

Validación del linaje de los datos de la colaboración LAGO. Instalaciones Sierra Negra y Chacaltaya

Data Provenance Validation for the LAGO stations Sierra Negra and Chacaltaya

Luis A. Núñez

PhD. en Física, Profesor Emérito,
Investigador Grupo GIRG, Escuela de Física, Facultad de Ciencias,
Universidad Industrial de Santander UIS, Bucaramanga, Colombia,
y Departamento de Física, Facultad de Ciencias,
Universidad de los Andes, Mérida, Venezuela, e Investigador de
la Colaboración LAGO.
lnunez@uis.edu.co

Fernando Quiñónez

PhD. en Física, Docente Tiempo Completo,
Investigador Grupo SIGMMA, Facultad de Ingeniería Industrial,
Universidad Santo Tomás USTA, Bucaramanga, Colombia, e
Investigador de la Colaboración LAGO, Universidad Industrial de
Santander, e Investigador de la Colaboración ATLAS,
Pontificia Universidad Católica de Chile.
fernando.andres.quinonez.granados@cern.ch

Christian Andrés Sarmiento Cano

Físico, Universidad Industrial de Santander UIS,
Estudiante de Maestría, Investigador Grupo GIRG,
Escuela de Física, Facultad de Ciencias, Universidad Industrial de
Santander UIS, Bucaramanga, Colombia, e
Investigador de la Colaboración LAGO.
christian.sarmiento@correo.uis.edu.co

Resumen— Este trabajo muestra los resultados obtenidos luego de validar la calidad (linaje) de los datos registrados para la búsqueda de explosiones de rayos gamma en el marco de la colaboración LAGO. Para este propósito se analizaron las marcas de GPS en los datos recolectados por los detectores ubicados en las estaciones de Sierra Negra (México) y Chacaltaya (Bolivia). Durante esta validación se encontraron inconsistencias entre la frecuencia de detección esperada, y la frecuencia de detección registrada por el sistema de adquisición de los datos LAGO durante los años 2006 hasta el 2009 en Sierra Negra, y desde el 2008 hasta el 2011 en Chacaltaya. Luego se propone una solución para poder realizar el análisis de estos datos recolectados, se rescatan las partes que presentan la frecuencia de detección óptima, catalogándolos en listas de tipo ROOT-C++ que se han llamado “listas de buenos datos recolectados (GRDL)”. Se da una receta de uso de las GRDL para que los futuros investigadores puedan hacer análisis sobre los datos LAGO buenos. Finalmente, hacemos una búsqueda de explosiones de rayos gamma sobre los datos LAGO buenos, utilizando el método de los excesos.

Palabras clave— LAGO: Large Aperture Gamma-Ray-Burst Observatory. GPS: Global Positioning System. GRDL: Good Recollected Data Lists. GRB: Gamma Ray Bursts.

Abstract— Some results on the search of Gamma Ray Burst from the LAGO Collaboration are presented. These are obtained after validating the quality

(provenance) of the data recorded at two LAGO installation (Sierra Negra (Mexico) and Chacaltaya Bolivia). We analyze the GPS marks at data collected. During this validation process we have found inconsistencies between the expected detection rate and the one recorded by the acquisition system. This was found at Sierranegra in 2008 and at Chacaltaya from 2006 to 2009. Then, we propose an strategy to perform the analysis of the data collected in order to recover those parts with optimum detection frequency. Next, we label them in ROOT-C++ format as Good Run Data List (GRDL). In order to perform a more correct data analysis at LAGO repositories, a GRDL recipe is provided. Finally, do a search for gamma-ray bursts on the validated data, by using the method of the excesses.

Keywords— LAGO: Large Aperture Gamma-Ray-Burst Observatory. GPS: Global Positioning System. GRDL: Good Recollected Data Lists. GRB: Gamma Ray-Bursts.

I. INTRODUCCIÓN

Las explosiones de rayos gamma (GRB), provenientes del espacio exterior (galáctico y/o extragaláctico) están entre los fenómenos más energéticos del universo. Estas GRB fueron descubiertas en los años 60 por satélites militares estadounidenses encargados de rastrear actividad de posibles ensayos nucleares de sus

países enemigos. Las GRB llegan a al planeta Tierra como haces de fotones muy energéticos ($E > 1$ keV) que ocurren pocas veces al día y en períodos muy cortos de tiempo (0,1s - 100 s). El modelo más aceptado para explicarlos, tiene su origen en la hipótesis de que son originados en explosiones estelares, ver [1],[2], y las referencias allí citadas.

Hasta el día de hoy, el mayor aporte al conocimiento de las GRB ha provenido de observaciones de telescopios satelitales (*HETE*, *INTEGRAL*, *Swift* y *GLAST*, ahora *Fermi Gamma-Ray Space Telescope*). Al aumentar la energía de los fotones su flujo disminuye como una ley de potencias y las superficies que ofrecen los detectores satelitales se hacen imprácticas. Por ello, el espectro en energía de esos destellos se conoce bien a bajas energías, pero se tiene muy poca información sobre las regiones de altas energías (> 1 GeV). Esto ha cambiado con la operación del Telescopio Espacial Fermi [3], pero aún así, los costos de las observaciones satelitales son aún elevados [4].

