

Estimación de Energía de Rayos Cósmicos para el Observatorio HAWC aplicando algoritmos de Inteligencia Artificial



Jorge Alfredo Jaimes Teherán – jjalfredo68@gmail.com

Ibrahim Torres, Tomás Capistrán, Christian Sarmiento Cano, Luis Nuñez.

Universidad Industrial de Santander
Instituto Nacional de Astrofísica, Óptica y Electrónica
Università degli Studi di Torino

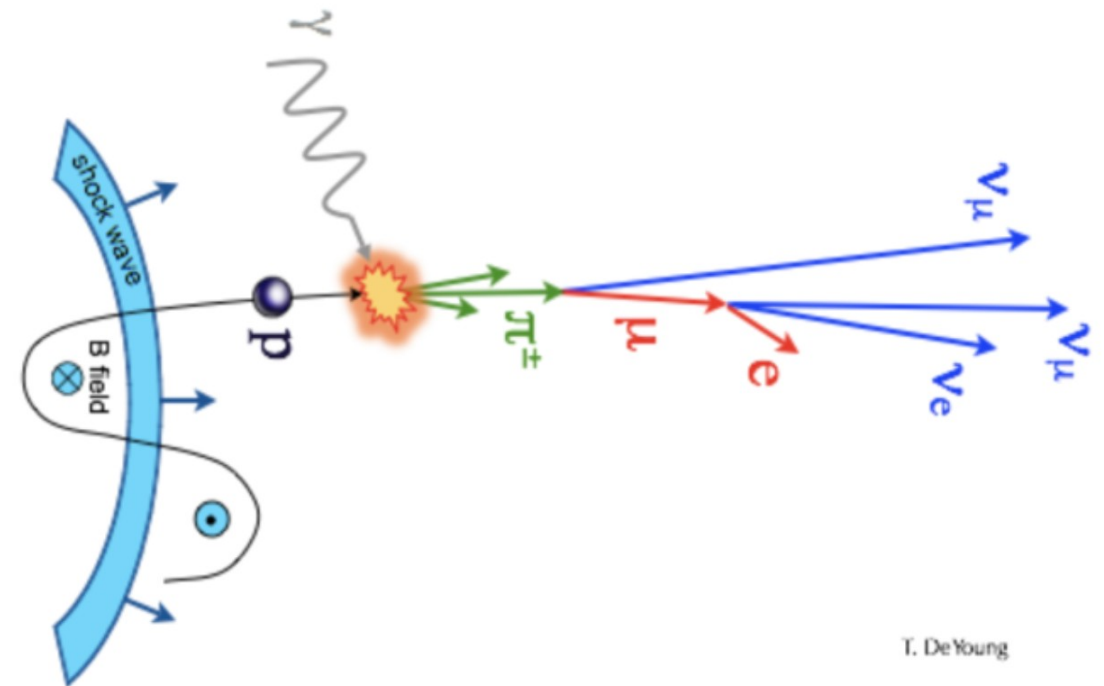


Rayos cósmicos

- Núcleos (hidrógeno,..., hierro)
- Remanentes de supernovas, púlsares y agujeros negros
- Interactúan con campos magnéticos y eléctricos

Rayos gamma

- Son fotones
- Explosiones de supernovas, fusión de estrellas de neutrones, AGN's, entre otros
- Viajan en línea recta

