



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE CIENCIAS

Simulaciones Monte Carlo para optimizar la búsqueda de
señales de muones en el observatorio HAWC

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

Físico

PRESENTA:

José Roberto Angeles Camacho

TUTOR

Dr. Hermes León Vargas

Ciudad Universitaria, Cd. Mx. 2018





Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Hoja de datos del Jurado

1. Datos del Alumno

Angeles
Camacho
José
Roberto
Universidad Nacional Autónoma de México
Facultad de Ciencias
Física
310024952

2. Datos del Tutor

Dr.
Hermes
León
Vargas

3. Datos del Sinodal 1

Dr.
Ernesto José María de la Salette
Belmont
Moreno

4. Datos del Sinodal 2

Dr.
Nissim Illich
Fraiya
Cabrera

5. Datos del Sinodal 3

Dr.
David
García
Gudiño

6. Datos del Sinodal 4

Dr.
Raúl Josué
Hernández
Hernández

7. Datos del trabajo escrito

Simulaciones Monte Carlo para optimizar la búsqueda de señales de muones en el observatorio
HAWC
130 p.
2018

Agradecimientos

Me gustaría agradecer a Dios y a toda mi familia. A mi padre Roberto Angeles Hernández y a mi madre Candelaria Camacho López que siempre me ha apoyado y es la razón principal de que terminara mis estudios universitarios. También quisiera mencionar el apoyo de mis tíos Alejandro Nieto y Carmen Camacho López que siempre han estado conmigo.

Agradezco a mi hermano Jesús Emmanuel Angeles Camacho, en especial por dejarme su computadora. Esta fue una pieza fundamental para la elaboración de este proyecto.

Además quisiera agradecer a mi asesor Hermes León Vargas por su paciencia y por la oportunidad de trabajar con él desde el servicio social. También agradezco a todos los sinodales por leer este trabajo y hacer las correcciones pertinentes.

Este trabajo fue realizado gracias al apoyo de los proyectos CONACyT 254964 y UNAM-DGAPA-PAPIIT IA102715.

Resumen

El observatorio HAWC (High Altitud Water Cherenkov) es un instrumento de investigación conformado por tanques que detectan rayos gamma de alta energía para estudiar al Universo. Una línea de investigación alternativa que se está estudiando en el Instituto de Física de la UNAM es utilizar a este observatorio como un detector de neutrinos ultra energéticos. Para ello se buscan señales de leptones cargados que vengan en la dirección en donde se encuentra el Pico de Orizaba (que rodea a este observatorio). Estas partículas se forman por la interacción entre los átomos de la montaña y los neutrinos que vienen en esa dirección. Pero antes de intentar buscar señales inducidas por neutrinos ultra energéticos, se debe caracterizar al ruido que proviene de cascadas atmosféricas generadas en la misma dirección. Este ruido en mayor parte está formado por muones que viajan de forma cuasi-horizontal y cuya dirección se puede aproximar si nos fijamos en los tanques activados durante su paso por el observatorio. En este trabajo se realizaron dos análisis empleando simulaciones Monte Carlo para estudiar a estas señales de ruido.

En el primer análisis, se encontró la región mas probable del tanque por la cual pudo haber pasado un muón que viaje de manera cuasi-horizontal. Esta región se definió para el caso de dos energías: 1 TeV y 10 TeV. Con esta información se puede mejorar el algoritmo que se emplea para aproximar la dirección de este tipo de señales hasta en un 63 %.

En el otro análisis se utilizó una red neuronal artificial para dar información sobre la energía inicial de muones individuales que se propagan de manera horizontal. Con el resultado de la red neuronal, se pudo dar un intervalo en donde se encuentra la energía del muón y que tiene una eficiencia de hasta un 96 %.

Índice general

Agradecimientos	III
Resumen	IV
Lista de figuras	VII
Lista de tablas	X
Introducción	1
1. Marco teórico	3
1.1. Rayos Cósmicos	3
1.1.1. Historia de los Rayos Cósmicos en México	5
1.2. El observatorio HAWC	6
1.2.1. Cascadas atmosféricas	6
1.2.2. Proceso de Detección en HAWC	8
1.2.3. Algunos resultados destacados obtenidos en HAWC	11
1.2.4. Detección de neutrinos en HAWC	11
2. Simulaciones en HAWCSIM	13
2.1. Software utilizado en HAWC	13
2.2. HAWCSIM: Aspectos Generales	16
2.2.1. GEANT4	16
2.2.2. CORSIKA	18
2.3. Descripción de cascadas	19
2.3.1. Método Monte Carlo	19
2.3.2. Modelo de Heitler	20
2.3.3. Ejemplo de cascada electromagnética en GEANT4	22
2.4. Instalación de HAWCSIM	24
2.4.1. Configuración de HAWCSIM	25
2.5. Entradas y Salidas en HAWCSIM	26
2.6. Ejemplo de una simulación en HAWCSIM	27
2.6.1. Procedimiento para hacer la simulación	27
2.6.2. Ejecución de la simulación	31
2.6.3. Ruido y número de eventos en la reconstrucción	33
3. Red Neuronal Artificial	35
3.1. Reconocimiento de números escritos a mano	35
3.2. Red Neuronal Biológica	36
3.3. Red Perceptrón	37
3.4. Red Perceptrón Multicapa	39
3.4.1. Estructura	42

3.4.2. Funcionamiento	45
3.4.3. Un ejemplo en Root	52
4. Primer análisis: Volumen Efectivo de Detección	57
4.1. Motivación: Reconstrucción de trazas	57
4.2. Descripción de las simulaciones	61
4.2.1. Simulación de muones horizontales	62
4.2.2. Simulación de muones verticales	63
4.2.3. Número y energías	65
4.2.4. Posición inicial	65
4.3. Variables a observar en la simulación	66
4.4. Resultados	67
4.4.1. Altura y radio del cilindro en una sola dirección	67
4.4.2. Altura del cilindro en todas las direcciones	70
4.4.3. Radio del cilindro en todas las direcciones	73
4.4.4. Volumen efectivo de detección	75
4.4.5. Prueba para el volumen efectivo de detección	78
4.4.6. Discusión de Resultados	80
5. Segundo análisis: Estimador de Energía para muones horizontales	84
5.1. Descripción de la prueba	84
5.1.1. Descripción de las simulaciones	84
5.1.2. Descripción de la red neuronal artificial	85
5.1.3. Datos para entrenar y probar a la red neuronal	93
5.2. Resultados y Análisis	94
5.2.1. Energía para traza de muones individuales	99
Conclusiones Generales	103
A) Número de muones usados en la simulación	104
B) Diferencia entre la carga del PMT central y uno periférico	106
C) Código para una red neuronal en ROOT	108
D) Método de propagación hacia atrás (backpropagation)	113
E) Volumen efectivo de detección μ^+ y consideraciones del porcentaje P_H	117
F) Tablas de intervalos de energías y volumen efectivo de Detección.	121
Bibliografía	130

Índice de figuras

1.1. Llegada de rayos cósmicos a la Tierra	3
1.2. Simposio de rayos cósmico de 1939 en Chicago	5
1.3. Imagen del observatorio HAWC	6
1.4. Esquema de una cascada atmosférica	8
1.5. Polarización de un dieléctrico por una partícula a baja velocidad	9
1.6. Polarización de un dieléctrico por una partícula a alta velocidad	9
1.7. Emisión de radiación Cherenkov	10
1.8. Mapa del cielo en rayos gamma	11
2.1. Reconstrucción de cascadas con AERIE	15
2.2. Definición de clases en GEANT4	18
2.3. Esquema del modelo de Heitler	20
2.4. Esquema de simulación en GEANT4	22
2.5. Gráfica de \hat{l} en función de la energía	23
2.6. Perfil longitudinal de cascada EM	24
2.7. Opciones disponibles en HAWCSIM	25
2.8. Representación gráfica de un ejemplo en HAWCSIM	28
2.9. Sistema de referencia del archivo survey.xml	30
2.10. Ejemplo del archivo de entrada en HAWCSIM	31
2.11. Representación del tiempo de detección en una simulación de HAWCSIM	32
2.12. Ventana que aparece al ejecutar TBrowser tb;	33
3.1. Números escritos a mano para entrenar una red neuronal	35
3.2. Neurona biológica	36
3.3. Neurona artificial Perceptrón	37
3.4. Neurona artificial para una compuerta NAND	39
3.5. Red neuronal para sumar 2 bits	40
3.6. Gráfica de la función sigmoide	41
3.7. Estructura de una red Perceptrón Multicapa	42
3.8. Red para reconocer números escritos a mano	44
3.9. Imagen de 28×28 pixeles de un número escrito a mano	44
3.10. Red neuronal para el reconocimiento del número 7 y 8	45
3.11. Representación de la activación de dos neuronas	46
3.12. Ejemplo de función de costo para una sola imagen	48
3.13. Ejemplo para minimizar la función de costo en 2D	49
3.14. Ejemplo para calcular $-\nabla C$	50
3.15. Descenso de gradiente estocástico en 3D	51
3.16. Estructura de red neuronal del ejemplo 3.3.4	53

3.17. Datos para entrenar y probar la red del ejemplo 3.3.4	54
3.18. Comparación entre las funciones $f(x, y)$ y $f(x, y)_{NN}$	55
4.1. Ejemplo de trazas de muones	58
4.2. Representación para el ángulo azimutal	58
4.3. Representación gráfica del ángulo β y α	59
4.4. Representación para el ángulo de elevación	60
4.5. Representación de muones verticales y horizontales atravesando un tanque	62
4.6. Dirección para la simulación de muones horizontales	63
4.7. Alturas para la simulación de muones horizontales	63
4.8. Dirección para la simulación de muones verticales	64
4.9. Separación respecto al eje de simetría para muones verticales	65
4.10. Región en el tanque en donde es más probable que se activen 4,3, ó 2 PMTs para 1 TeV	76
4.11. Región en el tanque en donde es más probable que se activen 4,3, ó 2 PMTs para 10 TeV	76
4.12. Región en el tanque en donde es más probable que se activen al menos 3 ó 2 PMTs para 1 y 10 TeV	77
4.13. Trayectoria del muón para la simulación de prueba para el volumen efectivo de detección	78
4.14. Vista desde el plano $x - y$ para los tanques activados en la simulación de prueba	79
4.15. Vista desde el plano $x - z$ para los tanques activados en la simulación de prueba	79
4.16. Representación del cono de luz para un muón enviado en H.I a 2 diferentes alturas	80
4.17. Representación del cono de luz para los tanques I20 y G20 de la simulación de prueba	82
5.1. Red neuronal para estimar la energía de muones horizontales	85
5.2. Ejemplo para explicar las variables de entrada de la red neuronal	86
5.3. Histograma de la carga dejada por el muón en la línea de tanques para 1 TeV	87
5.4. Gráfica del promedio de carga en toda la línea del tanque vs la energía	88
5.5. Histograma de la diferencia de carga entre tanques consecutivos para 1 TeV	90
5.6. Gráfica de diferencia de carga entre tanques consecutivos en contra de la energía	90
5.7. Histograma del promedio de detectores activados en un tanque para 1 TeV	92
5.8. Gráfica del promedio de tanques activados en contra de la energía	92
5.9. Histograma de la carga dejada por el muón en la línea de tanques para eventos pares e impares	94
5.10. Histograma de la energía calculada por la red neuronal para los datos correspondientes a la energía de 1 TeV	94
5.11. Comparación entre la energía que arrojo la red neuronal y la energía que se esperaba	95
5.12. Gráfica de la diferencia entre la energía esperada con la energía que calculo la red neuronal	96
5.13. Histograma del error porcentual para los datos de prueba que corresponden a una energía de 1 TeV	97
5.14. Gráfica que muestra el comportamiento de la desviación estándar para la energía promedio calculada por la red	98
5.15. Esquema de la simulacion usada en el apendice B)	106
5.16. Histograma del apéndice B)	106
5.17. Red con una sola neurona en cada capa	113
5.18. Relación entre las cantidades de la función de costo	114

5.19. Notación para una red neuronal	115
5.20. Región del tanque en donde es más probable que se activen al menos 3 ó 2 PMTs para μ^+	118
5.21. Región del tanque en donde es más probable que se activen 4,3 ó 2 PMTs para μ^+	118
5.22. Datos de Porcentajes de μ^- horizontales de 1 TeV	122
5.23. Datos de Porcentajes de μ^- verticales de 1 TeV	123
5.24. Datos de Porcentajes de μ^- horizontales de 10 TeV	124
5.25. Datos de Porcentajes de μ^- verticales de 10 TeV	125
5.26. Datos de Porcentajes de μ^+ horizontales de 1 TeV	126
5.27. Datos de Porcentajes de μ^+ verticales de 1 TeV	127

Índice de tablas

1.1. Principales acontecimientos para el descubrimiento de los rayos cósmicos.	4
2.1. Modelos hadrónicos usados en CORSIKA.	19
3.1. Tablas de verdad de compuertas AND y OR.	39
3.2. Tablas de verdad de la compuerta NAND.	40
3.3. Resultados para la suma de 2 bits	41
4.1. Ejemplo de la información observada para el volumen de detección	67
4.2. Porcentaje de 4 PMTs activados para muones en la dirección H.II con 1 TeV de energía	68
4.3. Porcentaje de 3 PMTs activados para muones en la dirección V.II con 1 TeV de energía	69
4.4. Posición inicial y final para definir la altura del cilindro	70
4.5. Posición inicial y final para definir el radio del cilindro	74
4.6. Ubicación de la región en donde es más probable que se activen 4, 3 ó 2 PMTs	75
4.7. Ubicación de la región en donde es más probable que se activen mínimo 3 ó 2 PMTs	77
4.8. Porcentajes de detectores activados para la simulación de prueba	78
4.9. Reconstrucción de trazas usando las regiones en donde mínimo se activa 3 ó 2 PMTs	83
5.1. Promedio y RMS para los histograma de la variable C_a	87
5.2. Intervalos de incertidumbre para el inciso b) de la figura 5.4	89
5.3. Intervalos de incertidumbre para la figura 5.6	91
5.4. Intervalos de incertidumbre para la figura 5.8	93
5.5. Energía promedio calculada por la red neuronal	95
5.6. Error porcentual entre la energía calculada por la red y la energía esperada.	97
5.7. Intervalos de incertidumbre para la energía calculada por la red	99
5.8. Porcentaje de eficiencia para los resultados de la red, en el intervalo que va de -0.5097 a 0.9613	100
5.9. Energía para muones individuales usando los resultados de la red neuronal	100
5.10. Extrapolación para estimar la energía de muones individuales	101
5.11. Comparación de las variables del capítulo 5 para diferente número de partículas simuladas	104
5.12. Comparación de las variables del capítulo 4 para diferente número de partículas simuladas	104
5.13. Porcentaje de veces que se activan 4 PMTs para muones positivos y negativos	117
5.14. Porcentaje de 3 PMTs activados para muones en la dirección H.II con 1 TeV de energía	119
5.15. Intervalos de incertidumbre para la energía calculada con la ecuación 5.1	121

Introducción

Núcleos activos de galaxias, explosión de supernovas, y destellos de rayos gamma son algunos de los fenómenos más violentos en el Universo. Los rayos gamma de altas energías (en órdenes de TeV) son indicadores de este tipo de sucesos [1]. Así que el estudio de estas señales, proporciona información sobre estos acontecimientos. El observatorio HAWC es un instrumento de investigación que se dedica a estudiar este tipo de señales. Éste observatorio se ubica en las laderas del volcán Sierra Negra en el estado de Puebla, y está conformado por tanques llenos de agua, de aproximadamente 5 m de altura y 7.3 m de diámetro. Cuando una partícula (ya sea un hadrón o un rayo gamma), llega del espacio exterior hacia la Tierra, ésta interactúa con las partículas de la atmósfera, y forma una cascada de partículas secundarias. Éstas partículas llegan a los tanques del observatorio, y por medio del efecto Cherenkov se genera luz que es captada por los detectores que hay en cada uno de los tanques (cada tanque tiene 4 fotomultiplicadores). Por medio del tiempo de detección y la distribución de carga en estos dispositivos, es posible determinar la dirección de la partícula inicial [2], y con ello generar mapas del cielo que nos dan información sobre las fuentes de rayos gamma en el Universo.

Al igual que los rayos gamma los neutrinos de alta energía también apuntan hacia a las fuentes más grandes de aceleración de partículas que hay en el Universo, y para altas energías estas partículas son una mejor fuente de estudio que los rayos gamma, ya que los fotones con energías superiores a los 100 TeV sufren una absorción por producción de pares con fotones de la radiación cósmica de fondo y con ello se reduce su camino libre medio [3]. La desventaja que tienen los neutrinos es que son difíciles de detectar y su tasa es prácticamente despreciable en órdenes de TeV. Sin embargo, hay varios métodos que se usan para la detección de estas partículas. Uno de ellos es la técnica *Earth-skimming* en donde se estudia la interacción de un neutrino y un nucleón al intercambiar un bosón W^\pm dando como resultado la producción de un leptón cargado del mismo sabor del neutrino procedente. Debido a que la sección eficaz del neutrino y el nucleón es muy pequeña es necesario utilizar un objetivo de gran masa, como el suelo de la superficie terrestre o una montaña.

Actualmente se está investigando la detección de neutrinos usando el observatorio HAWC por medio de la técnica *Earth-skimming*, en donde se usa al volcán Pico de Orizaba (que rodea a este observatorio) como objetivo de gran masa para producir la reacción neutrino-nucleón. El leptón con carga eléctrica que se produce viajará en la misma dirección del neutrino que lo originó, y para interacciones con montañas la dirección será cuasi-horizontal. Se propone reconstruir directamente la trayectoria del leptón cargado, o el producto de su descomposición, cuando pase por el observatorio. Es importante mencionar, que para encontrar este tipo de señales, primero se debe hacer una caracterización del ruido generado por cascadas atmosféricas en la misma dirección. Estas señales en su mayor parte serán muones que pasan por los tanques del observatorio, y al igual que los leptones, su ubicación se podrá reconstruir a partir de la posición de los tanques en donde se

detectaron señales.

En este trabajo se realizaron dos análisis empleando simulaciones Monte Carlo para estudiar estas señales de ruido. El primer análisis consistió en simular muones horizontales y verticales a través de un tanque del observatorio, esto con la finalidad de encontrar la región del tanque en donde es más probable que se activen 4,3 ó 2 fotomultiplicadores. Así, cuando se detecte un partícula que viaje por el observatorio, se podrá conocer la sección más probable del tanque que atravesó, y con ello se podrá definir mejor su trayectoria. El segundo análisis consistió en usar una red neuronal artificial para elaborar un estimador de energía para muones horizontales. Los datos que se usaron para entrenar a esta red neuronal provienen de simulaciones Monte Carlo. El objetivo de este análisis consiste en dar información sobre la energía de trazas de muones que el observatorio detecte.

Para explicar este trabajo, en los primeros dos capítulos se presenta toda la información referente al observatorio HAWC y a la forma en la que se realizan las simulaciones que se emplearon en este proyecto. En el tercer capítulo, se muestran los aspectos generales de una red neuronal artificial . Después, en los siguientes dos capítulos, se explica toda la información referente a los análisis realizados, en donde se describen los métodos utilizados y se discuten los resultados (el capítulo 4 para el volumen efectivo de detección y el capítulo 5 para el estimador de energía). Por último, se da una conclusión general de los resultados obtenidos en este trabajo.

Capítulo 1

Marco teórico

En este trabajo se analizó el paso de muones en los detectores de HAWC para diferentes energías y trayectorias dentro de los tanques por medio de simulaciones. Pero es importante conocer la fuente natural que proporciona estas partículas a la atmósfera y la forma en que se pueden detectar. Es por ello que en este capítulo se dará un repaso general de rayos cósmicos, así como un resumen del funcionamiento y de las principales características del observatorio HAWC, que es un instrumento de investigación que detecta este tipo de señales.

1.1. Rayos Cósmicos

Los rayos cósmicos se pueden entender como partículas y radiación que llegan a la Tierra procedentes del espacio exterior. Con esta definición se incluye a la radiación electromagnética en todas sus longitudes de onda, neutrinos y todo tipo de partícula eléctricamente cargada como electrones, positrones, protones, y núcleos atómicos, que van desde el hidrógeno, hasta núcleos más pesados. Estos rayos cósmicos llegan constantemente a nuestro planeta originados en distintas partes del Universo, donde son acelerados y llegan a la atmósfera de la Tierra con energías superiores a las que actualmente se pueden alcanzar en cualquier acelerador de partículas [4].

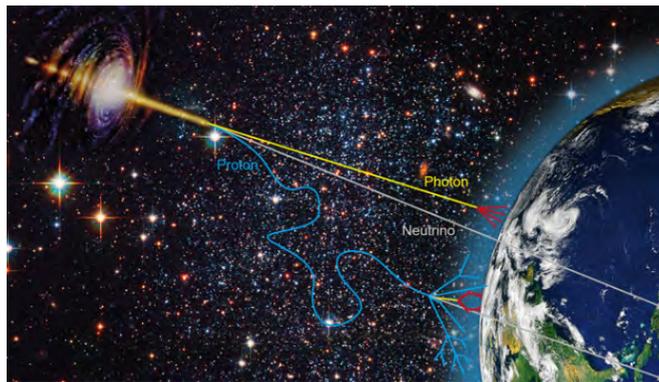


Figura 1.1: Representación gráfica de la llegada de rayos cósmicos a la Tierra. Imagen tomada de [5].

Actualmente la radiación cósmica es objeto de estudio de gran interés, y muchos laboratorios en el mundo se encuentran trabajando en este campo. Pero para llegar a ese punto, hubo varios acontecimientos que contribuyeron al descubrimiento de esta radiación, motivo por el cual en la

siguiente tabla se mencionarán de forma breve, los principales sucesos para el descubrimiento de los rayos cósmicos.

	Acontecimiento
Inención del electroscopio	En 1600 William Gilbert inventa el primer electroscopio [6]. Este instrumento consiste en un recipiente transparente con un par de láminas de metal sostenidas por una barra en su interior. Cuando un cuerpo con carga se acerca a la barra, las láminas de metal se separan debido a que el exceso de carga se distribuye en ellas con la misma polaridad. Cuando el electroscopio se descarga, las láminas vuelven a juntarse [7].
Descarga espontánea del electroscopio	En 1785 Coulomb observó la descarga de un electroscopio que se encontraba aislado. Él explicó este fenómeno a través de la interacción de las moléculas del aire con las láminas del electroscopio, pero dicha explicación no convenció a la comunidad científica de la época [7].
Partículas cargadas en el aire	En 1835 Faraday explicó la descarga para un electroscopio aislado, por la interacción de partículas cargadas que se encontraban en el aire y que provocaban la descarga de este aparato. Aunque no menciona el origen de dichas partículas [7].
Descubrimiento de la radiactividad	Con el descubrimiento de la radiactividad por el físico francés Henri Becquerel en marzo de 1896 [8], se tuvo un gran desarrollo en esta área por científicos de la época. Robert Duncan observó la interacción entre un electroscopio y una fuente radiactiva, encontró que esta radiación podría descargar este aparato de forma rápida [7].
Hipótesis de Rutherford	Rutherford explicaría el fenómeno de descarga del electroscopio diciendo que el suelo emite radiación que interactúa con las láminas del electroscopio y provocan su descarga; esta hipótesis fue probada durante los siguientes años haciendo experimentos en globos aerostáticos, barcos, y establecimientos subterráneos, que tenían como propósito alejarse o acercarse de la superficie terrestre, y su objetivo era ver qué tan rápido se descargaba un electroscopio. A mayor altura esperaban que éste se descargara más lento en comparación con un electroscopio en el suelo, debido a la baja interacción de las láminas del aparato con la radiación del suelo. A pesar de estos intentos no se encontraron resultados concluyentes [7].
Experimentos de Victor Hess	Fue hasta 1911 que el físico Victor Hess hizo un vuelo en globo alcanzando una altura de 5350 m [7], encontrando que el electroscopio se descargaba más rápido que cuando se encontraba en el suelo. Concluyendo que esta radiación no provenía de la Tierra sino del espacio exterior. Sus resultados fueron aceptados por varios científicos que replicaban su experimento y más tarde ganaría el premio Nobel.

Tabla 1.1: Principales acontecimientos para el descubrimiento de los rayos cósmicos.

Hoy en día se tiene un buen conocimiento de las reacciones que ocurren entre los rayos cósmicos y la atmósfera. Dando como resultado una lluvia de partículas que se pueden detectar cuando llegan a la Tierra. Más adelante se hablará en general de este mecanismo.

1.1.1. Historia de los Rayos Cósmicos en México

El estudio de rayos cósmicos en México fue promovido en sus orígenes por Manuel Sandoval Vallarta quien fue un físico mexicano graduado del MIT, que formó un grupo interesado en el estudio de la radiación cósmica. Entre los miembros de este equipo se encontraba Jaime Lifshitz, Nabor Carrillo, Alberto Barajas, Manuel L. Perrusquia, Juan de Oyarzabal, Hector Uribe, Fernando Alba Andrade, Alfredo Baños y Carlos Graef. Estos investigadores en el periodo de 1933 y 1939 escribieron una serie de artículos sobre rayos cósmicos que fueron publicados en *Nature*, *Physics Review*, *The Review of Modern Physics* [9].

En 1940 Alfredo Baños y Manuel I. Perrusquia fabricaron un sistema para medir la intensidad de rayos cósmicos en función del ángulo cenital. Este aparato fue colocado en la estación meteorológica del Palacio de Minería (en donde la Facultad de Ciencias de la UNAM estaba ubicada en ese entonces) y recolectó datos de forma ininterrumpida durante 100 días. Con la información obtenida, Alfredo Baños escribió el primer artículo de investigación en rayos cósmicos publicado en una revista profesional en México [9].



Figura 1.2: Imagen del simposio de rayos cósmico de 1939 en Chicago. Que cuenta con la presencia de Alfredo Baños y Sandoval Vallarta (6 y 9), así como físicos ganadores de un premio Nobel. Hans Bethe (1), Arthur Compton (4), Edward Teller (5), S. Goudsmit (8), Carl Andersen (12), Victor Hess (15), Wilhelm Bothe (18) y Werner Heisenberg (19). Imagen tomada de [9].

Entre otros trabajos que se han hecho en este ámbito se encuentran los siguientes: los eventos solares (Miroshnichenko y Pérez-Peraza, 2008), el transporte coronario solar azimutal de rayos cósmicos (Pérez-Peraza, 1986), la propagación de rayos cósmicos interplanetaria (Valdés-Galicia, 1992), eventos de protones solares relativistas (Pérez-Peraza et al. 2005) y la propagación de neutrones solares en la atmósfera terrestre (Dorman y Valdés-Galicia, 1997) [9].

Este repaso histórico de los rayos cósmicos en México refleja el interés de los investigadores nacionales en este ámbito que incluso han trabajado en proyectos de gran complejidad, como el observatorio de rayos gamma Milagro, el Gran Colisionador de Hadrones (LHC) del CERN en Suiza y los observatorios de rayos cósmicos Pierre Auger y *Telescope Array* que se encuentran en Argentina y Estados Unidos respectivamente [11]. Actualmente en México hay centros de investigación dedicados al estudio de esta radiación, como el observatorio HAWC (High Altitude Water Cherenkov) ubicado en las laderas del volcán Sierra Negra en Puebla; de éste se hablará en la siguiente sección.

1.2. El observatorio HAWC

Es un observatorio conformado por 300 tanques, que se dedica a estudiar el Universo en altas energías (entre 100 GeV y 100 TeV). Se encuentra ubicado a una altura aproximada de 4100 metros sobre el nivel del mar en una de las laderas del volcán Sierra Negra en Puebla y se inauguró en marzo del 2015 [10]. Cabe destacar que este proyecto tardó más de 8 años desde que se hicieron las primeras pruebas hasta que se inició operaciones. En un principio había otros candidatos que pudieron ser su sede; por ejemplo Chacaltaya en Bolivia y el Tibet en China, pero al final el volcán Sierra Negra fue elegido en 2007 gracias al interés de los investigadores mexicanos y a su experiencia en proyecto similares [11]. Actualmente HAWC es una colaboración internacional en la que participan más de 30 instituciones de México, Estados Unidos y Europa [10].



Figura 1.3: Imagen del observatorio HAWC. Tomada de [12].

El observatorio HAWC durante un ciclo de 24 horas puede observar dos terceras parte del cielo [10]; ahora se dirá con un poco más de detalle el tipo de señales que se estudian.

1.2.1. Cascadas atmosféricas

Los rayos cósmicos que llegan del espacio hacia la Tierra son principalmente rayos gamma y hadrones, así que tendremos dos mecanismos de cascadas.

Caso I: Cascada atmosférica generada por un fotón:

En general podemos visualizar este procedimiento mediante los siguientes pasos:

1. Un rayo gamma primario entra a la atmósfera y a través del mecanismo de producción de pares formará un electrón y un positrón [13].
2. Con la formación del electrón y el positrón sucederá cualquiera de las siguientes interacciones:
 - a) *Bremsstrahlung*: En este proceso puede intervenir tanto como un electrón como un positrón, en donde la partícula perderá energía al ser desviada por un campo eléctrico de un núcleo atómico que esté en el aire, y como consecuencia se emitirán fotones [13].
 - b) *Aniquilación*: El positrón y electrón se aniquilan para formar rayos gamma; en esta interacción se convierten en energía [13].

También gran parte de los electrones y positrones perderán energía a través de colisiones con partículas del aire [14].

3. Los rayos gamma formados en el proceso anterior van a generar más electrones y positrones por producción de pares y se irá repitiendo el proceso descrito en el paso dos. Cabe destacar que estos fotones también pierden energía por dispersión de Compton [14].

4. La energía de las partículas involucradas en estas interacciones disminuye hasta que la dispersión de Compton sea el proceso que domine en los fotones (superando a la producción de pares), y la pérdida de energía a través de colisiones sea mayor que la producción de gammas por Bremsstrahlung y por aniquilación, para los electrones y positrones [14]. Así que la pérdida de energía por ionización empieza a dominar la cascada que se irá disipando en la atmósfera.

Caso II: Cascada atmosférica generada por un Hadrón:

En general podemos visualizar este procedimiento mediante los siguientes pasos:

1. Un Hadrón (H) con bastante energía (en ordenes de TeV) llega a la Tierra y colisiona con las moléculas que componen el aire (N_2, O_2, Ar); como consecuencia de esta interacción se crean un gran número de partículas secundarias. Este proceso se puede representar mediante la siguiente expresión:

$$H + N_2, O_2, Ar \rightarrow H^* + (N_2, O_2, Ar)^* + \pi^\pm + \pi^0 + K^\pm + K^0 + \text{Otros Mesones.}$$

En donde H^* es un fragmento del rayo cósmico original que sigue interactuando con la atmósfera y $(N_2, O_2, Ar)^*$ son fragmentos de un núcleo atmosférico en un estado excitado [15]. El símbolo π^\pm representa a un pión de carga positiva o negativa y π^0 es un pion de carga nula. Por último K^\pm denota a un kaón de carga positiva o negativa y K^0 es un kaón de carga neutra.

2. El proceso de producción de partículas secundarias continúa debido a la desintegración de piones, y que se representa mediante las siguientes ecuaciones [15]:

$$\begin{aligned} \pi^+ &\rightarrow \mu^+ + \nu_\mu \\ \pi^- &\rightarrow \mu^- + \bar{\nu}_\mu \\ \pi^0 &\rightarrow 2\gamma \end{aligned}$$

3. Los muones y neutrinos con suficiente energía llegan a la Tierra en donde pueden ser detectados [15], mientras que los fotones siguen las interacciones mencionadas anteriormente y forman una componente electromagnética de esta cascada hadrónica.

En la figura 1.4 se puede observar un esquema que representa el proceso de formación de una cascada atmosférica causada por un rayo cósmico. En donde se aprecia el gran número de partículas secundarias y las diferentes componentes que conforman a la cascada.

Con los tanques del observatorio HAWC, por medio de técnicas especializadas se puede estimar la dirección inicial y la energía del rayo cósmico primario (algunos de estas se explican con mayor detalle en [2] y [15]). En particular las cascadas iniciadas por un fotón son de gran interés, debido a que los rayos gamma no son afectados por los campos magnéticos que se encuentran en el espacio, y si se conoce la dirección de llegada del rayo gamma primario se podrá ubicar la fuente que lo emitió. Con ello se podrán estudiar e identificar nuevas fuentes de rayos gamma.

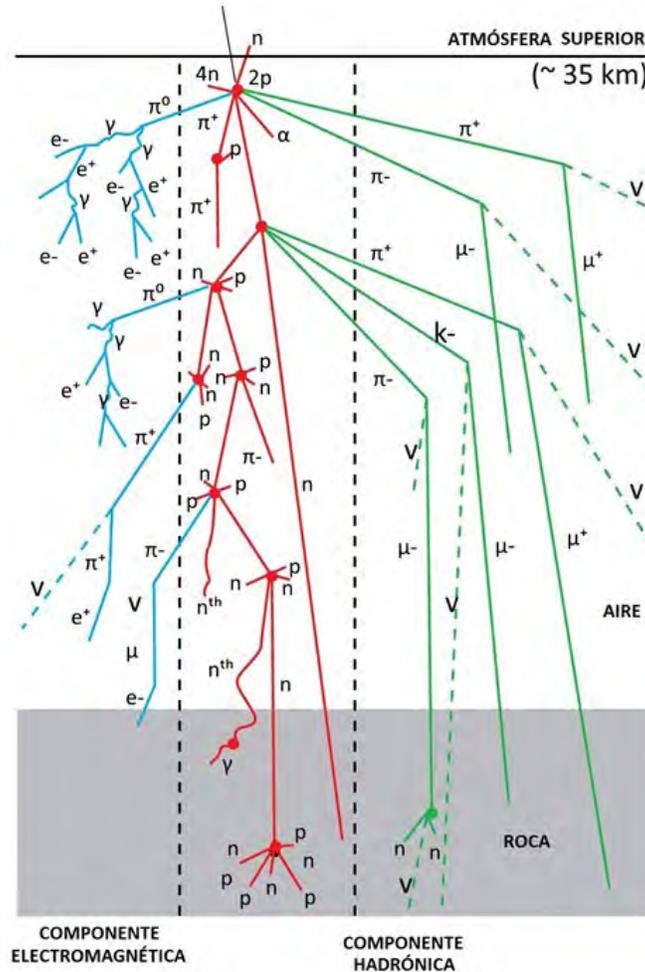


Figura 1.4: Representación gráfica de una cascada atmosférica. Imagen adaptada de [16].

1.2.2. Proceso de Detección en HAWC

Las partículas generadas en la cascada atmosférica inciden en los tanques de HAWC y son detectadas mediante el efecto Cherenkov. Este fenómeno se explicará a continuación.

Efecto Cherenkov

El efecto Cherenkov ocurre cuando una partícula cargada se mueve en un medio dieléctrico más rápido que la velocidad de la luz en el mismo medio. Como resultado se emite radiación Cherenkov y las longitudes de onda de los fotones que se arrojan están dentro y alrededor del espectro visible [17]. Para entender un poco más a fondo que es lo que sucede en este fenómeno, supongamos que tenemos una partícula cargada que se mueve a través de un medio dieléctrico, entonces tendremos cualquiera de los siguientes casos:

- **La partícula se mueve a una velocidad mucho menor que la velocidad de la luz en el medio ($v \ll c$).**

En la figura 1.5, los círculos representan a los átomos que forman el medio y el punto P representa la posición inicial de la partícula cargada. Cuando esta partícula entra al medio va a distorsionar a los átomos cercanos debido a su campo eléctrico y va a causar una polarización totalmente simétrica alrededor del punto P. Si la partícula tiene carga negativa el lado positivo de cada átomo es el que se orienta en la vecindad del punto P, mientras que el lado negativo quedará en el otro extremo, así que cada átomo se podrá modelar como un dipolo. Conforme la partícula cargada va avanzando por los átomos (hasta llegar al punto P'), estos irán regresando a su forma original y debido a la completa simetría en la polarización alrededor de la partícula no habrá campo resultante [17].

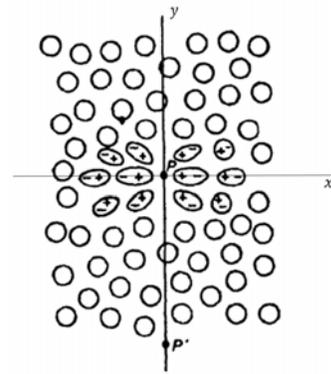


Figura 1.5: Polarización de un dieléctrico por una partícula a baja velocidad. Tomada de [17].

- **La partícula se mueve a una velocidad cercana o igual a la velocidad de la luz en el medio ($v \sim c$).**

Para este caso la polarización no es completamente simétrica, como se ve en la figura 1.6; hay una simetría en la posición de los átomos cargados en el eje y pero no en el eje x . Esto causará un campo dipolar resultante que se irá formando conforme va avanzando la partícula y que va a generar un breve impulso electromagnético que se irá irradiando en cada segmento de la trayectoria de la partícula. Pero en general estas pequeñas ondas que se irán emitiendo por todo el camino de la partícula se irán interfiriendo destructivamente, así que la intensidad del campo resultante será igual a cero [17].

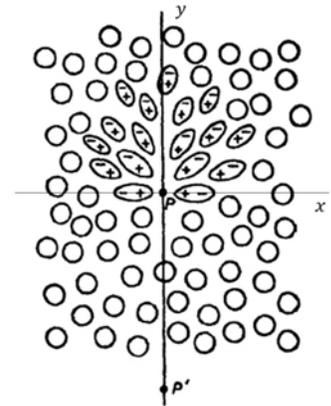


Figura 1.6: Polarización de un dieléctrico por una partícula a alta velocidad. Tomada de [17].

- **La partícula se mueve a una velocidad mayor a la velocidad de la luz en el medio ($v > c$).**

En este caso, como la partícula se mueve más rápido que la luz en el medio se forma una onda de choque, es decir, las ondas que se irán emitiendo por todo el camino de la partícula estarán en fase, y se tendrá una interferencia constructiva, dando como resultado un campo distinto de cero. Esto se puede entender mediante el principio de Huygens, en donde la radiación emitida solo se ve en un ángulo en particular, tal y como se ve en el inciso a de la figura 1.7, en donde las ondas generadas en los puntos P_1, P_2 y P_3 son coherentes y se combinan en un frente de onda plana en el segmento BC. Esta coherencia ocurre cuando la partícula recorre el segmento AB en el mismo tiempo en que la luz recorre el segmento AC [17].

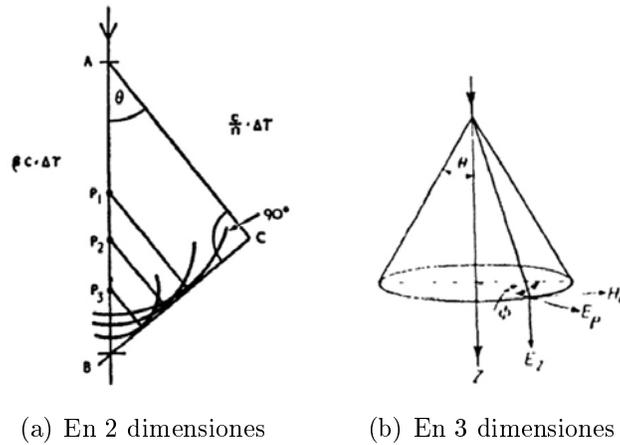


Figura 1.7: Emisión de radiación Cherenkov. Imagen tomada de [17].

Teniendo en cuenta lo anterior podemos expresar la longitud de dichos segmentos de la siguiente forma:

$$AB = (\text{Velocidad de la Partícula})(\text{Tiempo}) = (\beta c)(\Delta t)$$

$$AC = (\text{Velocidad de la luz})(\text{Tiempo}) = (c/n)(\Delta t)$$

Por lo tanto se tendrá que:

$$\cos \theta = \frac{AC}{AB} = \frac{1}{\beta n} \quad (1.1)$$

En donde β es la velocidad relativa a la de la luz . Así que el efecto Cherenkov tendrá las siguientes características:

- Para un medio con índice de refracción n habrá una velocidad umbral en la que se pueda emitir radiación $\beta_{min} = \frac{1}{n}$
- Para una partícula relativista con $\beta = 1$ habrá un ángulo característico de emisión $\theta = \cos^{-1}(1/n)$
- La emisión de esta radiación formará un cono, tal y como se observa en el inciso b de la figura 1.7.

Los tanques del observatorio HAWC están llenos de agua y por medio del efecto Cherenkov detectan señales de partículas secundarias con carga eléctrica generadas en la cascada atmosférica. Las características de estos tanques son las siguientes:

- *Dimensiones:* Tienen una altura de 5 m y un radio de 7.3 m [10].
- *Volumen de agua:* En cada tanque hay una bolsa negra y hermética que impide el paso de luz exterior, ésta contiene 180,000 litros de agua ultra purificada [18].
- *Detectores:* Hay 4 fotomultiplicadores (PMTs) en cada tanque, uno colocado en el centro con alta eficiencia cuántica¹ de 10 pulgadas de diámetro, y otros tres fotomultiplicadores periféricos de 8 pulgadas de diámetro [10].

¹Esta cantidad es el cociente entre los electrones obtenidos en el cátodo del PMT y el número de fotones que inciden en él.

Las partículas generan luz al entrar en los tanques y ésta es captada por los detectores que hay en ellos. En general lo que pasa en el detector es que la luz recibida pasa a un cátodo que emitirá electrones debido al efecto fotoeléctrico y estos pasarán por un sistema de dínodos que amplifican la señal que llegará a un sistema electrónico en donde se podrá guardar y analizar.

1.2.3. Algunos resultados destacados obtenidos en HAWC

Entre algunos de los resultados publicados por parte de la colaboración se encuentran los siguientes:

- En el año 2013 se observó la sombra de la luna en rayos cósmicos, que fue un estudio en donde únicamente se utilizaron 30 tanques ya que la construcción del observatorio no había terminado y que atrajo la atención de medios internacionales como la BBC (*British Broadcasting Corporation*) [10].
- Se observó un fenómeno transitorio, en donde se encontró un aumento en la emisión de rayos gamma en la galaxia Markarian 501 el 6 de abril del 2016 [10].
- Presentación de un mapa del cielo de rayos gamma en altas energías (ver figura 1.8), en donde se observan nuevas fuentes de emisión a lo largo de la Vía Láctea, así como los objetos extragalácticos Markarian 421 y 501 [10].



Figura 1.8: Mapa del cielo en rayos gamma. Imagen tomada de [10].

Los investigadores en HAWC se encuentran trabajando en muchos temas de interés; como la búsqueda de nuevas fuentes de emisión de rayos gamma, la búsqueda de señales de aniquilación de materia oscura y diferentes estudios sobre el Sol (por mencionar algunas líneas de investigación). Además existe la propuesta de utilizar este observatorio como un detector de neutrinos, que es el tema de la siguiente sección.

1.2.4. Detección de neutrinos en HAWC

En la introducción de este proyecto, se planteó la idea de utilizar la técnica *Earth-skimming* para la detección de neutrinos en el observatorio HAWC. En donde se utiliza al volcán Pico de Orizaba como blanco para producir la interacción entre neutrinos y nucleones. Para esta interacción, se debe tomar en cuenta que hay tres tipos de neutrinos (asociados a cada una de las familias leptónicas): electrón neutrino (ν_e), muón neutrino (ν_μ) y tau neutrino (ν_τ). Para la técnica *Earth-skimming*, la detección de un electrón neutrino es desfavorable, ya que los electrones de altas

energías producidos en este proceso, iniciarán una cascada electromagnética, que será fácilmente absorbida por el volcán [3]. Para un muón neutrino, se debe tomar en cuenta que tiene una vida media relativamente grande, así que solo se tendrá una señal detectable para un neutrino ultra energético [3]. Por otro lado, la vida media de una tau neutrino, es relativamente corta (aunque sea muy energético), así que los leptones producidos por la interacción entre neutrinos y nucleones, van a decaer en partículas secundarias, que podrán generar una señal detectable en los tanques del observatorio. Por lo tanto, el estudio de neutrinos ultraenergéticos en el observatorio HAWC, se llevará a cabo por medio de la búsqueda y el análisis de señales de leptones tau cargados, o el producto de su descomposición en los tanques de este observatorio [3].

Se ha hecho una estimación de la tasa de detección usando un modelo geométrico para calcular el área efectiva del observatorio, en donde se llegó a la conclusión de que es factible realizar mediciones del flujo de neutrinos de ultra alta energía de origen cósmico durante la vida útil esperada del observatorio HAWC [3].

Para buscar una señal de un leptón cargado o el producto de su descomposición, primero se tiene que hacer una caracterización del ruido generado por partículas que provienen de cascadas atmosféricas. Estas partículas en mayor parte son muones, y su dirección se puede estimar a partir de los tanques que activaron, se hablará con un poco mas a detalle sobre este tipo de reconstrucción en el capítulo 4, en donde se realizan pruebas utilizando simulaciones Monte Carlo para tratar de mejorar este algoritmo.

Es importante mencionar que la información que se obtiene en los detectores de este observatorio se va almacenando de forma ininterrumpida, y se va guardando en archivos denominados *run*, que contienen los datos obtenidos durante un día [2]. Estos archivos contienen divisiones, que corresponden a 2 minutos de datos y se denominan *subrun*. En estos archivos, por cada 2 minutos de datos solo se almacenan 250 milisegundos, ya que se tiene un sistema de adquisición de datos (*trigger*) que considera ciertos criterios antes de almacenar cualquier tipo de información. El *trigger* de HAWC, se basa en que cierto número de fotomultiplicadores sean activados, y registren una carga mayor a cierto valor umbral en una ventana de tiempo de aproximadamente 150 ns [3]. Es importante mencionar que no se requiere un *trigger* especializado para la búsqueda de muones horizontales [3]. Así que en los datos del observatorio, se buscarán señales en donde se activen varios fotomultiplicadores en al menos dos tanques vecinos (con una propagación consistente con la velocidad de la luz), en donde se irán almacenando los datos correspondientes al número de fotomultiplicadores activados, el tiempo promedio registrado y la suma de la carga en cada detector. Con esta información se puede identificar trazas de muones en los datos de este observatorio, que son el primer paso para encontrar trazas de leptones cargados.