Para aumentar el área de registro se ha comenzado a instalar detectores terrestres. Estos observatorios complementan y refrendan la información obtenida por los satélites. Un ejemplo de ello puede verse en la observación simultánea de emisiones gamma en las galaxias M82 AND NGC 253 por el Telescopio Espacial Fermi y la colaboración HESS [5], [6], [7].

La llegada al detector de las cascadas de partículas de alta energía de eventos candidatos a GRB, provenientes desde dentro y fuera de nuestra galaxia, nos puede dar información sobre cómo está formado el universo mediante campos magnéticos galácticos, nubes gaseosas densas y agujeros negros, a los cuales se les conoce como aceleradores cósmicos.

Para resolver el misterio de los orígenes de los GRB, se debe tener una capacidad de detección de fotones y neutrinos en tierra para poder resolver el rompecabezas de si el fotón viene de un electrón dispersado o desde el decaimiento de un pión[8], [9]. Es muy difícil detectar neutrinos que detectar fotones, en ese sentido, se empieza a atacar el problema de la determinación de los orígenes de los GRB, desde su lado más factible, *i.e.* detectar fotones.

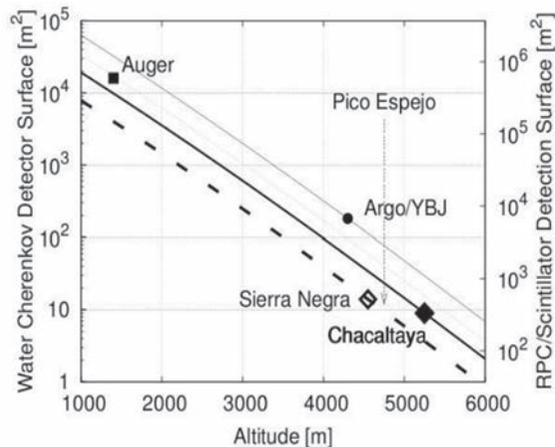
Luego se quiere contar el número de fotones que llegan a los foto-multiplicadores contenidos en cada Detector Cherenkov de Agua (WCD) a lo largo de la serie del tiempo.

En América hay dos macro iniciativas de observatorios gamma: el Observatorio Pierre Auger y el Proyecto HAWC (High-Altitude Water Cherenkov Observatory [10]). El Observatorio Pierre Auger[11], cuenta con 24 telescopios de fluorescencia UV y 1600 WCD que cubre una superficie de recolección de datos de 16000m². Sin embargo, su ubicación a 1400msnm limita su eficiencia para la detección de los GRB más energéticos a causa de la absorción atmosférica [12]. El proyecto HAWC tendrá 300 WCD en total con 4 PMT/WCD que cubrirán un área de 12556m², y estarán ubicados a una altura de 4100msnm, habrá un sistema de aislamiento óptico entre cada tanque para reducir la tasa de ruido entre tanques. Los tanques Cherenkov en HAWC son tanques de acero corrugado de 4m de alto y 7.3m de diámetro.

Otra iniciativa de medir GRB en América Latina es la colaboración LAGO (*Large Aperture Gamma-Ray-Burst Observatory*). LAGO reúne científicos de Venezuela, Colombia, Ecuador, Perú, Bolivia, México, Guatemala y Argentina y apunta a instalar detectores de GRB en cada uno de estos países. La colaboración LAGO inició la toma de datos hacia el año 2006 en Sierra Negra (SN) y en 2008 en Chacaltaya (CHA). Durante este tiempo se han recolectado aproximadamente 2 TB de datos que corresponden a 1600 días de recolección entre las dos estaciones. El buen funcionamiento de los instrumentos de recolección es lo que garantiza la veracidad de los resultados que se puedan obtener luego de analizar los datos. Por lo tanto, se hace necesario revisar minuciosamente los datos y así validar la calidad de su linaje (o proveniencia).

Para comprender la importancia del experimento LAGO en el estudio de fotones de alta energía es importante resaltar que un área efectiva de detección de 20m² que opera en Chacaltaya a 5200msnm, es suficiente para equiparar la sensibilidad instrumental de la superficie del experimento Pierre Auger[13] con 16000m², esto se puede observar en la Fig. 1.

Fig.1. RELACIÓN ENTRE LA ALTURA Y EL ÁREA NECESARIA PARA LA DETECCIÓN DE FOTONES DE ALTA ENERGÍA, COMO SE OBSERVA, 20M2 OPERANDO EN CHACALTAYA ES SUFICIENTE PARA EQUIPARAR LA SENSIBILIDAD INSTRUMENTAL DE LA SUPERFICIE DEL EXPERIMENTO PIERRE AUGER



Fuente: Colaboración LAGO

II. DATOS LAGO

LAGO empieza a adquirir datos mediante sus dos detectores más antiguos SN y CHA en 2006 y 2009 respectivamente. En los principios de la colaboración, cada quien empezaba a analizar los datos creyendo, de manera inocente, que ellos estaban bien tomados. En algunos casos las personas encargadas de la adquisición de los datos, eran las mismas encargadas de analizarlos, luego ellos sabían que en conjunto de datos estaban defectuosos y, por ende, no los consideraban en sus análisis, ó corregían partes de los datos en sus análisis sin socializarlo con el resto de científicos de la colaboración. No quedaba registro para que otros miembros de la colaboración estuvieran conscientes de la deficiencia de los datos en algunos lapsos de tiempo. Más tarde al crecer la colaboración, se empezaron a sumar científicos a LAGO, muchos novatos y otros más especializados en analizar datos que en tomar experimentalmente los mismos. Antes de empezar a hacer algoritmos sofisticados de búsqueda de GRB, se han planteado las preguntas: ¿Estarán bien tomados estos datos? ¿Vamos a considerar los errores sistemáticos conservadoramente? Y fue hasta ese entonces en que se empezó a dar cuenta de ciertas inconsistencias que presentaban los datos.

