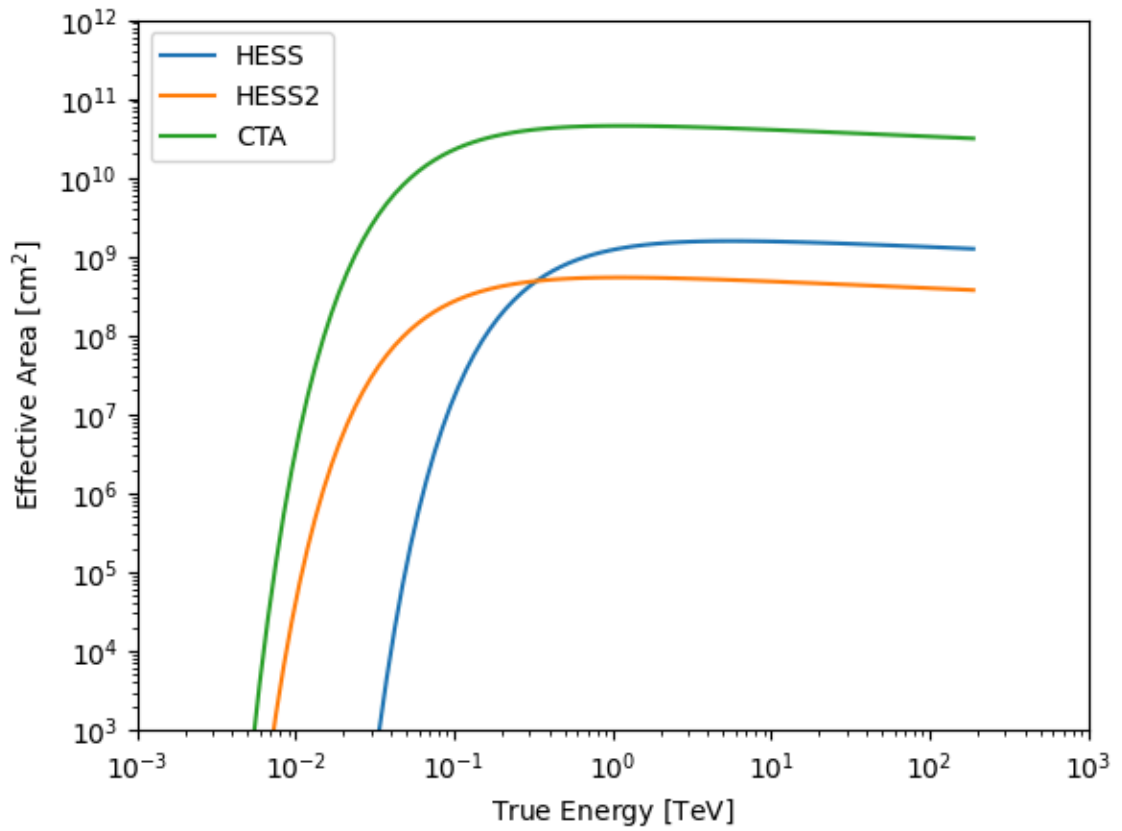


Imagen 9

(Gammapy Development Team, 2024) 10

Respecto a los datos de CTA North, a continuación se muestra un ejemplo de visualización de los datos y también de una interpolación aritmética sencilla.

```
✓ 0s [▶] aeff = EffectiveAreaTable2D.read("irf_file.fits", hdu="EFFECTIVE AREA")
      irf_filename = "irf_file.fits"
      print(aeff)

      ↗ EffectiveAreaTable2D
      -----

      axes : ['energy_true', 'offset']
      shape : (42, 6)
      ndim : 2
      unit : m2
      dtype : >f4
```

Imagen 10

```
✓ 4s [5] # To see the IRF axes binning, eg, offset
      print(aeff.axes["offset"])

      # To get the IRF data
      #print(aeff.data)

      # the aeff is evaluated at a given energy and offset
      print(aeff.evaluate(energy_true=[1, 10] * u.TeV, offset=[0.2, 2.5] * u.deg))

      # The peek method gives a quick look into the IRF
      aeff.peek()

      ↗ MapAxis

      name      : offset
      unit      : 'deg'
      nbins     : 6
      node type : edges
      edges min : 0.0e+00 deg
      edges max : 6.0e+00 deg
      interp    : lin
```

Imagen 11

Finalmente, usando la función peek se pueden obtener visualizaciones de alto valor para entender el comportamiento de las IRF:

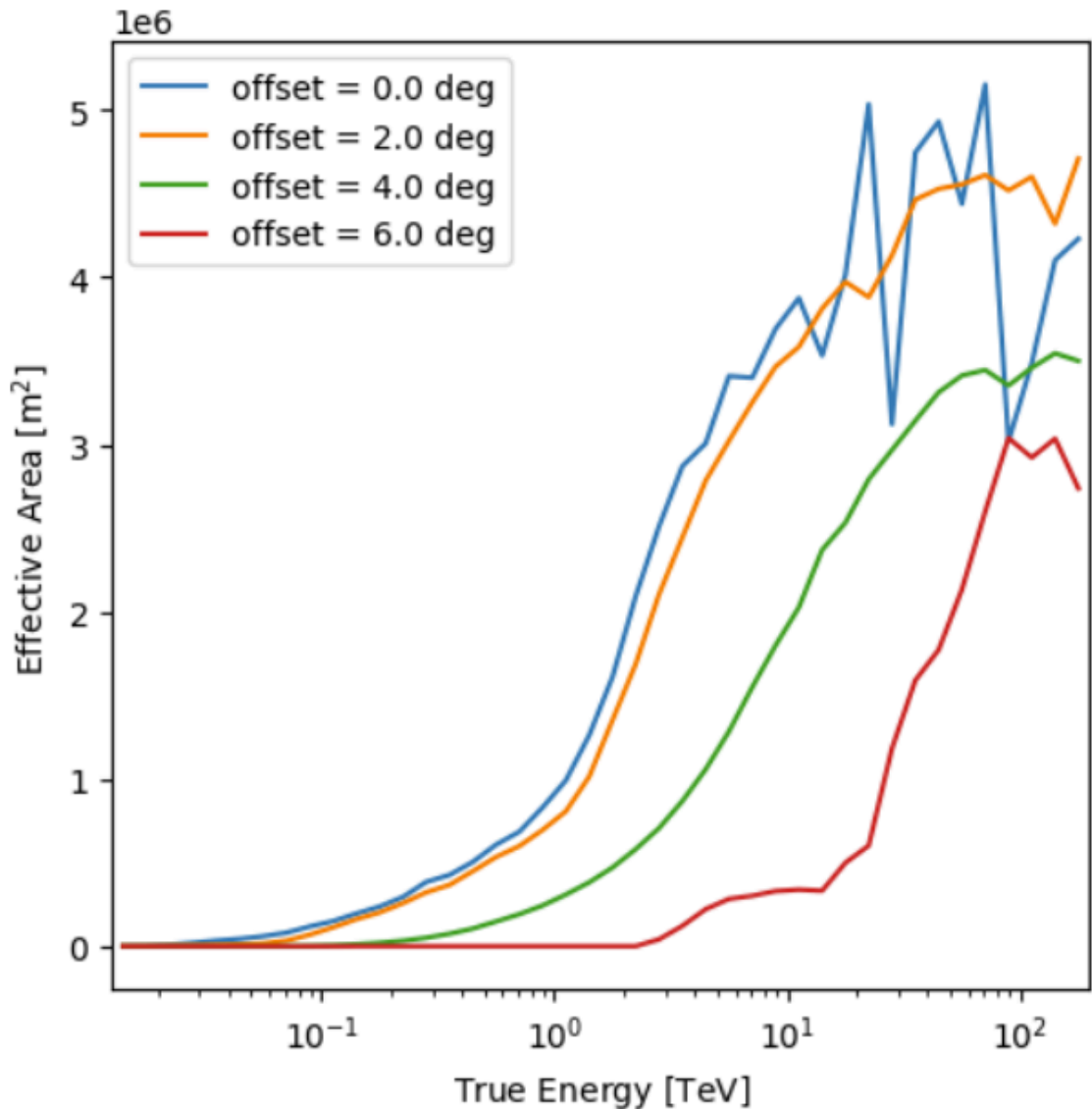


Imagen 12

Respecto a este primer gráfico, es posible concluir que entre mayor es el offset (desajuste respecto a la tangente del centro de la cámara), el área efectiva se reduce. También, que se necesita mayor área efectiva para encontrar registros de True Energy más altos.

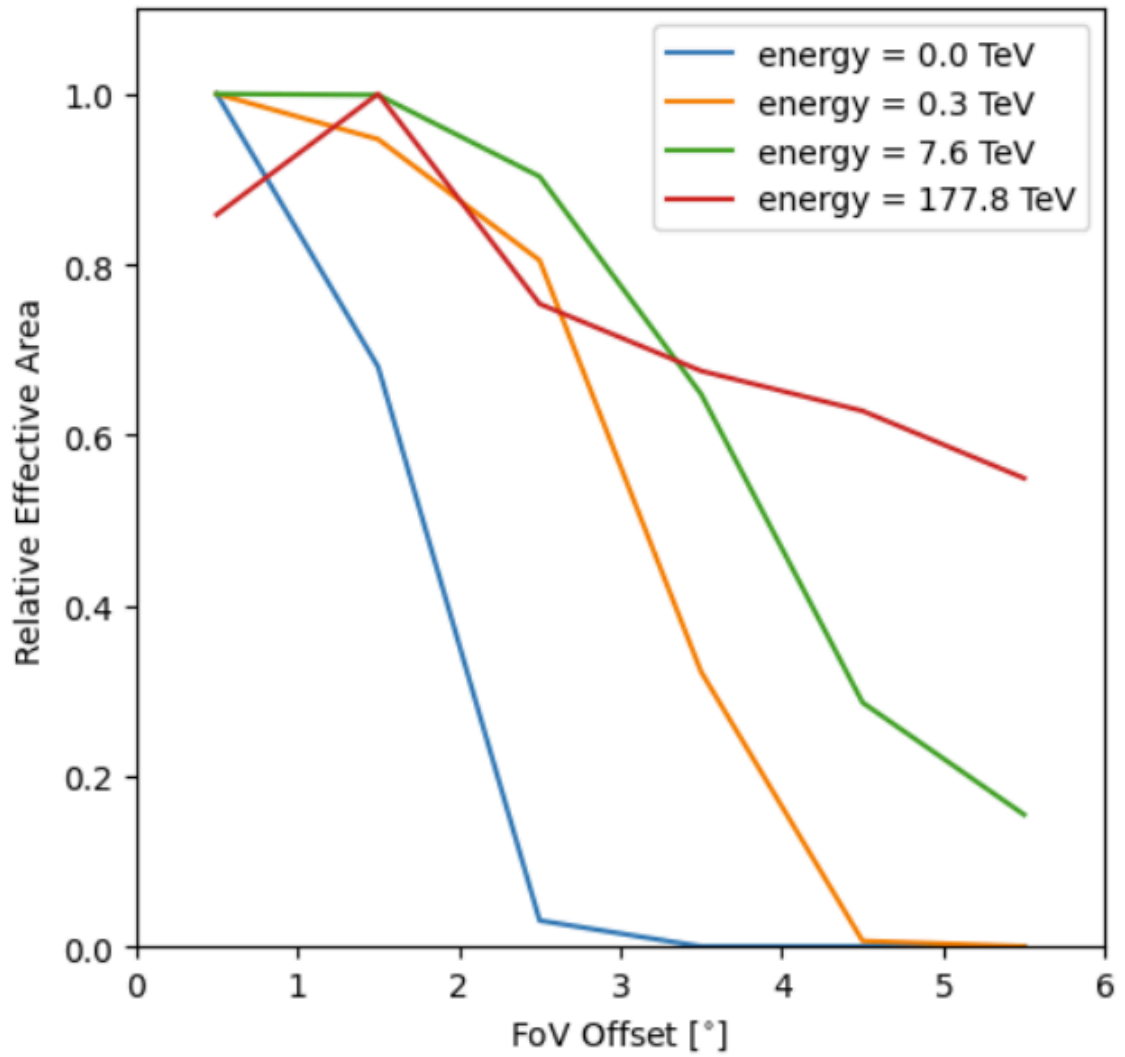


Imagen 13

Este segundo gráfico señala también que altos niveles de desajuste reducen el área efectiva relativa.

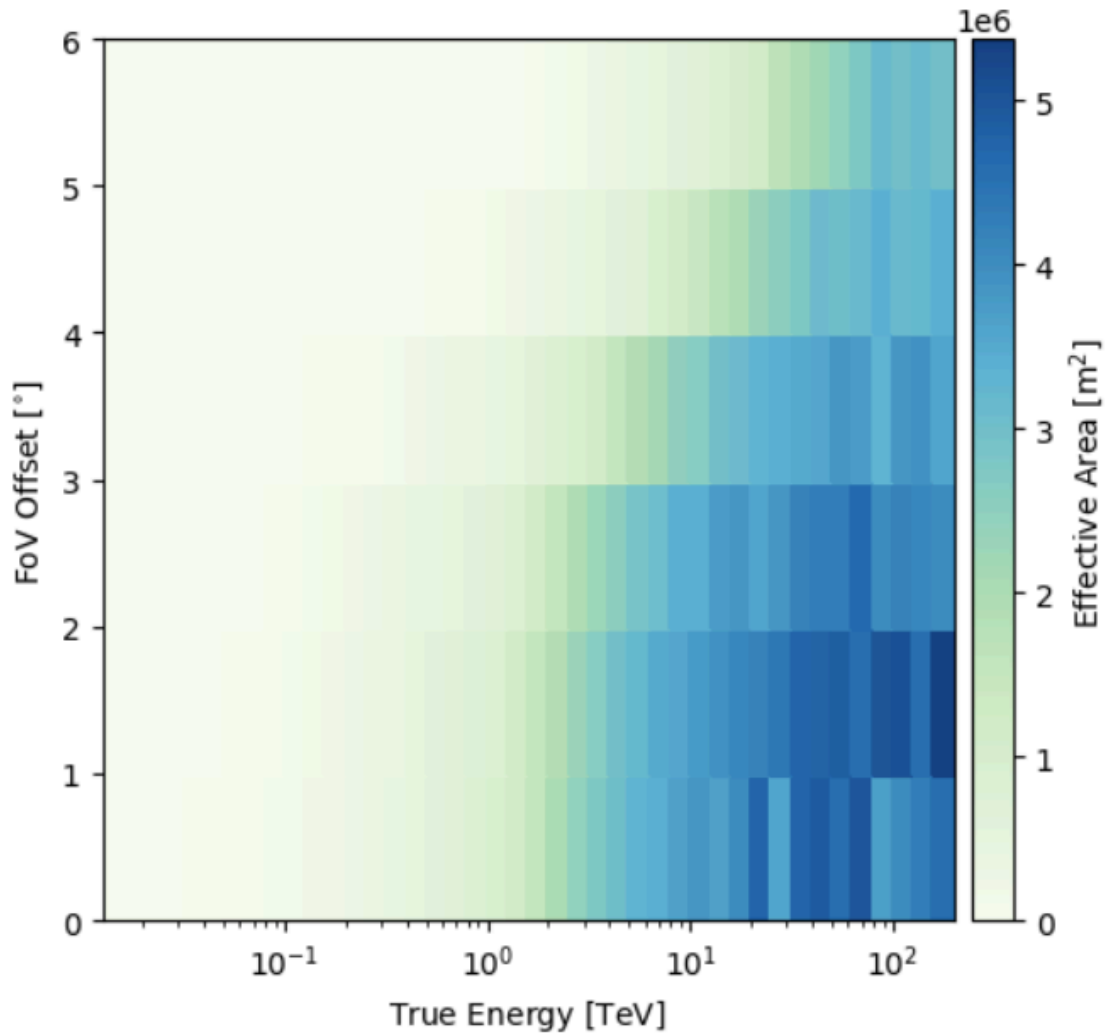
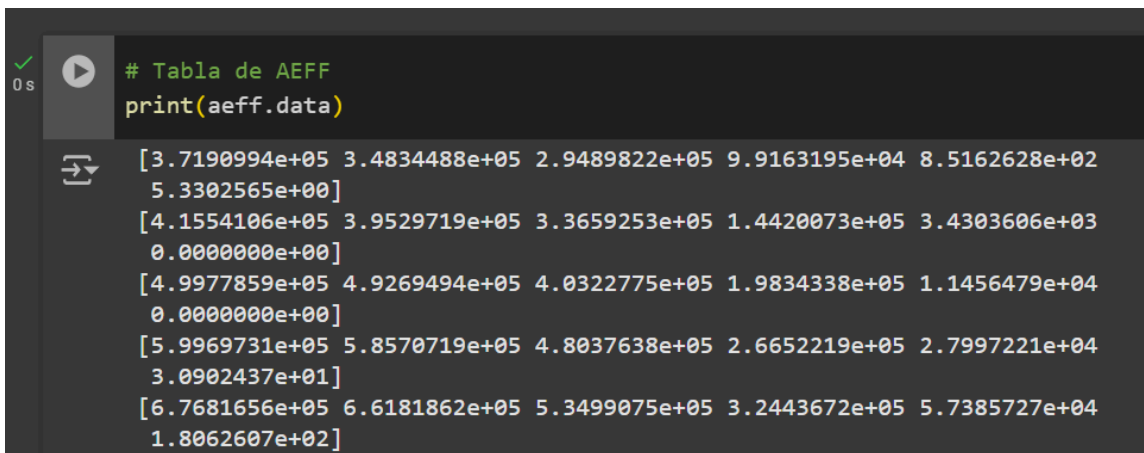


Imagen 14

En la tercera visualización, de tipo mapa de calor, se aprecia que las áreas efectivas más altas tienen bajos desajustes y logran percibir registros de energía más altos.

En conclusión, el desajuste siempre merma la calidad de la medición, aunque si se trata de registros de muy alta energía, pese a que se afecta de manera negativa, lo hace en alguna menor medida.

A continuación se pueden ver los datos a partir de los que se constituyeron los gráficos y posteriormente se realiza una interpolación sencilla para probar la usabilidad de estos. Es muy importante señalar que los datos no se pueden interpolar infinitamente, ya que los datos son válidos o seguros para rangos de energía dados y hoy en día aún no existe un esquema adecuado para definir con exactitud los rangos de validez. (Gamma-ray astronomy community, 2018)¹²



```
0s # Tabla de AEFF
print(aeff.data)

[3.7190994e+05 3.4834488e+05 2.9489822e+05 9.9163195e+04 8.5162628e+02
5.3302565e+00]
[4.1554106e+05 3.9529719e+05 3.3659253e+05 1.4420073e+05 3.4303606e+03
0.0000000e+00]
[4.9977859e+05 4.9269494e+05 4.0322775e+05 1.9834338e+05 1.1456479e+04
0.0000000e+00]
[5.9969731e+05 5.8570719e+05 4.8037638e+05 2.6652219e+05 2.7997221e+04
3.0902437e+01]
[6.7681656e+05 6.6181862e+05 5.3499075e+05 3.2443672e+05 5.7385727e+04
1.8062607e+02]
```

Imagen 15

```

# Assuming 'aeff.data' is a NumPy array with the data from the EffectiveAreaTable2D object
# and has shape (n_energy_bins, n_offset_bins)

# Create a DataFrame from the 'aeff.data'
df = pd.DataFrame(aeff.data)

# Rename the columns
df = df.rename(columns={0: 'a', 1: 'b', 2: 'c', 3: 'd', 4: 'e', 5: 'f'})

df
# Assuming 'df' is the DataFrame from the previous code

# Create a new DataFrame with interpolated values
df_interpolated = df.copy()

# Interpolate 5 times forward (adds 5 new rows)
for _ in range(5):
    # Calculate the difference between consecutive values
    diffs = df_interpolated.iloc[-1] - df_interpolated.iloc[-2]

    # Create the new row by adding the difference to the last row
    new_row = df_interpolated.iloc[-1] + diffs

    # Append the new row to the DataFrame using pd.concat
    df_interpolated = pd.concat([df_interpolated, pd.DataFrame([new_row.values],
                                                                columns=df_interpolated.columns)], ignore_index=True)

# Repeat the process until you have 46 rows
while len(df_interpolated) < 46:
    diffs = df_interpolated.iloc[-1] - df_interpolated.iloc[-2]
    new_row = df_interpolated.iloc[-1] + diffs
    df_interpolated = pd.concat([df_interpolated, pd.DataFrame([new_row.values],
                                                                columns=df_interpolated.columns)], ignore_index=True)

df_interpolated

```

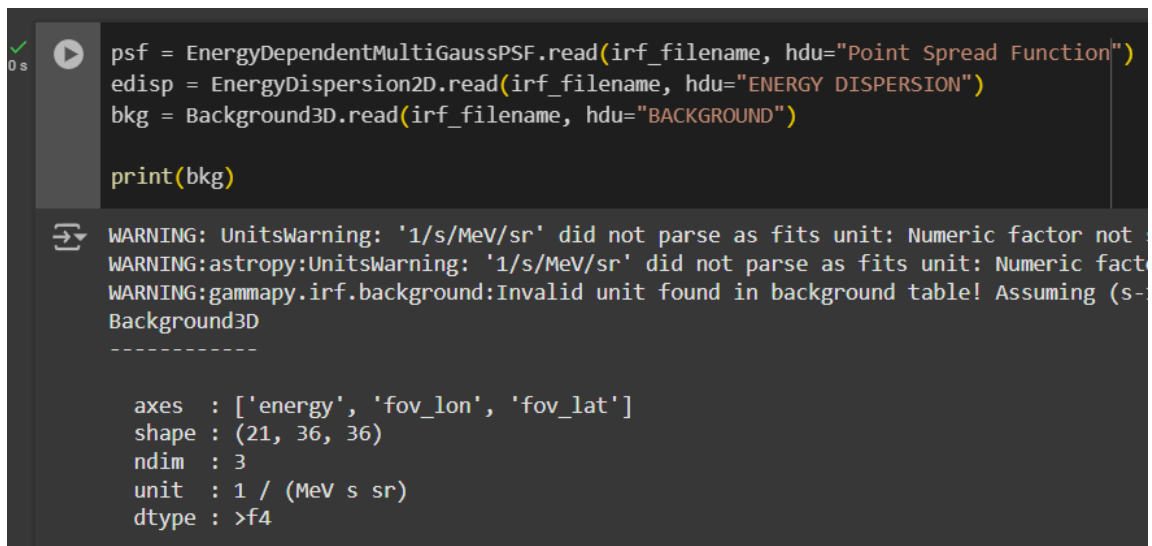
	a	b	c	d	e	f
0	3.939136e+02	2.677224e+02	1.195581e+01	0.000000e+00	0.000000e+00	0.000000e+00
1	2.197657e+03	1.371078e+03	6.601817e+01	0.000000e+00	0.000000e+00	0.000000e+00
2	7.282799e+03	4.493366e+03	1.927749e+02	0.000000e+00	0.000000e+00	0.000000e+00

Imagen 17

Luego de analizar en profundidad la IRF de AEF, es posible pasar a estudiar otras nuevas, en este caso, previamente se habían explicado en qué consistían el Point Spread Function y el Energy Dispersion, por lo tanto se presenta como novedad el Background.