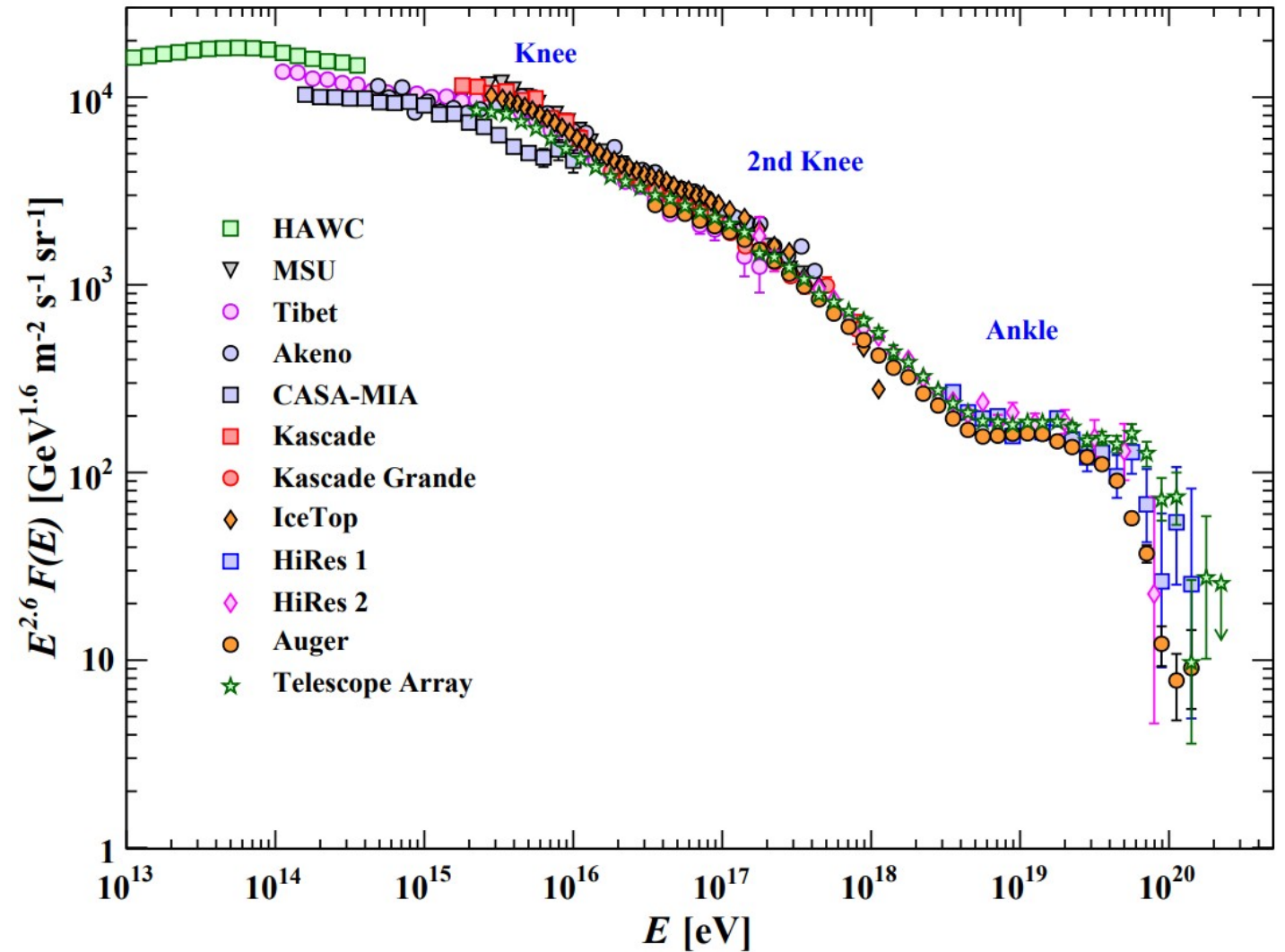


T. DeYoung



Espectro de rayos cósmicos

- La rodilla
- Segunda rodilla
- Tobillo
- Supresión

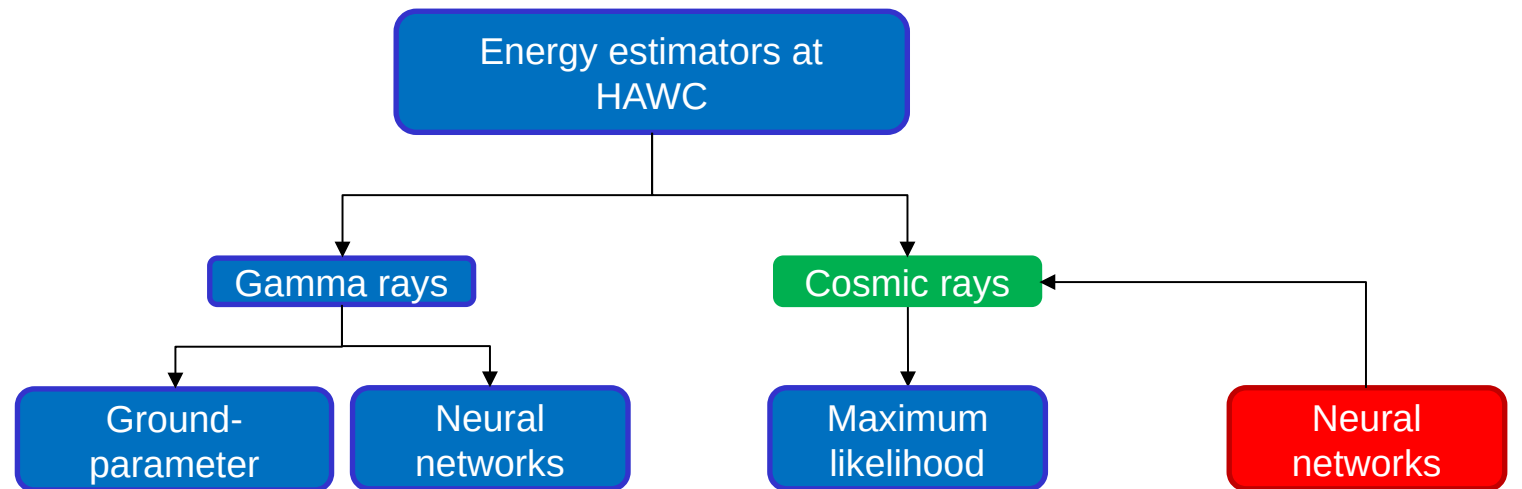
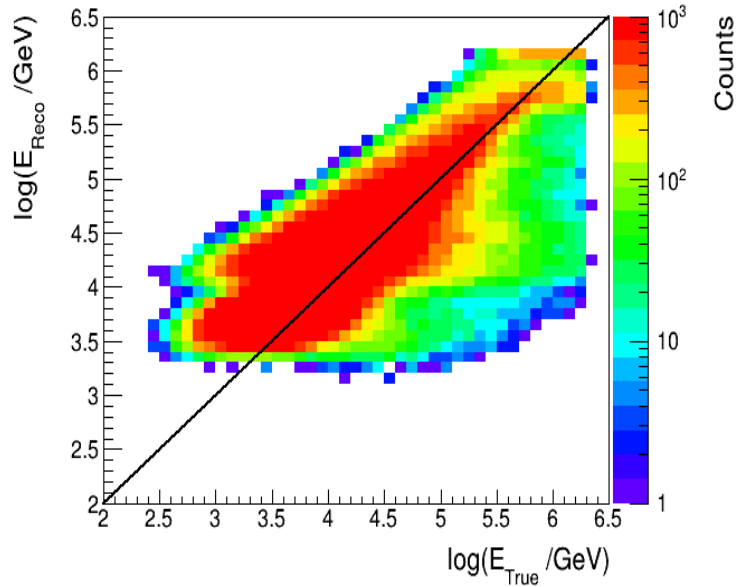


HAWC

- Situado en el volcán Sierra Negra
- 100 GeV hasta ~ 1 PeV (RC)
- 300 Detectores Cherenkov en agua
- 4 foto multiplicadores por tanque



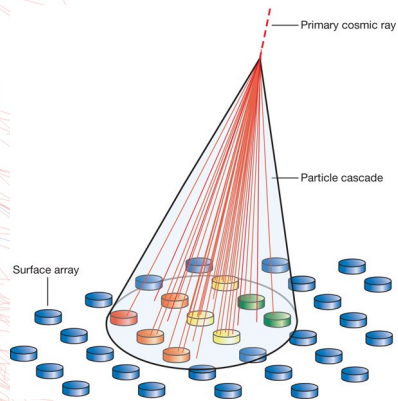
Importancia de la estimación de energía de los rayos cósmicos



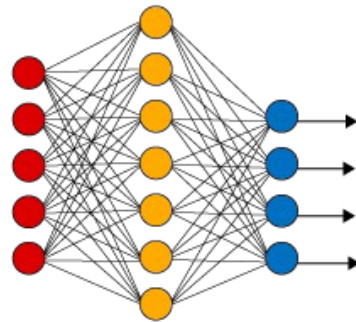
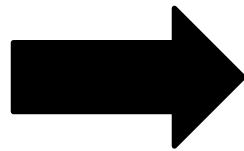
Construcción del modelo

Entrenamiento del modelo con eventos simulados

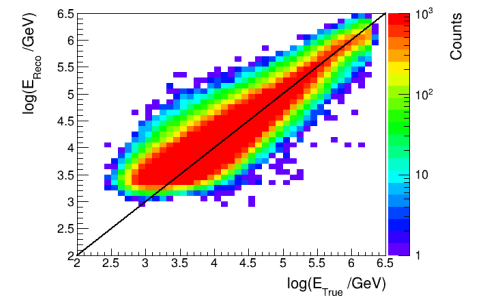
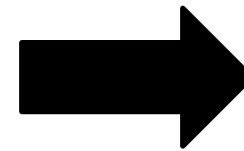
Entrenamos en **simulaciones**, pero usamos en **datos reales**



