Building an effective sea-level cosmic ray generator

8254R

1 Abstract

A program written in the Python [1] language was designed and created to be able to quickly produce realistic sea-level cosmic rays for use in other projects and experiments, without the need to simulate them directly. The generator is based upon theoretical calculation as well as parametrised sea level data from many experiments. A Monte-Carlo simulator was then used to validate some of the generator's results and provide confidence in its output.

2 Introduction

This project aims to produce a program that will output realistic sea-level cosmic rays, such that they can be used in simulations or experiments in the future. This is useful as simulating the creation of the comic rays requires significant time and computational power. I will start by summarising the necessary background of cosmic rays and air showers, before outlining the theoretical background for the models used. A summary of cosmic-ray experimental data is then given to aid the understanding and creation of the generator. I will then motivate the design of the generator, and give some results of its output. Finally, simulations using CORSIKA, a program for simulating cosmic-ray air showers, will be presented in an attempt to validate the output of the generator and the theory and data behind it.

3 Background

3.1 Primary cosmic rays and secondary formation

Cosmic rays are particles originating extra-terrestrially that collide with the atmosphere. Primary cosmic rays ('primaries') are produced by astrophysical objects such as stars – including the sun, while 'secondaries' are synthesised from interactions between primaries and particles in the Interstellar Medium (ISM), and (more importantly for this project) Earth's atmosphere. Primaries consist of all charged particles with a lifetime of order 10^6 years or greater [2], including electrons, positrons, etc. though they are most commonly bare atomic nuclei.

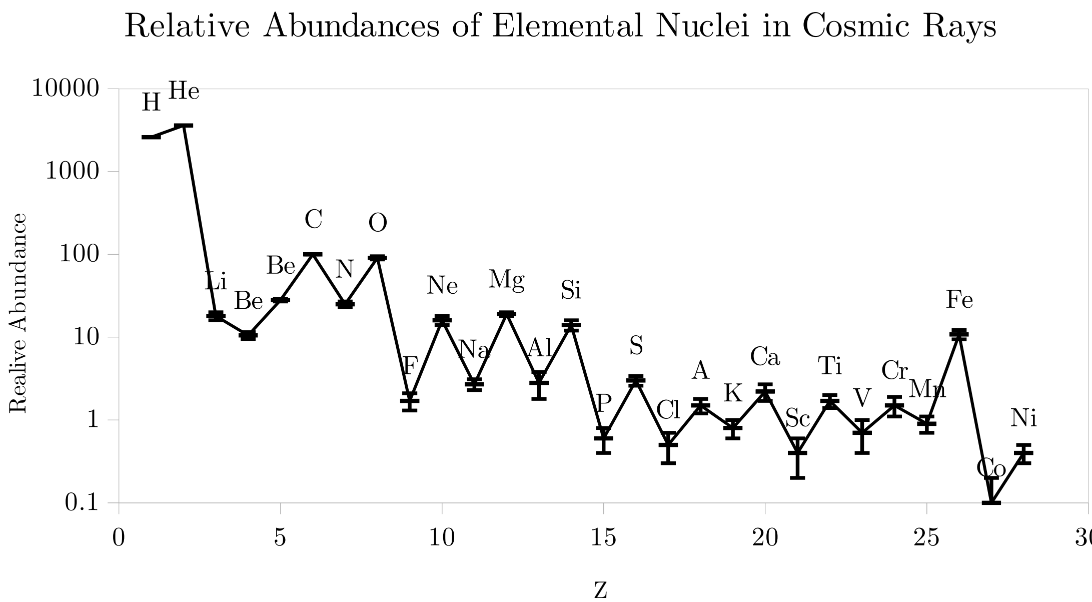

Figure 1: The relative abundances of nuclei in cosmic rays arriving at Earth's atmosphere, normalised to Carbon = 100. Note the logarithmic y scale.

Figure 1 (data from [3]) shows relative abundances of nuclei from Hydrogen up to Zinc in cosmic rays above the atmosphere. The majority of cosmic rays that arrive at Earth's atmosphere are protons and α particles, with heavier nuclei up to Nickel being of order 10^2 to 10^3 less abundant. Large nuclei with $A > 65$ are present in cosmic rays but their abundance is suppressed [4], [5]. Many experiments have measured the flux and energies of primary cosmic rays at the top of the atmosphere. The resulting data have shown simple power-law dependencies in different energy regions of the cosmic ray spectrum, with shifts in power-law indices at 10^{15} and 10^{18} eV per particle, and a high-energy cut-off around 10^{20} eV [6].

Primaries are understood to be essentially isotropic in deep space, with deviations from this very small (though of some astrophysical interest) [7].

Upon colliding with the atmosphere, the nuclei produce hadronic showers, largely comprised of pions and kaons, in events called Extensive Air Showers (EAS) [6]. These in turn form sub-showers until their energy is low enough that they decay before interacting, with charged hadrons producing muons, $K^\pm \longrightarrow \mu^\pm$ and $\pi^\pm \longrightarrow \mu^\pm$. As such, muons are the most abundant cosmic ray particle at sea level [2], and will be the main focus of this project.

3.2 Cascade equations

The behaviour of a cascade of particles in an EAS can be modelled by the use of the very general cascade equations – Equation 1. The following equations are defined as in [6].

$$\frac{dN_i(E_i,X)}{dX} = -\frac{N_i(E_i,X)}{\lambda_i} - \frac{N_i(E_i,X)}{d_i} + \sum_{j-i}^{J} \int_{E}^{\infty} \frac{F_{ji}(E_i,E_j)}{E_i} \frac{N_j(E_j,X)}{\lambda_j} dE_j$$

Equation 1 The cascade equations. All quantities in lab frame. $dN_i(E_i,X)dE_i$ is the flux of type i at depth X (measured along the line of the cascade) in energy interval [E, dE]. The probabilities of particles of type j (having interaction and decay lengths λ_i and d_i) interacting in distance dX are dX/λ_i and dX/d_i. The dimensions of X, λ_i and d_i are all [mass][length]$^{-2}$ where decay and interaction lengths are defined as the density multiplied by the appropriate decay length. This provides a density-independent interaction length, but means the decay lengths d_i are now proportional to density.

Most of the relevant particle physics is contained in the last term, where F_{ij} gives the 'dimensionless particle yield' that corresponds to a particle of energy E_j producing a particle i of energy $E_i < E_j$ on collision with an atmospheric particle.

In subsequent analysis, it is often assumed that the nuclei can be treated as a set of independent nucleons, called the superposition model. This assumption is supported by the the large interaction energies relative to binding energies in the nuclei [4]. Taking the average over many showers, this assumption “distribution of nucleon interaction points in the atmosphere coincides with that of more realistic calculations that account for nucleus interactions and for breakup into remnant nuclei” [4].

An important solution to these equations is that of a power law, which occurs when we assume that "for energies that are large compared to the critical energy, collisional losses and Coulomb scattering can be neglected" and "the radiation length is independent of Energy" [6]. These assumptions are in fact not well motivated for data acquired from most sea-level detector experiments, with relatively low particle energies (see Figure 3). The details of this analysis will not be repeated here, but it motivates the form for Hadronic flux given in Equation 2.

$$N(E,X) \propto \exp(-X/\Lambda) E^{-(\gamma+1)}$$

Equation 2 Hadronic flux given in at slant depth X and energy E, with γ being the integral spectral index and Λ being some characteristic decay length in the atmosphere.

This is a simple power law spectrum. Secondary cosmic ray flux data are sometimes transformed to E^p log(dN/dE) for some power p (note dN/dE "differential flux") which makes it easy to relate them to power law spectra.

However, the flattening of the distribution at higher energies shows that the assumptions hold better here. It is noteworthy that the *primary* cosmic ray spectra are described by power laws quite precisely in certain energy regions (as mentioned in 3.1), but the astrophysical mechanisms of this and any deviation from it are beyond the scope of this project.