Los estudios realizados en [14] encontraron tres posibles candidatos a GRB en los datos tomados entre el año 2006 y 2009 en la estación

de SN. Sin embargo, dejan claro que hasta entonces no se ha podido verificar (por métodos estadísticos) una detección de fotones de alta energía en superficie terrestre provenientes de GRB. Además, proponen estudios detallados para confirmar o descartar dicho hallazgo.

A partir del 2010 comienza la instalación de la red de Repositorios LAGODATOS [15], [16], y [17]. Esta propuesta se basa en desarrollar una infraestructura repositorio de datos que utilicen *DSpace*. A este *software* de código abierto, se le han realizado adaptaciones con el fin de convertirlo en la herramienta capaz de preservar los datos de la colaboración. LAGODATOS funciona actualmente en un servidor principal en el nodo de la Universidad Industrial de Santander, el cual se encarga de catalogar y almacenar los datos adquiridos en cada nodo del sistema. Actualmente se tienen disponibles 14385 archivos de las instalaciones de Chacaltaya y Sierra Negra.

El acceso a los datos y la posibilidad de incorporar algoritmos inteligentes para hallar posibles GRB en los datos registrados, más allá de emprender su búsqueda a través la identificación de excesos respecto a la media, ha impulsado a la comunidad a preguntarse sobre la calidad de los datos medidos en estas dos instalaciones. Garantizar la calidad y vigilar el linaje de los datos es indispensable para un experimento que involucra el registro distribuido e independiente de un volumen significativo de datos[18].

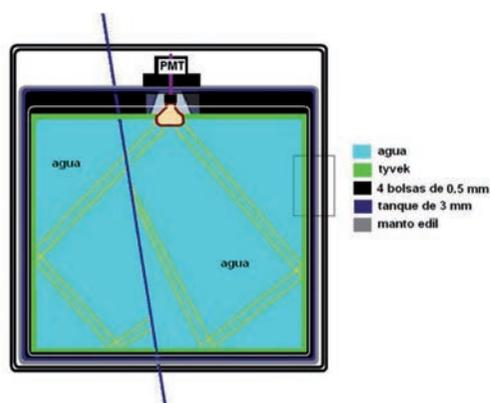
Este trabajo evalúa la calidad de los datos almacenados en el repositorio LAGODATOS e identifica cuáles segmentos de esos datos pueden ser analizados en la búsqueda de señales de GRB. El análisis iniciará por revisar la continuidad en la toma de datos por parte de las estaciones de Sierra Negra (SN) y Chacaltaya (CHA). Seguidamente, se seleccionará los archivos que cumplan con las condiciones mínimas para ser analizados y luego revisar el funcionamiento de la etiqueta de tiempo del GPS (*Global Position System*) que aparece en los datos tomados. Finalmente, se buscará excesos en la tasa de conteos de los datos mediante el método de los excesos, que será explicado en la Sección III.

A continuación se describe la forma como se detecta la radiación, el instrumento utilizado, y la digitalización de los datos adquiridos por estas estaciones.

A. Detector Cherenkov de Agua

Los detectores *Cherenkov* de agua, WCD, usados en el proyecto LAGO son tanques con una geometría cilíndrica con un foto-multiplicador (PMT) en el centro de la tapa superior. Este instrumento contiene varios recubrimientos. El más exterior es opaco y es el encargado de aislar la luz visible. El más interno es de un material conocido como *tyvek* el cual es altamente difusivo y reflectante; lo que permite a las partículas que ingresan al tanque propagarse y rebotar hasta llegar al PMT, ver Fig. 2.

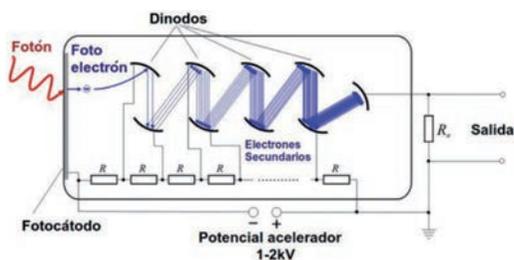
Fig.2. ESQUEMA DE UN DETECTOR *CHERENKOV* DE AGUA. EL TRAZO AZUL REPRESENTA UN RAYO CÓSMICO QUE LLEGA AL TANQUE Y LOS AMARILLOS SON LAS POSIBLES TRAYECTORIAS DE LOS FOTONES GENERADOS POR RADIACIÓN *CHERENKOV*



Fuente: Autores

El PMT es un dispositivo de alta sensibilidad que usa el efecto fotoeléctrico para generar un flujo de corriente, en función de un número de fotones incidentes. Consta de un fotocátodo donde se producen electrones que se aceleran por una serie de electrodos, o dinodos, conectados a una serie de diferencias de potencial en cascada, ver Fig. 3. Estos electrones, al llegar al último dinodo, denominado ánodo, se convierten en un pulso de tensión[19].