Datos simulados



Red neuronal

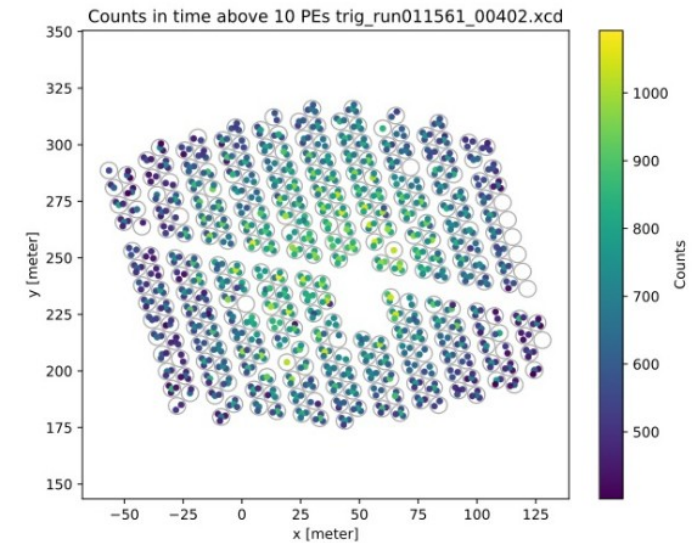
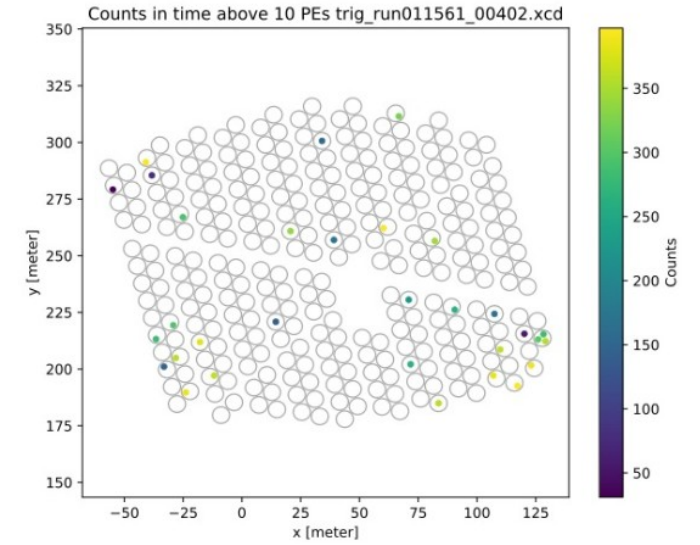
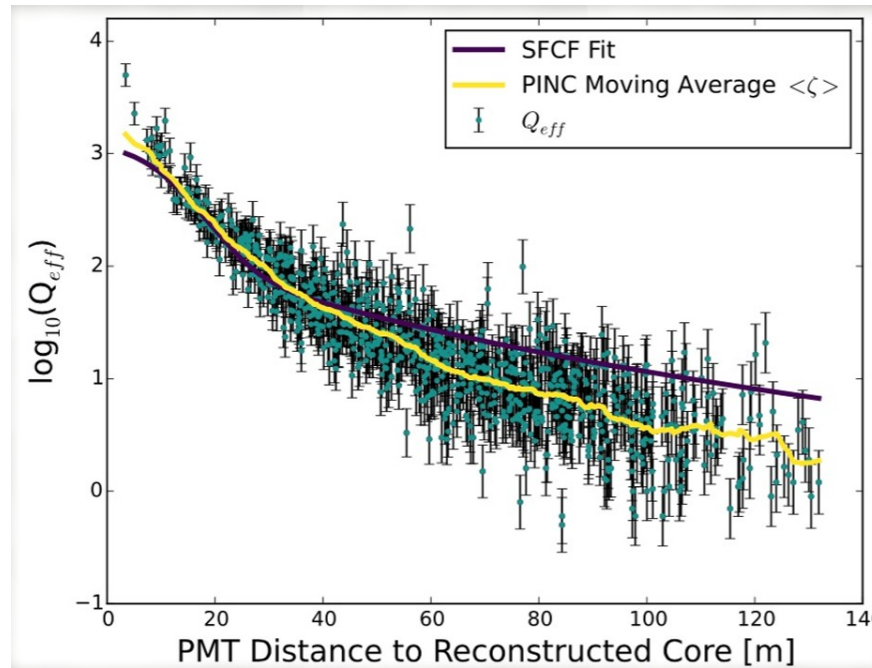


Reconstrucción de energía

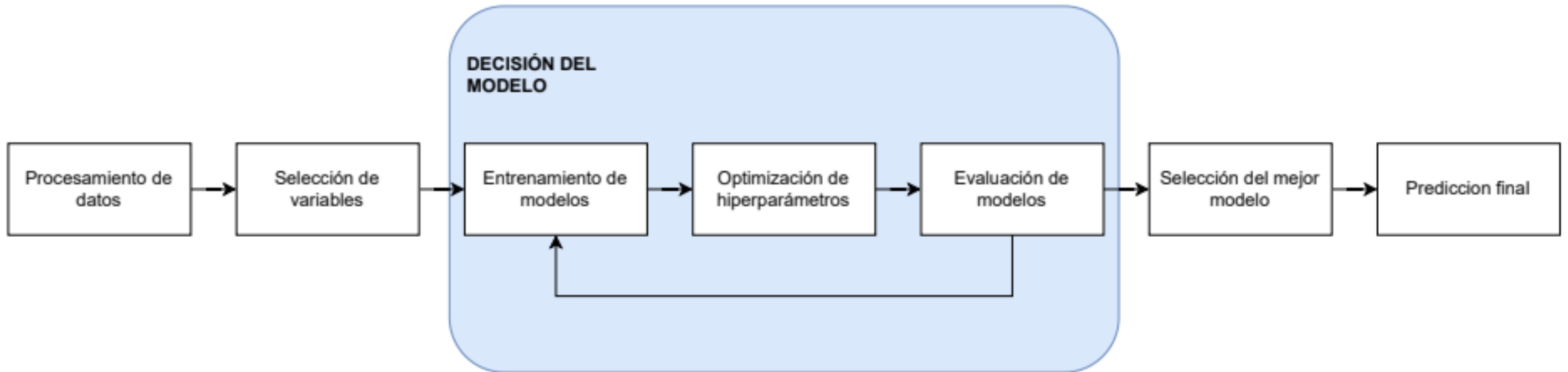
Variables físicas



- fHit
- fTank
- Lateral_Amp
- CoreD
- CZenith
- fAnnulusCharge (9 variables)
- 22 variables +



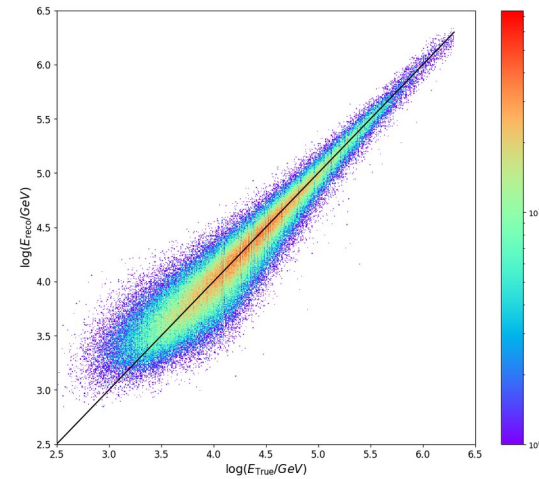
Pipeline de entrenamiento



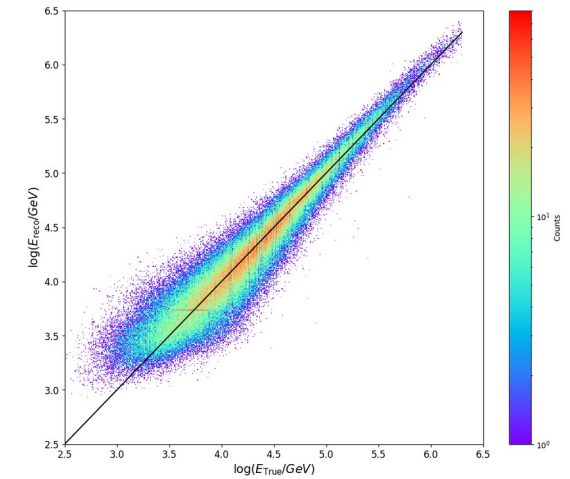


Modelo	Conjunto	MSE	MAE	R^2
MLP	Entrenamiento	0.0276	0.1212	0.9182
	Prueba	0.0279	0.1218	0.9109
XGBoost	Entrenamiento	0.0250	0.1165	0.9201
	Prueba	0.0281	0.1225	0.9101
LightGBM	Entrenamiento	0.0262	0.1192	0.9163
	Prueba	0.0282	0.1228	0.9099
CNN	Entrenamiento	0.0365	0.1518	0.8832
	Prueba	0.0367	0.1522	0.8825