Capítulo 2

Simulaciones en HAWCSIM

El tema central de este trabajo consistió en analizar simulaciones de partículas viajando a través de los tanques del observatorio. Para hacer estas simulaciones se utilizó HAWCSIM, que es un programa desarrollado de manera interna por la colaboración HAWC [19]. En este capítulo se darán las principales características, y la forma en la que se usa este programa para hacer simulaciones. Empezando por la descripción de AERIE (*Analysis and Event Reconstruction Integrated Environment*) que es el software utilizado en HAWC para analizar datos, pasando por la instalación y configuración de HAWCSIM en AERIE, hasta llegar a un ejemplo en concreto de como hacer una simulación usando este programa.

2.1. Software utilizado en HAWC

AERIE es el software utilizado en HAWC para realizar el análisis de datos obtenidos en el experimento. Para poder utilizar todas las funciones de este programa es necesario instalar una serie de paquetes adicionales (a los cuales nos referimos como *externals*). Así que el proceso de instalación del software usado en HAWC se puede dividir en dos partes: la instalación de los paquetes *externals* y la compilación de AERIE.

Los pasos a seguir para la instalación de este software se pueden resumir en los siguientes puntos, en donde se hará énfasis en los pasos que nos ayudarán a explicar la instalación de HAWCSIM en las sección 2.4:

- **Instalación de los paquetes *externals*.**

Estos paquetes son programas y librerías necesarias para compilar el código base en HAWC. Actualmente boost, fftw, gsl, root, xerces, cmake, cfitsio, cealpix-cxx y photospline conforman a los *externals* [1].

Método de instalación: *Auger Package Environment* (APE) es un software desarrollado para el experimento Auger y que fue implementado en HAWC como una herramienta que permite instalar paquetes y establecer un entorno estandarizado de forma sencilla. Por ello la instalación de los *externals* se hará por medio del sistema APE. Los pasos de instalación son los siguientes:

- Tanto el software APE como los programas que componen a los paquetes *externals*, necesitan ciertos programas y librerías para funcionar de manera correcta. Así que en primer lugar se tienen que instalar estos programas como prerrequisito antes de usar el

sistema APE. Una lista completa y la forma de instalación de estos programas se puede consultar en la referencia [20].

- El siguiente paso consiste en descargar el archivo con la versión más reciente del APE y guardarlo en un directorio específico para la instalación del software. La carpeta `ape-hawc-2.02.02` es un ejemplo del archivo APE después de descomprimirlo.
- Después se debe hacer un archivo que tendrá la ubicación en donde el software será instalado. El nombre del archivo será `APERC` junto con la versión APE descargada (`aperc_2.02.02` por ejemplo). La estructura del archivo es la siguiente:

```
[DEFAULT]
base = <carpeta>/externals/2.02.02/
build = /tmp/<usuario>-apebuild
[APE]
jobs = 2
mirrors = mx us
```

La sección `[DEFAULT]` contiene la ubicación en donde se instalarán los programas, `jobs` indicará el número de núcleos que se usarán en paralelo para realizar la compilación y dependerá de la potencia de la computadora en donde se esté realizando la instalación; `mirrors` es donde se especifica el orden en que se conectará con los sitios de distribución (uno en México y otro en Estados Unidos) para descargar los programas a instalar.

- Se establece una variable de entorno de nombre `APERC`, que apuntará² al archivo creado en el paso anterior, por último se entrará a la carpeta APE que se descargó y se ejecutará el comando de instalación por medio de la terminal.

■ Compilación de AERIE

Método de instalación: En esta parte se describe el proceso de instalación de AERIE al descargar una versión de este programa de las páginas internas de HAWC (svn). Los pasos a seguir son los siguientes:

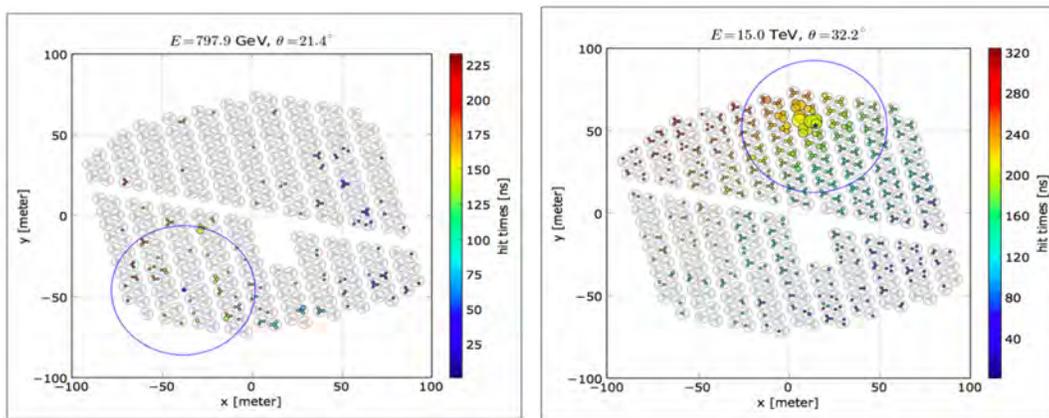
- Como primer paso se debe hacer una configuración del entorno de instalación, en donde se comprueba que la variable de entorno `APERC` siga funcionando, además de ejecutar un script con la configuración `ape` para colocar rutas y variables de entorno que apunten a los *externals*.
- Después se descargarán una versión de AERIE de las páginas internas de HAWC y se usará el programa `CMAKE` para llevar a cabo la compilación y la instalación de AERIE.
- Por último se descargará de las páginas internas de la colaboración, los archivos que contienen las características del observatorio para tener una correcta configuración a la hora de usar AERIE en el análisis de datos.

Es importante destacar que aquí solo se presentó una visión general de la instalación del software utilizado en HAWC, que nos servirá para describir la instalación y la configuración de HAWCSIM en una sección posterior. Para una explicación más detallada de la instalación de AERIE y los paquetes *externals* se puede consultar la referencia [2].

²Se refiere a que la variable de entorno contiene la ruta para acceder al archivo mencionado

A continuación se mencionan algunas de las las funciones que tiene AERIE para el análisis de datos:

- Reconstrucción de Cascadas:** Utilizando los datos obtenidos por el observatorio y que han sido clasificados como posibles candidatos a cascadas atmosféricas, se puede realizar una reconstrucción de esta información por medio de este software; se obtendrá tanto el núcleo de la cascada como la dirección en la que llega. En la figura 2.1 se muestra un ejemplo de esta reconstrucción, en donde la estrella que está en el centro del círculo (con un radio de 40 m.) representa el núcleo de la cascada, la escala de colores indica el tiempo en que se detecta la señal y el tamaño de cada marcador es proporcional al número de fotoelectrones detectados. Por medio de estos diagramas se pueden distinguir las cascadas generadas por un hadrón o por un fotón. Una cascada generada por un hadrón se caracteriza por tener señales grandes lejos del núcleo de la cascada, debido a que algunas de las partículas que aquí se generan son muy penetrantes (figura 2.1a), y las cascadas generadas por un fotón tiene la características de ser compactas y tener señales grandes cerca del núcleo de la cascada debido a que los fotones y electrones creados aquí no tienen un gran momento transversal (figura 2.1b).



(a) Cascada generada por un hadrón.

(b) Cascada generada por un fotón.

Figura 2.1: Reconstrucción de cascadas con AERIE. Imagen tomada de [10].

- Creación de Gráficos e Histogramas:** Con el paquete externo ROOT se pueden realizar diversos gráficos e histogramas de algunas de las variables del archivo reconstruido o de alguna simulación. Por ejemplo, se pueden hacer histogramas del número de eventos, el tiempo registrado en los fotomultiplicadores y la distribución de rayos cósmicos que llegan al detector.
- Mapas del Cielo:** Se pueden hacer mapas del cielo a partir de los datos de cascada que han sido reconstruidos, como el que se muestra en la figura 1.8, que nos da información de las fuentes en el Universo que generan rayos gamma.
- Simulaciones:** También se pueden realizar simulaciones de cascadas atmosféricas que llegan al observatorio, o de la respuesta de los detectores al paso de una partícula específica.

En la siguiente sección se hablará un poco más a detalle sobre este tema, en donde nos enfocaremos en dar las principales características de HAWCSIM, un programa diseñado para hacer simulaciones.

2.2. HAWCSIM: Aspectos Generales

Como se mencionó anteriormente, HAWCSIM es un programa desarrollado de manera interna por la colaboración, que simula a los detectores del observatorio tomando en cuenta sus características específicas; por ejemplo su ubicación y dimensiones, así como los materiales utilizados en su fabricación. Con este programa se puede estudiar la respuesta del observatorio al simular cascadas atmosféricas, o el paso de alguna partícula en específico que viaje a través del arreglo. Es importante mencionar que este programa está basado en GEANT4 [21], que es una herramienta especializada en la simulación de detectores. Además, la entrada de una simulación en HAWCSIM puede ser un archivo en CORSIKA, que es un programa especializado en simular cascadas atmosféricas. Por este motivo, en la siguiente sección hablaremos en general de las principales características de GEANT4 y de CORSIKA a fin de tener una idea en general del funcionamiento de HAWCSIM a la hora de hacer una simulación.

2.2.1. GEANT4

GEometry ANd Tracking (GEANT4), es un conjunto de librerías basadas en C++ [22], que usa el método Monte Carlo para simular el paso de partículas a través de materia. También ofrece gran flexibilidad para el diseño de detectores, así como la información que se quiere obtener del mismo, dando al usuario una gran variedad de experimentos físicos que pueden ser simulados. Actualmente, esta herramienta se utiliza en muchos campos de investigación, entre los más conocidos están: la física de altas energías, física nuclear, ciencia espacial y medicina [23].

● **Funcionamiento: Librerías de GEANT4**

GEANT4 es una herramienta desarrollada en C++, que nos brinda un conjunto de librerías para la construcción de simulaciones. Una librería en C++ se puede entender como un archivo que contiene el código de muchos programas que permiten realizar una función en común [24]. Estas librerías nos permiten usar diferentes tipos de clases, que son plantillas que nos permiten crear objetos de datos y que tienen como objetivo cumplir con alguna tarea en específico a la hora de hacer una simulación; con esto se hace más sencillo la elaboración de estos programas.

Las clases en GEANT4 se pueden dividir en diferentes categorías de acuerdo a sus características, y éstas operan mediante la metodología Booch, en donde hay una relación entre estas clases que operan de manera dependiente, es decir, que algunas clases necesitan de otras para poder funcionar. Para entender mejor esta metodología y el diagrama del funcionamiento de las diferentes clases de GEANT4 que operan mediante este algoritmo se puede consultar la referencia [25].

Estas clases representan el componente principal para hacer una simulación y se pueden dividir en las siguientes categorías:

- **Global:** La declaración de esta clase es de suma importancia en una simulación, ya que sirve para definir las diferentes constantes físicas, las unidades del sistema y para crear números aleatorios.
- **Materiales y partículas:** Esta clase contiene toda la información necesaria para describir las propiedades físicas de los materiales y las partículas usadas en una simulación.
- **Geometría:** Proporciona las clases necesarias para definir la geometría del arreglo experimental que se desea simular.

- **Procesos físicos:** Incluye los modelos que permiten simular los procesos físicos involucrados en la interacción de las partículas con la materia, por ejemplo, interacciones electromagnéticas y hadrónicas.
- **Rastreo (*Track*):** Nos muestra la evolución del estado de la partícula durante la simulación, en donde se guardará la información de la partícula (como su dirección, su energía, su posición, etc.) en un instante determinado y se va ir actualizando conforme vaya cambiando al siguiente paso.
- **Impacto (*Hit*):** Estas clases nos sirven para indicar la sensibilidad en la geometría del detector y en la categoría de rastreo [26].
- **Evento (*Event*):** Aquí se almacena la información de todas las partículas que se crearon a partir de una interacción inicial [26].
- **Corrida (*Run*):** Secuencia de eventos en donde la interacción inicial es la misma para todos los eventos y pasan por el mismo arreglo experimental.
- **Visualización:** Nos permite realizar una visualización de la simulación, en donde se muestra al detector simulado y la interacción entre las partículas simuladas y el material por donde pasaron.

Teniendo en cuenta las diferentes clases que ofrece GEANT4, ahora veremos los principales componentes para hacer una simulación.

● Elaboración de una simulación

Las simulaciones en GEANT4 van a depender de las tareas o las funciones que el usuario esté buscando realizar, y estas pueden ser simulaciones de forma sencillas o muy complejas, en donde la cantidad y las diferentes clases que se usen sean muy variadas. En términos generales toda simulación contiene 3 clases fundamentales para poder operar [26], estas son las siguientes:

- **DetectorConstruction:** Indica los materiales del detector, así como su geometría.
- **PhysicsList:** Se indican los modelos físicos que se tomarán en cuenta para la interacción entre las partículas y la materia en la simulación.
- **PrimaryGeneratorAction:** Aquí se establecen las partículas a simular considerando su posición, energía y momento.

Cabe destacar que cada clase necesita dos archivos para funcionar [27], un archivo que indique los métodos de la clase y que llevarán la extensión .hh, y otro archivo para implementar estos métodos con extensión .cc. Por ejemplo, para las clases principales estos archivos serán los siguientes:

- DetectorConstruction.hh y DetectorConstruction.cc
- UserPhysicsList.hh y UserPhysicsList.cc
- UserPrimaryGeneratorAction.hh y UserPrimaryGeneratorAction.cc

Estas clases se declaran al inicio de la simulación, para que el programa funcione de manera correcta.

```
#include DetectorConstruction.hh
#include UserPhysicsList.hh
#include UserPrimaryGeneratorAction.hh
```

Figura 2.2: Ejemplo de como se declaran las clases en una simulación.

En el caso de HAWCSIM, estas tres clases se encuentran ubicadas en la carpeta de hawcsim/include para los archivos con extensión .hh, y en la carpeta de hawcsim/src para los archivos con extensión .cc. Además estas carpetas incluyen todas las clases que se utilizaron para la elaboración de esta simulación. Es importante mencionar que todos estos archivos ya vienen incluidos a la hora de instalar AERIE, pero para poder usarlos hace falta hacer una configuración e instalar GEANT4.

Además de las tres clases mencionadas anteriormente, toda simulación lleva los siguientes dos archivos [27]:

- **Ejecutable.cc:** Archivo en donde se realizó la estructura de la simulación, e indica todas las clases utilizadas y el orden conforme se van utilizando.
- **CMakeLists.txt:** Este archivo sirve para hacer la compilación de la simulación en CMAKE, y generar el archivo ejecutable.

En HAWCSIM, estos archivos están en la carpeta aerie/src/hawcsim que ya viene incluida a la hora de instalar AERIE.

Se pueden consultar las referencias [25] y [27] para una descripción más detallada del uso de GEANT4 a la hora de hacer una simulación; en esta sección solo se mencionaron los aspectos fundamentales para entender el funcionamiento de HAWCSIM.

2.2.2. CORSIKA

CORSIKA, es un software especializado en la simulación de cascadas atmosféricas que emplea el método Monte Carlo³ para modelar la interacción de los rayos cósmicos con la atmósfera. El funcionamiento de este programa se puede dividir en 4 partes [28]:

- **Parte I:** Se toman en cuenta las partículas inestables que decaen, y se va siguiendo el estado de cada partícula tomando en cuenta la pérdida de energía por ionización y el cambio de posición debido a las dispersiones.
- **Parte II:** En esta parte se toma en cuenta la interacción en altas energías de los núcleos y hadrones de la cascada con los elementos de la atmósfera.
- **Parte III:** Aquí se consideran todas las interacciones hadrónicas de baja energía.
- **Parte IV:** En esta parte se incluyen las interacciones electromagnéticas de la cascada.

Los modelos usados en CORSIKA para las interacciones electromagnéticas son EGS4 y NKG, y para los procesos hadrónicos se usan modelos que dependen de la energía [28].

³Éste método se explica con un poco más de detalle en la sección 2.3.1.

Modelo para altas energías	Modelos para bajas energías
VENUS QGSJET DPMJET SIBYLL HDPM	GHEISHA ISOBAR

Tabla 2.1: Modelos hadrónicos usados en CORSIKA.

La explicación y el funcionamiento de los modelos presentados en la tabla 2.1, así como los modelos EGS4 y NKG, se pueden consultar en la referencia [29]. En términos generales, HAWCSIM toma la información de las partículas generadas en CORSIKA, y simula la interacción de estas partículas al pasar por el arreglo del observatorio.

2.3. Descripción de cascadas

Para estudiar las principales características de una cascada electromagnética, hadrónica o una combinación de ambas, comúnmente se utilizan dos métodos que no requieren una medición experimental. El primero de ellos consiste en usar técnicas Monte Carlo para realizar simulaciones que nos dan información detallada de la cascada, con la desventaja de que los requerimientos computacionales son altos. El otro método consiste en estudiar a la cascada por medio de un modelo analítico, que nos da una descripción cualitativa de la cascada permitiendo una mejor comprensión física.

2.3.1. Método Monte Carlo

Hemos mencionado anteriormente que tanto GEANT4 como CORSIKA utilizan el método Monte Carlo para poder funcionar. En esta parte se dará la idea en general de como se emplea este método para simular las cascadas provocadas por una interacción inicial al pasar por un medio.

• Idea general del funcionamiento de este método

Las cascadas generadas por una partícula al interactuar con la materia, se pueden considerar como procesos estocásticos [30], y pueden ser tratados desde el punto de vista estadístico para ser simulados mediante una computadora. Para ello se debe configurar un modelo que tome en cuenta el movimiento y la interacción de las partículas en la cascada.

La idea base de este modelo consiste en representar a la sección eficaz como una distribución de probabilidad y usar números aleatorios que nos permitan tomar muestras de esta distribución para poder simular esta sección eficaz. Es importante mencionar que este tipo de técnicas, en donde se usan números aleatorios para simular procesos estocásticos típicamente son llamados métodos Monte Carlo [30].

La distribución de probabilidad mencionada anteriormente tiene parámetros determinados por la partícula inicial y la composición del medio. Las partículas son representadas en la computadora como un conjunto de números que nos dan el tipo, la energía, y la posición. Estos números se basan en modelos que siguen el estado de las partículas mientras están interactuando; se va

guardando esta información y da como resultado la simulación de la cascada.

Los métodos Monte Carlo son una rama de la estadística ampliamente estudiada y que actualmente se sigue desarrollando; ya se mencionó la idea fundamental de como son implementados estos métodos en simulaciones de cascadas. Para tener una referencia completa de estos métodos y su implementación en la simulación de cascadas atmosféricas se puede consultar en el trabajo hecho por Messel, H. y Crawford, D. [30].

2.3.2. Modelo de Heitler

El modelo analítico más conocido para modelar cascadas electromagnéticas es el modelo de Heitler. En este modelo solo se toman en cuenta dos interacciones: bremsstrahlung y producción de pares. La descripción de este modelo puede resumirse en los siguientes pasos y su representación gráfica se puede observar en la figura 2.3:

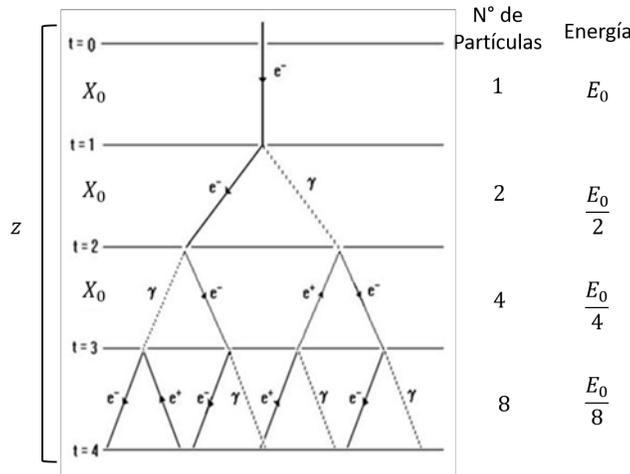


Figura 2.3: Esquema del modelo de Heitler. Imagen modificada de [33].

1. La partícula inicial llegará al medio con una energía inicial E_0 .
2. Esta partícula viajará una cierta distancia en el medio, para después producir un par electrón-positrón (por producción de pares); en el caso de que la partícula inicial sea un fotón. Si la partícula inicial es un electrón, por bremsstrahlung se producirá un fotón. A esa distancia se le conoce como longitud de radiación y se representa mediante el símbolo X_0 . Una definición más formal para longitud de radiación es la siguiente:
 - Distancia para la cual un electrón energético pierde un 63 % de su energía por bremsstrahlung [28].
 - Distancia en donde el fotón recorre aproximadamente $7/9$ del camino libre medio de producción de pares [28].

En cualquiera de los dos casos, cada vez que una partícula viaje en promedio una distancia igual a la longitud de radiación, producirá un par de partículas.

3. El proceso descrito en el paso anterior se irá repitiendo, siempre y cuando la energía de las partículas de la cascada sea mayor a la energía crítica, que es el valor de la energía en el cual

el proceso de ionización es el mecanismo principal para la pérdida de energía superando a la producción de pares y al bremsstrahlung. En el momento en que las partículas alcancen la energía crítica, también se alcanzará el número máximo de partículas producidas en la cascada.

● **Información relevante del Modelo de Heitler**

A partir de este modelo, se puede conocer la siguiente información de la cascada:

- **La profundidad de la cascada:** Supongamos que la cascada electromagnética empieza en el origen y se desplaza en el eje z positivo. La distancia correspondiente a la evolución de la cascada en el punto z , se puede expresar multiplicando t veces la longitud de interacción X_0 :

$$z = X_0 t. \tag{2.1}$$

A la variable t se le conoce como profundidad de la cascada. Esta variable es adimensional y toma valores enteros. Al despejar el valor de t en la ecuación anterior se tendrá que:

$$t = \frac{z}{X_0}. \tag{2.2}$$

La profundidad de la cascada también se ve representada en la figura 2.3.

- **El número de partículas generadas:** Considerando que se producen un par de partículas en cada longitud de interacción, el número de partículas generadas en el paso t será:

$$N = 2^t. \tag{2.3}$$

- **Energía de cada partícula:** Suponiendo que la energía de la partícula inicial se divide proporcionalmente en cada una de las partículas generadas. La energía en la profundidad de la cascada t será igual a:

$$E = \frac{E_0}{2^t}. \tag{2.4}$$

- **Profundidad máxima de la cascada:** La profundidad máxima de la cascada se alcanzará cuando la energía de las partículas sea igual a la energía crítica:

$$E_c = \frac{E_0}{2^{t_{max}}}. \tag{2.5}$$

Entonces la profundidad máxima de la cascada será:

$$t_{max} = \frac{1}{\ln(2)} \ln\left(\frac{E_0}{E_c}\right). \tag{2.6}$$

- **Número máximo de partículas generadas:** El número máximo de partículas se alcanzará en la profundidad máxima de la cascada:

$$N_{max} = 2^{t_{max}} = \frac{E_0}{E_c}. \tag{2.7}$$

- **Edad de la cascada:** Un parámetro de suma importancia que proporciona el modelo de Heitler, es la edad de la cascada (s) y se define como:

$$s = \frac{3t}{t + 2t_{max}}. \tag{2.8}$$

Este parámetro nos da información sobre la evolución de la cascada:

- Para $s < 1$ el número N de partículas en la cascada aumenta.
- Para $s = 1$ el número N de partículas llega a su máximo.
- Para $s > 1$ el número N de partículas en la cascada disminuye.

Recordemos que en esta sección solo se presentó el modelo de Heitler que describe la evolución de una cascada electromagnética. Pero este modelo se puede extender para la descripción de una cascada hadrónica o la combinación de ambas, en donde la cascada incluye una parte hadrónica y una parte electromagnética [31].

2.3.3. Ejemplo de cascada electromagnética en GEANT4

A continuación se mostrará un ejemplo de una cascada electromagnética generada en dos medios distintos: hielo y agua. Esta simulación se elaboró en GEANT4 y fue tomada del artículo hecho por L. Rädcl y C. Wiebusch [32]. En este artículo se estudian algunas propiedades de la cascada, como la cantidad de luz generada por efecto Cherenkov, el perfil de la cascada y la distribución angular de las partículas, por mencionar algunas variables. El objetivo de mostrar este ejemplo es observar la información que se puede obtener de la cascada electromagnética a partir del método Monte Carlo.

Descripción de la simulación

En el ejemplo se simulan tres partículas diferentes: un positrón, un electrón y un fotón; cada una de estas genera una cascada electromagnética. La partícula inicial se sitúa en la parte de abajo y en el centro de un cilindro de 40 m de altura y 60 m de diámetro (con estas dimensiones se tendrán bien confinadas a las partículas secundarias). La cascada se propagará hacia arriba en dirección del eje positivo z (ver inciso a) de la figura 2.4) y el cilindro podrá estar lleno de agua o de hielo. El modelo de la cascada se puede observar en el inciso b) de la figura 2.4, y en cada segmento i de la cascada, se va guardando la siguiente información: la longitud del segmento l_i , el factor de Lorentz β_i , la posición z_i y la dirección α_i . Así al sumar sobre todos los segmentos se obtendrá la cantidad de luz producida y la distribución angular. Ese estudio se realizó para diferentes energías de la partícula inicial abarcando desde 1 GeV hasta 10 TeV.

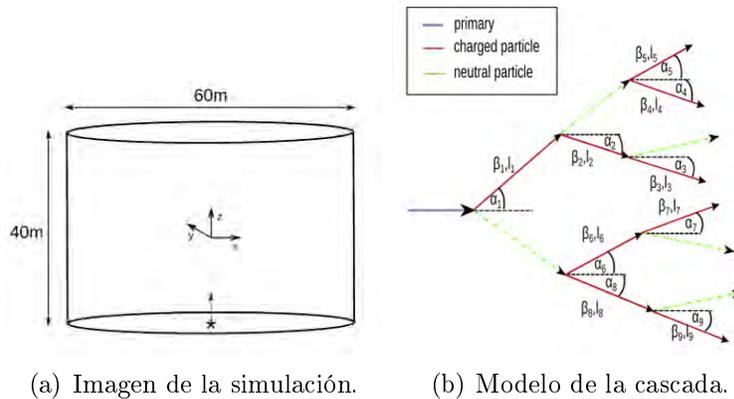


Figura 2.4: Ejemplo de simulación en GEANT4. Imagen tomada de [32].

Parametrización de la cantidad de luz producida

Las partículas generadas en la simulación y que se están propagando en un medio pueden emitir luz debido al efecto Cherenkov. Para una partícula con $\beta = 1$ la cantidad de fotones producidos será proporcional a la longitud del segmento l y se podrá calcular mediante la ecuación de Frank-Tamm (ec. 2.9). Esta ecuación describe la cantidad de fotones (dN) producidas en una distancia (dx). En donde λ es la longitud de onda, α es la constante de estructura fina y z es el número de carga (q/e).

$$\frac{d^2N}{dx d\lambda} = \frac{2\pi\alpha z^2}{\lambda^2} \text{sen}^2(\theta_c) \quad (2.9)$$

Pero para una partícula con $\beta < 1$ la cantidad de fotones producidos será más pequeña y proporcional al factor $\text{sen}^2(\theta_c)$. Teniendo esto en cuenta, la ecuación 2.10 muestra la longitud del segmento l dividido por dicho factor. Con este nuevo valor (\hat{l}) se puede contabilizar de forma adecuada la menor cantidad de fotones producidos.

$$\hat{l} = \frac{\text{sen}^2(\theta_c)}{\text{sen}^2(\theta_{c,0})} l \quad (2.10)$$

En esta ecuación se tendrá que:

$$\text{sen}^2(\theta_{c,0}) = 1 - \cos^2(\theta_{c,0}) = 1 - \frac{1}{n} \quad (2.11)$$

Que es el caso para $\beta = 1$. Con esta consideración, el nuevo valor para la longitud del segmento \hat{l} será equivalente a l y la producción de fotones será la misma, con la ventaja de que los resultados presentados en la referencia [32] pueden ser reescalados para diferentes índices de refracción y son independientes del intervalo de longitud de onda que se asume para el detector.

Después, mediante histogramas se estudió la distribución de \hat{l} para diferentes partículas primarias con diferentes energías; se observó que los datos se ajustaban a una distribución Gaussiana. Por medio del promedio y la desviación estándar que arrojaba esta distribución se realizó la gráfica mostrada en la figura 2.5.

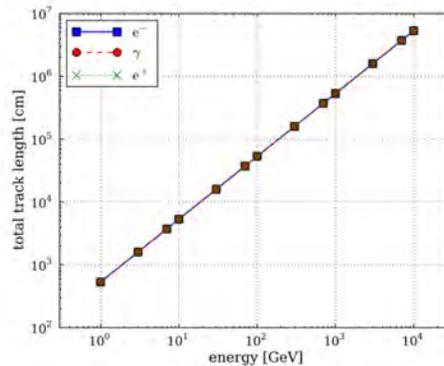


Figura 2.5: Gráfica de \hat{l} en función de la energía. Tomada de [32].

Los puntos de la gráfica anterior se pueden ajustar a una ley de potencia del estilo:

$$\hat{l}(E_0) = \alpha E_0^\beta, \quad \sigma_i(E_0) = \alpha E_0^\beta \quad (2.12)$$

En donde β es cercano a uno, y el valor de ambos coeficientes no cambia mucho al cambiar el tipo de partícula primaria. Como la producción de fotones para \hat{l} y l son iguales, a partir de los datos obtenidos con la simulación se encontró un parámetro que relaciona de manera lineal la cantidad de luz que se produce con la energía de la partícula primaria.

Perfil de la cascada

El perfil longitudinal $\frac{d\hat{l}}{dz}$ a lo largo del eje de la cascada, se puede observar en la figura 2.6. En donde el inciso a) muestra el perfil para rayos gamma de diferentes energías y en el inciso b) se muestra el caso para positrones y electrones de 100 GeV.

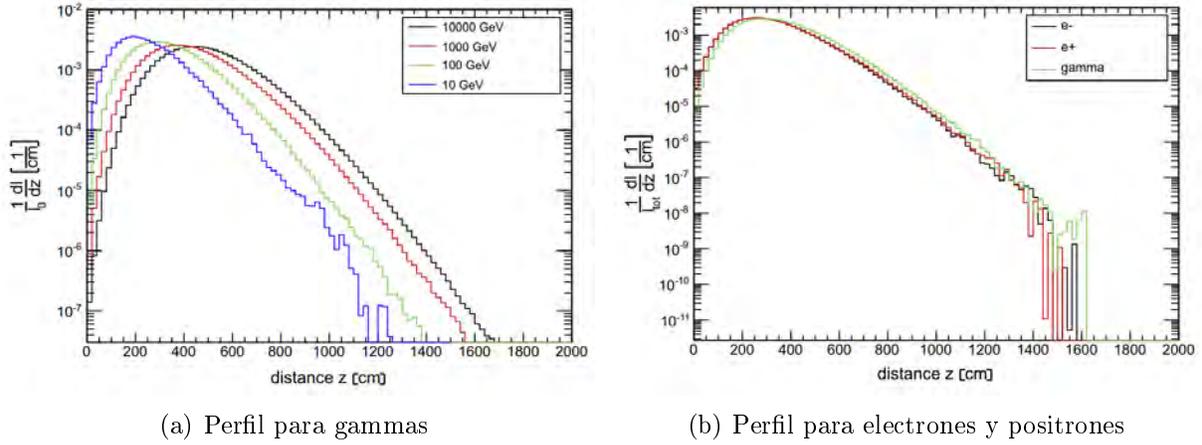


Figura 2.6: Perfil longitudinal de cascada electromagnética (EM). Imagen tomada de [32]

En la imagen anterior se puede observar que la curva crece de forma logarítmica con la energía de la partícula primaria, hasta llegar a la profundidad máxima de la cascada, tal y como es predicho por el modelo de Heitler. A partir de los datos de la simulación se puede hacer una parametrización de este perfil utilizando una distribución gamma que se muestra en la ecuación 2.13.

$$\hat{l}_{tot}^{-1} \frac{d\hat{l}}{dt} = b \frac{(bt)^{a-1} e^{-bt}}{\Gamma(a)} \quad (2.13)$$

En donde t es la profundidad de la cascada y las constantes a y b son determinadas a partir de los datos de la simulación.

Con esto se puede observar, que los modelos basados en técnicas Monte Carlo nos dan información detallada de cascadas electromagnéticas, y nos permiten hacer una parametrización tanto de luz Cherenkov emitida, como del perfil longitudinal de la cascada. Por lo tanto aunque los métodos Monte Carlo se basen en el uso de números aleatorios para poder funcionar, si en la simulación se incluye una cantidad suficiente de eventos podemos entender lo que está sucediendo en la simulación, tal y como se hizo en ambas parametrizaciones, esto se puede extender para el caso de una simulación en HAWCSIM, ya que el funcionamiento de esta herramienta se basa en GEANT4.

2.4. Instalación de HAWCSIM

Los archivos que necesita HAWCSIM para poder funcionar ya vienen incluidos en AERIE, y para poder usar este programa solo hace falta agregar el programa de GEANT4 a los paquetes

externals. Así que el procedimiento para la instalación de HAWCSIM es prácticamente el mismo al que se mencionó en la sección 2.1. La única diferencia consiste en especificar la inclusión de GEANT4 en el archivo APERC, para que este software pueda ser instalado junto con los otros programas que conforman a los paquetes *externals*. La estructura del archivo APERC que incluye a GEANT4 es la siguiente:

```
[DEFAULT]
base = <carpeta>/externals/2.02.02/
build = /tmp/<usuario>-apebuild
[package externals]
dependencias.append = geant4
[APE]
jobs = 2
mirrors = mx us
```

Después solo hace falta instalar los archivos *externals* y compilar AERIE, tal y como se menciona en la sección 2.1. Una vez finalizada la instalación, se puede ejecutar el siguiente comando en la terminal:

```
usuario@maquina: $ hawcsim-exe - -help
```

Este comando nos mostrará las opciones disponibles de ejecución a la hora de hacer una simulación. Si la instalación tuvo éxito se ejecutará ese comando sin problemas y se mostrará un mensaje similar al que se muestra en la figura 2.7.

```
jose@jose-Aspire-E51-131:~$ hawcsim-exe --help
INFO [hawcsim.cc, main:76]:
  hawcsim v4.2 2015-12-18
  HAWC Tank Simulation
  contact: Brian Baughman (bbaugh@undgrb.umd.edu)
INFO [CommandLineConfigurator.cc, ParseCommandLine:124]:
Allowed options:
Generic options:
-h [ --help ]           Write this message
-v [ --version ]       Print Aerie version info
-V [ --verbosity ] arg (=2) Set the verbosity of the logging system:
                        0=trace, 1=debug, 2=info, 3=warn, 4=error,
                        5=fatal
-x [ --fpexcept ] arg (=0) Enable floating-point exceptions for
                        debugging: 1=invalid-arg, 4=div-by-zero,
                        8=overflow, 13=all
--timeLog              Include current UT time in AERIE logs.
Configuration:
-i [ --input ] arg (=./DAT000001) Input base filename.
                        Will be appended to for certain output
                        options.
--itype arg (=corsika) Define input type:
                        corsika or ascii
-o [ --output ] arg (=test.dat) Output filename.
--otype arg (=ascii)   Define output type:
                        xcdf or ascii
--macro arg            Macro to run.
-R [ --nReuse ] arg (=0) Number times events will be reused from
                        input file.
--Nstart arg (=0)      Starting event in input file.
--Nrd arg (=1)         Number of events to read from inputfile.
--weight arg (=1)     If >0 overrides setting in settings files.
                        in tanks given in meters.
--seed arg (=1513121734) Random number generator seed.
--aseed arg (=15131217340) Random number generator seed for placement
                        of events.
--onlyhits            Flag to save ONLY events with hits.
```

Figura 2.7: Resultado al ejecutar el comando `hawcsim-exe -help`

2.4.1. Configuración de HAWCSIM

Para que HAWCSIM pueda funcionar de manera adecuada, es necesario definir la variable de entorno `HAWCSIM_CONFIG`, que apuntará a la carpeta “config” que contiene los archivos de configuración de este programa. Esta variable se define mediante el siguiente comando:

```
usuario@maquina: $ exportHAWCSIM_CONFIG="$Carpeta/aerie/src/hawcsim/config"
```

La carpeta “config” se encuentra adentro de la carpeta de AERIE y su ruta se especifica en el comando anterior. En particular HAWCSIM necesita de 8 archivos para que se ejecute de forma correcta; éstos son los siguientes:

- **settings.dat:** Este archivo contiene los parámetros principales para la simulación, por ejemplo los valores para la distribución del núcleo de cascada.
- **geometry.dat:** Incluye toda la información de la geometría que será simulada y es en donde se mencionan las características de los detectores de HAWC.

El archivo geometry.dat apunta a otros tres archivos que necesita el programa para funcionar de forma correcta, estos son:

- **Survey.xml:** En este archivo están las coordenadas de cada uno de los detectores que conforman al observatorio.
- **ChannelStatus.xml:** Contiene el estado del funcionamiento de los detectores de HAWC.
- **materials.dat:** Aquí se especifican las características de los materiales de los detectores que fueron tomados en cuenta para la simulación.

Tanto el archivo Survey.xml como el archivo ChannelStatus.xml necesitan de tres archivos para funcionar: DetectorLayout.xsd, HAWCStatusMap.xsd y HAWCSchemaTypes.xsd [34]. Con esto se completan los 8 archivos usados en la configuración.

Hay algunas ocasiones en donde se presentan problemas al correr una simulación debido a que los archivos que contiene la carpeta “config” no están actualizados o tuvieron algún problema al ser cargados; esto dependerá de la versión de AERIE que se esté utilizando. Una versión de esta carpeta que funciona sin problemas se puede descargar del repositorio de HAWC mediante el siguiente comando:

```
usuario@maquina: $ svn checkout https://private.hawc-  
observatory.org/svn/hawc/workspaces/aerie/branches/v2-02/src/hawcsim/config/
```

Por último, se tiene que reemplazar el archivo Survey.xml por el archivo HAWCRealSurvey-newnorth.xml que está en la carpeta config-hawc/survey/gen, el cual contiene la posición real de los detectores del observatorio. El otro archivo tiene las posiciones de los detectores del observatorio cuando fue diseñado y por lo tanto no están actualizadas. Una vez que el archivo HAWCRealSurvey-newnorth.xml esté en la carpeta “config”, se debe cambiar su nombre al de Survey.xml.

2.5. Entradas y Salidas en HAWCSIM

Para llevar a cabo una simulación en HAWCSIM se pueden utilizar dos tipos de entradas, en donde se especifican las características de la simulación que se desea realizar. Estas entradas son las siguientes:

- Se usa un archivo CORSIKA para especificar las características de la simulación. Por lo general este tipo de entrada se usa para simular cascadas atmosféricas. El comando para iniciar la simulación tiene la siguiente estructura:

```
usuario@maquina: $ hawcsim-exe - -input /direccion/de/DAT000756.gz - -itype corsika -  
-output /direccion/de/output.dat - -otype ascii - -seed 1284129413
```

En donde:

- hawcsim-exe: Indica que se usará el programa HAWCSIM.
- -input /direccion/de/DAT000756.gz: Indica el archivo CORSIKA a utilizar.
- -itype corsika: Declara el tipo del archivo de entrada.
- -output /direccion/de/output.dat: Indica el nombre del archivo de salida.
- -otype ascii: Declara el tipo del archivo de salida.
- -seed 1284129413: Semilla para la generación de números aleatorios.

Como resultado de la simulación se obtiene un archivo de salida de tipo ascii con el formato específico del observatorio MILAGRO y HAWC [34].

- La otra opción consiste en especificar los datos de la simulación en un archivo de entrada en formato ascii, en donde se coloca la siguiente información: el número y el tipo de partícula a simular, la posición y el vector momento de la partícula. El comando para iniciar la simulación tiene la siguiente estructura:

```
usuario@maquina: $ hawcsim-exe - -input /direccion/de/Entrada.txt - -itype ascii - -output  
/direccion/de/salida.xcdf - -otype xcdf - -seed 1284129413
```

La estructura de este comando es básicamente la misma que en el caso anterior, y en este caso se obtendrá un archivo de salida en formato xcdf que podrá ser reconstruido y transformado a formato root, en donde se agrupará la información más relevante de la simulación en diferentes ramas.

En este trabajo se utilizaron simulaciones en donde el archivo de entrada estaba en formato ascii, ya que nos da la posibilidad de estudiar un tipo de partícula en específico que interactuaba con los tanques del observatorio. En la siguiente sección se dará un ejemplo de este tipo de simulación, que nos servirá como referencia a la hora de explicar las características de las simulaciones principales que se usaron en este trabajo.

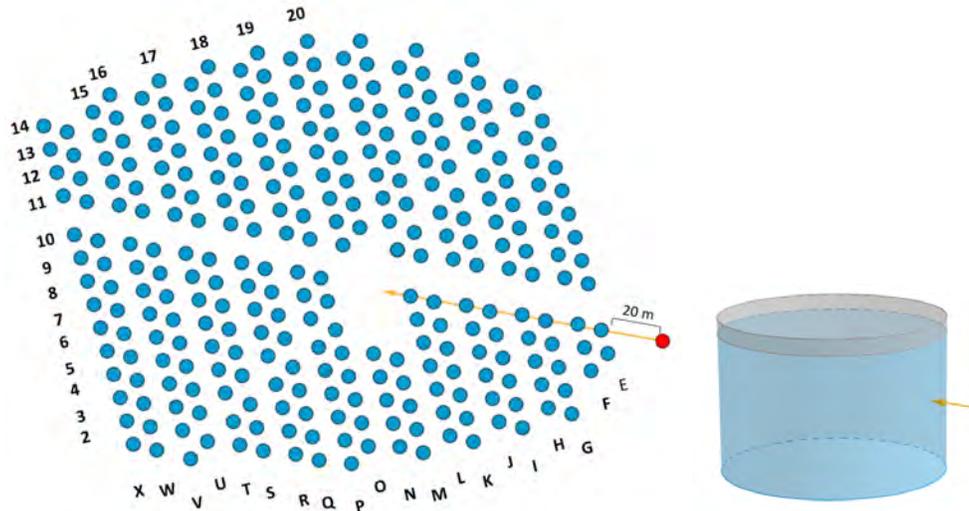
2.6. Ejemplo de una simulación en HAWCSIM

En este ejemplo se mostrará la forma en la que se simulan 10 muones que se propagan de manera horizontal con una energía de 100 GeV. La posición inicial de las partículas simuladas deberá estar a una distancia de 20 m de separación del PMT central del tanque E10. Los muones simulados deberán pasar por la mitad del tanque (en altura) y por encima del PMT central tanto del tanque E10 como del tanque L10 (ver figura 2.8).

2.6.1. Procedimiento para hacer la simulación

Recordemos que el archivo de entrada para realizar la simulación está en formato ascii, y en este caso la información que se requiere para hacer la simulación es la siguiente:

- Tipo de partícula a simular.



(a) Posición de partícula inicial (punto rojo) respecto al arreglo de HAWC. (b) Altura por donde pasarán las partículas.

Figura 2.8: Esquema del ejemplo anterior en donde se ilustra la posición inicial de las partículas y la dirección de su propagación.

- Coordenadas del vector de momento de la partícula.
- Posición de inicio de la partícula.

Así que el procedimiento para realizar esta simulación consiste en encontrar esos tres valores.

Procedimiento para encontrar el tipo de partícula a simular

Los tipos de partículas que se pueden simular con este programa y la forma en la que se declaran se pueden encontrar en el archivo ParticleIDManager.hh ubicado en la carpeta hawcsim/src. Los muones de carga negativa se declaran en la simulación como mu^- .

Procedimiento para encontrar las coordenadas del vector de momento (P)

En este ejemplo se busca que las partículas simuladas pasen por encima del PMT central, tanto del tanque E10 como del tanque L10, así que nos fijamos en la posición de estos detectores. Según el archivo survey.xml la posición del PMT central del tanque E10 y del tanque L10 es la siguiente:

E10	cm.	L10	cm.
x	12568.04	x	6329.86
y	22120.77	y	23255.99
z	4.83	z	2.37

Para indicar que el muón vaya del tanque E10 al tanque L10, obtenemos las coordenadas del vector P al restar las coordenadas del tanque E10 a las coordenadas del tanque L10. Así que el vector de momento se podrá expresar como:

$$P = (P_x, P_y, P_z).$$

En donde $P_x = 6329.86 - 12568.04 = -6238.18$ y $P_y = 23255.99 - 22120.77 = 1135.22$. Para definir que el muón sea horizontal, la coordenada z será igual a cero ($P_z = 0$). Entonces este vector se podrá expresar como:

$$P = (-6238.18, 1135.22, 0) \text{ GeV}/c^2.$$

Ahora normalizamos el vector anterior al dividir cada entrada por la norma del vector.

$$\hat{P} = (-0.9838, 0.1790, 0) \text{ GeV}/c^2, \text{ tal que } \|\hat{P}\| = 1.$$

Para altas energías el momento y la energía de las partículas son prácticamente el mismo valor. Para simular un muón de 100 GeV, multiplicamos por 100 cada una de las coordenadas de \hat{P} . Entonces:

$$\hat{P} = (-98.38, 17.90, 0) \text{ GeV}/c^2, \text{ tal que } \|\hat{P}\| = 100.$$

Así que ya tenemos el vector de momento para un muón horizontal que va del tanque E10 al tanque L10 con una energía de 100 GeV. Es importante destacar que las unidades de este vector deben estar en GeV/c^2 , ya que estas son unidades que utiliza el programa para la simulación.

Procedimiento para encontrar la posición inicial

Se debe tener en cuenta que las unidades de las coordenadas que indican la posición inicial de la partícula deben estar en centímetros. En este ejemplo se pide que la partícula inicial esté a una distancia de 20 m. (2000 cm.) del PMT central de E10, así que multiplicamos por 2,000 las coordenadas del vector $-\|\hat{P}\|$, para que la posición de la partícula esté en la línea que pasa por el PMT central del tanque E10 y del tanque L10; el signo menos se debe a que el vector \hat{P} apunta en dirección al tanque L10, y al multiplicar por un signo menos éste apuntará en la dirección opuesta. Con esto, los muones iniciarán a una distancia de 20 m del tanque E10, pasarán por encima del PMT central y llegarán hasta el tanque L10.

$$-2000\hat{P} = (1967.68, -358.08, 0.0) \text{ GeV}/c^2.$$

Por último, solo sumamos las coordenadas de este vector a la ubicación del PMT central del tanque E10. Entonces la posición de la partícula primaria será:

$$\begin{aligned} x &= (1967.60 + 12568.04) \text{ cm} = 14535.64 \text{ cm}. \\ y &= (-358.08 + 22120.77) \text{ cm} = 21762.69 \text{ cm}. \end{aligned}$$

Para que esté a la mitad del tanque, sumamos 225 cm. a la coordenada z del PMT central del tanque E10. Esta es la altura de la mitad del tanque que contiene agua.

$$z = (4.83 + 225) \text{ cm} = 229.83 \text{ cm}.$$

Con este último dato ya tendremos la ubicación de la partícula inicial.

