# Correlating Primary Type with the Longitudinal Profile

**Paul T. Mikulski**
*Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA*

**Abstract**

The features of an air shower generator which simulates longitudinal profiles are described. The generator captures basic interaction features and contains a single adjustable parameter. It has been used to create sets of proton and oxygen induced showers at $E_0 = 3 \times 10^{20} eV$. These showers are investigated using a neural network in an effort to understand to what degree intrinsic fluctuations affect the correlation between primary type and the longitudinal profile.

## 1 Introduction:

Future experiments, OWL (Streitmatter et al., 1998) and Airwatch (Marzo et al., 1998) will use the air fluorescence technique to obtain a high statistics observation of the highest energy cosmic rays by monitoring the atmosphere from orbitting detectors. If such an observation is achieved, it may be possible to infer not only an energy spectrum but also information about composition. Simulation of longitudinal profiles will play a central role in the interpretation of detected showers.

The traditional approach to the simulation of longitudinal profiles (Gaisser et al., 1993) describes a shower as a sum of subshowers all with energies less than a prescribed simulation threshold. These subshowers are described by parameterizations which have been built up in a bootstrap approach starting at low energy. Section 2 describes an air shower generator which utilizes this approach to construct parameterizations for a set of models which can be interpolated between through the adjustment of a single parameter. Section 3 compares proton and oxygen showers generated at $E_0 = 3 \times 10^{20} eV$ where the models have been tuned to give the same average depth at maximum for both primary types.

## 2 The Air Shower Generator:

The free parameter which fixes the interaction model specifies a choice of extrapolation for the inelastic p-p cross section. So that the user of the generator can think in terms of an effective model, this parameter has been recast as the average depth at maximum for protons at $E_0 = 3 \times 10^{20} eV$. What follows is a brief sketch of main features. Many rough approximations are made, but care is taken so that fluctuations are not underestimated. Though simulation of showers at the highest energies requires extrapolation well beyond energies achieved at accelerators, the longitudinal profile is only sensitive to grosse features of hadronic interactions. Assuming that the physics evolves smoothly as the energy is raised, it is unlikely that a different event generator would imply sets of showers which differ appreciably from an effective model which produces the same average depth at maximum for showers initiated by protons at the energy scale of interest. For example, compare the analyses of the mass composition for Fly's Eye data in (Gaisser, 1993) and (Ding, 1997).

**2.1 Particle Types:** Since the aim is to capture general features, the only particle types treated are nucleons, pions, electrons and photons (the energy loss due to muons and neutrinos is also tracked). So that all subshowers have energies near the simulation threshold, a few pseudo-particles which represent collections of pions are also utilized. In the context of the Hillas splitting algorithm (Hillas, 1981), these represent fragments which have yet to go through presplittings and fragments that are in some stage of hadronization (hereafter referred to as *presplits* and *fragments* respectively).

**2.2 Cross Sections:** The choice of extrapolation of the inelastic p-p cross section is the model's adjustable parameter, though all choices have the same low energy limit. The cross section for pions is taken to be two thirds of the p-p cross section. The parameterization due to Kopeliovich (1989) is used to relate cross sections on nucleon targets to cross sections on air. Nucleus-air cross sections are given by a standard geometric parameterization (Westfall et al., 1979), $\sigma_{A_1A_2} = \pi R_0^2 (A_1^{1/3} + A_2^{1/3} - \delta)^2$ where $R_0 = 1.47 fm$ and the value of $\delta$ is energy dependent to roughly account for the growth in the underlying nucleon-nucleon cross sections.

**2.3 Hadron-Nucleon Interactions:** Hadron-nucleon collisions are generated in a manner similar to the original statement of Hillas' splitting algorithm though the number of presplittings is chosen from a poisson distribution. The leading particle is sampled from a flat distribution, but prior to sampling a fraction of the energy representing a valence quark fragment which is not part of the leading hadron is removed by sampling the fractional energy distribution $p(x) \propto (1-x)^{2n_s - 1}/\sqrt{x}$ where $n_s$ is one less than the number of valence quarks in the projectile. The implementation is rough, but it gives a natural distinction between nucleons and pions since these hadrons differ in the number of valence quarks. This method is similar to that employed in Sibyll (Fletcher et al., 1994).