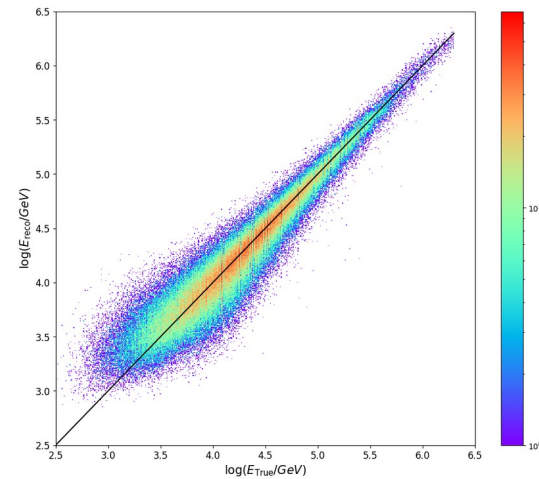
MLP



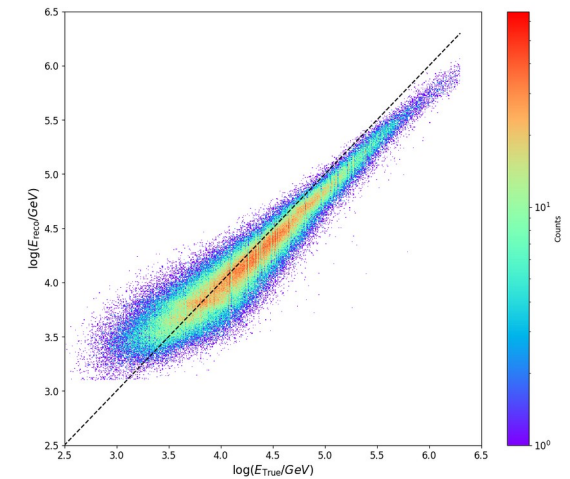
XGBOOST



LGBM



CNN

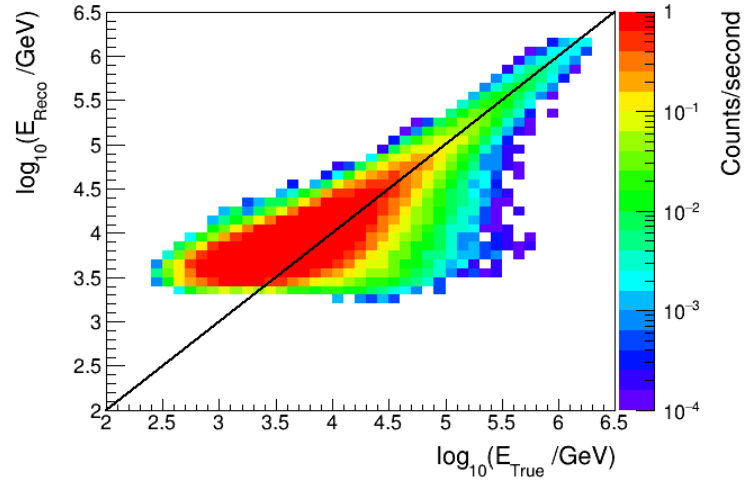


Energía reconstruida

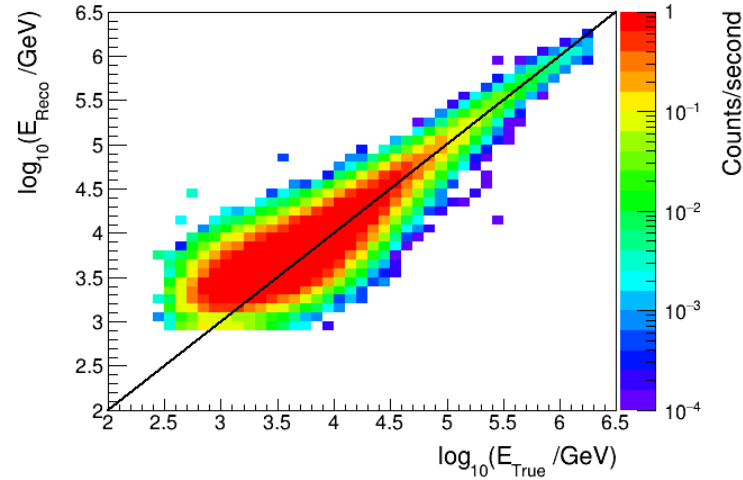


Protón

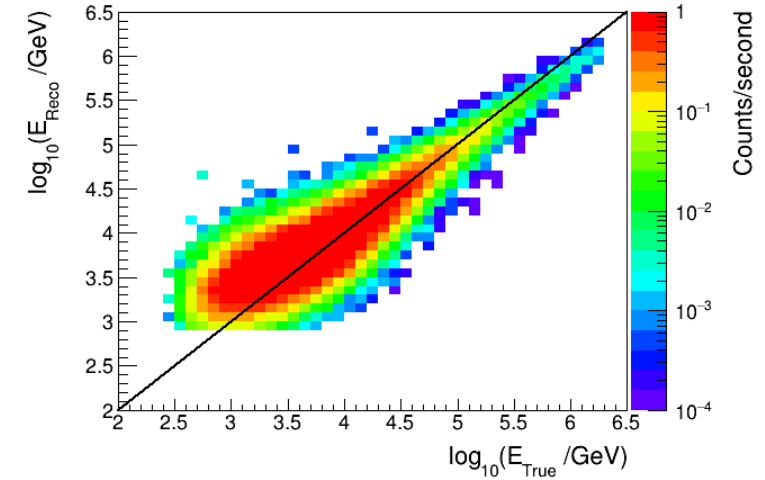
Likelihood



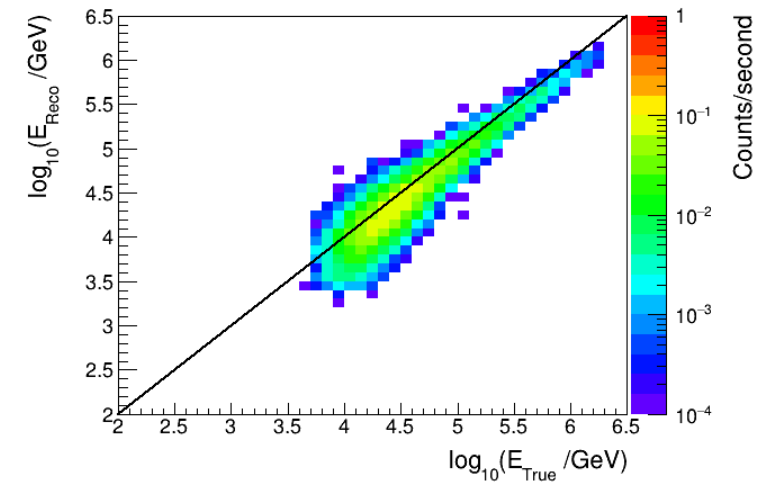