Sistema de referencia del archivo de entrada

Es de gran importancia mencionar, que la posición de los detectores descrita en el archivo `survey.xml` es diferente a la posición de los detectores que el programa utiliza para hacer la simulación. Así que tendremos dos sistemas de referencia: el sistema de referencia del archivo `survey.xml`, al cual llamaremos O , y el sistema de referencia que usa el programa para funcionar, al cual llamaremos O' . El centro del sistema O' visto desde el sistema O viene especificado en el archivo `survey.xml`. Por lo tanto hay una traslación entre ambos sistemas de referencia, y se ve representado en la figura 2.9.

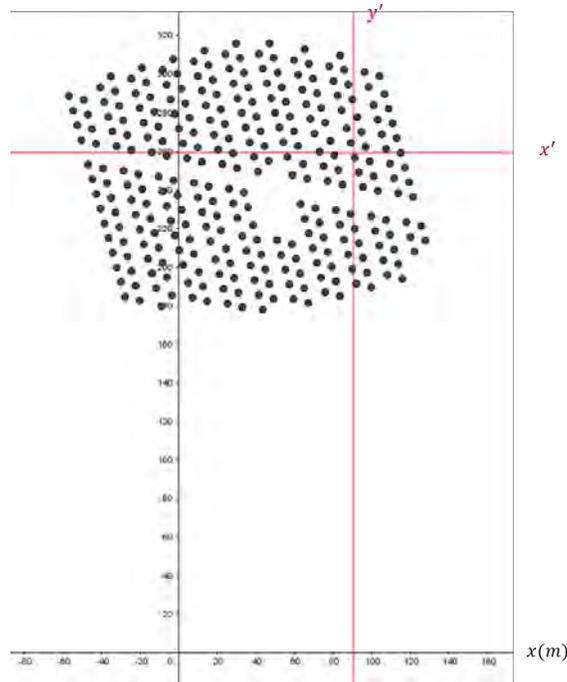


Figura 2.9: Ubicación del sistema de referencia O' (ejes color rojo) visto desde el sistema de referencia O (ejes color negro).

Las coordenadas del centro del sistema O' visto desde el sistema O son las siguientes:

$$c = (9065.23, 25936.63, 0.0) \text{ cm.}$$

Por lo tanto, para describir la posición de la partícula inicial en el sistema O' , se deberá restar la posición del centro que viene en el archivo `survey.xml` a las coordenadas de la ubicación inicial de la partícula a simular.

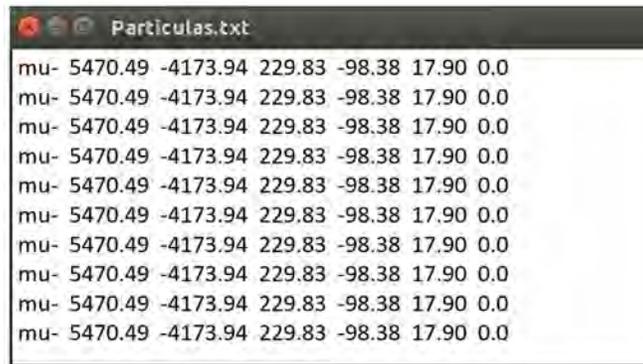
$$\begin{aligned} x' &= x - c_x = (14535.64 - 9065.23) \text{ cm} = 5470.49 \text{ cm.} \\ y' &= y - c_y = (21762.69 - 25936.63) \text{ cm} = -4173.94 \text{ cm.} \\ z' &= z - c_z = (229.83 - 0.0) \text{ cm} = 229.83 \text{ cm.} \end{aligned}$$

Con esta última consideración ya tendremos toda la información necesaria para hacer la simulación en el sistema de referencia correcto.

Entrada de la simulación

La entrada de la simulación se coloca en un archivo de texto, por ejemplo `Particulas.txt`. Primero se coloca el tipo de partícula, después las coordenadas que indican su posición inicial y por último

se pondrán las coordenadas del vector de momento. Un ejemplo del contenido de este archivo y que indica la información a simular del ejemplo mostrado anteriormente se muestra en la figura 2.10.



```

Particulas.txt
mu- 5470.49 -4173.94 229.83 -98.38 17.90 0.0

```

Figura 2.10: Ejemplo del archivo de entrada en HAWCSIM.

El número de entradas en el archivo indica el número de partículas que se van a simular; como en el ejemplo se pide simular 10 muones, este archivo tendrá 10 líneas.

2.6.2. Ejecución de la simulación

Para ejecutar la simulación el comando a utilizar es el mismo que se mencionó en la sección 2.5. Para el ejemplo mostrado anteriormente el comando es el siguiente:

```

usuario@maquina: $ hawcsim-exe - -input Particulas.txt - -itype ascii - -output salida.xcdf -
-otype xcdf - -seed 1284129413

```

Como resultado tendremos el archivo de salida en formato xcdf. Después se utiliza la herramienta *offline-reconstructor* que viene incluida en AERIE, para hacer la reconstrucción de los datos. El comando a utilizar es el siguiente:

```

usuario@maquina: $ offline-reconstructor - -extended - -config-file HAWCSimConfig.xml - -input
salida.xcdf - -output reco.xcd

```

En el comando anterior se puede observar que para llevar a cabo la reconstrucción de la simulación, es necesario tener en la misma carpeta en la que se está ejecutando el comando el archivo HAWCSimConfig.xml o en su defecto indicar la dirección en donde está. Este archivo se puede encontrar en la carpeta `erie/build/hawcsim/config`.

Por último pasamos el archivo reconstruido a formato root mediante el siguiente comando:

```

usuario@maquina: $ xcdf-root -o reco.root -input reco.xcd

```

En este formato la información de la simulación será almacenada en diferentes ramas de un árbol binario (este formato comúnmente se utiliza en C++). Entre la información que nos proporciona la simulación podemos encontrar el número de detectores activados, la carga registrada y el tiempo de detección. La lista completa de las variables que nos proporciona la simulación se puede encontrar en el archivo *output* que se encuentra en la carpeta “hawcsim”.

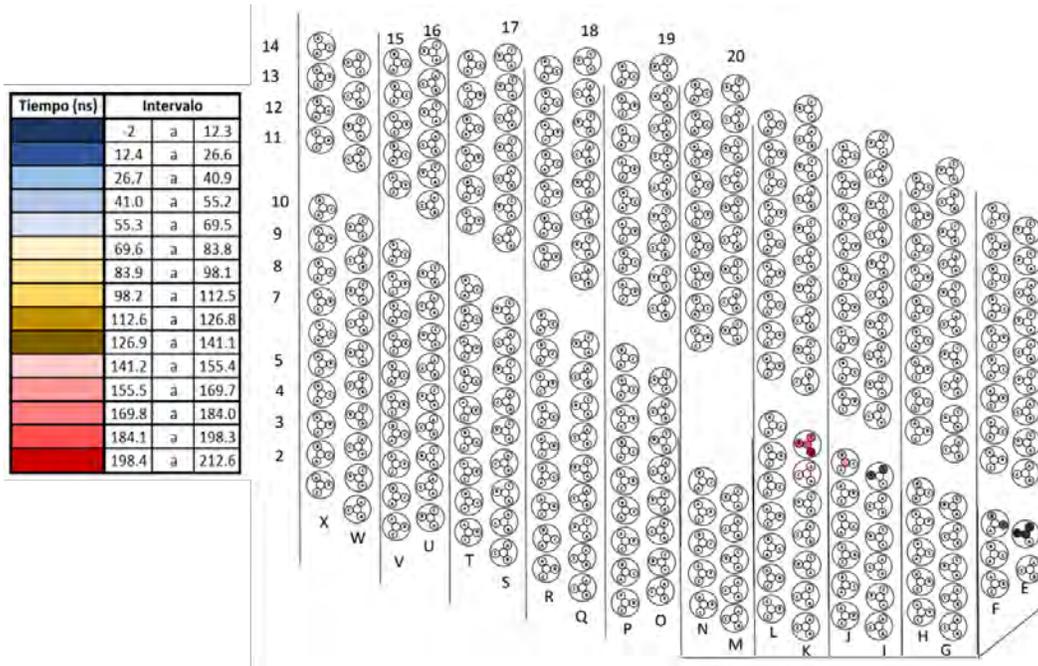


Figura 2.11: Representación del tiempo de detección utilizando los datos de la rama event.hit.time del archivo en formato root para la simulación del ejemplo anterior, en donde solo se muestran los datos de un evento y la escala de colores indican el tiempo conforme se fueron activando los fotomultiplicadores.

Para abrir el archivo de la simulación en formato root se usa el siguiente comando:

```
usuario@maquina: $ root -l reco.root
```

y para ver el contenido del archivo se usa el comando:

```
.ls
```

la simulación está conformada por eventos y el número de éstos depende de las partículas que se simularon. Para entrar a un evento en específico y observar todas las variables que corresponden a ese evento se usa el comando:

```
XCDF->Show(5,10)
```

el número 5 del comando anterior indica el número de evento que se desea visualizar y el número 10 indica el número de datos que se encuentran en las ramas que se quieren ver.

Una visualización rápida de los resultados de la simulación se puede obtener mediante el comando:

```
TBrowser tb;
```

que nos abrirá una ventana en donde se muestran todas las ramas que hay en el árbol binario. Al dar clic en una rama en específico, se podrán visualizar los datos que conforman a esta rama en forma de histograma. Éste se llenará con todos los datos que contiene la rama de cada uno de los eventos que conforman al archivo. Un ejemplo se puede observar en la figura 2.12.

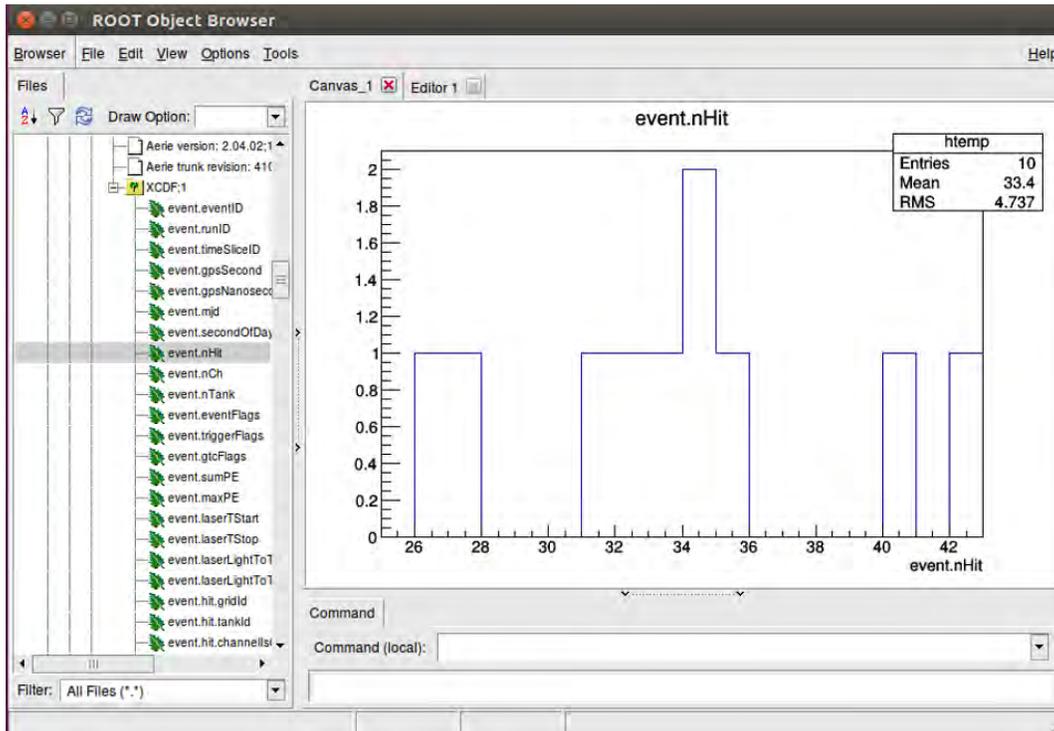


Figura 2.12: Esta es la ventana que aparece al ejecutar el comando `TBrowser tb`; para el ejemplo descrito en esta sección, y en particular se está observando el histograma correspondiente a la rama `event.nHit`, es decir, que el histograma se llenó con los datos correspondientes al número de detectores que se activaron en cada evento.

2.6.3. Ruido y número de eventos en la reconstrucción

En HAWCSIM, es posible modificar los parámetros para el mínimo número de hits necesarios para reconstruir un evento, y de la tasa del ruido que produce señales adicionales en los fotomultiplicadores.

Para ello se debe ingresar al archivo `mc-param.xml` que se encuentra en la carpeta `config-hawc/reconstruction`, y modificar los valores de la tasa del ruido (`noiseRate`) y del mínimo número de hits que se requieren para la reconstrucción (`minimumHits`).

En todas las simulaciones realizadas en este proyecto, se tomaron los siguientes valores para estas variables:

$$\begin{aligned} \text{noiseRate} &= 0 \text{ kHz.} \\ \text{minimumHits} &= 1. \end{aligned}$$

Con este valor para la tasa del ruido, se ahorra tiempo en la duración de cada simulación, además de que se prefirió trabajar de esta manera debido a que el ruido que se observa en los detectores no ha sido aún caracterizado satisfactoriamente. Al considerar el valor de 1 como el mínimo de hits necesarios para hacer la reconstrucción, se toma en cuenta a todos los eventos que por lo menos activaron a un detector; con esto el número de eventos en la simulación era igual al número de partículas simuladas.

En términos generales, las simulaciones Monte Carlo se llevan a cabo en este trabajo por medio del programa HAWCSIM, en donde se simula el paso de una partícula a través de los tanques del observatorio HAWC. Los resultados de estas simulaciones se estudian por medio de histogramas, observando una variable en particular de la simulación. En capítulos posteriores se darán más detalles sobre estas simulaciones.

Capítulo 3

Red Neuronal Artificial

En el capítulo 5, se dará la estructura de la red utilizada así como todos los detalles involucrados en esta prueba para hacer una estimación de la energía de muones horizontales. Por este motivo, para comprender y para poder explicar de manera más sencilla este capítulo, en esta parte se darán todas las características sobre una red neuronal artificial, empezando por la descripción de una red neuronal biológica que motivó el desarrollo de una red neuronal artificial, pasando por la estructura y el modelo matemático que emplea este algoritmo para funcionar, hasta llegar a un ejemplo en concreto en donde se utiliza una red neuronal en ROOT, que es el programa utilizado para el análisis de datos en este proyecto.

3.1. Reconocimiento de números escritos a mano

Una red neuronal artificial es un modelo matemático inspirado en sistemas neuronales biológicos, que puede ser implementado en sistemas computacionales. Para ser más específicos, una red neuronal artificial es una función de R^n a R^m , en donde R^n son los datos de entradas y R^m son los datos de salida [34]. El usuario proporciona una lista de datos de entradas y salidas que tienen como objetivo entrenar a la red neuronal. Como resultado se podrá dar cualquier dato de entrada y la red calculará el dato de salida en base al entrenamiento previo. Por ejemplo, se puede alimentar a una red neuronal artificial con imágenes de números escritos a mano que van del cero al nueve, estas imágenes serán los datos de entrada. Los datos de salida serán las etiquetas que vienen en cada figura y que indican el número al cual corresponde cada imagen (ver 3.1). La idea es que el usuario proporcione una cierta cantidad de datos para entrenar a la red neuronal.

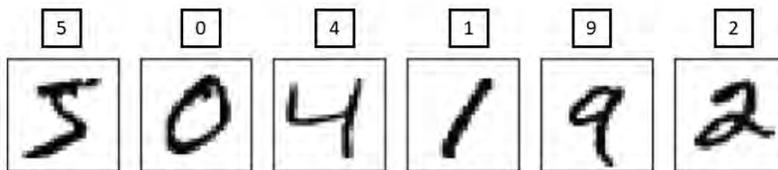


Figura 3.1: Ejemplo de imágenes de números escritos a mano para alimentar a una red neuronal artificial; el cuadro en la parte superior de cada imagen indica el número que se trató de escribir a mano en cada figura. Imagen modificada de [35].

Después se probará este algoritmo utilizando datos de entradas y salidas que la red neuronal no haya visto antes, es decir, que son distintos a los datos de entrenamiento. Con esta información podemos comparar el resultado que arroja la red neuronal con la etiqueta que acompaña a cada

imagen, así se podrá definir un porcentaje de error para este método.

Con este ejemplo se observó en términos generales el funcionamiento de una red neuronal artificial y como se puede emplear para resolver un problema. En este caso se utilizó para el reconocimiento de imágenes, pero realmente su campo de aplicaciones es bastante amplio. Una red neuronal artificial se puede utilizar en problemas de clasificación y reconocimiento de patrones de voz, imágenes, señales o en problemas más específicos, como es el caso de este trabajo, en donde se usará una red neuronal para realizar la siguiente tarea:

- Realizar un estimador de energía para muones horizontales que pasan por los tanques del arreglo de HAWC.

Es importante destacar, que los datos para entrenar y probar la red serán obtenidos por medio de simulaciones.

3.2. Red Neuronal Biológica

Para empezar a hablar de una red neuronal artificial un buen punto de partida es explicar de manera general el funcionamiento de una red neuronal biológica, ya que estos modelos sirvieron de inspiración para la implementación y desarrollo de modelos matemáticos que simulan el funcionamiento de una red neuronal en un ordenador [36].

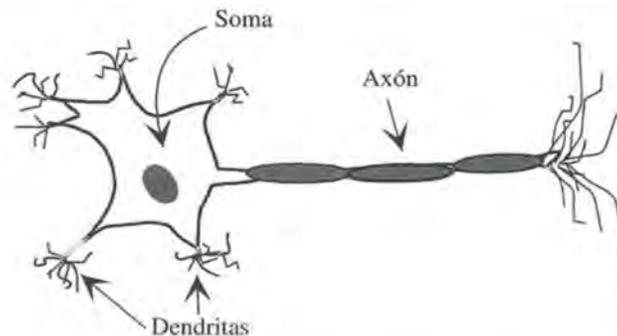


Figura 3.2: Esquema de una neurona biológica, en donde se muestran sus tres componentes principales: Dendritas, Soma y Axón. Imagen tomada de [34].

La unidad elemental de una red neuronal biológica es una neurona (ver figura 3.2), que está conformada por tres componentes que cumplen con una tarea en particular. Estos componentes son los siguientes:

- **Dendritas:** Son las encargadas de recibir la información procedente de otra neurona.
- **Soma:** Se encarga de procesar la información recibida y genera la información de salida.
- **Axón:** Se encarga del envío de la información de salida.

En el caso de los humanos una red neuronal biológica está conformada por cien mil millones de neuronas interconectadas entre sí que se comunican entre ellas por medio de sinapsis [34]; que es el mecanismo en donde las neuronas intercambian información a través de impulsos nerviosos. Es importante destacar que en este procedimiento no existe un contacto físico entre neuronas vecinas, pero la distancia de separación entre ellas es muy pequeña (del orden de 0.3 micras) [34].

La estructura de esta red neuronal se divide en conjuntos conformados por miles de neuronas, y usualmente se suele aceptar que el conocimiento que obtenemos está más relacionado con las conexiones neuronales que con las mismas neuronas. Es decir, que el conocimiento se distribuye a través de la sinapsis de la red [37]. Esta estructura biológica nos permite adquirir conocimiento y tomar decisiones en base a nuestras experiencias pasadas.

Tomando en cuenta las principales componentes de una red neuronal biológica, se empezaron a realizar modelos matemáticos que emulaban el funcionamiento de conexiones neuronales. Uno de los primeros modelos de una red neuronal artificial fue el Perceptrón, que tuvo buena aceptación en su época y sirvió de base para la elaboración de algoritmos neuronales más complejos.

3.3. Red Perceptrón

El Perceptrón es un modelo de una red neuronal artificial desarrollado por Frank Rosenblatt durante las décadas de 1950 y 1960 [35]. En este modelo se define la estructura de una neurona artificial que trata de emular el funcionamiento de una neurona real. La estructura de este modelo se ve representado en la figura 3.3.

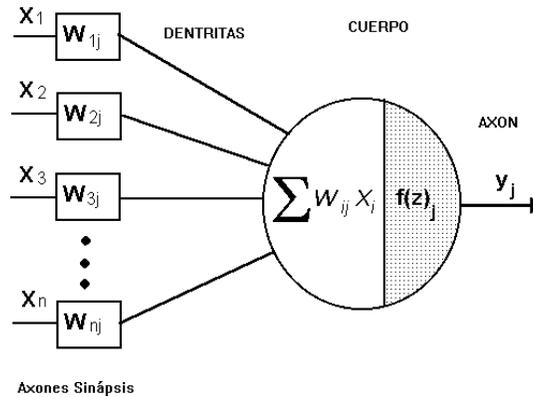


Figura 3.3: Estructura de una neurona artificial utilizada en el modelo de Rosenblatt. Imagen tomada de [38].

Funcionamiento: Esta neurona artificial, al igual que en el caso de una neurona real recibe información, la procesa y genera una señal de salida. En este modelo, los datos de entrada son números y se representan por medio de letras $x_1, x_2, x_3, \dots, x_n$ (ver figura 3.3). La forma en la que se genera una señal de salida es por medio de pesos $w_1, w_2, w_3, \dots, w_n$ que son números reales que indican la importancia de cada entrada para generar el dato de salida. Esta salida solamente podrá elegir entre dos valores, el 0 y el 1. La neurona artificial calculará la suma del producto de los datos de entrada con sus pesos ($\sum_{i=1}^n x_i w_i$), y el dato de salida dependerá de qué el valor de esta suma sea mayor o menor a cierto valor umbral. Tomando en cuenta lo anterior, la señal de salida podrá ser representada mediante la siguiente función:

$$Dato\ de\ salida = \begin{cases} 0 & \text{si } \sum_{i=1}^n x_i w_i \leq \text{valor umbral} \\ 1 & \text{si } \sum_{i=1}^n x_i w_i > \text{valor umbral} \end{cases} \quad (3.1)$$

Básicamente así es como funciona este modelo, y se puede pensar como un dispositivo que puede tomar decisiones al colocar pesos a los datos de entrada. Para ver esta analogía tomaremos un

ejemplo parecido al que se muestra en la referencia [35].

Ejemplo: Supongamos que una persona quiere tomar la decisión de subir o no a una montaña rusa. Y para tomar esta decisión solo toma en cuenta los siguientes tres factores que representan las entradas de nuestra neurona artificial:

- x_1 : Si hay algún amigo que quiera subirse con él.
- x_2 : Si la persona se marea con facilidad.
- x_3 : Si el boleto de acceso es barato.

El valor que tomará cada dato de entrada será un valor binario y dependerá de las condiciones en la que se encuentre la persona, por ejemplo:

- $x_1 = 1$. Si en efecto hay un amigo que quiera subirse con él.
- $x_2 = 1$. Si la persona se marea con facilidad.
- $x_3 = 0$. Si el boleto de entrada es caro.

El valor de los pesos dependerá de qué tan importante es cada dato de entrada para la persona que se subirá a la montaña rusa, por ejemplo:

- $w_1 = 11$. Si le importa mucho ir acompañado.
- $w_2 = 2$. Si no le importa mucho sentirse mal durante el juego.
- $w_3 = 3$. Si no le importa mucho el costo del boleto.

Por último solo hace falta definir un valor umbral por ejemplo 5. Con este último dato podemos calcular la decisión que tomó la red neuronal artificial. Para nuestro ejemplo:

$$\sum_{i=1}^3 x_i w_i = (1)(11) + (1)(2) + (0)(3) = 13.$$

Como la suma anterior es mayor al valor umbral, el dato de salida será igual a 1, que indicará que la persona sí se subirá a la montaña rusa. Tomando en cuenta este ejemplo, si se varía el valor de los pesos y del valor umbral, se habrá modelado un mecanismo simple para tomar decisiones. Por lo tanto, la forma en la que se modifican los pesos y los valores umbrales resulta la pieza fundamental del funcionamiento de este modelo.

Usualmente se suele simplificar la función de salida mostrada en la ecuación 3.1 haciendo los siguientes cambios:

1. Si consideramos que los datos de entrada y sus correspondientes pesos son elementos de un vector, podemos escribir la suma descrita anteriormente como el producto punto entre ambos vectores:

$$x \cdot w = \sum_{i=1}^n x_i w_i.$$

2. El segundo cambio consiste en pasar del otro lado al valor umbral en la desigualdad mostrada en la ecuación 3.1. Haciendo la siguiente definición:

$$b = - \text{valor umbral.}$$

Normalmente a este valor “ b ” se le conoce en todas las referencias de redes neuronales en inglés como “*bias*”, y se puede pensar como un valor que determina qué tan difícil es obtener una salida igual a uno.

Tomando en cuenta estos cambios, la función que determina la salida de la neurona artificial quedará como:

$$\text{Dato de salida} = \begin{cases} 0 & \text{si } x \cdot w + b \leq 0 \\ 1 & \text{si } x \cdot w + b > 0 \end{cases} \quad (3.2)$$

El Perceptrón fue un modelo muy popular en su época que atrajo la atención de muchos investigadores, pero lamentablemente este modelo no podía usarse para simular algunas compuertas lógicas, por ejemplo las compuertas AND y XOR (ver tabla 3.1) que son muy usadas en el ámbito computacional [36].

Entrada		Salida	
A	B	A AND B	A XOR B
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	0

Tabla 3.1: Tablas de verdad de compuertas AND y OR.

Por este motivo se desarrolló un modelo de una red neuronal artificial más sofisticada que podía resolver ese tipo de problemas, y que será el tema central de la siguiente sección.

3.4. Red Perceptrón Multicapa

El Perceptrón Multicapa es un modelo de red neuronal artificial que fue desarrollado tomando como base el modelo de una neurona artificial Perceptrón, pero a diferencia de este, puede simular cualquier compuerta lógica. Para llevar a cabo esta tarea, se debe tomar en cuenta que el modelo de neurona artificial descrito en la sección anterior puede simular una compuerta lógica NAND [35]. Una forma para simular este tipo de compuerta es considerar una neurona artificial con dos entradas, cada una con un peso de -2 y una *bias* de 3. El esquema de esta neurona se puede observar en la figura 3.4.

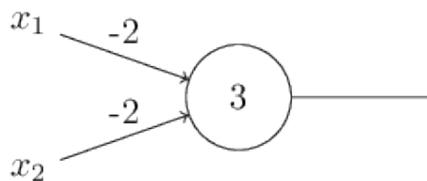


Figura 3.4: Estructura de una neurona artificial para simular una compuerta NAND. Imagen tomada de [35].

Con la estructura anterior se puede simular una compuerta NAND. Por ejemplo si $x_1 = 0$ y $x_2 = 1$ entonces:

$$\sum_{i=1}^2 x_i w_i = (0)(-2) + (1)(-2) + 3 = 1.$$

Como el valor de la suma anterior es mayor a cero, la salida de la función sera igual 1, tal y como ocurre en una compuerta NAND. Si colocamos los otros datos de entrada $(0, 0)$, $(1, 0)$ y $(1, 1)$ y calculamos el dato de salida se obtendrían los mismos resultados que en una compuerta NAND (ver tabla 3.2).

Entrada		Salida
x_1	x_2	NAND
0	0	1
0	1	1
1	0	1
1	1	0

Tabla 3.2: Tablas de verdad de la compuerta NAND.

Por lo tanto el modelo de neurona artificial Perceptrón puede simular una compuerta lógica NAND. Tomando en cuenta que un arreglo de varias compuertas NAND pueden simular cualquier compuerta lógica [35], entonces una red conformada por varias neuronas de este tipo, podrá simular las compuerta lógicas AND y XOR. Por ejemplo, una red neuronal que simula la suma de dos bits se puede construir a partir de un arreglo entre neuronas que simulan una compuerta NAND. La estructura de esta red se puede observar en la figura 3.5.

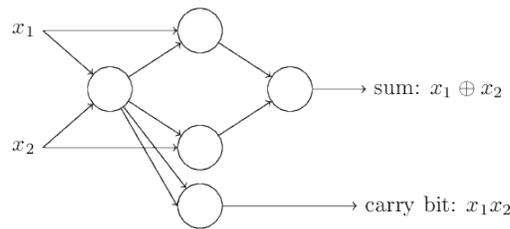


Figura 3.5: Red neuronal para sumar dos bits. Imagen tomada de [35].

Esta red neuronal puede realizar la suma de dos bits (en la imagen a esta suma se le representa mediante el símbolo \oplus). Se puede observar que esta red neuronal tiene dos entradas y dos salidas. Los datos de entradas son los bits a sumar, y los datos de salida son la suma normal entre los dos números que representan a cada bit (que se denotan por la letra S) y el número de acarreo (representado por la letra C), que es una cantidad que usualmente se maneja en este tipo de cálculos. Por ejemplo al sumar el bit $x_1 = 1$ con el bit $x_2 = 1$ la suma normal entre estas dos cantidades será igual a 0 (ya que el número 2 no se encuentra en el sistema binario) y el número de acarreo será igual a 1, por lo tanto $x_1 \oplus x_2 = 10$. Al considerar los mismos pesos y bias que se definieron para la compuerta NAND, los resultados de esta red neuronal para la suma de los bits $(0, 0)$, $(1, 0)$, $(1, 0)$ y $(1, 1)$ se muestran en la tabla 3.3.

En esta tabla se puede observar que los resultados para la suma (S) y para el número de acarreo (C) son los mismos que para las compuertas XOR y AND respectivamente. Con este ejemplo

Entrada		Salida	
x_1	x_2	S	C
0	0	0	0
0	1	1	0
1	0	1	0
1	1	0	1

Tabla 3.3: Resultados para la suma de 2 bits

se puede observar que al agregar más neuronas artificiales y agruparlas en diferentes capas, se soluciona el problema de simular cualquier compuerta lógica.

Ahora retomemos el ejemplo mencionado en la parte inicial de este capítulo, en donde se dijo que una red neuronal artificial se puede utilizar para el reconocimiento de números escritos a mano. Se había comentado que la idea de este modelo consiste en entrenar una red neuronal por medio de imágenes de números escritos a mano, que ya se habían clasificado, y por ende se tenía una etiqueta que indicaba el número al que corresponde cada imagen. Lo que hace la red neuronal para lograr este reconocimiento, es que irá modificando el valor de los pesos y los *bias* que conforman a la red neuronal a medida que el usuario vaya presentando datos de entrenamiento. Supongamos que utilizamos como dato de entrada la imagen de un 5, y como resultado la red neuronal nos dice que esa imagen corresponde a un 7. Se puede ver que la red neuronal aun no funciona bien, y para que la próxima vez no se equivoque, se deberán hacer pequeños cambios a los valores de pesos y *bias*. Recordemos que la función que define al dato de salida de una neurona artificial solo toma dos valores, el 0 y el 1 (ver ecuación 3.2). Por lo tanto, si realizamos pequeños cambios a los valores de pesos y *bias*, causarían saltos abruptos entre 0 y 1 que no ayudaran a la red neuronal. Para solucionar este problema se opta por utilizar la función sigmoide $\sigma(z)$ para producir el dato de salida de cada neurona artificial.

$$\sigma(z) = \frac{1}{1 + e^{-z}}. \quad (3.3)$$

La gráfica de la función 3.3 se puede observar en la figura 3.6, en donde se puede ver su comportamiento. Esta función solo permite valores de salida entre 0 y 1 y su centro se encuentra en $\sigma(0) = 0.5$. Mientras más grande sea el valor de z esta función se ira acercando a 1, y mientras más pequeño sea el valor de z la función se acercará a 0.

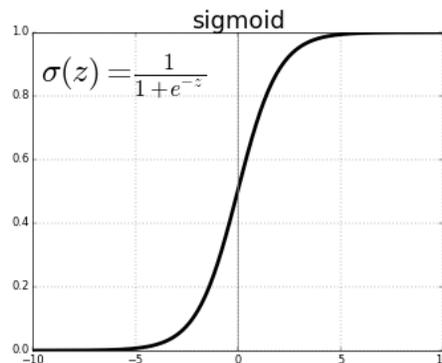


Figura 3.6: Gráfica de la función sigmoide. Imagen tomada de [39].

Si se considera la función $\sigma(z)$ para calcular el dato de salida en cada neurona artificial, se habrá solucionado el problema anterior, ya que el dato de salida será más sensible, y representará mejor estos cambios a la hora de hacer pequeñas modificaciones en los valores de pesos y *bias*, ya que esta función arroja valores continuos entre 0 y 1.

Por lo tanto el modelo Perceptrón Multicapas se obtiene al considerar diversas capas de neuronas artificiales en la estructura de la red, y al usar la función $\sigma(z)$ para calcular el dato de salida en cada neurona artificial. Con estas dos consideraciones se obtiene un modelo de red neuronal artificial más complejo, que puede simular cualquier compuerta lógica y que es más sensible a cambios en pesos y *bias*. Esto causará un mejor aprendizaje a la hora de hacer el entrenamiento en la red neuronal.

3.4.1. Estructura

En la sección anterior solamente se revisó el concepto de Perceptrón Multicapa, ahora procederemos a explicar la estructura que acompaña a este modelo y que se puede representar en la figura 3.7.

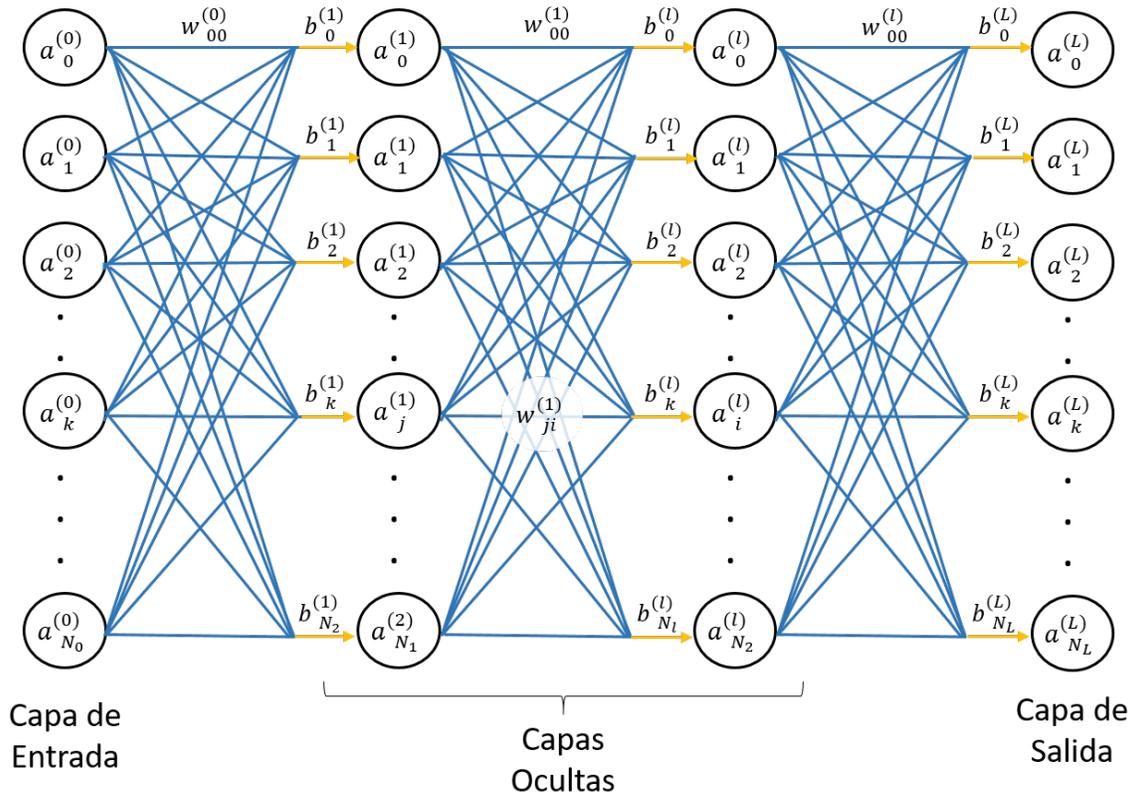


Figura 3.7: Estructura de una red Perceptrón Multicapa

Los componentes de esta red son las siguientes:

- **Neurona:** La neurona artificial es la unidad elemental de la red, y consiste en una función que recibe un dato de entrada y proporciona un dato de salida entre 0 y 1 ($f : \mathcal{R} \rightarrow [0, 1]$). Normalmente a este dato de salida se le conoce como valor de activación, por ejemplo si el valor de una neurona es cercano a 1, se dice que la neurona es muy activa. En la figura 3.7,

a cada neurona se le representa por medio de un círculo, y la activación de cada neurona es representada mediante el símbolo a_k^l , en donde el índice l indica la capa en donde se encuentra ese valor de activación, y el símbolo k hace referencia a la neurona de la capa l de la cual se está hablando. El valor de l estará entre 0 y L , en donde L indica el número de la última capa. Por otro lado, el valor de k estará entre 0 y N_l , en donde N_l indica el número de la última neurona en la capa l .

- **Pesos:** Los pesos son representados en la imagen por las líneas de color azul. Estos valores se van modificando a medida que el usuario presente datos de entrenamientos, y son responsables de que la red funcione de manera correcta. Recordemos que podemos visualizar a estos valores como cantidades que nos indican que tanta influencia tienen los datos de entrada, o de una capa previa en el dato de salida de una neurona en la siguiente capa. A estos pesos los denotaremos mediante el símbolo w_{ji}^l , el índice i estará entre 0 y N_l , y el valor de j estará entre 0 y N_{l+1} . Más adelante se calculará de forma explícita la activación de una neurona artificial, en donde quedará un poco más claro el uso de estos índices.
- **Bias:** Los *bias* son representados en la imagen por medio de flechas amarillas y se pueden visualizar como valores que determinan que tan activa o inactiva es una neurona. Estas se representa mediante el símbolo b_k^l .
- **Capa de Entrada:** Es el conjunto de neuronas que contiene a los datos de entrada que el usuario proporcione, ya sea para entrenar o para probar la red neuronal. En la imagen las letras $a_0^0, a_1^0, a_2^0, \dots, a_{N_0}^0$ representan a estos datos.
- **Capas Ocultas:** Recordemos que el punto clave para que una red funcione de manera correcta está en modificar los valores de pesos y *bias* a través del entrenamiento de la red. Pero estas modificaciones se verán reflejadas en el valor de activación de cada neurona. Las neuronas se pueden agrupar en conjuntos de neuronas que llamaremos capas. Anteriormente se observó que el agregar una capa extra entre los datos de entrada y los datos de salida, ayuda a simular compuertas lógicas que antes no se podían recrear usando solamente una neurona. El Perceptrón Multicapa considera capas de neuronas entre los datos de entrada y los datos de salida, a las que usualmente se le llaman capas ocultas. Estas capas son implementadas en este modelo para resolver problemas más sofisticados. Tanto el número de capas ocultas como el número de neuronas que se utiliza en cada capa, dependerá del problema que el usuario este intentando resolver. Realmente no hay una regla que nos indique el número de capas ocultas que debemos usar. Comúnmente el usuario decide el número de capas a utilizar en base a su experiencia y a los resultados que va obteniendo [40].
- **Capa de Salida:** Es el conjunto de neuronas que contiene los datos de salida de la red. En la imagen las letras $a_0^L, a_1^L, a_2^L, \dots, a_{N_L}^L$ representan a estos datos.

Para el ejemplo que hemos estado mencionando, en donde se emplea una red neuronal para el reconocimiento de números escritos a mano, se tendrá una red Perceptrón Multicapa con la siguiente estructura (ver figura 3.8).

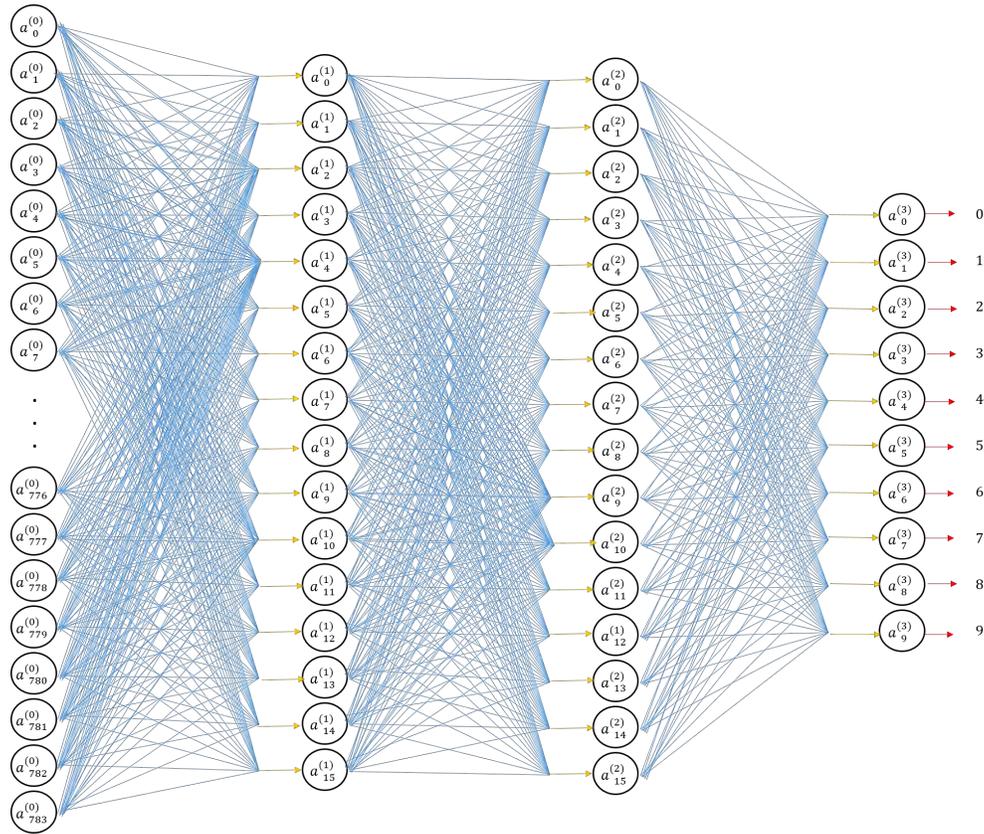


Figura 3.8: Estructura de una red Perceptrón multicapa empleada para el reconocimiento de números escritos a mano. Imagen basada en [41]

- Capa de Entrada (a_k^0):** Como dato de entrada, se usará la imagen de un número entre 0 y 9 escrito a mano. Este número estará escrito en una imagen de 28×28 píxeles (un total de 784 píxeles). Cada neurona en la capa de entrada tendrá un número entre 0 y 1, y este representará el valor de escala de grises que alberga cada píxel. Por ejemplo, si un píxel tiene un valor muy cercano a uno, indicará que el color de ese píxel es cercano al negro, si el píxel tiene un valor muy cercano a cero, indicará que el color de ese píxel es cercano al blanco. Se tendrán un total de 784 entradas ($N_0 = 783$), que representan a cada uno de los píxeles que conforman a la imagen del número escrito a mano.

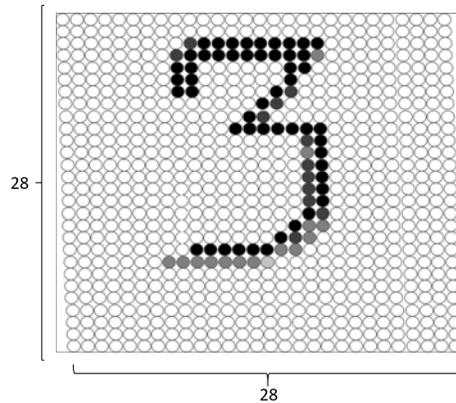


Figura 3.9: Ejemplo de una imagen de 28×28 píxeles de un número escrito a mano. Imagen basada en [41].

- **Capa de Salida** ($a_k^{L=3}$): La capa de salida para este ejemplo solo estará conformada por diez neuronas ($N_{L=3} = 9$), cada neurona representa a uno de los números que se esta intentando reconocer (en nuestro caso son 0,1,2,3,4,5,6,7,8 y 9). Al igual que en la capa de entrada, cada neurona de la capa de salida solo tomará un valor entre 0 y 1, este valor representará que tan segura está la red de que la imagen de entrada corresponda a cierto número. Por ejemplo, si la neurona en la capa de salida que corresponde al nueve tiene un valor cercano a uno, y las otras neuronas de la capa de salida tienen valores cercanos a cero, indicará que red identificó a la imagen de entrada como un nueve.
- **Capas Ocultas** (a_k^1) y (a_k^2): Para este ejemplo se utilizaron 2 capas ocultas, cada una con 16 neuronas $N_1 = N_2 = 15$. Realmente no hay una razón de peso para la elección de esta estructura, en donde se consultó este ejemplo [41], solo se menciona que se obtienen buenos resultados con este diseño.

Una vez que conocemos la estructura de esta red neuronal, procederemos a mencionar la idea en general del mecanismo que utiliza esta red para funcionar.

3.4.2. Funcionamiento

Para describir el funcionamiento de esta red neuronal, se usará el ejemplo de reconocimiento de números escritos a mano que se ha estado usando en este capítulo, y se tomará como base la referencia [41]. Recordemos que nuestro objetivo final consiste en introducir a la red neuronal una imagen de un número escrito a mano y que la red pueda identificarlo, es decir, que en la capa de salida el valor de activación sea igual a 1 para la neurona que corresponde al número de la imagen que se introdujo, siempre y cuando el valor de activación sea cercano a cero para las otras neuronas que se encuentra en esa misma capa. Para cumplir con esta tarea, se debe tomar en cuenta que los datos de entrada van a influir en el valor de activación de las neuronas que se encuentren en la siguiente capa (primera capa oculta a_k^1), y a su vez estas neurona van a influir en el valor de activación de las neuronas en la siguiente capa (a_k^2). Por último, el valor de activación de la ultima capa (que es el que determina el reconocimiento de la imagen), dependerá del valor de activación de la capa previa (a_k^2), cuyo valor de activación se vio afectado por las capas anteriores. Para entender esta última parte, podemos observar la figura 3.10.