3.3 Response Curves

Response curves give an estimate of the range of energies of primary particles that produce a secondary particle with a given energy (in fact a range from some E up to ∞). They are defined in detail in [6]. The important result from this type of analysis is to calculate the mean primary energy $\langle E_0 \rangle$ given a secondary energy E. This is given in Equation 3.

$$\langle E_0 \rangle = \frac{E\, n'(E, X)}{n(E, X)}$$

where n' is the flux if primary spectrum were $E'_0 = [E_0/E]\, N_0$

Equation 3: Mean primary energy for production of secondary of energy E with flux n.

This equation can be solved numerically, and an important result given in [6] is that for muons at ~20 GeV (an overestimate for sea-level muons), the mean primary cosmic ray energy is ~750 GeV. This uses the assumption of negligible muonic decay and importantly the power-law spectrum of primaries and secondaries as above.

This will become important when simulating muon production, as a large range of primary energies must be used to allow accurate simulation of muon fluxes at sea-level. This will to some extent motivate the simulation parameters used later, though the constraints set by simulation software will be more significant.

3.4 Atmospheric muon parametrisation

Given the production channels of muons in the atmosphere, with $K^{\pm} \longrightarrow \mu^{\pm}$ and $\pi^{\pm} \longrightarrow \mu^{\pm}$, is is possible to derive a simple parametrised formula for the energy spectrum of muons at sea level, as set out in [6]. This is shown in Equation 4 and Equation 5. The first and second terms on the RHS of Equation 4 represent the pionic and kaonic contributions to the muon flux. Note that at higher energies, the Kaon channel becomes more significant in the total flux, but never surpasses the pion channel.

$$\frac{dN_\mu}{dE_\mu} = S_\mu \frac{14000\, E_\mu^{-2.7}}{\mathrm{m^2\ sec\ sr\ GeV}} \left\{ \frac{1}{1 + \frac{1.11\, E_\mu \cos(\theta)}{115\ \mathrm{GeV}}} + \frac{0.054}{1 + \frac{1.11\, E_\mu \cos(\theta)}{850\ \mathrm{GeV}}} \right\}$$

Equation 4: Parametrised formula for differential muon cross section at sea-level, with zenith angle ϑ and muon energy E_μ. The suppression factor S_μ is given below.

The cos(ϑ) prefactors and other constants are functions of the relevant branching ratios, masses and physical constants that are measured using accelerator data.

$$S_\mu(E_\mu) = \int_0^{X_0/\cos(\theta)} \frac{dX}{\Lambda_N} \left(\frac{X\cos(\theta)}{X_0}\right)^{p1} \left(\frac{E_\mu}{E_\mu + \alpha(X_0/\cos(\theta) - X)}\right)^{p_1+\gamma+1} \exp\left[-\frac{X}{\Lambda_N}\right]$$

$$\text{where } p_1 = \epsilon_\mu \Big/ E_\mu \cos(\theta) + \alpha X_0$$

Equation 5: Suppression factor for muon production given muon energy E_μ, a is rate of energy loss through atmosphere (defined in the same way as in the cascade equations). Λ_N and γ are the decay length and spectral index as in the power law equation. ε_μ is the decay constant as defined in [5].

This suppression factor may be simplified if one assumes muon production occurs high in the atmosphere (upper integral limit to ∞ and second term has no X dependence). This leads to Equation 6.

$$S_\mu(E_\mu) \approx \left(\frac{\Lambda_N \cos(\theta)}{X_0}\right)^{p1} \left(\frac{E_\mu}{E_\mu + 2\,\text{GeV}/\cos(\theta)}\right)^{p_1+\gamma+1} \Gamma[p_1+1]$$

Equation 6: Simplified suppression factor. Note that $aX_0 \cong 2GeV$ at sea level.

It is clear to see from Figure 2 that the "no muonic decay" (i.e. $S_\mu = 1$) assumption breaks down at lower energies, as the theoretically predicted muon flux is much higher than the various experimental values. The unsuppressed theory seems to fit well above ~200 GeV/c, while the suppressed theory is better at lower energies but still systematically higher than the data. The results in the figure were produced using the `theory()` and `suppressed_theory()` functions in the codebase.

The angular dependence of Equation 4 is difficult to analyse, as it is a differential spectrum not an integral spectrum. In [6] this integration is performed using the approximation of low energy muons (justified by experimental data similar to that in Figure 3) and that p_1 is small; it is shown that the approximate theoretical muon flux scales as approximately $\cos^2\vartheta$. This is also motivated by experimental data, see [8] which measures the distribution to be $\cos^n\vartheta$ with $n = 2.33 \pm 0.11$.

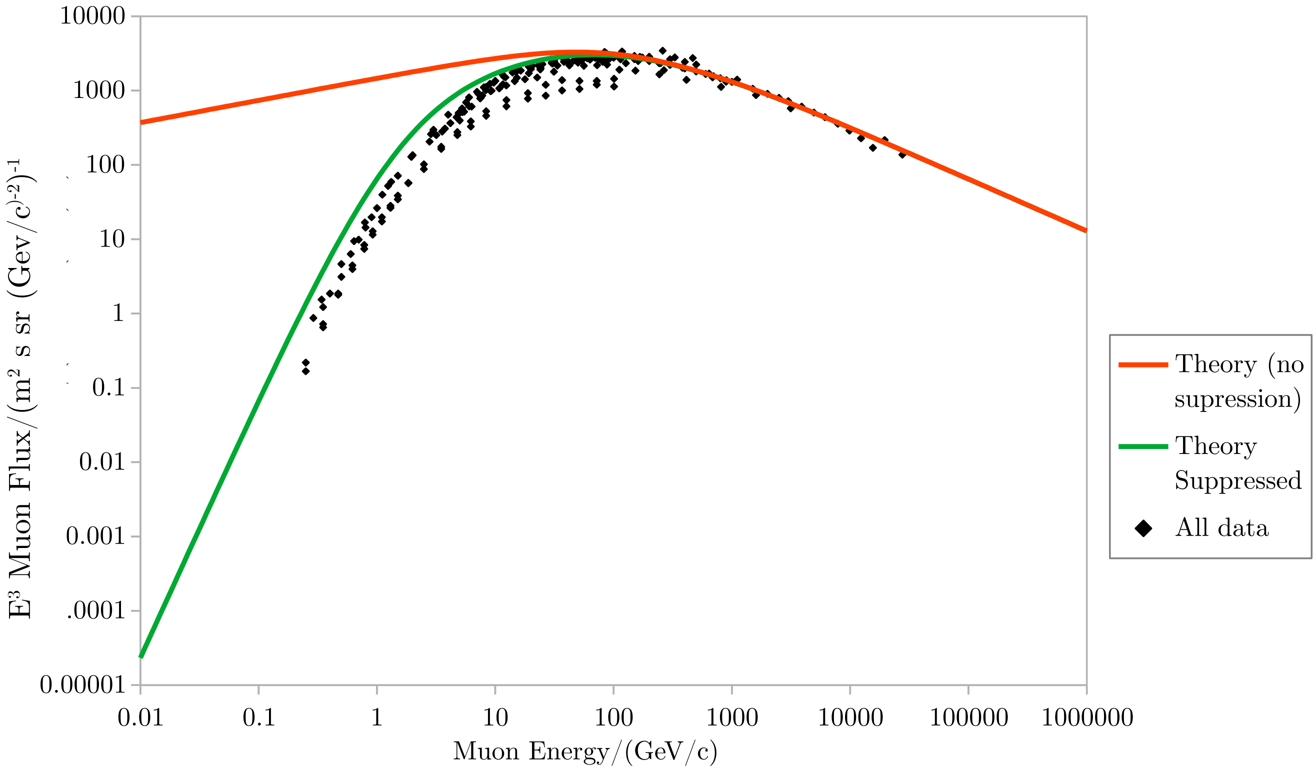

Figure 2: All muon flux data (besides BESS data) plotted against the theoretical formula for muon flux at the zenith. All data sources shown in Appendix B. Good agreement with no suppression is shown above 200 GeV but below this the non-suppression assumption breaks down. Even the suppressed theory appears systematically higher below 200 GeV using parameters taken from [5].