BACKGROUND

La tabla para la función Background está dada en coordenadas FoV, midiendo la energía según un fov_lon y fov_lat. Actualmente esta IRF en Gammapy soporta alineamientos con el sistema de coordenadas horizontal (ALTAZ) y con el ecuatorial (RADEC). (Gammapy Developers, n.d.)¹⁰



```
psf = EnergyDependentMultiGaussPSF.read(irf_filename, hdu="Point Spread Function")
edisp = EnergyDispersion2D.read(irf_filename, hdu="ENERGY DISPERSION")
bkg = Background3D.read(irf_filename, hdu="BACKGROUND")

print(bkg)
```

```
WARNING: UnitsWarning: '1/s/MeV/sr' did not parse as fits unit: Numeric factor not
WARNING:astropy:UnitsWarning: '1/s/MeV/sr' did not parse as fits unit: Numeric fact
WARNING:gammapy.irf.background:Invalid unit found in background table! Assuming (s-
Background3D
-----
axes  : ['energy', 'fov_lon', 'fov_lat']
shape : (21, 36, 36)
ndim  : 3
unit  : 1 / (MeV s sr)
dtype : >f4
```

Imagen 18

A continuación se realizará una interpolación, donde los ceros se remplazan por nulos, esto con el propósito de mejorar la visualización.

```
✓ 0 s # Evaluate background
# Note that evaluate functions support numpy-style broadcasting
energy = [1, 10, 100, 1000] * u.TeV
fov_lon = [1, 2] * u.deg
fov_lat = [1, 2] * u.deg
ev = bkg.evaluate(
    energy=energy.reshape(-1, 1, 1),
    fov_lat=fov_lat.reshape(1, -1, 1),
    fov_lon=fov_lon.reshape(1, 1, -1),
)

bkg.interp_kwargs["fill_value"] = np.nan

ev2 = bkg.evaluate(
    energy=energy.reshape(-1, 1, 1),
    fov_lat=fov_lat.reshape(1, -1, 1),
    fov_lon=fov_lon.reshape(1, 1, -1),
)
print(ev2)
```

```
[[[1.27093776e-04 1.12423250e-04]
  [1.12423250e-04 9.85188496e-05]]

 [[7.67759236e-07 9.01235137e-07]
  [9.01235137e-07 1.05901956e-06]]

 [[5.74465974e-09 5.24685677e-09]
  [5.24685677e-09 4.65357608e-09]]

 [[          nan          nan]
 [          nan          nan]]] 1 / (MeV s sr)
```

Imagen 19

```
7s [▶] print(
    "Interpolation scheme for energy axis is: ",
    bkg.axes["energy"].interp,
    "and for the fov_lon axis is: ",
    bkg.axes["fov_lon"].interp,
)

# Evaluate energy dispersion
ev = edisp.evaluate(energy_true=1 * u.TeV, offset=[0, 1] * u.deg, migra=[1, 1.2])
print(ev)

edisp.peak()

print(psf)

⇒ Interpolation scheme for energy axis is: log and for the fov_lon axis is: lin
[189.80982039  54.28382603]
EnergyDependentMultiGaussPSF
-----

axes      : ['energy_true', 'offset']
shape     : (25, 6)
ndim      : 2
parameters: ['sigma_1', 'sigma_2', 'sigma_3', 'scale', 'ampl_2', 'ampl_3']
```

Imagen 20

A continuación las visualizaciones realizadas con registros de la IRF Background:

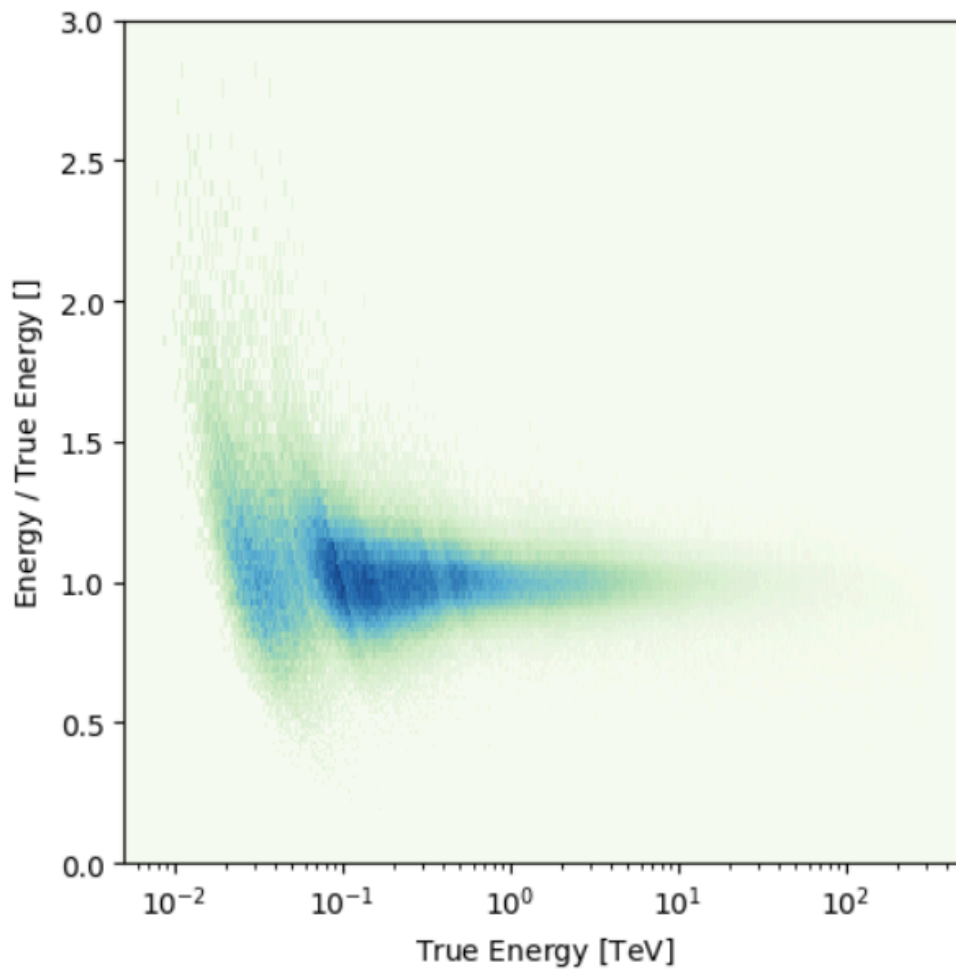


Imagen 21

Este primer gráfico muestra la relación entre el eje y Y de energía reconstruida dividido por energía verdadera, que por lo tanto es una suerte de error o precisión de la reconstrucción. En el eje X está la energía verdadera o real del fotón gamma incidente.

Este gráfico muestra claramente que las regiones más densas, es decir las que son más precisas, están entre $10e-2$ y 10 TeV. Tanto para energías más bajas como más altas que el intervalo antes mencionado estarían sujetas a una menor precisión.

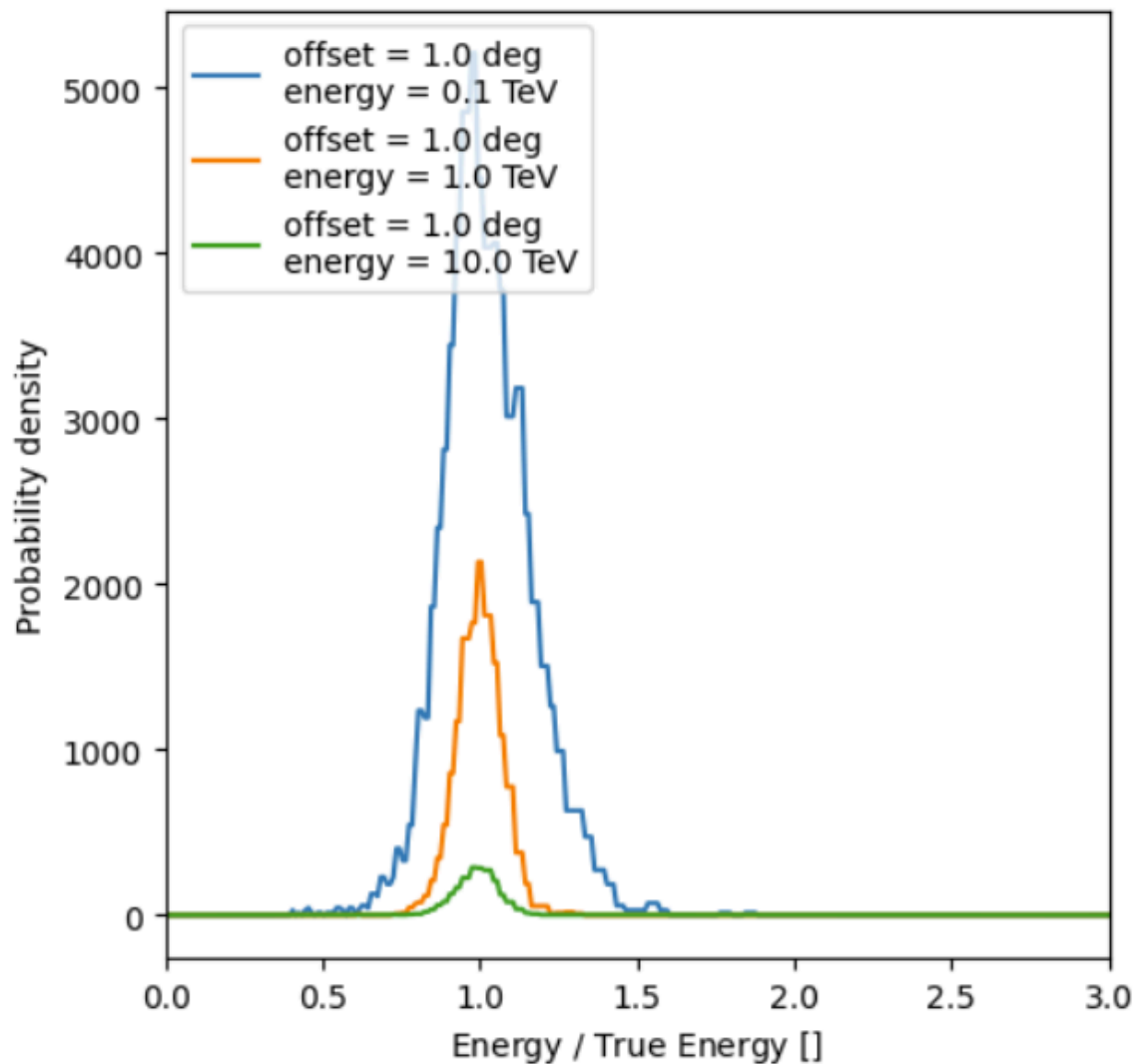


Imagen 22

Este segundo gráfico podría denominarse como histogramas de resolución energética para diferentes energías. En el eje Y está la densidad de probabilidad (frecuencia de la ocurrencia) y en el eje X la proporción entre energía reconstruida y energía verdadera. En cada leyenda se muestran curvas para distintos valores de energía y podría interpretarse que, entre más baja es la energía, mayor es la dispersión, mientras que si la energía es mayor la dispersión descende.

Se puede notar que las curvas están concentradas en 1 (es decir sin error) y que la resolución energética mejora significativamente con energías más altas.

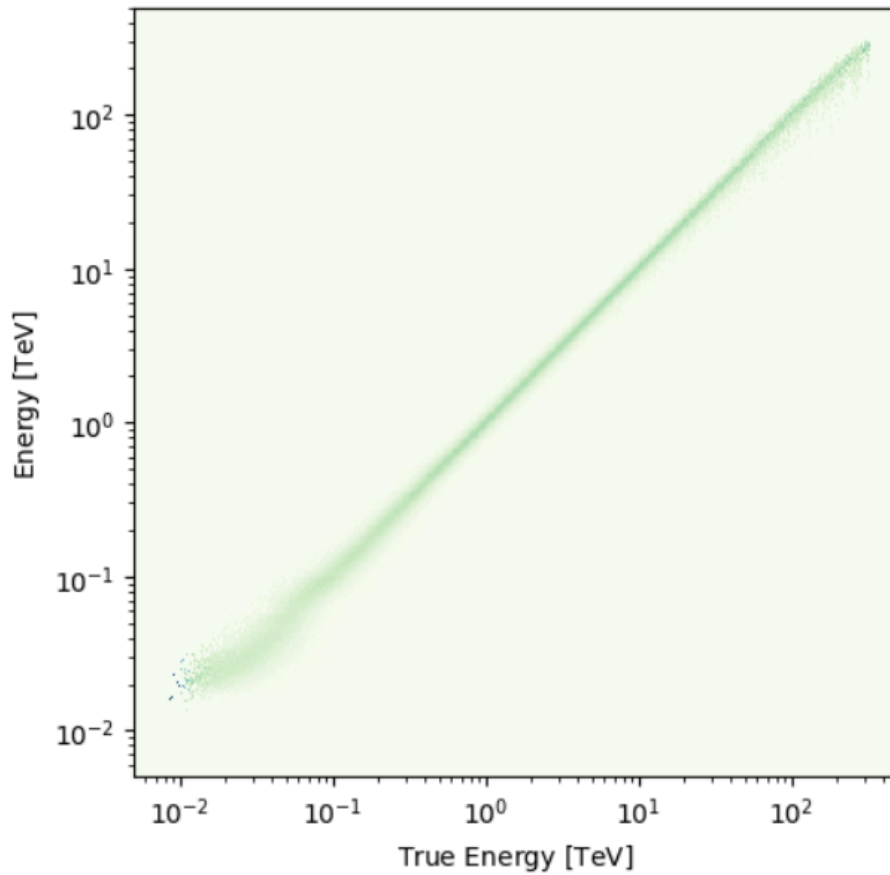


Imagen 23

Este tercer gráfico muestra cómo la energía reconstruida se compara directamente con la verdadera, donde existe un patrón 1:1 desde los $10e-1$ hasta 10 TeV aproximadamente. Antes y después de dicho intervalo la recta se difumina.

SOPORTE DE GAMMAPY PARA IRF ASIMÉTRICAS

En ciertos casos, las IRFs no son simétricas en el campo de visión (Field of View, FOV), lo que significa que el rendimiento del instrumento puede variar en diferentes direcciones.

Aunque Gammapy no tiene clases integradas para admitir funciones de respuesta del instrumento asimétricas (excepto para Background3D), se pueden crear clases personalizadas. Para que estas funcionen correctamente con el MapDatasetMaker, solo se deben permitir variaciones en `fov_lon` (longitud en el campo de visión) y `fov_lat` (latitud en el campo de visión).

La idea principal es que la lista de ejes requeridos debe definirse correctamente en la definición de la clase.

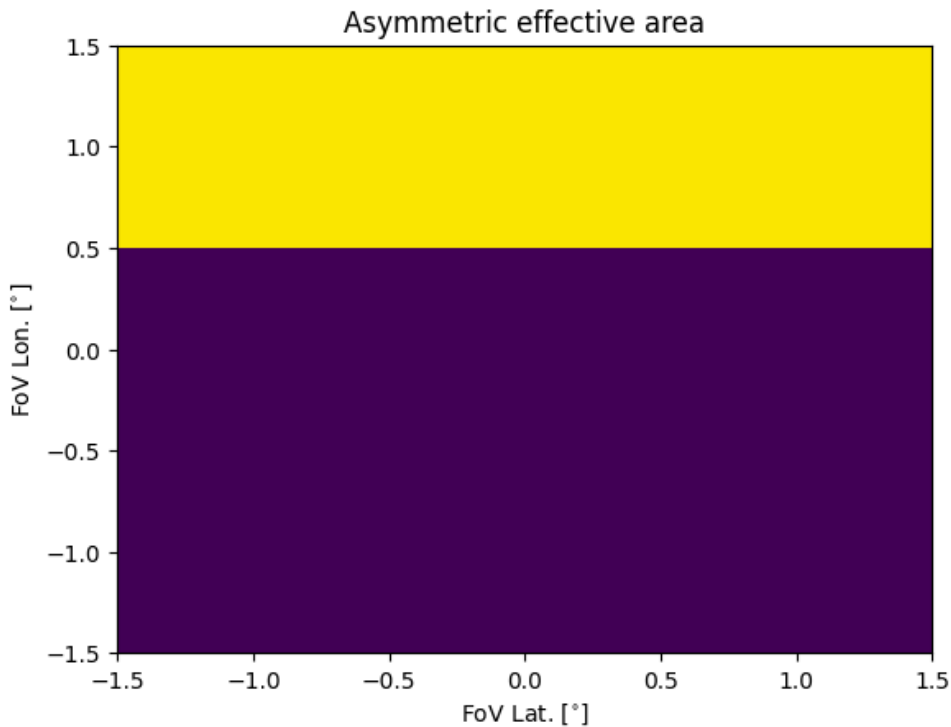


Imagen 24

A raíz del gráfico, se puede concluir que el área efectiva no es uniforme en todo el campo de visión. Esto se refleja en el cambio de color entre las diferentes regiones (amarillo y morado). Se puede alcanzar la máxima sensibilidad en la parte superior del campo de visión, el área efectiva es más alta (amarilla), mientras que la mínima sensibilidad está en la parte inferior donde el área efectiva es significativamente menor (morada). Esto puede deberse a la configuración del telescopio, limitaciones instrumentales o efectos atmosféricos.

MAPAS DE EXPOSICIÓN

El objetivo de los mapas es mostrar cómo varía la exposición acumulada del telescopio en diferentes bandas de energía. Los mapas confirman que el telescopio tiene mayor exposición (y, por ende, sensibilidad) a energías más altas, mientras que su campo de visión está limitado a una región central del cielo. Estos mapas son útiles para planificar observaciones y analizar datos astrofísicos, asegurando que los resultados estén calibrados para la exposición del instrumento.

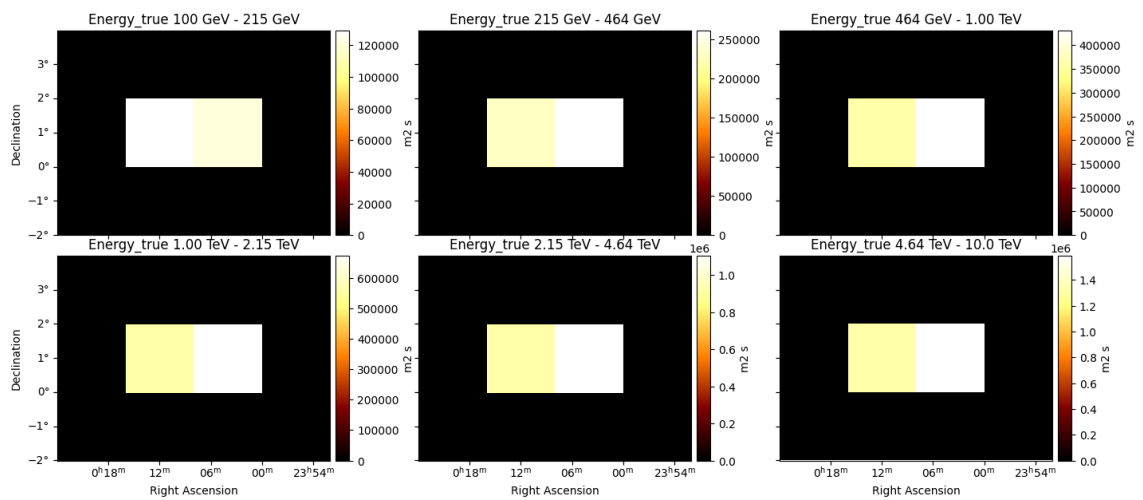


Imagen 25

CONCLUSIÓN

Las Funciones de Respuesta del Instrumento (IRF) son elementos fundamentales en la astronomía de rayos gamma, especialmente en la interpretación de datos obtenidos de telescopios Cherenkov como los del proyecto CTA North. Estas funciones reflejan las limitaciones inherentes de los instrumentos de detección, proporcionando herramientas matemáticas y prácticas para calibrar y analizar los datos observacionales. Dentro de las IRFs más relevantes se encuentran el Área Efectiva (AEFF), la Dispersión de Energía (EDISP) y la Función de Dispersión Puntual (PSF), que juntas caracterizan la capacidad de los telescopios para medir con precisión eventos astrofísicos.

Las capacidades de la librería Gammapy han demostrado ser esenciales para procesar, modelar y visualizar datos de rayos gamma. Herramientas como la interpolación, los mapas de calor y la integración de datos en distintas dimensiones han facilitado la comprensión de los fenómenos subyacentes. Esto incluye la creación de clases personalizadas para representar IRFs asimétricas, evidenciando variaciones en el rendimiento según la región del campo de visión.

Finalmente, las IRFs no solo son críticas para interpretar datos, sino también para avanzar en la comprensión de fenómenos cósmicos extremos y abordar preguntas fundamentales como el origen de los rayos cósmicos, la materia oscura y las condiciones del universo primitivo. En este sentido, proyectos como el CTA, Fermi y el SWGO amplían significativamente nuestra capacidad de explorar el cosmos, posicionando a la astronomía de rayos gamma como un campo clave en la astrofísica moderna. Estas herramientas y metodologías garantizan que los resultados científicos sean precisos, confiables y útiles para enfrentar los desafíos de la investigación de alta energía.

4.2. Predicción y Análisis Temporal de Eventos en Datos del Telescopio Fermi usando Machine Learning (Por Jorge Morgado)

Contexto

En el ámbito de la astronomía y la ciencia de datos, el telescopio espacial Fermi (NASA, n.d.)¹⁵ ha sido una fuente de información fundamental para estudiar algunos de los fenómenos más energéticos del universo, como estallidos de rayos gamma y núcleos galácticos activos. Tal como se analiza en la sección 4.1, se presentan en detalle el funcionamiento y las principales características de este observatorio, reforzando su relevancia en la detección de eventos de alta energía. Este proyecto, cuyo código y detalles de implementación pueden consultarse libremente en [GitHub](#), tiene como objetivo principal guiar a los estudiantes en el proceso de recolección, limpieza y análisis de datos espaciales, tomando en consideración los datos recolectados de entrevistas realizadas a estudiantes (disponibles en Anexo I y II) para proporcionar un acercamiento inicial al análisis con ML. Durante 400 semanas de observaciones, se han reunido 2.6 millones de eventos, cada uno conteniendo información clave como Ascensión Recta (RA), Declinación (DEC) y energía registrada.

Para el almacenamiento optimizado de los datos en formato columna, se eligió el formato Parquet, caracterizado por su eficiencia en la comprensión y lectura selectiva (Apache Software Foundation, n.d.)¹⁶. Esta elección facilita el manejo de grandes volúmenes de datos astronómicos, al reducir el espacio en disco y acelerar las operaciones de entrada y salida. Además, se aplicaron métodos de clustering (KMeans) (Pedregosa et al., 2011)¹⁷ para agrupar los eventos en sectores específicos del cielo, por otro lado modelos de series de tiempo (ARIMA y SARIMAX) (Box & Jenkins, 1976; Seabold & Perktold, 2010)^{18,19} para analizar patrones de variación energética, y

herramientas de detección de anomalías (autoencoders) (Hinton & Salakhutdinov, 2006)²⁰ para identificar episodios atípicos. A lo largo del estudio, también se incluyen análisis estadísticos (ANOVA) (Montgomery, 2020)²¹ que permiten evaluar la significancia de los cambios energéticos a lo largo del tiempo. El propósito de este trabajo es ofrecer un ejemplo práctico que inspire a otros estudiantes a explorar los datos astronómicos con confianza, a diseñar sus propios experimentos analíticos y a formar una visión más amplia de la investigación científica.

El análisis incluye:

1. Cruce y validación de datos espaciales.
2. Modelos predictivos basados en características espaciales y energéticas.
3. Identificación de patrones temporales mediante métodos estadísticos.