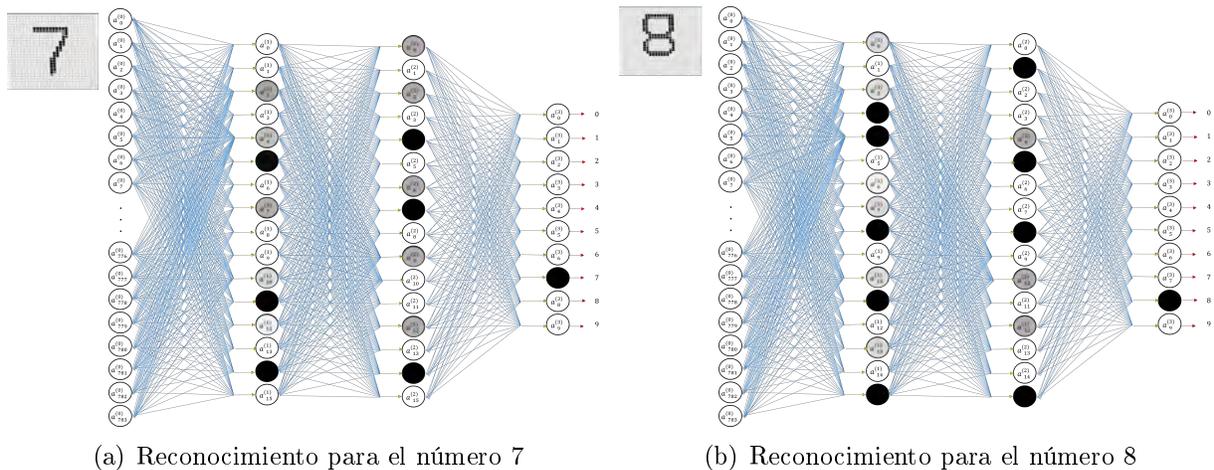


Figura 3.10: Red neuronal para el reconocimiento del número 7 y 8. Imagen basada en [41].

Esta figura muestra a la red neuronal con la que hemos estado trabajando, el inciso a) usa la imagen de un 7 como dato de entrada, y el inciso b) usa la imagen de un 8. Como se puede observar, la red ya ha sido entrenada y clasifica a ambas imágenes de forma correcta. Pero se puede ver que la activación de las neuronas en las dos capas ocultas es diferente para ambos casos. Esto se debe a que los datos de entrada son diferentes en ambos ejemplos, y el valor de activación de las neuronas en una capa previa influye en el valor de activación de las neuronas en capas posteriores [41]. Es importante mencionar que el valor de los pesos y los *bias* también influyen en el valor de activación de cada neurona. Para ser más específico, se calculará el valor de activación para una neurona en la primera capa oculta de este ejemplo. Este esquema se puede observar en el inciso a) de la figura 3.11.

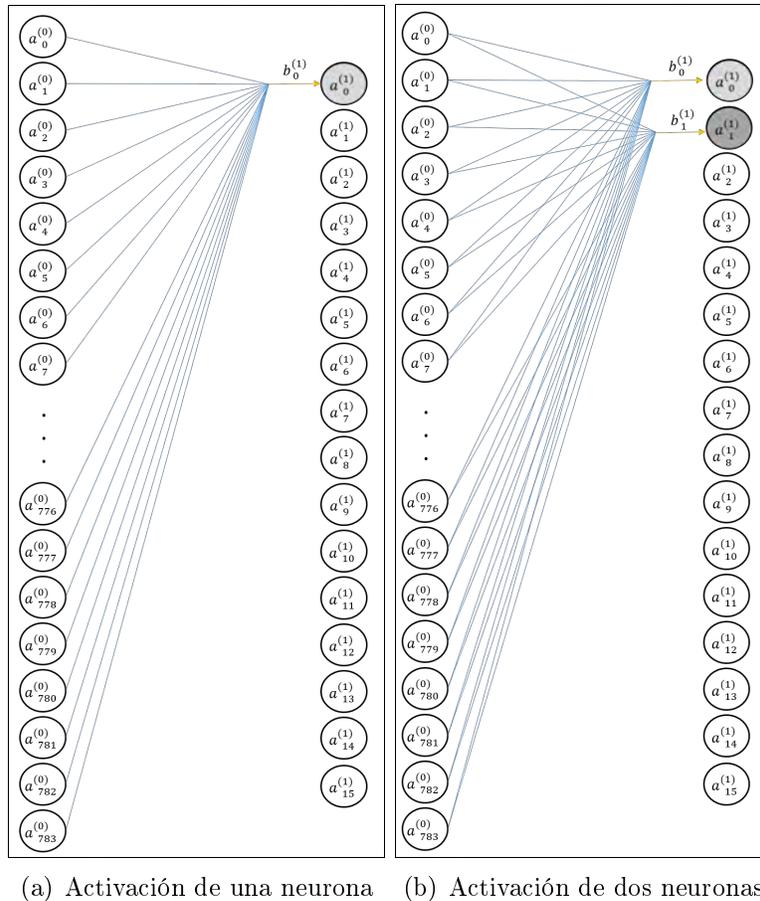


Figura 3.11: Representación de la activación de una y dos neuronas de la primera capa oculta. Imagen basada en [41].

En este caso la capa de entrada contará con 784 neuronas, y por ende 784 pesos. Cada uno de estos pesos se representan por medio de líneas azules (ver inciso a) de la imagen anterior). Como solo calculamos la activación de una neurona, solo se tendrá un valor de *bias*. Para realizar este cálculo, primero realizamos la suma que habíamos estado manejando en el modelo Perceptrón, en donde sumamos la multiplicación del valor de activación con su correspondiente peso y al final le sumamos el valor de *bias*.

$$\left(\sum_{i=0}^{783} a_i^{(0)} w_{0,i}^{(0)} \right) + b_0 = (a_0^{(0)} w_{0,0}^{(0)} + a_1^{(0)} w_{0,1}^{(0)} + \dots + a_{783}^{(0)} w_{0,783}^{(0)}) + b_0.$$

En donde:

- $a_i^{(0)}$: Datos de entrada con $0 \leq i \leq 783$.
- $w_{0,i}^{(0)}$: Pesos correspondientes a la primera neurona de la segunda capa.

Por último, para calcular la activación de la primera neurona de la segunda capa ($a_0^{(1)}$) aplicamos la función sigmoide a la suma anterior.

$$a_0^{(1)} = \sigma((a_0^{(0)}w_{0,0}^{(0)} + x_1^{(0)}w_{0,1}^{(0)} + \dots + a_{783}^{(0)}w_{0,783}^{(0)}) + b_0).$$

Para calcular la activación de la siguiente neurona de esa misma capa, y cuyo diagrama se encuentra en el inciso b) de la figura 3.11, es necesario considerar otros 784 pesos, que no necesariamente son iguales a los primero pesos que se utilizaron para calcular la activación de la neurona anterior ($a_0^{(1)}$). Por este motivo, se usa una notación de dos índices en los pesos ($w_{j,i}$) con $0 \leq j \leq 16$, para indicar que se utilizan diferentes valores para calcular la activación de cada una de las neuronas de la segunda capa. Considerando también otro valor de *bias*, la activación de la segunda neurona será:

$$a_1^{(1)} = \sigma((a_0^{(0)}w_{1,0}^{(0)} + x_1^{(0)}w_{1,1}^{(0)} + \dots + a_{783}^{(0)}w_{1,783}^{(0)}) + b_1).$$

En general se puede calcular el valor de activación de todas las neuronas en la segunda capa, mediante la siguiente multiplicación de matrices y vectores columna:

$$\begin{bmatrix} a_0^{(1)} \\ a_1^{(1)} \\ \vdots \\ a_{783}^{(1)} \end{bmatrix} = \sigma \left(\begin{bmatrix} w_{0,0}^{(0)} & w_{0,1}^{(0)} & \dots & w_{0,783}^{(0)} \\ w_{1,0}^{(0)} & w_{1,1}^{(0)} & \dots & w_{1,783}^{(0)} \\ \vdots & \vdots & \ddots & \vdots \\ w_{16,0}^{(0)} & w_{16,1}^{(0)} & \dots & w_{16,783}^{(0)} \end{bmatrix} \begin{bmatrix} a_0^{(0)} \\ a_1^{(0)} \\ \vdots \\ a_{783}^{(0)} \end{bmatrix} + \begin{bmatrix} b_0^{(0)} \\ b_1^{(0)} \\ \vdots \\ b_{783}^{(0)} \end{bmatrix} \right).$$

En donde:

$$\sigma \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \sigma(x) \\ \sigma(y) \\ \sigma(z) \end{pmatrix}.$$

Si representamos a la matriz que contiene a todo los pesos para calcular la activación de la primera capa con la letra M^1 , y a los dos vectores columna como a^0 y b^0 para los valores de activación y para los *bias* respectivamente. La ecuación anterior se puede escribir como:

$$a^1 = \sigma(M^1 a^0 + b^1).$$

El valor de activación para las otras capas se puede representar mediante las siguientes dos ecuaciones:

$$\begin{aligned} a^2 &= \sigma(M^2 a^1 + b^2). \\ a^3 &= \sigma(M^3 a^2 + b^3). \end{aligned}$$

Es importante mencionar que el número de pesos (w) depende tanto de la capa previa como de la capa en donde se calculan las activaciones. Por ejemplo en la capa a^1 hay un total 12544 pesos, debido a que en en la primera capa hay 784 neuronas y en la segunda capa hay 16 ($784 \times 16 = 12544$). Todos estos valores están incluidos en la matriz M^1 que justamente tiene 784 columnas y 16 filas.

Teniendo en cuenta lo anterior, la matriz M^2 tendrá 16 columnas y 16 filas, y la matriz M^3 tendrá 16 columnas y 10 filas, cada una con 256 y 160 valores respectivamente. Por lo tanto, en toda la red habrá un total de 12960 pesos. Teniendo en cuenta que en la primera y en la segunda capa oculta hay 16 *bias*, y en la última capa hay 10 *bias*, entonces en toda la red habrá un total de 42 *bias*. Por lo tanto en toda la red neuronal habrá un total de 13002 parámetros (12960 pesos + 42 *bias*). Todos estos valores son representados de forma gráfica en la figura 3.8, los pesos con líneas azules y los *bias* con flechas amarillas. Recordando que una red neuronal es una función; para este ejemplo, en donde se intenta usar una red neuronal para el reconocimiento de números escritos a mano, estaremos hablando de una función con 13002 parámetros que nos arroja 10 cantidades, que son los valores de activación de las neuronas en la última capa. Esta función se puede escribir como:

$$f(a_0, \dots, a_{783}) = \begin{bmatrix} y_0 \\ \vdots \\ y_9 \end{bmatrix}.$$

Así que el objetivo principal de una red neuronal artificial consiste en encontrar los valores más adecuados, tanto de los pesos como de los *bias*, que generen un dato de salida apropiado. Estos valores se encontraran a través de los datos de entrenamiento que se van presentando a la red. Cuando se presenta la primera imagen de entrenamiento, la red no sabe que valores colocar, así que elige valores al azar tanto de los pesos como de los *bias*. Por obvias razones los datos de salida de la última capa no son nada buenos, y para ver que tan malo son estos resultados, nos fijamos en la diferencia al cuadrado entre los datos de salida y los datos que se esperaban. Un ejemplo de esta situación, se puede visualizar en la figura 3.12.

	Dato de salida	Valr que se espera	Diferencia al cuadrado
	$a_0^{(3)} = 0.95$	$y_0 = 0.0$	$(a_0^{(3)} - y_0)^2 = 0.9025$
	$a_1^{(3)} = 0.50$	$y_1 = 0.0$	$(a_1^{(3)} - y_1)^2 = 0.25$
	$a_2^{(3)} = 0.45$	$y_2 = 0.0$	$(a_2^{(3)} - y_2)^2 = 0.2025$
	$a_3^{(3)} = 0.02$	$y_3 = 1.0$	$(a_3^{(3)} - y_3)^2 = 0.9604$
	$a_4^{(3)} = 0.54$	$y_4 = 0.0$	$(a_4^{(3)} - y_4)^2 = 0.2916$
	$a_5^{(3)} = 0.05$	$y_5 = 0.0$	$(a_5^{(3)} - y_5)^2 = 0.0025$
	$a_6^{(3)} = 0.65$	$y_6 = 0.0$	$(a_6^{(3)} - y_6)^2 = 0.4225$
	$a_7^{(3)} = 0.26$	$y_7 = 0.0$	$(a_7^{(3)} - y_7)^2 = 0.0676$
	$a_8^{(3)} = 0.56$	$y_8 = 0.0$	$(a_8^{(3)} - y_8)^2 = 0.3136$
	$a_9^{(3)} = 0.87$	$y_9 = 0.0$	$(a_9^{(3)} - y_9)^2 = 0.7569$
			Suma = 4.1701

Figura 3.12: Ejemplo de la diferencia al cuadrado entre los datos de salida de la red y el valor que se esperaba para la imagen de un tres. Imagen basada en [42].

En el ejemplo mostrado en la figura anterior, se utiliza la imagen de un tres como primer dato de entrenamiento, se calcula la diferencia al cuadrado que se había comentado anteriormente, y después se suman todas estas diferencias. Al resultado de esta suma se le conoce como función de costo de una sola imagen y se denota como C_0 [42].

$$C_0 = \sum_{i=0}^9 (a_i^{(3)} - y_i)^2 = 4.1701.$$

Es importante destacar que la función anterior solo toma en cuenta a la primera imagen que se uso para el entrenamiento de la red, pero supongamos que tenemos 1000 imágenes para realizar todo el entrenamiento. Entonces vamos a tener 1000 funciones de costo (C_0, C_1, \dots, C_{999}). Al promedio de todas estas funciones se le conoce como función de costo promedio (C) [42].

$$C = \frac{1}{n} \sum_{i=0}^{n-1} C_i = \frac{1}{1000} (C_0 + C_1 + \dots + C_{999}).$$

En nuestro ejemplo de reconocimiento de números escritos a mano, esta función tendrá 13002 datos de entrada y solamente arrojará un dato de salida, en donde los parámetros de esta función serán las 1000 imágenes que se están usando como datos de entrenamiento. Tomando en cuenta lo anterior, tanto los pesos como los *bias* se pueden visualizar como grados de libertad de la función de costo.

En la figura 3.12 se calculó la función de costo para el primer dato de entrenamiento de la red, en donde se obtuvo un valor relativamente grande de $C_0 = 4.1701$. Si el valor de la función de costo es cercano a cero, indicará que los datos de salida de la red cada vez se parecen más a los datos que se esperan, por el contrario, si el valor de la función de costo es bastante elevado, indicará que los datos de salida de la red no se parecen en nada a los datos que se esperan. Por lo tanto, para hacer que la red funcione de manera correcta, se tendrá que ir acercando a cero el valor de la función de costo a medida que el usuario vaya presentado imágenes de entrenamiento. En otras palabras se tendrá que minimizar esta función.

Para visualizar este procedimiento, supongamos que solo tenemos una variable en nuestra función de costo $C(w)$. La gráfica y los pasos a seguir para minimiza esta función, se ven representados en la figura 3.13.

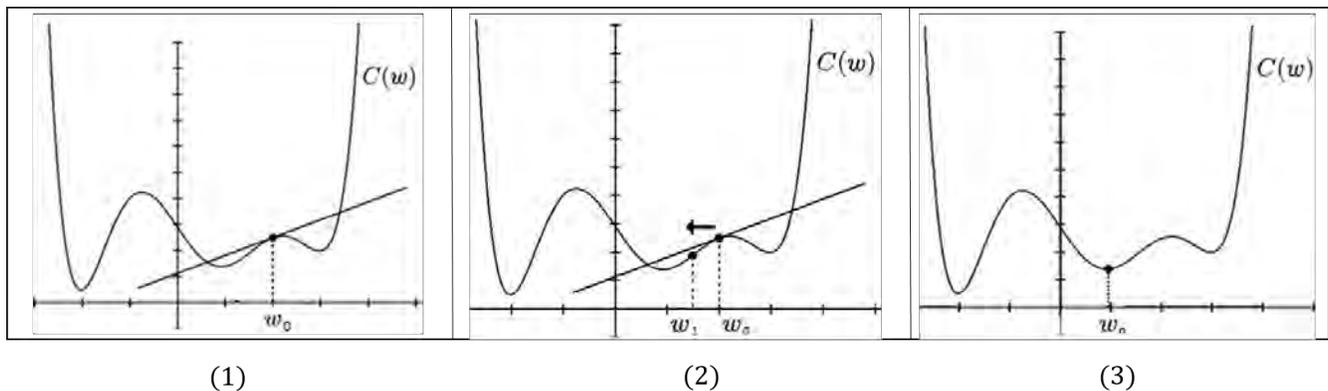


Figura 3.13: Pasos para minimizar la función de costo en dos dimension. Imagen modificada de [42].

Estos pasos son:

1. Se escoge un punto al azar (w_0) en el dominio de la función de costo, y nos fijamos en la pendiente de la recta tangente en ese punto, es decir, en la derivada de la función de costo.
2. Después se sumará el valor de esta derivada a nuestro punto inicial, que nos dará un punto más cercano al mínimo de la función.

$$w_1 = w_0 + \left. \frac{dC}{dw} \right|_{w_0}.$$

3. Se irá repitiendo el proceso descrito anteriormente: ir calculando y sumando el valor de la derivada a la posición w en donde se encuentre ese punto, hasta llegar al valor mínimo de la función.

Es importante mencionar que encontrar el mínimo total de toda la función es una tarea muy complicada, pero con este procedimiento se puede encontrar un mínimo local.

Considerando el valor negativo del gradiente de la función de costo ($-\nabla C$) se puede extender este procedimiento a más dimensiones, ya que esta expresión nos indica la dirección en donde la función disminuye con más rapidez. En el caso de nuestro ejemplo, en donde se trata de usar una red neuronal para el reconocimiento de números escritos a mano, tendremos una función de costo en 13002 dimensiones que no podemos graficar. Pero podemos considerar a las coordenadas del vector $-\nabla C$, como pequeños cambios que se deben hacer a los pesos y *bias* después de haber mostrado una imagen de entrenamiento. Supongamos que las 1000 imágenes de entrenamiento que se usan en este ejemplo, ya se han mostrado a la red neuronal (que en un inicio tomaba valores al azar de pesos y *bias*). Los cambios que se desean hacer para estos valores en cada imagen de entrenamiento se muestran en la figura 3.14.

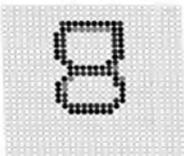
<i>Imágenes de entrenamiento</i>					
	(1)	(2)	(1000)	
<i>Pesos y bias</i>				<i>Promedio</i>
$w_0 = 0.12$	$dw_0^1 = -0.11$	$dw_0^2 = 0.23$		$dw_0^{1000} = 0.14$	$dw_0 = 0.78$
$w_1 = 0.01$	$dw_1^1 = 0.02$	$dw_1^2 = 0.56$		$dw_1^{1000} = 0.78$	$dw_1 = 0.045$
$w_2 = 0.41$	$dw_2^1 = 0.54$	$dw_2^2 = 0.78$	$dw_2^{1000} = 0.45$	$dw_2 = 0.78$
$w_3 = 0.54$	$dw_3^1 = -0.31$	$dw_3^2 = -0.09$		$dw_3^{1000} = 0.75$	$dw_3 = 0.87$
⋮	⋮	⋮	⋮	⋮	⋮
$w_{13001} = 0.74$	$dw_{13001}^1 = -0.32$	$dw_{13001}^2 = 0.78$	$dw_{13001}^{1000} = 0.78$	$dw_{13001} = 0.78$

Figura 3.14: Imagen que muestra todos los valores de pesos y *bias* (ambos denotados por la letra w), así como los cambios que se desean hacer a estos valores en cada imagen de entrenamiento. Imagen modificada de [43]

Todos los cambios que desea cada una de las imágenes de entrenamiento, son tomados en cuenta al realizar un promedio. Por ejemplo, el cambio promedio que se desea para la primera variable será:

$$dw_0 = \frac{1}{1000} (dw_0^1 + dw_0^2 + \dots + dw_0^{1000}).$$

Justamente el promedio de estos cambios deseado para las 13002 variables, representan a las coordenadas del vector $-\nabla C$ [43].

$$-\nabla C = \begin{bmatrix} dw_0 \\ dw_1 \\ dw_2 \\ \vdots \\ dw_{13001} \end{bmatrix}.$$

Podemos observar que para calcular las coordenadas del vector $-\nabla C$, es necesario tomar en cuenta a las 1000 imágenes de entrenamiento, pero este cálculo puede ser muy complicado y tardado, incluso para una computadora. Por este motivo, en vez de considerar a todos los datos de entrenamiento, simplemente se suele tomar una muestra al azar de estas imágenes, y calcular las coordenadas del vector $-\nabla C$ con esta información. A este procedimiento se le conoce como “descenso de gradiente estocástico”[43]. En tres dimensiones, este procedimiento se puede visualizar como una persona borracha que desciende una colina hasta llegar a la parte más baja (ver inciso a de la figura 3.15). Al procedimiento que toma en cuenta a las 1000 imágenes de números escritos a mano que son usadas para entrenar la red, simplemente se le conoce como “descenso del gradiente”[43], y se puede visualizar como una persona que va tomando el camino más corto para llegar a la parte baja de la colina (ver inciso b de figura 3.15).

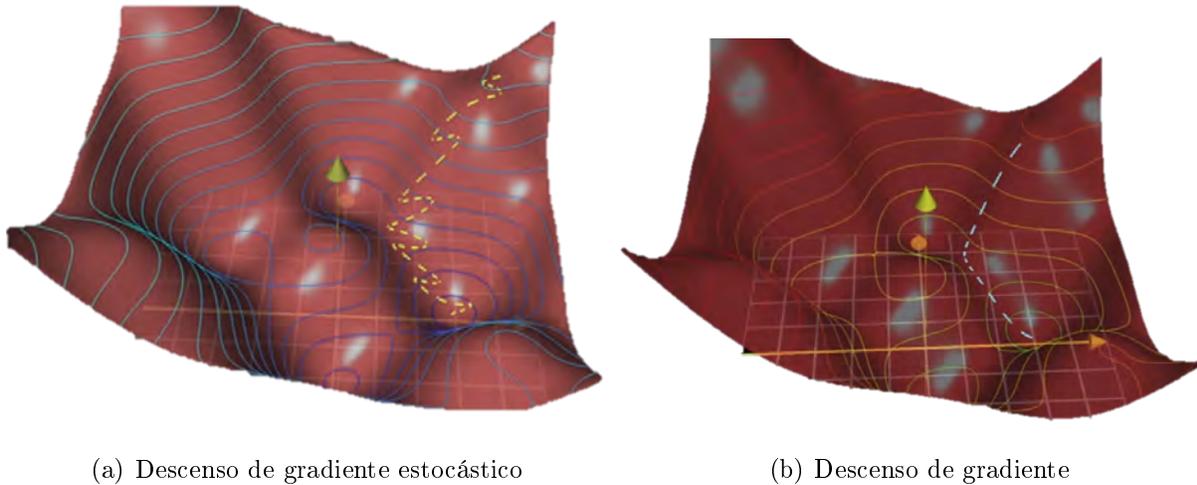


Figura 3.15: Representacion en 3 dimensiones en donde se minimiza la funcion de costo por medio de dos métodos. Imagen modificada de [43].

En resumen el funcionamiento de una red neuronal se puede visualizar mediante los siguientes pasos:

- Se toman al azar todos los valores de pesos y *bias* de la red neuronal, es decir, que elegimos un punto cualquiera w_0 del dominio de la función de costo, que nos servirá como punto de partida.

- Se lleva a cabo el procedimiento del descenso de gradiente estocástico para encontrar las coordenadas del vector $-\nabla C$ que toma en cuenta a una muestra de todos los datos de entrenamiento.

$$-\nabla C |_{w_0} = \begin{bmatrix} dw_0 \\ dw_1 \\ dw_2 \\ \vdots \\ dw_{13001} \end{bmatrix}.$$

- Se suman las coordenadas del vector $-\nabla C |_{w_0}$ a los valores de pesos y *bias*, para que la función de costo se vaya acercando a cero.

$$w_1 = w_0 + -\nabla C |_{w_0} = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \vdots \\ w_{13001} \end{bmatrix} + \begin{bmatrix} dw_0 \\ dw_1 \\ dw_2 \\ \vdots \\ dw_{13001} \end{bmatrix}.$$

- Se irá repitiendo el proceso descrito anteriormente, de ir calculando y sumando el valor de las coordenadas del vector $-\nabla C$ a la posición del punto w , hasta llegar al valor mínimo de la función.

Así que por medio de este ejemplo de reconocimiento de números escritos a mano, se pudo explicar el procedimiento que utiliza una red neuronal artificial para funcionar. Este procedimiento es el que utiliza el programa que se empleo para programa a la red neuronal presentada en este trabajo.

Para calcular las coordenadas de $-\nabla C$ en donde solo se toma en cuenta una imagen de entrenamiento, normalmente se utiliza un método llamado propagación hacia atrás (*backpropagation* en inglés). En el apéndice D se da una breve explicación de como funciona este procedimiento.

3.4.3. Un ejemplo en Root

Recordemos que la activación de cada una de las neuronas en la red se puede calcular a través de una multiplicación entre matrices y vectores columnas. Este tipo de operaciones son muy usuales en computación, y muchos programas incluyen diversos paquetes que facilitan el uso de estas operaciones. Por este motivo, no es difícil imaginarse que el algoritmo descrito anteriormente de una red neuronal artificial se pueda implementar a una computadora, y al igual que en el caso anterior, diversos programas de computo ya incluyen paquetes que modelan este tipo de procedimientos.

En este trabajo, se uso ROOT, un programa de computo instalado junto con AERIE en los paquetes *externals*, que sirve para el análisis, el almacenamiento y la visualización de datos en el experimento HAWC. En general este programa se basa principalmente en C++, pero integra algunas funciones de Python y R [46].

En ROOT se puede modelar una red neuronal Perceptrón multicapa, por medio de la clase “TMultiLayerPerceptron”. Esta clase permite al usuario elegir la estructura de red neuronal que más le convenga, al poder modificar las siguientes variables:

- El número de datos de entradas de la red.
- El número de capas ocultas de la red.
- El número de neuronas que hay en cada capa oculta.
- El número de datos de salida de la red.

En general, esta clase permite utilizar 6 métodos para entrenar a la red neuronal. Nosotros usaremos el método “kStochastic ” que se basa tanto en el algoritmo de descenso de gradiente estocástico como en el método de propagación hacia atrás que se describió anteriormente. En realidad, los otros 5 métodos de entrenamiento se basan en el procedimiento anterior, pero aquí no se entrará en detalle sobre la descripción y el funcionamiento de cada uno de estos procedimientos, ya que en este trabajo no se utilizaron. Si se desea profundizar en estos algoritmos se puede consultar la referencia [47].

Ahora veremos un ejemplo de una red neuronal artificial modelada en ROOT, para dar una mejor descripción de como se implementa este algoritmo en este lenguaje de programación.

Ejemplo: Se utilizó una red neuronal para aproximar la función $f(x, y) = x^2 + y^2$. En este caso tendremos dos datos de entradas, que son las variables (x, y) , y el dato de salida será el valor de la función $z = f(x, y)$. En este ejemplo se usará una sola capa oculta con 10 neuronas, debido a que hay un teorema que menciona que una sola capa oculta es suficiente para aproximar cualquier función; a este teorema se le conoce en el ámbito de redes neuronales como *universal approximation theorem* y se puede consultar en la referencia [35]. La estructura de esta red se puede visualizar en la figura 3.16.

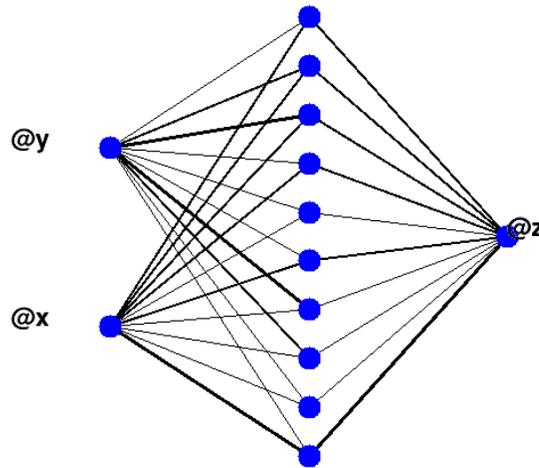


Figura 3.16: Estructura de red neuronal para este ejemplo. Este diagrama se realizó en ROOT usando la clase “TMultiLayerPerceptron”.

Los datos para esta red neuronal se generaron mediante las siguientes ecuaciones:

$$x = y = (j)(0.1).$$

$$z = x^2 + y^2.$$

En donde $j = \{0, 1, 2, \dots, 2000\}$. Los datos para entrenar la red neuronal, son aquellos en donde el valor de j es par. Mientras que los datos de prueba son aquellos en donde el valor de j es impar. Debemos aclarar que no hay ningún motivo en especial para tomar números pares como datos de entrenamiento y números impares para probar a la red neuronal. Esto solamente se hizo para que los datos de entrenamiento sean diferentes a los datos que se utiliza a la hora de verificar el funcionamiento de la red.

Esta información se guarda en dos archivos de texto *entrenamiento.dat* y *prueba.dat*. En la siguiente figura se muestra un fragmento de los datos que contienen estos archivos y la forma en la que se distribuyen en 3 columnas.

The figure shows two side-by-side screenshots of text editors. The left editor, titled '*entrenamiento.dat', displays a table with three columns labeled 'x', 'y', and 'z'. The data rows are: (0, 0, 0), (0.2, 0.2, 0.08), (0.4, 0.4, 0.32), (0.6, 0.6, 0.72), (0.8, 0.8, 1.28), (1, 1, 2), (1.2, 1.2, 2.88), (1.4, 1.4, 3.92), (1.6, 1.6, 5.12), (1.8, 1.8, 6.48), and (2, 2, 8). The status bar at the bottom indicates 'Ln 15, Col 8'. The right editor, titled 'prueba.dat (~red/tes', displays a table with three columns labeled 'x', 'y', and 'z'. The data rows are: (0.1, 0.1, 0.02), (0.3, 0.3, 0.18), (0.5, 0.5, 0.5), (0.7, 0.7, 0.98), (0.9, 0.9, 1.62), (1.1, 1.1, 2.42), (1.3, 1.3, 3.38), (1.5, 1.5, 4.5), (1.7, 1.7, 5.78), (1.9, 1.9, 7.22), and (2.1, 2.1, 8.82). The status bar at the bottom indicates 'Ln 2, Col 2'.

(a) Datos de entrenamiento (b) Datos de prueba

Figura 3.17: Fragmento de los datos que contienen los archivos *entrenamiento.dat* y *prueba.dat*.

En términos generales la estructura del programa es la siguiente:

- *Parte I:* Se define una clase que permite colocar los datos del archivo *entrenamiento.dat* en ramas de un árbol binario. En nuestro caso tendremos 3 ramas (x, y, z).
- *Parte II:* Se utiliza la clase “TMultiLayerPerceptron” que define la estructura de la red neuronal, además se especifican los datos que se usarán en el entrenamiento de la red y que son tomados de las ramas del árbol binario que se definió anteriormente. Recordemos que los valores de activación de las neuronas de entrada, deben estar entre cero y uno, por este motivo, esta clase va a normalizar a todos los datos que se utilicen en el entrenamiento. Estos datos serán normalizados por medio de su desviación estándar (en la siguiente pagina se da un ejemplo de esta normalización).
- *Parte III:* Se especifica el método que se usará para el entrenamiento, en nuestro caso se usará el procedimiento *kStochastic* mencionado anteriormente y descrito en la referencia [47]. Además se define el número de veces que la red neuronal modificara los valores de pesos y bias para alcanzar el punto mínimo de la función de costo.
- *Parte IV:* Por último se tomarán los datos del archivo *prueba.dat*, para verificar el funcionamiento de la red neuronal. En donde se usarán los valores de x y de y del archivo anterior para calcular el valor de la función ($f(x, y)_{NN}$) en base al entrenamiento previo.

El código completo que se utilizó para este ejemplo se basa en el programa mostrado en [45], y se puede consultar en el apéndice C.

El resultado de esta red neuronal para aproximar la función $f(x, y) = x^2 + y^2$ se muestra en la figura 3.18 (en donde los datos que se observan no están normalizados). Al inicio de esta gráfica los puntos no se ajustan de buena manera, esto se debe a un efecto de borde en donde la red neuronal no sabe que valores ajustar debido a que no tiene información para estos datos. Si no se toman en cuenta los primeros 10 puntos de esta gráfica, se obtiene un error porcentual promedio menor al 10%; que en términos generales es un buen ajuste para esta gráfica.

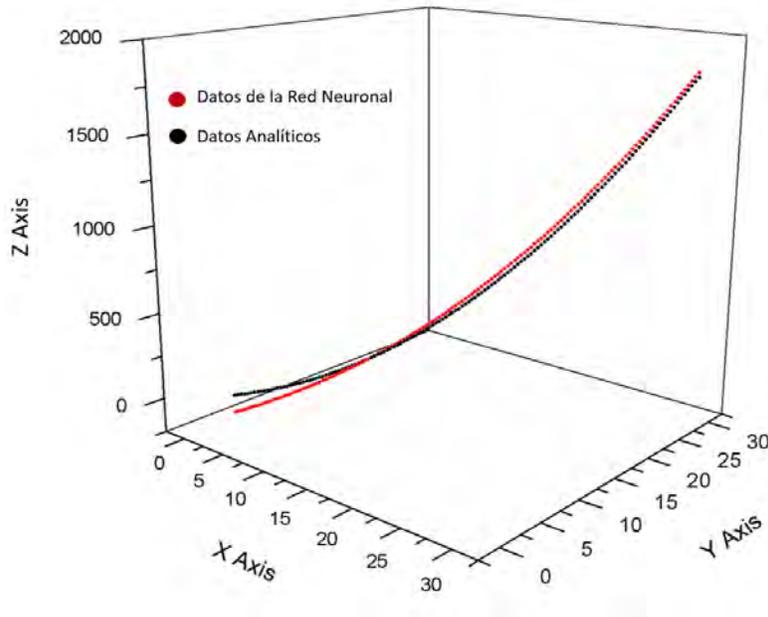


Figura 3.18: Comparación entre las funciones $f(x, y)$ y $f(x, y)_{NN}$.

Para normalizar los datos de prueba utilizando el valor de su desviación estándar, los pasos a seguir son los siguientes:

- Calculamos el promedio (z_p) de todos los datos utilizados.

$$z_p = \frac{z_1 + z_2 + \dots + z_{1000}}{1000} = \frac{0 + 0.8 + \dots + 79840.08}{1000} = 26626.68.$$

- Calculamos el valor de su desviación estándar (σ)

$$\sigma = \sqrt{\frac{(z_1 - z_p)^2 + (z_2 - z_p)^2 + \dots + (z_{1000} - z_p)^2}{1000}} = 23829.02.$$

- Así que el valor normalizado \hat{z} para cada dato de entrada se puede calcular como:

$$\hat{z}_i = \frac{(z_i - z_p)}{\sigma}$$

Al normalizar estos valores por medio de su desviación estándar, se tendrá que el promedio de estos datos normalizados será igual a cero, y el valor de su desviación estándar será igual a uno.

En resumen este fue un ejemplo de como se implementa el método de una red neuronal artificial en ROOT, y en términos generales la estructura del funcionamiento de este programa servirá de base para las pruebas realizadas en el capítulo 5.

Capítulo 4

Primer análisis: Volumen Efectivo de Detección

En este capítulo se estudió por medio de simulaciones Monte Carlo, el volumen efectivo de detección de los tanques del observatorio HAWC, en otras palabras, se buscó la región del tanque por donde tienen que pasar las partículas para que puedan ser detectadas. Para llevar a cabo esta tarea, se simularon muones horizontales y verticales, que pasaban a través de un tanque de este observatorio. Estas simulaciones tenían la finalidad de encontrar la región del tanque en donde es más probable que se activen 4, 3 ó 2 fotomultiplicadores. En particular se utilizaron muones como partícula de prueba, ya que en un futuro se espera utilizar este estudio para mejorar el algoritmo de reconstrucción de trazas de muones cuasi-horizontales, que son parte de las señales de ruido en la búsqueda de detección de neutrinos utilizando al observatorio HAWC. Por esta razón, y como parte de la motivación, en la primera sección de este capítulo, se mencionará de forma general, como se lleva a cabo esta reconstrucción. Posteriormente se hablará de la descripción y el método utilizado para encontrar esta región, para después mostrar los resultados y hacer la discusión de los mismos.

4.1. Motivación: Reconstrucción de trazas

La motivación de esta prueba, nace de la idea de mejorar el algoritmo de reconstrucción de trazas utilizando el método propuesto en la referencia [3]. Anteriormente se había comentado que este método consiste en buscar señales de leptones cargados o el producto de su descomposición en los tanques de este observatorio. Estos leptones son producidos mediante la interacción entre neutrinos y nucleones que se encuentran en el volcán Pico de Orizaba. Pero antes de estudiar estas señales es importante caracterizar el ruido provocado por muones generados en cascadas atmosféricas que vienen en la misma dirección. Como estas señales tendrían que pasar por la montaña, se espera que su dirección sea cuasi-horizontal. Tomando en cuenta que el sistema de referencia en HAWC ubica la posición de cada uno de sus detectores, cuando una de estas señales genere una traza de muones, es decir, que se activen una serie de tanques vecinos consistente con la velocidad de la luz, será posible reconstruir su dirección. Para observar como se realiza esta reconstrucción, veamos un ejemplo que se basa en la referencia [48].

Ejemplo: Supongamos que tenemos una traza de muones que activó 4 tanques del observatorio. Los tanques que se activaron son los siguientes: M13, L12, K12 y J11 (se puede ver la figura 2.11 para observar la ubicación de estos tanque en el arreglo de HAWC, y en la figura 4.1 se

representan a estos 4 tanques).

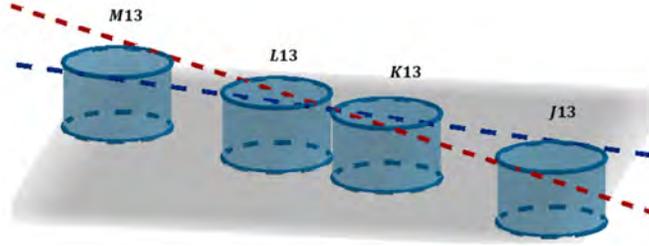


Figura 4.1: Representación gráfica de los tanques activados para este ejemplo. Las líneas de color rojo y azul, representan a muones que alcanzan a dejar una señal en los 4 tanques.

La dirección de esta señal, se ubicará en el sistema de referencia del observatorio por medio de las coordenadas (ϕ, θ) . Por lo tanto, esta reconstrucción se basa en encontrar a estos dos valores.

Ángulo azimutal (ϕ)

En la figura 4.2 se puede observar una representación de los tanques M13 y J11 en el plano (x, y) . Las coordenadas del centro para cada uno de estos tanques son las siguientes:

$$M13 = (113.393, 341.05) m.$$

$$J11 = (140.03, 314.9) m.$$

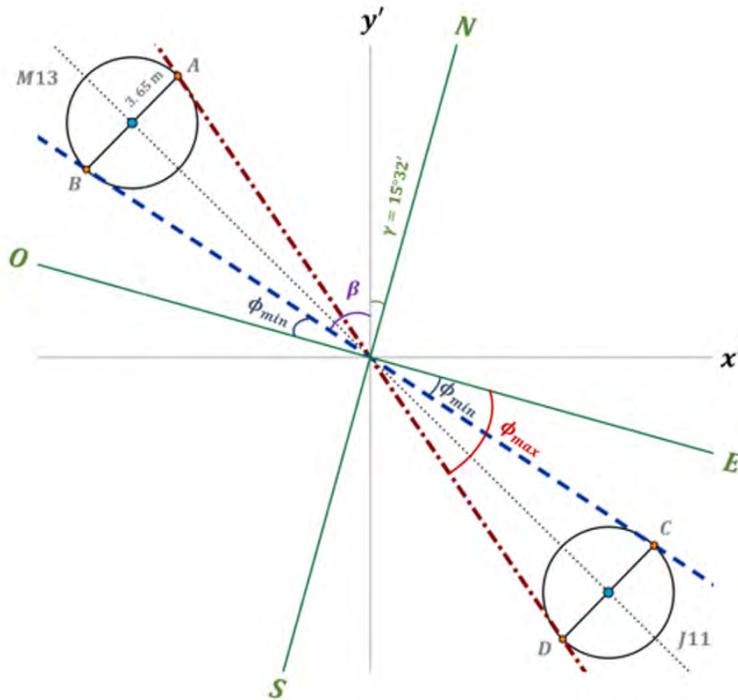


Figura 4.2: Representación para el ángulo azimutal, en donde se muestra a partir de donde se mide a los ángulos ϕ_{min} y ϕ_{max} . Imagen basada en [48].

Se dibujó una recta que pase por el centro de estos tanques, después se trazó una línea perpendicular a la recta anterior y que pasa por el centro de los tanques (se hicieron dos líneas

perpendiculares). Como resultado se van a generar 4 puntos (A, B, C y D), que son la intersección entre estas líneas perpendiculares y la circunferencia de cada tanque (ver figura 4.2). Para este ejemplo, las coordenadas de estos puntos son las siguientes:

m	A	B	C	D
x	115.95	110.84	142.59	137.47
y	343.65	338.45	317.5	312.3

La línea roja se trazó al unir el punto A con el punto D , y la línea azul se trazó al unir el punto B con el punto C . Para una partícula con esa trayectoria, estas líneas representan dos direcciones en donde apenas se alcanzó a dejar una señal. En otras palabras, la dirección de los muones que activaron los tanques estará delimitada por estas dos direcciones.

Como se puede observar en la figura 4.2, los ejes del sistema de referencia en HAWC se han trasladado (sistema (x', y')), ahora el origen se encuentra en la intersección de las líneas de color azul y rojo. Los ejes de color verde representan la orientación de los puntos cardinales, hay un ángulo de inclinación entre el eje y' y la orientación del norte geográfico de $\gamma = 15^{\circ}33'$. Tomando en cuenta lo anterior, se definen dos ángulos para la reconstrucción azimutal ϕ_{min} y ϕ_{max} , que se ven representados en la figura 4.2.

Cálculo del ángulo azimutal

Para encontrar el valor del ángulo ϕ_{min} los pasos son los siguientes:

- Tomando a la línea azul como hipotenusa, se puede definir el siguiente triángulo:

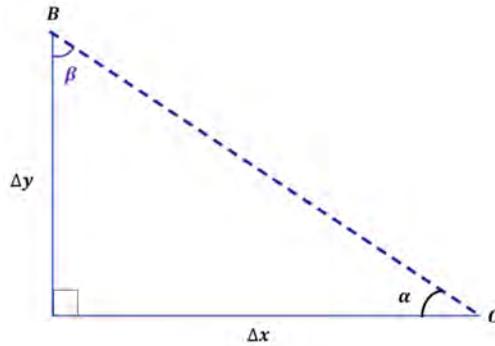


Figura 4.3: Representación gráfica del ángulo β y α .

Para este caso:

$$\Delta x = x_B - x_C.$$

$$\Delta y = y_B - y_C.$$

En donde

$$\alpha = \tan^{-1} \left(\frac{\Delta x}{\Delta y} \right).$$

$$\beta = 90^{\circ} - \alpha .$$

- En la figura 4.2 se muestra a los ángulos β y γ . A partir de ese diagrama, es fácil notar la siguiente relación:

$$\phi_{min} + \beta + \gamma = 90^{\circ}.$$

Por lo tanto:

$$\phi_{min} = 90^{\circ} - (\beta + \gamma).$$

Para encontrar el valor del ángulo ϕ_{max} , el procedimiento es prácticamente el mismo al que se mostró anteriormente. La única diferencia recae en la definición de los siguientes valores:

$$\Delta x = x_A - x_D.$$

$$\Delta y = y_A - y_D.$$

En donde se utiliza a los puntos A y D, para definir el triángulo anterior. En el caso particular de este ejemplo, se obtuvieron los siguientes resultados:

$$\phi_{min} = 50.83^{\circ}.$$

$$\phi_{max} = 60.02^{\circ}.$$

Por lo tanto, la dirección de la partícula que generó la señal, estará acotada por estos dos ángulos en la dirección azimutal.

Ángulo de elevación(θ)

En la figura 4.4 se puede observar una representación de los tanques M13 y J11 en el plano (x, z) . Las coordenadas del centro para cada uno de estos tanques son las siguientes:

$$M13 = (x, z) = (113.393, 0) \text{ m.}$$

$$J11 = (x, z) = (140.03, 0) \text{ m.}$$

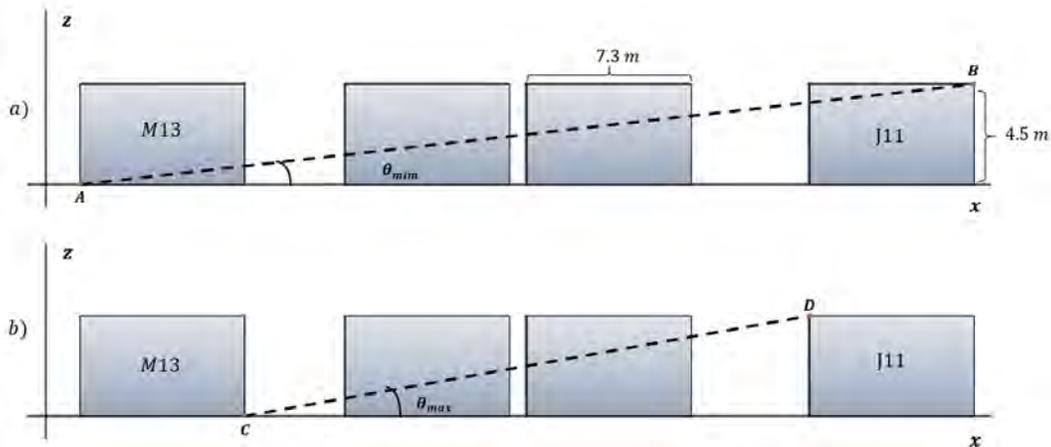


Figura 4.4: Representación para el ángulo de elevación, en donde se muestra a partir de donde se mide a los ángulos θ_{min} y θ_{max} . Imagen basada en [48].

