LIP-STUDENTS-21-20

Development of an interface to analyze events at the Pierre Auger Observatory, for use in Masterclasses

Ricardo Correia[1,a]

[1] *Universidade do Minho, Braga, Portugal*

Project supervisors: Raul Sarmento and Henrique Carvalho

October 30, 2021

Abstract. The Pierre Auger Observatory has made 10% of cosmic ray data available to the public.
Using this data, the Unity3D development platform and the Pierre Auger event viewer made in previous LIP Internships, it was possible to create an interface aimed at Masterclasses, intended to be used by students. With a dynamic and educational character, it is expected that these students acquire an understanding about cosmic rays and their reconstruction.

KEYWORDS: Pierre Auger Observatory, Cosmic rays, Masterclass, Interface

1 Introduction

1.1 IPPOG Masterclasses

A Masterclass is an activity led by an entity competent in a certain area, through digital platforms, in order to teach a certain audience.

The IPPOG (International Particle Physics Outreach Group) Masterclasses focuses on particle physics, in which every year, for one day, more than 13,000 students from 60 countries go to one of the about 225 universities, or associated research centers, to learn a little about this area of physics[1]. These activities are led by scientists specialized in the field, who provide an overview for the students. The students in turn carry out the proposed interactive activities, using a computer software that allows them to make measurements on real data from particle physics experiments. At the end of the day, the participants come together in a videoconference to have a taste on how it is to be a researcher.

Figure 1. Students in an IPPOG Physics Masterclass

1.2 Objectives

The idea for this project came from the fact that there are no Masterclasses with the Pierre Auger Observatory public data, which have recently been made available, at the IPPOG, and also because there is a 3D visualizer of Auger events developed at LIP that could be adapted for a Masterclass interface. To this end, a software has been developed, in which high school students will be able to reconstruct the properties of a cosmic ray by selecting certain parameters and better understand the physics behind them.

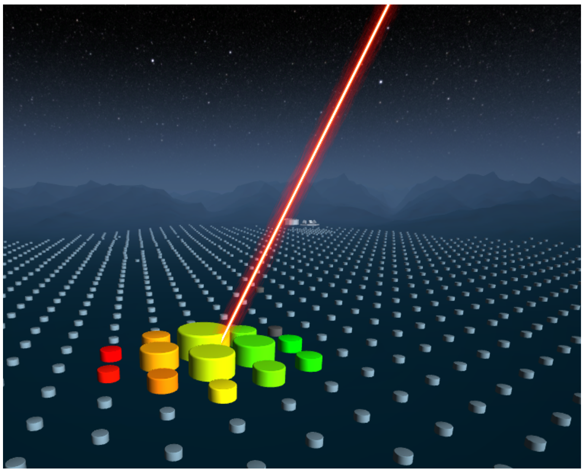

Figure 2. Cosmic ray event displayed with the Pierre Auger Observatory event viewer.

1.3 Pierre Auger Observatory

The Pierre Auger Observatory is located in the province of Mendoza, Argentina, and is mainly concerned with the detection and analysis of events associated with ultra-high energy cosmic rays. The observatory has about 1600 Cherenkov detectors spread over 3000 km^2 and four stations with fluorescence telescopes, which are responsible for detecting the shower of particles created after the interaction of the cosmic ray with the atmosphere. The Pierre Auger Observatory has made available to the public about 10 % of its official data, which are used in the preparation of this project.[2]

[a] e-mail: ricardocorreiaengenhariafisica@gmail.com

Figure 3. Fluorescence Detector station of the Pierre Auger Observatory.

2 Experimental activity

2.1 Overview

Initially, the student should open the application and load the file containing the event data.

Next, an event is chosen and its reconstruction is started. For this, the student must :

- Select the main Cherenkov station, i.e. the station with the highest signal;
- Select the tanks based on the distance to the main station;
- Select the tanks based on the time difference to the main station;
- Determine the azimuthal angle;
- Determine the first and last tank to detect the shower of particles originated by the cosmic ray.

With all this selection made by the students, the software proceeds with the reconstruction of the cosmic ray and performs:

- The calculation of the barycenter, i.e. the intersection of the shower axis with the surface;
- The calculation of the zenith angle;
- The fitting of a LDF (Lateral Distribution Function) to the data from the selected stations;
- The calculation of the $\frac{\chi^2}{ndf}$ of the fitted LDF;
- The calculation of the energy;
- The calculation of the galactic coordinates referring to the arrival direction of the cosmic ray.