Fig.3. ESQUEMA DE AMPLIFICACIÓN DE LA SEÑAL ENTRANTE EN UN PMT

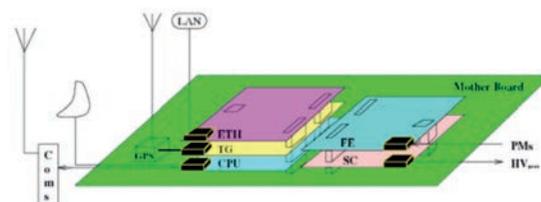


Fuente: Autores

B. Digitalización de la Señal

La entrada de una partícula al tanque produce radiación tipo *Cherenkov* que es captada por el PMT el cual está conectado a una placa electrónica llamada LS, (*Local Station*). Esta placa está compuesta por una tarjeta madre principal o *board* y cinco pequeñas tarjetas interconectadas entre sí, ver Fig. 4.

Fig.4. ESQUEMA DE LA ESTACIÓN LOCAL, LS, USADA POR LOS DETECTORES DE LAGO. LAS TARJETAS QUE COMPONEN EL DISPOSITIVO SON LAS SIGUIENTES: CONTROLADOR LOCAL (CPU), MARCADOR DE TIEMPO (TG), ETHERNET (ETH), CONEXIONES ELECTRÓNICAS (FG) Y CONTROL LENTO (SC)



Tomado de [20]

Los objetivos de este dispositivo son, por un lado controlar el PMT y, por el otro, adquirir, digitalizar, establecer el tiempo de exposición para la adquisición y enviar a un computador el conteo de señales producidas en el tanque. El proceso se realiza a una frecuencia de muestreo de 40Mhz y los histogramas son generados en función de la unidad de digitalización FADC, (*Flash Analog-to-Digital Converter*) de 10 bits [19].

Finalmente, la LS permite dos protocolos de adquisición *calib* y *DAQ*. El *calib* genera dos histogramas uno tipo pico y otro tipo carga, a partir de un mínimo de energía o umbral deseado. Además, el *DAQ* produce conteos para cuatro umbrales de energía diferentes y se le asigna una etiqueta de tiempo (año, mes, día, hora) a cada archivo de pulsos registrados. Este último protocolo es el usado por SN y CHA.

C. Caracterización de los Datos LAGO

Para iniciar con la caracterización de los datos recolectados por LAGO, es necesario enumerar algunas especificaciones técnicas de CHA y SN.

En el año 2006 se inició el montaje del arreglo de detectores en SN a 4650msnm, el cual contempló cuatro tanques de 4m² de área efectiva. Los detectores ubicados en SN tienen instalados PMT de 5 pulgadas modelo EMI 9030A. Cada tanque, independiente de la altura, está lleno solo

120cm para, de esta manera, asegurar el decaimiento del fotón en un par electrón - positrón en el tanque[21]. Los umbrales de energía usados para la generación de los datos dependen de las características del PMT, y en el caso de SN [14] son:

- 1) umbral 1: para partículas con energías $0 \leq E < 15$ MeV,
- 2) umbral 2: para partículas con energías $15 \leq E < 150$ MeV,
- 3) umbral 3: para partículas con energías $150 \leq E < 600$ MeV,
- 4) umbral 4: reservado para fallas de la electrónica.

Por otro lado, la estación Chacaltaya a 5200msnm cuenta con tres tanques detectores, dos de 4m² de área efectiva, y uno de 2m², cada uno con un PMT de 8 pulgadas tipo 9353 KB. En este caso los umbrales de energía tienen los siguientes valores:

- 1) umbral 1: para partículas con energías $0 \leq E < 5$ MeV,
- 2) umbral 2: para partículas con energías $5 \leq E < 10$ MeV,
- 3) umbral 3: para partículas con energías $10 \leq E < 50$ MeV,
- 4) umbral 4: reservado para fallas de la electrónica.

Un ejemplo de los datos recolectados en SN y generados por la LS se puede ver en la Fig. 5, donde se aprecian 25 columnas, 24 de ellas dedicadas a almacenar los conteos generados en los diferentes umbrales y la columna 25 destinada a la marca de tiempo o GPS (*Global Position System*). Es importante señalar que a la LS se le pueden conectar 3 PMT y cada uno de ellos en dos modos, alta y baja tensión. Además, cada conexión ocupa 4 columnas, que corresponden a los 4 umbrales de energía respectivos. En el caso específico de la estación de SN los 3 PMT están conectados por alta tensión, por lo cual en la Fig. 5 se observan 12 columnas con ceros.