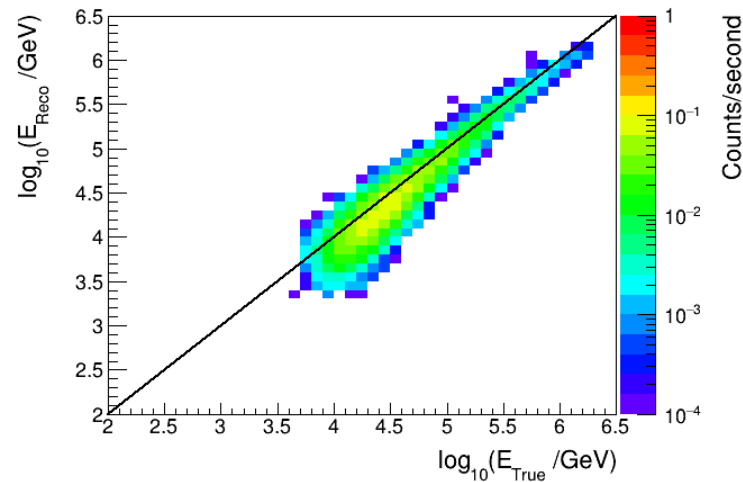
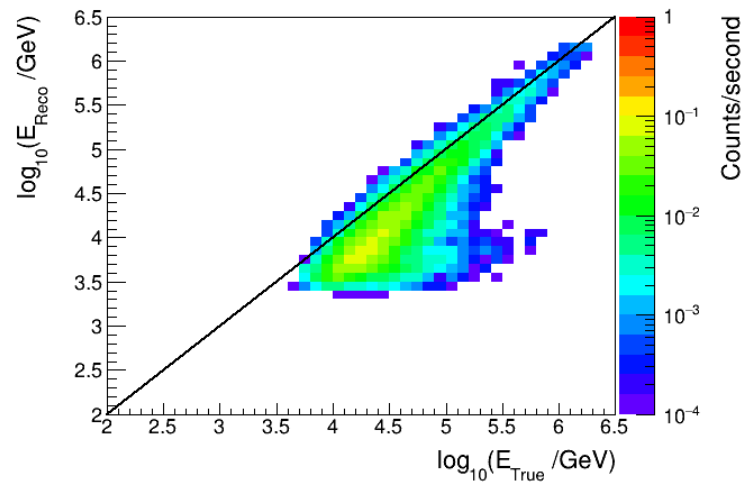
MP4CE



MP4SE



Hierro



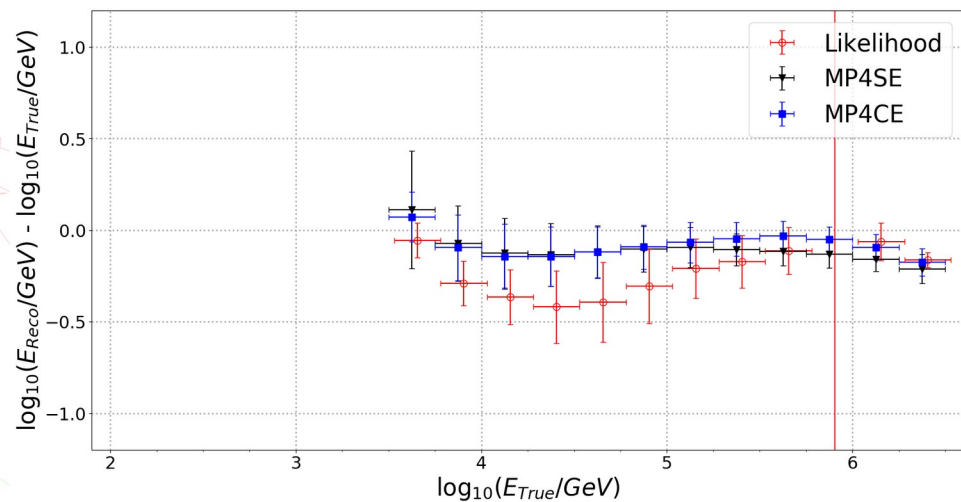
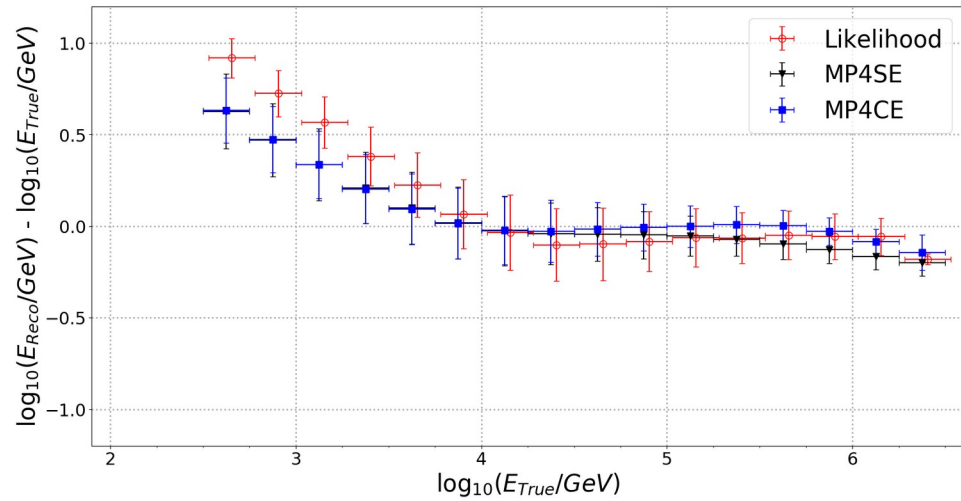
Comprobación de los métodos



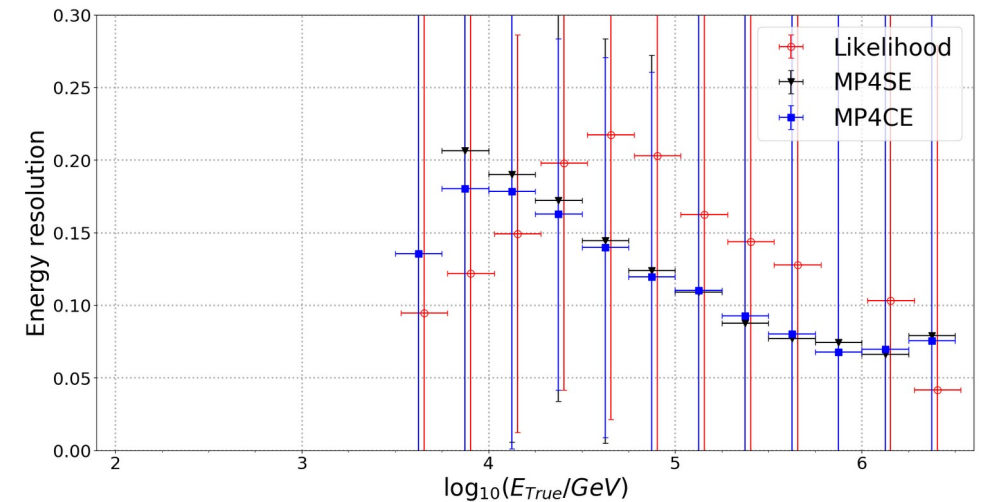
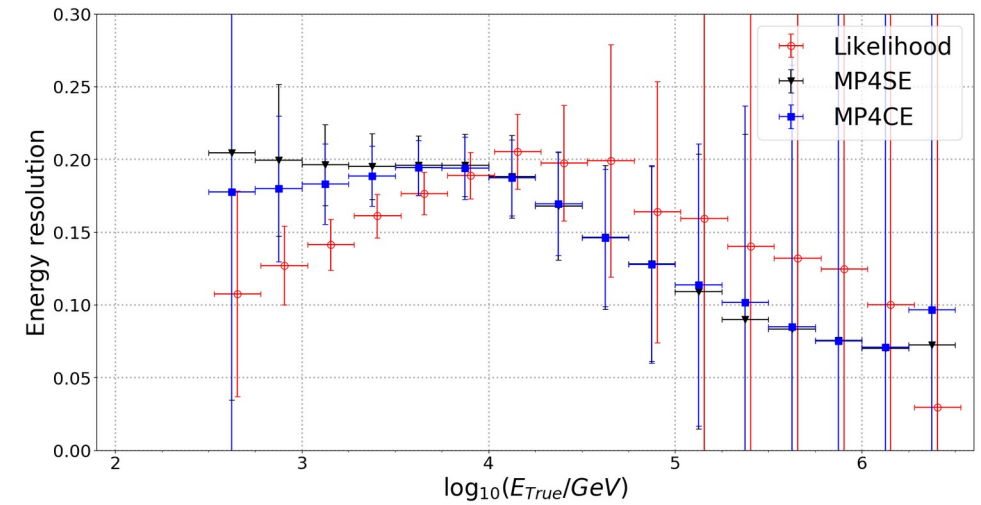
Protón