3.5 Primary cosmic ray data

As mentioned in 3.1, primary cosmic rays generally follow a power-law spectrum, valid for certain energy ranges. The differential energy spectrum for a certain particle can be described by Equation 7. Note the inclusion of the mass number of the nucleus as all energies are considered *per nucleon.*

$$N_i(E)\,\mathrm{dE} = A_i K_i E^{-\gamma_i}\,\mathrm{dE}$$

Equation 7: Flux of primary cosmic ray species i given mass number A_i.

Table 1 shows the power law indices and prefactors for elements Hydrogen to Iron in primary cosmic rays, taken from table 1 in [9]. This uses data from various experiments as listed in Appendix A. These data were considered in the lowest energy region, i.e. E < 10^{15} eV before the first shift in power law behaviour. It is appropriate to consider the range below 10^{15} eV given that our focus is on sea-level muon fluxes, with a mean energy around 4 GeV, and the results in 3.3 regarding response curves.

Table 1: Values of K_i and γ_i for primary cosmic ray particles.

Element	A_i	K_i(cm^2 s sr GeV/n)$^{-1}$	γ_i
H	1	8540.20e-8	1.63
He	4	507.07×e-8	1.66
C,N,O	14	46.96e-8	1.76
Ne–Si	28	0.8438e-8	1.63
Fe	56	0.012e-8	1.47

3.6 Sea-level muon data

Data on cosmic rays at sea level has been taken in many experiments. Some of these data are plotted in Figure 3, along with an order-3 polynomial best fit. This has not been done to test any particular theory, but to allow evaluation of the best-fit curve which was useful in creating the generator. Error bars are not included with the data as they were impossible to account for when creating the best-fit curve (the main aim of the plot), and are mostly too small to see on the plot.

The BESS datasets were excluded in the best-fit as they appeared systematically lower than others at all energies, and were very dense (small ΔE_μ), which would skew the fitting algorithm towards them. At high energies (>100 TeV), the best-fit polynomial begins to increase again, which does not seem to fit the trend of the data. The validity (more precisely lack thereof) of this extrapolation is discussed in 4.2.

The generally accepted overall integral intensity of muons above 1 GeV is approximately 70 s^{-1} sr^{-1} m^{-2} [2] [10].

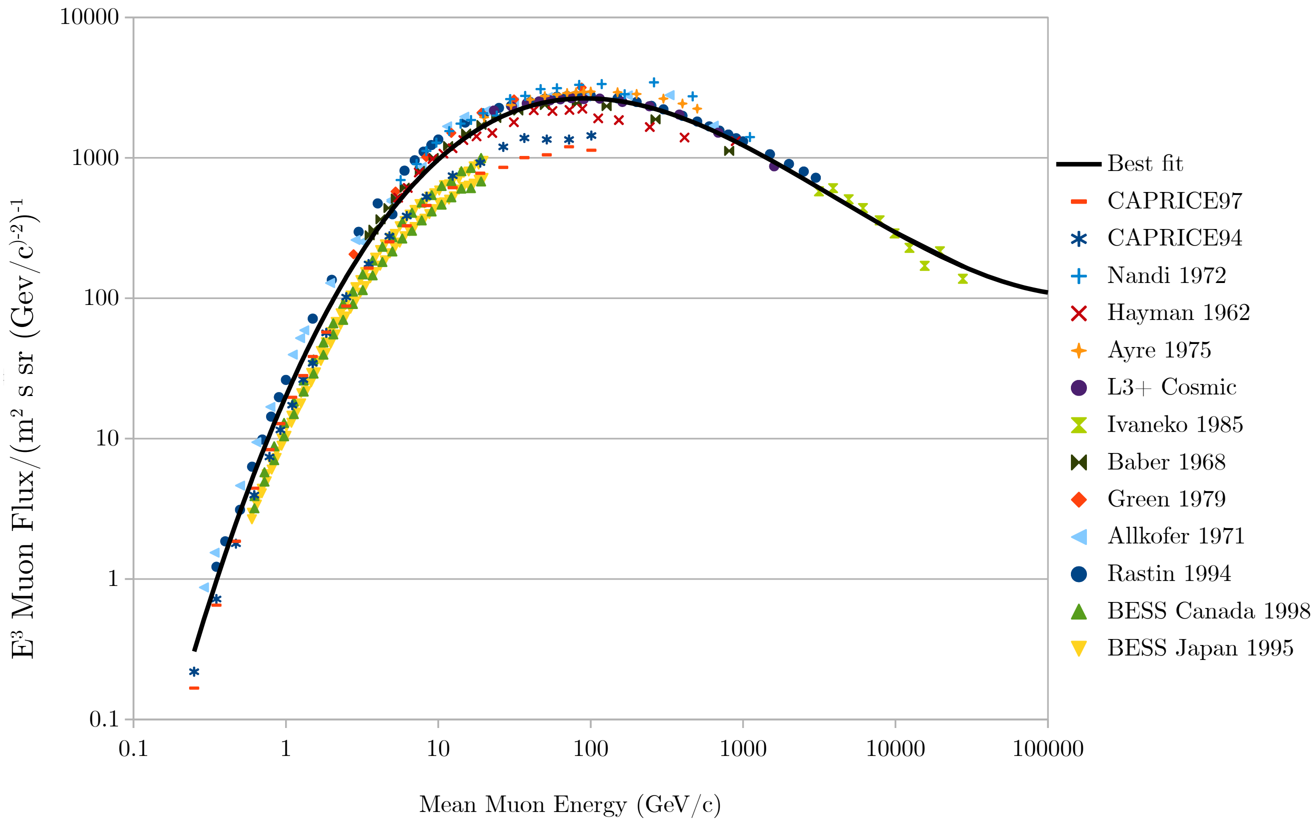

Figure 3: Plot of cosmic ray sea-level data and order 3 polynomial best-fit (excluding both BESS datasets). References for data used and trend-line formula can be found in Appendix B. $R^2 = 0.9978$. Both scales are logarithmic and the y-axis flux has been multiplied by the corresponding muonic energy to the third power, as is done in many publications and textbooks, to make the plot more readable. See Appendix A for best-fir trend-line.

4 Methods/Results

4.1 Requirements for the generator

The end-product generator had design requirements, constrained by its intended usage and purpose. The main points are summarised below:

1. **Be able to generate energies and zenith angles for cosmic rays** - these can be used for other simulations/programs. This is the main aim of the project.

2. **Low computational overhead** – The point of the generator is to be able to input realistic sea-level cosmic rays into a simulation or other program, without having to also simulate the formation of said particle. As such the generator should not have to perform significant computation to output the intended result.

3. **Universality** – The output should be easily ported to the relevant program that the generator is being implemented into, and easily customisable. A future developer should be able to easily transform the output of the generator to suit their needs.

4. **Portability** – The generator would ideally be cross-platform and useable in a wide range of computational environments. Some of the early lessons learned during this project involved the difficulties with distributing and installing software to a variety of environments. Given the relatively small scope and simplicity of this project this should be easily achievable.

Given these main requirements, it was decided that the Python [1] language would be used for the project. It is an interpreted language, so the difficulty of compilation/installation on many different systems would be negligible. Python is very widely used, and a version of Python is pre-installed on most modern Linux distributions. Coupled with the NumPy package, Python becomes a powerful tool for vectorised mathematics [11], and highly universal and easy to manipulate.

4.2 Modes of generation

The formulae given in Equation 4, Equation 5 and Equation 6 provides a useful distribution for certain circumstances. Figure 2 shows its output for $\vartheta = 0$ and a range of given energies. However, it is not appropriate as a probability distribution for all energies as it is clear that the curve plotted does not fit the data for E<200 GeV, with a systematic error above the data. The best-fit formula, plotted on Figure 3 has a relatively low R^2 value of 0.9978 and fits well to the low-energy data. However, the extrapolation of the best-fit to higher energies outside the dataset is unfounded, especially as it is not based on any theoretical background.