**2.4 Nuclear Target Effects:** The wounded nucleon picture is adopted. Given the cross sections on nucleon and air targets, the number of participating taget nucleons is sampled in a manner which produces the correct average, $N_W = A\sigma_{pp}^{inel}/\sigma_{p-air}^{inel}$, and also emphasizes peripheral collisions. This is accompished by sampling an intermediate average from the distribution $p(x)dx \propto (1-x)^{\alpha}$ where alpha is tuned to give the overall average. This intermediate average is used in sampling a binomial distribution with a maximum of 13 (taking a nitrogen target and always assuming at least 1 participant). For each participating target nucleon a presplit is generated. The last target nucleon is treated as in hadron-nucleon interactons. The distribution from which the presplit energies are sampled is a function of the number of participating target nucleons, and guarrantees that on average the leading hadron has an energy of the order of the presplit energies. This insures that the secondaries derived from the presplits have characteristic energies much less than the leading particle energy. This reflects the expectation that lead particles reflect valence constituents while secondaries reflect the sea.

**2.5 Nuclear projectiles:** The semi-superposition model is adopted. Using the same methodology as for target nucleons, the sampling of participants from the projectile implies the correct average while emphasizing peripheral collisions. Each participating nucleon from the projectile generates an independent nucleon-air interaction. All spectating nucleons from the projectile are lumped together as a single nucleus. This gives slightly larger fluctuations in shower development, but is fairly close to the more realistic picture where spectators group into fragments and individual nucleons (Engel et al., 1992).

**2.6 Electromagnetic Cascading:** Standard Bethe-Heitler formulas are used to propagate photons and generate electron-positron pairs, see (Gaisser, 1992). However in order to avoid the intricacies in handling the infrared divergence associated with bremsstrahlung, a simple splitting model is used to handle the cascading of electrons (the term also implies positrons). The electron is treated as an effective electromagnetic particle which propagates and splits (using a flat distribution) into two. To imply the correct elongation rate, the mean free path is taken to be half a radiation length. The LPM effect is very roughly implemented by simply taking the radiation length to be proportional to $\sqrt{E}$ above $E_{LPM} = 117 PeV (A_0/A)$, where $A$ is the column depth and $A_0$ is sea level, $1030 g/cm^2$, (Klein, 1998). A characteristic depth taken to be the creation depth of the lepton is used. This captures the main feature of introducing a long length scale at the highest energies.

**2.7 Parameterizing Subshowers:** Subthreshold particles are propagated to a point of interaction before its subshower is parameterized. All hadrons (including presplits and fragments) are parameterized using the Gaisser-Hillas profile function (Gaisser, Hillas, 1977) where $X_0$ is taken to be the depth of the initiating particle. This leaves three parameters($X_{max} - X_0$, $N_{max}$, and $\lambda$) which characterize the average profile of the subshower. These parameters vary as a function of model, particle type and energy. Neutral pions, photons and electrons are parameterized by a modified Greisen formula (Greisen, 1956). The modifications can consist of shifts in $X_0$ and $X_{max}$. The shifts are fitted to be consistent with the standard Greisen formula which describes photon initiated showers relative to the point of the photon's creation. Monte Carlo simulation is always conducted down to at least $E = 117PeV$, so that the LPM effect need not be considered when parameterizing electromagnetic cascades.

**2.8 Fitting of Showers:** To a good approximation, the area underneath the longitudinal profile is proportional to the energy deposited into the electromagnetic cascade (Sokolsky, 1989). This relation is assumed to hold exactly and is implemented by constructing subthreshold parameterizations which respect this. The deposited energy of the total shower is tracked along with the shower size at a number of discrete depths. When fitting the total shower, the deposited energy can be used to fix the area underneath the fitted profile.