Hierro

Bias



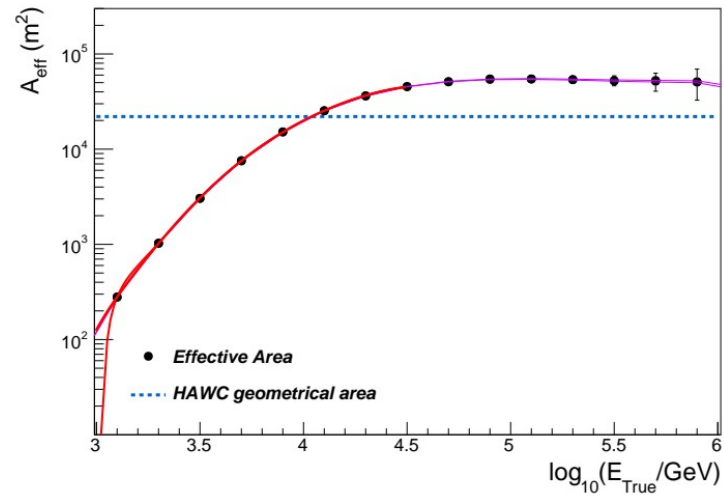
Energy resolution



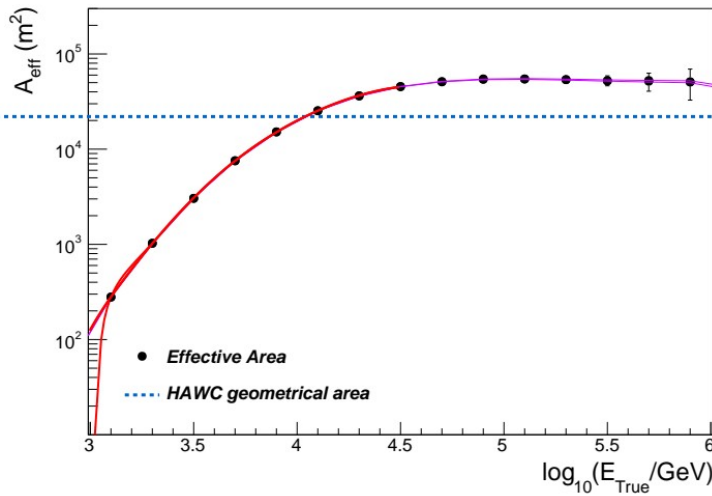
Área efectiva y matriz de respuesta



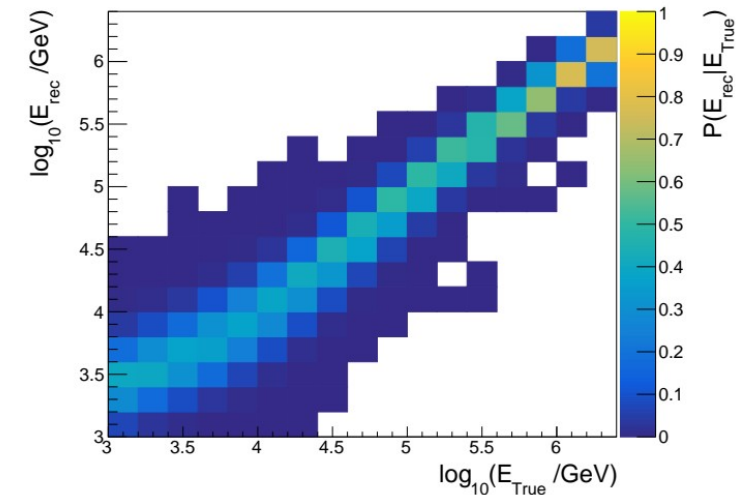
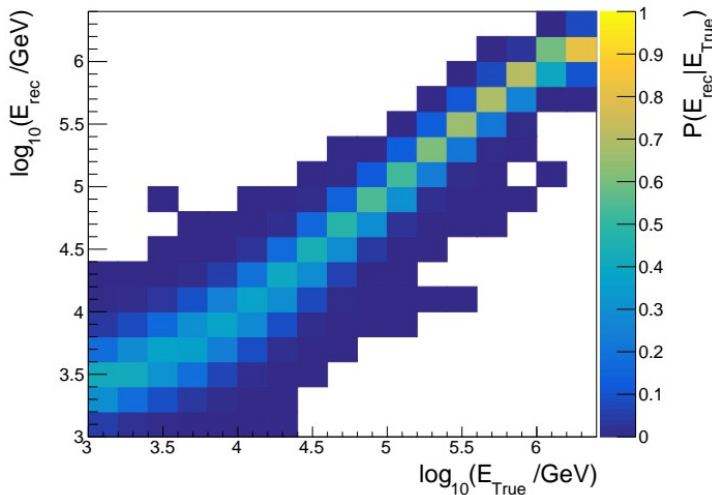
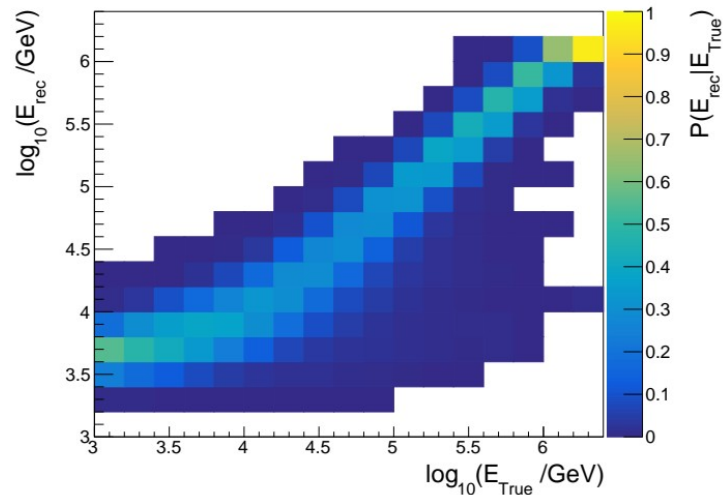
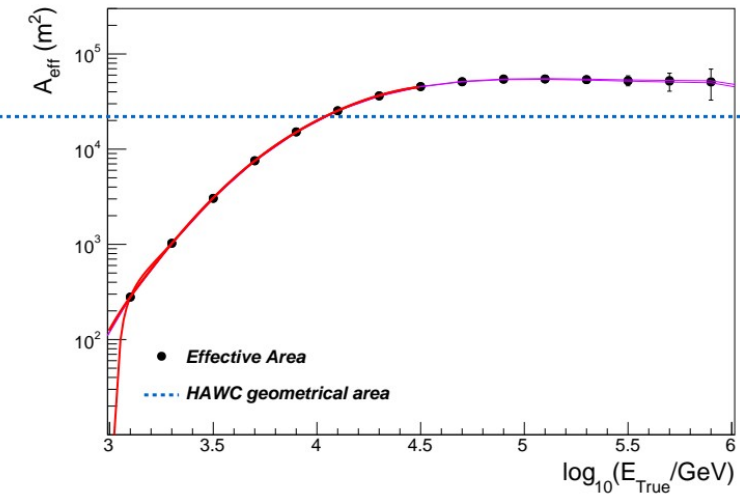
Likelihood