Metodología

Si bien es cierto, existen librerías que recomendamos usar que están en el paquete de Fermi Tools, sin embargo estas están asociadas a Conda (Anaconda Inc., n.d.)²², y debido a que en el último año los términos y condiciones de Conda han cambiado, muchos observatorios han dejado de usar Conda, por lo cual el enfoque de esta demostración y todos los análisis asociados a ML han sido realizados sin usar Fermi Tools, de esta manera poder proveer un enfoque diferente al estudiante que se interese en este estudio. El estudio se realizó en múltiples etapas, procesando datos semanales almacenados en formato Parquet. Los pasos clave incluyen:

Obtención del Dataset:

Los dataset se encuentran disponibles en la página de la NASA (HEASARC, n.d.)²³ en formato FTP, de la cual se puede descargar una a una o usando el siguiente código en SHELL: “`wget -m -P . -nH --cut-dirs=4 -np -e robots=off https://heasarc.gsfc.nasa.gov/FTP/fermi/data/lat/weekly/1s_spacecraft/”`, considerar que cada dataset tiene un tamaño aproximado de 80Mb. para trabajar los datos de manera más simple, se realizó una extracción de columnas importantes de cada dataset, filtrando además el sector específico que se decidió trabajar y luego se almacenó en formato parquet, una vez realizado esto para cada dataset, se procede a la unión de estos y la creación del dataframe con pandas para su posterior manejo.

Visualización y Limpieza:

Se descartaron valores nulos en RA, DEC y energía para garantizar la integridad de los datos. Se seleccionó una parte específica del cielo para reducir la carga innecesaria de datos y bajar la demanda de hardware. De este sector se seleccionó la agrupación con mayor eventos a lo largo de las 400 semanas de datos históricos, generando cerca de 2.6 millones de eventos a lo largo de este periodo de tiempo. Los gráficos generados permiten explorar tendencias espaciales y temporales. En un ejemplo, se usó Plotly para crear gráficos interactivos que representan la variación semanal de eventos energéticos en RA y DEC.

Preprocesamiento de Datos:

Los datos fueron agrupados en clusters espaciales mediante KMeans, utilizando un radio de 0.015 grados para la resolución angular. Esto asegura

la correcta identificación de eventos recurrentes en sectores específicos del cielo.

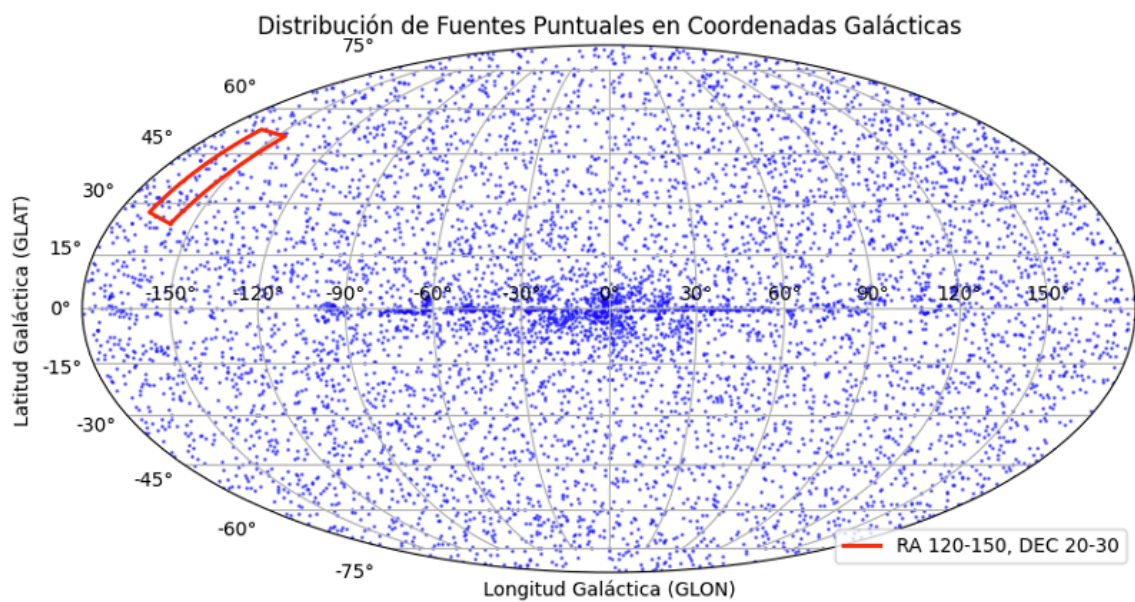


Imagen 26

En la imagen 26 se puede ver el sector seleccionado para trabajar con los eventos de esa zona.

Análisis Estadístico:

El análisis se centró en explorar correlaciones entre características clave como energía y tiempo de vida (lifetime). La matriz de correlación reveló una relación baja entre estas variables, lo que sugiere independencia entre ellas.

Modelos Predictivos:

Usando criterios como el menor valor de AIC (Akaike Information Criterion), se seleccionaron modelos ARIMA para estudiar patrones temporales. El modelo con parámetros (1,0,0) presentó el mejor ajuste con un AIC de 11362.69.

Resultados

El siguiente apartado resume los principales hallazgos en el procesamiento de un extenso dataset astronómico, compuesto por 2.6 millones de eventos recopilados a lo largo de 400 semanas de observación. Se describe el uso de KMeans para agrupar dichos eventos, así como el Análisis de Varianza (ANOVA) para evaluar la significancia estadística de los cambios energéticos. Además, se exponen los resultados de modelos de series de tiempo (ARIMA y SARIMAX) y la detección de anomalías con autoencoders, resaltando la complejidad y la naturaleza intermitente de los fenómenos observados. Finalmente, se destacan las principales correlaciones e independencia de variables, ofreciendo una visión práctica de cómo abordar grandes volúmenes de datos espaciales.

Datos Obtenidos

	ENERGY	RA	DEC	TIME	LIVETIME	WEEK	Cluster	Source_Name
0	362.749146	132.532700	20.052999	2.397752e+08	147.370070	010	0	132.53269958496094_20.05299949645996
1	45.746719	126.277100	20.652399	2.397754e+08	42.294502	010	1	126.277099609375_20.65239906311035
2	176.366364	128.560196	23.716299	2.397756e+08	249.884883	010	2	128.56019592285156_23.716299057006836
3	231.656982	143.024002	22.657801	2.397757e+08	60.290407	010	3	143.0240020751953_22.657800674438477
4	132.843567	132.658295	23.506300	2.397758e+08	163.291120	010	4	132.65829467773438_23.50629997253418
...
2668901	254.614441	133.319397	26.598499	3.970118e+08	31.737353	269	389953	133.31939697265625_26.598499298095703
2668902	210.217377	141.801193	23.884501	3.970118e+08	49.699023	269	198976	141.8011932373047_23.88450050354004
2668903	53.697163	139.590500	28.044300	3.970119e+08	159.471282	269	156169	139.5904998779297_28.044300079345703
2668904	110.054955	147.639999	25.790899	3.970120e+08	16.201075	269	332372	147.63999938964844_25.7908992767334
2668905	37.867615	145.971603	24.245300	3.970120e+08	29.746259	269	781196	145.9716033935547_24.24530029296875

2668906 rows x 8 columns

Tabla 2

En la captura de datos de las 400 semanas, se ha obtenido un dataset de 2.6 millones de eventos los cuales fueron agrupados por kMeans, el data frame construido tiene la estructura demostrada en la tabla 2.

Por otro lado, si se analiza las fuentes de las 400 semanas sin temporalidad, podemos ver que los eventos almacenados cubren todo el espacio designado para el estudio, cabe mencionar que el gráfico en la imagen 27 sólo tiene información de ubicación y energía no de repetibilidad de cada evento, es solo demostrativo para poder observar los eventos en el plano galáctico.

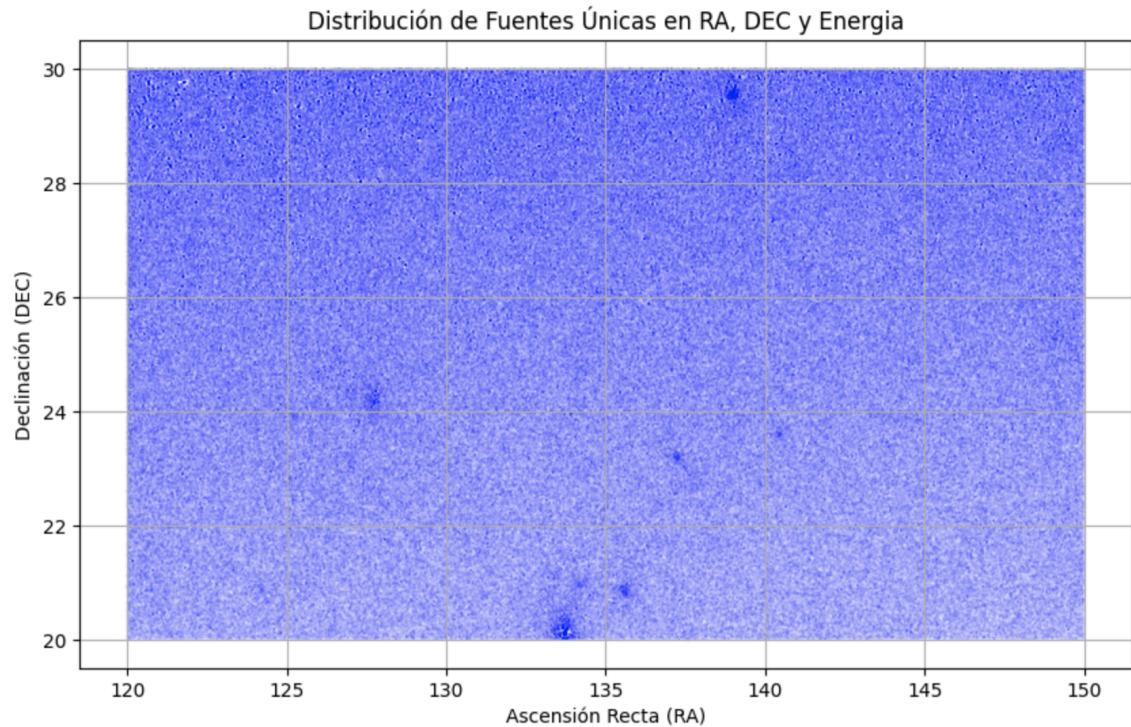


Imagen 27

Análisis Anova

Dada la extensa demanda de recursos para realizar este análisis a las 781197 agrupaciones de eventos que fueron obtenidos luego del clustering por KMeans, se ha hecho el trabajo de análisis en una sola fuente, el cual es escalable a las otras fuentes o a todas, según se tenga el poder de procesamiento. Para este caso se ha decidido usar el grupo con centroide en “133.722900390625_20.10759925842285”, nombre que fue compuesto por la ubicación del centroide en RA y DEC del grupo con una resolución angular de 0.015 grados. En este caso se obtiene que a lo largo de las 400

semanas hay 549 eventos en total siendo el centroide con más eventos en el sector seleccionado para el análisis.

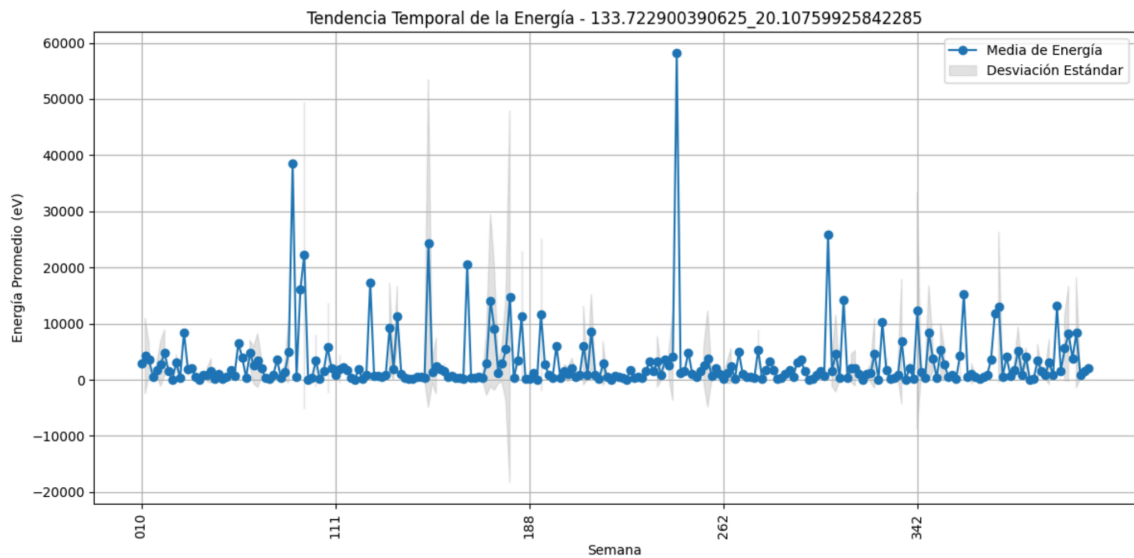


Imagen 28

La gráfica presentada en la imagen 28 ilustra la evolución temporal de la energía promedio registrada el centroide (en adelante “fuente”) con mayor número de ocurrencias, localizada en las coordenadas de Ascensión Recta (RA) 133.7229° y Declinación (DEC) 20.1076° . Este análisis se realiza a lo largo de las 400 semanas de observación y destaca las variaciones en la intensidad energética de la fuente, permitiendo identificar comportamientos característicos en su actividad.

En el eje horizontal se representan las semanas durante las cuales se llevaron a cabo las mediciones. Cada punto de esta escala temporal corresponde a un periodo semanal, permitiendo un seguimiento continuo de los cambios energéticos de la fuente. Por su parte, el eje vertical muestra la

energía promedio medida en cada semana, expresada en electronvoltios (eV). Esta métrica es fundamental para interpretar la intensidad de la emisión de la fuente durante cada periodo de observación.

La gráfica utiliza puntos azules para representar la media de energía semanal. Estos puntos están acompañados por barras verticales, que indican la desviación estándar de la energía en cada semana. Estas barras proporcionan información adicional sobre la variabilidad de las mediciones, permitiendo diferenciar entre semanas con fluctuaciones significativas y semanas de estabilidad.

Al observar el patrón general, se aprecia que durante la mayoría de las semanas la energía promedio se mantiene baja, con valores cercanos a cero y una variabilidad limitada. Sin embargo, en ciertas semanas destacan picos notables en la energía promedio, alcanzando valores significativamente elevados. El más prominente de estos picos supera los 60,000 eV, mientras que otros alcanzan entre 10,000 y 30,000 eV. Estas anomalías sugieren que durante estos periodos ocurrieron eventos astrofísicos relevantes, como destellos o explosiones de energía en la fuente puntual.

Además, se observa que las semanas con los picos más altos también presentan barras de desviación estándar mayores, lo que refleja una dispersión significativa en las mediciones de energía dentro de esos periodos. Esto podría indicar que la fuente experimentó fluctuaciones extremas en su intensidad, posiblemente asociadas a procesos dinámicos o transitorios. Por el contrario, en las semanas donde la energía promedio es baja, la desviación estándar también se reduce, lo que sugiere estabilidad o inactividad relativa de la fuente.

El análisis de estos patrones nos permite inferir que la fuente puntual en cuestión tiene una actividad mayormente estable, interrumpida ocasionalmente por periodos de alta emisión energética. Estos eventos pueden ser indicadores de procesos físicos intensos, como destellos gamma o fenómenos relacionados con la dinámica de objetos compactos (e.g.,

agujeros negros o estrellas de neutrones). Los periodos de baja energía, en cambio, reflejan posibles estados de reposo o emisiones continuas a un nivel basal.

También podemos analizar el mismo comportamiento con gráficos de cajas ilustrado en la imagen 29, los cuales muestran las mismas tendencias, pero a su vez se pueden observar datos que se consideran outliers.

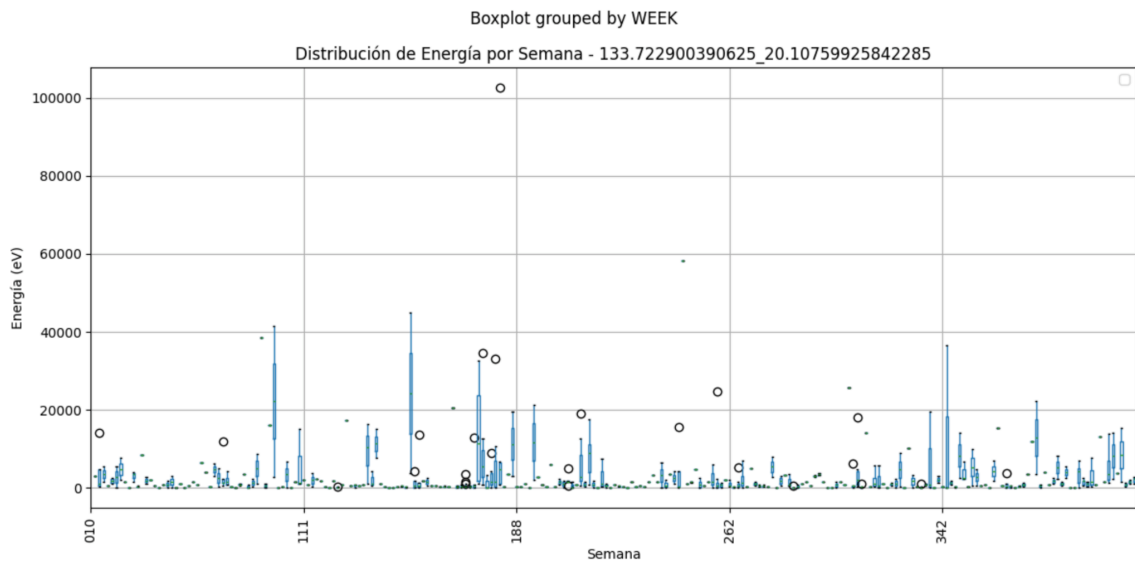


Imagen 29

Los resultados finales del análisis de varianza (ANOVA) realizado sobre los datos de esta fuente puntual arrojó los siguientes resultados:

Estadístico F: 0.9216

Valor p: 0.7470

Estos resultados indican que no hay diferencias estadísticamente significativas en la variabilidad de la energía entre las semanas analizadas. En otras palabras, aunque se observan picos y outliers en ciertas semanas, estos eventos extremos no generan una variación lo suficientemente consistente como para considerarse significativamente diferentes en el contexto general del comportamiento de la fuente.

Análisis Temporal y Modelos SARIMAX

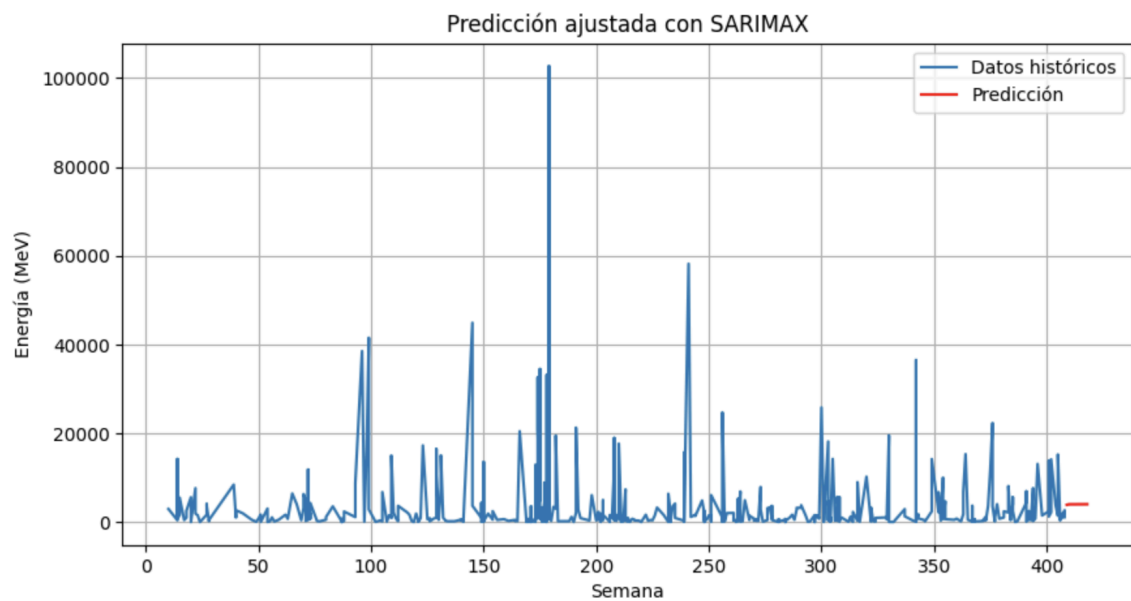


Imagen 30

La gráfica presentada en la imagen 30 muestra el resultado de un modelo de predicción ajustado con SARIMAX aplicado a los datos históricos de energía registrados para una fuente puntual específica. Este análisis abarca

400 semanas de observación, donde se modelan los datos históricos y se proyectan valores futuros basados en las tendencias aprendidas por el modelo. Para este modelo se usaron valores de p, q, d de 1, 0, 0 obteniendo un AIC de 11362.694921978042

Elementos del Gráfico

Eje Horizontal (Semana): Representa las semanas consecutivas de observación, donde los datos históricos abarcan las semanas iniciales y la predicción corresponde al intervalo proyectado al final del periodo.

Eje Vertical (Energía en MeV): Muestra los valores de energía registrados en megaelectronvoltios (MeV) para cada semana observada y proyectada.

Curva Azul: Representa los datos históricos de energía, mostrando las fluctuaciones observadas a lo largo de las 400 semanas.

Curva Roja: Corresponde a los valores predichos por el modelo SARIMAX en el intervalo de proyección. Este tramo ilustra cómo el modelo ajusta sus predicciones basándose en la dinámica de los datos pasados.

El comportamiento histórico muestra picos significativos de energía en semanas específicas, con un valor máximo cercano a los 100,000 MeV, observado alrededor de la semana 200. Estos picos reflejan eventos energéticos transitorios, como destellos gamma o fenómenos astrofísicos altamente energéticos, que ocurren de manera intermitente en la fuente.

En general, la energía promedio de las semanas oscila en un rango mucho más bajo, con variaciones más frecuentes en torno a los 10,000 - 20,000 MeV, lo que indica una actividad basal moderada de la fuente durante gran parte del periodo observado.

La gráfica resalta cómo el modelo SARIMAX ajusta las predicciones basándose en los datos históricos de energía. Si bien logra capturar las tendencias generales, las proyecciones son conservadoras, estabilizándose en valores bajos hacia el final del periodo. Esto puede reflejar las limitaciones inherentes del modelo para predecir eventos extremos en un dataset altamente variable. Este análisis es útil para entender el comportamiento promedio de la fuente, pero podría complementarse con otros enfoques para modelar los picos energéticos más significativos.

Relaciones Entre Energía y Livetime

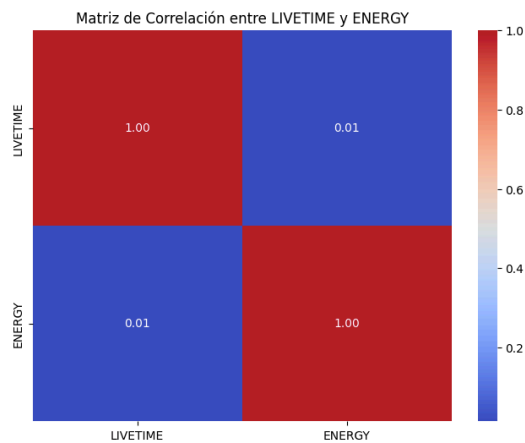


Imagen 31

Un análisis de correlación entre energía y tiempo de vida (livetime) muestra que no existe una relación lineal significativa entre ambas variables. Este resultado fue representado en la imagen 31, además podemos observar en la imagen 32, un gráfico de dispersión que confirma la independencia de estas características en el dataset.