Con las coordenadas del centro de cada tanque, y tomando en cuenta sus dimensiones (4.5 m de altura y 7.3 m de diámetro), se pueden encontrar las coordenadas de los puntos A,B,C y D. Estos puntos se muestran en la figura 4.4 y se encuentran en la esquina superior e inferior de cada tanque. Al igual que en el caso anterior, se formaron dos líneas, una al unir el punto A con el B y otra al unir el punto C con el D. Para el ángulo de elevación, estas serán las direcciones que delimitan el paso de la partícula para que alcance a dejar una señal en los tanques. Como en el caso anterior, se definen dos ángulos para esta reconstrucción θ_{min} y θ_{max} , que se ven representados en el inciso a) y en el inciso b) de la figura 4.4 respectivamente. Como se puede ver en esta imagen, tanto el ángulo θ_{min} como el ángulo θ_{max} , se definen a partir de un triángulo rectángulo, así que calcular sus valores no es muy difícil. En el caso de este ejemplo, se obtienen los siguientes resultados:

$$\begin{aligned}\theta_{min} &= 7.55^{\circ}. \\ \theta_{max} &= 13.10^{\circ}.\end{aligned}$$

Con estos valores, ya se tienen las coordenadas (θ, ϕ) para identificar la dirección de la partícula que genero la señal.

En el ejemplo anterior, a partir de una señal que activa varios tanques, se consiguen dos intervalos, uno para el ángulo azimutal y otro para el ángulo de elevación.

$$\begin{aligned}\phi &\in (\phi_{min}, \phi_{max}). \\ \theta &\in (\theta_{min}, \theta_{max}).\end{aligned}$$

En estos intervalos se encuentra la dirección de la partícula que genero la señal, y nos permite aproximar su posición. Para este ejemplo, el valor del ángulo máximo y el ángulo mínimo en ambas direcciones, se encuentran en función del volumen de agua en el tanque. Utilizando simulaciones Monte Carlo, el objetivo de esta prueba consiste en encontrar la región del tanque en donde es más probable que se activen 4, 3 ó 2 fotomultiplicadores (que en un principio sería una región menor al volumen total del tanque). Por lo tanto, al fijarnos en el número de detectores que se activan, se podrá conocer de forma más precisa la región del tanque por donde pasó la partícula, y con esa información se podría reconstruir mejor su trayectoria. En la siguiente sección se describe el procedimiento para encontrar estas regiones.

4.2. Descripción de las simulaciones

Como se ha estado mencionando, el objetivo de esta prueba consiste en encontrar la región del tanque en donde es más probable que se activen 4, 3 ó 2 detectores. Una primera aproximación para la forma de esta región, sería considerar a un cilindro más pequeño contenido en el volumen original del tanque. Así que en términos generales, el procedimiento de esta prueba consiste en simular muones horizontales (a diferentes alturas del tanque) para encontrar la altura de este cilindro, y simular muones verticales (a diferente distancia de separación del eje de simetría del tanque) para encontrar su radio. La figura 4.5 muestra un esquema a grandes rasgos de las simulaciones que se hicieron.

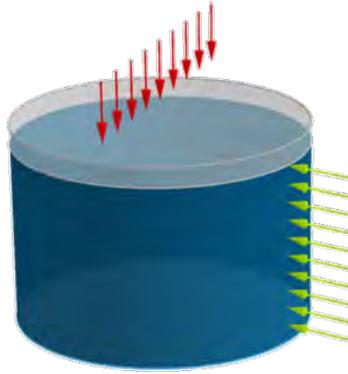


Figura 4.5: Imagen que representa las simulaciones de muones verticales y horizontales a través de un tanque del observatorio HAWC.

Estas simulaciones se realizaron en HAWCSIM, y se utilizó el tanque E10 como objetivo de prueba (en la figura 2.11 se muestra la ubicación del tanque E10 en el arreglo del observatorio). Debemos mencionar que no hubo un motivo en específico para la elección de este tanque, prácticamente cualquiera hubiera sido de utilidad. También es importante mencionar, que se utilizan muones en todas las simulaciones, ya que se quiere mejor la reconstrucción de trazas de estas partículas, que son señales de ruido para el método propuesto en la referencia [3] para la detección de neutrinos.

A continuación se darán todas las consideraciones que se tomaron para realizar las simulación de muones horizontales y verticales.

4.2.1. Simulación de muones horizontales

En esta parte se darán las especificaciones sobre la dirección y las diferentes alturas por donde se enviaron las partículas simuladas.

Dirección de muones horizontales a través del tanque

En la figura 4.6, se muestra una imagen en el plano (x, y) del tanque E10, en donde los 4 detectores son representados mediante puntos rojos. Debido a la ubicación de los fotomultiplicadores, es posible que la altura del cilindro no sea la misma en todas las direcciones. Por lo tanto, deberíamos enviar muones alrededor del tanque para observar los posibles cambios que se generan en la altura del cilindro. Las diferentes direcciones que se tomaron en cuenta para este trabajo, son representadas por líneas de diferente color en la figura 4.6, en donde se puede observar una separación angular de 30° entre cada trayectoria.

Tomando en cuenta la simetría en la posición de los fotomultiplicadores, algunas direcciones se repiten (estas direcciones se colocan del mismo color en la figura 4.6). Solamente se tendrán diferentes trayectoria en caminos que pasen por el centro del tanque y que se encuentren entre las líneas de color azul y verde (ver figura 4.6).

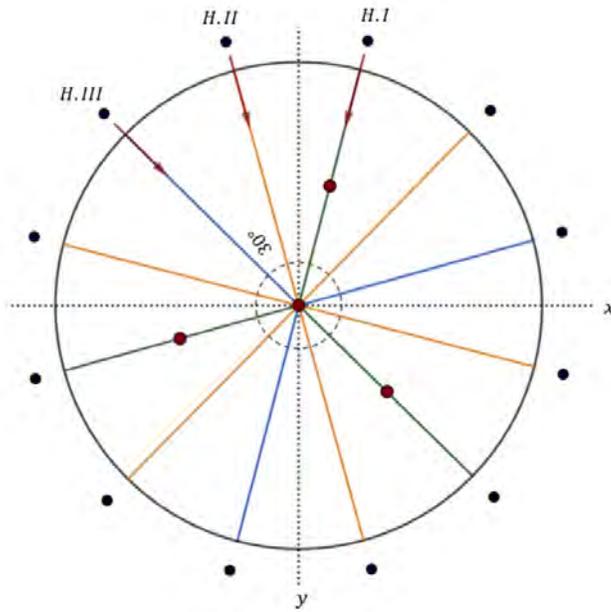


Figura 4.6: Dirección para la simulación de muones horizontales, en donde los puntos de color negro indican la posición inicial de estas partículas.

Así que para esta prueba, se harán simulaciones de muones horizontales a diferentes alturas del tanque en tres direcciones, a las que llamaremos $H.I$, $H.II$ y $H.III$ (ver figura 4.6).

Alturas

Las diferentes alturas por donde se simularon estos muones, se muestran de manera gráfica en la figura 4.7.

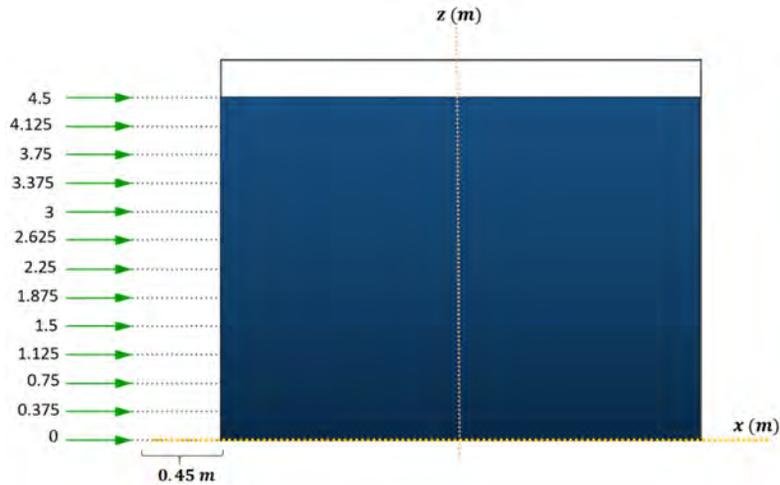


Figura 4.7: Alturas para la simulación de muones horizontales.

4.2.2. Simulación de muones verticales

En esta prueba, se enviaron muones de manera vertical, estas partículas se colocaron a diferente distancia de separación respecto al eje de simetría del tanque. A continuación, se describe la dirección y las diferentes distancias por donde se enviaron a estas partículas.

Dirección de muones verticales a través del tanque

Debido a la ubicación de los fotomultiplicadores, y al igual que en el caso anterior, se enviaron muones en diferentes direcciones alrededor del tanque, esto con la finalidad de estudiar los posibles cambios en el radio del cilindro. Las diferentes direcciones que se tomaron en cuenta para este trabajo, son representadas por líneas de diferente color en la figura 4.8, en donde se puede observar una separación angular de 15° entre cada trayectoria.

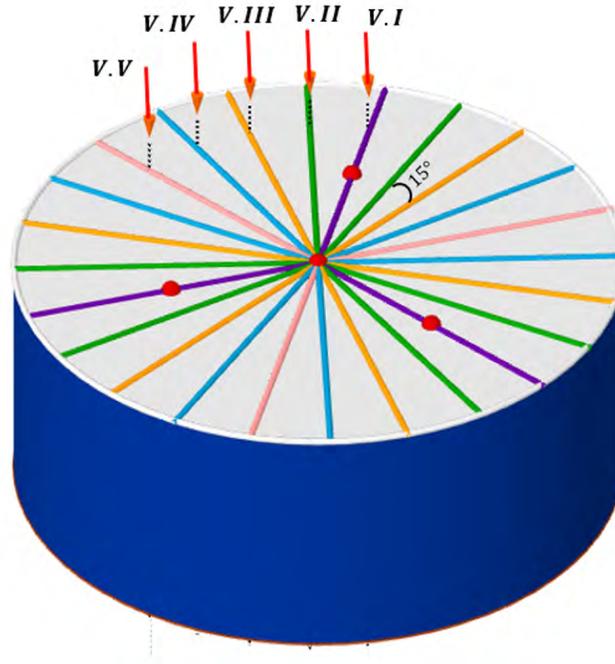


Figura 4.8: Dirección para la simulación de muones verticales, en donde las flechas de color rojo, indican la dirección de propagación de las partículas.

Tomando en cuenta la simetría en la posición de los fotomultiplicadores, algunas direcciones se repiten (estas direcciones se colocan del mismo color en la figura anterior). Solamente se tendrán diferentes trayectorias en caminos que pasen por el centro del tanque y que se encuentren entre las líneas de color morado y rosa.

Así que para esta prueba, se enviaron muones de manera vertical, estas partículas se colocaron a diferente distancia de separación respecto al eje de simetría del tanque, y se enviaron en cinco diferentes direcciones, a las que llamaremos *V.I*, *V.II*, *V.III*, *V.IV* y *V.V*.

Las diferentes separaciones por donde se enviaron a estas partículas, se pueden visualizar en la figura 4.9.

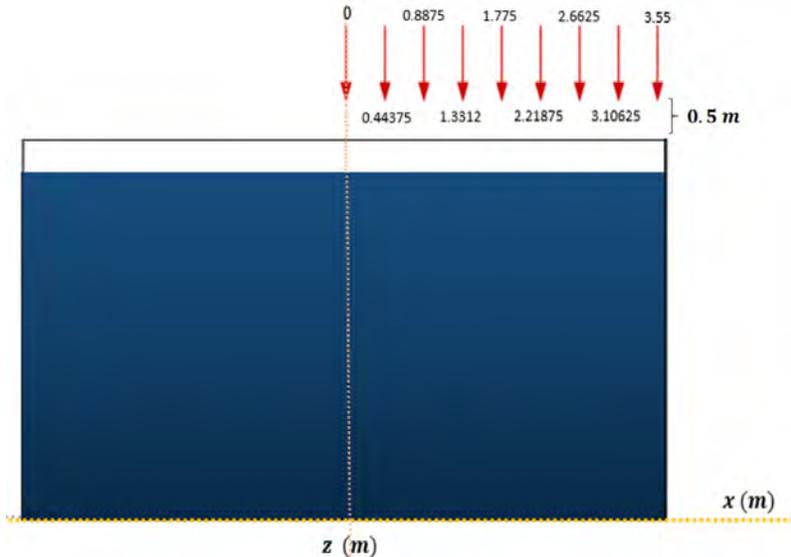


Figura 4.9: Separación respecto al eje de simetría para muones verticales

4.2.3. Número y energías

Para definir el volumen efectivo de detección, se utilizaron muones (μ^-) con dos diferentes energías: 1 TeV y 10 TeV. Se utilizan estas energías como parámetros, debido a que se esperan trazas de muones que estén dentro del rango de estos valores. Recordemos que este tipo de señales provienen de muones producidos en cascadas atmosféricas que atraviesan la montaña; por este motivo se estudian muones de alta energía. Además como se muestra en la gráfica de la figura 2.5; si se aumenta la energía en las simulaciones la cantidad de luz Cherenkov también aumenta. Esto se debe a que se producen más partículas secundarias. Por este motivo, se espera que la región correspondiente a una energía de 10 TeV sea más grande que la región para una energía de 1 TeV.

En cuanto al número de partículas a simular, debido a la desintegración aleatoria de los muones, se debe considerar un número elevado para estudiar su comportamiento. En esta prueba para cada simulación se utilizaron 10,000 muones. Esta cifra se utilizó a raíz de la cantidad de tiempo en que se tardaba cada simulación en terminar (en promedio se tardaban 7 horas para muones verticales y 1 día para muones horizontales⁴).

Tomando en cuenta en que se harían un total de 168 simulaciones aproximadamente, la cantidad de 10,000 partículas fue adecuada para realizar este trabajo. Además, es importante mencionar que la variable que se observa en cada simulación y que nos ayuda a definir el volumen efectivo de detección (ver sección 4.3), tiende a un valor asintótico a partir de los 10,000 muones, esto se muestra en el apéndice A, en donde se comparan los resultados de 4 simulaciones con diferente número de partículas simuladas.

4.2.4. Posición inicial

En las figuras 4.7 y 4.9, se puede observar la posición inicial de las partículas a simular (tanto para las simulaciones verticales como horizontales), en donde se muestra su ubicación en compa-

⁴Esto utilizando una computadora con 16 GB de RAM, un procesador AMD de 6 núcleos a 3.50 GHz, con un disco duro de 500 GB.

ración con el tanque.

Es importante mencionar que la distancia de vuelo promedio se puede calcular mediante la siguiente expresión:

$$d = \beta\gamma c\tau = v\gamma\tau = \tau c \sqrt{\frac{E^2}{m^2 c^4} - 1}. \quad (4.1)$$

En donde τ es la vida media de la partícula, β es la velocidad relativa a la de la luz v/c , γ es el factor de Lorentz, E es la energía del muón y m su masa.

Como se puede ver en esta ecuación, si la energía aumenta también la distancia de vuelo. Por ejemplo para el caso de muones (μ^-) con una energía de 100 GeV, se tendrá una distancia de vuelo promedio de 623.3 km. Considerando que el espacio para las simulaciones en HAWCSIM, abarca la región adentro de un cubo de 4 km de lado (todo proceso simulado afuera de esta región será descartado). La máxima distancia de separación para simular una partícula será de 4 km (de hecho debe ser menor debido a que los tanques del arreglo se encuentran cerca del origen en el sistema de referencia de HAWCSIM como se muestra en la figura 2.9). Tomando en cuenta que la distancia de vuelo promedio es mucho más grande que esta separación, y a que en las simulaciones no se toma en cuenta la interacción de las partículas con el aire durante su trayectoria, se pudo haber escogido cualquier distancia de separación con el tanque, y los resultados prácticamente serían los mismos. Por este motivo y como se muestra en las figuras 4.7 y 4.9, se escogió una distancia de separación respecto del tanque de 0.45 m y 0.5 m para muones horizontales y verticales respectivamente.

En resumen, para muones horizontales, se simulan 13 diferentes alturas en 3 posibles direcciones con 2 distintas energías, dando un total de 78 simulaciones. Para muones verticales, se simulan 9 diferentes distancias en 5 posibles dirección con 2 distintas energías, dando un total de 90 simulaciones. Obteniendo un total de 168 simulaciones consideradas en este trabajo.

Es importante mencionar que en algunas direcciones se realizaron simulaciones extras, para poder definir el volumen efectivo de detección.

4.3. Variables a observar en la simulación

En cada simulación, se almacena información de los detectores activados. Para definir el volumen efectivo de detección, nos fijamos en el número de fotomultiplicadores activados que registren una carga mayor o igual a 4 fotoelectrones (PE). Con esta condición se descartan las señales de ruido, que se caracterizan por dejar una carga menor a esta cantidad. Para ver de forma más clara la información que se obtiene en cada simulación, veamos el siguiente ejemplo.

Ejemplo: Se simulan 5 muones horizontales en la dirección $H.I$ a una altura de 2.25 m con una energía de 1 TeV. Los resultados que se obtienen se muestran en la tabla 4.1.

Como se observa en esta tabla, en todos los eventos se activan los 4 fotomultiplicadores, pero solamente en un caso no se cumple con la condición anterior. Tomando esto en cuenta, podemos definir en cada simulación los porcentajes en donde se activan 4, 3 ó 2 detectores y que además cumplen con esta condición. Para este ejemplo los porcentajes serán los siguientes:

$$\% \text{ en que se activan 4 detectores} = \frac{4}{5} \cdot 100 = 80 \%$$

Evento	Carga (PE)				Num. de detectores que cumplen la condición
	A	B	C	D	
1	2.5	4.5	5.2	4.3	3
2	6	6.7	7.6	4.7	4
3	7.2	7	6.2	10.1	4
4	8.9	8.5	5.9	9.6	4
5	9	10	5.8	6.9	4

Tabla 4.1: Datos de la simulación de ejemplo en donde se muestra la información observada para definir el volumen de detección.

$$\begin{aligned} \% \text{ en que se activan 3 detectores} &= \frac{1}{5} \cdot 100 = 20 \% \\ \% \text{ en que se activan 2 detectores} &= \frac{0}{5} \cdot 100 = 0 \% \end{aligned}$$

Por lo tanto, si un muón pasa por esa altura, el 80 % de las veces se activarán 4 detectores. En este trabajo para encontrar el volumen efectivo de detección se observaron estos porcentajes en todas las simulaciones; que en vez de 5 eventos tienen 10000 eventos. Por este motivo para hacer el cálculo de estos porcentajes se realizó un programa en ROOT, que pudiera manejar esta cantidad de datos en cada simulación.

4.4. Resultados

Antes de mostrar la tabla en donde se ubican las zonas en el tanque en donde es más probable que se activen entre 2 y 4 detectores. Primero mostraremos la información que se tomó en cuenta para llegar a estos resultados. Como punto de partida empezaremos por mostrar la forma en la que se define la zona en el tanque en donde es más probable que se activen 4 detectores, tomando en cuenta una sola dirección.

4.4.1. Altura y radio del cilindro en una sola dirección

En la tabla 4.2, se muestran los resultados para las simulaciones de muones horizontales enviados en la dirección H.II con una energía de 1 TeV. En la segunda columna de esta tabla, se muestran los porcentajes del número de veces que se activaron 4 fotomultiplicadores para cada altura simulada; cada valor de esta columna representa que tan probable es que se activen 4 detectores a una misma altura (a este porcentaje lo denotaremos como P_h). Con color amarillo se resalta a los resultados del porcentaje P_h mayores al 50 %. Por lo tanto, si un muón pasa a una altura mayor a 1.5 m y menor 1.875 m, al menos el 70.85 % de las veces se activaran cuatro detectores (ver tabla 4.2).

Entre cada altura que se tomó en cuenta para esta prueba hay cierta distancia de separación. En un principio se eligió una separación uniforme de 37.5 cm entre cada simulación, pero en algunas ocasiones se tomó una distancia más pequeña, esto con la finalidad de encontrar la altura en el tanque en donde se registraba un resultado mayor al 50 %. Para hacer notar esta proporción, a cada simulación se le asignó un valor o cierto peso, al que denotamos por w_h . Esta cantidad será igual a uno, si las separaciones entre dos simulaciones es igual 37.5 cm. Si esta separación es menor, w_h tomará un valor proporcional a que tan pequeña es esta separación en comparación a 37.5 cm.

Por ejemplo, entre 103.125 cm y 112.5 cm hay una separación de 9.375 cm, que representa una cuarta parte del valor de 37.5 cm, por lo tanto $w_h = 0.25$. Los números mostrados en la tercera columna de la tabla 4.2 representan a estos valores.

1 TeV					
Altura (cm)	Dirección H.II	Peso para la altura (w_h)	Porcentaje con peso	Porcentaje para esa altura	Porcentaje por región (P_R)
	4 PMTs %				
0	4.24	1	4.24	1.63	19.37
37.5	0.51	1	0.51	0.20	
75	3.30	1	3.30	1.27	
93.75	10.10	0.5	5.05	1.94	
103.125	19.39	0.25	4.85	1.87	
112.5	32.37	1	32.37	12.46	57.25
150	70.85	1	70.85	27.28	
187.5	77.81	1	77.81	29.96	
225	16.33	1	16.33	6.29	23.38
243.75	14.17	0.5	7.09	2.73	
262.5	11.59	1	11.59	4.46	
300	8.53	1	8.53	3.29	
337.5	4.73	1	4.73	1.82	
356.25	4.91	0.5	2.45	0.94	
375	4.50	1	4.50	1.73	
412.5	3.53	1	3.53	1.36	
450	1.97	1	1.97	0.76	
		Suma	259.69	100	100

Tabla 4.2: Porcentaje de 4 PMTs activados para muones en la dirección H.II con 1 TeV de energía.

La cuarta columna de esta tabla, muestra los valores que resultan de multiplicar los porcentajes de la segunda columna con los valores w_h de la tercera columna. Con esta información podemos encontrar a que altura del tanque es más probable que se activen 4 fotomultiplicadores. Para ello se suman todos los valores de la cuarta columna, y se calcula el porcentaje que aportó cada valor a este suma. En la quinta columna se muestran estos resultados. Por último, en la sexta columna se suman los porcentajes obtenidos en la quinta columna para ciertas regiones del tanque. Con esta información es posible determinar que tan probable es que se activen 4 fotomultiplicadores en diferentes alturas del tanque (a este valor lo llamaremos porcentaje por región y lo denotamos como P_R).

Para definir la altura del cilindro necesitamos una punto inicial (h_i) y un punto final (h_f), para este ejemplo:

$$h_i = 150 \text{ cm.}$$

$$h_f = 187.5 \text{ cm.}$$

Los porcentajes (P_h) que se registran en estas dos posiciones serán:

$$P_{h_i} = 70.85 \%$$

$$P_{h_f} = 77.81 \%$$

Por lo tanto, entre estas dos alturas (si se activan 4 detectores) habrá un porcentaje (P_R) del 57.25 % de que la partícula haya pasado por esta zona (ver tabla 4.2), y hay una probabilidad de al menos el 70.85 % de que se activen 4 fotomultiplicadores cuando el muón pase por esta región. Al combinar esta información, se obtiene de forma general la probabilidad de que el muón active 4 detectores entre dos alturas distintas del tanque, a esta probabilidad la denotaremos como (P_H) y se puede expresar mediante la siguiente expresión:

$$P_H = \min \{ \min \{ P_{hi}, P_{hf} \}, P_R \}. \quad (4.2)$$

Para el ejemplo anterior, este porcentaje será:

$$P_H^{H.II} = \min \{ \min \{ 70.85, 77.81 \}, 57.25 \} = 57.25 \%.$$

Por otro lado, la probabilidad de que se activen 4 detectores entre dos separaciones distintas se podrá expresar como:

$$P_V = \min \{ \min \{ P_{ri}, P_{rf} \}, P_R \} \quad (4.3)$$

En donde:

- P_{ri} = Es el porcentaje correspondiente a la posición inicial r_i .
- P_{rf} = Es el porcentaje correspondiente a la posición final r_f .

Separación (cm)	P_r
	3 PMTs
177.50	23.46 %
221.88	62.6 %
232.97	77.22 %
255.16	72.8 %
266.25	68.43 %
310.63	3.12 %

Tabla 4.3: Porcentaje de 3 PMTs activados para muones en la dirección V.II con 1 TeV de energía.

Es importante mencionar que para elegir al punto inicial y al punto final que definen al radio del cilindro, nos fijamos en la región del tanque en donde se obtenga un mayor porcentaje P_V ; para ver esto de una forma más clara observemos la tabla 4.3.

En esta tabla, se muestran los resultados para el porcentaje P_r de muones verticales de 1 TeV que se enviaron en la dirección V.II. En la región que va de 232.97 cm a 255.16 cm se obtiene un porcentaje P_r mínimo de 72.8 % y un porcentaje por región (P_R) de 17.1 %. Por otro lado, en la región que va de 221.88 cm a 266.25 cm se obtiene un porcentaje P_r mínimo de 62.6 % y un porcentaje por región de P_R de 76.85 %. De acuerdo con la ecuación 4.3, en la primera y en la segunda región se obtiene un porcentaje P_V del 17.1 % y del 62.6 % respectivamente. Como se obtiene un resultado mayor para este porcentaje en la segunda región, el radio del cilindro quedará definido por el punto inicial y final que corresponden a esta región.

De forma similar, para definir la altura del cilindro nos fijamos en la región del tanque en donde se obtenga un mayor porcentaje P_H .

4.4.2. Altura del cilindro en todas las direcciones

Recordemos que el objetivo de esta prueba es encontrar las siguientes tres regiones:

- Región más probable del tanque para que un muón active exactamente 4 detectores.
- Región más probable del tanque para que un muón active exactamente 3 detectores.
- Región más probable del tanque para que un muón active exactamente 2 detectores.

La forma de estas regiones se modelan en este trabajo como cilindros inscritos en el volumen total del tanque. En la tabla 4.4 se muestra la posición inicial y final que nos sirven para definir la altura de estas regiones. Como se puede observar en esta tabla, se muestran los resultados para todas las direcciones y energías que se tomaron en cuenta en este trabajo, así como su correspondiente porcentaje P_H .

Muones Horizontales						
	1 TeV			10 TeV		
	Dirección H.I			Dirección H.I		
	4 PMTs	3 PMTs	2 PMTs	4 PMTs	3 PMTs	2 PMTs
h_i (cm)	150	225	37.5	112.5	225	37.5
h_f (cm)	187.5	337.5	75	187.5	337.5	75
P_H %	59.16	67.09	74.64	51.15	50.19	53.99
Dirección H.II			Dirección H.II			
h_i (cm)	150	93.75	337.5	112.5	75	337.5
h_f (cm)	187.5	112.5	356.25	187.5	103.125	356.25
P_H %	57.25	67.09	74.64	47.67	50.19	53.99
Segunda Región			Segunda Región			
h_i (cm)		225			225	
h_f (cm)		262.5			300	
P_H %		67.09			50.19	
Dirección H.III			Dirección H.III			
h_i (cm)	150		37.5	112.5		37.5
h_f (cm)	187.5		75	187.5		75
P_H %	69.02		74.64	47.26		53.99
Segunda Región			Segunda Región			
h_i (cm)			262.5			262.5
h_f (cm)			337.5			337.5
P_H %			74.64			53.99

Tabla 4.4: Posición inicial y final para definir la altura del cilindro.

En esta tabla hay algunos espacios en blanco, esto se debe a que no en todas las direcciones se puede definir una región en donde es mas probable que se activen de 2 a 4 detectores. También en las direcciones H.II y H.III se define un espacio para una segunda región, debido a que en esas trayectorias hay dos zonas diferentes en el tanque en donde se puede establecer una misma región de probabilidad en la que se activen cierto número de fotomultiplicadores.

Con la posición inicial y final que mostramos en esta tabla, procedemos a definir la ubicación para la altura del cilindro que buscamos, y que toma en cuenta a cada una de las tres direcciones que consideramos. Como se mencionó anteriormente, no en todas las direcciones se encuentra una altura inicial y final para definir la región del tanque en donde es más probable que se activen 4, 3 ó 2 detectores. Incluso cuando se puede definir una de estas regiones en diferentes direcciones, no siempre se tienen valores iguales en la posición inicial y final. Al tomar en cuenta estas consideraciones, la forma en la que se definen la altura inicial y final para determinar la altura del cilindro que toma en cuenta a todas la direcciones que se usaron, se puede dividir en los siguientes casos:

- **Caso 1:** En todas las direcciones se encuentra una misma región.

En este caso, se debe tomar en cuenta a las 3 posibles direcciones por donde pudo haber llegado el muón. Estas direcciones se muestran en la figura 4.6. Cada dirección se representa en esta imagen por líneas de diferente color (un total de 12 de líneas). Como cada una de las tres direcciones se representa con un color diferente, se podrá definir la probabilidad de que el muón llegue en cierta dirección. Esto se hará al dividir el número de líneas del mismo color entre el número total de líneas.

$$\begin{aligned} P_{H.I} &= \frac{3}{12} = 0.25. \\ P_{H.II} &= \frac{6}{12} = 0.5. \\ P_{H.III} &= \frac{3}{12} = 0.25. \end{aligned}$$

Para este caso la altura inicial y final que considera a todas las direcciones, se podrán definir de la siguiente manera:

$$\begin{aligned} h_i &= h_i^{H.I} P_{H.I} + h_i^{H.II} P_{H.II} + h_i^{H.III} P_{H.III}. \\ h_f &= h_f^{H.I} P_{H.I} + h_f^{H.II} P_{H.II} + h_f^{H.III} P_{H.III}. \end{aligned}$$

Ejemplo: Un ejemplo de este caso, se puede observar a la hora de definir la altura del cilindro en donde es más probable que se activen 4 detectores para el caso de 1 TeV. La altura inicial y final para las tres direcciones en esta región se muestran en la tabla 4.4. Así que la ubicación del cilindro que representa a esta región será la siguiente:

$$\begin{aligned} h_i &= (150 \text{ cm})(0.25) + (150 \text{ cm})(0.5) + (150 \text{ cm})(0.25) = 150 \text{ cm}. \\ h_f &= (187.5 \text{ cm})(0.25) + (187.5 \text{ cm})(0.5) + (187.5 \text{ cm})(0.25) = 187.5 \text{ cm}. \end{aligned}$$

- **Caso 2:** En dos de las tres direcciones se encuentra una misma región.

La región que se encuentra en ambas direcciones debe estar en la misma zona. Por ejemplo, tanto en la dirección $H.I$ como en la dirección $H.II$ es posible determinar la región más probable en el tanque en donde se activan 3 detectores. Pero en la dirección $H.III$ se puede determinar esta región, en dos zonas diferentes del tanque (ver tabla 4.4). Por ello se especifica que la región debe estar por la misma zona.

Al igual que en el caso anterior, se define la altura y la posición del cilindro a partir de una altura inicial y una altura final. Para este caso, las expresiones serán las siguientes:

$$\begin{aligned} h_i &= h_i^{D.1} P_{D.1} + h_i^{D.2} P_{D.2}. \\ h_f &= h_f^{D.1} P_{D.1} + h_f^{D.2} P_{D.2}. \end{aligned}$$

En donde:

- $D.1$ = Es una de las dos direcciones en donde se encontró la región, y puede ser cualquiera de las 3 direcciones anteriores (HI , HII y $HIII$)
- $D.2$ = Es la otra dirección en donde se encontró la región.

Recordemos que en la figura 4.6, se representa a cada una de las tres direcciones con una línea de un color diferente. Para este caso, la probabilidad de que el muón llegue en ciertas direcciones, se calculará al dividir el número de líneas del mismo color entre la suma de las líneas de diferente color que representan a las 2 direcciones en donde se encontró una misma región. Para ver esto más claro, observemos el siguiente ejemplo.

Ejemplo: Para definir la altura del cilindro en donde es más probable que se activen 3 detectores, se debe recurrir al caso anterior. Esta región se encuentra en dos direcciones diferentes ($H.I$ y $H.II$), a una altura igual o superior de 225 cm. Para este caso solo se toma en cuenta a las trayectorias que representan a estas dos direcciones (un total de 9 líneas). Tomando estos datos, se tendrá que:

$$\begin{aligned} &: P_{H.I} = \frac{3}{9} = \frac{1}{3}. \\ &P_{H.II} = \frac{6}{9} = \frac{2}{3}. \end{aligned}$$

Entonces:

$$\begin{aligned} h_i &= (225 \text{ cm}) \left(\frac{1}{3}\right) + (225 \text{ cm}) \left(\frac{2}{3}\right) = 225 \text{ cm}. \\ h_f &= (337.5 \text{ cm}) \left(\frac{1}{3}\right) + (262.5 \text{ cm}) \left(\frac{2}{3}\right) = 287.5 \text{ cm}. \end{aligned}$$

- **Caso 3:** Solamente en una dirección se encuentra una misma región en una misma zona del tanque.

La altura inicial y final será exactamente igual a la que se encontró en la dirección en la que se definió esta región.

$$\begin{aligned} h_i &= h_i^{D.1} \\ h_f &= h_f^{D.1} \end{aligned}$$

Ejemplo: Un ejemplo para este caso, sería considerar a la zona del tanque en donde es más probable que se activen 3 detectores a una altura menor o igual a 112.5 cm. Esa región solamente se encontró en la dirección $H.II$, por lo tanto:

$$\begin{aligned} h_i &= 97.72 \text{ cm} \\ h_f &= 112.5 \text{ cm} \end{aligned}$$

Que son los mismos valores que se registraron en esa dirección.

En general estos tres casos se pueden resumir en las siguientes ecuaciones:

$$h_i = \sum_{j=1}^n h_i^{Dj} P_{Dj} \quad (4.4)$$

$$h_f = \sum_{j=1}^n h_f^{Dj} P_{Dj} \quad (4.5)$$

En donde n es el número de direcciones que registraron una misma región en una misma zona del tanque, y la probabilidad en cada dirección (P_{Dj}) depende de este valor, tal y como se mostró en los casos anteriores. El valor Dj hace referencia a las direcciones en donde se encontró a las regiones.

Cada región definida por los puntos h_i y h_f , se le asigna un probabilidad P_H calculada con la ecuación 4.2. Para el caso en donde se toma en cuenta a todas las direcciones, el porcentaje P_H se define mediante la siguiente ecuación:

$$P_H = \sum_{j=1}^n P_H^{D.j} P_{Dj} \quad (4.6)$$

Para el ejemplo que vimos en el primer caso este valor será:

$$P_H = (63\%)(0.25) + (57.25\%)(0.5) + (63.94\%)(0.25) = 60.36\%$$

Para las regiones en donde es mas probable que se activen 3 ó 2 fotomultiplicadores. Se toman algunas consideraciones extras para definir este porcentaje y que se explican con mas detalle en el apéndice E.

4.4.3. Radio del cilindro en todas las direcciones

En la tabla 4.5 se muestra la posición inicial y final que nos sirven para definir el radio de estas regiones. Como se puede observar, se muestran los resultados para todas las direcciones y energías que se tomaron en cuenta en este trabajo, así como su correspondiente porcentaje P_V . También se puede ver que hay algunos espacios en blanco, esto se debe a que no en todas las direcciones se puede definir una región en donde es mas probable que se activen de 2 a 4 detectores.

Para definir el radio del cilindro que toma en cuenta a todas las direcciones, el procedimiento es prácticamente el mismo al considerado para definir la altura del cilindro en la sección anterior. La única diferencia es que en este caso se contemplan cinco direcciones en lugar de tres. Para este caso:

$$r_i = \sum_{j=1}^n r_i^{Dj} P_{Dj}. \quad (4.7)$$

$$r_f = \sum_{j=1}^n r_f^{Dj} P_{Dj}. \quad (4.8)$$

$$P_V = \sum_{j=1}^n P_V^{D.j} P_{Dj}. \quad (4.9)$$

Muones Verticales						
	1 TeV			10 TeV		
cm	Dirección V.I			Dirección V.I		
	4 PMTs	3 PMTs	2 PMTs	4 PMTs	3 PMTs	2 PMTs
r_i	0		266.25	0		266.25
r_f	177.5		355	221.875		355
$P_V\%$	74.15		93.86	64.97		83.24
	Dirección V.II			Dirección V.II		
r_i	0	221.875	310.625	0	221.875	310.625
r_f	177.5	266.25	355	177.5	266.25	355
$P_V\%$	74.63	62.6	77.79	79.78	55.26	80.05
	Dirección V.III			Dirección V.III		
r_i	0	221.875	332.8125	0	221.875	332.8125
r_f	177.5	266.25	355	177.5	266.25	355
$P_V\%$	73.13	62.84	62.63	78.2	60.96	64.89
	Dirección V.IV			Dirección V.IV		
r_i	0	221.875		0	221.875	
r_f	177.5	355		177.5	355	
$P_V\%$	71.58	71.5		76.11	73.95	
	Dirección V.V			Dirección V.V		
r_i	0	221.875		0	221.875	
r_f	177.5	355		177.5	355	
$P_V\%$	69.69	77.41		73.91	78.55	

Tabla 4.5: Posición inicial y final para definir el radio del cilindro

4.4.4. Volumen efectivo de detección

En esta sección se darán los resultados para el volumen efectivo de detección, en otras palabras, se mostrará la ubicación de los puntos iniciales y finales que definen las alturas y los radios para los cilindros que mencionamos anteriormente. También se expresa que tan probable es que el muón haya pasado por esa región si solamente nos fijamos en los detectores activados. A esta probabilidad la denotaremos como P_C y se define como:

$$P_C = \min \{P_H, P_V\}. \quad (4.10)$$

Los resultados para el volumen efectivo de detección se encuentran en la tabla 4.7.

Posición	1 TeV			10 TeV		
	Región del tanque más probable de que se activen:					
	4 PMTs	3 PMTs	2 PMTs	4 PMTs	3 PMTs	2 PMTs
r_i	0	221.875	310.625	0	221.875	310.625
r_f	177.5	304.286	355	183.05	304.286	355
P_V	72.82	67.33	74.94	75.08	65.56	74.62
Primera región						
h_i	150	225	37.5	112.5	225	37.5
h_f	187.5	287.5	75	187.5	312.33	75
P_H	60.67	67.09	74.64	48.44	50.19	53.99
Segunda región						
h_i		93.75	312.5		75	312.5
h_f		112.5	350		103.125	350
P_H		67.09	74.64		50.19	53.99

Tabla 4.6: Ubicación de la región en donde es más probable que se activen 4, 3 ó 2 PMTs

A partir de la información de la tabla anterior, en la figura 4.10 se pone de forma gráfica las zonas en el tanque en donde es más probable que se activen 4, 3 ó 2 fotomultiplicadores.

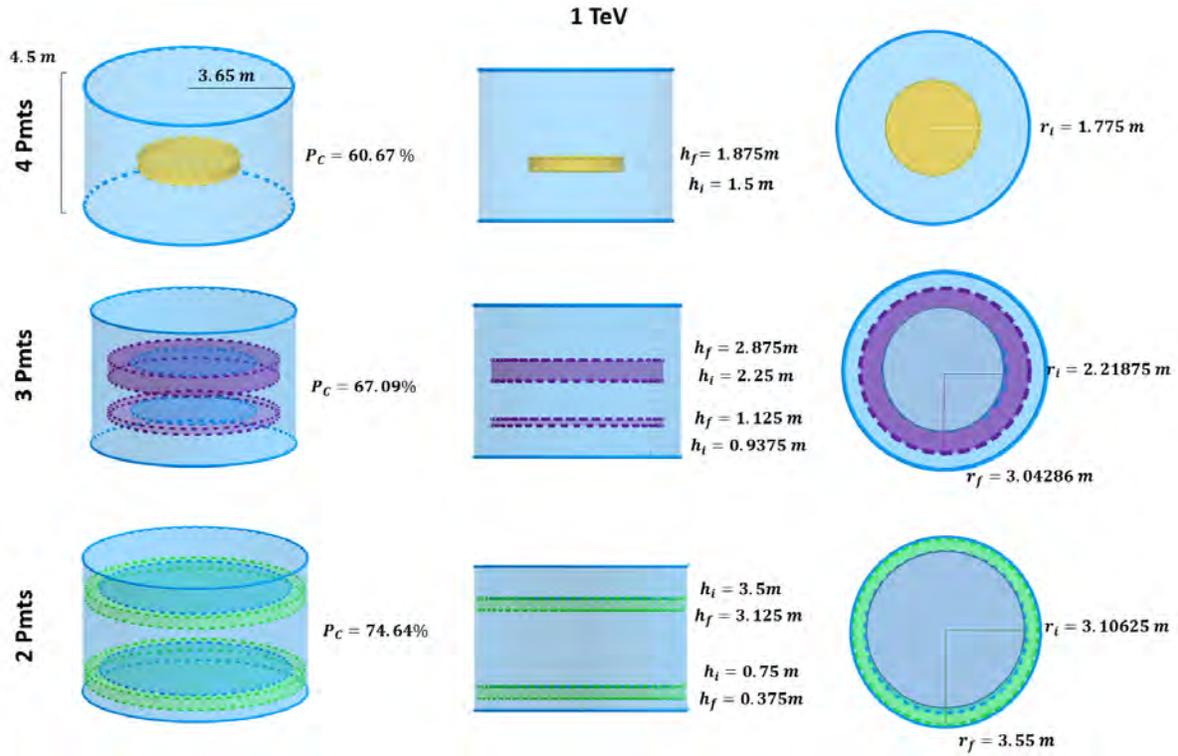


Figura 4.10: Imagen que representa la región en el tanque en donde es más probable que se activen 4,3, ó 2 PMTs para 1 TeV

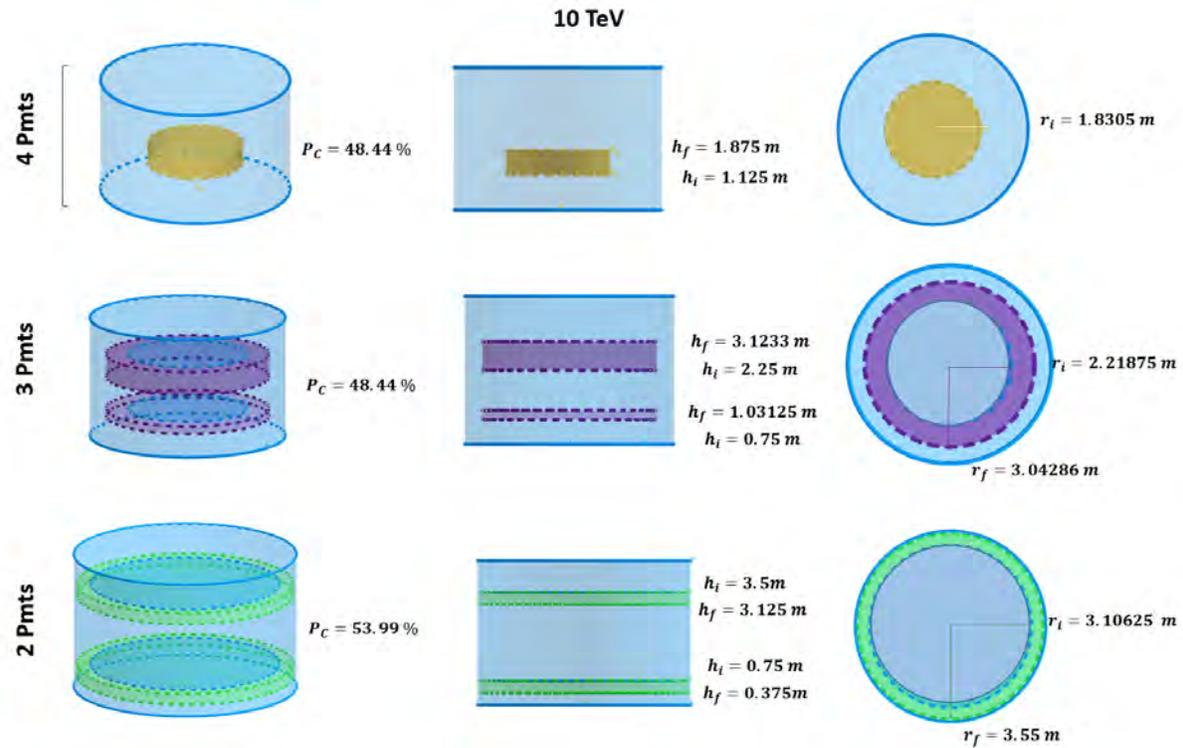


Figura 4.11: Imagen que representa la región en el tanque en donde es más probable que se activen 4,3, ó 2 PMTs para 10 TeV

Para complementar la información que se observó anteriormente. Al utilizar los datos que se obtuvieron en cada simulación, también se puede definir la región en el tanque en donde es más probable que se activen mínimo 3 ó 2 fotomultiplicadores. En la tabla 4.7, se colocan los puntos iniciales y finales para definir las alturas y los radios de estas dos regiones.

Posición	1 TeV		10 TeV	
	Región del tanque más probable de que se activen:			
	Mín. 3 PMTs	Mín. 2 PMTs	Mín. 3 PMTs	Mín. 2 PMTs
r_i	0	0	0	0
r_f	293.98	355	293.98	355
P_V	78.10	98.47	83.32	98.82
h_i	103.125	56.25	103.125	56.25
h_f	262.5	356.25	290.625	365.625
P_H	72.67	84.56	72.05	83.47

Tabla 4.7: Ubicación de la región en donde es más probable que se activen mínimo 3 ó 2 PMTs

A partir de la información de la tabla anterior, en la figura 4.12 se puede ver de forma gráfica las zonas en el tanque en donde es más probable que se activen mínimo 3 ó 2 fotomultiplicadores.