MP4CE



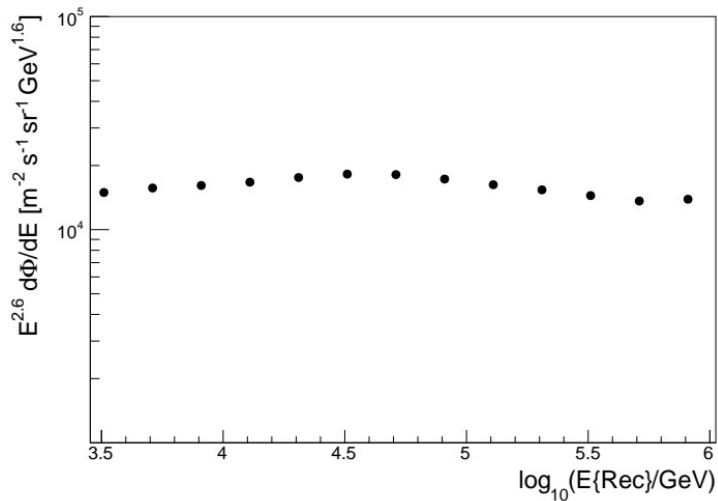
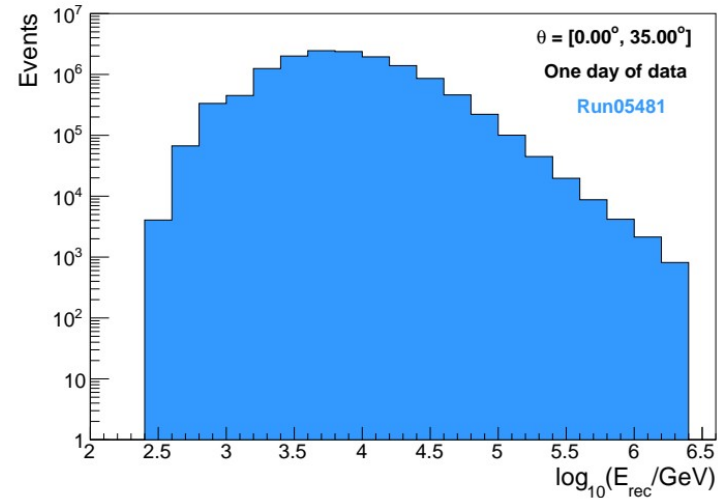
MP4SE



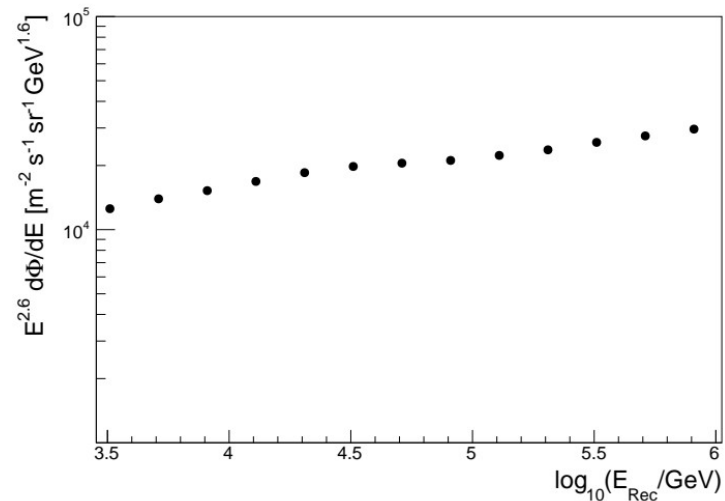
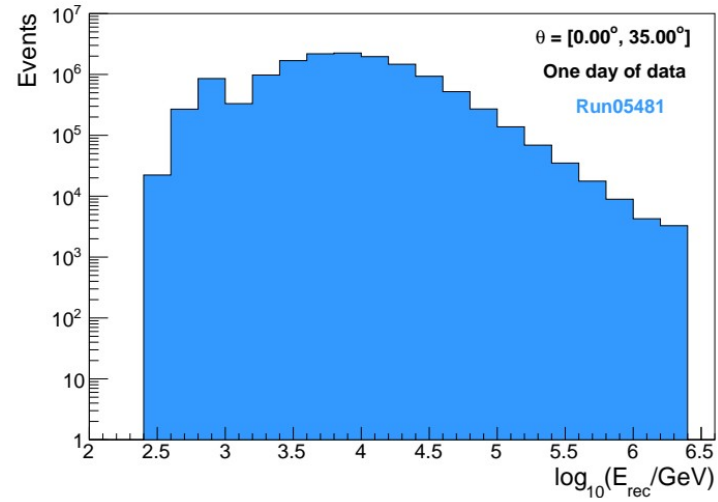
Un día de espectro



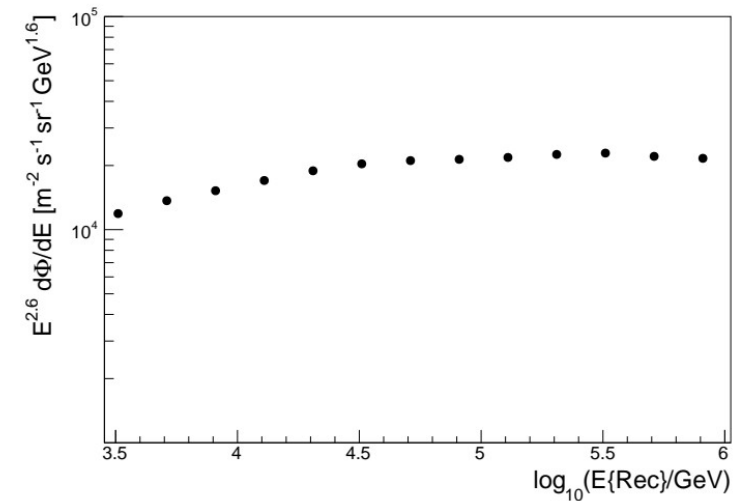
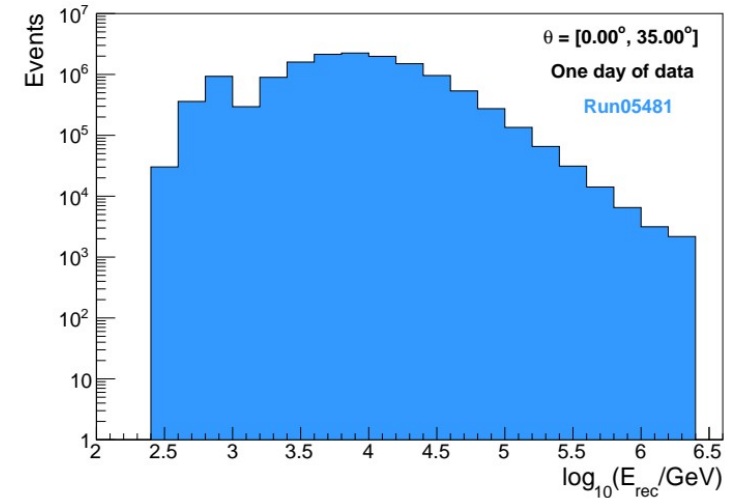
Likelihood



MP4CE



MP4SE





El modelo MP4CE demostró ser el más preciso y consistente en la estimación de energía de los rayos cósmicos, superando significativamente al método tradicional Likelihood y al modelo MP4SE.