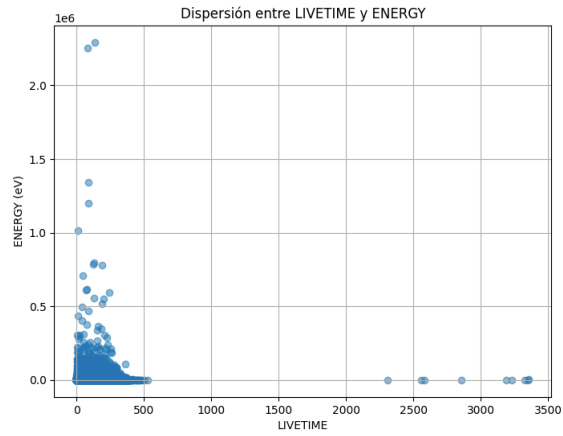


Imagen 32

Estos gráficos nos muestran que no hay ninguna relación entre estas variables, poniendo sobre la posibilidad de que no hay patrones en la detección de fotones.

Análisis de posibles anomalías con Autoencoder

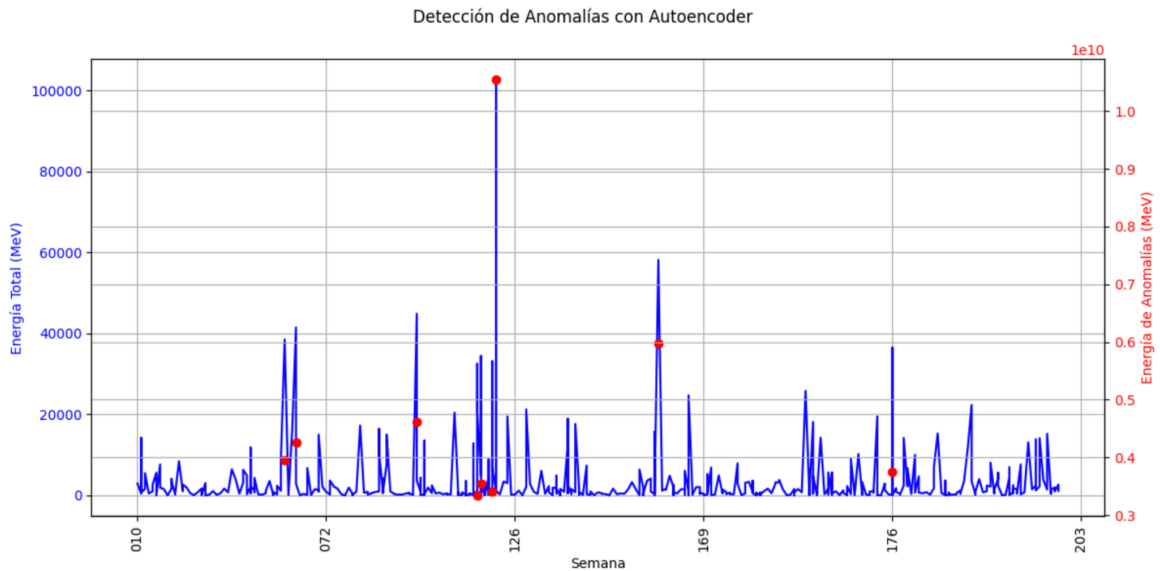


Imagen 33

El gráfico, titulado "Detección de Anomalías con autoencoder" de la imagen 33, representa la evolución temporal de la energía registrada en el dataset. En el eje horizontal se encuentra el tiempo, medido en semanas, mientras que en el eje vertical se muestran los valores de energía en megaelectronvoltios (MeV). La línea azul conecta los valores de energía registrados para cada semana, mostrando una dinámica energética con un alto nivel de variabilidad.

El comportamiento general, como se describió anteriormente en los otros análisis, revela un rango basal de energía en el que los valores permanecen bajos y estables para la mayoría de las semanas observadas. Sin embargo, se destacan varios picos pronunciados que alcanzan niveles extremadamente altos, superando los 10,000 MeV en algunos casos. Estos picos representan eventos de alta energía, posiblemente asociados con fenómenos astrofísicos transitorios o emisiones energéticas anómalas. La densidad de estos picos

disminuye con el tiempo, indicando que estos eventos son aislados y no forman parte de una tendencia sostenida.

En términos de análisis, la curva azul sugiere que la fuente presenta una actividad basal constante con episodios esporádicos de alta energía. Estos episodios podrían corresponder a destellos gamma o explosiones que ocurren debido a procesos energéticos extremos en la fuente puntual estudiada. La resolución semanal del gráfico permite observar los patrones temporales de forma precisa, destacando la irregularidad y la naturaleza impredecible de los picos energéticos.

Los puntos de color rojo del mismo gráfico, analiza los datos desde la perspectiva de las anomalías detectadas en el dataset original. En este caso, se consideran como anomalías aquellos puntos de datos que se desvían significativamente del rango típico de energía observado. Los puntos rojos representan estas anomalías, localizadas en semanas específicas y distribuidas en el tiempo.

En el eje horizontal se encuentran las semanas identificadas como anómalas, mientras que en el eje vertical se presentan los valores de energía correspondientes a estas anomalías. El gráfico indica que las anomalías alcanzan energías de magnitudes mucho mayores que el rango basal, algunas superando los $60 \times 10^8 \text{ MeV}$, lo cual es extremadamente elevado en comparación con los valores de energía estándar del resto de las semanas. La distribución de estas anomalías a lo largo del tiempo muestra que no se agrupan de forma consistente, lo que sugiere que estos eventos no siguen un patrón temporal recurrente.

Desde un punto de vista técnico, la detección de anomalías es crucial para identificar eventos extremos que podrían pasar desapercibidos en el análisis general del dataset. Estos valores atípicos podrían estar relacionados con fenómenos raros o inusuales, como explosiones de alta energía o eventos fuera del comportamiento típico de la fuente puntual. El enfoque independiente para graficar las anomalías resalta la importancia de tratarlas

de manera separada, ya que su magnitud y frecuencia no son representativas del comportamiento global del sistema estudiado.

A fin de cuentas, ambos gráficos proporcionan perspectivas complementarias del comportamiento energético de la fuente. Mientras que el primer gráfico muestra el contexto general de los datos energéticos con un enfoque en su evolución temporal, el segundo se centra en las características excepcionales del dataset, permitiendo un análisis más detallado de los eventos extremos y su posible origen físico. Juntos, estos análisis refuerzan la complejidad y la naturaleza esporádica de los eventos de alta energía en el sistema observado.

Conclusión

A lo largo de este apartado, se han demostrado las etapas fundamentales para manejar grandes volúmenes de datos astronómicos, desde su adquisición hasta la aplicación de modelos de predicción y la identificación de anomalías relevantes. Los datos provienen del repositorio oficial de la NASA (NASA, n.d.), lo que garantiza su alta calidad y relevancia para estudios relacionados con fenómenos de alta energía. Los resultados señalan que, si bien se observan picos de energía aislados en diferentes semanas, no existe una variación significativa que indique un patrón constante a lo largo del tiempo. Asimismo, la correlación entre la energía y el tiempo de vida (liveness) resulta ser prácticamente nula, lo que sugiere independencia entre estas dos variables. Por otro lado, el uso de modelos ARIMA y SARIMAX facilita la comprensión de tendencias generales, aunque los eventos excepcionalmente energéticos permanecen como desafíos para la predicción lineal.

Para los estudiantes interesados, este análisis no solo muestra la importancia de la metodología y la rigurosidad científica, sino que destaca cómo la colaboración interdisciplinaria (entre la astronomía y la ciencia de datos)

puede revelar información valiosa sobre el comportamiento del cosmos. Se invita a revisar el repositorio de datos de NASA (NASA, n.d.)¹⁵ y el código alojado en [GitHub](#), para profundizar en cada paso de la implementación, experimentar con parámetros propios y, en un futuro, extender este estudio a otras fuentes o dominios de interés. Esperamos que este apartado y proyecto en general sirva de inspiración para que más estudiantes se aventuren en la investigación astronómica respaldada por técnicas de análisis de datos, enriqueciendo su formación y contribuyendo al avance del conocimiento científico.

4.3. Análisis con Gammapy (Por Víctor Zúñiga)

Contexto

Los distintos instrumentos ya sean espaciales como el Fermi o terrestres como los IACT's o WCD, tienen distintos rangos de energía de captura de rayos gamma, y de visualización. Sin embargo, comparten la forma o formato de reconstrucción de la data en cuanto variables como ID del evento, energía, dirección, fecha y "gammaness".

(Indico. (2024, diciembre 3-4). Gammapy workshop)24

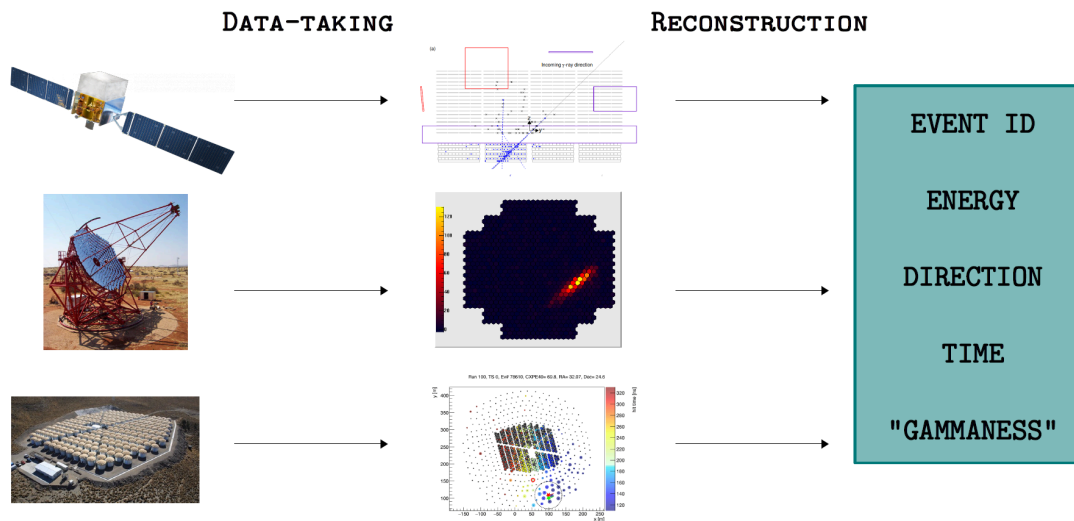


Imagen 34

En la última década, se ha trabajado en definir un formato estándar para almacenar datos astronómicos de rayos gamma, abarcando nombres de columnas, cantidades y otros aspectos, aprendiendo de estándares previos como los de rayos X y Fermi. Esto llevó a la creación de la comunidad de astronomía de rayos gamma y el desarrollo de estándares comunes, documentados en

<https://gamma-astro-data-formats.readthedocs.io/en/latest/general/index.html>. Este esfuerzo permite que toda la data tenga un formato uniforme, facilitando el uso de herramientas compartidas como **Gammapy**, el único paquete de Python dedicado completamente a la astronomía de rayos gamma, aunque no el único disponible.

¿Por qué usar Gammapy?

- En la última década, se ha hecho un gran esfuerzo de definir un formato estándar cuando se trata de almacenar datos astronómicos de rayos gamma.
- En formato nos referimos a nombres de columnas, cantidades, entre otras.
- Como resultado surgió la comunidad de astronomía de rayos gamma:
- <https://gamma-astro-data-formats.readthedocs.io/en/latest/general/index.html>
- Entonces si toda la data luce similar (como mencionamos anteriormente), podamos todos compartir una misma herramienta: "Gammapy".
- Si bien Gammapy no es el único paquete existente, es el único dedicado a rayos gamma como un todo. Dado esto, independientemente del instrumento que capture el evento, podemos usar gammapy como denominador común para estudiar, analizar y visualizar eventos, lo que la convierte en una librería ideal para ingenieros, científicos o investigadores.
- Los archivos de datos astronómicos se almacenan formato FITS (Flexible Image Transport System), que desde los 70 's han sido formato estándar para almacenamiento de estos datos. Gammapy tiene por tanto compatibilidad con este formato

Necesidad del campo educativo

De acuerdo a lo señalado en la encuesta de los mismos estudiantes como el caso de Rebeca Burgos, estudiante de astronomía de último año de la U de C. una de las mayores brechas señala se encuentra en el análisis de datos y programación, la cual se enseña muy a la ligera. Ella afirma que es la práctica finalmente la que termina formando a cada científico, considera que esta debiera fortalecerse mucho más en etapas tempranas de formación. De hecho su mismo docente, el Doctor Matias Jones aduce que “un astrónomo es un programador”. De esta forma propender a hacer más equilibrada la teoría con la práctica.

Relación con el nested model

Lo anterior lo podemos conjugar con la imperiosa necesidad de formación profesional de nuevos astrónomos que deban desempeñarse como científicos y analistas de fuentes de altas energías en el renombrado proyecto en desarrollo en Chile SWGO que estará operativo en 2030, el cual identificamos al ejecutar el modelo anidado, ya que ahí encontramos esta necesidad a nivel industria. De esta manera abordamos un proyecto de datos que atendiera tanto necesidades a nivel de industria educativa (estudiantes) fortaleciendo competencias analíticas y de programación, como profesional o investigativa (SWGO Chile) para fomentar el interés en el desarrollo del estudio de altas energías.

Objetivo del análisis

De acuerdo a lo mencionado anteriormente, y con el objetivo de promover desde etapas tempranas de la formación de astrónomos, especialmente en análisis y programación es que a continuación exponemos un caso práctico de uso de análisis de eventos o fotones (captados por el telescopio espacial Fermi) de rayos gamma de alto nivel con gammapy, único paquete de Python dedicado completamente a la astronomía de rayos gamma como un todo, con esto quiero decir que podemos usar gammapy como denominador común para estudiar, analizar y visualizar eventos de cualquier instrumento

(CTA, HESS y Fermi LAT), lo que la convierte en una librería ideal para ingenieros, científicos o investigadores.

Es muy relevante mencionar que trabajaremos dentro del rango de data level 3 (DL3), ya que utilizaremos datos ya preparados o procesados por los encargados técnicos de cada instrumento enfocados principalmente en el listado de eventos.

(Indico. (2024, diciembre 3-4). Gammapy workshop)²⁴

Cada evento es una detección individual de un fotón de rayos gamma. Estos eventos contienen información detallada sobre el tiempo, las coordenadas espaciales, la energía, y otras características relevantes de la detección, lo que permite a los astrónomos realizar un análisis profundo sobre las fuentes de rayos gamma.

(Python package for gamma-ray astronomy. *Astronomy & Astrophysics*)²⁵

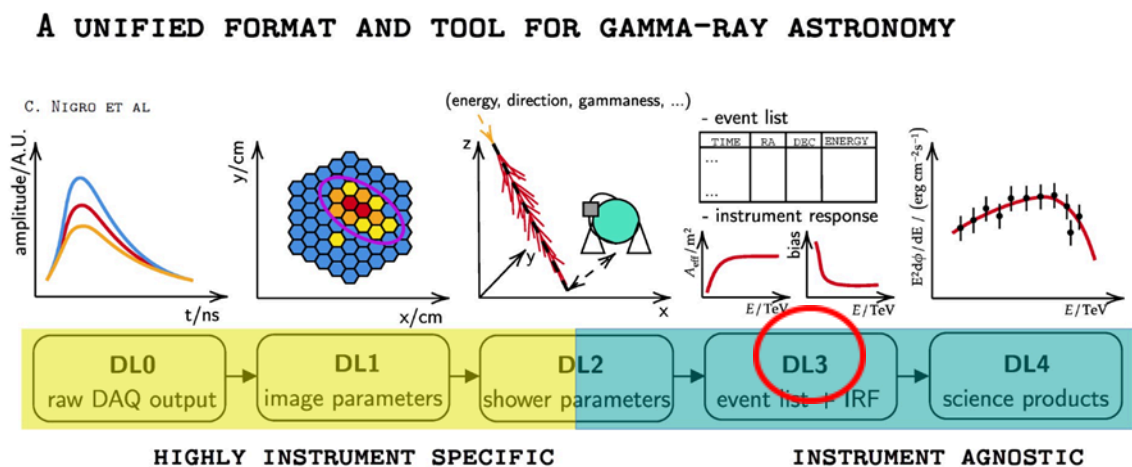


Imagen 34

Casi cualquier análisis de alto nivel de rayos gamma comienza en el nivel de datos DL3, donde los datos de rayos gamma se representan como listas de eventos tipo rayos gamma y sus correspondientes Funciones de

Respuesta del Instrumento (IRF's). (Python package for gamma-ray astronomy. *Astronomy & Astrophysics*)²⁵

En este acápite abordaremos paso a paso el cómo hacer en Gammapy los siguientes análisis:

- Análisis Espacial 3D
- Análisis de coincidencia espacial con catálogo Fermi
- Análisis de detección de fuentes y mapas de significancia

Estos 3 tipos de análisis si bien no son los únicos existentes que se pueden realizar con los Event Lists, sí representan ser análisis de alto nivel e introductorios a la astrofísica de altas energías. Por ejemplo, el análisis espacial permite identificar si un o unos eventos provienen de una fuente puntual (como un púlsar) o de una región más difusa (como el fondo galáctico). El análisis de coincidencia consiste en verificar si el conjunto de eventos (fuente) ya se encuentra identificado en el catálogo Fermi. Y finalmente el análisis de detección de fuentes y mapas de significancia ayuda a elegir un cúmulo de fotones como un potencial candidato de fuente verídica.

En los siguientes acápites explicaremos paso a paso cómo realizar cada uno de estos análisis en las Event Lists de tal manera de hacerlo lo más “user friendly” posible.

Manos a la obra (Step-by-Step)

Preparación y exploración de los datos

El siguiente código contempla la importación de las librerías necesarias, entre ellas `gammapy` y la lectura del listado de eventos subyacente a la misma librería `gammapy`

```
import numpy as np

from astropy import units as u

from astropy.coordinates import SkyCoord

# %matplotlib inline

import matplotlib.pyplot as plt

from IPython.display import display

from gammapy.data import EventList

from gammapy.datasets import Datasets, MapDataset

from gammapy.irf import EDispKernelMap, PSFMap

from gammapy.maps import Map, MapAxis, WcsGeom

from gammapy.modeling import Fit

from gammapy.modeling.models import (
```

```

Models,
PointSpatialModel,
PowerLawNormSpectralModel,
PowerLawSpectralModel,
SkyModel,
TemplateSpatialModel,
create_fermi_isotropic_diffuse_model,
)

from gammapy.utils.check import
check_tutorials_setup

check_tutorials_setup()

```

```

import os

os.environ['GAMMAPY_DATA'] =
'/content/gammapy-data/1.3'

```

```

events = EventList.read(

```

```
"/content/gammapy-data/1.3/fermi_3fhl/fermi_3fhl_e
vents_selected.fits.gz"
)
print(events)
```

EventList

Instrument : LAT

Telescope : GLAST

Obs. ID :

Number of events : 697317

Event rate : 0.003 1 / s

Time start : 54682.65603222222

Time stop : 57236.96833546296

Min. energy : 1.00e+04 MeV
Max. energy : 2.00e+06 MeV
Median energy : 1.59e+04 MeV

```
from astropy.time import Time

# Obtener el tiempo de inicio y fin de los eventos

time_start = Time(events.time[0], format='mjd') #
Primer evento como 'start'

time_stop = Time(events.time[-1], format='mjd') #
Último evento como 'stop'

# Convertir las energías de MeV a GeV
```

```

min_energy_geV    =    int((events.energy.min()    /
1000).value)    # Energía mínima

max_energy_geV    =    int((events.energy.max()    /
1000).value)    # Energía máxima

# Imprimir los resultados

print("Time start:", time_start.iso)    # Fecha en
formato ISO

print("Time stop:", time_stop.iso)    # Fecha en
formato ISO

print(f"Min. energy: {min_energy_geV:.0f} GeV")

print(f"Max. energy: {max_energy_geV:.0f} GeV")

```

Time start: 2008-08-04 15:49:26.782

Time stop: 2015-07-30 11:00:41.226

Min. energy: 10 GeV

Max. energy: 1998 GeV

El listado de eventos que cargamos corresponde al tercer catálogo de fuentes duras del Fermi Lat (Gamma-ray Large Area Space Telescope | GLAST), correspondiente a 7 años de data, de Agosto 2008 a Agosto 2015.

Esta posee cerca de 700 mil eventos con una frecuencia mínima de 10 GeV o 10,000 MeV y máxima de 2,000 GeV o 2,000,000 MeV, además de una mediana de 15,900 MeV o ~ 16 GeV.

A continuación vamos a aplicar la función `peek()` de la clase `EventList` que brinda una visión general amplia agrupando los eventos según las cantidades.

```
events.peek()
```

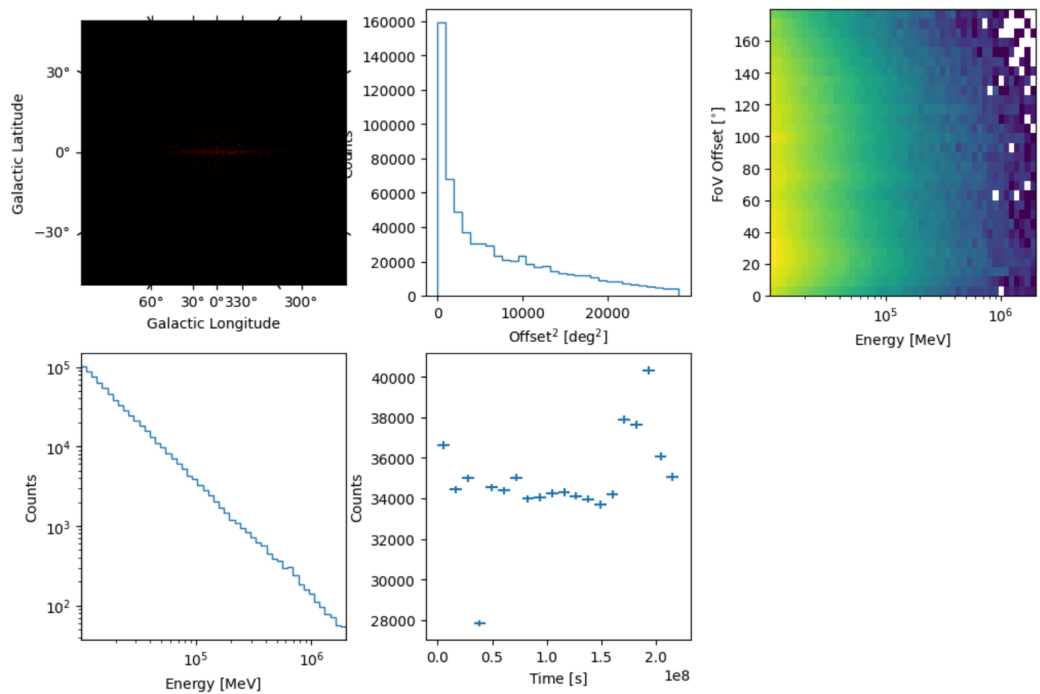


Imagen 36

Podemos ver en la fila superior, de izquierda a derecha, primero un mapa de conteo bruto (un simple histograma 2D en coordenadas celestes) de las

direcciones de llegada de los eventos, los cuales vemos se concentran en el centro galáctico.