Fig.5. ARCHIVO USUAL DE DATOS QUE RECOLECTA Y GENERA EN LA ESTACIÓN DE SN. COMO SE PUEDE APRECIAR EN ESTE EJEMPLO SE MUESTRAN 3 PMT CONECTADOS EN EL CANAL 1 (COLUMNAS 1 A 4), EN EL CANAL 2 (COLUMNAS 5 A 8) Y EN EL CANAL 5 (COLUMNAS 17 A 20). LAS COLUMNAS DONDE SE VEN CEROS, SON PORQUE ALLÍ NO HAY NADA CONECTADO

```
DAQ version 2.0
99 55 28 4 31 24 13 1 0 0 0 0 0 0 0 0 0 56 35 6 0 0 0 0 0 105052242
75 41 14 3 41 32 21 1 0 0 0 0 0 0 0 0 0 68 45 6 1 0 0 0 0 105052242
84 43 18 2 39 37 25 1 0 0 0 0 0 0 0 0 0 54 31 6 0 0 0 0 0 105052242
64 51 19 3 31 25 14 1 0 0 0 0 0 0 0 0 0 36 16 2 0 0 0 0 0 105052242
77 54 26 3 70 61 24 1 0 0 0 0 0 0 0 0 0 65 33 7 0 0 0 0 0 105052242
112 70 5 4 33 30 14 1 0 0 0 0 0 0 0 0 0 39 26 5 0 0 0 0 0 105052242
```

Por otra parte, aunque los datos recolectados en la estación de CHA son digitalizados por la misma LS las conexiones a ésta se distribuyeron de otra forma. En este caso, se encuentran conectados a la LS 3 PMT por alta tensión y uno de estos por baja tensión. Adicionalmente, se encuentra conectado por una de las entradas de baja tensión un detector de centelleo, la última entrada se encuentra desconectada por presencia de fallas. Finalmente, en estos datos se genera una línea cada 5m, es decir, 200 líneas/s, lo que da 720000 líneas/h.

III. VALIDACIÓN DEL LINAJE DE LOS DATOS Y BÚSQUEDA DE GRB

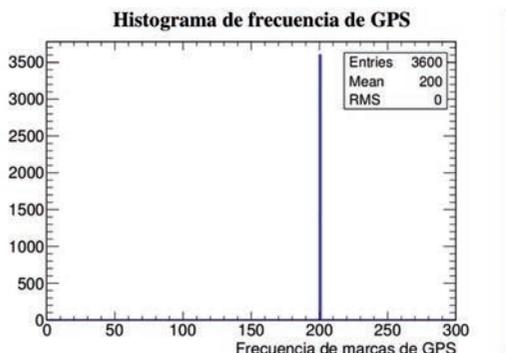
Aquí se muestra el análisis realizado sobre los datos descritos en la sección II. Se inicia con la validación de los datos, posteriormente se describe el algoritmo usado para la búsqueda de GRB en los datos de la colaboración y finalmente se examina otra metodología que se podría usar en análisis futuros.

Los fallos en la digitalización de los conteos se deben a la LS y son generados principalmente por perturbaciones electromagnéticas externas. Típicamente tormentas eléctricas, las cuales se observan como pulsos invertidos u oscilaciones en el voltaje.

Como se observó, los datos tienen una columna destinada al tiempo la cual está dada por una marca de GPS que cambia cada segundo, como se había dicho antes, se toman datos cada 5 ms, entonces habrá 200 líneas donde debe aparecer la misma marca de GPS (una marcas/s). Luego de adquirir datos durante una hora se encuentran 3600 marcas/h, ver sección II. Por lo tanto, un archivo ideal debe contener 720000 líneas y tener 3600 marcas diferentes de GPS. Una forma de verificar que la digitalización de los datos se ha realizado correctamente es revisar este com-

portamiento. Un histograma de frecuencias de un archivo de comportamiento óptimo, es decir, con 720000 líneas, donde se genera una línea cada 5 ms, con 3600 marcas de GPS diferentes, que cambian cada 200 líneas, se puede ver en la Fig.6.

Fig.6. HISTOGRAMA DE FRECUENCIAS DE LAS MARCAS DE GPS PRESENTES EN UN ARCHIVO IDEAL DE UNA HORA DE ADQUISICIÓN



Para dicho fin, se hizo un programa en ROOT/C++ [22], [23], [24] que cuenta el número de repeticiones que tiene cada marca de GPS, si la frecuencia de la marca de GPS es igual a 200 Hz, entonces etiquetamos dicha línea y damos cuenta de ella en una *lista de buenos datos recolectados* GRDL. Estas GRDL son clases de C++ y se pueden ver en [25]. Para acceder a alguna de estas clases, por ejemplo, si el usuario quiere trabajar con la GRDL de SN del 2009, debe utilizar la clase de la siguiente forma:

```
...
//inside a for that iterates over each
//acquisition of adetermined file.

GRLSN2009 objeto();
if(objeto.HasGoodRecollectedDataList(
"NombreDeArchivo.root", gps) == true
)DoAnalysis();
elsecontinue;
}
```

En el anterior ejemplo de código, se crea una instancia de la clase GRLSN2009 mediante el objeto al que se ha llamado objeto y luego se llama al método *Has Good Recollected Data List* que recibe dos argumentos, el primero es el nombre del archivo tipo *root*, y el segundo es la variable que da cuenta de la marca del GPS. *Had Good Recollected Data List* devuelve *true* si en esa marca de GPS existen 200 adquisiciones, de lo contrario devuelve *false*. Si se devuelve *true* la función *DoAnalysis()* es llamada, ésta es una función que contiene el análisis que se le quiere hacer a los datos

en esa determinada marca de GPS. Si se devuelve *false* continúa hacia la siguiente marca de GPS.