Finally, the student checks that the event meets the following validation criteria, from the large-scale anistropy study [3]:

- Energy greater than 8 EeV;
- 5 or more stations around the main Cherenkov station;
- $\frac{\chi^2}{ndf} < 3$.

The student must validate the event only if all the criteria are fulfilled.

After dozens of events have been validated (typically, each group of two students is expected to analyze about fifty events), a flux map is constructed like the one in the Fig.14.

Finally, the interpretation and discussion of the results is made, at first in a small group and afterwards with all the students and the participating scientists.

2.2 Reconstruction of the primary cosmic ray

To reconstruct the primary cosmic ray, it is necessary to know what the azimuthal angle, zenith angle, the barycenter, the energy of the primary cosmic ray and its galactic coordinates are. Once the azimuthal angle is selected by the students, the method of determining the other variables will be presented, also mentioning how the LDF graph and reduced χ^2 was obtained.

2.2.1 Zenith angle

Required parameters:

- The azimuthal angle;
- The first and last detector associated with the shower.

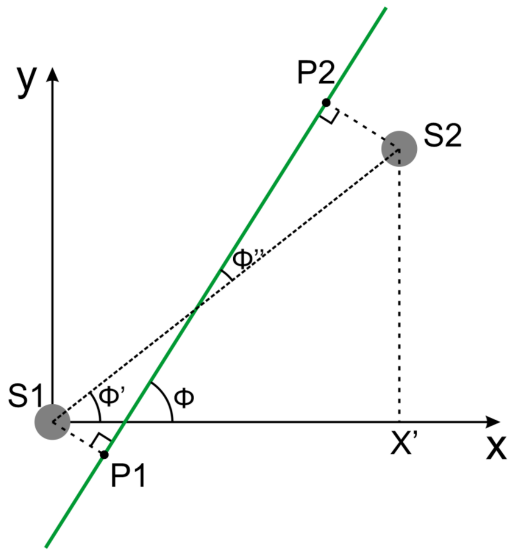

Figure 4. Horizontal projection of the cosmic ray.

Legend:
green line: cosmic ray projection on the earth's surface
S1: station 1;
S2: station 2;
P1: shortest distance point between ray and S1;
P2: shortest distance point between ray and S2

Let's consider as an example the projection of the cosmic ray on the Earth's surface represented in the image above.

Knowing the position of the first (S1) and last (S2) station associated with the cosmic ray, move the origin of the referential to the position of the first station and calculate the distance between them (h). Having these parameter calculated, we determine the value of Φ' by the following expression:

$$\Phi' = \cos^{-1}\left(\frac{X'}{h}\right) \quad (1)$$

Where X' represents the x-coordinate of the last station.

Once we have determined Φ' and know the azimuthal angle, we trivially calculate the value of Φ'':

$$\Phi'' = \Phi - \Phi' \quad (2)$$

Then, from trignometric relations, the length of a portion of the cosmic ray projection ($dist$) is determined:

$$dist = \cos(\Phi'') \times h \quad (3)$$

Knowing the $dist$, we can now view the cosmic ray from the side:

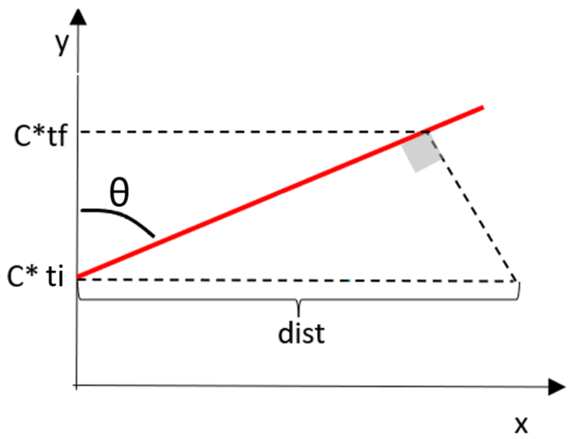

Figure 5. Lateral projection of the cosmic ray.