En el segundo gráfico tenemos los eventos histogramados según su desplazamiento (su distancia cuadrada en grados), corresponde a una distribución de los fotones detectados en función de su distancia cuadrada respecto al centro del campo de visión (FOV) del telescopio. A medida que nos alejamos del centro del campo de visión, es común que el número de fotones detectados disminuya. Esto se debe a que la eficiencia de detección del telescopio generalmente disminuye a medida que los fotones llegan desde ángulos más alejados del centro.

Luego en el tercer gráfico vemos cómo varía la eficiencia de detección de los fotones (claro amarillo menor eficiencia y azul oscuro que representa mayor eficiencia) dependiendo de su energía y su desviación angular desde el centro del campo de visión (FOV), de hecho se aprecia que a más altas energías la eficiencia de detección de fotones es mayor, tal como lo expone Carlos en la explicación de la IRF's.

En la segunda fila el primer gráfico nos muestra el conteo de fotones o eventos en función de su energía en el eje x el cual vemos decrece a mayor energía, esto ya que los eventos altamente energéticos son más inusuales.

El segundo gráfico nos muestra como varía el conteo de fotones a lo largo del tiempo, donde podemos apreciar que hay peaks importantes de actividad los últimos rangos de tiempo.

Selección de datos y reducción de datos

Una vez realizada la exploración general de datos de nuestra EventList, corresponde ahora hacer una reducción de datos y selección de área de interés en la cual queremos realizar los próximos análisis.

Como primer filtro seleccionaremos los eventos energéticos que sean compatibles con los rangos con los que trabajará el proyecto SWGO, o sea sobre los 100 GeV.

```
from astropy.units import Quantity

energy_cut = Quantity(1e5, 'MeV') # 100 GeV = 1e5 MeV

filtered_events = events.select_energy([0.1*u.TeV, 2*u.TeV])

print("Numero de eventos detectados en el
rango:", len(events.table[events.table['ENERGY'] >= energy_cut]))

print("-----")

print("Primeros 10 registros de tabla de eventos filtrada >100 GeV")

print(filtered_events.table[:10])
```

Numero de eventos detectados en el rango: 23628

Primeros 10 registros de tabla de eventos filtrada >100 GeV

ENERGY	RA	DEC	L	B	...	DIFRSP0	DIFRSP1	DIFRSP2	DIFRSP3	DIFRSP4
MeV	deg	deg	deg	deg	...					
145927.23	291.6619	42.234055	74.54367	11.86778	...	0.0	0.0	0.0	0.0	0.0
221273.36	46.985752	-40.638943	247.48888	-58.873917	...	0.0	0.0	0.0	0.0	0.0
221223.75	83.56263	-4.217438	207.7828	-19.077145	...	0.0	0.0	0.0	0.0	0.0
698996.9	320.89468	-1.327888	51.221825	-33.97183	...	0.0	0.0	0.0	0.0	0.0
119158.54	318.81128	12.30284	62.63608	-24.41602	...	0.0	0.0	0.0	0.0	0.0
1418123.2	100.31091	-47.44805	256.468	-21.264088	...	0.0	0.0	0.0	0.0	0.0
412811.53	107.60061	13.661594	202.77013	10.322091	...	0.0	0.0	0.0	0.0	0.0
276105.47	86.58336	-65.0223	274.63168	-31.242601	...	0.0	0.0	0.0	0.0	0.0
141472.95	255.3199	-26.424526	356.55865	9.530827	...	0.0	0.0	0.0	0.0	0.0
315729.47	238.16652	35.242012	56.364143	50.74429	...	0.0	0.0	0.0	0.0	0.0

Imagen 37

```
for e_min in [100, 500, 1000, 1500, 1990, 2000] *  
u.GeV:  
  
    n = (filtered_events.energy > e_min).sum()  
  
    print(f"Events above {e_min:4.0f}: {n:5.0f}")
```

Events above 100 GeV: 23628

Events above 500 GeV: 2179

Events above 1000 GeV: 544

Events above 1500 GeV: 164

Events above 1990 GeV: 5

Events above 2000 GeV: 0

Como primer análisis de nuestro filtro energético (>100 GeV) en base a los distintos rangos aplicados en la imagen anterior, los eventos sobre los 100 GeV's son alrededor de 24.000. Además nos percatamos que a medida que vamos amplificando el rango, cada vez son menores los eventos energéticos, en donde son solo 5 eventos los que se encuentran sobre los 1990 GeV's. Esto se condice con la primera vista con la función `.peek()` donde el conteo iba decantando a medida que aumentaba el rango energético.

Análisis espacial 3D

Ahora vamos a estudiar cómo se distribuyen los eventos de rayos gamma en el cielo para identificar y caracterizar las fuentes de rayos gamma.

Una vez seleccionados los eventos de interés, en este caso todos aquellos sobre los 100 GeV's de energía (para replicar el rango de tratamiento del SWGO), nos aprestamos a definir una geometría de mapa, esto es, un espacio de coordenadas con píxeles y proyección adecuada.

En el siguiente ejemplo seleccionaremos la coordenada galáctica en el centro de la vía láctea 0° de longitud galáctica y 0° de latitud galáctica, los bins de ejes de energía dados en el MapAxis. También usamos interpolación en el eje de energía en escala logarítmica, lo cual es común cuando se trabaja con espectros de rayos gamma, ya que estas suelen tener una distribución logarítmica de energía. Y finalmente se crea el mapa de conteo de eventos en el Map. create, en el sistema de coordenadas “galáctico” y proyección del mapa tangencial (TAN), la cual es muy utilizada para representar áreas pequeñas del cielo de manera precisa.

```

gc_pos      =      SkyCoord(0,      0,      unit="deg",
frame="galactic")

energy_axis = MapAxis.from_edges(
    [1e4, 3e4, 1e5, 3e5, 2e6], name="energy",
unit="MeV", interp="log"
)

counts = Map.create(
    skydir=gc_pos,
    npix=(300, 200),
    proj="TAN",
    frame="galactic",
    binsz=0.1,
    axes=[energy_axis],
    dtype=float,
)

```

```
counts.fill_events(filtered_events)

print(counts.geom.axes[0])

counts.sum_over_axes().smooth(2).plot()

#counts.sum_over_axes().smooth(2).plot(stretch="sqrt", vmax=30)

plt.title("Distribución eventos de rayos Gamma  
(100 - 2000 GeV's) | Centro galáctico")

plt.show()
```

MapAxis

```
name      : energy
unit      : 'MeV'
nbins     : 4
node type : edges
edges min : 1.0e+04 MeV
edges max : 2.0e+06 MeV
interp    : log
```

Distribución eventos de rayos Gamma (100 - 2000 GeV's) | Centro galáctico

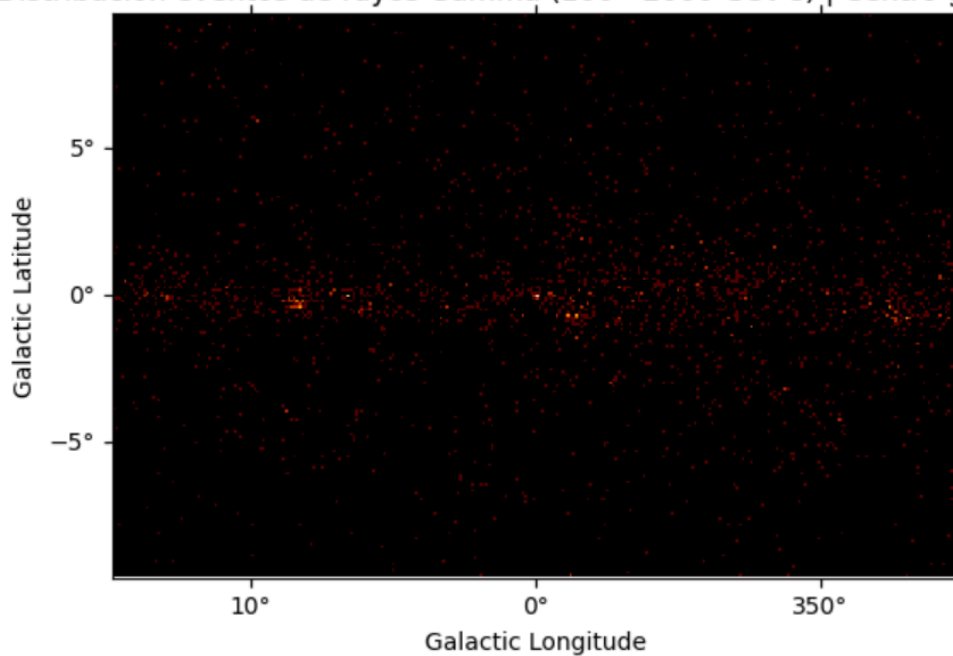


Imagen 38

En este mapa vemos la distribución de los eventos o fotones de rayos gamma en el centro galáctico.

Distribución eventos de rayos Gamma (100 - 2000 GeV's) | Centro galáctico

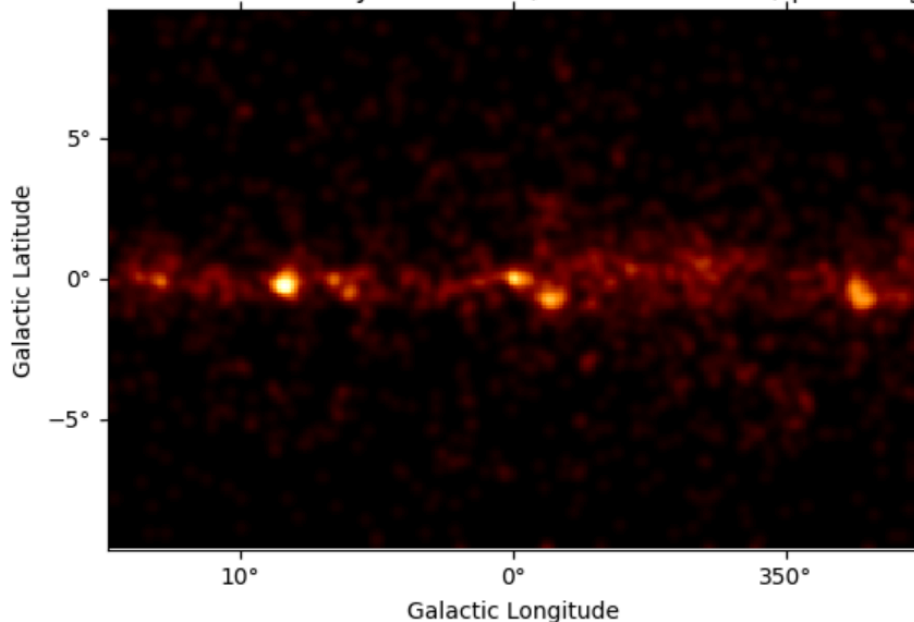


Imagen 39

En este nuevo mapa al suavizar los eventos con la función `smooth(2)` es más fácil discernir eventuales fuentes, objetos o incluso regiones que emiten rayos gamma.

Podemos apreciar la cantidad de eventos que conforman el mapa que definimos por medio del siguiente comando:

```
print(f"Numero total de conteos en el mapa:
{counts.data.sum():.0f}")
```

Número total de conteos en el mapa: 2094

Dado el resultado del despliegue del mapa podríamos comentar que los eventos de rayos gamma podrían estar provenientes de esta región densa de

la galaxia (centro de nuestra galaxia), donde hay una mayor actividad de rayos cósmicos y otras fuentes de emisión gamma.

De la misma forma que expusimos el conteo de eventos sobre cierto rangos energéticos anteriormente, podemos también exponerlo en la distribución espacial en el mapa que generamos en el centro galactico:

```
gc_pos      =      SkyCoord(0,      0,      unit="deg",
frame="galactic")

energy_axis = MapAxis.from_edges(
    [1e4, 3e4, 1e5, 3e5, 2e6], name="energy",
unit="MeV", interp="log"
)

# Ciclo sobre los diferentes umbrales de energía
for e_min in [100, 500, 1000, 1500, 1990, 2000] *
u.GeV:
    # Filtrar los eventos por el umbral de energía
    i = events.table[events.table['ENERGY'] >=
e_min]
```

```

    i = EventList(i) # Es necesario pasarla a
eventlist

# Crear un mapa vacío para los conteos

counts = Map.create(

skydir=gc_pos,

npix=(400, 200),

proj="TAN",

frame="galactic",

binsz=0.1,

axes=[energy_axis],

dtype=float,

)

counts.fill_events(i)

counts.sum_over_axes().smooth(2).plot()

# Agregar título

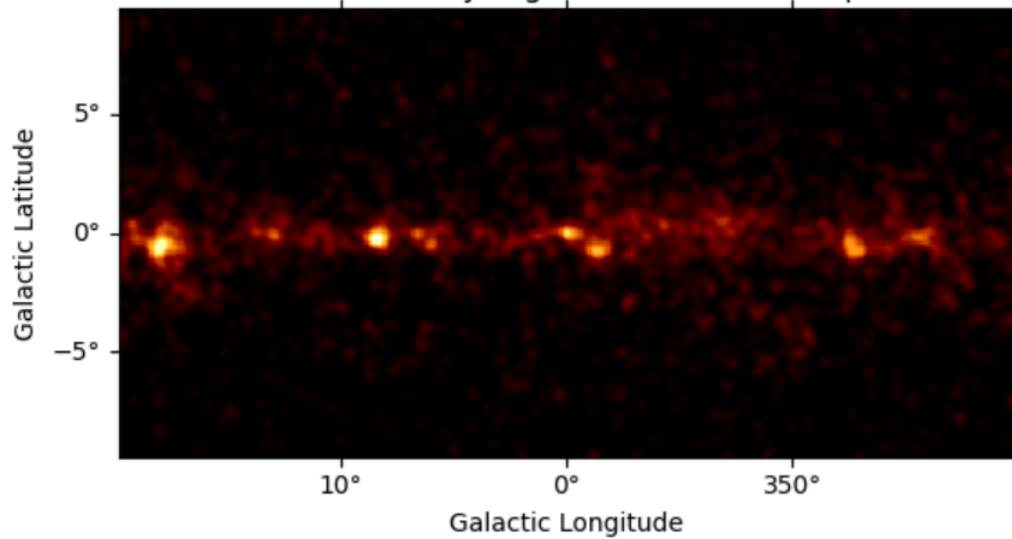
plt.title(f"Distribución de eventos de rayos
gamma (> {e_min:.0f}) | Centro Galáctico")

```

```
plt.show()

print(f"Numero total de conteos en la imagen:
{counts.data.sum():.0f}")
```

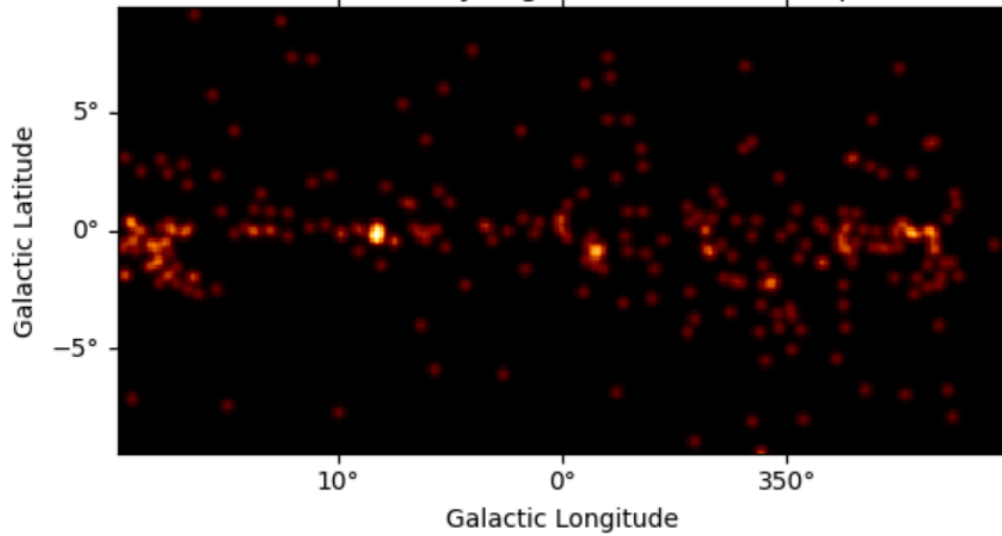
Distribución de eventos de rayos gamma (> 100 GeV) | Centro Galáctico



Numero total de conteos en la imagen: 2849

Imagen 40

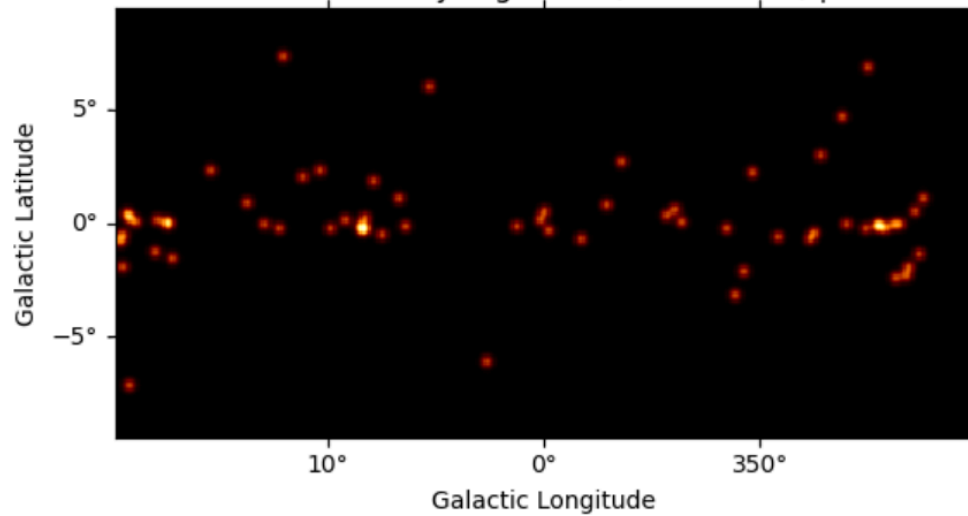
Distribución de eventos de rayos gamma (> 500 GeV) | Centro Galáctico



Numero total de conteos en la imagen: 272

Imagen 41

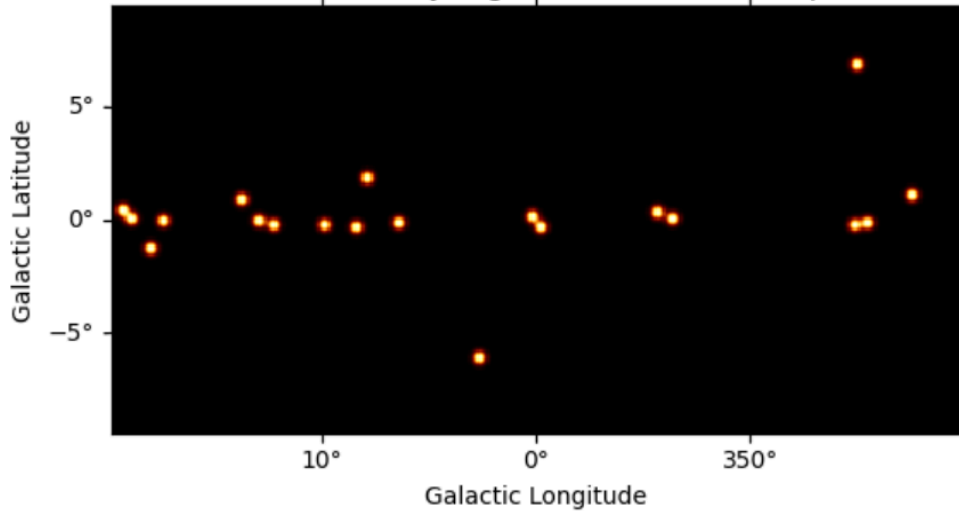
Distribución de eventos de rayos gamma (> 1000 GeV) | Centro Galáctico



Numero total de conteos en la imagen: 64

Imagen 42

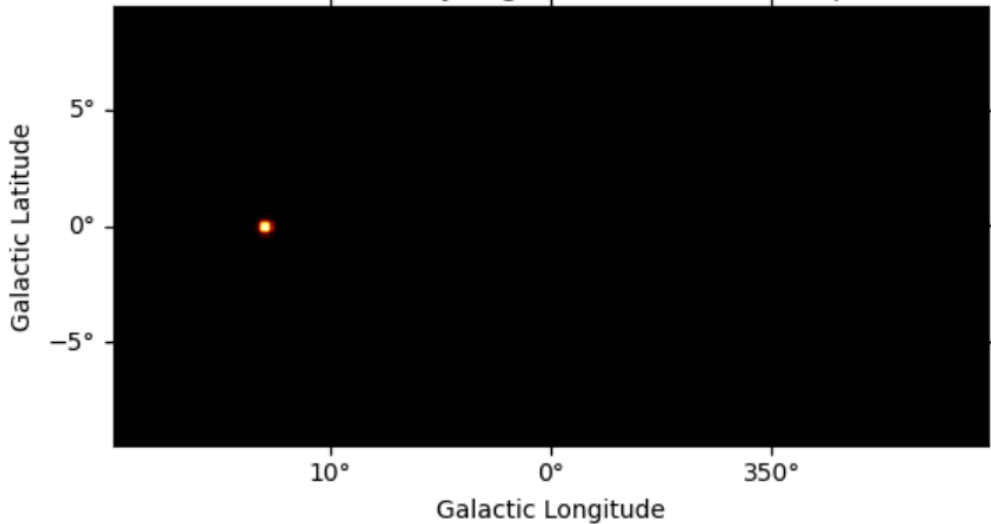
Distribución de eventos de rayos gamma (> 1500 GeV) | Centro Galáctico



Numero total de conteos en la imagen: 20

Imagen 43

Distribución de eventos de rayos gamma (> 1990 GeV) | Centro Galáctico



Numero total de conteos en la imagen: 1

Imagen 44

Las vistas anteriores refuerzan lo visto inicialmente, donde existe una relación decreciente entre el conteo de eventos y la energía, los cuales son muy aislados en rangos energéticos extremos sobre los 1,500 GeV's por

ejemplo. De hecho sobre los 1990 GeV's solo se puede constatar 1 evento en el mapa galáctico.