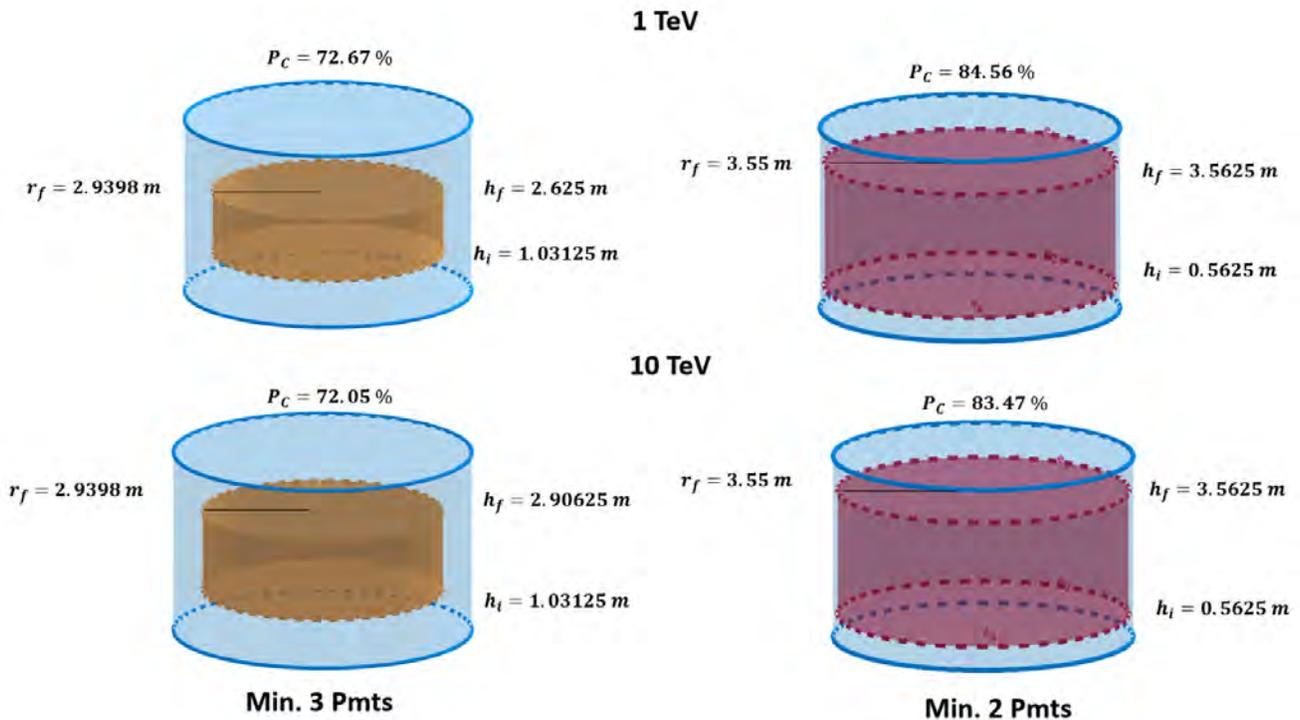


Figura 4.12: Imagen que representa la región en el tanque en donde es más probable que se activen al menos 3 ó 2 PMTs para 1 y 10 TeV

4.4.5. Prueba para el volumen efectivo de detección

Recordemos que el objetivo de las regiones que definimos en la sección anterior, consiste en mejorar el algoritmo de reconstrucción de trazas (que se comentó al inicio de este capítulo). En esta sección se realizó una primera simulación para probar a estas regiones. Esta prueba consiste en simular a un muón con trayectoria cuasi-horizontal que dejará una señal en una serie de tanques vecinos. Al fijarnos en los detectores activados (con una carga mayor o igual a 4 PE), se podrá conocer la región del tanque más probable por la cual paso esta partícula (en base a los diagramas que definimos en la sección anterior). La idea de esta simulación consiste en observar si realmente esta partícula pasa por estas regiones.

Descripción de la simulación: Se simularon 100 muones con una energía de 1 TeV. Estas partículas dejaron una señal en tres diferentes tanques: G20, I20 y K20 (La figura 2.11 muestra la ubicación de estos tanques en el arreglo del observatorio). La trayectoria de las partículas simuladas se puede observar en la figura 4.13.

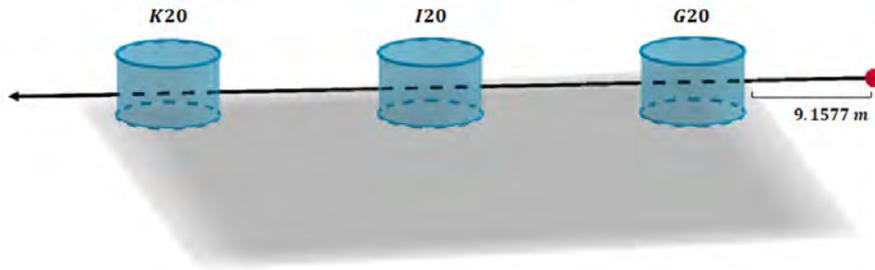


Figura 4.13: Trayectoria del muón para la simulación de prueba para el volumen efectivo de detección.

Resultados: En la tabla 4.8, se ven los resultados del porcentaje de detectores activados para esta simulación. A partir de esta información, se puede observar que es más probable que se activen 3 detectores en los tanques I20 y G20, y que se activen 2 detectores en el tanque K20. Por lo tanto, si el observatorio detecta la señal de un muón individual que venga en esa dirección, lo más probable es que se obtengan estos resultados en los tanques activados.

Detectores Activados	Tanque		
	G20	I20	K20
4 PMTs	17 %	20 %	1 %
3 PMTs	73 %	71 %	9 %
2 PMTs	10 %	9 %	89 %

Tabla 4.8: Porcentajes de detectores activados para la simulación de prueba

En la figura 4.14, se muestra un diagrama en el plano $x - y$ de la trayectoria del muón al pasar por los tanques mencionados. El inciso *a*) de esta imagen, presenta a las regiones en donde es más probable que se activen 3 ó 2 fotomultiplicadores (que son las zonas de color morado y verde respectivamente), y el inciso *b*) muestra las regiones en el tanque en donde es más probable que se

activen mínimo 3 ó 2 fotomultiplicadores (que son las zonas de color rojo y naranja). En la figura 4.15, se ve exactamente a estas mismas regiones (tanto para el inciso a como para el inciso b), pero vistas desde el plano $x - z$. Con estos dos diagramas se puede observar si realmente la trayectoria del muón paso por las regiones que definimos en la sección anterior.

Plano X - Y

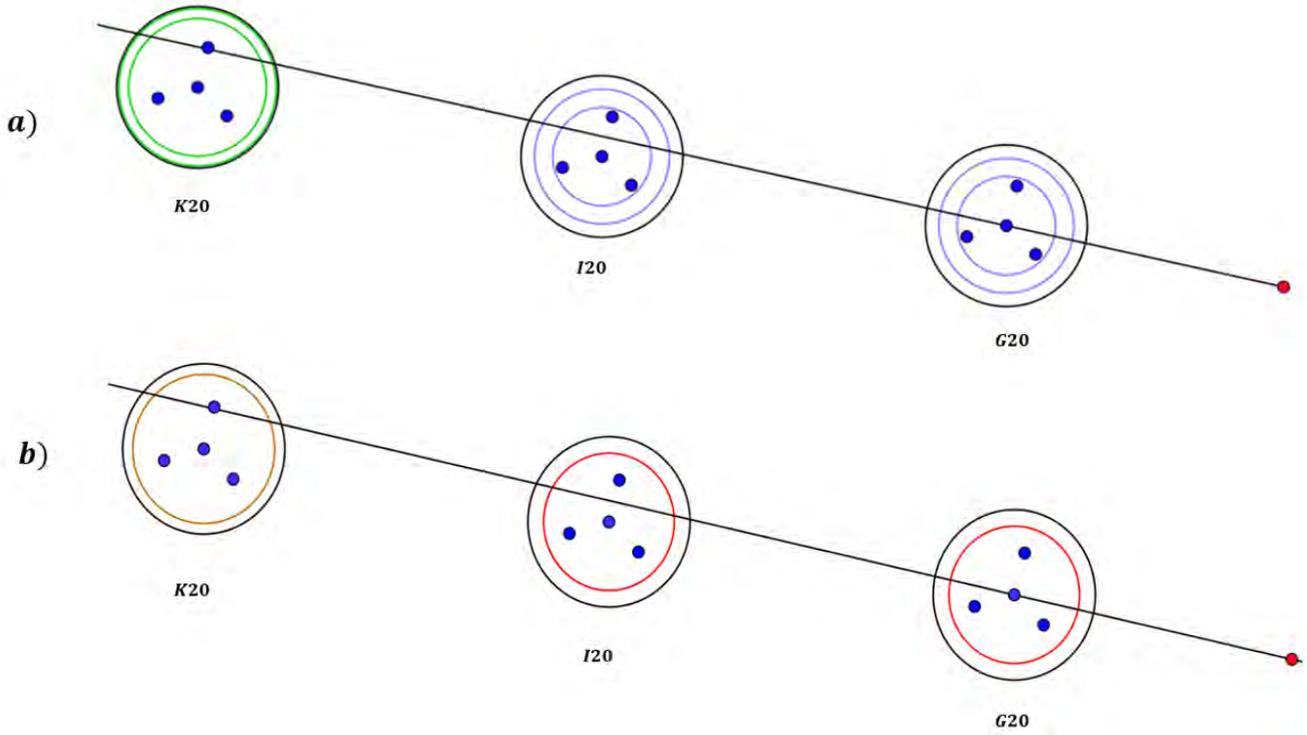


Figura 4.14: Vista desde el plano $x - y$ para los tanques activados en la simulación de prueba, en donde se compara la trayectoria del muón con las regiones que se definieron en la sección anterior.

Plano X - Z

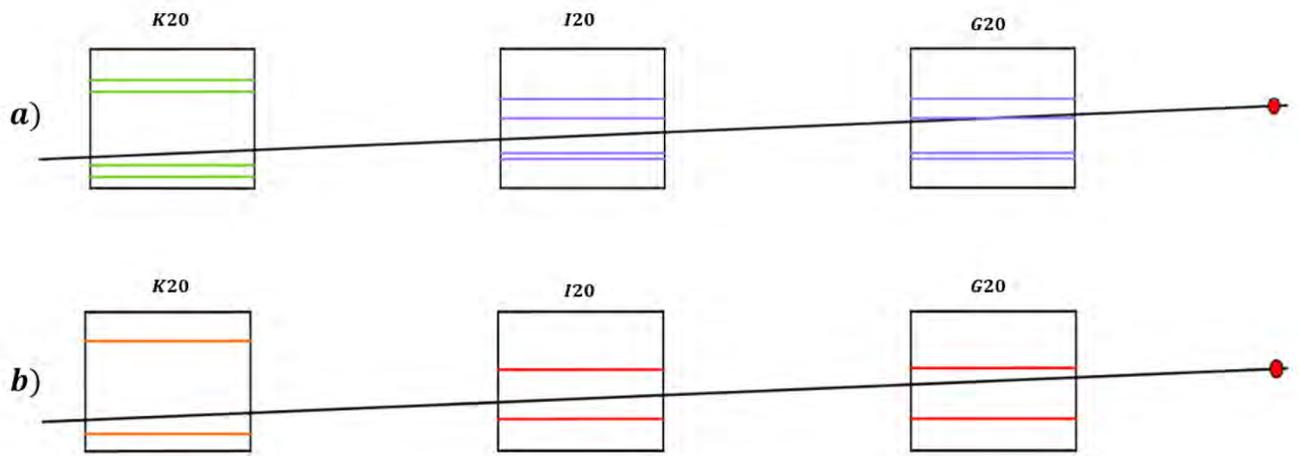


Figura 4.15: Vista desde el plano $x - z$ para los tanques activados en la simulación de prueba, en donde se compara la trayectoria del muón con las regiones que se definieron en la sección anterior.

En la siguiente sección se discutirán los resultados correspondientes a los diagramas que definimos como volumen efectivo de detección, así como los resultados obtenidos en esta simulación de prueba.

4.4.6. Discusión de Resultados

Como se puede observar en la figura 4.10, se definen las regiones más probables en el tanque por donde tiene que pasar un muón para que este active de 4 a 2 fotomultiplicadores. En todas estas regiones hay una probabilidad de al menos el 60 % de que el muón haya pasado por esa región. En las zonas en donde es más probable que se activen 3 ó 2 fotomultiplicadores se generan 2 cilindros.

Para definir al cilindro superior, se simularon muones que pasan por esa parte del tanque. En este caso, el cono de luz generado por la partícula que pasa por esa región, no alcanza a llegar a todos los fotomultiplicadores. Por ejemplo, para la simulación de muones horizontales enviados en la dirección H.I a una altura de 3 m, el 69.99 % de las veces se activaron 3 fotomultiplicadores. Para este caso, el fotomultiplicador A, fue el dispositivo que menos veces se activó en todos los eventos. En el inciso *a)* de la figura 4.16, se muestra un diagrama para el cono de luz de la partícula que se utilizó en esta simulación. Como se puede ver, la luz generada por esta partícula no alcanza a llegar al fotomultiplicador A. En general, esto sucede al simular muones en la parte superior del tanque, donde el cono de luz no logra cubrir a todos los detectores o apenas alcanza a cubrirlos, esto genera un cilindro en la parte superior del tanque en donde es más probable que se activen 3 ó 2 fotomultiplicadores (que depende de la dirección por la cual se simuló el muón).

Para el caso del cilindro inferior, se puede observar en el inciso *b)* de la imagen 4.16, una representación del cono de luz generado por la partícula simulada. Esta se envió en la dirección H.I a una altura de 0.375 m y el 96.26 % de las veces se activaron 2 fotomultiplicadores. Para este caso, los detectores A y B, fueron los dispositivos que menos veces se activaron en todos los eventos. Esto se debe a que todos los detectores están acoplados a un tubo de PVC de aproximadamente 30 cm de largo, y para este caso, la luz Cherenkov que se genera no puede acceder a la ubicación de los detectores restantes. Por este motivo se forma un cilindro en la parte inferior del tanque en donde es más probable que se activen 2 ó 3 fotomultiplicadores.

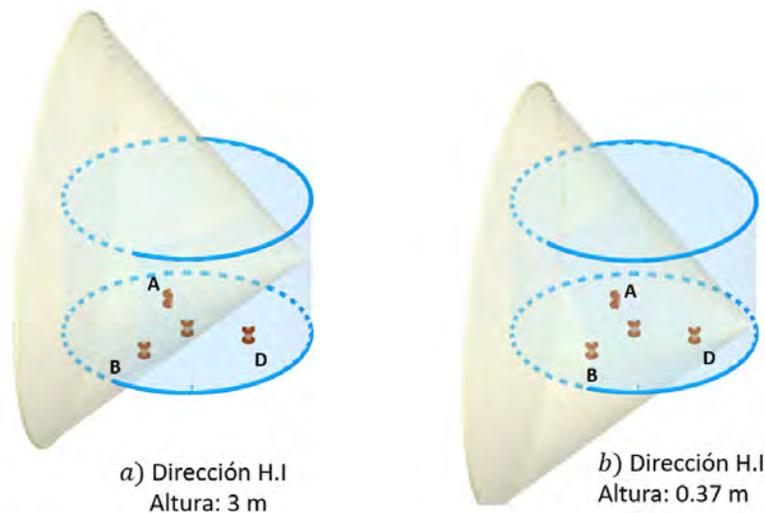


Figura 4.16: Representación del cono de luz para un muón enviado en H.I a 2 diferentes alturas.

Para la prueba del volumen efectivo de detección, en el inciso *a*) de las figuras 4.14 y 4.15, se muestra la región más probable del tanque por la cual pudo haber pasado el muón en base al método descrito en este capítulo. Para ser precisos, se muestra la región en donde es más probable que se activen 3 ó 2 fotomultiplicadores (tres para los tanques I20 y G20, y dos para el tanque K20) que corresponden a los detectores activados en cada tanque para esta prueba. Para la dirección azimutal, se puede observar que en efecto el muón simulado pasa por estas regiones, sin embargo para el ángulo de elevación, este muón solamente pasa por la región que se estableció para el tanque G20, para los otros 2 tanques, la trayectoria del muón queda fuera de las zonas establecidas (ver inciso *a*) de la figura 4.15). Aunque se observaron 3 ó 2 detectores en los tanques I20 y K20, el muón simulado no pasa a través de las regiones que definimos para estos casos.

En primer instancia, se simularon 100 muones que solamente pasaban por el tanque K20 y que llevaban la misma dirección de la primera simulación. Para esta prueba, al igual que en el caso anterior, se obtuvo que en ese tanque es más probable que se activen 2 detectores. Con esta información podemos descartar la idea de que el número de fotomultiplicadores activados en los tanques I20 y K20 sea menor debido a la posible pérdida de energía de la particular simulada al interactuar con los tanques anteriores, y que pudiera afectar en los volumen establecidos en cada tanque.

Tomando en cuenta lo anterior, la posible discrepancia entre los resultados de la simulación y los volúmenes que establecimos para cada tanque, se pueden explicar si consideramos que para definir la altura en todas las regiones, solamente se tomaron en cuenta a muones que pasaban por el centro y recorrían todo el diámetro del tanque. Para el tanque G20 el muón de esta prueba si pasa por el centro del tanque, y en este caso su trayectoria si pasa por el volumen establecido. Por otro lado, esta partícula no pasa por el centro y recorre una distancia menor al diámetro (en el plano $x - y$) en los tanques I20 y K20. De acuerdo a la trayectoria del muón, en los tanques I20 y G20 se deberían activar 4 y 3 fotomultiplicadores respectivamente, con esta condición, los volúmenes que se establecen en estos tanques estarían acorde con la zona por donde paso el muón. Pero en el caso de la simulación, debido a la sección del tanque que atravesó el muón, el cono de luz en el tanque I20 solamente alcanza a cubrir a tres fotomultiplicadores (ver inciso *a*) de la figura 4.17), y para el caso del tanque K20, el cono de luz solamente pasa de manera apropiada en 2 detectores (ver inciso *b*) de la figura 4.17). Esto explicaría la discrepancia en los resultados de estos tanques.

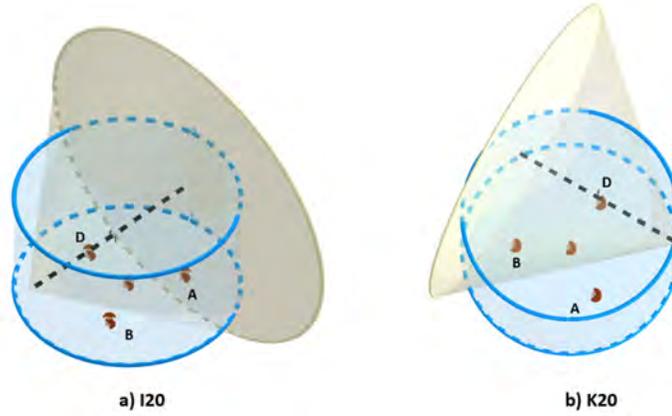


Figura 4.17: Representación del cono de luz para los tanques I20 y G20 de la simulación de prueba.

En el inciso *b*) de las figuras 4.14 y 4.15 se muestran las regiones en el tanque en donde es más probable que se activen mínimo 3 ó 2 fotomultiplicadores (tres para los tanque I20 y G20, y dos para el tanque K20). Como se puede observar en estos dos diagramas, el volumen que se definió en cada tanque corresponde a la zona por donde pasó el muón.

También podemos mencionar que para el caso de una energía de 10 TeV se encontró un volumen de detección más grande que para el caso de una energía de 1 TeV. Anteriormente se había comentado que se esperaba este resultado, debido a que una simulación con mayor energía también va a generar mayor cantidad de luz (debido a que se generan más partículas secundarias). Así que esto incrementa en el número de detectores en el tanque, dando como resultado un volumen mayor para una energía de 10 TeV.

Recordemos que la dirección del muón que generó esta señal estará entre dos intervalos (uno para el ángulo azimutal y otro para el ángulo de elevación). Para esta simulación de prueba, el tamaño de estos intervalos son los siguientes:

$$\begin{aligned}\phi_{max} - \phi_{min} &= 22.31^{\circ}. \\ \theta_{max} - \theta_{min} &= 2.92^{\circ}.\end{aligned}$$

Al tomar en cuenta a las regiones en el tanque en donde es más probable que se activen mínimo 3 ó 2 fotomultiplicadores (que se ven representadas en la figura 4.12), el tamaño de estos intervalos serán los siguientes:

$$\begin{aligned}\phi_{max} - \phi_{min} &= 19.89^{\circ}. \\ \theta_{max} - \theta_{min} &= 1.38^{\circ}.\end{aligned}$$

Por lo tanto en este ejemplo, se mejora un 10.85% la reconstrucción del ángulo zenital, y un 52.74% para el ángulo de elevación.

Región		Porcentaje que mejora para 1 TeV		Porcentaje que mejora para 10 TeV	
tanque inicial	tanque final	azimutal	Elevación	azimutal	Elevación
		ϕ	θ	ϕ	θ
min 2	min 2	2.67 %	32.07 %	2.67 %	32.07 %
min 2	min 3	10.85 %	52.97 %	10.84 %	42.52 %
min 3	min 3	19.11 %	63.03 %	19.11 %	57.12 %

Tabla 4.9: Reconstrucción de trazas usando las regiones en donde mínimo se activa 3 ó 2 PMTs

En general, la tabla 4.9 muestra que tanto mejora la reconstrucción al considerar los diferentes casos en donde al tanque inicial y al tanque final se le asigna una región en donde al menos se activan 3 ó 2 detectores.

Capítulo 5

Segundo análisis: Estimador de Energía para muones horizontales

En este capítulo se presenta un estudio para estimar la energía de muones horizontales. La respuesta de los detectores para estos muones con diferente energía se obtuvo a partir de simulaciones Monte Carlo. Con esta información se alimentó a una red neuronal artificial, que nos sirvió como mecanismo para estimar la energía de los muones incidentes en base a las señales detectadas.

Anteriormente se había comentado que este proyecto tiene la finalidad de estudiar a las señales de ruido involucradas en la búsqueda de neutrinos en el observatorio HAWC. Recordemos que este tipo de señales son muones que viajan de manera horizontal que activan una serie de tanques vecinos (trazas de muones) y que se pueden identificar a través de los datos de este observatorio. Por lo tanto, la motivación de esta prueba consiste en estimar la energía de los muones involucrados en estas trazas.

Así que en este capítulo se explica a detalle en que consiste esta prueba, empezando por la descripción de la red y de las simulaciones utilizadas, hasta presentar los resultados y la discusión de los mismos.

5.1. Descripción de la prueba

La descripción de esta prueba se divide en dos partes. En la primera parte, se mencionan todas las características de las simulaciones realizadas para obtener la información que nos servirán como datos de entrada para la red neuronal. Y en la segunda parte se especifican todos los detalles sobre la arquitectura de la red que se utiliza en este trabajo.

5.1.1. Descripción de las simulaciones

Para este caso (al igual que en el capítulo anterior) se utilizó HAWCSIM para realizar las simulaciones Monte Carlo. Los datos de la simulación son los siguientes:

- **Partículas a simular:** Se simularon muones con carga negativa (μ^-).
- **Dirección y posición inicial:** En esta prueba se simularon muones horizontales que pasan por la mitad del tanque E10 y se propagan hacia el tanque X10 (ver figura 2.8). De hecho en la sección 2.6 ya se había mencionado la dirección y la posición inicial de estas partículas, y esta información se puede ver de forma gráfica en la figura 2.8.

- **Energía:** Para esta prueba, se realizaron 11 simulaciones con diferente energía. Las energías que se tomaron en cuenta para esta prueba, son las siguientes: 100 GeV, 1 TeV, 2 TeV, 3 TeV, 4 TeV, 5 TeV, 6 TeV, 7 TeV, 8 TeV, 9 TeV y 10 TeV. Recordemos que las señales que estudiamos provienen de muones producidos en cascadas atmosféricas que atraviesan la montaña, por este motivo se simularon muones de alta energía.
- **Número de partículas:** Para cada energía se simularon 2000 partículas. Esta cantidad se escogió debido al tiempo en que se tardaba en terminar la simulación. Para una energía de 10 TeV, dos mil muones se simulaban en 5 días (con la misma computadora descrita en el capítulo 4). Para 10 mil muones de 4 TeV, la simulación se tardaba una semana y media. Por esta razón 2 mil partículas parecía un número adecuado para realizar esta prueba sin la necesidad de usar un cluster.

5.1.2. Descripción de la red neuronal artificial

Recordemos que una red neuronal artificial requiere de una lista de entradas y salidas. Con esta información se realiza el entrenamiento utilizando el método mencionado en el capítulo 3. Como resultado, se podrá dar cualquier dato de entrada y la red neuronal calculará el dato de salida en base al entrenamiento previo.

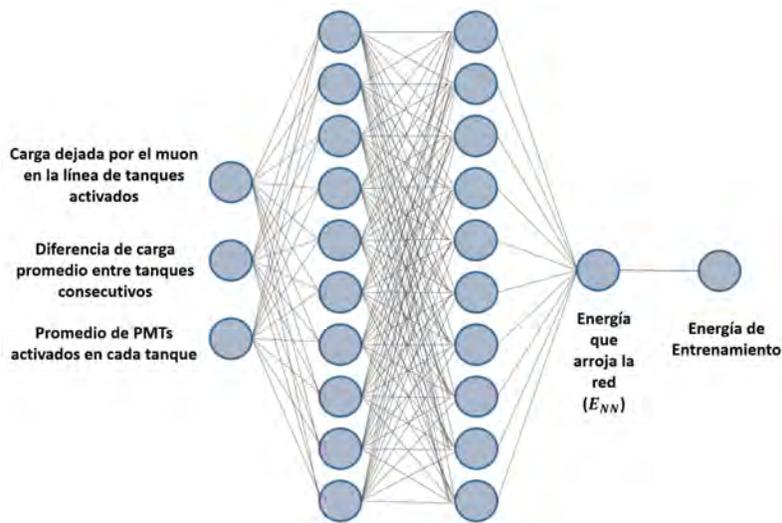


Figura 5.1: Estructura de la red neuronal utilizada para estimar la energía de muones horizontales

Anteriormente se habían mencionado dos ejemplos en donde se emplea una red neuronal artificial. En el primero de ellos se utiliza este algoritmo para identificar números escritos a mano, en donde los datos de entrada son los pixeles que conforman a esta imagen y que contienen diferentes tonos en la escala de grises. En el otro caso se emplea una red neuronal para aproximar una función de dos variables.

Para esta prueba la red neuronal se entrenó con los datos de las simulaciones realizadas, tomando diferentes variables correspondientes a la respuesta de los detectores, junto con la energía de los muones que dejaron esta señal. En la figura 5.1, se puede observar la estructura de la red neuronal utilizada, en donde se consideraron tres variables de entrada y dos capas ocultas (con diez neuronas cada capa). La red neuronal nos da un valor de la energía (E_{NN}) que va ir mejorando cuando se

compare con la energía que estamos utilizando para entrenar a la red neuronal; que corresponde a la energía de los muones simulados.

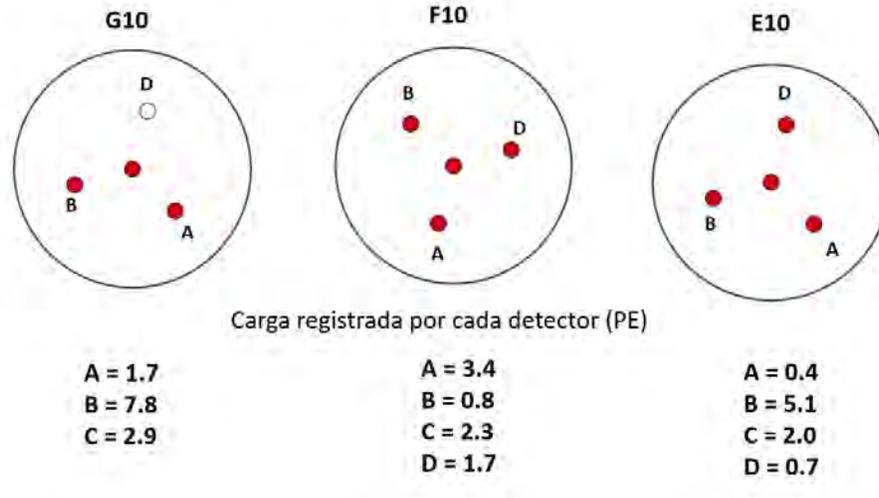


Figura 5.2: Ejemplo para explicar las variables de entrada de la red neuronal.

Primera variable: Carga dejada por el muón en la línea de tanques activados (C_a)

Esta variable se escogió, debido a que este valor representa la cantidad de carga promedio que deja el muón en un tanque. Y se quiere observar esta cantidad, ya que hay una relación entre la energía inicial de la partícula y la carga que se encuentra en los fotomultiplicadores.

Para obtener este valor, primero se obtiene la carga registrada en cada tanque individual. Para ello se suma la carga registrada en el tanque y se divide por el número de detectores activados. Después se van sumando estos valores y se divide entre el número de tanques que tuvieron señal.

Es importante mencionar que la carga que registra el PMT central se divide entre dos, ya que este detector registra aproximadamente el doble de carga que registran los otros detectores (ver apéndice B). A continuación mostraremos un ejemplo de como se calcula esta variable.

Ejemplo: En la figura 5.2, se muestra el caso en donde un muón horizontal pasa por tres tanques. Los detectores que se activaron se muestran en color rojo.

La carga en el tanque E10 sera:

$$E_{10} = \frac{0.4+5.1+(2/2)+0.7}{4} = 1.8 \text{ PE.}$$

En donde la carga del PMT central se divide entre dos. Después se suman las cargas que se obtuvieron en cada tanque y se dividen entre el número de tanques activados.

$$C_a = \frac{E_{10}+F_{10}+G_{10}}{3} = \frac{1.8+2.03+3.65}{3} = 2.49 \text{ PE.}$$

En cada evento simulado se obtiene este valor. Así que podemos realizar un histograma de esta variable para cada energía simulada.

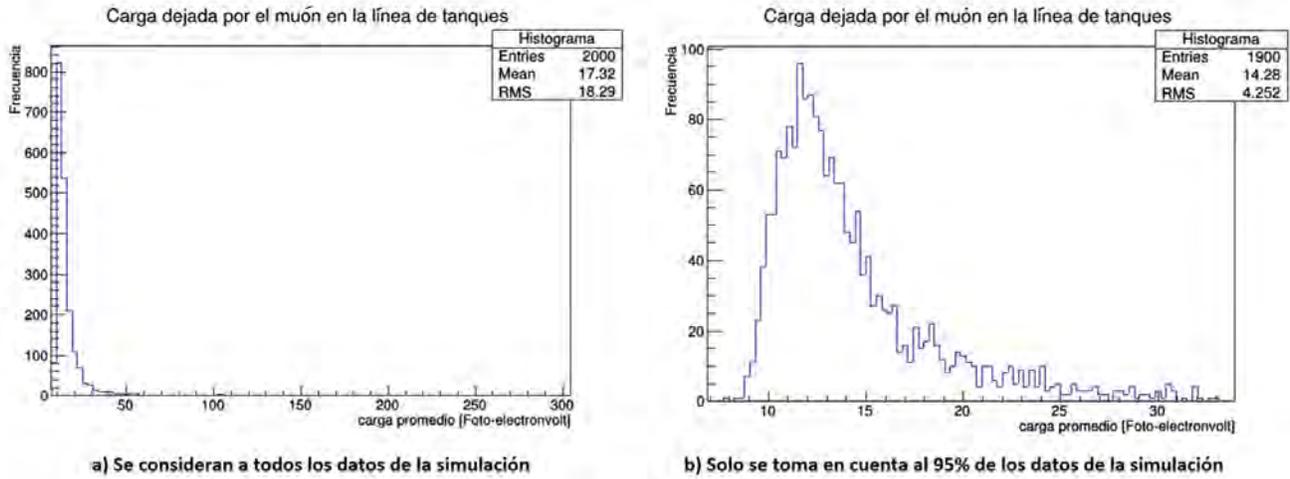


Figura 5.3: Histograma de la carga dejada por el muón en la línea de tanques para una energía de 1 TeV.

En la figura 5.3 se muestra un ejemplo para este tipo de histograma, en donde los datos utilizados corresponden a las simulaciones de muones con 1 TeV de energía. En el inciso a) de esta imagen, se muestra el histograma de la variable C_a para el caso en donde se toma en cuenta a todos los eventos de la simulación. En el inciso b) primero se realizó una lista de los valores de C_a ordenados de menor a mayor, después se llenó el histograma para este inciso con los primeros 1900 valores que conformaban a esta lista (el 95% de todos los eventos). Con esta consideración se puede apreciar de mejor manera la forma del histograma resultante.

En la tabla 5.1 se muestran los promedios y el valor del RMS (en los histogramas que muestra ROOT, este valor corresponde a la desviación estándar) que arroja cada histograma de las energías simuladas.

Energía (TeV)	Todos los eventos		95% de los eventos	
	Promedio	RMS	Promedio	RMS
0.1	10.32	2.355	9.911	1.198
1	17.32	18.29	14.28	4.252
2	24.85	36.34	18.8	7.75
3	32.54	57.82	23.14	11.16
4	38.97	71.86	27.58	14.21
5	46	82.57	31.87	18.02
6	57.28	120.3	36.09	21.67
7	62.3	124.2	40.73	24.63
8	70.19	152.2	45.73	28.87
9	78.98	171.8	51.68	32.27
10	90	201.9	54.69	38.48

Tabla 5.1: Promedio y RMS para los histograma de la variable C_a tomando en cuenta a todas las energías simuladas.

Con esta información, se puede observar en la figura 5.4 una gráfica de la energía en contra del promedio que arroja cada histograma de la variable (C_a). En donde se puede ver el comportamiento de esta variable en función de la energía.

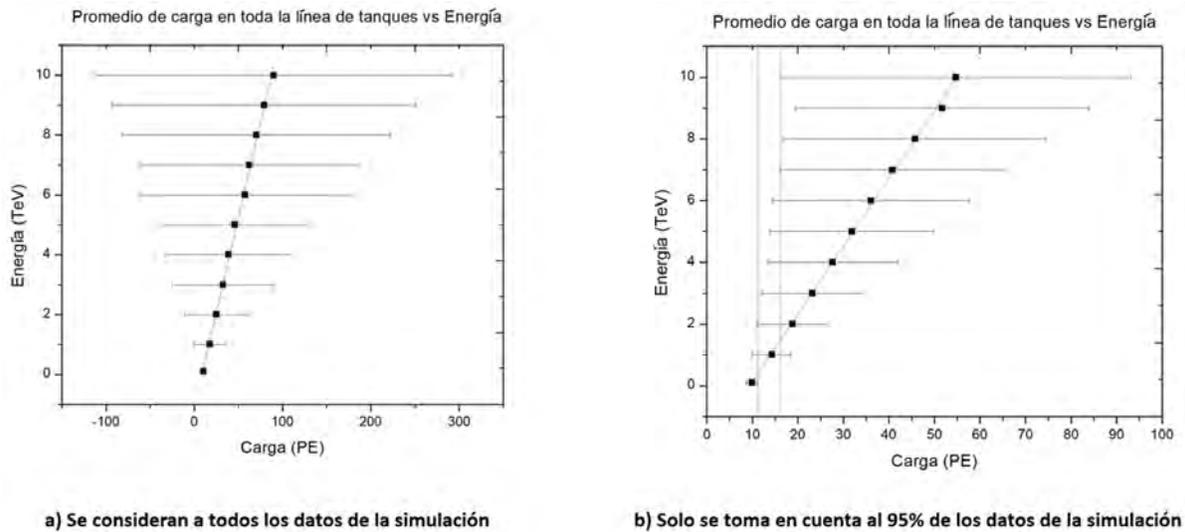


Figura 5.4: Gráfica del promedio de carga en toda la línea del tanque vs la energía.

En base a esta imagen podemos realizar las siguientes observaciones:

Observaciones del inciso a) de la figura 5.4

- La variable C_a sigue un comportamiento lineal al variar la energía del muón simulado.
- Las barras de incertidumbre (que corresponde al valor del RMS) para cada punto graficado se traslapan en todas las energías. Esto hace imposible identificar la energía del muón incidente en base a la variable C_a .

Observaciones del inciso b) de la figura 5.4

- La variable C_a sigue un comportamiento lineal al variar la energía del muón simulado.
- Las barras de incertidumbre en el caso de la energía de 0.1 TeV y 10 TeV no se traslapan, en otras palabras los datos del intervalo $[x_p - RMS, x_p + RMS]$ en ambas energías son diferentes. En la figura 5.4 se trazan dos líneas verticales para ver esto de forma gráfica y en la tabla 5.2 se presentan los valores de estos intervalos. Por lo tanto podemos decir lo siguiente de la energía del muón:

1) Si el valor C_a que se registra es menor o igual a 11.109 PE ($x_p + RMS$), entonces la energía del muón que genero esta señal será menor a 3 TeV. Esto se debe a que el intervalo de error en esa energía es mayor a 11.05 PE.

2) Si el valor C_a que se registra es mayor a 11.109 PE, entonces la energía del muón que genero esta señal será mayor a 100 GeV. Debido a que el intervalo de error en esa energía es menor a 11.05 PE.

Energía (TeV)	Promedio x_p	RMS	$x_p - RMS$	$x_p + RMS$
0.1	9.911	1.198	8.713	11.109
1	14.28	4.252	10.028	18.532
2	18.8	7.75	11.05	26.55
3	23.14	11.16	11.98	34.3
4	27.58	14.21	13.37	41.79
5	31.87	18.02	13.85	49.89
6	36.09	21.67	14.42	57.76
7	40.73	24.63	16.1	65.36
8	45.73	28.87	16.86	74.6
9	51.68	32.27	19.41	83.95
10	54.69	38.48	16.21	93.17

Tabla 5.2: Intervalos de incertidumbre para el inciso b) de la figura 5.4

Así que podemos dar esta información sobre la energía que genero esta señal con una efectividad del 95 %.

Segunda variable: Diferencia de carga promedio entre tanques consecutivos (D_c)

Esta variable nos muestra la diferencia promedio entre la carga de tanques vecinos. Y se eligió debido a que se quiere observar otra variable en donde intervenga la carga que el detector reciba (ya que este valor esta relacionado con la energía de la partícula). Además esta variable puede dar información diferente de C_a .

Para calcular este dato de entrada, primero se calculó la carga de dos tanques consecutivos. Después nos fijamos en la diferencia (en valor absoluto) de carga entre estos dos tanques, y se fue repitiendo este procedimiento con todos los tanques activados. Por último se calculó un promedio de esta diferencia entre tanques.

Ejemplo: Daremos un ejemplo de esta variable utilizando los datos de la figura 5.2. La carga de cada tanque se calculó al sumar la carga de cada detector activado (con la misma consideración de dividir la carga del PMT central entre dos). La carga del tanque E10 será la siguiente:

$$E_{10} = 0.4 + 5.1 + (2/2) + 0.7 = 7.2 PE.$$

La carga para los otros tanques será:

$$\begin{aligned} F_{10} &= 6.1 PE. \\ G_{10} &= 10.95 PE. \end{aligned}$$

Como sólo se activan tres tanques, la diferencia entre tanques consecutivos (D) serán las siguientes:

$$\begin{aligned} D_{E_{10}-F_{10}} &= |7.2 - 6.1| = 1.1 PE. \\ D_{F_{10}-G_{10}} &= |6.1 - 10.95| = 4.85 PE. \end{aligned}$$

Asi que la variable D_c será el promedio de estos valores.

$$D_c = \frac{1.1+4.85}{2} = 2.975 PE .$$

Al igual que en caso anterior, podemos llenar un histograma con la variable D_c de cada evento simulado. En la figura 5.5, se muestra uno de estos histogramas para una energía de 1 TeV. Para observar de mejor manera el comportamiento de este histograma, solamente se lleno con el 95 % de los datos.

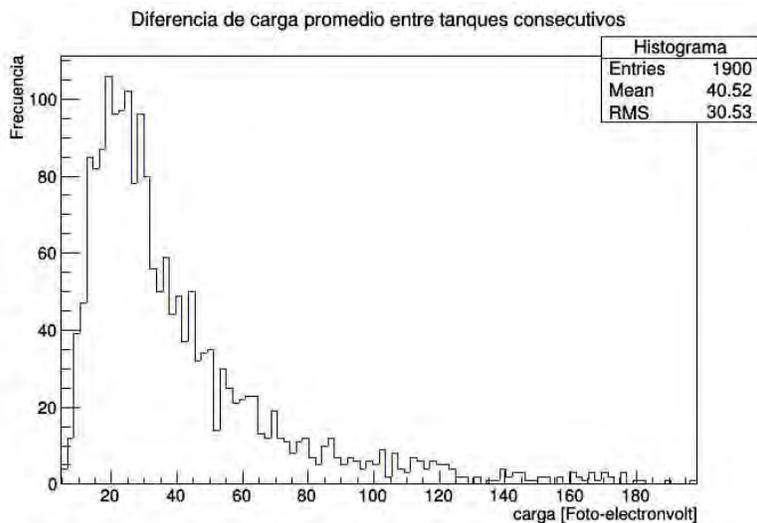


Figura 5.5: Histograma de la diferencia de carga entre tanques consecutivos para una energía de 1 TeV.

Utilizando el promedio y el valor RMS de cada histograma con una energía diferente, podemos realizar una gráfica de la variable D_c en contra de la energía. Esta gráfica se muestra en la figura 5.6 y se puede observar el comportamiento de esta variable en función de la energía.

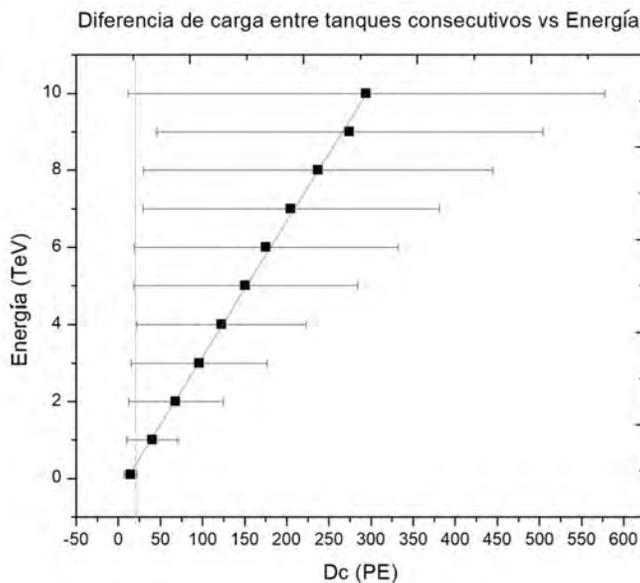


Figura 5.6: Gráfica de diferencia de carga entre tanques consecutivos en contra de la energía.

En base a esta imagen podemos realizar las siguientes observaciones:

Observaciones de la figura 5.6

- La variable D_c sigue un comportamiento lineal al variar la energía del muón simulado.
- La barra de incertidumbre en el caso de la energía de 0.1 TeV no se traslapa con las energías de 7, 8 y 9 TeV. Esto se puede observar en la tabla 5.3, que muestra el punto inicial (promedio - RMS) y el punto final (promedio + RMS) que definen a estas barras de incertidumbre para cada energía.

Energía (TeV)	Promedio x_p	RMS	$x_p - RMS$	$x_p + RMS$
0.1	14.4	7.737	6.663	22.137
1	40.52	30.53	9.99	71.05
2	68.27	56.02	12.25	124.29
3	96.1	80.7	15.4	176.8
4	122.4	100.3	22.1	222.7
5	151.1	132.6	18.5	283.7
6	175.4	156.3	19.1	331.7
7	205	176	29	381
8	236.9	207.1	29.8	444
9	274.6	229.5	45.1	504.1
10	294.4	282.8	11.6	577.2

Tabla 5.3: Intervalos de incertidumbre para la figura 5.6

Por lo tanto podemos decir lo siguiente de la energía del muón:

- 1) Si el valor D_c que se registra es menor o igual a 22.137 PE ($x_p + RMS$), entonces la energía del muón que genero esta señal será menor a 7 TeV o mayor o igual a 10 TeV.
- 2) Si el valor D_c que se registra es mayor a 22.137 PE, entonces la energía del muón que genero esta señal será mayor a 100 GeV.

Así que podemos dar esta información sobre la energía que genero esta señal con una efectividad del 95 %.

Tercera variable: Promedio de PMTs activados en cada tanque (P_a)

Esta variable se escogió debido a que el número de detectores que una partícula active, esta relacionado con la luz que esta genere. Y este valor también se relaciona con la energía de la partícula.

Para calcular esta variable, nos fijamos en la suma del número de detectores que se activan en cada tanque y lo dividimos entre el número de tanques activados.

Ejemplo: Daremos un ejemplo de esta variable utilizando los datos de la figura 5.2. Primero nos fijamos en la suma de los detectores que se activaron en cada tanque.

$$PMT_{s_{E10}} + PMT_{s_{F10}} + PMT_{s_{G10}} = 4 + 4 + 3 = 11.$$

Así que la variable P_a se obtiene al dividir el resultado anterior entre el número de tanques activados.

$$P_a = \frac{11}{3} = 3.66.$$

Al igual que en caso anterior, podemos llenar un histograma con la variable P_a de cada evento simulado. En la figura 5.7, se muestra uno de estos histogramas para una energía de 1 TeV. En este caso se utilizan todos los eventos simulados.

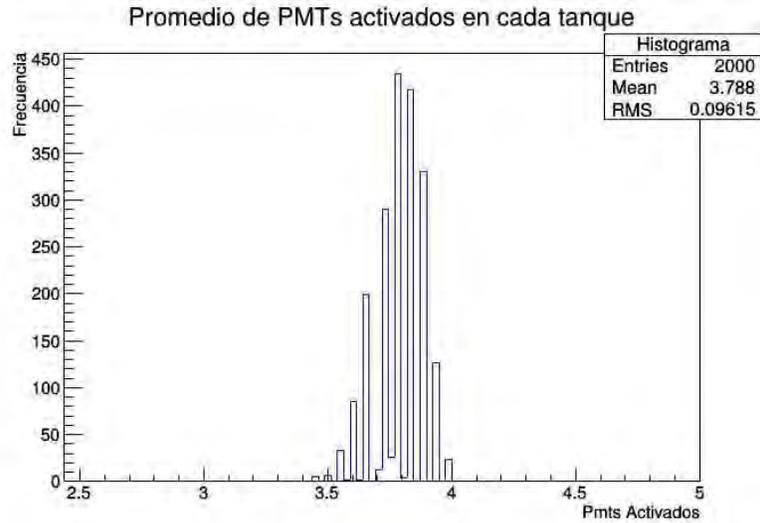


Figura 5.7: Histograma del promedio de detectores activados en un tanque para 1 TeV.

Utilizando el promedio y el valor RMS de cada histograma con una energía diferente, podemos realizar una gráfica de la variable P_a en contra de la energía. Esta gráfica se muestra en la figura 5.8 y se puede observar el comportamiento de esta variable en función de la energía.

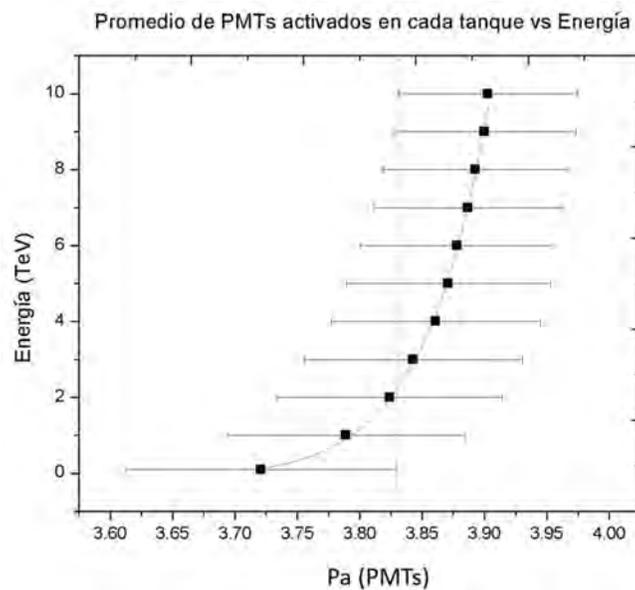


Figura 5.8: Gráfica del promedio de tanques activados en contra de la energía.