Additionally, collection of muon flux data at non-zenith angles is more difficult, and finding enough experimental data to be able to parametrise the muon angular distribution for *all relevant energies* is challenging. Data are often given only for large bins of 10° or larger, and detectors have varying nuance with respect to sensitivity at different angles, etc. Hence, different method of generation are used if the user wants to generate muons with zenith angle $\vartheta \neq 0$.

Note on statistics: it may not be immediately obvious that quantities like dN_μ/dE_μ can be used as probability distributions for muonic energies. However, integrating said equation dE_μ up to some E_0 gives the number of particles produced with energy less than E_0 i.e. $N_\mu(E_\mu < E_0)$. When normalised, this is exactly a Cumulative Distribution Function (CDF) for E_μ; hence the original function (again when normalised) fits the definition of a Probability Distribution Function (PDF) for E_μ. The normalisation will not necessarily be important for our purposes given my implementation of rejection sampling. This also explains why they are not proper probability distributions for ϑ.

4.2.1 $\vartheta = 0$

It seems most sensible to use the theoretical formula at higher energies, where the no-suppression assumption is more valid as the muons have a lower interaction cross-section at more relativistic energies. The best-fit formula will then be used at lower energies. Both functions are plotted on Figure 4. Their coincidence was found numerically to be at 18857 GeV. This, along with the "composite" function, are also plotted in Figure 4. These are evaluated by the `theory()`, `theory_suppressed()`, `best_fit()` and `composite()` functions in the `functions.py` script.

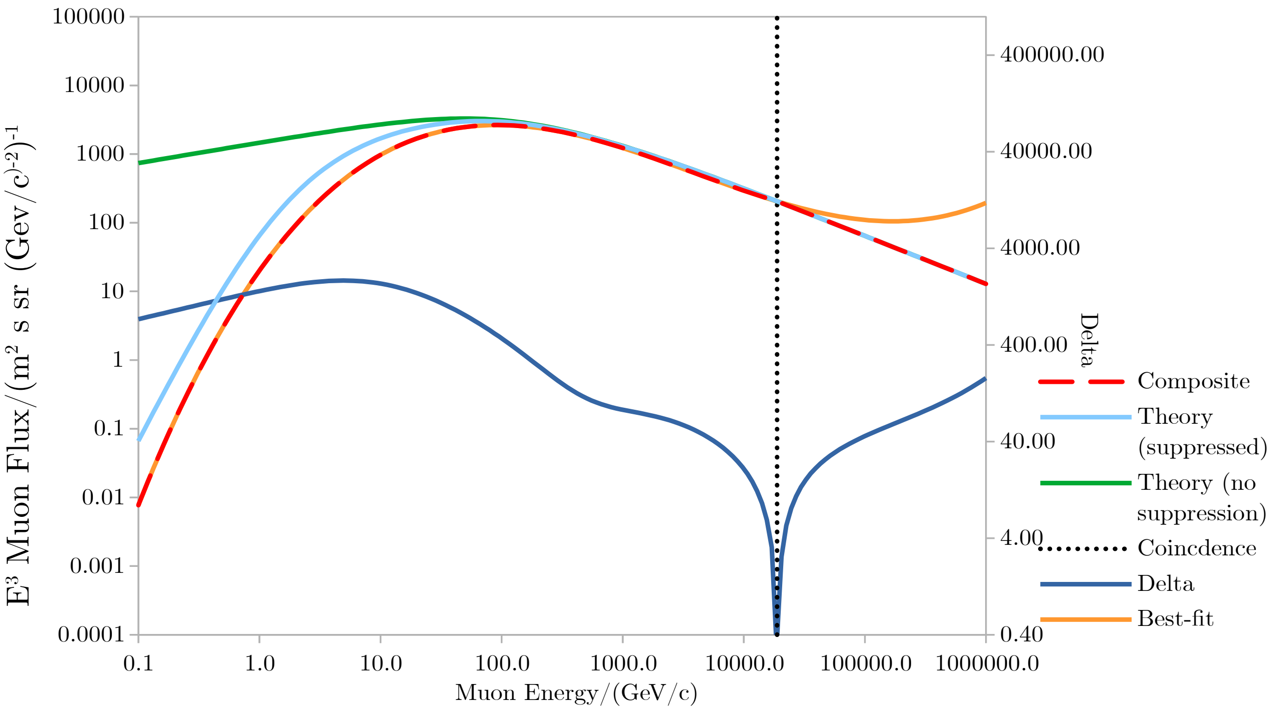

Figure 4: The theoretical (both suppressed and not) and best-fit functions plotted against energy along with the Δ of theory and best-fit. Also shown is the proposed composite function. The composite, theory, theory (suppressed) and best-fit data are simply the outputs of ***`composite()`, `theory()`, `theory_suppressed()` and `best_fit()`*** *in the codebase, with the usual E^3 factor for consistency. The 'coincidence' line shows the point at which the theory and best-fit meet.*

The mean of the `composite()` function is calculated numerically by integration of $E_\mu \cdot \text{composite}(E_\mu)$ between 0 and ∞ (normalised), and is 4.7360 GeV (error of order 10^{-7}). Likewise the mean of `theory_suppressed()` for $\vartheta = 0$ is 3.0887 GeV. As expected this is smaller as the theory PDF is systematically larger than the experimental data best-fit for low energies.

The functions are sampled using a simple rejection sampler. For $\vartheta = 0$ this simply trials the function in a rectangular range between $E_{mu} = 0$ and some upper limit, and the maximum of the distribution, calculated at runtime. This is *extremely* inefficient given that the sampled area for the uniform distribution is so much larger than the area under the curve. This is discussed in 5.2.

Timestamps for muon production can be created by the generator given a horizontal detector area by a simple Poisson distribution, using the value given for overall intensity in 3.6.

4.2.2 $\vartheta \neq 0$

For non-zero zenith angles, we must rely on the theoretical formula as it gives a parametrisation of the muonic flux for any ϑ. The generator generates n random angles according to a $\cos^2\vartheta$ distribution via the angle_generator() function. This is motivated by the theory in 3.4 but also much experimental work, see [8]. The method of generation is then the same as described above except that the `theory_suppressed()` function is used, and the n angles are passed to it in turn. Figure 5 shows histograms for the zenith angles and energies of 20,000 produced muons. The angular distribution follows $\cos^2\vartheta$ as expected. Optimisations are again discussed in 5.2.