Distribución temporal de eventos

A continuación queremos visualizar la frecuencia de los eventos en el tiempo, o sea, cómo se distribuyen conforme pasan los días en el intervalo de tiempo desde 2008 a 2015 en la frecuencia energética de 100 a 2,000 GeV's y encontrar peaks de actividad en alguna temporalidad.

```
#from astropy.time import Time

events_df = pd.DataFrame({

                                'TIME':
Time(filtered_events.time,format='mjd').datetime,
# Tiempo del evento (debe estar en formato float,
no string)

                                'RA': filtered_events.radec.ra.value,      #
Coordenada de ascensión recta

                                'DEC': filtered_events.radec.dec.value,    #
Coordenada de declinación

                                'ENERGY':
filtered_events.energy.value.astype('float64')      #
Energía del fotón

})
```

```

#events_df[['TIME','RA','DEC','ENERGY']]

events_df

events_df['time_bin'] =
events_df['TIME'].dt.floor('d') #Aqui jugar con H
:horas u otro

# 2. Contar el número de eventos por intervalo de
tiempo

event_counts =
events_df.groupby('time_bin').size()

# 3. Análisis básico de variabilidad: Calcular la
desviación estándar y la media

mean_events = event_counts.mean()

std_events = event_counts.std()

# 4. Identificar peaksde actividad: Por ejemplo,
peaks que superen dos desviaciones estándar

```

```

threshold = mean_events + 2 * std_events # Umbral
para peakos

peaks = event_counts[event_counts > threshold] #
peaks por encima del umbral

# 5. Visualización: Graficar la distribución
temporal de eventos

plt.figure(figsize=(10, 6))

event_counts.plot(kind='line', marker='o',
label='Número de eventos')

plt.axhline(y=mean_events, color='r',
linestyle='--', label=f'Media
({mean_events:.2f})')

plt.axhline(y=threshold, color='g',
linestyle='--', label=f'Threshold
({threshold:.2f})')

# Resaltar los picos detectados

plt.scatter(peaks.index, peaks.values,
color='purple', label='Peaks', zorder=5)

```

```
plt.title('Distribución Temporal de Eventos de  
Fermi-3FHL (100 - 2000 GeV)')  
  
plt.xlabel('Tiempo')  
  
plt.ylabel('Número de Eventos')  
  
plt.legend()  
  
plt.xticks(rotation=45)  
  
plt.grid(True)  
  
plt.show()  
  
# 6. Imprimir los peaksdetectados  
print("Peaks detectados:")  
  
print(peaks)
```

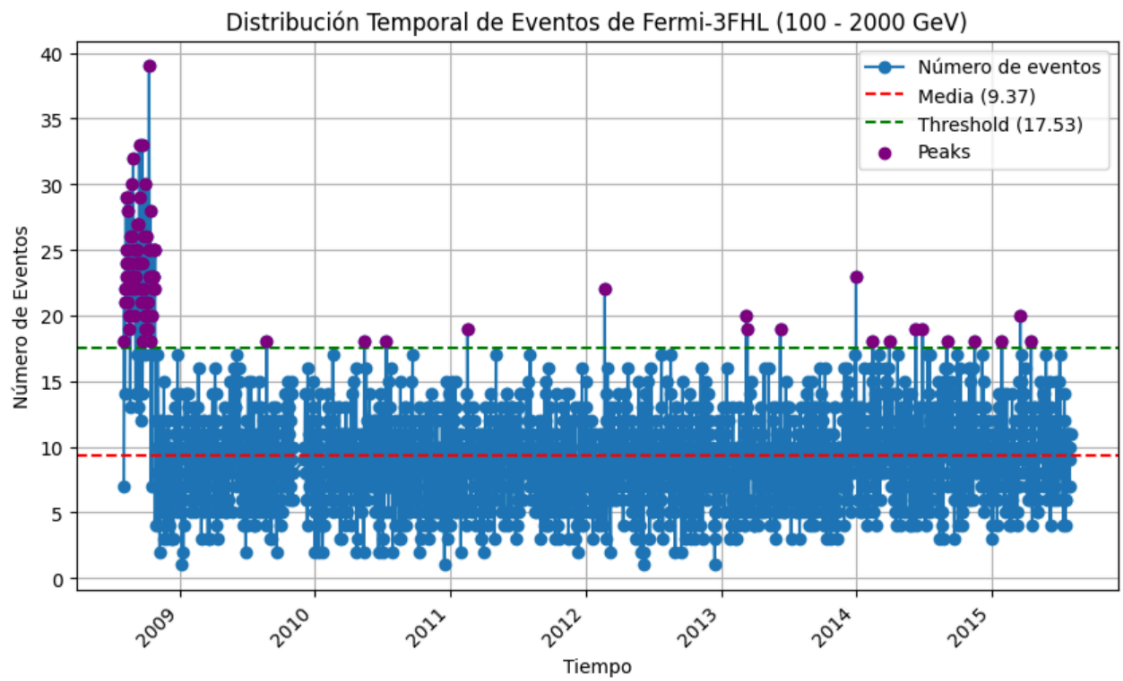


Imagen 45

Un insight importante del resultado del gráfico de la temporalidad es que se aprecia una actividad bastante inusual en el año 2008, donde está concentrada la mayor cantidad de eventos o fotones de forma diaria.

Análisis en base a una región de interés (ROI)

Si queremos hacer un análisis de un cúmulo de eventos (fuente de rayos gamma) en alguna zona o región que nos llame la atención o nos interese visualmente, podemos hacerlo acotando a una región de interés que definiremos a continuación.

En este caso la zona encerrada en amarillo corresponde a la región de interés que queremos estudiar.

Lo primero que hacemos es importar los mismos eventos que cargamos inicialmente sobre los 100 GeV's de energía.

Distribución eventos de rayos Gamma (100 - 2000 GeV's) | Centro galáctico

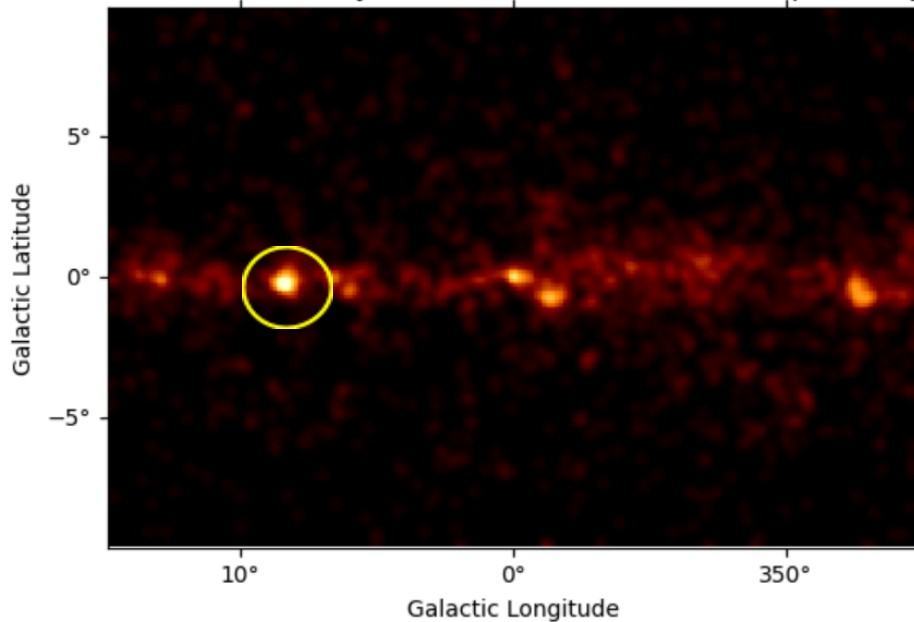


Imagen 46

Luego, estudiaremos y analizaremos los eventos comprendidos en el círculo amarillo para posteriormente ver si es que en dicha región existe o no una fuente de rayos gamma.

```
# select all events within a radius of 0.1 deg
around center

from          gammapy.utils.regions          import
SphericalCircleSkyRegion

filtered_events_2 =
events.select_energy([0.1*u.TeV, 2*u.TeV])

center = SkyCoord(8, 0, unit="deg",
frame="galactic")

region = SphericalCircleSkyRegion(center, radius=1
* u.deg)

events_gc_2fh1 =
filtered_events_2.select_region(region)

events_gc_2fh1.energy.to("GeV")
```

```
# sort events by energy

events_gc_2fhl.table.sort("ENERGY")

# and show highest energy photon

events_gc_2fhl.energy[-1].to("GeV")
```

Energía más alta identificada en el Event List: 1829.351 GeV

```
print(len(events_gc_2fhl.table))
```

```
events_gc_2fhl.table
```

Table length=99

ENERGY	RA	DEC	L	B	THETA	PHI	ZENITH_ANGLE	EARTH_AZIMUTH_ANGLE
MeV	deg	deg	deg	deg	deg	deg	deg	deg
float32	float32	float32	float32	float32	float32	float32	float32	float32
100329.52	271.40952	-21.564693	8.645785	-0.18880378	17.84215	157.30751	15.073135	190.7156
101066.33	271.5826	-21.781975	8.53477	-0.43529925	42.033054	8.8914795	70.32646	248.26418
101680.31	271.45508	-21.735386	8.51758	-0.30918947	35.453808	193.54512	10.496282	282.87872
102372.15	271.84436	-21.689478	8.734031	-0.6025432	51.05598	191.62651	58.25722	257.21912
103185.88	271.46698	-21.606936	8.635045	-0.2560509	34.28453	25.280195	55.447216	142.42116
103632.02	271.46655	-21.909153	8.371169	-0.40338606	8.390311	250.31274	54.325634	236.97026
106827.516	271.1994	-22.048355	8.128531	-0.25547642	26.716974	248.28587	46.482807	243.26077
106991.4	271.59506	-21.731266	8.58468	-0.42065373	58.360016	354.3488	76.73899	237.96423
108421.43	271.5164	-21.912731	8.39064	-0.44548428	31.08411	105.4147	46.373764	154.1359
...
631501.06	270.9089	-21.708805	8.291982	0.14607579	49.003212	25.972027	70.344795	240.94011
660174.5	271.36795	-21.95162	8.289413	-0.34438288	27.645044	359.7821	91.17501	246.90823
774617.94	271.08966	-22.698746	7.5119596	-0.48603866	18.74637	145.74567	59.263054	102.34943
930836.9	271.45926	-22.061625	8.234835	-0.47199404	51.259186	202.57497	21.079771	102.319725
1096903.5	271.2469	-21.881199	8.295825	-0.21195436	51.36735	197.77272	43.875687	111.92674
1209242.0	271.2467	-21.66629	8.483098	-0.10652971	58.34826	289.59494	46.533035	235.93837
1264046.9	271.20877	-22.723944	7.543881	-0.59415054	58.04395	158.80681	33.502056	114.16288
1437969.9	270.8469	-21.688179	8.281592	0.20638508	9.958978	203.9197	63.67567	107.72244
1829351.0	271.38434	-21.836084	8.397618	-0.301123	29.280958	357.80603	60.267723	224.94708

Imagen 47

Con el código anterior enmarcamos los eventos > 100 GeV que se encuentran circundantes a la región de interés en coordenadas galácticas (en este caso 8° de longitud galáctica y 0° de latitud galáctica). En total equivalen a 99 fotones de rayos gamma en dicha región de interés cuya máxima energía alcanza los 1,829.351 GeV's. Además los fotones según la tabla inmediatamente anterior nos muestra que las coordenadas RA (Ascensión recta) y DEC (Declinación), del sistema ecuatoriano de referencia celeste (equivalente de la longitud y latitud respectivamente), se encuentran en torno a los RA: 271 y DEC: -21, coordenadas que nos servirán luego para buscar coincidencias de existencia de fuentes de rayos gamma en el catálogo Fermi-LAT.

De estos 99 eventos recogidos podemos comprobar de una forma más explícita la distribución de la energía de dichos eventos o fotones:

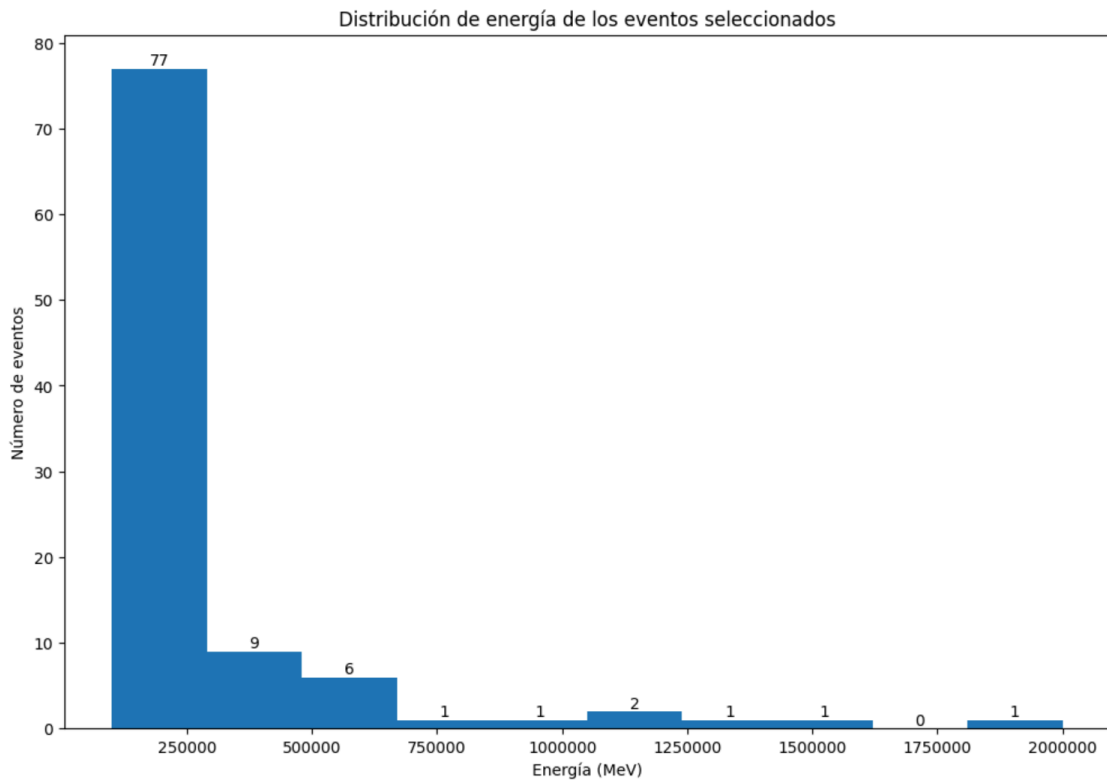


Imagen 48

Vemos que solamente 5 fotones se encuentran en el rango de energía sobre los 1,000 GeV's, mientras que la gran mayoría de los fotones se sitúan bajo los 250 GeV's, algo bastante esperable.

```
events_gc_2fhl.peek()
```

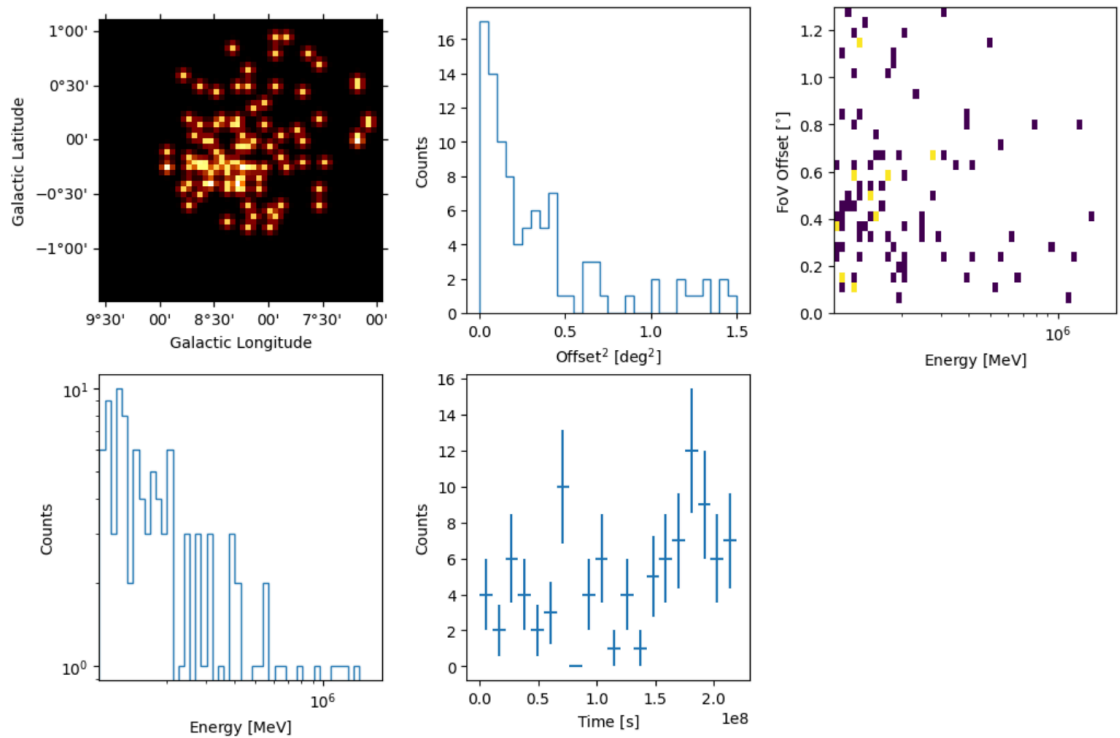


Imagen 49

Con estas vistas generales apreciamos en el primer gráfico la dispersión de los fotones alrededor de las latitudes y longitudes galácticas. En el segundo gráfico la detección de fotones va decantando a medida que nos alejamos del centro de campos de visión. En el tercer gráfico vemos que aparece menos colorido el mapa debido a que estamos estudiando un acotado número de eventos de una zona específica, donde de todas formas comprobamos que la eficiencia de detección de las altas energías es mayor, las cuales en este caso son bastante pocas.

Abajo en el primer gráfico vemos también que el conteo de eventos decrece a medida que aumentamos en el rango energético. Y finalmente vemos la variación del conteo en el tiempo, la cual no muestra una variabilidad tan significativa en el tiempo de la actividad de eventos.

Análisis de coincidencia espacial catálogo Fermi

Antes de iniciar el análisis, un poco de contexto respecto al catálogo Fermi.

Un catálogo es un producto procesado (Data Level 5) que sintetiza información sobre fuentes identificadas de rayos gamma como posiciones y extensiones de las fuentes, espectros de energía y curvas de luz, clasificaciones tentativas de las fuentes (púlsares, blazares, supernovas, etc.)

Utilizaremos el catálogo Fermi 3FHL, el cual fue construido en base al listado de eventos procesado en el punto anterior.

A continuación vamos a verificar usando el catálogo de fuentes de Fermi 3FHL si este conjunto de fotones estudiados anteriormente se encuentra mapeado como una fuente u objeto de rayos gamma ya conocida.

Primero cargamos el catálogo directamente de la librería Gammapy:

```
import os

os.environ['GAMMAPY_DATA'] =
    '/content/gammapy-data/1.3'
```

```
from gammapy.catalog import SourceCatalog3FHL

fermi_3fhl = SourceCatalog3FHL()

print(fermi_3fhl)

#print(fermi_3fhl.table.colnames)

fermi_3fhl.table[:5]
```



```
# Pasamos a DataFrame

df = filtered_table.to_pandas()

# Contamos la ocurrencias de la columna "CLASS"
(tipo de evento energetico)

class_counts = df['CLASS'].value_counts()

print(class_counts)

print("-----")

print("Total fuentes del catalogo:", len(df))
```

```

CLASS
b'b11      '      731
b'bcu     '      290
b'        '      177
b'fsrq    '      142
b'PSR     '       53
b'FSRQ    '       30
b'unknown'       23
b'BLL     '       19
b'snr     '       17
b'SNR     '       13
b'rdg     '        9
b'spp     '        9
b'PWN     '        9
b'pwn     '        8
b'psr     '        6
b'RDG     '        4
b'HMB     '        4
b'sbg     '        4
b'g1c     '        2
b'hmb     '        1
b'NLSY1   '        1
b'sfr     '        1
b'BIN     '        1
b'agn     '        1
b'SFR     '        1
Name: count, dtype: int64
-----
Total fuentes del catalogo: 1556

```

Imagen 51

En anexo se encuentra el diccionario de los nemotécnicos de las clases.

TABLE 2
3FHL SOURCE CLASSES

Description	Identified		Associated	
	Designator	Number	Designator	Number
Pulsar	PSR	53	psr	6
Pulsar Wind Nebula	PWN	9	pwn	8
Supernova remnant	SNR	13	snr	17
Supernova remnant / Pulsar wind nebula	spp	9
High-mass binary	HMB	4	hmb	1
Binary	BIN	1
Globular cluster	glc	2
Star-forming region	SFR	1	sfr	1
Starburst galaxy	sbg	4
BL Lac type of blazar	BLL	19	bll	712
Flat spectrum radio quasar type of blazar	FSRQ	30	fsrq	141
Non-blazar active galaxy	agn	1
Narrow-line Seyfert 1	NYLS1	1
Radio galaxy	RDG	4	rdg	9
Blazar candidate of uncertain type	bcu	309
Total	identified	136	associated	1219
Unclassified	unknown	24
Unassociated	177
Total in the 3FHL	1556

Imagen 52

La captura del paper del Fermi 3FHL nos constata que se lograron identificar significativamente 136 fuentes de rayos gamma donde los púlsares son los más identificados con certeza.

(3FHL: The Third Catalog of Hard Fermi-LAT Sources)26

Una vez cargado el catálogo Fermi 3FHL, buscamos las fuentes de acuerdo a los rangos captados anteriormente por los eventos seleccionados que estaban en torno a las 271 RA y -21 DEC:

```
# Definir los límites del rango de RA y DEC

ra_min = 271 # RA mínima (en grados)

ra_max = 272 # RA máxima (en grados)

dec_min = -22 # DEC mínima (en grados)

dec_max = -21 # DEC máxima (en grados)

# Filtrar las fuentes que están dentro del rango de RA y DEC

sources_in_range = []

# Acceder a las columnas RA y DEC de la tabla

ra_column = fermi_3fhl.table['RAJ2000']

dec_column = fermi_3fhl.table['DEJ2000']

name_column = fermi_3fhl.table['Source_Name'] # Nombre de la fuente

# Filtrar las fuentes
```

```

for i in range(len(fermi_3fhl.table)):

    ra = ra_column[i] # Obtener RA de la fuente

    dec = dec_column[i] # Obtener DEC de la fuente

    # Verificar si la fuente está dentro del rango de RA y DEC

    if ra_min <= ra <= ra_max and dec_min <= dec <= dec_max:

        sources_in_range.append((name_column[i], ra, dec))

# Mostrar las fuentes que están dentro del rango

for source in sources_in_range:

    print(f"Fuente: {source[0]}, RA: {source[1]}, DEC: {source[2]}")

```

```

Fuente: 3FHL J1804.7-2144e, RA: 271.1910095214844, DEC:
-21.735000610351562

```

Finalmente encontramos la fuente en el catálogo, de la cual podemos obtener mayor detalle:

```
fermi_3fhl["3FHL J1804.7-2144e"]
```

Catalog row index (zero-based) : 1435
Source name : 3FHL J1804.7-2144e
Extended name : FGES J1804.7-2144
Associations : W 30, HESS J1804-216, 3FGL J1805.6-2136e
ASSOC_PROB_BAY : --
ASSOC_PROB_LR : --
Class : snr
TeVcat flag : E

*** Other info ***

Significance (10 GeV - 2 TeV) : 22.714
Npred : 472.5

HEP Energy : 1829.350 GeV
HEP Probability : 0.874
Bayesian Blocks : 1
Redshift : --
NuPeak_obs : 0.0 Hz

*** Position info ***

RA : 271.191 deg
DEC : -21.735 deg
GLON : 8.397 deg
GLAT : -0.095 deg

Semimajor (95%) : 0.0000 deg
Semiminor (95%) : 0.0000 deg
Position angle (95%) : 0.00 deg
ROI number : 37

Imagen 53

Aquí vemos que dicha fuente corresponde a un “snr” esto es un remanente de supernova.

ADD-ON: Gamma-Sky

Existe otra forma de identificar la existencia de fuentes en el catálogo de una forma más amigable para el usuario, y esta es accediendo al gamma-sky.net, plataforma web e interactiva del catálogo de los distintos observatorios. Su objetivo es ofrecer una herramienta accesible para los astrónomos y científicos interesados en estudiar fuentes de rayos gamma, así como eventos astrofísicos relacionados con la radiación de muy alta energía.