En base a esta imagen podemos realizar las siguientes observaciones:

Observaciones de la figura 5.8

- A los puntos de esta gráfica se les ajusto una función exponencial del tipo:

$$y = y_0 + Ae^{R_0x}.$$

- Las barras de incertidumbre en el caso de la energía de 0.1 TeV y 10 TeV no se traslapan; aunque para este caso estos intervalos están muy cercanos 3.8298 y 3.83133 para estas energías respectivamente. Esto se puede observar en la tabla 5.4, que muestra el punto inicial (promedio -RMS) y el punto final (promedio + RMS) que definen a estas barras de incertidumbre para cada energía.

Energía (TeV)	Promedio x_p	RMS	$x_p - RMS$	$x_p + RMS$
0.1	3.721	0.1088	3.6122	3.8298
1	3.789	0.0955	3.6935	3.8845
2	3.824	0.09051	3.73349	3.91451
3	3.843	0.0874	3.7556	3.9304
4	3.861	0.08359	3.77741	3.94459
5	3.871	0.08176	3.78924	3.95276
6	3.878	0.07769	3.80031	3.95569
7	3.887	0.07584	3.81116	3.96284
8	3.893	0.07452	3.81848	3.96752
9	3.9	0.07305	3.82695	3.97305
10	3.903	0.07167	3.83133	3.97467

Tabla 5.4: Intervalos de incertidumbre para la imagen 5.8

Por lo tanto podemos decir lo siguiente de la energía del muón:

- 1) Si el valor P_a que se registra es menor o igual a 3.8298 ($x_p + RMS$), entonces la energía del muón que genero esta señal será menor a 10 TeV.
- 2) Si el valor P_a que se registra es mayor a 3.8298, entonces la energía del muón que genero esta señal será mayor a 100 GeV.

Estas son las tres variables que se utilizan para entrenar a la red neuronal. Se usaron estos datos debido a que se ve (al menos en el promedio de eventos) una dependencia con la energía del muón que genero la señal. La idea de esta red neuronal consiste en que a partir de las variables C_a , D_c y P_a se puede decir algo sobre la energía del muón individual.

5.1.3. Datos para entrenar y probar a la red neuronal

Anteriormente se había mencionado que cada simulación cuenta con 2 mil eventos. Estos ya vienen enumerados en el archivo que se obtuvo a partir de la simulación. Así que para entrenar a la red neuronal se utilizó a los eventos con número par y para probar la red se utilizó a los eventos con número impar.

En la figura 5.9, se muestran dos histogramas para la variable C_a con una energía de 1 TeV. El inciso *a*) de esta imagen se llena con los eventos impares de la simulación, y el inciso *b*) con los eventos pares. Como se puede observar, el valor del promedio y de RMS no cambia mucho en ambos casos, incluso es parecido al que se obtiene cuando se consideran todos los eventos (ver inciso *a*) de la figura 5.3). Por este motivo se considero adecuado usar la selección anterior para entrenar y probar a la red neuronal, ya que el comportamiento de las variables que se observan se mantienen.

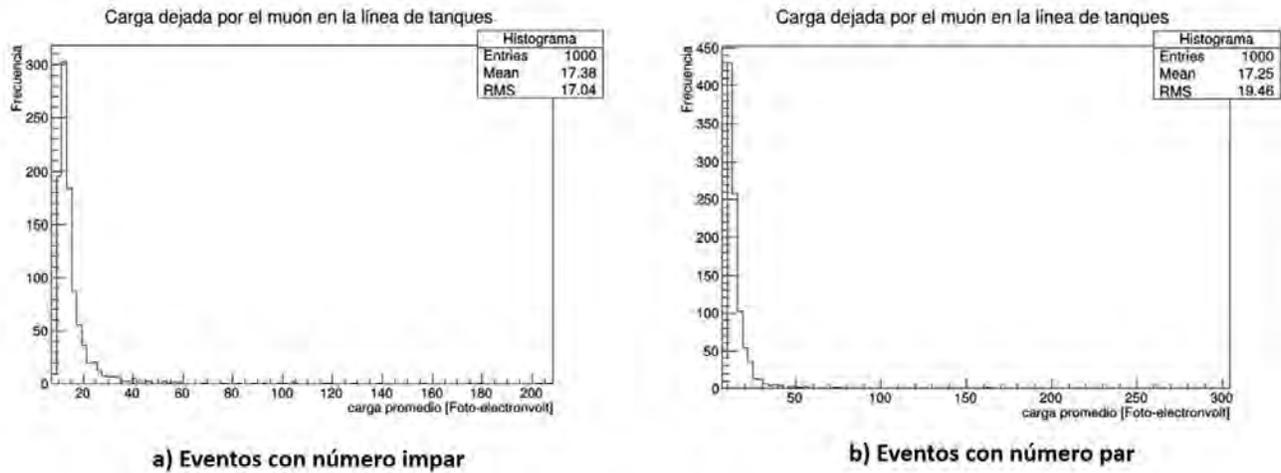


Figura 5.9: Histograma de la carga dejada por el muón en la línea de tanques para eventos pares e impares.

Para entrenar y probar a la red neuronal, se utilizan a las 11 diferentes energías simuladas. Por lo tanto se utilizan un total de 11,000 datos para entrenar a la red neuronal, y otros 11,000 datos para probar a esta red neuronal.

5.2. Resultados y Análisis

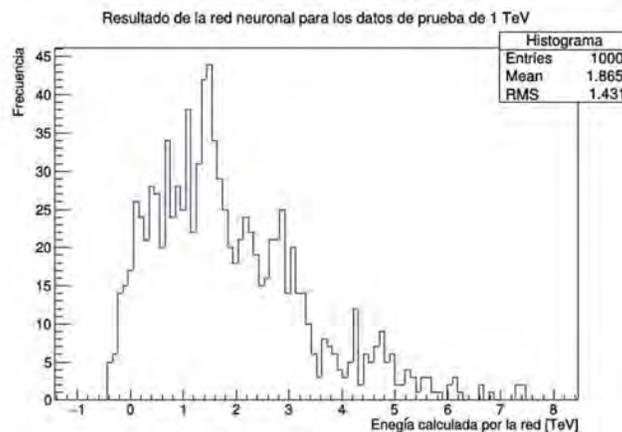


Figura 5.10: Histograma de la energía calculada por la red neuronal para una energía de 1 TeV.

Después de entrenar a la red neuronal, se usaron los datos de prueba para calcular las variables C_a , D_c y P_a . Estas variables se pusieron en la red neuronal y se observó el resultado que este procedimiento arrojaba. En la figura 5.10, se muestra un histograma de la energía que calculó la red neuronal utilizando los datos de prueba correspondientes a una energía de 1 TeV.

La tabla 5.5, muestra el promedio y el valor de RMS de los histogramas correspondientes a la energía que calculó la red neuronal en base a los datos de prueba de las diferentes energías utilizadas.

Energía de los datos de prueba utilizados (TeV)	Energía promedio calculada por la red E_{NN} (TeV)	RMS
0.1	0.2258	0.7355
1	1.865	1.431
2	3.174	1.67
3	4.039	1.662
4	4.7	1.67
5	5.134	1.548
6	5.462	1.478
7	5.85	1.416
8	6.118	1.324
9	6.425	1.266
10	6.57	1.217

Tabla 5.5: Energía promedio calculada por la red neuronal

Utilizando los datos de la tabla anterior, se puede realizar una gráfica que compare la energía que calculó la red neuronal y la energía que se esperaba. Esta gráfica se muestra en la figura 5.11.

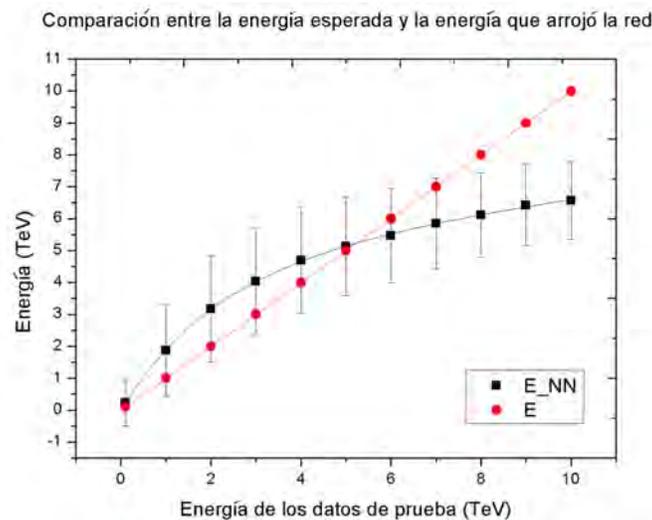


Figura 5.11: Comparación entre la energía que arrojó la red neuronal (E_{NN}) y la energía que se esperaba (E).

Observaciones de la figura 5.11

- A los puntos correspondientes a la energía que calculó la red neuronal (E_{NN}) se le ajustó una función exponencial del tipo:

$$E_{NN} = A_1 e^{\frac{-E}{t_1}} + A_2 e^{\frac{-E}{t_2}} + y_0. \tag{5.1}$$

Como se puede observar, esta función ajusta de buena manera a los puntos graficados.

- La diferencia entre la energía que se espera y la energía promedio que calcula la red neuronal ($E - E_{NN}$), se va incrementando de forma lineal a partir de la energía de 5 TeV. En la figura 5.12, se puede observar este incremento, en donde se gráfica la diferencia de estas energías. Los puntos negros de esta imagen representan a los valores de $(E - E_{NN})$ calculados con los resultados de E_{NN} obtenidos con los datos de prueba y la red neuronal, y los puntos de color rojo representan a los valores de $(E - E_{NN})$ calculados con los valores de E_{NN} obtenidos con la ecuación 5.1. Esta extrapolación se hizo debido a que en las condiciones normales de GEANT4 no se pueden simular energías mayores a 10 TeV. Y con la ecuación 5.1 se puede dar información para energías más altas.

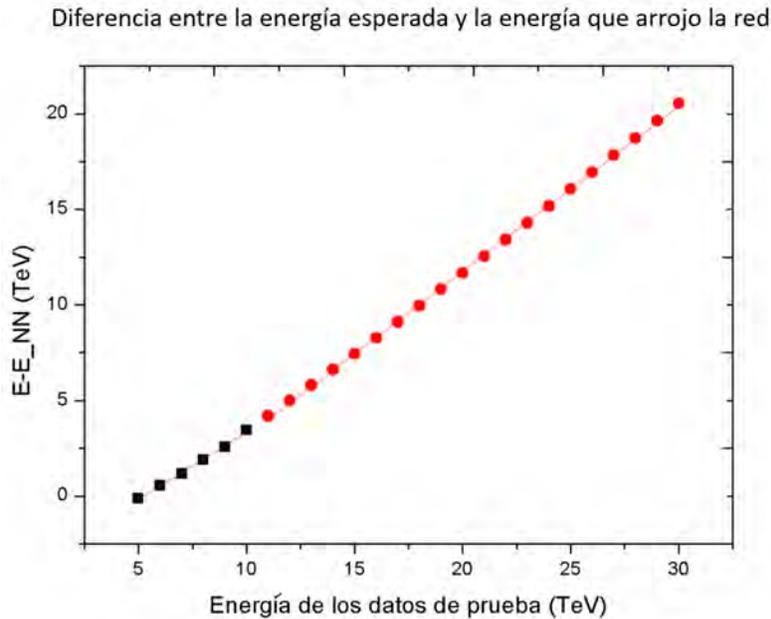


Figura 5.12: Gráfica de la diferencia entre la energía esperada con la energía que calculo la red neuronal.

Este incremento para la diferencia $E - E_{NN}$ parece razonable, recordemos que para las variable C_a y D_c el valor de la desviación estándar también incrementa al aumentar la energía de los muones (ver figuras 5.4 y 5.6). Así que para altas energías, el valor de estas cantidades será muy disperso, esto va a repercutir a la red neuronal y hará más difícil identificar a los muones con altas energías. Por este motivo la red neuronal sigue la forma mostrada en la ecuación 5.1, en donde empieza a clasificar a muones muy energéticos con valores similares de E_{NN} hasta que llega a un valor asintótico.

También nos fijamos en el error porcentual para cada energía utilizada en los datos de prueba. Este error se expresa como:

$$\frac{|E - E_{NN}|}{E} 100.$$

En la figura 5.13, se muestra un histograma que se llenó con este valor. En este caso solo se tomo en cuenta a los datos de prueba correspondiente a una energía de 1 TeV.

La tabla 5.6, muestra al error porcentual promedio y su desviación estándar para todos los datos de prueba, esta información esta separada por las diferentes energías que se utilizaron en las simulaciones.

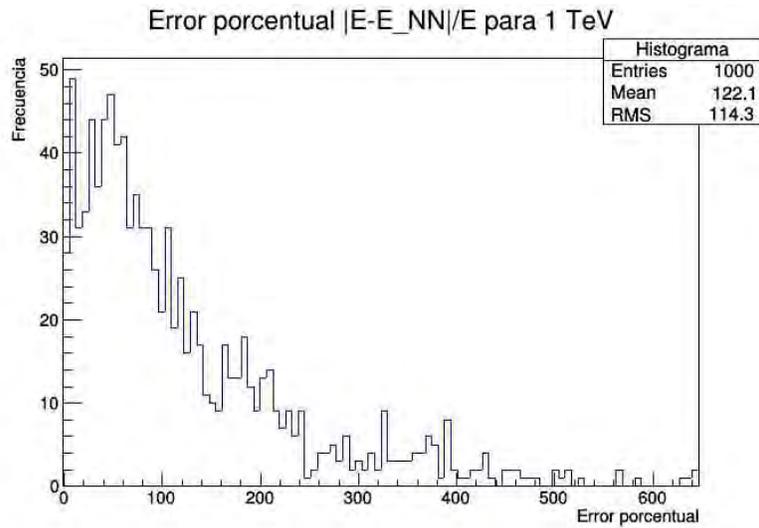


Figura 5.13: Histograma del error porcentual para los datos de prueba que corresponden a una energía de 1 TeV.

Energía de los datos de prueba utilizados (TeV)	Error porcentual promedio $ E - E_{NN} /E$ (%)	RMS
0.1	489.4	563.3
1	122.1	114.3
2	80.59	62.6
3	52.49	38.91
4	36.67	24.41
5	25.43	17.86
6	20.31	16.58
7	19.91	16.82
8	23.81	16.14
9	28.61	14.07
10	34.3	12.17

Tabla 5.6: Error porcentual entre la energía calculada por la red y la energía esperada.

Esta variable nos indicará la precisión de la medida y el error estará indicado por el valor RMS. En esta tabla se puede observar que para bajas energías la precisión de la medición es mala. Pero esto va mejorando conforme la energía aumenta (ya que este error empieza a disminuir). Y llega

a sus valores más bajos en las energías que van de 5 a 7 TeV. Después de esto el error porcentual empieza a aumentar. Este comportamiento está de acuerdo con la información mostrada en la gráfica de la figura 5.11. Para bajas energías el valor promedio que calcula la red neuronal para las energías de 0.1 TeV y 1 TeV son de 0.2258 TeV y 1.865 TeV respectivamente, por este motivo se tiene un porcentaje de incertidumbre más alto en esos casos. Para muones muy energéticos, se había comentado que la diferencia entre la energía esperada y la energía que arroja la red neuronal va aumentando conforme a la energía del muón. Por este motivo el error porcentual también empieza a aumentar si la energía se incrementa.

Ahora nos fijaremos en el comportamiento del error asociado a la energía promedio que calcula la red neuronal. Para ello usaremos los valores de RMS mostrados en la tabla 5.6, con esta información se realiza una gráfica de la variación de la desviación estándar conforme la energía de los datos de prueba aumenta. Este gráfica se muestra en la figura 5.14. Como se puede observar el valor de la desviación estándar va disminuyendo.

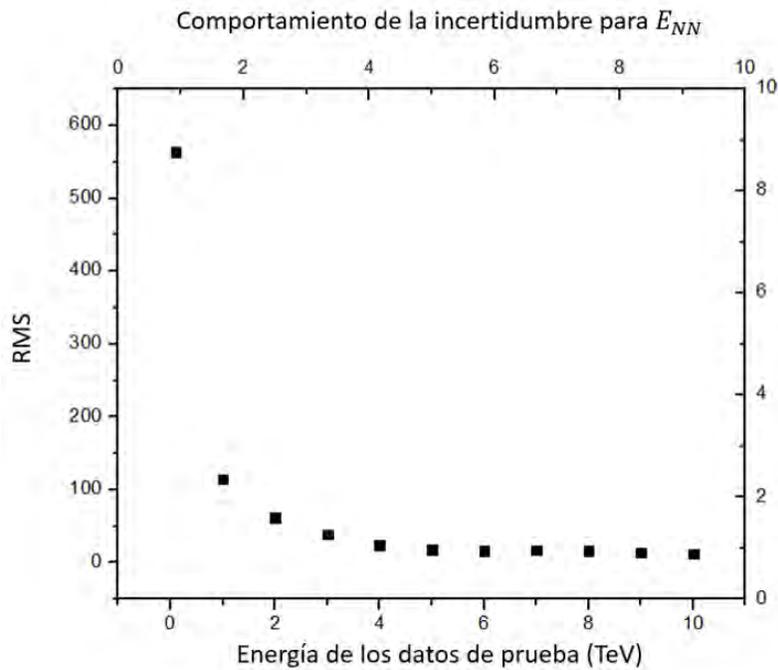


Figura 5.14: Gráfica que muestra el comportamiento de la desviación estándar para la energía promedio calculada por la red.

Esto se debe al comportamiento asintótico que tiene la red neuronal para estimar la energía de los muones, y que está descrito por la ecuación 5.1. De acuerdo a esta ecuación, para energías muy elevadas el resultado promedio de la energía que calcula la red neuronal será prácticamente el mismo. Por ejemplo:

$$\begin{aligned}
 E_{NN}(140 \text{ TeV}) &= 11.71. \\
 E_{NN}(200 \text{ TeV}) &= 11.73. \\
 E_{NN}(1000 \text{ TeV}) &= 11.74.
 \end{aligned}$$

Esto explicaría porque la desviación estándar empieza a disminuir y también tiende a un valor

asintótico conforme la energía aumente; ya que la energía promedio que estima la red neuronal también lo hace.

5.2.1. Energía para traza de muones individuales

Hasta ahora solo hemos revisado los promedios de la energía que calcula la red neuronal al tomar en cuenta los datos de prueba para diferentes energías. Para dar información sobre la energía de muones individuales, nos fijaremos en la energía promedio que calculó la red neuronal y en su correspondiente barra de incertidumbre (calculada con el valor de su desviación estándar). Esta información se muestra en la tabla 5.7.

Energía de los datos de prueba utilizados (TeV)	Energía promedio calculada por la red E_{NN} (TeV)	RMS	$E_{NN} - RMS$	$E_{NN} + RMS$	Probabilidad (%)
0.1	0.2258	0.7355	-0.5097	0.9613	87.1
1	1.865	1.431	0.434	3.296	70.8
2	3.174	1.67	1.504	4.844	64.4
3	4.039	1.662	2.377	5.701	67
4	4.7	1.67	3.03	6.37	68.5
5	5.134	1.548	3.586	6.682	65.4
6	5.462	1.478	3.984	6.94	65.9
7	5.85	1.416	4.434	7.266	66.3
8	6.118	1.324	4.794	7.442	68.5
9	6.425	1.266	5.159	7.691	73.2
10	6.57	1.217	5.353	7.787	72.5

Tabla 5.7: Intervalos de incertidumbre para la energía calculada por la red.

Recordemos que para probar a esta red neuronal, en cada una de las energías que se muestra en la tabla anterior se utilizaron mil datos de prueba. Como se ve en la figura 5.10, no todos los eventos están en el intervalo que se definió en la tabla 5.7. Por este motivo, la sexta columna de esta tabla indica el porcentaje de eventos que se encuentra adentro del intervalo $[E_{NN} - RMS, E_{NN} + RMS]$.

Para dar información sobre la energía de muones individuales, nos fijaremos en los intervalos que no se traslapen en la tabla 5.7. Por ejemplo, el intervalo de error correspondiente a los datos de prueba de 0.1 TeV no se traslapan con los intervalos correspondientes a energías mayores o iguales a 2 TeV. Por lo tanto si la red neuronal arroja un resultado menor o igual a 0.9613 la energía del muón individual será menor a 2 TeV. Para verificar esta afirmación, nos fijamos en el número de eventos con un resultado menor a 0.9613 en las diferentes energías que se tomaron en cuenta para esta prueba (la sexta columna de la tabla anterior nos indica que puede haber muones que registren este resultado y que tengan una energía superior a 2 TeV). Esta información se presenta en la tabla 5.8. Por lo tanto, si la red neuronal arroja un resultado menor a 0.9613 la energía del muón será menor a 2 TeV, y esta afirmación tendrá una eficiencia del 89.95 %.

La tabla 5.9, muestra la información completa para estimar la energía de muones individuales.

Energía de los datos de prueba utilizados (TeV)	Número de Eventos	Porcentaje de Eficiencia (%)
0.1	871	89.95
1	292	
2	76	10.05
3	34	
4	11	
5	7	
6	2	
7	0	
8	0	
9	0	
10	0	
suma	1293	100

Tabla 5.8: Porcentaje de eficiencia para los resultados de la red, en el intervalo $[E_{NN} - RMS, E_{NN} + RMS]$ que va de -0.5097 a 0.961 que corresponde a una energía de 0.1 TeV

Esta tabla se realizó en base a los intervalos que no se traslapan en los resultados mostrados en la tabla 5.7. Con la tabla anterior, se puede dar información para traza de muones individuales que

Condición para el resultado de la red (R_{NN})	Energía del muón individual (E)	Eficiencia (%)
$R_{NN} \leq 0.9613$	$E < 2 \text{ TeV}$	89.95
$0.9613 < R_{NN} \leq 3.296$	$0.1 \text{ TeV} < E < 5 \text{ TeV}$	76.55
$3.296 < R_{NN} \leq 4.844$	$1 \text{ TeV} < E < 9 \text{ TeV}$	85.75
$R_{NN} > 4.844$	$E > 2 \text{ TeV}$	96.2

Tabla 5.9: Energía para muones individuales utilizando los resultados de la red neuronal

el observatorio detecte. Solo basta con observar las tres variables que comentamos anteriormente. Por ejemplo, supongamos que se detecta una traza para un muón horizontal con la siguiente información:

$$C_a = 11.7919 \text{ PE.}$$

$$D_c = 10.3853 \text{ PE.}$$

$$P_a = 3.77778$$

Con estos datos, el resultado que arrojo la red neuronal fue el siguiente:

$$R_{NN} = 1.53.$$

Por lo tanto, de acuerdo a los resultados presentado en la tabla 5.9. La energía para este muón será mayor a 100 GeV y menor a 5 TeV (con una eficiencia del 76.55 %).

Considerando que los resultados arrojados por la red neuronal se ajustan a la función mostrada en la ecuación 5.1. Se pueden extrapolar los resultados presentados en la tabla 5.9. Para ello se toman en cuenta los siguientes puntos:

- El error de dispersión de la energía calculada para el resultado de la red neuronal (R_{NN}) va disminuyendo (ver figura 5.14).
- Como el resultado estimado por la red neuronal va tendiendo a un valor asintótico conforme la energía aumenta, se espera que el error de dispersión que lo acompaña también tienda a un valor asintótico conforme la energía vaya aumentando.

Condición para el resultado de la red (R_{NN})	Energía del muón individual (E)
$R_{NN} \leq 0.9613$	$E < 2 \text{ TeV}$
$0.9613 < R_{NN} \leq 3.296$	$0.1 \text{ TeV} < E < 5 \text{ TeV}$
$3.296 < R_{NN} \leq 4.844$	$1 \text{ TeV} < E < 9 \text{ TeV}$
$4.844 < R_{NN} \leq 5.701$	$2 \text{ TeV} < E < 14 \text{ TeV}$
$5.701 < R_{NN} \leq 6.37$	$3 \text{ TeV} < E < 18 \text{ TeV}$
$6.37 < R_{NN} \leq 6.682$	$4 \text{ TeV} < E < 21 \text{ TeV}$
$6.682 < R_{NN} \leq 6.941$	$5 \text{ TeV} < E < 22 \text{ TeV}$
$6.941 < R_{NN} \leq 7.266$	$6 \text{ TeV} < E < 24 \text{ TeV}$
$7.266 < R_{NN} \leq 7.442$	$7 \text{ TeV} < E < 26 \text{ TeV}$
$7.442 < R_{NN} \leq 7.691$	$8 \text{ TeV} < E < 28 \text{ TeV}$
$7.691 < R_{NN} \leq 7.787$	$9 \text{ TeV} < E < 29 \text{ TeV}$
$7.787 < R_{NN} \leq 8.361$	$10 \text{ TeV} < E < 37 \text{ TeV}$
$8.361 < R_{NN} \leq 8.562$	$11 \text{ TeV} < E < 39 \text{ TeV}$
$8.562 < R_{NN} \leq 8.752$	$12 \text{ TeV} < E < 43 \text{ TeV}$
$8.752 < R_{NN} \leq 8.933$	$13 \text{ TeV} < E < 47 \text{ TeV}$
$8.933 < R_{NN} \leq 9.107$	$14 \text{ TeV} < E < 51 \text{ TeV}$
$9.107 < R_{NN} \leq 9.273$	$15 \text{ TeV} < E < 54 \text{ TeV}$
$9.273 < R_{NN} \leq 9.432$	$16 \text{ TeV} < E < 60 \text{ TeV}$
$9.432 < R_{NN} \leq 9.584$	$17 \text{ TeV} < E < 64 \text{ TeV}$
$9.584 < R_{NN} \leq 9.730$	$18 \text{ TeV} < E < 74 \text{ TeV}$
$9.730 < R_{NN} \leq 9.871$	$19 \text{ TeV} < E < 82 \text{ TeV}$
$9.871 < R_{NN} \leq 10.006$	$20 \text{ TeV} < E < 111 \text{ TeV}$
$R_{NN} > 10.006$	$E > 20 \text{ TeV}$

Tabla 5.10: Extrapolación para estimar la energía de muones individuales.

Si se toman en cuenta los puntos anteriores, se pueden realizar intervalos de error ($R_{NN} \pm RMS$) tomando como base el resultado de la desviación estándar para una energía de 5 TeV (ver tabla 5.7). Este valor se escogió debido a que a partir de esa energía el error de dispersión disminuye de forma asintótica, así que este valor representa el peor escenario posible. En el apéndice E, se muestra la tabla para los intervalos ($R_{NN} \pm RMS$) calculados a partir de la ecuación 5.1 y del error asociado a una energía de 5 TeV. Si nos fijamos en los intervalos que no se traslapan de esta tabla, podemos extender los resultados para estimar la energía de muones individuales; estos datos se

presentan en la tabla 5.10. Con estos resultados se pueda dar más información sobre la energía de muones individuales, aunque para estos intervalos no se pueda dar un porcentaje de eficiencia. Por el momento la tabla 5.9 es la única referencia presentada en este trabajo para estimar la energía de muones individuales que presenta este tipo de porcentaje y que se recomienda usar para datos reales.

Conclusiones Generales

En este trabajo se presentaron dos análisis de simulaciones Monte Carlo para optimizar la búsqueda de muones en el observatorio HAWC.

En el primer análisis se estudiaron dos diferentes propuestas. La primera fue identificar la zona en donde es más probable que se activen exactamente 4, 3 ó 2 fotomultiplicadores, y en la segunda, la zona en donde es más probable que se activen al menos 3 ó 2 detectores. Para mejorar la reconstrucción en la dirección de muones horizontales, se llegó a la conclusión de que es más apropiado utilizar la zona del tanque en donde es más probable que se activen un mínimo 3 ó 2 detectores. Aunque utilizando los datos de la primera propuesta se puede reconstruir la trayectoria del muón con mayor precisión, se observó en la simulación de prueba del capítulo 4 que para este caso no es suficiente observar solamente el porcentaje de PMTs activados para determinar con una buena eficiencia la región del tanque por donde pasó el muón. En dicha simulación, las zonas definidas en esa propuesta no coincide la trayectoria del muón. En cambio para las zonas definidas en la segunda propuesta si coinciden con las partes en el tanque por donde se enviaron dichas partículas, y la eficiencia en los resultados para este caso es mayor que para las zonas definidas en la primera propuesta. Por lo tanto si se consideran las zonas en el tanque en donde al menos se activan 3 ó 2 detectores, se puede mejorar la reconstrucción hasta en un 19.11 % para el ángulo azimutal y un 63.03 % para el ángulo de elevación.

En el segundo análisis, a través de las variables C_a , D_c y P_a que se usaron para entrenar y probar a la red neuronal, se llegó a determinar los valores mostrados en la tabla 5.9, que nos dan información sobre la energía de trazas de muones individuales. Debido a la forma en la que se definieron los valores de C_a , D_c y P_a , este algoritmo podría implementarse en el observatorio HAWC con el propósito de detectar trazas de muones reales.

Por último diremos que estos dos análisis son complementarios, ya que si tenemos una idea de la energía del muón que generó la señal, se podrán implementar los volúmenes que se definieron en el capítulo 4 para mejorar la reconstrucción de su dirección. Por ejemplo, supongamos que un muón registra una energía menor a 9 TeV. Entonces podemos utilizar los resultados de los diagramas definidos en el capítulo 4 para una energía correspondiente a 10 TeV. Esto tomando en cuenta que el volumen efectivo de detección aumenta si la energía también aumenta. Así que si consideramos las regiones definidas con una energía de 10 TeV se cubre a las regiones que se puedan definir para energías menores a 9 TeV (que fue la energía que se uso para este ejemplo).

En un futuro se espera hacer más pruebas con simulaciones y datos reales, para observar los resultados que se obtienen con los estudios realizados en este proyecto.

A) Número de muones usados en la simulación

Anteriormente se había comentado de forma breve los motivos utilizados para explicar el número de partículas que se emplean en cada simulación. Una de las razones tenía que ver con el tiempo en que se tardaba la computadora en terminar cada proceso. Ya que entre más partículas utilizadas, más tiempo se requería para la simulación. El otro motivo era que las variables utilizadas en este trabajo tiende a no cambiar tanto si el número de partículas aumenta. Para ver esto se realizaron cuatro simulaciones con un número diferente de partículas: mil, cinco mil, diez mil y veinte mil. En estas simulaciones se enviaron muones (μ^-) de forma horizontal que pasaron a través del tanque E10 a una altura de 4 m sobre el PMT central. La dirección de propagación de estas partículas era hacia el PMT central del tanque X10 y llevaban una energía de 100 GeV. Es importante mencionar que estas simulaciones son prácticamente iguales a la que se presentaron en el capítulo 5, la única diferencia es la altura a la que se envían estas partículas. Así que el inciso a) de la figura 2.8 sirve muy bien para representar a estas simulaciones.

Número de Partículas	C_a		D_c		P_a	
	Promedio	RMS	Promedio	RMS	Promedio	RMS
1000	3.067	0.8145	6.252	5.819	2.972	0.2683
5000	3.166	0.8282	6.213	5.287	3.001	0.2665
10000	3.182	0.8878	6.34	5.636	2.999	0.264
20000	3.177	0.9141	6.326	5.982	2.999	0.2653

Tabla 5.11: Comparación de las variables del capítulo 5 para diferente número de partículas simuladas.

Número de partículas	Porcentaje en que se activan			
	4 PMTs	3 PMTs	2 PMTs	1 PMT
1000	0.78 %	1.83 %	26.73 %	70.66 %
5000	0.83 %	2.37 %	22.30 %	74.50 %
10000	0.69 %	2.49 %	23.98 %	72.85 %
20000	0.78 %	2.38 %	23.59 %	73.24 %

Tabla 5.12: Comparación de las variables del capítulo 4 para diferente número de partículas simuladas.

La tabla 5.11, muestra tanto el promedio como la desviación estándar de los histograma llenado con los datos C_a , D_c y P_a presentadas en el capítulo 5. Y la tabla 5.12 muestra el porcentaje de veces que se activaron 4, 3 ó 2 fotomultiplicadores en el tanque E10 (recordemos que esta información es la que se uso en el capítulo 4 para definir al volumen de detección).

Recordemos que los muones son partículas inestables que por lo general se suelen estudiar en grandes cantidades para conocer algunas de sus propiedades (por ejemplo su tiempo de vida media). Por este motivo se consideran una gran cantidad de estas partículas en las diferentes variables utilizadas en este trabajo, ya que en cierta forma estas dependen del número de detectores activados antes de que el muón decaiga. Como se puede observar en las tablas anteriores, las variables que se usan en ambos capítulos no presentan gran diferencia al incrementar el número de partículas simuladas. Por este motivo se utilizo la cantidad de dos mil partículas (para el capítulo 5) y diez partículas (para el capítulo 4) en las pruebas que se realizaron en este proyecto, ya que en un tiempo menor se obtiene resultados parecidos al caso en el que se simulan una mayor cantidad de muones. Todo esto considerado que en cada prueba se utilizó el mismo número de partículas, por ejemplo en el capítulo 5 todas las simulaciones se llevaron a cabo con 2 mil partículas simuladas.

B) Diferencia entre la carga del PMT central y uno periférico

Las variables C_a y D_c presentadas en el capítulo 5, se definieron al considerar que la carga registrada por el PMT central es aproximadamente el doble en comparación a la carga que registra un PMT periférico. Esto se debe a que el detector central tiene mayor superficie y eficiencia cuántica que los otros fotomultiplicadores. Por este motivo, en esta sección se presenta una simulación para verificar este hecho. Para ello se enviaron 10,000 muones verticales de 100 GeV a una altura de 2.1 m. Esta altura se escogió para que la apertura del cono de luz solamente le diera al PMT central y a uno periférico. En la figura 5.15 se presenta la posición inicial por donde se enviaron a estas partículas.

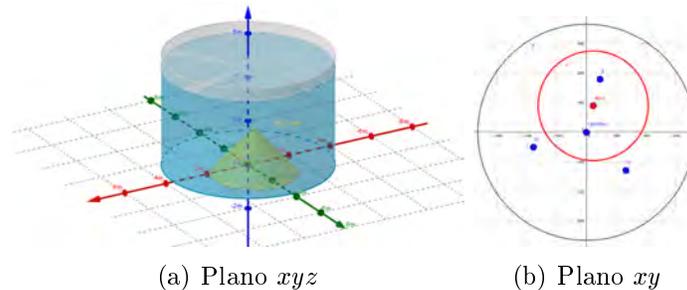


Figura 5.15: Representación de la posición inicial de la partícula simulada (punto rojo).

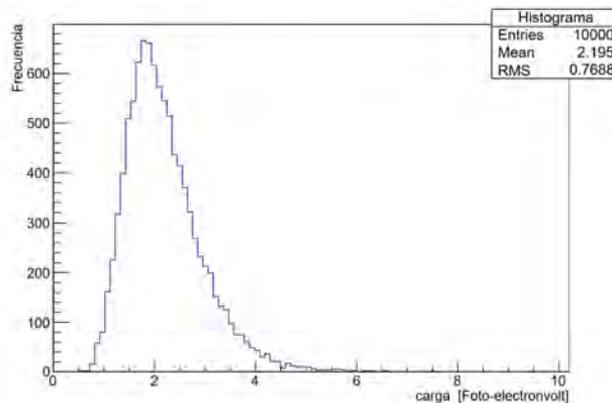


Figura 5.16: Histograma que muestra la proporción entre la carga que detecta el PMT central y uno periférico.

Después se lleno un histograma en donde se divide la carga del PMT central entre la carga del PMT periférico. Si efectivamente la carga que detecta el PMT central es aproximadamente el doble con respecto a uno periférico los datos del histogramas se llenaran con valores cercanos a 2. El resultado se puede ver en la figura 5.16.

Con esto se puede observar que la carga que registra el PMT central es aproximadamente el doble en comparación a los otros fotomultiplicadores.

C) Código para una red neuronal en ROOT

En esta parte se muestra el código para el ejemplo de una red neuronal mostrado en el capítulo 3. Que sirve como complemento para los interesados en empezar en el tema de programación de redes neuronales.

```
1
2 // PARTE I: Definición del árbol binario
3 class TDataPreparation {
4
5 private:
6     TTree* fDataTree;
7
8 public:
9
10    TDataPreparation() {
11
12        fDataTree = new TTree( "datatree", "datatree");
13        Int_t nlines = fDataTree->ReadFile("entrenamiento.dat","x:y:z"); // Aquí
14        se especifica el archivo de donde se obtendrá la información
15    }
16
17    TTree* getDataTree() { return fDataTree; }
18 };
19
20 //=====
21
22 void Ejemplo_NN() {
23     // Declaramos a nuestro árbol binario
24     TDataPreparation myFeedingData;
25     TTree* t = myFeedingData.getDataTree();
26 // PARTE II: Definición de la clase TMultiLayerPerceptron
27     TMultiLayerPerceptron *mlp =
28     // Declaramos las variables de entradas y salidas, además del número de capas
29     // ocultas en la red
30     new TMultiLayerPerceptron( "@x,@y:10:@z", t);
31
32     /*
33     | Tendremos dos entradas x, y
34     | Una capa oculta con 10 neuronas.
35     | Solo se tendrá una salida, z=f(x,y).
36     | el símbolo "@" indica que normalizamos la lista de datos con su desviación
37     | estandar.
38     */
39 }
40 //=====
```

```

38 // PARTE III: Metodo de entrenamiento
39 // Se especifica el metodo que utiliza la red para trabajar.
40 mlp->SetLearningMethod( TMultiLayerPerceptron::kStochastic );
41 /*
42 Se puede elegir entre los siguientes metodos:
43     kStochastic // default
44     kBatch
45     kSteepestDescent
46     kRibierePolak
47     kFletcherReeves
48     kBFGS
49 */
50
51 mlp->Train( 5000, "text ,update=100" );
52 /*
53 | Se especifica el numero de iteraciones de la red. En este caso se uso 5000.
54 | Estas iteraciones son los pasos que se van dando para llegar al minimo de la
55 | funcion de costo.
56 */
57 //=====
58 // PARTE III: Probar la red neuronal.
59
60 // Colocamos los datos para probar la red neuronal, como coordenadas de
61 // vectores.
62 Int_t nEvent = 2000; // numero de datos del archivo de prueba.
63 Double_t* x = new Double_t [ nEvent ];
64 Double_t* y = new Double_t [ nEvent ];
65 Double_t* z = new Double_t [ nEvent ];
66 Double_t* z_cal = new Double_t [ nEvent ];
67 Double_t* z_norm = new Double_t [ nEvent ];
68 ifstream readIn;
69 readIn.open( "prueba.dat", ios::in ); // Aqui se indica el archivo de los datos
70 // para probar la red.
71 char buffer[200];
72 readIn.getline( buffer, 200 );
73 readIn.getline( buffer, 200 );
74 for( int i=0; i<nEvent; i++) {
75     readIn >> x[i] >> y[i] >> z[i];
76 }
77 readIn.close();
78
79 // Calculamos el promedio y la desviacion estandar de los datos de entrenamiento
80 // Para que los datos de salida de la red, y los datos que se usaron para probar
81 // la red esten en la misma escala
82
83 Double_t suma=0;
84 Double_t promedio;
85 Float_t ze;
86 t->SetBranchAddresses("z",&ze);
87 for( Int_t i=0; i<nEvent; i++){
88     t->GetEntry(i);
89     suma = ze+suma;
90 }
91 promedio = suma/nEvent;
92 Double_t contador=0;
93 Double_t desviacion;
94 for( Int_t i=0; i<nEvent; i++){

```

```

91     t->GetEntry(i);
92     contador = ((ze-promedio)*(ze-promedio)) + contador;
93 }
94 desviacion = sqrt(contador/nEvent);
95
96 // Calculamos los datos de salida de la red.
97 Double_t inputs[2];
98 Double_t costo=0;
99 for( Int_t i=0; i<nEvent; i++ ){
100
101     inputs[0] = x[i];
102     inputs[1] = y[i];
103     z_cal[i] = mlp->Evaluate( 0, inputs ); // dato de salida de la red neuronal
104     z_norm[i] = (z_cal[i]*desviacion)+promedio ; // Se vuelve a escalar el dato
105     de salida de la red
106     costo = (z[i] - z_norm[i])*(z[i] - z_norm[i]); // Calculamos la funcion de
107     costo
108     ofstream ficheroSalida;
109     ficheroSalida.open("datos.txt", ios::app);
110     ficheroSalida << z_norm[i]<<endl; // Ponemos estos datos en un archivo de
111     texto
112     ficheroSalida.close();
113 }
114 cout<< costo/1000 <<endl; // imprimimos en pantalla el valor de la funcion de costo
115 promedio
116 mlp->Draw();
117 // Cambiamos el nombre del archivo con los datos de salida.
118 time_t now = time(0);
119 char* dt = ctime(&now);
120 rename("datos.txt", dt);
121 }

```

Ahora mostraremos el código utilizado para la red neuronal que se empleo en el capítulo 5 para estimar la energía de muones horizontales.

```

1 // PARTE I: Definicion del arbol binario
2 class TDataPreparation {
3
4 private:
5     TTree* fDataTree;
6
7 public:
8
9     TDataPreparation() {
10
11         fDataTree = new TTree( "datatree", "datatree");
12         Int_t nlines = fDataTree->ReadFile("entrenamiento.dat","ca:pm:dc:en"); //
13         Aqui se especifica el archivo de donde se obtendra la informacion
14
15     }
16
17     TTree* getDataTree() { return fDataTree; }
18 };
19

```

```

20 //=====
21
22 void Energia() {
23     // Declaramos a nuestro arbol binario
24     TDataPreparation myFeedingData;
25     TTree* t = myFeedingData.getDataTree();
26 // PARTE II: Definicion de la clase TMultiLayerPerceptron
27     TMultiLayerPerceptron *mlp =
28     // Declaramos las variables de entradas y salidas , ademas del numero de capas
29     // ocultas en la red
30     new TMultiLayerPerceptron( "@ca,@pm:dc:10:10:@en", t);
31
32     /*
33     |   Tendremos tres entradas ca,pm y dc
34     |   Dos capas ocultas con 10 neuronas.
35     |   Solo se tendra una salida , En.
36     */
37 //=====
38 // PARTE III: Metodo de entrenamiento
39 // Se especifica el metodo que utiliza la red para trabajar.
40 mlp->SetLearningMethod( TMultiLayerPerceptron::kStochastic );
41
42 mlp->Train( 5000, "text ,update=100");
43 //=====
44 // PARTE III: Probar la red neuronal.
45
46 // Colocamos los datos para probar la red neuronal , como coordenadas de
47 // vectores.
48 Int_t nEvent = 11000;
49 Double_t* ca = new Double_t [ nEvent ];
50 Double_t* pm = new Double_t [ nEvent ];
51 Double_t* dc = new Double_t [ nEvent ];
52 Double_t* en = new Double_t [ nEvent ];
53 Double_t* en_cal = new Double_t [ nEvent ];
54 Double_t* en_norm = new Double_t [ nEvent ];
55 ifstream readIn;
56 readIn.open( "prueba.dat", ios::in );// Aqui se indica el archivo de los datos
57 // para probar la red.
58 char buffer[200];
59 readIn.getline( buffer , 200 );
60 readIn.getline( buffer , 200 );
61 for( int i=0; i<nEvent; i++) {
62     readIn >> ca[i] >> pm[i] >> dc[i] >> en[i];
63 }
64 readIn.close();
65
66 // Calocamos el promedio y la desviacion estandar de los datos de entrenamiento
67 //=====
68 Double_t promedio= 5.0090909091;
69 Double_t desviacion= 3.1480021212;
70 //=====
71 // Calculamos los datos de salida de la red.
72 Double_t inputs [3];
73 int dim=0;
74 for( Int_t i=0; i<nEvent; i++ ){

```

```
74     inputs[0] = ca[i];
75     inputs[1] = pm[i];
76     inputs[2] = dc[i];
77     en_cal[i] = mlp->Evaluate( 0, inputs ); /// look here.
78     en_norm[i] = ((en_cal[i])*(desviacion))+(promedio);
79
80 // Colocamos el resultado que arroja la red neuronal en un archivo de texto
81     ofstream ficheroSalida;
82     ficheroSalida.open("en_norm.txt", ios::app);
83     ficheroSalida << en_norm[i] <<endl;
84     ficheroSalida.close();
85 }
86 }
```

Como se puede ver en este código, su estructura tomo como base al primer programa presentado en esta sección.

D) Método de propagación hacia atrás (backpropagation)

El objetivo de este procedimiento es encontrar las coordenadas del vector $-\nabla C$ (gradiente negativo de la función de costo) cuando solamente se ha presentado una imagen de entrenamiento. Para llevar a cabo esta tarea, se tiene que observar que tan sensible es la función de costo a pequeños cambios que se hagan a los valores de pesos y *bias*. Con esta información, se podrá hacer modificaciones a estos valores, y minimizar la función de costo de la forma más adecuada. Para observar como funciona este procedimiento, nos fijaremos en el caso más sencillo, en donde solamente hay una neurona en cada capa (ver figura 5.17).

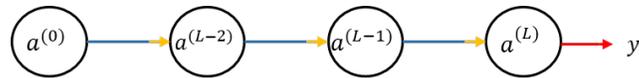


Figura 5.17: Ejemplo de una red neuronal que solamente tiene una neurona en cada capa. Imagen modificada de [44].

Pondremos nuestra atención en los dos últimos valores de la imagen anterior. La letra y hace referencia a la cantidad que el usuario espera obtener, y el símbolo a^L es el valor de activación de la neurona en la última capa y que representa el resultado que arroja la red. En este caso, la función de costo para una sola imagen de entrenamiento será:

$$C_0 = (a^L - y)^2. \quad (5.2)$$

En donde la activación de la última capa se podrá expresar como:

$$a^L = \sigma(a^{L-1}w^L + b^L). \quad (5.3)$$

Para hacer más simple la notación, definimos el siguiente valor:

$$z^L = a^{L-1}w^L + b^L \quad (5.4)$$

Se puede ver en estas tres ecuaciones, que la función de costo depende de la activación de la neurona en la última capa, y a su vez esta activación depende de z . La relación de dependencia que hay en cada uno de estas cantidades se ve representada en el diagrama de la figura 5.19.