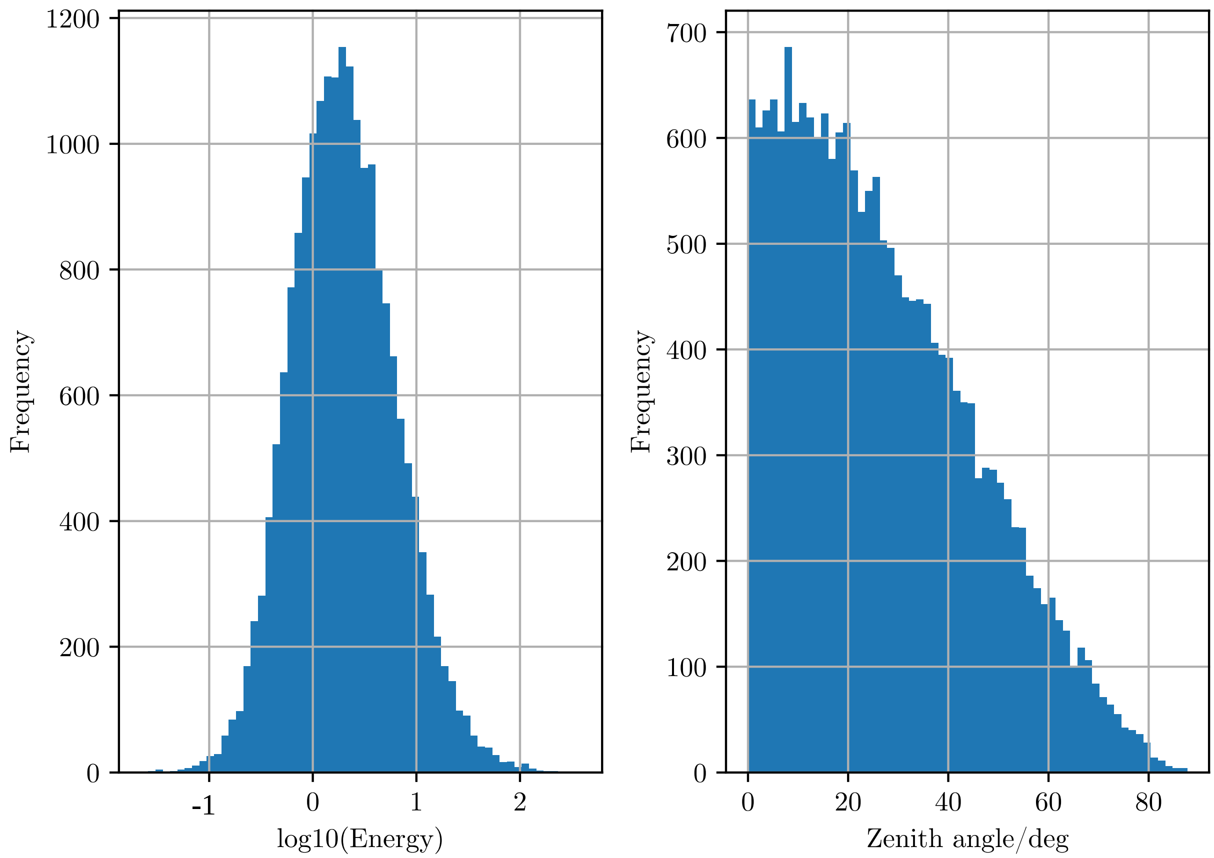

Figure 5: Histograms of log10(Energy/(GeV/c)) and zenith angle produced by the generator. 20,000 muons produced in 345.78s.

4.3 Validating the Results

4.3.1 CORSIKA

CORSIKA (COsmic Ray SImulations for KAskade) is a Monte-Carlo simulator that can be used to simulate primary cosmic ray-initiated air showers [12].

CORSIKA is primarily written in Fortran, with some of the data-processing utility programs written in C++, using their COAST (COrsika data AccesS Tools) package. CORSIKA is a large and comprehensive software package, with many tools and options to allow much versatility. The program distinguishes between 'high' and 'low' energy hadronic-nuclei interaction models to allow different methods to be used for particle interactions at different energies. By default CORSIKA will use the HDPM [13] model at high energies and FLUKA (FLUktuierende KAskade) [14] at low energies, and was the case for the initial simulations. The details of the installation options and simulation parameters are given in Appendix C, but the main points are summarised here.

The program was set to simulate air showers produced by Hydrogen and Helium nuclei arriving at the top of the atmosphere (initially just H, see below), as these are by far the most abundant primary particles. The nuclei are assigned energies according to a power-law curve, using data from Table 1. The atmosphere model used was the US Standard Atmosphere [15]. The particle interactions are then tracked and all particles arriving at the "detector level" (at 110m above sea level) are logged.

Figure 6 shows the muon properties for 577,528 primary proton showers, with energies between 10 and 10^9 GeV producing 257,667 Muons reaching detector level. Importantly, CORSIKA requires a lower cut-off to the energies of the particles being produced. For all simulations this was left at 0.3 GeV, especially given that the generator function `combined()` has a lower cut-off at a higher 0.5 GeV. However, this may skew some of the statistics of the produced particles. This hard cut-off is shown clearly on the left of Figure 6.

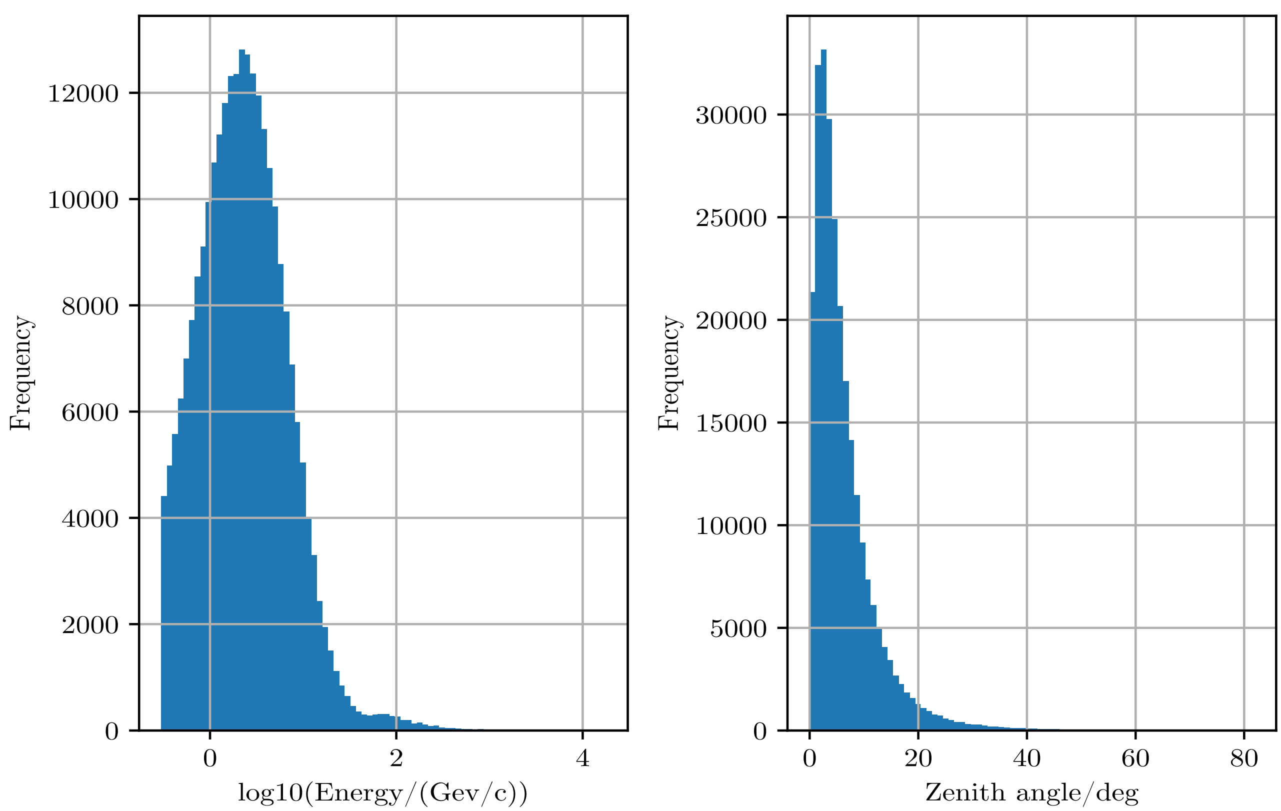

Figure 6: Histograms showing the kinetic energies and angular distribution of muons for the HDPM simulation of . The mean energy was 5.8 ± 6.1 GeV, and the mean zenith angle was 6.63°. Each histogram has 50 bins. Note small secondary peak/tail around 100 GeV.

The mean energy of the muons at near sea level was slightly higher than the literature suggests, and the variance on the mean is very large (5.8 ± 6.1 GeV), which is discussed in

5.1. Given the prediction by the response curve in [6], this mean should not be sensitive to high energy primaries. An increase in maximum primary energy to 10^{12} GeV on re-running the simulation gave a mean E_μ of 5.9 ± 4.2 GeV, which supports this, though the variances are very large. All simulations so far were performed using only H nuclei as primary cosmic ray particles, which could also have affected the mean.