Once we know the value of the speed of light in vacuum (c) and the time recorded at the first and last station, we can easily calculate the value of the zenith angle (θ):

$$\theta = 90^o - \cos^{-1}\left(\frac{c(tf - ti)}{dist}\right) \quad (4)$$

These calculations only work properly if the Φ and the angle that the two stations make with the horizontal are in the 1st or 3rd quadrant. If either of them were not, it was a matter of placing them in one of those quadrants.

2.2.2 Barycenter

Definition: the barycenter represents the center of the shower that hit the ground at time $t0$ (see Figure 6).

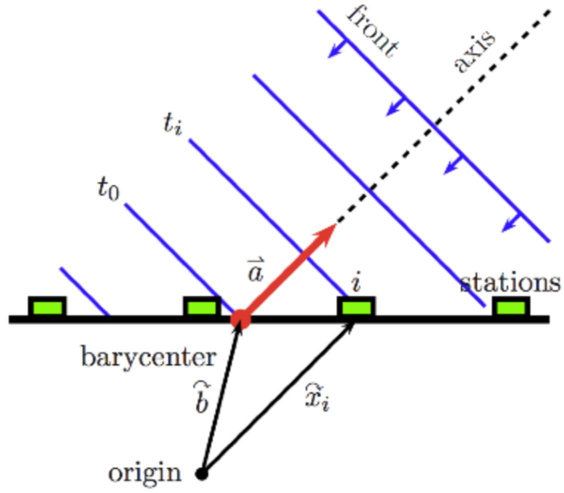

Figure 6. Barycenter.

To calculate the barycenter, the logarithm of the positions of the stations associated with the given event (except saturated ones) was weighed based on the logarithm of the signal recorded by each tank.

2.2.3 LDF Graphic

The LDF graphic relates the signal recorded in each tank (y-axis) with its distance from the shower's axis (x-axis). The signal recorded in each tank was acquired from those 10 % of data provided to the public by the Pierre Auger Observatory, but the distance to the axis of the shower had to be calculated. To determine the distance to the axis of the shower for each tank, we change the original coordinate system (see Figure 7) of the tanks to the plane perpendicular to the axis (see Figure 8), and calculate the distance to the axis using the X and Y of this new reference (see Figure 9).

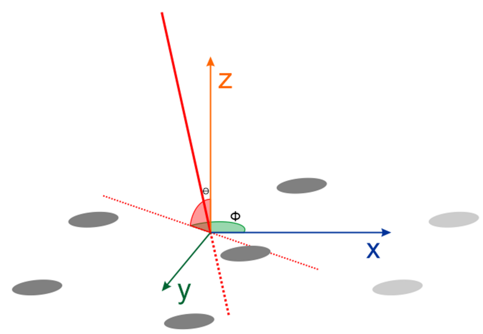

Figure 7. Old axis system.

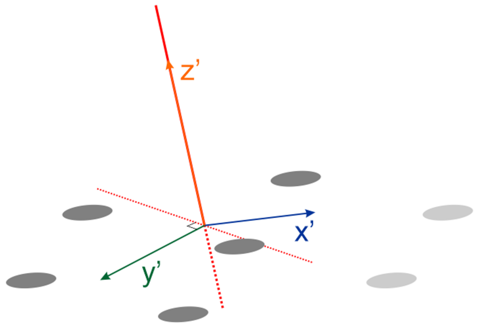

Figure 8. New axis system.

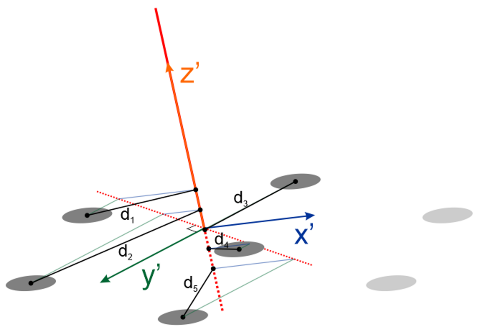

Figure 9. Shortest distance from stations to the shower axis in the new axis system.