Aunque MP4CE mostró excelentes resultados en precisión, presentó una ligera sobreestimación del flujo a altas energías, lo que indica la necesidad de ajustes adicionales en simulaciones y calibraciones.

Este trabajo valida el uso de redes neuronales profundas como herramientas clave para mejorar la reconstrucción de energía en experimentos como HAWC, abriendo oportunidades para explorar arquitecturas avanzadas como GNN, PINN y transformadores en estudios futuros.



¡Gracias!



GRUPO
HALLEY





rec.nChAvail: Número de canales disponibles para la detección de eventos.

rec.nHitTot: Número total de hits detectados en todo el observatorio.

rec.nHit: Número de hits detectados en el evento actual.

rec.nHitSP10: Número de hits dentro de una ventana temporal de 10 ns.

rec.nHitSP20: Número de hits dentro de una ventana temporal de 20 ns.

rec.nTankAvail: Número de tanques de agua disponibles para la detección.

rec.nTankHitTot: Número total de tanques que detectaron hits durante el evento.

rec.nTankHit: Número de tanques que detectaron hits en el evento actual.

rec.windowHits: Número de hits detectados en una ventana temporal definida.

rec.planeNDOF: Grados de libertad en el ajuste del plano de la onda del evento.

rec.mPFnHits: Número de hits asociados con el método de partición de planos (PlaneFit).

rec.mPFp0nAssign: Número de hits asignados al primer paso de reconstrucción del ajuste de plano.

rec.mPFp1nAssign: Número de hits asignados al segundo paso del ajuste de plano.

rec.coreFiduScale: Escala de fiducialidad para la posición del núcleo del evento.

rec.zenithAngle: Ángulo cenital del evento, indicando la dirección del rayo cósmico con respecto a la vertical.

rec.planeChi2: Valor de Chi-cuadrado para la calidad del ajuste del plano de la onda.

rec.fAnnulusCharge0: Carga detectada en el primer anillo concéntrico del detector.

rec.fAnnulusCharge1: Carga detectada en el segundo anillo concéntrico.

rec.fAnnulusCharge2: Carga detectada en el tercer anillo concéntrico.

rec.fAnnulusCharge3: Carga detectada en el cuarto anillo concéntrico.

rec.fAnnulusCharge4: Carga detectada en el quinto anillo concéntrico.

rec.fAnnulusCharge5: Carga detectada en el sexto anillo concéntrico.

rec.fAnnulusCharge6: Carga detectada en el séptimo anillo concéntrico.

rec.fAnnulusCharge7: Carga detectada en el octavo anillo concéntrico.

rec.fAnnulusCharge8: Carga detectada en el noveno anillo concéntrico.

rec.logMaxPE: Logaritmo del máximo número de electrones fotoemisores (PE) detectados en el evento.

rec.logNPE: Logaritmo del número total de electrones fotoemisores detectados.

rec.CxPE40: Número estimado de electrones fotoemisores en un área de 40 metros alrededor del núcleo.

rec.logGP: Logaritmo del ajuste de la función de probabilidad geométrica.

rec.mPFp0Weight: Peso asignado al primer paso en el ajuste de partición de planos.

rec.mPFp1Weight: Peso asignado al segundo paso en el ajuste de partición de planos.

rec.disMax: Distancia máxima entre el núcleo reconstruido y los hits detectados.

rec.TankLHR: Likelihood ratio asociado a los tanques que detectaron hits.

rec.LHLatDistFitXmax: Parámetro de ajuste para la distancia lateral máxima en el evento.

rec.LHLatDistFitEnergy: Ajuste de la energía basado en la distribución lateral del evento.

rec.LHLatDistFitGoF: Calidad del ajuste de la distribución lateral del evento.



MP4CE

1. rec.LDFamp
2. rec.fAnnulusCharge0
3. rec.fAnnulusCharge1
4. rec.fAnnulusCharge2
5. rec.fAnnulusCharge3
6. rec.fAnnulusCharge4
7. rec.fAnnulusCharge5
8. rec.fAnnulusCharge6
9. rec.fAnnulusCharge7
10. rec.fAnnulusCharge8
11. rec.nHitSP20
12. rec.nTankHitTot
13. rec.logGPV2
14. rec.coreFitUnc
15. rec.nHitTot
16. rec.logNNEnergyV2
17. rec.logCoreAmplitude
18. rec.LDFChi2
19. rec.nHit
20. rec.logNPE
21. rec.gammalheEnergy
22. rec.logMaxPE
23. rec.nTankHit
24. rec.windowHits
25. rec.nHitSP10
26. rec.PINC

27. rec.CxPE40
28. rec.protonlheLLH
29. rec.planeNDOF
30. rec.coreFiduScale
31. rec.planeChi2
32. rec.LDFAge
33. rec.SFCFNDOF
34. rec.SFCFChi2
35. rec.gammalheLLH

MP4SE

1. rec.LDFamp
2. rec.fAnnulusCharge0
3. rec.fAnnulusCharge1
4. rec.fAnnulusCharge2
5. rec.fAnnulusCharge3
6. rec.fAnnulusCharge4
7. rec.fAnnulusCharge5
8. rec.fAnnulusCharge6
9. rec.fAnnulusCharge7
10. rec.fAnnulusCharge8
11. rec.nHitSP20
12. rec.nTankHitTot
13. rec.coreFitUnc
14. rec.nHitTot
15. rec.logCoreAmplitude
16. rec.LDFChi2
17. rec.nHit
18. rec.logNPE
19. rec.logMaxPE
20. rec.nTankHit
21. rec.windowHits
22. rec.nHitSP10
23. rec.PINC
24. rec.CxPE40

25. rec.planeNDOF
26. rec.coreFiduScale
27. rec.planeChi2
28. rec.LDFAge
29. rec.SFCFNDOF
30. rec.SFCFChi2
- rec.LDFamp



$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$$

La función de verosimilitud para un conjunto de datos $X = \{x_1, x_2, \dots, x_n\}$ es:

$$L(\mu, \sigma^2 | X) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right)$$

En este caso, podemos derivar y resolver para μ y σ^2 que maximizan la verosimilitud. Esto nos llevará a los estimadores de máxima verosimilitud para una distribución normal:

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i \quad (\text{media muestral})$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{\mu})^2 \quad (\text{varianza muestral})$$