Problems were experienced attempting to use He nuclei as primaries with the simple HDPM high-energy interaction model. Subsequently, the simulation was repeated using different interaction models, specifically DPMJET III [16] for high energy interactions and GHEISHA (as distributed with the Geant4 package [17]) at low energies. Note that for He interactions the superposition model is used, as explained in 3.2.

The histograms for the DPMJET-based simulations appear much the same as those above, for both He and H nucleus primaries. The H nuclei produced a mean muonic energy of 9.8 ± 2.3 GeV while the He produced 10.8 ± 2.2 GeV. These variances are much smaller than before.

The zenith angle distributions don't at all resemble the $\cos^2\vartheta$ distribution that was predicted in 3.4. They are much narrower nearer the zenith. This is discussed in 5.1.

5 Discussion

5.1 Mean energies & simulation models

The mean of both the PDFs used for muon generation and the simulations performed with CORSIKA were slightly different from that quoted in the literature. A review from the Particle Data Group PDG [1] suggests the mean Energy of sea-level muons is approximately 4 GeV. The closest distribution mean to this was the composite() function, with a mean of ~4.7 GeV. The HDPM simulations gave means of between 5 and 6 GeV, with very large variances. This is likely due to the small secondary peak (or large tail) shown on the right of the energy distribution, around 100 GeV; see Figure 6. I do not understand the origin of this peak, as it is not predicted by the theory.

As mentioned in 4.3.1, the high- and low-energy simulation models were changed in order to be able to simulate He primaries. There is some evidence that the simple HDPM model for high energies becomes less applicable at higher energies, such as the erroneous peak observed in Figure 6. This is further illustrated by the data in Figure 7, (taken directly from [14]), which shows the mean energy fraction in a H-N collision that is lost to EM radiation. Clearly, at higher energies, the HDPM model begins to diverge from the others in this metric. I expect that this exact mechanism could also reduce the mean muonic energy as mentioned above, giving lower means.

Hence DPMJET III [12] was chosen instead. It is also able to simulate He primaries, while some issues were experienced attempting this with the HDPM model. The mean muonic energies were larger but the variances were much smaller, likely due to the absence of the secondary peak in the distribution.

The high mean energy is probably explained by the use of only vertical incoming primary particles for the simulation. This was an oversight that should have been altered very early on. This also allows very little interpretation of the angular results of the simulation, and more work could easily be done to validate this aspect of the simulation. One final simulation was performed to attempt to characterise the angular distribution, with incoming particle angles between 0 and 70°, but the angular distributions were very strange and hard to analyse, with strong peaks at specific angles. I believe that only a few incoming showers had muonic components that reached the ground, producing these strong peaks for each shower. Hence more simulations would need to be done to analyse and verify the angular distribution (10,824 showers analysed with over 12hrs of computing time).

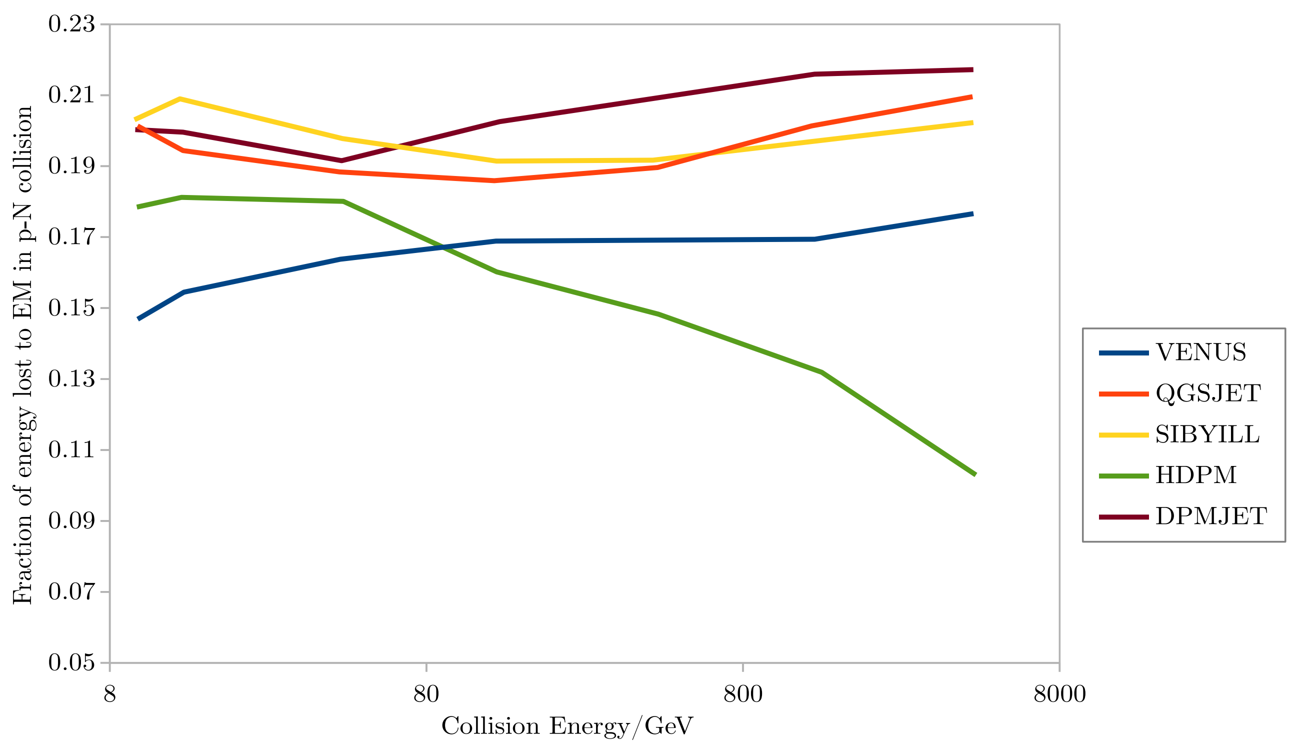

Figure 7: Comparison of the mean Energy fraction lost to photons after a proton-Nitrogen nucleus collision as a function of COM energy. Taken from [14]. HDPM shows significant deviation from the other models at higher energies.

5.2 Optimisation

The generator created is highly unoptimised. Given that the distribution being sampled has the form of a negative power in E_μ, using a uniform distribution to sample it is very inefficient given that the area under the uniform distribution curve is so much larger than that under the PDF curve. This is shown in Figure 8. It would be ideal to use a sampling distribution that fits the data much better to increase efficiency, a simple example of this is also shown on Figure 8. However, performing a simple test revealed that using the uniform distribution or the triangular distribution to sample 10,000 muons with $\vartheta = 0$ and maximum energy 10^4 GeV took 61.8s and 17.5s respectively. I expect that this is due to the algorithm used by NumPy to obtain random values for the triangular distribution, but this could be looked into further. Clearly much better distributions could be sampled, such as some sort of power law, exponential or even a normal distribution if they were scaled correctly. This would be a simple way to optimise the code.