Since to construct the LDF graph we do not take into accountr the silent stations (stations with no signal that are used in the official LDF algorithm), sometimes leading to considerable differences from the standard Auger reconstruction, we implemented the following thresholds:

- The farthest 30% of stations are removed if the number of stations was greater than 3;
- When a satured station is present, and if the distance between the first and last station is <300m, saturated stations are selected, otherwise not;
- All stations that are more than 1800m away from the axis are discarded;

The following figure represents an example of an LDF plot:

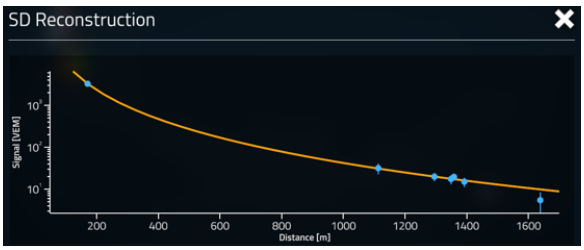

Figure 10. LDF plot.

2.2.4 Energy of the primary cosmic ray

For the following calculations, the stations that were considered in the elaboration of the LDF chart were used.

Using the NKG (Nishimura-Kamata-Greisen) function (which characterizes the LDF curve):

$$S(r) = S_{1000}\left(\frac{r}{1000}\right)^{\beta}\left(\frac{r+700}{1700}\right)^{\beta+\gamma} \tag{5}$$

In which $\gamma = 0$ by approximation :

$$S(r) = S_{1000}\left(\frac{r^2+700r}{1000\times 1700}\right)^{\beta} \tag{6}$$

The linearization was performed:

$$ln(S(r)) = ln(S_{1000}) + \beta ln\left(\frac{r^2+700r}{1000\times 1700}\right) \tag{7}$$

In which: $Y = ln(S(r)), B = ln(S_{1000}), A = \beta$ e $X = ln(\frac{r^2+700r}{1000\times 1700})$

Staying:

$$Y = B + AX \tag{8}$$

This way we can already do a linear regression using the method of least squares and determine B.

At this point, it is worth mentioning that the Unity platform was not conceived to have dedicated data analysis tools (e.g. fitting, etc), hence the need to devise and implement simple numerical algorithms.

Once we know B, we know S_{1000}:

$$S_{1000} = e^B \tag{9}$$

Then we used the following formulas taken from the Pierre Auger Observatory website[4]:

$$S_{38} = S_{1000}/F(h) \tag{10}$$

$$F(h) = 1 + 0.92h - 1.13h^2 \tag{11}$$

$$h = cos^2(\theta) - 0.62096 \tag{12}$$

Using the following table and the θ determined earlier we arrive at the value of the energy of the primary ray:

	0°<θ<30°	30°<θ<45°	45°<θ<60°
$A/10^{17}eV$	1.89±0.08	1.86±0.04	1.83±0.04
B	1.029±0.012	1.030±0.006	1.034±0.006

Figure 11. Parameters used for energy calibration, at different zenith angle intervals.

$$E = A \times (S_{38})^B \tag{13}$$

A and B are obtained from the table above.

2.2.5 Reduced χ^2

Initially the difference between the signal value at each station (official value coming from those 10% data provided) and the value from the NKG equation (equation 6) was performed:

$$\delta S(r) = Signal - S(r)_{NKG} \tag{14}$$

Then all the stations' values were summed, using the following expression:

$$\frac{\delta S(r)^2}{DSignal^2} \tag{15}$$

Where $DSignal$ was also taken from the 10% data.

The result of the sum of the equation 15 is the intended error.

2.2.6 Galactic coordinates

The galactic coordinates indicate the location in the universe where the cosmic ray came from.

For its calculation, we used the conversion from local coordinates to galactic coordinates[5].

To determine them it is necessary to use the azimuthal and zenithal angle of the ray, the date of its detection, and the coordinates of the observatory.

The date of detection and the observatory coordinates are used because the earth moves around its axis and around the sun, and it is necessary to determine where it has been in the universe in order to efficiently determine the galactic coordinates.

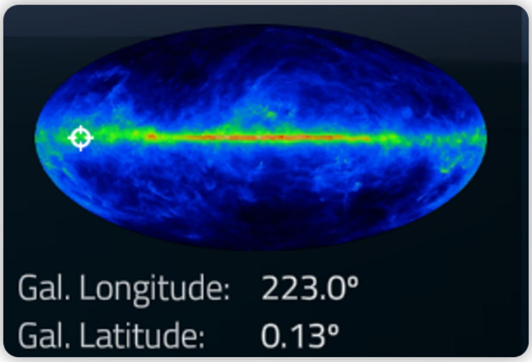

Figure 12. Universe map with the galactic coordinates of an event.