La validación fue hecha sobre la totalidad de los datos de SN y CHA albergados en el repositorio de datos de la colaboración LAGODATOS. Dr obtuvo histogramas que dejaron ver el comportamiento anual de la adquisición de los datos y así, saber en cuáles regiones el análisis puede ser más fructífero.

Luego de realizar dicho análisis se encontró que en la estación de SN los datos recolectados entre los años 2006 a 2008 (Fig. 7 a Fig. 9) y parte de 2009 (Fig.10) tienen una frecuencia de marcas entre 128 y 256 marcas/s, lo cual muestra que la adquisición no funcionó como se esperaba.

Fig.7. HISTOGRAMA DE LA FRECUENCIA DE LAS MARCAS DE GPS PARA LOS DATOS ADQUIRIDOS EN SIERRA NEGRA DURANTE EL AÑO 2006

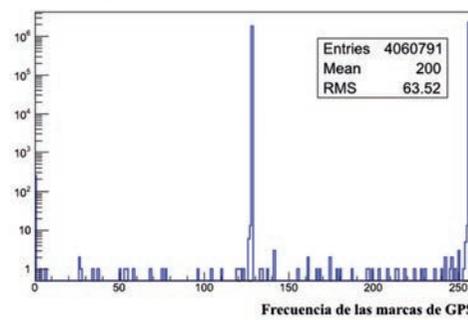


Fig.8. HISTOGRAMA DE LA FRECUENCIA DE LAS MARCAS DE GPS PARA LOS DATOS ADQUIRIDOS EN SIERRA NEGRA DURANTE EL AÑO 2007

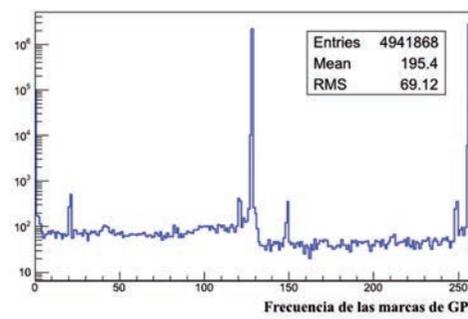


Fig.9. HISTOGRAMA DE LA FRECUENCIA DE LAS MARCAS DE GPS PARA LOS DATOS ADQUIRIDOS EN SIERRA NEGRA DURANTE EL AÑO 2008

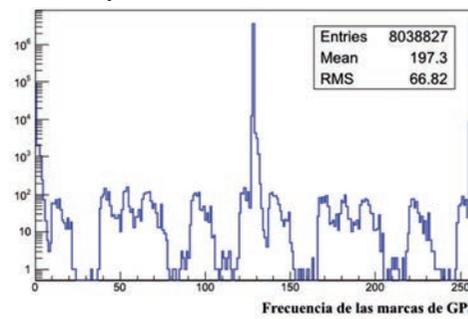
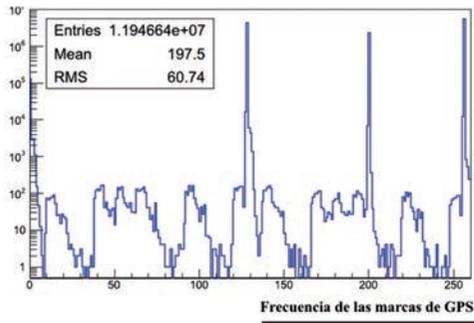


Fig.10. HISTOGRAMA DE LA FRECUENCIA DE LAS MARCAS DE GPS PARA LOS DATOS ADQUIRIDOS EN SIERRA NEGRA DURANTE EL AÑO 2009



Lo mismo sucedió en la estación de CHA en el año 2008 y parte de 2009 (Fig.11 y Fig.12). Se ve que hay marcas en todos los valores posibles entre 0 y 256, lo que es una clara muestra de irregularidad en la adquisición de los datos.

Fig.11. HISTOGRAMA DE LA FRECUENCIA DE LAS MARCAS DE GPS PARA LOS DATOS ADQUIRIDOS EN CHACALTAYA DURANTE EL AÑO 2008

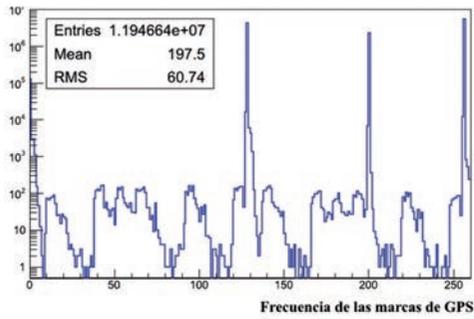
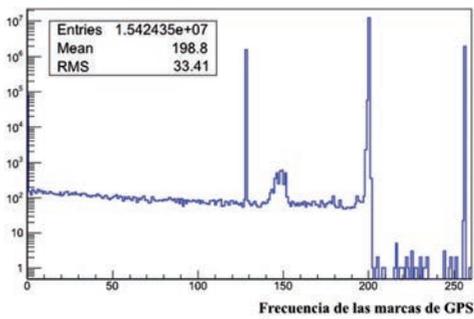