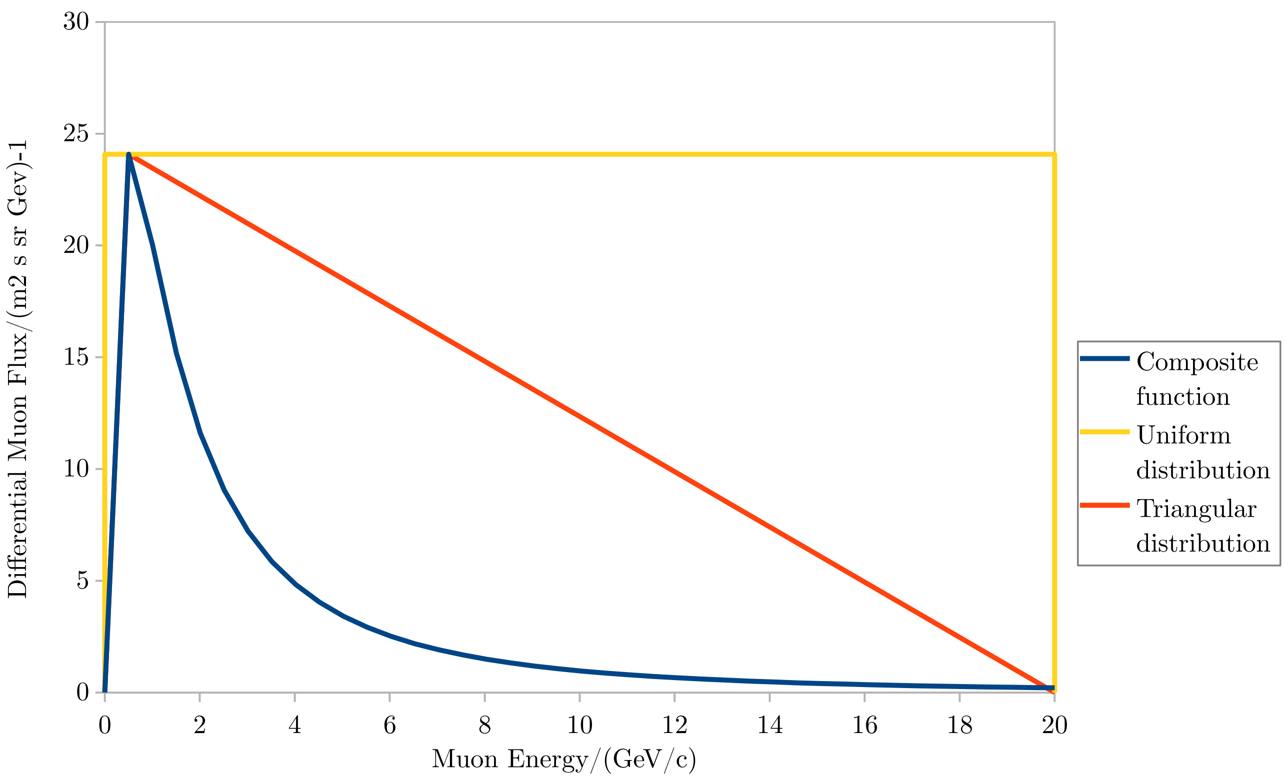

Figure 8: The `composite()` distribution, uniform distribution and triangular distribution against muonic energy for sampling muons with a maximum energy of 20 GeV (in reality this would be much higher).

Additionally, calculating the maximum of each distribution at runtime is computationally inefficient, it would be much more efficient to have the maxima stored as a constant, but this is less versatile with respect to the input function being sampled. I opted to keep the generator as versatile as possible. The maximum of the PDF varies with the ϑ chosen (shown in), so if $\vartheta \neq 0$ is specified then the generator must compute this maximum for each ϑ chosen, which is very time consuming. The maximum shifts away from $E_\mu = 0$ for large ϑ, as shown in Figure 9.

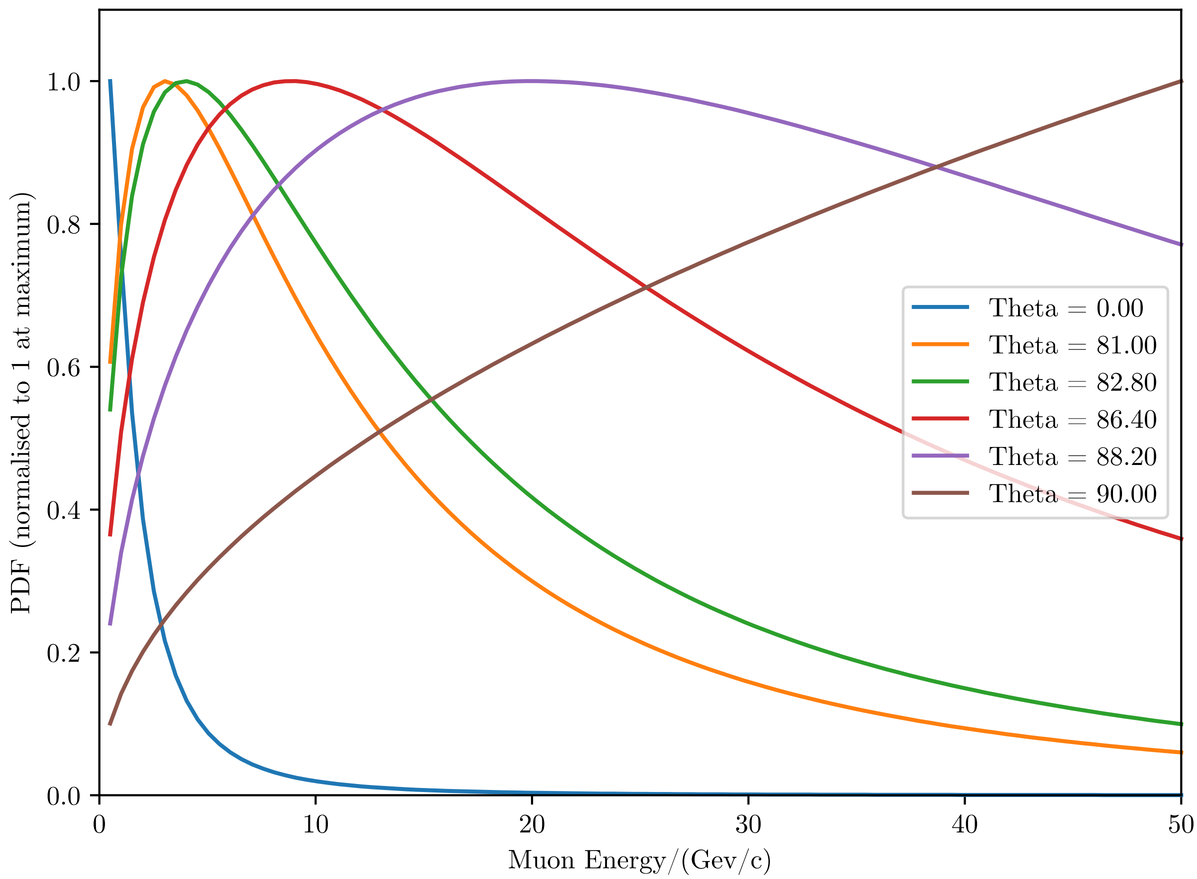

Figure 9: Comparison of PDF for different zenith angles, all normalised to 1 at their respective maxima.

6 Conclusions

As was set out at the start of the project, I wrote a software generator that produces realistic sea-level cosmic rays, in much less time than it would take to directly simulate them. The output of the generator fits well with experimental data (and is somewhat based on it). Simulations were performed to further validate the results and also provided interesting comparisons with the experimental data.

7 Bibliography

[1] Python Software Foundation, "Python Language Reference v 3.10.4." [Online]. Available: http://www.python.org/

[2] P. A. Zyla *et al.*, "Review of particle physics," *Prog. Theor. Exp. Phys.*, vol. 2020, no. 8, pp. 1–2093, 2020, doi: 10.1093/ptep/ptaa104.

[3] R. Hillier, *Gamma ray astronomy.* Clarendon Press, 1984.

[4] R. Engel, D. Heck, and T. Pierog, "Extensive air showers and hadronic interactions at high energy," *Annual Review of Nuclear and Particle Science*, vol. 61. pp. 467–489, 2011. doi: 10.1146/annurev.nucl.012809.104544.

[5] J. R. Hörandel, "Cosmic-ray composition and its relation to shock acceleration by supernova remnants," *Adv. Sp. Res.*, vol. 41, no. 3, pp. 442–463, Jan. 2008, doi: 10.1016/j.asr.2007.06.008.

[6] T. K. Gaisser, R. Engel, and E. Resconi, *Cosmic rays and particle physics.* Cambridge University Press, 2016. doi: 10.1017/CBO9781139192194.

[7] A. A. Abdo *et al.*, "The large-scale cosmic-ray anisotropy as observed with milagro," *Astrophys. J.*, vol. 68, no. 2, pp. 2121–2130, Jun. 2009, doi: 10.1088/0004-637X/698/2/2121.