2.3 Flux map

To create the flux map it was necessary for users to perform several event reconstructions, and since there were no users who could do this it was necessary to implement an algorithm that was able to make a selection that would be close to the human selection, that is, that would introduce errors in the selection of the parameters.

To have a means of comparison, an official flux map was created (Fig.13), in which the parameters used coming from the 10% data provided by the Pierre Auger observatory , and a flux map obtained (Fig.14), in which the parameters that should be selected by the users, were selected by the algorithm above. For the graphical representation of both flow maps, the algorithm from website was used [6].

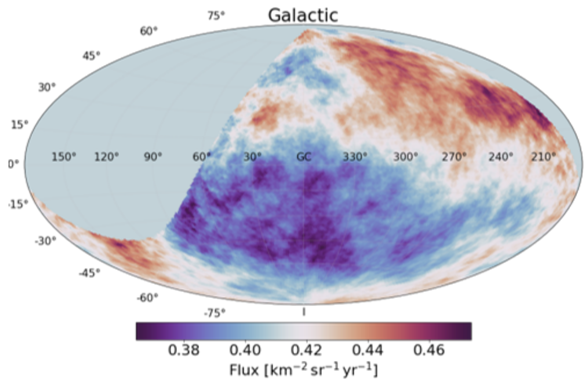

Figure 13. Official flux map.

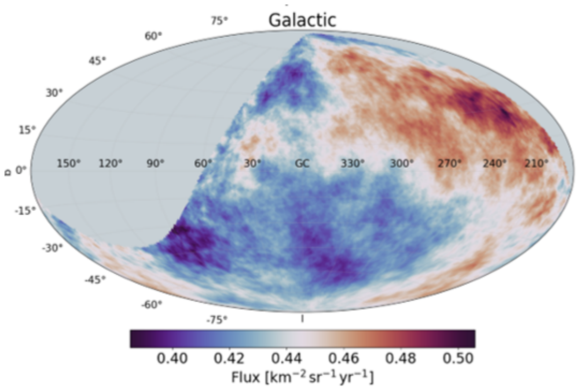

Figure 14. Obtained flux map.

You can see that the flux maps are similar, and the darker areas have a higher density of cosmic rays.

3 Performance

3.1 Reconstruction performance

To get an idea whether the algorithms for determining the variables calculated in chapter 2.2 introduced an acceptable error in the calculations, it was necessary to make tests with official parameters. Even though we know that some parameters are chosen by the users (for example the azimuth angle), and may introduce a large deviation from the official parameters,but the algorithm must be properly adapted to give results close to the official ones, when the parameters are close to the original ones. As such, the following tests were made to validate the performance of the algorithms implemented in chapter 2.2:

3.1.1 Performance of the zenith angle algorithm

To verify that the algorithm of chapter 2.2.1 introduced a low error, it was enough to use as parameters of the function: the official azimuth angle and the first and last station of the event in question.

Using these official parameters, the result was compared (for several events) with the zenith angle of those 10% of data, in which the difference has:

- a mean of 2.96º;
- a median of 1.43º;
- a standard deviation of 4.91º;
- θ<1º:38%;
- θ<2º:62%;
- θ<3º:76% ;

Since the error introduced is relatively low, the algorithm can be used without concern.

3.1.2 Performance of the barycenter algorithm

To determine the barycenter, and since this did not require any parameters that were chosen by the students, they only compared it to the official value (value of the 10% data).

As for this algorithm, the error introduced was significant. Because of this, this algorithm will be subject to improvements in the future.

3.1.3 Performance of the LDF graph algorithm

To construct the LDF graph, the "obtained" graph was superimposed on the standard Auger graph. The "obtained" graphic was determined using the zenith angle calculated from the official parameters, the official azimuthal angle and the official barycenter (because the calculated one introduced a considerable error), while the "official" one was determined using a more advanced algorithm present on Pierre Auger's website, in which all the parameters were the official ones.

After several comparisons, it was found that the graphs were very similar, and because of this the LDF graph construction algorithm performed well.

3.1.4 Performance of the energy algorithm

To calculate the energy of the primary cosmic ray, the official value was also compared.