(GammaSky. (n.d.). Gamma sky map)²⁷

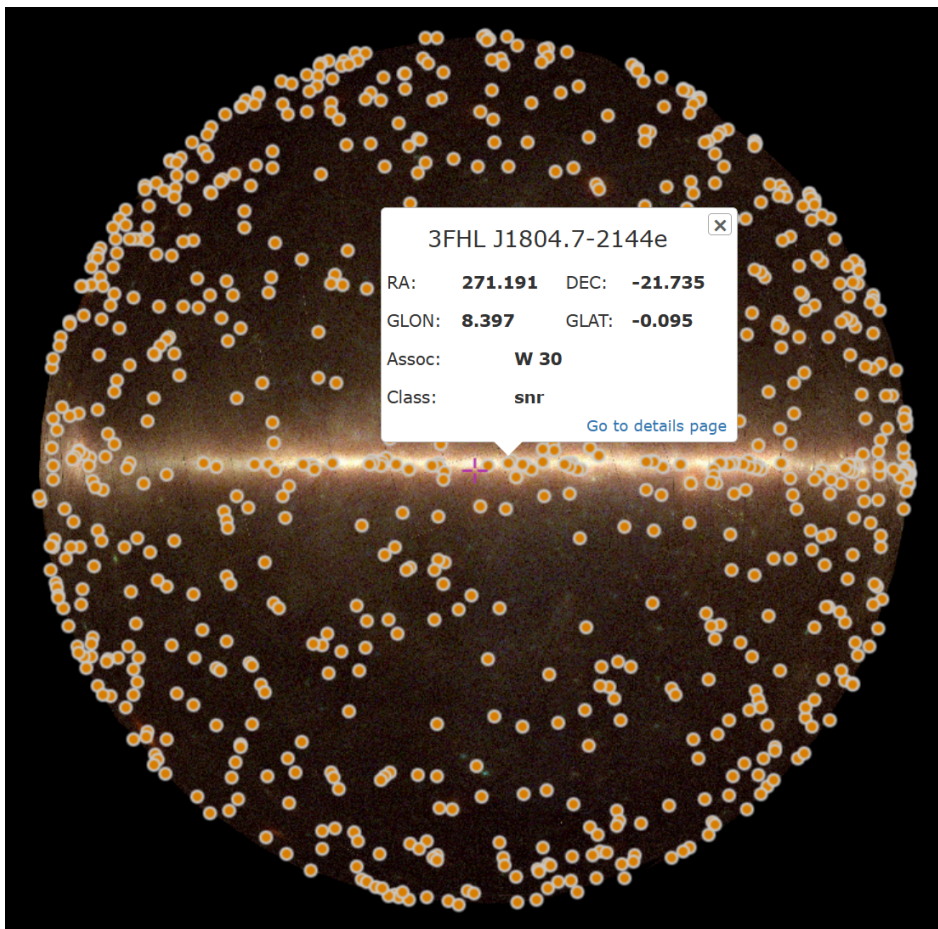


Imagen 54

Con la coordenadas identificadas es posible lograr identificar la fuente más o menos con la misma precisión que al hacerlo con gammapy.

Análisis de detección de fuentes y mapas de significancia

La primera tarea en la producción de un catálogo de fuentes es identificar los excesos significativos en los datos que se puedan asociar eventualmente a fuentes desconocidas y proporcionar una parametrización preliminar en términos de posición, extensión y flujo. Para esto vamos a presentar cómo detectar fuentes candidatas en mapas de conteos.

(Gammapy. (2025). Detecting gamma-ray sources: 2D analysis tutorial)28

En este apartado aplicaremos algunas de las IRF's (de las que en este apartado no profundizaremos ya que se abordan en el acápite de las IRF's) incluidas en el dataset de gammapy de Fermi 3FHL para ajustar el conteo de eventos por la sensibilidad del instrumento. También aplicaremos técnicas para poder discriminar potenciales fuentes estadísticamente significativas.

Lo primero que hacemos es cargar los mapas necesarios de fermi 3FHL. Como inputs, primero está el mapa de conteos de eventos (usaremos el mapa de conteos en el rango de energías definido de 100 a 2000 GeV's), luego el mapa de fondo conocido que nos entrega el dataset fermi-3FHL que nos permite aislar los “eventos” que forman parte de fondo galáctico. Tenemos también el cubo de exposición (IRF) de tal manera de corregir el conteo de eventos en función de la sensibilización del telescopio. También tenemos el PSF(point spread function) (IRF) que representa la resolución angular del instrumento. Y finalmente el energy dispersion (IRF) que describe cómo se dispersa la energía de los fotones detectados dada las limitaciones del detector mismo.

```
import numpy as np
```

```

import astropy.units as u

# %matplotlib inline

import matplotlib.pyplot as plt

from IPython.display import display

from gammapy.datasets import MapDataset

from gammapy.estimators import ASmoothMapEstimator, TSMAPEstimator

from gammapy.estimators.utils import find_peaks

from gammapy.irf import EDispKernelMap, PSFMap

from gammapy.maps import Map

from gammapy.modeling.models import PointSpatialModel, PowerLawSpectralModel, SkyModel

# MAPA DE FONDO

counts = Map.read("$GAMMAPY_DATA/fermi-3fhl-gc/fermi-3fhl-gc-counts-cube.fits.gz")

# MAPA DE FONDO

background = Map.read(

```

```

"$GAMMAPY_DATA/fermi-3fhl-gc/fermi-3fhl-gc-background-cube.fits.gz
"
)

#MAPA DE EXPOSICIÓN

exposure =
Map.read("$GAMMAPY_DATA/fermi-3fhl-gc/fermi-3fhl-gc-exposure-cub
e.fits.gz")

psfmap = PSFMap.read(
    "$GAMMAPY_DATA/fermi-3fhl-gc/fermi-3fhl-gc-psf-cube.fits.gz",
    format="gtpsf",
)

# ENERGY DISPERSION

edisp = EDispKernelMap.from_diagonal_response(
    energy_axis=counts.geom.axes["energy"],
    energy_axis_true=exposure.geom.axes["energy_true"],
)

```

```
#AJUSTE DEL DATASET CONFORME A LOS MODELOS
```

```
dataset = MapDataset(  
    counts=counts,  
    background=background,  
    exposure=exposure,  
    psf=psfmap,  
    name="fermi-3fhl-gc",  
    edisp=edisp,  
)
```

En la última sección del código anterior ajustamos el mapa de conteo con los modelos de respuesta del instrumento que nos entrega `gammapy` por defecto.

Algoritmo de suavizado adaptativo

Un primer enfoque en la detección de posibles fuentes es aplicar el llamado “algoritmo de suavizado adaptativo” para suavizar el mapa de conteos. El algoritmo básicamente busca mejorar la resolución y la precisión en la detección de estructuras de interés en datos de alta energía, particularmente en el contexto de la astrofísica de rayos gamma y la observación de cúmulos de galaxias.

(ASMOOTH: A simple and efficient algorithm for adaptive kernel smoothing of two-dimensional imaging data)²⁹

Siguiendo con el dataset ajustado anterior realizamos el algoritmo de suavizado.

En el siguiente código el argumento `threshold` indica la significancia mínima esperada, donde los valores bajo este valor simplemente se recortan.

```
scales = u.Quantity(np.arange(0.05, 1, 0.05), unit="deg")

smooth = ASmoothMapEstimator(threshold=3, scales=scales,
energy_edges=[100, 2000] * u.GeV)

images = smooth.run(dataset)

plt.figure(figsize=(9, 5))

images["flux"].plot(add_cbar=True, stretch="asinh")
```

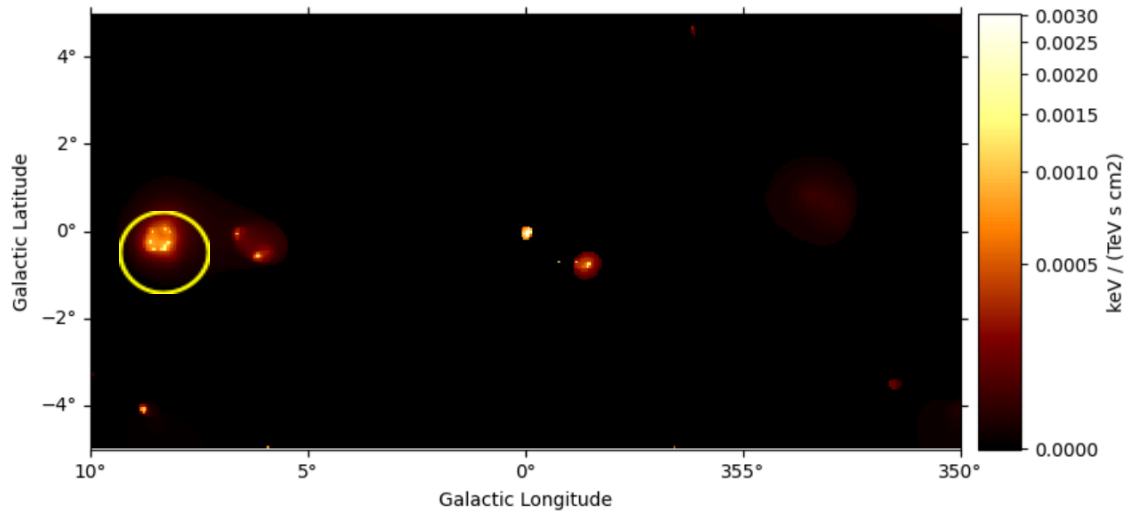


Imagen 55

Visualmente podemos constatar con cierta claridad en el gráfico las potenciales fuentes candidatas ya que resaltan gracias a la significancia del algoritmo de suavizado.

De hecho la zona marcada en amarillo corresponde al área definida en el acápite anterior de análisis espacial, donde se aprecia una fuerte significancia.

Estimación del mapa Test Estadístico (TS)

Otro enfoque en la medición de la significancia es aplicar la estimación del test estadístico.

El test estadístico $TS = 2 \Delta \log L$ de Mattox 1996 compara la función de verosimilitud L optimizada con y sin una fuente dada. El mapa TS se calcula ajustando por un único parámetro de amplitud en cada píxel, como se describe en el Apéndice A de Stewart (2009). El ajuste se simplifica encontrando raíces de la derivada de las estadísticas de ajuste (la configuración predeterminada utiliza el método de Brent).

(Gammapy. (2025). Detecting gamma-ray sources: 2D analysis tutorial)28

La prueba estadística tiene como objetivo identificar regiones del cielo que presentan excesos significativos de eventos, lo cual indica la presencia de una fuente, mientras se compara con el ruido de fondo (eventos no asociados a ninguna fuente conocida).

Para emprender la estimación del mapa del test estadístico lo primero es definir el modelo que se usará para probar la existencia de una fuente. En este caso usaremos “point source” que representa una fuente puntual, donde se asume que la fuente está localizada en un punto específico del cielo, sin distribución espacial.

```
spatial_model = PointSpatialModel()

# We choose units consistent with the map units here...
spectral_model = PowerLawSpectralModel(amplitude="1e-22 cm-2 s-1 keV-1", index=2)
model = SkyModel(spatial_model=spatial_model,
spectral_model=spectral_model)
```

Se crea el estimador que calculará el mapa de prueba estadística (TS) para la detección de fuentes en el cielo, utilizando el modelo previamente definido y con un rango de energías de 100 - 2000 GeV's

```
estimator = TSMAPEstimator(  
    model,  
    kernel_width="1 deg",  
    energy_edges=[100, 2000] * u.GeV,  
)  
maps = estimator.run(dataset)
```

Aquí generamos 3 gráficos, el primero es el mapa de significancia, el segundo el mapa de flujo y el tercero el mapa de iteraciones durante el ajuste del modelo.

```
fig, (ax1, ax2, ax3) = plt.subplots(  
    ncols=3,  
    figsize=(15, 3),  
    subplot_kw={"projection": counts.geom.wcs},  
    gridspec_kw={"left": 0.1, "right": 0.98},  
)  
  
maps["sqrt_ts"].plot(ax=ax1, add_cbar=True)  
ax1.set_title("Significance map")  
maps["flux"].plot(ax=ax2, add_cbar=True, stretch="sqrt", vmin=0)  
ax2.set_title("Flux map")
```

```
maps["niter"].plot(ax=ax3, add_cbar=True)
ax3.set_title("Iteration map")
```

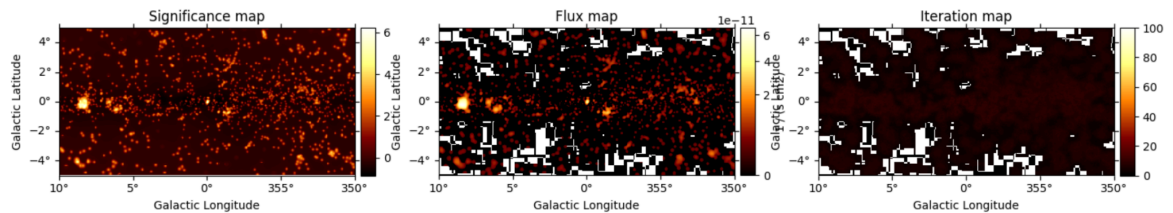


Imagen 56

En el primer gráfico de mapa de significancia, los colores representan la significancia estadística de las fuentes, donde las zonas más intensas representan más evidencia de una posible fuente, mientras que las zonas más oscurecidas podrían indicar ruido de fondo.

El segundo gráfico mapa de flujo, que muestra la intensidad de la emisión de rayos gamma en el cielo. El flujo describe la cantidad de fotones detectados por unidad de área y tiempo, las áreas de colores más brillantes o cálidos indican un mayor flujo de fotones, lo que sugiere una fuente más intensa.

El tercer gráfico indica cuántas iteraciones se han realizado en cada punto del mapa. Las áreas donde se han realizado más iteraciones pueden ser aquellas donde el modelo de la fuente tiene más incertidumbre, o donde se necesita mayor refinamiento en el ajuste de los parámetros.

, ahora podemos demarcar las zonas estadísticamente significativas ejecutando un buscador de peaks. La función `find_peaks` realiza una búsqueda máxima local en una ventana deslizante, el argumento

`min_distance` es la distancia mínima en píxeles entre picos (el valor más pequeño posible y el valor predeterminado es 1 píxel).

```
sources = find_peaks(maps["sqrt_ts"], threshold=5,
min_distance="0.25 deg")

nsou = len(sources)

display(sources)

# Plot sources on top of significance sky image

plt.figure(figsize=(9, 5))

ax = maps["sqrt_ts"].plot(add_cbar=True)

ax.scatter(
    sources["ra"],
    sources["dec"],
    transform=ax.get_transform("icrs"),
    #transform=plt.gca().get_transform("icrs"),
    color="none",
    edgecolor="w",
```

```

marker="o",

s=600,

lw=1.5,

)

plt.show()

# sphinx_gallery_thumbnail_number = 3

```

Table length=3

value	x	y	ra deg	dec deg
float64	int64	int64	float64	float64
6.2159	34	93	271.34260	-21.95444
6.0046	201	98	266.43351	-29.03927
5.1837	24	17	275.21776	-23.33290

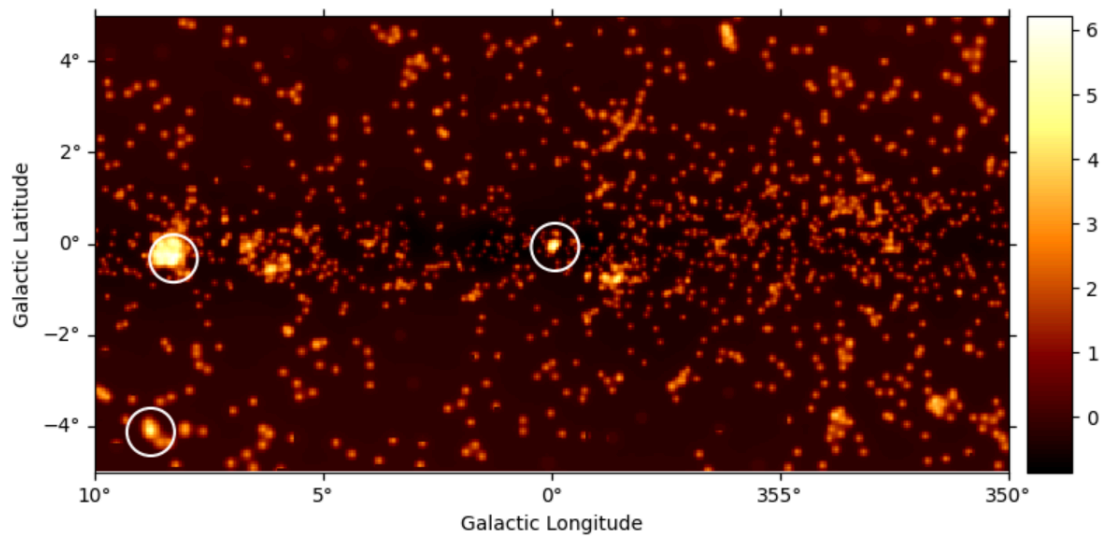


Imagen 57

Las zonas encerradas en círculo corresponden a fuentes con un threshold mayor o igual a 5 de la significancia estadística. Esto quiere decir que estas 3 fuentes del mapa de conteo cargado en los rangos energético 100 a 200 GeV's, son potenciales fuentes verídicas de rayos gamma.

Conclusión

En este apartado logramos poner en práctica 3 tipos de análisis alto nivel que son relevantes en el campos de la astronomía de altas energías, por una parte el análisis espacial que nos permitió lograr aprender a desplegar en un mapa galáctico un listado de eventos de tal manera de visualizar cómo se distribuyen en función de sus coordenadas y energía.

Vimos también la aplicación del catálogo Fermi para poder contrastar y hacer match de un cúmulo de eventos (potencial fuente) del listado de eventos y así verificar la coincidencia de fuentes.

Finalmente aplicamos técnicas de pruebas estadísticas y suavizado para poder identificar con mayor seguridad potenciales fuentes de rayos gamma.

Con estos ejercicios prácticos esperamos poder ayudar y motivar al estudiante en etapa formativa de astronomía (interesado en programación y alta energía) fortaleciendo justamente capacidades necesarias constatadas en la entrevistas del estudiantado; En el análisis y la programación de datos astronómicos.

5. Conclusiones

El estudio y análisis de las fuentes de rayos gamma es un área que se ha visto enriquecida por la creciente disponibilidad de herramientas y recursos tecnológicos para la captura y análisis de datos.

La construcción del observatorio SWGO en Chile, previsto para su funcionamiento en 2030, abre nuevas oportunidades para la investigación. Con el propósito de contribuir a la formación de profesionales en el área, este trabajo presenta una guía orientada a facilitar el entendimiento y manejo de datos de fuentes de rayos gamma, un campo que se caracteriza por su complejidad y la necesidad de habilidades técnicas avanzadas. La guía, que abarca tres áreas clave; Funciones de Respuesta del Instrumento (IRF's), análisis con la librería Gammapy, y el uso de Machine Learning. En base a esto busca ser un primer paso accesible para aquellos interesados en profundizar en el análisis de datos astrofísicos y eventuales prospectos trabajadores en el SWGO.

Vimos que las **Funciones de Respuesta del Instrumento (IRFs)** son cruciales para corregir y calibrar las mediciones de telescopios como Fermi-LAT, permitiendo estimar con precisión las propiedades de las fuentes de rayos gamma, como energía, flujo y localización. Estas funciones abordan las limitaciones de los instrumentos y mejoran la interpretación de los datos, especialmente en fuentes débiles o eventos transitorios.

Las simulaciones de Monte Carlo son clave para optimizar la calibración de los instrumentos y mejorar la precisión de las observaciones al modelar el comportamiento del telescopio frente a partículas de alta energía.

Los análisis de fuentes de rayos gamma con Gammapy que vimos en este trabajo corresponden a análisis de alto nivel en un nivel de datos 3 ya

que trabajamos con datos ya procesados (lista de eventos), análisis que todo astrónomo en etapas formativas debiese manejar y tener conocimiento.

Aprendimos a importar listas, realizar una vista y entendimiento general de los eventos de fotones de rayos gamma. También a analizar por medio de la distribución espacial 3D el comportamiento de los eventos, y también a detectar potenciales eventos que pueden ser fuente de rayos gamma.

Con esto se espera propender a incentivar y fortalecer los conocimientos en cuanto a análisis y programación en futuros astrónomos de altas energías.

El uso de Machine Learning en el análisis de rayos gamma, ha permitido explorar de manera integral los eventos astrofísicos de alta energía, combinando técnicas avanzadas de análisis espacial, energético y temporal con métodos estadísticos y de aprendizaje automático. Estas herramientas no solo han demostrado ser esenciales para desentrañar patrones ocultos en los datos, sino que también han facilitado la identificación de eventos transitorios y anomalías que ofrecen una visión más profunda de los procesos físicos detrás de estos fenómenos.

El uso de Machine Learning, aplicado tanto a la predicción temporal como a la detección de anomalías, ha mejorado significativamente la precisión y eficiencia en el manejo de grandes volúmenes de datos. Esto no solo contribuye a comprender mejor las características y comportamientos de las fuentes de rayos gamma, sino que también establece un precedente valioso para futuros análisis en proyectos como el Observatorio de Rayos Gamma del Hemisferio Sur (SWGGO). En este contexto, el presente trabajo sienta una base sólida para el entrenamiento de profesionales que trabajarán con datos de rayos gamma, ofreciendo una guía práctica y accesible para estudiantes y nuevos investigadores que buscan incursionar en este campo.

Además, se resalta la importancia de contar con herramientas de visualización interactivas y metodologías claras que permitan no solo analizar, sino también interpretar y comunicar efectivamente los resultados

obtenidos. La correcta interpretación de los datos es crucial para garantizar que los avances científicos derivados del análisis sean comprensibles y aplicables en contextos más amplios.

Este trabajo representa una contribución significativa al campo de la astrofísica de rayos gamma, al proporcionar una guía práctica que combina teoría, herramientas avanzadas y metodologías accesibles. Este enfoque asegura que los interesados puedan abordar de manera efectiva los desafíos que presentan proyectos como el SWGO, promoviendo así el desarrollo de un área incipiente en Chile. Al capacitar a futuros científicos y astrónomos con las habilidades necesarias, este trabajo no solo apoya la investigación actual, sino que también impulsa el crecimiento del talento local en un campo de relevancia global.

Limitaciones

Dentro de las principales limitaciones o dificultades destacamos las siguientes:

- Falta de conocimiento técnico en astronomía y astrofísica, lo que dificulta el entendimiento de muchos conceptos.
- Desafío en la dificultad de la representación gráfica de datos de rayos gamma.
- Cumplir con la expectativa del eventual estudiante target al que va dirigido este artículo.
- Acceso público de información limitado para fuentes de alta energía tipo rayos gamma (por ejemplo en IRF's)
- Material de información web en general poco amigable o intuitivo para el usuario incipiente.

Trabajo futuro

- Revisión con la academia de la idoneidad del material
- Publicación y despliegue oficial.

6. Bibliografía

"Introduction to High-Energy Astrophysics" de Rosswog y Bruggen: Una introducción amigable a los conceptos fundamentales de la astrofísica de alta energía, incluyendo rayos gamma, sin requerir un conocimiento profundo de física avanzada.

"High-Energy Astrophysics" de Fulvio Melia: Profundiza en las observaciones y los instrumentos, incluyendo el telescopio LAT, útil para entender cómo funcionan los detectores de rayos gamma.

Artículos de Fermi-LAT (disponibles en arXiv): Artículos como "The Fermi Large Area Telescope On Orbit: Event Classification, Instrument Response Functions, and Calibration" ofrecen una visión directa de cómo se manejan y procesan los datos reales.