[8] M. Bahmanabadi, "A method for determining the angular distribution of atmospheric muons using a cosmic ray telescope," *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 916, pp. 1–7, Feb. 2019, doi: 10.1016/J.NIMA.2018.11.010.

[9] T. K. Haldar, M. Mitra, and D. P. Bhattacharyya, "Estimation of sea level muon energy spectrum at high latitude from the latest primary nucleon spectra near the top of the atmosphere," *Czechoslov. J. Phys.*, vol. 53, no. 7, pp. 613–623, 2003, doi: 10.1023/A:1024822518795.

[10] P. K. F. Grieder, *Cosmic Rays at Earth.* Elsevier, 2001. doi: 10.1016/b978-0-444-50710-5.x5000-3.

[11] C. R. Harris *et al.*, "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357–362, Sep. 2020, doi: 10.1038/S41586-020-2649-2.

[12] D. Heck, J. Knapp, J. N. Capdevielle, G. S C Hatz, and T. Thouw, "CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers," 1998.

[13] J. N. C. et Al, "HDPM," *Kernforschungszentrum Karlsruhe*, no. 4998, 1992.

[14] C. Ahdida *et al.*, "New Capabilities of the FLUKA Multi-Purpose Code," *Front. Phys.*, vol. 9, p. 705, Jan. 2022, doi: 10.3389/FPHY.2021.788253/BIBTEX.

[15] N. Center, "US standard atmosphere (1976)," 1992. [Online]. Available: http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19770009539.pdf%5Cnhttp://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:US+Standard+Atmosphere+1976#0%5Cnhttp://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:US+standard+atmosphere+(19

[16] S. Roesler, R. Engel, and J. Ranft, "The Monte Carlo Event Generator DPMJET-III," in *Advanced Monte Carlo for Radiation Physics, Particle Transport Simulation and Applications*, 2001, pp. 1033–1038. doi: 10.1007/978-3-642-18211-2_166.

[17] S. Agostinelli *et al.*, "GEANT4 - A simulation toolkit," *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 506, no. 3, pp. 250–303, Jul. 2003, doi: 10.1016/S0168-9002(03)01368-8.

[18] P. J. Green *et al.*, "Absolute intensities of medium-energy muons in the vertical and at zenith angles 55°-85°," *Phys. Rev. D*, vol. 20, no. 7, pp. 1598–1607, 1979, doi: 10.1103/PHYSREVD.20.1598.

[19] B. C. Rastin, "An accurate measurement of the sea-level muon spectrum within the range 4 to 3000 GeV/c," *J. Phys. G Nucl. Phys.*, vol. 10, no. 11, p. 1609, Nov. 1984, doi: 10.1088/0305-4616/10/11/017.

[20] U. Ivanenko, I.P.; Ivanova, M.A.; Kuzmichev, L.A.; Ilyina, N.P.; Mandritskaya, K.V.; Osipova, E.A.; Rakobolskaya, I.V.; Zatsepin, G.T. (Academy of Sciences, Moscow, "Results of investigation of muon fluxes of superhigh energy cosmic rays with X-ray emulsion chambers," *Ninet. Int. Cosm. Ray Conf. HE Sess. Vol. 8*, 1985.

[21] M. Motoki *et al.*, "Precise measurements of atmospheric muon fluxes with the BESS spectrometer," *Astropart. Phys.*, vol. 19, no. 1, pp. 113–126, Apr. 2003, doi: 10.1016/S0927-6505(02)00195-0.

[22] O. C. Allkofer, K. Carstensen, and W. D. Dau, "The absolute cosmic ray muon spectrum at sea level," *Phys. Lett. B*, vol. 36, no. 4, pp. 425–427, Sep. 1971, doi: 10.1016/0370-2693(71)90741-6.

[23] S. R. Baber, W. F. Nash, and B. C. Rastin, "The momentum spectrum of muons at sea-level in the range 3-1000 GeV/c," *Nucl. Physics, Sect. B*, vol. 4, no. 6, pp. 539–548, Mar. 1968, doi: 10.1016/0550-3213(68)90112-0.

[24] P. Le Coultre, "L3+C - The atmospheric muon momentum spectrum," *Nucl. Phys. B - Proc. Suppl.*, vol. 145, no. 1-3 SPEC. ISS., pp. 136–140, Aug. 2005, doi: 10.1016/J.NUCLPHYSBPS.2005.03.052.

[25] P. J. Hayman and A. W. Wolfendale, "The momentum spectrum of cosmic ray muons near sea level in the momentum range 5-1000 GeV/c," *Proc. Phys. Soc.*, vol. 80, no. 3, pp. 710–728, Sep. 1962, doi: 10.1088/0370-1328/80/3/315.

[26] B. C. Nandi and M. S. Sinha, "The momentum spectrum of muons at sea level in the range 5-1200 GeV/c," *J. Phys. A. Math. Gen.*, vol. 5, no. 9, pp. 1384–1394, Sep. 1972, doi: 10.1088/0305-4470/5/9/011.

[27] J. Kremer *et al.*, "Measurements of Ground-Level Muons at Two Geomagnetic Locations," 1999.

[28] C. A. Ayre, J. M. Baxendale, C. J. Hume, B. C. Nandi, M. G. Thompson, and M. R. Whalley, "Precise measurement of the vertical muon spectrum in the range 20-500 GeV/c," *J. Phys. G Nucl. Phys.*, vol. 1, no. 5, pp. 584–600, Jun. 1975, doi: 10.1088/0305-4616/1/5/010.

[29] W. Mckinney, "Data Structures for Statistical Computing in Python," 2010.

8 Source Code

The files `functions.py` and `rejection_sampler.py` are available from:
`https://github.com/condimentman/cosmic_ray_generator`

Appendix A

Order 3 polynomial best fit rend-line:

$$f(x) = 0.0802084387362498\, x^3 - 0.865540849689325\, x^2 - 0.529814522816601\, x + 1.30273127415449$$

with f(x) being the $\log_{10}$(Flux) and $x = \log_{10}(E_\mu)$

Appendix B

Data used for muon vertical flux plot.

- Green 1979 - Measurments of Muons made at a variety of spectrometer zenith angles, with energies between 2 – 300 GeV/c. [18]
- Rastin 1984 – “near-vertical” direction, measuring Muons with energies from 4 to 3000 GeV/c. [19]
- Ivananko 1985 - [20]
- BESS 1995– Use of the BESS (Balloon-Borne Experiment with a Superconducting Spectrometer) at two near-sea level locations (after use on balloon in 1993), with measurements taken from 0.6 to 20 GeV/c [21]
- Alkofer 1971 - [22]
- Baber 1968 – [23]
- L3+ Cosmics – CERN underground experiment that used simulations to calculate surface muon spectrum from underground data [24]
- Hayman 1962 [25]
- Nandi 1972 [26]
- CAPRICE97 & CAPRICE98 [27]
- Ayre 1975 [28]

Appendix C

Details of the usage of CORSIKA. The default installation options were used, except for the switch to DPMJET and GHEISHA models, and the usage of the "INCLINED" option, which allows the user to select an inclined detector array. This was not used but is required at installation to also install the COAST package for analysis of data.

A C++ script `CorsikaRead` was used to extract the detected-particle data from the output. `CorsikaRead` was a slightly modified version of an example script provided by the CORSIKA authors. The raw data is written to binary output files which mist be translated before analysis.

A Python script `postprocess.py` was then used to extract the desired particles' (in this case muons') properties into a Pandas [29] `DataFrame` and calculate some relevant parameters (e.g. Zenith angle). This data is then saved as a `.csv` file for plotting or analysis.

Use of the PARALLEL option may have been useful in speeding up high energy simulations, as it can allocate sub-showers between different CPU cores, but this only allows one shower to be simulated per-run, which was unideal for this use case. An example of the input file used for a CORSIKA simulation can be found with the source code.