For this comparison, the calculated energy received as a parameter the zenith angle calculated with official parameters, in which the NKG function had been simplified, while the official energy used the S_{1000}, S_{38} and the zenith angle from such data.

Calculation of the percentage difference of energies:

- mean: 11.9%
- median: 7.51%
- standard deviation: 79.58%

Making the comparison for several events, it turned out that both energies were quite close in the vast majority.

3.1.5 Performance of the error algorithm

As for the error, only the calculated error was compared with the error coming from the 10% data.

For the calculated error, the S_{1000} and the β determined to calculate the energy in the previous chapter were used.

Comparing for several events, it was found that the error calculation introduced a low error.

3.1.6 Performance of the galactic coordinates algorithm

The galactic coordinates were also compared with the official galactic coordinates.

As parameters of the calculated galactic coordinates, we used the official azimuthal angle and the zenithal angle calculated in chapter 3.1.1, while in the official one we used only official parameters.

In some events, namely the very low energy events, the difference between the values was high, but there was no problem since the very low energy events were discarded in the datasets.

3.2 Student testing

To test whether the program worked properly, it was decided to conduct a test with high school students in Lisbon.

On the day of the test, the students received a script and the teacher involved explained the physics behind this type of cosmic events.

After the students solved the script (again, it was for them to understand the physics behind it), the teacher started the program and solved two events in order for the students to understand how it worked.

Then the students solved a considerable number of events, and in most of them they obtained results very similar to the official values.

With this test, it turned out that this program can be implemented in a Pierre Auger Masterclass, because the students were able to perform all the tasks and steps of the reconstruction quite successfully.

4 Conclusion

In conclusion, the project has received good feedback from those who have tried it, although there are some improvements to be made and some additions to the project that are intended to be made during this academic year. When the project is fully completed, it is hoped that it can be used in Masterclasses around the world and increase the enjoyment of cosmic ray physics by those who use it.

Moreover, the realization of this project allowed me to learn the use of the UNITY platform, the C# language, the elaboration of algorithms and to improve my ability to solve problems, because during the realization of this project appeared many unforeseen obstacles that I had to solve.

Annex: Interface presentation and description

This chapter will show how the program works and what important features it contains.

When we start the program we get something like the figure 15:

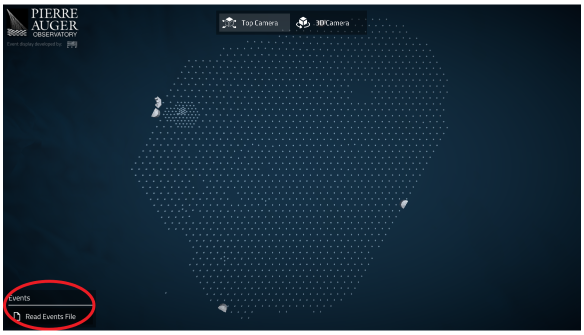

Figure 15. Initial screen

To open the event file, click on the "Read Events File" button, as shown on Fig.15, and select the events file.

Then click on an event and start the reconstruction.

As mentioned in chapter 2.1, the user starts by determining the main Cherenkov station, taking into account it's position and signal, and select that station cicking over it:

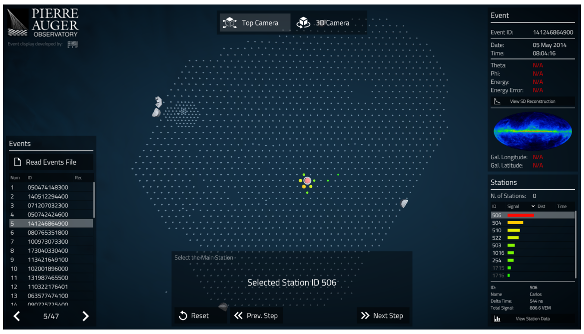

Figure 16. Selection of the main Cherenkov station

Then the user select the remaining event stations, by distance and by time, using their corresponding sliders:

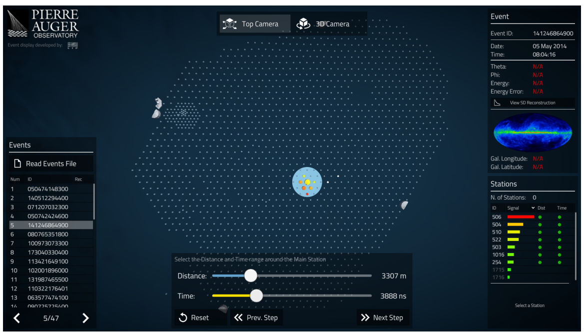

Figure 17. Selection of the stations by time and distance

Next, the user selects the azimuthal angle aligning the arrow colors with the stations colors, using the angle slider:

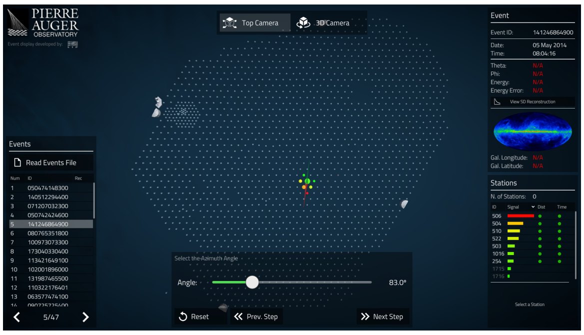

Figure 18. Selection of the azimuthal angle

Following the user has to select two stations, being the best option to select the oposite stations aligned to the arrow defined in the previous step, and being the first station the one with least time and the last station the one with most time:

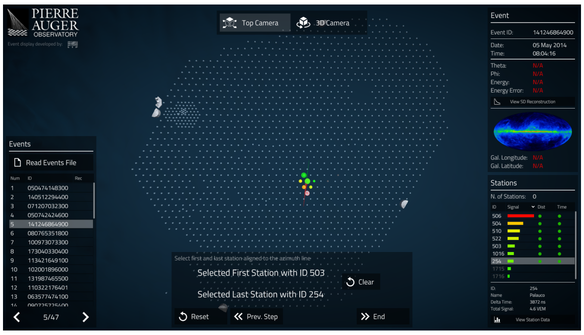

Figure 19. Selection of the first and last station

At the end of selecting all these parameters,the user simply cheks the validation criteria described in chapter 2.1. and validate or not the event:

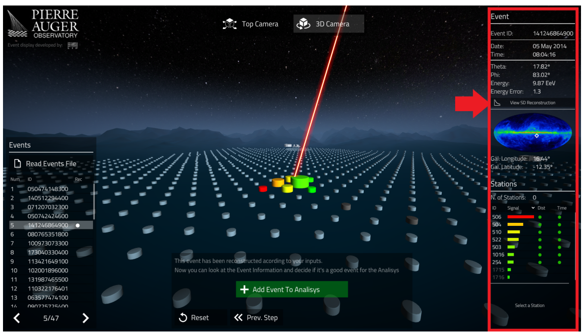

Figure 20. Reconstruction result and event validation

In the red square of the figure 20 you will find the event ID, the date and time of the detection, the selected azimuthal angle, the calculated zenith angle, the calculated error, the calculated energy, the calculated galactic coordinates, the LDF plot and the values (time and signal) recorded in each selected tank.

To see the LDF plot just click on the"View SD reconstruction" button, pointed by the red arrow on Fig.20.

Acknowledgements

A special thanks to the teacher Raul Sarmento, who was always very present in this project, giving tips, formulas, explaining the theory behind certain issues, and also to Henrique Carvalho who helped me in learning the language and the Unity platform, helping in the development of algorithms.

I also thank the high school students from Lisbon who tested the functioning of this program, the teacher Sofia Andringa and the teacher Luís Afonso who made this test possible, also contributing with tips for the project

References

[1] IPPOG, International Masterclasses, https://physicsmasterclasses.org/ (2021)

[2] Pierre Auger Collaboration, Auger Open Data Realise, doi:10.5281/zenodo.4487613 (2021), [Online]

[3] A. Aab, P. Abreu, M. Aglietta, I.A. Samarai, I.F.M. Albuquerque, I. Allekotte, A. Almela, J.A. Castillo, J. Alvarez-Muñiz, G.A. Anastasi et al., Science **357**, 1266 (2017)

[4] Pierre Auger Observatory, Internal Note **GAP-2005-035** (2005)

[5] W.M. Smart, Text-book on spherical astronomy (1965)

[6] Auger Open Data, Large-scale anisotropy, https://www.kaggle.com/augeropendata/large-scale-anisotropy (2021)