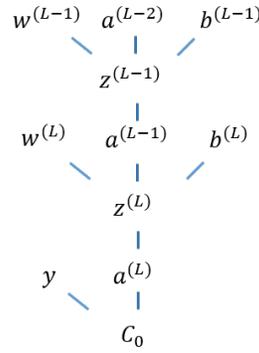


Figura 5.18: Relación entre las cantidades que definen a la función de costo. Imagen modificada de [44].

Teniendo en cuenta lo anterior, la idea será calcular que tan sensible es la función de costo a pequeños cambios del peso $w^{(L)}$, en otras palabras, calcularemos la derivada parcial de la función de costo con respecto al peso $w^{(L)}$. Esta cantidad, se puede expresar mediante el uso de la regla de la cadena, y tendrá el siguiente valor:

$$\frac{\partial C_0}{\partial w^{(L)}} = \frac{\partial z^{(L)}}{\partial w^{(L)}} \frac{\partial a^{(L)}}{\partial z^{(L)}} \frac{\partial C_0}{\partial a^{(L)}}. \quad (5.5)$$

Teniendo en cuenta a las ecuaciones 5.2, 5.3 y 5.4, podemos calcular las siguientes derivadas parciales:

$$\frac{\partial z^{(L)}}{\partial w^{(L)}} = a^{(L-1)}, \quad \frac{\partial a^{(L)}}{\partial z^{(L)}} = \sigma'(z^{(L)}), \quad \frac{\partial C_0}{\partial a^{(L)}} = 2(a^{(L)} - y).$$

Entonces

$$\frac{\partial C_0}{\partial w^{(L)}} = a^{(L-1)} \sigma'(z^{(L)}) 2(a^{(L)} - y). \quad (5.6)$$

De forma similar podemos calcular que tan sensible es la función de costo, para un pequeño cambio en el valor de $b^{(L)}$:

$$\frac{\partial C_0}{\partial b^{(L)}} = \frac{\partial z^{(L)}}{\partial b^{(L)}} \frac{\partial a^{(L)}}{\partial z^{(L)}} \frac{\partial C_0}{\partial a^{(L)}} = \sigma'(z^{(L)}) 2(a^{(L)} - y). \quad (5.7)$$

Recordemos que los valores de $\frac{\partial C_0}{\partial w^{(L)}}$ y $\frac{\partial C_0}{\partial b^{(L)}}$ solamente se obtuvieron al considerar un dato de entrenamiento (y). En el caso que toma en cuenta a todos los ejemplos de entrenamiento, estas cantidades se pueden representar por medio de un promedio:

$$\begin{aligned} \frac{\partial C}{\partial w^{(L)}} &= \frac{1}{n} \sum_{i=0}^{n-1} \frac{\partial C_i}{\partial w^{(L)}}. \\ \frac{\partial C}{\partial b^{(L)}} &= \frac{1}{n} \sum_{i=0}^{n-1} \frac{\partial C_i}{\partial b^{(L)}}. \end{aligned}$$

Estas dos ultimas cantidades, solamente representan dos coordenadas del vector $-\nabla C$.

$$-\nabla C = \begin{bmatrix} \frac{\partial C}{\partial w^0} \\ \frac{\partial C}{\partial b^0} \\ \vdots \\ \frac{\partial C}{\partial w^L} \\ \frac{\partial C}{\partial b^L} \end{bmatrix}.$$

Para encontrar el valor de las otras coordenadas de este vector, tendremos que fijarnos tanto en los valores de pesos y *bias* de una capa previa. Por ejemplo, ahora nos fijaremos en la derivada parcial de la función de costo con respecto al peso $w^{(L-1)}$ que se puede expresar como:

$$\frac{\partial C_0}{\partial w^{(L-1)}} = \frac{\partial z^{(L-1)}}{\partial w^{(L-1)}} \frac{\partial a^{(L-1)}}{\partial z^{(L-1)}} \frac{\partial z^{(L)}}{\partial a^{(L-1)}} \frac{\partial a^{(L)}}{\partial z^{(L)}} \frac{\partial C_0}{\partial a^{(L)}}.$$

En donde:

$$\frac{\partial C_0}{\partial a^{(L-1)}} = \frac{\partial z^{(L)}}{\partial a^{(L-1)}} \frac{\partial a^{(L)}}{\partial z^{(L)}} \frac{\partial C_0}{\partial a^{(L)}} = w^{(L)} \sigma'(z^{(L)}) 2(a^{(L)} - y)$$

Al fijarnos que tan sensible es la función de costo al valor de activación de una capa previa ($L - 1$), como capaces de calcular el valor de $\frac{\partial C_0}{\partial w^{(L-1)}}$, de forma similar se puede calcular el valor de $\frac{\partial C_0}{\partial b^{(L-1)}}$. Por lo tanto, si vamos repitiendo este procedimiento de manera iterativa, podemos calcular los cambios que debemos realizar a todos los pesos y *bias* en cada capa de la red para encontrar el mínimo de la función de costo.

Las ecuaciones que vimos anteriormente, se pueden generalizar para el caso de una red neuronal que tenga más de una neurona en cada una de sus capas. Esto se hará al agregar un subíndice que señala a la neurona con la que se este trabajando en cada capa (ver figura 5.19).

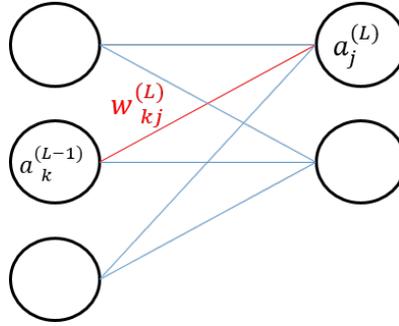


Figura 5.19: Notación para una red neuronal. Imagen modificada de [44].

Esta notación ya se había revisado anteriormente, y si consideramos estos cambios, tanto la función de costo, como la activación de las neuronas de la capa j , se podrán expresar mediante las siguientes ecuaciones:

$$C_0 = \sum_{j=0}^{N_L-1} (a_j^{(L)} - y_j)^2, \quad a_j = \sigma(z_j^{(L)}).$$

Entonces la derivada parcial de la función de costo con respecto al peso w_{jk} será:

$$\frac{\partial C_0}{\partial w_{jk}^{(L)}} = \frac{\partial z_j^{(L)}}{\partial w_{jk}^{(L)}} \frac{\partial a_j^{(L)}}{\partial z_j^{(L)}} \frac{\partial C_0}{\partial a_j^{(L)}}.$$

Ahora solo nos falta ver como cambia la derivada de la función de costo para las activaciones de una capa previa (en nuestro caso es la capa k).

$$\frac{\partial C_0}{\partial a_k^{(L-1)}} = \sum_{j=0}^{N_L-1} \frac{\partial z_j^{(L)}}{\partial a_k^{(L-1)}} \frac{\partial a_j^{(L)}}{\partial z_j^{(L)}} \frac{\partial C_0}{\partial a_j^{(L)}}.$$

En este caso se toma en cuenta a todas las neuronas de la capa, por este motivo se agrega el símbolo de suma a esta expresión.

Tomando en cuenta a todas estas expresiones, y al igual que en el ejemplo anterior (en donde se tenía una red sencilla con una sola neurona en cada capa) , se pueden ir calculando los cambios que se desean hacer a todos los valores de pesos y *bias* que se encuentran en cada una de las capas de la red. De esta manera es como funciona el método de propagación hacia atrás.

E) Volumen efectivo de detección μ^+ y consideraciones del porcentaje P_H

Volumen efectivo de detección para μ^+

El volumen mostrado en el capítulo 4 se definió a partir de simulaciones de muones negativos. En esta sección se muestra que estas regiones son de utilidad para reconstruir la dirección en el caso de muones positivos, ya que este volumen no presenta ninguna diferencia al utilizar partículas positivas para su construcción. Para ver esto, se calculó el volumen efectivo de detección utilizando muones positivos de 1 TeV.

Altura (cm)	1 TeV	
	μ^-	μ^+
	Dirección H.I 4 PMTs %	Dirección H.I 4 PMTs %
0	5.66	2.42
37.5	0.97	1.23
75	3.61	3.65
112.5	40.00	39.56
150	77.81	78.13
187.5	86.25	86.55
225	17.78	17.70
262.5	12.47	12.12
300	9.35	9.53
337.5	7.27	7.69
375	6.14	6.25
412.5	5.58	5.23
450	4.45	5.30

Tabla 5.13: Porcentaje de veces que se activan 4 PMTs para muones positivos y negativos

En la tabla 5.13 se muestra una comparación entre los resultados obtenidos para el caso de muones horizontales simulados con diferente carga. Esta se puede apreciar que el porcentaje de veces que se activan 4 detectores es prácticamente el mismo para ambos casos. Utilizando esta información junto con las tablas del apéndice F, realizamos los diagramas correspondientes al volumen efectivo de detección para este caso y que se pueden observar en las imágenes 5.20 y 5.21.

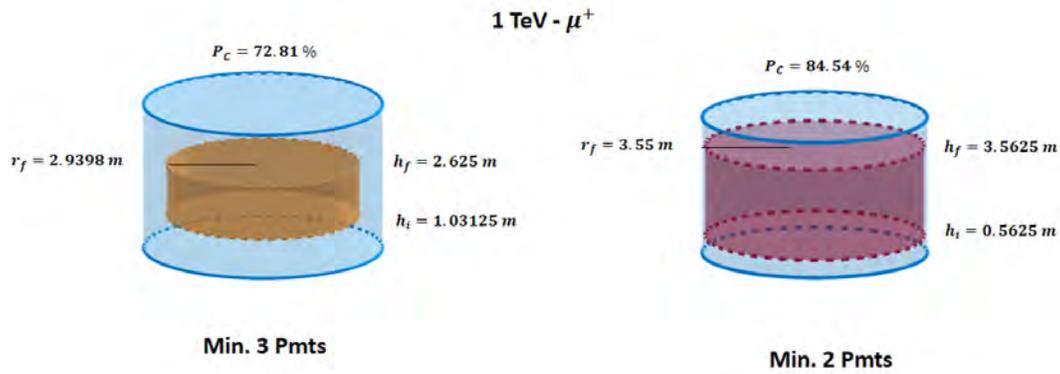


Figura 5.20: Región del tanque en donde es más probable que se activen al menos 3 ó 2 PMTs para μ^+

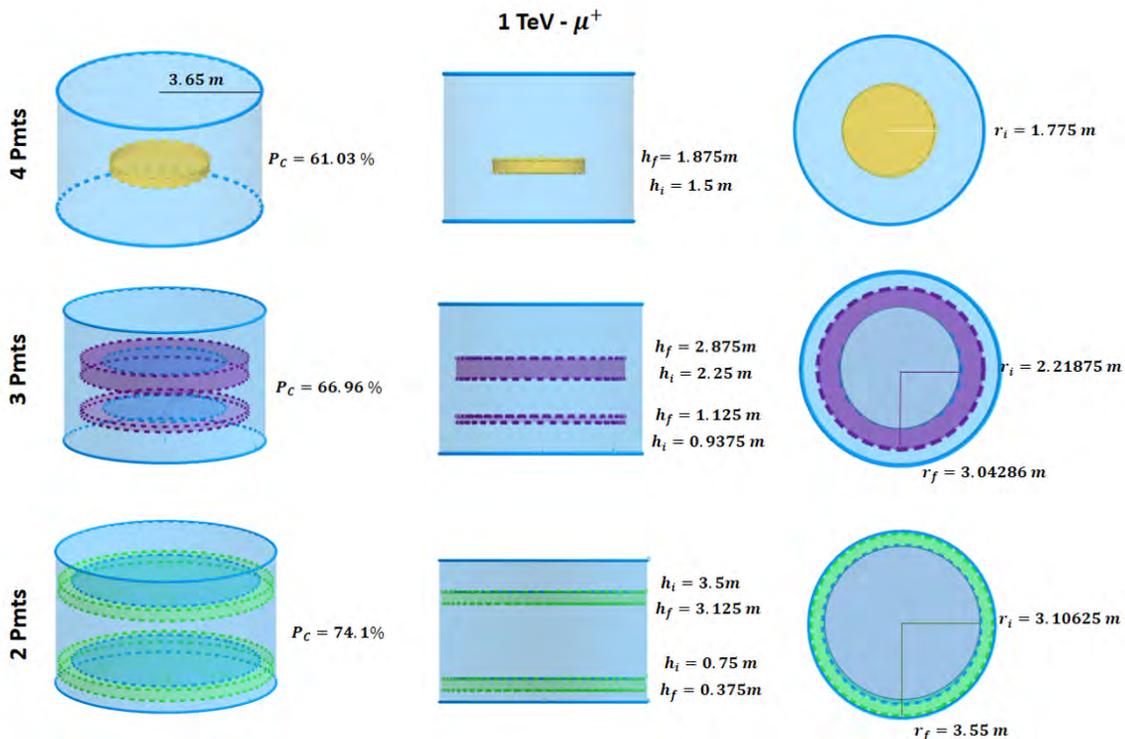


Figura 5.21: Región del tanque en donde es más probable que se activen 4,3 ó 2 PMTs para μ^+

Las dimensiones de estos diagramas son exactamente iguales a las que se presentaron en las figuras del capítulo 4. Para los porcentajes que se manejan en ambos casos, hay una diferencia menor al 2% (que se atribuye al número de partículas simuladas). Así que para ambas partículas (μ^- y μ^+) se obtienen las mismas regiones. Esto se debe a que la luz Cherenkov generada en este proceso es la misma para ambos casos. Como se comentó en el capítulo 1, las partículas del agua se pueden modelar como dipolos, que son perturbados por el campo eléctrico de la partícula incidente. Cuando una partícula de carga negativa pasa por el medio, atrae la parte positiva del dipolo, y de manera opuesta cuando una partícula de carga positiva pasa por el medio atrae la parte negativa del dipolo. En ambos casos se genera una onda de choque (si la partícula se propaga más rápido que la luz en el medio) que genera la luz registrada por el detector. Por este motivo se registra la misma cantidad de luz en ambos casos.

Tomando lo anterior en cuenta, los volúmenes del capítulo 4 se pueden utilizar tanto para el caso de muones negativos como para el caso de muones positivos.

Consideraciones para el porcentaje P_H

En la tabla 5.14 se muestran en términos de porcentaje el número de veces que se activan 3

1 TeV		
Altura (cm)	Dirección H.II 3 PMTs %	Porcentaje por región (P_R)
0	4.24	10.43
37.5	0.76	
75	47.47	
93.75	85.98	25.82
103.125	79.12	
112.5	67.09	
150	28.97	9.93
187.5	20.95	
225	69.41	34.81
243.75	70.05	
262.5	70.62	
300	55.23	19.01
337.5	13.58	
356.25	11.70	
375	9.74	
412.5	7.14	
450	4.05	

Tabla 5.14: Porcentaje de 3 PMTs activados para muones en la dirección H.II con 1 TeV de energía.

PMTs en el caso de muones horizontales enviados en la dirección H.II. Para esta dirección se pueden definir 2 regiones en el tanque en donde es mas probable que se activen 3 fotomultiplicadores; estas zonas se resaltan con color amarillo en la tabla anterior. Utilizando la ecuación 4.2 calculamos el porcentaje P_H para el caso del cilindro inferior:

$$P_H = \min \{ \min \{ 85.98, 67.09 \}, 25.82 \} = 25.82.$$

Si ahora consideramos una altura inicial de 75 cm para este cilindro; el valor de P_H será:

$$P_H = \min \{ \min \{ 47.47, 67.09 \}, 25.82 \} = 25.82.$$

Como se puede observar en ambos casos el valor de P_H es el mismo, por lo tanto se puede escoger entre estas dos opciones para definir la altura inicial y final de este cilindro. Para eliminar este problema y evitar considera porcentajes menores al 50 %. El porcentaje P_H se define de la siguiente manera para el caso en donde es mas probable que se activen exactamente 3 ó 2 fotomultiplicadores:

- Paso 1: Se calcula el porcentaje P_R que toma en cuenta a todas las direcciones mediante siguiente ecuación:

$$P_R = \sum_{j=1}^n P_R^{D.j} P_{Dj}.$$

En donde n es el número de direcciones que registraron una misma región en una misma zona del tanque, Dj hace referencia a las direcciones en donde se encontró a estas regiones y P_{Dj} es la probabilidad de que la partícula llegue por esa dirección.

- Paso 2: Se calcula el valor del mínimo entre P_{hi} y P_{hf} que toma en cuenta a todas las direcciones mediante siguiente ecuación:

$$\min \{P_{hi}, P_{hf}\} = \sum_{j=1}^n \min \{P_{hi}, P_{hf}\}^{D.j} P_{Dj}$$

- Paso 3: Sumamos los porcentaje P_R del cilindro superior e inferior, para tomar en cuenta a toda la región del tanque en donde es mas probable que se activen 3 PMTs. A este valor lo llamaremos porcentaje combinado P_B .
- Paso 4: Encontramos el valor del mínimo entre P_{hi} y P_{hf} que toma en cuenta a todas las direcciones para el caso del cilindro superior.
- Paso 5: Se define el porcentaje P_H como:

$$P_H = \min \left\{ P_B, \min \{P_{hi}, P_{hf}\}^{Sup.}, \min \{P_{hi}, P_{hf}\}^{Inf.} \right\}$$

Ejemplo: Nos fijamos en el cilindro inferior que define la región en donde es mas probable que se activen 3 detectores. Esta zona solo se encontró en la dirección HII, por lo tanto:

$$P_R = P_R^{HII} P_{HII} = 25.82.$$

$$\min \{P_{hi}, P_{hf}\} = \min \{67.09, 85.98\} = 67.09.$$

Los porcentajes para el cilindro superior son los siguientes:

$$\min \{P_{hi}, P_{hf}\} = 67.67.$$

$$P_R = 44.69.$$

Estos datos ya incluyen a las otras direcciones. Entonces:

$$P_B = 44.69 + 25.82 = 70.51.$$

Por lo tanto:

$$P_H = \min \{70.51, 67.67, 67.09\} = 67.09$$

Con esta definición el cilindro superior tendrá el mismo porcentaje P_H que el cilindro inferior. Además se soluciona el problema mostrado en el inicio de esta sección. Ya que para el caso en donde h_i es igual a 75 cm se tiene un valor del porcentaje P_H igual a 47.47%. Así que el punto inicial y final para estas regiones queda bien definido tomando estas consideraciones.

F) Tablas de intervalos de energías y volumen efectivo de Detección.

Intervalos de energía

En esta parte se muestra los intervalos de error para la energía calculada con la ecuación 5.1, en donde se toma al valor de la desviación estándar (asociado a los datos de prueba para la energía de 5 TeV) para calcular dichos intervalos.

Energía de los datos de prueba utilizados (TeV)	Energía promedio calculada por la red E_{NN} (TeV)	RMS	$E_{NN} - RMS$	$E_{NN} + RMS$
11	6.8129	1.548	5.2649	8.3609
12	7.0137	1.548	5.4657	8.5617
13	7.2041	1.548	5.6561	8.7521
14	7.3854	1.548	5.8374	8.9334
15	7.5587	1.548	6.0107	9.1067
16	7.7245	1.548	6.1765	9.2725
17	7.8835	1.548	6.3355	9.4315
18	8.0360	1.548	6.4880	9.5840
19	8.1823	1.548	6.6343	9.7303
20	8.3228	1.548	6.7748	9.8708
21	8.4577	1.548	6.9097	10.0057
22	8.5872	1.548	7.0392	10.1352
23	8.7116	1.548	7.1636	10.2596
24	8.8310	1.548	7.2830	10.3790
25	8.9458	1.548	7.3978	10.4938
26	9.0560	1.548	7.5080	10.6040
27	9.1618	1.548	7.6138	10.7098
28	9.2635	1.548	7.7155	10.8115
29	9.3611	1.548	7.8131	10.9091
30	9.4549	1.548	7.9069	11.0029
31	9.5450	1.548	7.9970	11.0930
32	9.6315	1.548	8.0835	11.1795
33	9.7147	1.548	8.1667	11.2627
34	9.7945	1.548	8.2465	11.3425

Tabla 5.15: Intervalos de incertidumbre para la energía calculada con la ecuación 5.1

Volumen efectivo de detección Se muestran los datos de los porcentajes que definen las regiones mostradas en el capítulo 4.

Dirección HI - 1 TeV					
Altura (cm)	4 pmt	3 pmt	2 pmt	1 pmt	suma
0	5.66	4.72	7.55	82.08	100
37.5	0.97	2.77	96.26	0.00	100
75	3.61	8.58	87.81	0.00	100
112.5	40.00	40.68	19.32	0.00	100
150	77.81	20.08	2.11	0.00	100
187.5	86.25	13.03	0.72	0.00	100
225	17.78	70.47	11.08	0.67	100
262.5	12.47	72.47	13.96	1.10	100
300	9.35	69.99	18.75	1.91	100
337.5	7.27	64.20	25.04	3.49	100
375	6.14	35.28	39.02	19.56	100
412.5	5.58	18.16	30.26	46.00	100
450	4.45	9.48	16.86	69.22	100

Dirección HII - 1 TeV					
Altura (cm)	4 pmt	3 pmt	2 pmt	1 pmt	suma
0	4.24	4.24	13.56	77.97	100
37.5	0.51	0.76	4.18	94.55	100
75	3.30	47.47	40.27	8.96	100
93.75	10.10	85.98	3.89	0.03	100
103.125	19.39	79.12	1.49	0.00	100
112.5	32.37	67.09	0.54	0.00	100
150	70.85	28.97	0.18	0.00	100
187.5	77.81	20.95	1.24	0.00	100
225	16.33	69.41	14.04	0.22	100
243.75	14.17	70.05	15.53	0.25	100
262.5	11.59	70.62	17.36	0.43	100
300	8.53	55.23	34.19	2.05	100
337.5	4.73	13.58	74.30	7.39	100
356.25	4.91	11.70	73.56	9.84	100
375	4.50	9.74	41.84	43.93	100
412.5	3.53	7.14	32.87	56.46	100
450	1.97	4.05	26.93	67.06	100

Dirección HIII - 1 TeV					
Altura (cm)	4 pmt	3 pmt	2 pmt	1 pmt	suma
0	2.68	2.68	10.07	84.56	100
37.5	0.48	1.19	98.33	0.00	100
75	1.67	4.84	93.49	0.00	100
112.5	30.47	43.51	26.02	0.00	100
150	69.94	26.46	3.60	0.00	100
168.75	75.87	22.24	1.89	0.00	100
187.5	78.05	20.20	1.75	0.00	100
206.25	74.36	23.10	2.54	0.00	100
225	21.93	38.83	39.18	0.06	100
262.5	6.62	14.92	77.74	0.72	100
300	5.72	12.95	78.92	2.41	100
337.5	4.66	10.79	79.59	4.96	100
375	4.12	7.58	44.41	43.90	100
412.5	3.55	6.56	33.93	55.96	100
450	1.57	3.81	26.63	68.00	100

Figura 5.22: Datos de Porcentajes de μ^- horizontales de 1 TeV

Dirección VI - 1 TeV					
Radio (cm)	4 pmts	3 pmts	2 pmts	1 pmts	suma
%					
44.375	97.9	2.07	0.03	0	100
88.75	93.7	6.2	0.1	0	100
133.125	86.14	13.22	0.64	0	100
177.5	74.15	23.53	2.32	0	100
221.875	57.59	35.5	6.91	0	100
266.25	1.79	4.3	93.86	0.05	100
310.625	0.96	2.35	96.44	0.25	100
355	0.52	1.49	96.74	1.25	100

Dirección VII - 1 TeV					
Radio (cm)	4 pmts	3 pmts	2 pmts	1 pmts	suma
%					
44.375	97.62	2.37	0.01	0	100
88.75	93.85	6.01	0.14	0	100
133.125	85.83	13.68	0.49	0	100
177.5	74.63	23.46	1.91	0	100
221.875	24.76	62.6	12.64	0	100
232.96875	5.3	77.22	17.48	0	100
255.15625	2.95	72.8	24.24	0.01	100
266.25	2.45	68.43	29.12	0	100
310.625	0.86	3.38	95.53	0.23	100
355	0.41	2.11	96.18	1.3	100

Dirección VIII - 1 TeV					
Radio (cm)	4 pmts	3 pmts	2 pmts	1 pmts	suma
%					
88.75	93.37	6.56	0.07	0.00	100
133.125	85.63	14.15	0.22	0.00	100
177.5	73.13	25.72	1.15	0.00	100
221.875	4.22	86.80	8.98	0.00	100
266.25	1.92	81.89	16.18	0.01	100
310.625	0.97	43.43	55.12	0.48	100
332.8125	0.80	11.99	86.42	0.79	100
355	0.47	3.89	93.77	1.87	100

Dirección IV - 1 TeV					
Radio (cm)	4 pmts	3 pmts	2 pmts	1 pmts	suma
%					
44.375	97.79	2.21	0.00	0.00	100
88.75	92.83	7.15	0.02	0.00	100
133.125	85.02	14.90	0.08	0.00	100
177.5	71.58	27.98	0.44	0.00	100
221.875	3.38	93.08	3.54	0.00	100
266.25	1.81	90.27	7.88	0.04	100
310.625	0.97	83.33	15.40	0.30	100
355	0.64	71.50	26.32	1.54	100

Dirección V - 1 TeV					
Radio (cm)	4 pmts	3 pmts	2 pmts	1 pmts	suma
%					
44.375	97.92	2.07	0.01	0.00	100
88.75	93.29	6.71	0.00	0.00	100
133.125	85.43	14.53	0.04	0.00	100
177.5	69.69	30.10	0.21	0.00	100
221.875	3.17	95.23	1.60	0.00	100
266.25	1.47	93.70	4.75	0.08	100
310.625	0.94	87.69	10.92	0.45	100
355	0.65	77.41	20.12	1.82	100

Figura 5.23: Datos de Porcentajes de μ^- verticales de 1 TeV.

Dirección HIII - 10 TeV					
Altura	4 pmt	3 pmt	2 pmt	1 pmt	suma
(cm)	%				
0	4.76	7.34	17.86	70.04	100
37.5	4.82	5.56	89.62	0.00	100
75	10.41	10.69	78.90	0.00	100
112.5	47.26	34.98	17.76	0.00	100
150	79.68	18.14	2.18	0.00	100
168.75	83.81	14.73	1.46	0.00	100
187.5	85.35	13.71	0.94	0.00	100
206.25	82.27	16.32	1.40	0.01	100
225	42.12	32.74	25.11	0.03	100
262.5	25.94	18.57	55.00	0.49	100
300	23.74	17.29	57.61	1.36	100
337.5	22.58	16.01	58.46	2.95	100
375	20.93	14.17	39.28	25.62	100
412.5	20.24	12.53	33.00	34.22	100
450	13.84	10.02	30.96	45.18	100

Dirección HIII - 10 TeV					
Altura	4 pmt	3 pmt	2 pmt	1 pmt	suma
(cm)	%				
0	5.96	6.88	17.20	69.95	100
37.5	5.05	4.56	12.21	78.18	100
75	16.10	50.19	29.18	4.53	100
93.75	29.53	68.33	2.14	0.00	100
103.125	39.44	59.76	0.79	0.01	100
112.5	50.15	49.40	0.45	0.00	100
150	80.57	19.36	0.07	0.00	100
187.5	85.60	13.66	0.74	0.00	100
225	35.44	55.76	8.71	0.09	100
243.75	33.40	57.44	9.04	0.12	100
262.5	31.15	58.03	10.55	0.27	100
300	27.32	50.27	21.54	0.87	100
337.5	22.59	19.90	53.49	4.02	100
356.25	22.27	18.16	54.12	5.45	100
375	21.28	16.71	36.16	25.85	100
412.5	20.13	14.85	31.62	33.40	100
450	0.00	0.00	0.00	0.00	0

Dirección HI - 10 TeV					
Altura	4 pmt	3 pmt	2 pmt	1 pmt	suma
(cm)	%				
0	6.7114094	6.48769575	17.2259508	69.5749441	100
37.5	9.77	8.37	81.86	0.00	100
75	19.31	14.13	66.56	0.00	100
112.5	60.96	27.90	11.14	0.00	100
150	87.12	11.83	1.05	0.00	100
187.5	91.71	7.94	0.35	0.00	100
225	37.06	56.23	6.42	0.29	100
262.5	29.80	60.83	8.73	0.64	100
300	27.25	60.34	11.47	0.94	100
337.5	26.22	57.07	15.06	1.65	100
375	24.83	39.13	24.83	11.20	100
412.5	25.01	25.51	23.99	25.49	100
450	23.06	17.76	19.10	40.08	100

Figura 5.24: Datos de Porcentajes de μ^- horizontales de 10 TeV.

Dirección VI - 10 TeV					
Radio (cm)	4 pmts	3 pmts	2 pmts	1 pmts	suma
	%				
44.375	98.58	1.4	0.02	0	100
88.75	95.52	4.38	0.1	0	100
133.125	89.37	10.17	0.46	0	100
177.5	79.94	18.33	1.73	0	100
221.875	64.97	30.29	4.74	0	100
266.25	7.97	8.75	83.24	0.04	100
310.625	5.73	6.18	87.92	0.17	100
355	3.84	4.93	90.25	0.98	100

Dirección VII - 10 TeV					
Radio (cm)	4 pmts	3 pmts	2 pmts	1 pmts	suma
	%				
44.375	98.41	1.56	0.03	0	100
88.75	95.42	4.53	0.05	0	100
133.125	89.27	10.36	0.37	0	100
177.5	79.78	18.69	1.53	0	100
221.875	35.01	55.26	9.73	0	100
232.96875	14.5	70.96	14.54	0	100
255.15625	9.86	70.22	19.91	0.01	100
266.25	8.8	67.43	23.76	0.01	100
310.625	5.39	7.04	87.45	0.12	100
355	3.87	5.84	89.43	0.86	100

Dirección VIII - 10 TeV					
Radio (cm)	4 pmts	3 pmts	2 pmts	1 pmts	suma
	%				
44.375	98.55	1.45	0	0	100
88.75	95.49	4.44	0.07	0	100
133.125	88.9	10.9	0.2	0	100
177.5	78.2	20.97	0.83	0	100
221.875	11.76	81.43	6.81	0	100
266.25	8.09	79.67	12.23	0.01	100
310.625	4.9	47.71	47.12	0.27	100
332.8125	4.49	18.59	76.35	0.57	100
355	3.8	8.41	86.13	1.66	100

Dirección IV - 10 TeV					
Radio (cm)	4 pmts	3 pmts	2 pmts	1 pmts	suma
	%				
44.375	98.25	1.75	0.00	0.00	100
88.75	95.10	4.90	0.00	0.00	100
133.125	88.41	11.53	0.06	0.00	100
177.5	76.11	23.55	0.34	0.00	100
221.875	10.30	86.86	2.84	0.00	100
266.25	6.99	87.11	5.90	0.00	100
310.625	5.13	82.86	11.82	0.19	100
355	3.79	73.95	21.29	0.96	100

Dirección V - 10 TeV					
Radio (cm)	4 pmts	3 pmts	2 pmts	1 pmts	suma
	%				
44.375	98.29	1.71	0.00	0.00	100
88.75	94.93	5.06	0.01	0.00	100
133.125	88.24	11.76	0.00	0.00	100
177.5	73.91	26.00	0.09	0.00	100
221.875	9.94	88.92	1.14	0.00	100
266.25	6.58	89.80	3.53	0.09	100
310.625	4.63	86.70	8.30	0.37	100
355	3.90	78.56	16.05	1.49	100

Figura 5.25: Datos de Porcentajes de μ^- verticales de 10 TeV.

Dirección HI - 1 TeV					
Altura (cm)	4 pmt	3 pmt	2 pmt	1 pmt	suma
	%				
0	2.42	2.42	9.68	85.48	100
37.5	1.23	2.59	96.18	0.00	100
75	3.65	8.64	87.71	0.00	100
112.5	39.56	40.78	19.66	0.00	100
150	78.13	19.68	2.19	0.00	100
187.5	86.55	12.75	0.70	0.00	100
225	17.70	70.56	10.99	0.75	100
262.5	12.12	72.52	14.48	0.88	100
300	9.53	70.29	18.25	1.93	100
337.5	7.69	64.02	24.93	3.36	100
375	6.25	34.42	39.20	20.12	100
412.5	5.23	17.95	31.86	44.97	100
450	5.30	9.12	19.32	66.26	100

Dirección HII - 1 TeV					
Altura (cm)	4 pmt	3 pmt	2 pmt	1 pmt	suma
	%				
0	0.98	0.98	13.73	84.31	100
37.5	0.48	0.78	4.47	94.27	100
75	3.18	48.29	39.91	8.62	100
93.75	9.73	86.19	4.02	0.06	100
103.125	20.00	78.57	1.42	0.01	100
112.5	32.26	66.96	0.76	0.02	100
150	70.46	29.34	0.20	0.00	100
187.5	77.71	21.01	1.26	0.02	100
225	16.32	69.19	14.21	0.28	100
243.75	13.72	70.49	15.40	0.39	100
262.5	12.29	69.82	17.43	0.46	100
300	8.06	56.44	33.34	2.16	100
337.5	5.30	14.05	73.42	7.22	100
356.25	4.64	11.55	73.67	10.14	100
375	4.69	8.83	41.89	44.59	100
412.5	3.82	7.70	32.87	55.60	100
450	1.85	4.71	26.76	66.68	100

Dirección HIII - 1 TeV					
Altura (cm)	4 pmt	3 pmt	2 pmt	1 pmt	suma
	%				
0	5.00	2.50	10.00	82.50	100
37.5	0.44	1.44	98.12	0.00	100
75	1.77	4.49	93.74	0.00	100
112.5	29.74	43.83	26.43	0.00	100
150	69.46	27.17	3.37	0.00	100
168.75	75.61	22.72	1.67	0.00	100
187.5	78.73	19.52	1.75	0.00	100
225	21.32	40.06	38.50	0.12	100
262.5	6.75	14.64	77.85	0.76	100
300	5.57	13.31	78.99	2.13	100
337.5	4.10	11.17	79.81	4.92	100
375	4.24	7.63	45.46	42.66	100
412.5	3.78	6.39	34.09	55.75	100
450	1.82	3.43	26.10	68.66	100

Figura 5.26: Datos de Porcentajes de μ^+ horizontales de 1 TeV.

Dirección VI - 1 TeV					
Radio	4 pmts	3 pmts	2 pmts	1 pmts	suma
(cm)	%				
44.375	97.87	2.12	0.01	0	100
88.75	93.98	5.89	0.13	0	100
133.125	85.96	13.26	0.78	0	100
177.5	74.21	23.64	2.15	0	100
221.875	57.59	35.66	6.75	0	100
266.25	1.83	4.69	93.47	0.01	100
310.625	0.84	2.61	96.3	0.25	100
355	0.53	1.61	96.76	1.1	100

Dirección VII - 1 TeV					
Radio	4 pmts	3 pmts	2 pmts	1 pmts	suma
(cm)	%				
44.375	97.77	2.21	0.02	0.00	100
88.75	93.70	6.20	0.10	0.00	100
133.125	85.75	13.58	0.67	0.00	100
177.5	74.22	23.83	1.95	0.00	100
221.875	23.95	62.93	13.12	0.00	100
232.96875	5.58	76.43	17.99	0.00	100
255.15625	2.89	73.36	23.74	0.01	100
266.25	2.76	68.34	28.88	0.02	100
310.625	0.78	3.28	95.70	0.24	100
355	0.59	2.00	96.10	1.31	100

Dirección VIII - 1 TeV					
Radio	4 pmts	3 pmts	2 pmts	1 pmts	suma
(cm)	%				
44.375	97.84	2.14	0.02	0.00	100
88.75	93.93	6.05	0.02	0.00	100
133.125	85.81	13.79	0.40	0.00	100
177.5	72.75	25.95	1.30	0.00	100
221.875	3.99	87.65	8.36	0.00	100
266.25	1.97	82.29	15.74	0.00	100
310.625	0.81	42.76	55.98	0.45	100
332.8125	0.71	12.30	85.97	1.02	100
355	0.51	3.27	94.14	2.08	100

Dirección IV - 1 TeV					
Radio	4 pmts	3 pmts	2 pmts	1 pmts	suma
(cm)	%				
44.375	97.47	2.53	0.00	0.00	100
88.75	94.05	5.95	0.00	0.00	100
133.125	85.43	14.50	0.07	0.00	100
177.5	70.58	28.94	0.48	0.00	100
221.875	3.03	93.24	3.72	0.01	100
266.25	1.44	90.43	8.12	0.01	100
310.625	0.99	83.83	14.92	0.26	100
355	0.53	70.66	27.32	1.49	100

Dirección V - 1 TeV					
Radio	4 pmts	3 pmts	2 pmts	1 pmts	suma
(cm)	%				
44.375	97.58	2.42	0.00	0.00	100
88.75	92.96	7.03	0.01	0.00	100
133.125	84.97	15.01	0.02	0.00	100
177.5	68.92	30.95	0.13	0.00	100
221.875	3.27	94.99	1.74	0.00	100
266.25	1.67	93.07	5.16	0.10	100
310.625	0.99	87.07	11.35	0.59	100
355	0.61	76.89	20.68	1.82	100

Figura 5.27: Datos de Porcentajes de μ^+ verticales de 1 TeV.

Bibliografía

- [1] https://private.hawc-observatory.org/wiki/index.php/Install_HAWC_software. Página interna de HAWC. Recuperado el 28/11/2017.
- [2] ROSALES, G. (2017). Caracterización de la modulación azimutal del flujo de cascadas atmosféricas con el observatorio HAWC. (Tesis de licenciatura). UNAM, México.pp.44-85.
- [3] LEÓN V. HERMES ET AL. *Capability of the HAWC Gamma-Ray Observatory for the Indirect Detection of Ultrahigh-Energy Neutrinos*. En: Hindawi volumen 2017, pp.1-13.
- [4] AGUILAR, M. B. (2011). *Cien años de Rayos Cósmicos*. Rev. R. Acad. Cienc. Exact. Fís.Nat. (Esp), Vol. 105, N°. 1, pp 107-128.
- [5] ALVARO, D. (2016). *Detectar rayos cósmicos con agua resolverá el mayor misterio del Universo*. Recuperado el 26-10-2017, de Omicrono Sitio web: <http://omicrono.espanol.com/2016/05/detectar-rayos-cosmicos-con-agua-resolvera-el-mayor-misterio-del-universo/>
- [6] KESSLER, C. (2012). Aluminum-Leaf Electroscope. En *A Project Guide to Electricity and Magnetism* (p.13). USA: Mitchell Lane.
- [7] DIAZ, J. (2015). *Mensajeros cósmicos I: rayos cósmicos*. Recuperado el 20-10-2017, de Conexión Casual Sitio web: <https://conexioncausal.wordpress.com/2015/12/12/mensajeros-cosmicos-i-rayos-cosmicos/>
- [8] MOULD, L. (1993). Discovery of Radioactivity and Radium. En *A century of X-Ray and Radioactivity in medicine*. (p.10). Gran Bretaña: Institute of Physics,London Publishing.
- [9] DORMAN, L. & DORMAN, I. (2014). History of Cosmic Ray Research in Mexico. En *Cosmic Ray History*. (pp.662-665).New York: Nova.
- [10] <https://www.hawc-observatory.org/>. Recuperado el 22/10/2017.
- [11] GONZÁLEZ, MARÍA MAGDALENA. *HAWC: un nuevo observatorio de rayos gamma en México*. Revista Digital Universitaria [en línea]. 10 de octubre de 2009, Vol. 10, No. 10 [Consultada: 26 de octubre de 2017]. Disponible en Internet: <<http://www.revista.unam.mx/vol.10/num10/art66/int66.htm>>
- [12] GUADALUPE, R. (2015). *Se inauguró el observatorio de rayos gamma HAWC* . Recuperado el 26-10-2017,de INAOE Sitio web: <http://www.inaoep.mx/noticias/?noticia=276&anio=2015>
- [13] HART, R. (2009). *Electromagnetic Cascade*. Recuperado el 24-10-2017, de Hardhack Sitio web: http://hardhack.org.au/electromagnetic_cascades

- [14] MINAYA, I. (2011). Composición de la radiación cósmica de ultra-alta energía a partir de los datos del Observatorio Pierre Auger. (Tesis de Maestría). Facultad de Ciencias Físicas de la Universidad Complutense de Madrid, España.
- [15] GRIEDER, P. (2010). Introduction, Facts and Phenomenology. En *Extensive Air Shower*. (pp.3-24). Berlín: Springer.
- [16] DAVIES, B. (2016). *Cosmic Rays*. Recuperado el 26-10-2017, de Antarctic Glaciers Sitio web: <http://www.antarcticglaciers.org/glacial-geology/dating-glacial-sediments-2/cosmic-rays/>
- [17] JELLEY, J. (1958). Introduction, En *Cerenkov Radiation and its applications*. (pp.3-6). Gran Bretaña: United Kingdom Atomic Energy Authority.
- [18] https://es.wikipedia.org/wiki/Observatorio_de_Rayos_Gamma_HAWC. Recuperado el 22/10/2017.
- [19] AHRON S. BARBER ET AL. *Simulation of Near Horizontal Muons and Muon Bundles for the HAWC Observatory with CORSIKA*. En: Proceeding of science ICRC (Julio 2017), pp. 5-6.
- [20] <https://svn.auger.unam.mx/trac/projects/ape/wiki/ApePreRequirements>. Recuperado el 28/11/2017.
- [21] S. AGOSTINELLI ET AL. GEANT4 - a simulation toolkit. NIM A 506 (2003) pp.250-303.
- [22] RAMOS, M. *Geant4 tutorial*. Recuperado el 29-11-2017, de FisMedFCFM Sitio <https://fismedfcfm.wikispaces.com/Geant4+tutorial>
- [23] <http://geant4.web.cern.ch/geant4/> . Recuperado el 30/11/2017.
- [24] MEZA, G. *Bibliotecas o librerías en C++*. Recuperado el 2-12-2017, de Programa Ya. Sitio <https://www.programarya.com/Cursos/C++/Bibliotecas-o-Librerias>
- [25] THE GEANT4 COLLABORATION. *Geant4 User's Guide for Application Developers*. 2017
- [26] GALINDO, T.. (2017). Calibración e instalación de un detector Cherenkov de agua del sitio lago Sierra Negra. (Tesis de Doctorado). INAOE, México.pp.62-74.
- [27] HERNÁNDEZ, M. (2017). Caracterización de la interacción de rayos cósmicos con volcanes por medio de una simulación Monte Carlo con GEANT4. (Tesis de licenciatura). UNAM, México. pp. 26-28.
- [28] AGUILAR, T. (2015). Simulación de Cascadas Aéreas Extensas para el proyecto LAGO, Guatemala. (Tesis de licenciatura). Universidad de San Carlos de Guatemala, México.pp.13-20.
- [29] D. HECK Y T. PIEROG. *Extensive Air Shower Simulations with CORSIKA: A User's Guide*. 2017
- [30] MESSEL, H Y CRAWFORD, D. (1970). *Electron-photon Shower Distribution Function: Tables for Lead, Copper, and Air Absorbers*. (pp. 3-7). Londres: Pergamon Press.
- [31] MATTHEWS, J. *A Heitler model of extensive air showers*. En: Science direct. Astroparticle Physics 22 (2005) pp.387-397.

- [32] L. RÄDEL Y C. WIEBUSCH.. *Calculation of the Cherenkov light yield from electromagnetic cascades in ice with Geant4*. En: Astroparticle Physics 44 (Apr. 2013), pp. 102-113.
- [33] WEISHI,S ET AL. *Spallation Backgrounds in Super-Kamiokande Are Made in Muon-Induced Showers*.En:Phys.Rev. D91 (2015) no.10, p.4.
- [34] CAPISTRÁN, R. (2014). Aplicación de redes neuronales para la separación de gammas y hadrones en HAWC. (Tesis de maestría). INAOE, México.pp.22-34.
- [35] MICHAEL A. NIELSEN, *Neuronal Networks and Deep Learning*, Determination Press, 2015.
- [36] LARA, F. *Fundamento de redes neuronales artificiales*, Centro de instrumentos,UNAM.
- [37] PALMER P. ET AL. *¿Qué son las redes neuronales artificiales? Aplicaciones realizadas en el ámbito de las adicciones*. En: Adicciones 1999 Vol. 11 Num. 3, pp.243-255.
- [38] Recuperado el 05-01-2018, de Electronica Sitio web: <http://electronica.com.mx/neural/informacion/perce>
- [39] Recuperado el 07-01-2018,de Github Sitio web: https://ml4a.github.io/ml4a/neural_networks/
- [40] HEATON, J. *The Number of Hidden Layers* . Recuperado el 07-01-2018,de Heaton Research Sitio <http://www.heatonresearch.com/2017/06/01/hidden-layers.html>
- [41] SANDERSON, G. [3Blue1Brown]. (5 de octubre 2017). But what *is* a Neural Network? | Chapter 1, deep learning. [Archivo de video]. Recuperado de <https://www.youtube.com/watch?v=aircAruvnKk&t=465s>
- [42] SANDERSON, G. [3Blue1Brown]. (16 de octubre 2017). Gradient descent, how neural networks learn | Chapter 2, deep learning. [Archivo de video]. Recuperado de <https://www.youtube.com/watch?v=IHZwWFHWa-w&t=262s>
- [43] SANDERSON, G. [3Blue1Brown]. (3 de noviembre 2017). What is backpropagation really doing? | Chapter 3, deep learning. [Archivo de video]. Recuperado de <https://www.youtube.com/watch?v=llg3gGewQ5U>
- [44] SANDERSON, G. [3Blue1Brown]. (3 de octubre 2017). Backpropagation calculus | Appendix to deep learning chapter 3. [Archivo de video]. Recuperado de <https://www.youtube.com/watch?v=tIeHLnjs5U8&t=511s>
- [45] [Dai xination]. (25 de julio 2017). CERN ROOT tutorial for beginners – topic: neural network by TMultiLayerPerceptron and TMLPAnalyzer. [Archivo de video]. Recuperado de <https://www.youtube.com/watch?v=WhzduU9ZXgo&t=1100s>
- [46] <https://root.cern.ch/>. Recuperado el 10/01/2018.
- [47] DELAERE, C. *TMultiLayerPerceptron: Designing and using Multi-Layer Perceptrons with ROOT*. Recuperado el 07-01-2018, de UCL Sitio: <http://cp3.irmp.ucl.ac.be/delaere/MLP/>
- [48] LEON,H. *Finding Muon Tracks in (10 seconds of) Raw Data*. Documento interno de HAWC, 2016.