Fig.12. HISTOGRAMA DE LA FRECUENCIA DE LAS MARCAS DE GPS PARA LOS DATOS ADQUIRIDOS EN CHACALTAYA DURANTE EL AÑO 2009



Por otra parte, durante parte del 2009 en SN (Fig.10) y entre 2010 a 2012 (Fig.13-15) en Chacaltaya, gran parte de los datos muestran el comportamiento esperado, lo cual les da un factor de confiabilidad a estos datos al momento de realizar búsquedas de un fenómeno tan fino como los GRB. Como se observa el número de líneas que tiene un comportamiento irregular ha disminuido,

aunque aún hay gran cantidad de frecuencias de marcas de GPS entre 0 y 200.

Fig.13. HISTOGRAMA DE LA FRECUENCIA DE LAS MARCAS DE GPS PARA LOS DATOS ADQUIRIDOS EN CHACALTAYA DURANTE EL AÑO 2010

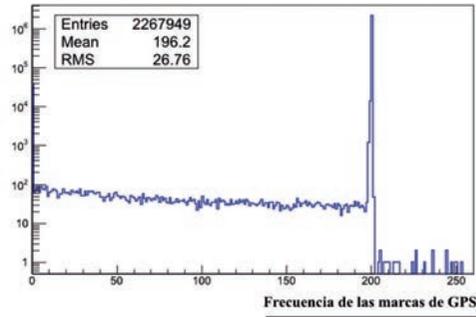


Fig.14. HISTOGRAMA DE LA FRECUENCIA DE LAS MARCAS DE GPS PARA LOS DATOS ADQUIRIDOS EN CHACALTAYA DURANTE EL AÑO 2011

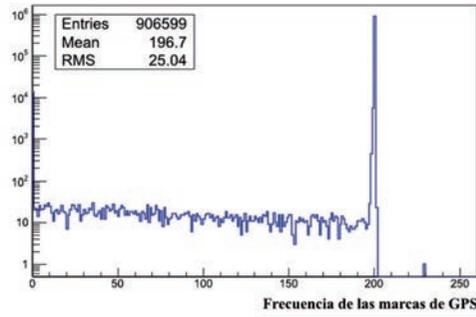
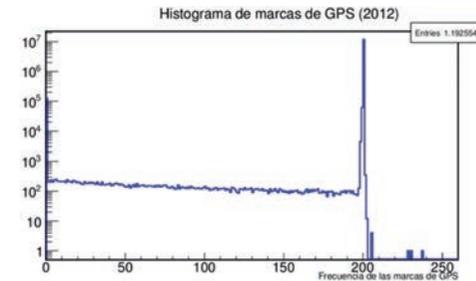


Fig.15. HISTOGRAMA DE LA FRECUENCIA DE LAS MARCAS DE GPS PARA LOS DATOS ADQUIRIDOS EN CHACALTAYA DURANTE EL AÑO 2012



Aquí se puede apreciar la verdadera importancia de las GRDL, ya que sólo tienen en cuenta los datos con frecuencia de 200 marcas/s, por lo tanto, el analizador que utilice estas GRDL va a poder implementar el mismo algoritmo sobre todos los segundos de adquisición de los datos de la colaboración LAGO y podrá asumir errores sistemáticos conservadoramente.

Con estas GRDL se hizo una búsqueda preliminar con el método de los excesos [14], con este método se buscan regiones donde ocurren excesos, siendo este:

$$Exceso = \frac{n_s - \bar{n}_h}{\sigma_h},$$

Con n_s igual al número de cuentas en un segundo, \bar{n}_h el promedio de todas las cuentas en una hora, y σ_h es la desviación estándar de las cuentas en una hora. Una cuenta es el número de destellos gamma en determinada conexión y en determinado umbral que mide el PMT. La idea es encontrar regiones de tiempo continuas en los datos en donde el exceso sea mayor o igual a 3.0.

Luego de realizar nuestro análisis sobre las GRDL de todos los años en SN y en CHA, no se encontró ninguna adquisición (5 ms) con el exceso requerido, por lo tanto, no se encontró ningún candidato a GRB.

En futuros análisis se podrían aplicar algoritmos más avanzados sobre las GRDL, entre ellos redes neuronales artificiales (ANN), árboles de decisión acelerada (BDT) y Matriz de co-ocurrencia, entre otros, para ello se necesitan variables cinemáticas sobre las cuales se puedan realizar cortes.

También sería bueno hacer una revisión de los registros de GRB de los satélites de la NASA y calcular si estos GRB pueden llegar los detectores en tierra de LAGO.

IV. CONCLUSIONES

La importancia de esta revisión está dada por la necesidad de mejorar día a día la recolección de datos, pues en este caso son ellos los que validan o derogan algún hallazgo realizado con los WCD.

Las GRDL que se crearon en este trabajo serán de gran utilidad para los futuros analizadores que quieran trabajar sobre los datos de la colaboración LAGO.

Tal y como se mencionó arriba, no se encontraron señales que dieran cuenta de GRB en los datos analizados. La dificultad de encontrar este tipo de señales en observatorios terrestres ha sido reportada previamente.