1. Centro de Astrofísica y Tecnologías Afines (CATA). (n.d.). Estudio de rayos gamma: El desafío de la nueva generación de astrónomos en Chile. CATA.
<https://cata.cl/estudio-de-rayos-gamma-el-desafio-de-la-nueva-generacion-de-astronomos-en-chile/>
2. Donath, A., Terrier, R., Remy, Q., Sinha, A., Nigro, C., Pintore, F., Khélifi, B., Olivera-Nieto, L., Ruiz, J. E., Brügge, K., Linhoff, M., Contreras, J. L., Acero, F., Aguasca-Cabot, A., Berge, D., Bhattacharjee, P., Buchner, J., Boisson, C., Carreto Fidalgo, D., Chen, A., de Bony de Lavergne, M., de Miranda Cardoso, J. V., Deil, C., Füssling, M., Funk, S., Giunti, L., Hinton, J., Jouvin, L., King, J., Lefaucheur, J., Lemoine-Goumard, M., Lenain, J. P., López-Coto, R., Mohrmann, L., Morcuende, D., Panny, S., Regeard, M., Saha, L., Siejkowski, H., Siemiginowska, A., Sipocz, B. M., Unbehauen, T., van Eldik, C., Vuillaume, T., & Zanin, R. (2023). Gammapy: A Python package for gamma-ray astronomy. *Astronomy & Astrophysics*, 678, A157. <https://doi.org/10.1051/0004-6361/202346488>

3. Wood, M., Caputo, R., Charles, E., Di Mauro, M., Magill, J., & Perkins, J. (2017). Fermipy: An open-source Python package for analysis of Fermi-LAT Data. *Proceedings of the 35th International Cosmic Ray Conference (ICRC2017)*, 824. <https://doi.org/10.48550/arXiv.1707.09551>
4. Conceição, R., & SWGO Collaboration. (2023). The Southern Wide-field Gamma-ray Observatory. *Proceedings of the 38th International Cosmic Ray Conference (ICRC2023)*, 963. <https://doi.org/10.48550/arXiv.2309.04577>
5. Pasaje a la Ciencia. (2021, septiembre 3). *La Astronomía de rayos gamma de alta energía: los telescopios MAGIC*. Recuperado de <https://pasajealaciencia.es/la-astronomia-de-rayos-gamma-de-alta-energia-los-telescopios-magic/>
6. Orellana, M. (s.f.). *Radiación Gamma en Binarias con Acreción*. Instituto Argentino de Radioastronomía. Recuperado de <http://tux.iar.unlp.edu.ar/divulgacion/art-difu-31.htm>
7. Universidad Técnica Federico Santa María. (2024, agosto 21). *USM tendrá destacada participación en el primer observatorio de rayos gamma del hemisferio sur*. <https://usm.cl/noticias/usm-tendra-destacada-participacion-en-el-primer-observatorio-de-rayos-gamma-del-hemisferio-sur/>
8. Cherenkov Telescope Array Observatory. (s.f.). *Diccionario*. <https://www.ctao.org/news-resources/outreach-and-education/ctao-for-educators/diccionario/>
9. ctools. (s.f.). *What are IRFs?*. ctools User Manual. https://cta.irap.omp.eu/ctools/users/user_manual/irf_intro.html
10. Gammapy Development Team. (s.f.). *Instrument Response Functions*. Gammapy Documentation. <https://docs.gammapy.org/0.9/irf/index.html>

11. Gammapy Development Team. (s.f.). *CTA first data challenge (1DC) with Gammapy*. Gammapy Documentation. Recuperado de https://docs.gammapy.org/0.9/notebooks/cta_1dc_introduction.html
12. Gamma-ray astronomy community. (2018). *Effective Area vs true energy*. En *Data formats for gamma-ray astronomy* (versión 0.3). Recuperado de https://gamma-astro-data-formats.readthedocs.io/en/v0.3/irfs/full_enclosure/aeff/index.html#effective-area-vs-true-energy
13. Donath, A., Terrier, R., Remy, Q., Sinha, A., Nigro, C., Pintore, F., ... & Zanin, R. (2023). Gammapy: A Python package for gamma-ray astronomy. *Astronomy & Astrophysics*, 678, A157. <https://doi.org/10.1051/0004-6361/202346488>
14. H.E.S.S. Collaboration. (s.f.). *About H.E.S.S.*. H.E.S.S. - The High Energy Stereoscopic System. <https://www.mpi-hd.mpg.de/HESS/pages/about/>
15. NASA. (n.d.). *Fermi Gamma-ray Space Telescope*. <https://fermi.gsfc.nasa.gov/>
16. Apache Software Foundation. (n.d.). *Apache Parquet documentation*. <https://parquet.apache.org/documentation/latest/>
17. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). *Scikit-learn: Machine Learning in Python*. *Journal of Machine Learning Research*, 12, 2825-2830. <http://jmlr.org/papers/v12/pedregosa11a.html>
18. Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control*. Holden-Day.

19. Seabold, S., & Perktold, J. (2010). *Statsmodels: Econometric and statistical modeling with python*. In *9th Python in Science Conference*.
<https://www.statsmodels.org/stable/index.html>
20. Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, 313(5786), 504-507.
<https://doi.org/10.1126/science.1127647>
21. Montgomery, D. C. (2020). *Design and analysis of experiments* (10th ed.). John Wiley & Sons.
22. Conda documentation. <https://docs.conda.io/>
23. HEASARC. (n.d.). *Fermi LAT Data - Weekly Spacecraft Files*. NASA's High Energy Astrophysics Science Archive Research Center.
https://heasarc.gsfc.nasa.gov/FTP/fermi/data/lat/weekly/1s_spacecraft/
24. Indico. (2024, dicembre 3-4). Gammapy workshop:
<https://indico.in2p3.fr/event/34684/timetable/#20241203>
25. Python package for gamma-ray astronomy. *Astronomy & Astrophysics*
https://www.aanda.org/articles/aa/full_html/2023/10/aa46488-23/aa46488-23.html
26. 3FHL: The Third Catalog of Hard Fermi-LAT Sources

<https://arxiv.org/abs/1702.00664>

27. GammaSky. (n.d.). Gamma sky map

<http://gamma-sky.net/#/map?target=12.909,-1.004&fov=180.000&marker=tev>

28. Gammapy. (2025). Detecting gamma-ray sources: 2D analysis tutorial

<https://docs.gammapy.org/1.0/tutorials/analysis-2d/detect.html#>

29. ASMOOTH: A simple and efficient algorithm for adaptive kernel smoothing of two-dimensional imaging data

<https://doi.org/10.1111/j.1365-2966.2006.10135.x>

- Astropy:

<https://docs.astropy.org/en/stable/>

- FITS Standard

https://fits.gsfc.nasa.gov/fits_standard.html

- Ackermann 2012

Ackermann, M., Ajello, M., Albert, A., Allafort, A., Atwood, W. B., Axelsson, M., Baldini, L., Ballet, J., Barbiellini, G., Bastieri, D., Bechtol, K., Bellazzini, R., Bissaldi, E., Blandford, R. D., Bloom, E. D., Bogart, J. R., Bonamente, E., Borgland, A. W., Bottacini, E., ... Zimmer, S. (2012). The Fermi Large Area Telescope on orbit: Event

classification, instrument response functions, and calibration. The Astrophysical Journal Supplement Series, 203(1), 4. <https://doi.org/10.1088/0067-0049/203/1/4>

- Atwood 2009

Atwood, W. B., Abdo, A. A., Ackermann, M., Anderson, B., Axelsson, M., Baldini, L., Ballet, J., Band, D. L., Barbiellini, G., Bartelt, J., Bastieri, D., Baughman, B. M., Bechtol, K., Bédèrède, D., Bellardi, F., Bellazzini, R., Berenji, B., Bignami, G. F., Bisello, D., ... Ziegler, M. (2009). The Large Area Telescope on the Fermi Gamma-Ray Space Telescope Mission. arXiv preprint arXiv:0902.1089v1. <https://arxiv.org/abs/0902.1089>

- Ruiz Velasco 2024

Ruiz Velasco, E. L., Schoorlemmer, H., & SWGO Collaboration. (2024). Reference Analysis Chain Description. Southern Wide-field Gamma-ray Observatory. Max Planck Institute for Nuclear Physics, Heidelberg, Germany. Retrieved from https://www.swgo.org/SWGOWiki/lib/exe/fetch.php?media=swgopla_nv1.3.pdf

7. Anexo I: Encuesta Rebeca Burgos estudiante de pregrado en Astronomía

1. ¿Cuál es tu background académico?

Soy estudiante de astronomía de la UdeC, me encuentro cursando el penúltimo semestre de mi carrera, actualmente estoy trabajando en mi tesis, que está enfocada en estudiar las perturbaciones periódicas del flujo de un telescopio que monitorea Blazars. Durante el verano estuve trabajando en un internship en la ESO con el doctor Matias Jones, dónde trabajé modificando parámetros de calibración del instrumento SPHERE en el telescopio UT3, y anteriormente participé de una pasantía en la Universidad de Concepción con el profesor Neil Nagar adaptando códigos para analizar espectros de quasars usando ppxf. Además me apasiona mucho la divulgación astronómica, soy parte del Equipo de Divulgación Astronómica

(EDA) de la UdeC, y recientemente comencé una página en instagram de divulgación por mi cuenta @astro.beca en colaboración a mi compañero @yercu.contreras.astro.

2. ¿Qué conocimientos previos tienes sobre los rayos gamma y su uso en la astronomía?

Los rayos gamma en la astronomía son producidos por eventos que liberan electrones de muy alta energía, estos pueden ser por explosiones de supernovas, formación estelar, fusión entre dos estrellas de neutrones, entre otros. Estas frecuencias son las más energéticas del espacio y casi alcanzan

la velocidad de la luz. Mediante el estudio de estos rayos se puede conocer todos estos eventos o fenómenos que no son vistos a simple vista.

3. ¿Conoces algo acerca del satélite Fermi?

Es un telescopio diseñado para estudiar los rayos gamma del espacio, siendo su objetivo principal realizar un mapa detallado del Universo en esta frecuencia. Está compuesto por dos instrumentos principales: LAT (Large Area Telescope), el cual se usa para mapear el cielo buscando distintos tipos de fenómenos, como AGNs y pulsares; y GBM (Gamma-Ray Burst Monitor) el cual monitorea cualquier brote de rayos gamma en rayos x. Este telescopio ha dado pistas claves respecto a la evolución de los Agujeros Negros Supermasivos en AGNs, y a detectado nuevos púlsares.

4. ¿Conoces algo sobre el proyecto SWGO 2030?

Es un proyecto para construir en Chile el primer telescopio de rayos gamma en el hemisferio sur. SWGO viene de Southern Wide-field Gamma-ray Observatory y a diferencia de los observatorios tradicionales, este no necesitará antenas ni telescopios, ya que este obtendrá información a través de rayos gamma que serán captados en tanques de agua sellados.

5. ¿Consideras que los recursos educativos que tuviste disponibles son suficientes para aprender sobre astrofísica de altas energías y procesar datos en Python o R?

Creo que tanto como estudiar astronomía no es una tarea fácil, debe ser muy

complejo al mismo tiempo preparar alumnos desde 0 para convertirlos en científicos capacitados en solo 5 años, resumir la información y las habilidades de toda una vida de estudio para traspasarla a jóvenes inexpertos debe ser muy difícil de completar. En base a esto creo que no es suficiente lo que se aprende de forma teórica, ya que la práctica al final termina realmente formando a cada científico. Aún así, creo que la enseñanza desde temprano en el procesamiento de datos está muy débil, es la herramienta principal y básica para un astrónomo, incluso un poco más

que tener habilidades matemáticas, el doctor Matias Jones siempre le dice a sus alumnos “Un astrónomo es un programador” cosa que es muy cierta pero al mismo tiempo se enseña muy a la ligera. Creo que sería ideal que se enseñara la física de la mano con el procesamiento de datos, que podamos ver lo teórico y lo práctico al mismo tiempo para poder ir formando un hilo continuo, y al mismo tiempo acostumbrarnos, y desarrollar el instinto de investigación. Ahora, al menos en mi Universidad y en los proyectos en los que he participado no he usado R ni me han enseñado a usarlo. Pero sí son muy comunes los lenguajes como Python y Fortran.

6. ¿Qué herramientas o recursos consideras que son aún necesarios en los programas académicos de Chile respecto a la materia señalada en el punto anterior?

Creo que se debería profundizar en el análisis de datos y en los lenguajes de programación, las mallas de astronomía deberían compartir mayores características con carreras como informática. Añadir más práctica entre la teoría también es muy necesario, la mayor parte de la carrera se siente en gran parte teórica, y recién en la tesis debes aprender a llevar todo a la práctica y eso se siente como un gran salto al agua sin flotador. Creo que es muy importante que la carrera de astronomía a nivel nacional se plantee como un balance entre lo teórico y lo práctico, que no se le de mayor

importancia a la cantidad de información, sino que vaya a la par de mayor cantidad de recursos y herramientas.

8. Anexo II: Encuesta Mauricio Baeza estudiante de pregrado en Astronomía

1. ¿Cuál es tu background académico?

Astrónomo y Licenciado en astronomía con minor en Ciencia de Datos. Colaborador de investigación en astrofísica extragaláctica.

2. ¿Qué conocimientos previos tienes sobre los rayos gamma y su uso en la astronomía?

Es el tipo de emisión electromagnética más energética, emitida exclusivamente por los objetos y eventos más extremos del universo, por ejemplo quásares, púlsares y estrellas de neutrones.

Se observa mediante telescopios espaciales especializados en óptica y sensores, y mediante arreglos en tierra que captan la lluvia de partículas que generan en la atmósfera.

3. ¿Conoces algo acerca del satélite Fermi?

Sólo lo general dado que estoy interesado en bajas energías, en particular que es un telescopio espacial dedicado a rayos gamma.

4. ¿Conoces algo sobre el proyecto SWGO 2030?

Escuché de este mediante la prensa, entiendo que es un arreglo de rayos gamma a instalarse en la puna de Atacama junto con ALMA, el primero de su tipo en suelo nacional.

5. ¿Consideras que los recursos educativos que tuviste disponibles son suficientes para aprender sobre astrofísica de altas energías y procesar datos en Python o R?

Sí, considero que mis conocimientos son suficientes para aprender sobre altas energías y procesar datos en Python. Sin embargo nunca he visto o trabajado con observaciones de rayos gamma.

6. ¿Qué herramientas o recursos consideras que son aún necesarios en los programas académicos de Chile respecto a la materia señalada en el punto anterior?

Se podría incluir en los programas de licenciatura el trabajar de forma práctica con un espectro más amplio de observaciones, incluyendo rayos gamma o rayos X.

9. Anexo III: Descripción de atributos EventList

ENERGY: Energía del evento en electronvoltios (eV).

RA: Ascensión recta (Right Ascension) del evento en coordenadas J2000.

DEC: Declinación (Declination) del evento en coordenadas J2000.

L: Longitud galáctica del evento.

B: Latitud galáctica del evento.

THETA: Ángulo de incidencia del evento respecto a la dirección del detector.

PHI: Ángulo azimutal del evento.

ZENITH_ANGLE: Ángulo cenital del evento, es el ángulo entre la dirección vertical y la línea de visión del evento.

EARTH_AZIMUTH_ANGLE: Ángulo azimutal en relación con la Tierra, medido desde el centro del campo de visión del telescopio.

TIME: Momento en que ocurrió el evento, generalmente en segundos desde una referencia temporal específica (por ejemplo, el tiempo en el que comenzó la observación).

EVENT_ID: Identificador único asignado a cada evento.

RUN_ID: Identificador del "run" o conjunto de datos, que corresponde a una secuencia de observaciones o eventos.

RECON_VERSION: Versión del algoritmo de reconstrucción utilizado para procesar los datos del evento.

CALIB_VERSION: Versión del software de calibración utilizado para procesar los datos.

EVENT_CLASS: Clase del evento, que puede describir diferentes categorías de eventos (como muones o fotones).

EVENT_TYPE: Tipo del evento (por ejemplo, un fotón o un evento de partículas cargadas).

CONVERSION_TYPE: Tipo de conversión del evento, que puede hacer referencia a cómo la partícula detectada se convirtió en señal dentro del detector.

LIVETIME: Tiempo efectivo de observación durante el cual se detectaron eventos, considerando la eficiencia del instrumento.

DIFRSP0: Respuesta diferencial en el canal 0 del detector, que representa la cantidad de señal detectada para ese canal específico.

DIFRSP1: Respuesta diferencial en el canal 1 del detector.

DIFRSP2: Respuesta diferencial en el canal 2 del detector.

DIFRSP3: Respuesta diferencial en el canal 3 del detector.

DIFRSP4: Respuesta diferencial en el canal 4 del detector.

10. Anexo III: Descripción de atributos catálogo Fermi 3FHL

Source_Name: Nombre de la fuente según el catálogo 3FHL. Este es un identificador único para cada fuente.

RAJ2000: Ascensión recta (Right Ascension) de la fuente en el sistema de coordenadas J2000, medida en grados.

DEJ2000: Declinación (Declination) de la fuente en el sistema de coordenadas J2000, medida en grados.

GLON: Longitud galáctica de la fuente, medida en grados en el sistema de coordenadas galácticas.

GLAT: Latitud galáctica de la fuente, medida en grados en el sistema de coordenadas galácticas.

Conf_95_SemiMajor: Radio semilargo de la región de confianza del 95%, en grados. Describe la extensión de la fuente en la dirección más larga de la elipse de confianza.

Conf_95_SemiMinor: Radio semimensual de la región de confianza del 95%, en grados. Describe la extensión de la fuente en la dirección más corta de la elipse de confianza.

Conf_95_PosAng: Ángulo de posición de la región de confianza del 95%, medido en grados. Este ángulo describe la orientación de la elipse de confianza.

ROI_num: Número de la Región de Interés (ROI) en la que la fuente está localizada. Esto puede corresponder a una determinada región del cielo donde se agrupan fuentes.

Signif_Avg: Significancia media de la fuente, generalmente expresada en unidades de sigma. Indica cuán probable es que la fuente sea real y no un exceso aleatorio en el mapa.

Pivot_Energy: Energía pivote en unidades de GeV. Es la energía en la cual el modelo espectral (por ejemplo, un modelo de potencia) se define, proporcionando un punto de referencia para la descripción espectral de la fuente.

Flux_Density: Densidad de flujo de la fuente, generalmente expresada en unidades de $\text{cm}^{-2} \text{s}^{-1} \text{GeV}^{-1}$. Representa el número de fotones por unidad de área, tiempo y energía en la energía de referencia.

Unc_Flux_Density: Incertidumbre en la densidad de flujo. Mide la incertidumbre asociada a la estimación de la densidad de flujo de la fuente.

Flux: Flujo total de la fuente en unidades de $\text{cm}^{-2} \text{s}^{-1}$. Es el número total de fotones detectados por unidad de área y tiempo, integrando sobre todas las energías observadas.

Unc_Flux: Incertidumbre en el flujo total de la fuente.

Energy_Flux: Flujo de energía de la fuente, representado en unidades de $\text{erg cm}^{-2} \text{s}^{-1}$. Es el flujo total integrado sobre las energías.

Unc_Energy_Flux: Incertidumbre en el flujo de energía.

Signif_Curve: Significancia de la curva de luz. Indica si la fuente presenta variabilidad en su emisión, basada en el análisis de la curva de luz.

SpectrumType: Tipo de espectro de la fuente (por ejemplo, potencia, log-parabólica, etc.).

Spectral_Index: Índice espectral de la fuente, que describe cómo la intensidad de la fuente varía con la energía. En un modelo de potencia, el índice espectral es la pendiente de la distribución de fotones.

Unc_Spectral_Index: Incertidumbre en el índice espectral.

beta: Parámetro beta asociado al modelo espectral (usualmente en modelos log-parabólicos), que describe la curvatura del espectro.

Unc_beta: Incertidumbre en el parámetro beta.

PowerLaw_Index: Índice de la ley de potencias. Este índice describe la pendiente de un espectro de potencia, y es una medida fundamental en astrofísica de rayos gamma.

Unc_PowerLaw_Index: Incertidumbre en el índice de la ley de potencias.

Flux_Band: Flujo en el rango de energías específico. Esto indica el flujo medido en un rango de energía particular, lo cual puede ser útil para fuentes de energía variable.

Unc_Flux_Band: Incertidumbre en el flujo en el rango de energía.

nuFnu: $\nu F(\nu)$ es una medida del flujo energético. Representa el valor de $\nu F(\nu)$, donde $F(\nu)$ es el flujo de energía y ν es la frecuencia. Es útil para describir la distribución de energía de la fuente.

Sqrt_TS_Band: Raíz cuadrada de la estadística de prueba (TS) en el rango de energía. Esta es una medida de la significancia de la fuente en un rango de energía específico.

Npred: Número predicho de fotones de la fuente en una región dada o en el mapa de observación.

HEP_Energy: Energía de alta energía probable de la fuente. Esto podría referirse a la energía en la que la fuente muestra una emisión predominante de rayos gamma.

HEP_Prob: Probabilidad de alta energía de la fuente. Relacionado con la probabilidad de que la fuente emita en altas energías.

Variability_BayesBlocks: Variabilidad de bloques de Bayes. Representa una medida estadística de la variabilidad temporal de la fuente, usando un análisis de bloques de Bayes para detectar cambios en el flujo de la fuente.

Extended_Source_Name: Nombre de una fuente extendida asociada, si es relevante. Esto indica si la fuente observada tiene una extensión significativa en el cielo.

ASSOC_GAM: Asociación con una fuente de rayos gamma conocida.

TEVCAT_FLAG: Indicador TEVCAT. Puede indicar si la fuente está asociada con una fuente catalogada en el TEVCAT, que es un catálogo de fuentes de rayos gamma de muy alta energía.

ASSOC_TEV: Asociación con una fuente TeV. Indica si la fuente está asociada con una fuente catalogada en el dominio de rayos gamma de muy alta energía (TeV).

CLASS: Clasificación de la fuente (por ejemplo, AGN, pulsar, etc.).

ASSOC1, ASSOC2: Fuentes asociadas (posiblemente múltiples) con la fuente en cuestión.

ASSOC_PROB_BAY: Probabilidad de asociación de Bayes. La probabilidad de que la fuente observada esté asociada con una fuente conocida basada en un análisis bayesiano.

ASSOC_PROB_LR: Probabilidad de asociación de likelihood ratio. La probabilidad de asociación calculada mediante un enfoque de razón de verosimilitud.

Redshift: Desplazamiento al rojo de la fuente, que es una medida de la expansión del universo y la velocidad a la que la fuente se aleja de la Tierra.

NuPeak_obs: Frecuencia observada de $\nu F(\nu)$. La frecuencia en la que el flujo energético $\nu F(\nu)$ alcanza su valor máximo para la fuente observada.

11. Autoevaluación/Coevaluación

Integrante	Coevaluación Carlos	Coevaluación Jorge	Coevaluación Víctor	Comentarios del alumno (autoevaluación)
Carlos Bustamante	6,5 (autoevaluación)	7	7	Creo que mis compañeros hicieron un excelente trabajo. Respecto a mi desempeño, no quedé del todo conforme con la calidad de mi parte del entregable, podría ser porque las IRF son un tema complejo y grande.
Jorge Morgado	7	6,5 (autoevaluación)	7	La coordinación fue muy buena, un buen team, por mi lado, subestimé la cantidad de tiempo necesario para entrenar los modelos, lo que me llevó a modificar sobre la marcha el trabajo, por falta de poder de procesamiento.
Víctor Zúñiga	7	7	6,5 (autoevaluación)	Creo que logramos sortear gracias al trabajo en equipo las dificultades técnicas inherentes del proyecto, sin embargo tuve que acotar bastante mi análisis por lo complejo que se tornó.