Se recalca la necesidad de tener GRDL de los nuevos datos que la colaboración LAGO vaya obteniendo. Sería excelente si se automatizara la validación de los datos, de manera tal que se supiera qué datos están buenos o malos, incluso antes de guardarlos en disco de manera permanente.

AGRADECIMIENTOS

Los autores agradecen el apoyo financiero de la Vicerrectoría de Investigación y Extensión de la Universidad Industrial de Santander a través del proyecto VIE5541. También agradecen al grupo GIRG por permitir usar el servidor Mac Pro Obatala. Fernando Quiñónez, agradece también a la Vicerrectoría Académica de la Universidad Industrial de Santander, por el soporte financiero al inicio del proyecto que generó este artículo, también agradece a la Universidad Santo Tomás por el apoyo financiero.

REFERENCIAS

- [1] P. Sommers and S. Westerhoff, «Cosmic ray astronomy,» *New Journal of Physics*, vol. 11, n° 5, p. 055004, 2009.
- [2] F. Halzen, «The Search for the Sources of the Cosmic Rays One Century after their Discovery,» de *Proceedings of XVI International Symposium on Very High Energy Cosmic Ray Interactions*, Batavia, IL, USA, 2010.
- [3] <http://www-glast.stanford.edu>.
- [4] J. Hinton, «Ground-based gamma-ray astronomy with Cherenkov telescopes,» *New Journal of Physics*, p. 055005, 2009.
- [5] <http://www.mpi-hd.mpg.de/hfm/HESS/>.
- [6] Colaboración HESS, «Detection of Gamma Rays from a Starburst Galaxy,» *Science*, vol. 326, No. 5956, pp. 1080-108, 2009.
- [7] Colaboración Fermi, «Detection of Gamma Ray Emission from Starburst Galaxies M82 and NGC 253 with the Large Area Telescope on Fermi,» *The Astrophysical Journal Letters*, Vol. 709, No. 2, p. L152, 2010.
- [8] W. Baade and B. Zwicky, «On Super-Novae,» de *Proceedings of the National and Academic Sciences of the USA*, Washington, 1934.
- [9] W. Baade and F. Zwicky, «Remarks on Super-Novae and Cosmic Rays,» *Physical Review*, vol. 46, pp. 76-77, 1934.
- [10] <http://hawc.umd.edu/>.
- [11] <http://www.auger.org>.
- [12] P. Mantsch, «The Pierre Auger Observatory progress and first results,» de *Proceedings 29th ICRC 2005*, Pune, India, 2006.
- [13] Colaboración LAGO, «The Large Aperture GRB Observatory,» *ArXiv e-prints*, p. 0906.0816, 2009.

- [14] Grajales, «Búsqueda de GRB en Sierra Negra,» de *Facultad de Ciencias Físico Matemáticas. Benemérita Universidad Autónoma de Puebla*, Puebla, México, 2010.
- [15] <http://halley.uis.edu.co/LAGOVirtual/>.
- [16] R. Camacho, R. Chacón, G. Díaz, C. Guada, V. Hamar, H. Hoeger, A. Melfo, L.A. Núñez, Y. Pérez, C. Quintero, M. Rosales y R. Torrens, «LAGOVirtual. A Collaborative Environment for the Large Aperture GRB Observatory,» de *Proceedings of the Second EELA-2 Conference*, Choroní, Venezuela, 2009.
- [17] L.A. Núñez, L. A. Torres, R. Torrén-Heeren, y E. Barrios-Hernandez, «Implementación de un Repositorio de Datos Científicos usando Dspace,» *e-colabora* <http://publicaciones.renata.edu.co/index.php/RCEC>, vol. 1, pp. 101-117, 2011.
- [18] L. Moreau, P. Groth, S. Miles, J. Vazquez-Salceda, J. Ibbotson, S. Jiang, S. Munroe, O. Rana, A. Schreiber, V. Tan, and L. Varga, «The provenance of electronic data,» *Communications of the ACM*, Vol. 51, No. 4, 2008.
- [19] M. Suárez y L. Núñez, «Instalación de un detector Cherenkov de agua para la detección de trazas de rayos cósmicos a 956 metros sobre el nivel del mar,» de *Escuela de Física, Univeridad Industrial de Santander*, Bucaramanga, Colombia, 2011.
- [20] A. Velarde and R. Ticona and P. Miranda, «PROYECTO LAGO BOLIVIA,» *Revista Boliviana de Física*, vol. 15, n° 15, pp. 32-38, 2009.
- [21] P. Allison and others, «Observing muon decays in water Cherenkov detectors at the Pierre Auger Observatory,» de *Proceedings of the 29th ICRC*, Pune, India, 2005.
- [22] R. Brun and F. Rademakers, «ROOT: An object oriented data analysis framework,» *Nuclear Instruments and Methods in physics research.*, vol. A389, pp. 81-86, 1997.
- [23] B. Stroustrup, *The C++ Programming Language*, Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 2000.
- [24] <http://code.google.com/p/lago-grdl/source/browse/#svn%2Ftrunk%2FGRDLGenerators>.
- [25] <https://sites.google.com/site/lagogrdl/>.