# 3 Comparison of Proton and Oxygen Initiated Showers:

A total of 2000 proton showers and 2000 oxygen showers have been generated at $E_0 = 3 \times 10^{20} eV$. The oxygen showers were generated using a fairly penetrating model (average depth at maximum of $960 g/cm^2$ for protons). The average depth at maximum for the oxygen showers ($829 g/cm^2$), was used to fix the model for the set of proton showers. The models conservatively bracket current models in use: the model for the oxygen primaries implies showers which are probably too penetrating while the model for protons implies showers which are probably not penetrating enough. The upper graph of Figure 1 shows resulting $X_{max}$ distribution for the two sets of showers. The distribution indicates that $X_{max}$ can be used to identify primary type with about 60% efficiency.

To see if information about the shape of the profile can be used to increase the correct classification rate, a three-layered feed-forward neural network has been used to analyze the showers (Bishop, 1997). Input to the network consists of the three fitted parameters to the Gaisser-Hillas profile function ($X_{max}$, $X_0$ and $N_{max}$) and an error estimate which represents the degree to which the simulated profile and the fitted profile do not overlap. The fitted value of $X_0$ is only weakly correlated with the actual value of $X_0$, and in fact the fitted value of $X_0$ is often a number less than zero. This is a reflection of the limitations of the fitting function which was developed based on simulations of proton initiated showers using a scaling hadronic interaction model. So that the network does not pick up on the slight systematic

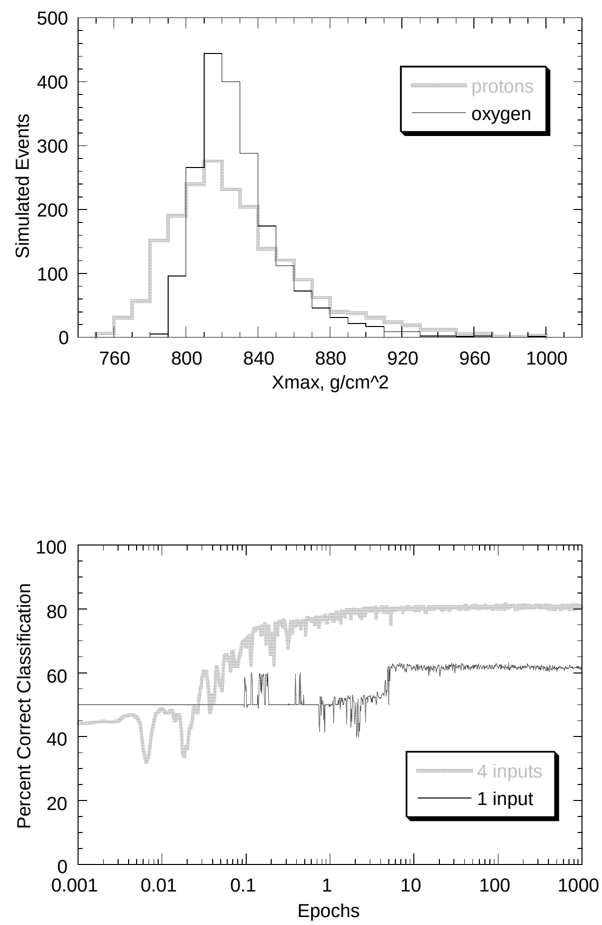


**Figure 1:** Upper graph: The $X_{max}$ distribution for the 2000 proton showers and 2000 oxygen showers, all at $E_0 = 3 \times 10^20 eV$. Lower graph: The percentage of showers correctly identified as proton or oxygen for the two trained networks. An epoch corresponds to the presentation of 2000 randomly chosen profiles from the training set.

differences in the deposited energy between protons and oxygen, $N_{max}$ is the size at maximum after the profile has been scaled to unit area. The shower sets were split equally into a training set and a test set, and the network was trained using a stochastic gradient descent. For means of comparision a separate network was also trained using only $X_{max}$ as an input.

The lower graph of Figure 1 shows the correct classification rate as the networks are trained. This rate is that obtained by presenting the entire test set to the network. A single epoch corresponds to the presentation of 2000 randomly picked profiles from the training set. As expected, the network trained with just $X_{max}$ correctly classifies about $60\%$ while the rate for the four-input network is about $80\%$. While reconstruction errors for actual events are quite large, this simple investigation suggests that it may be useful to classify individual events in a manner which also accounts for the shape of the profile when trying to infer information about the mass composition. Extensive investigations need to be done with detector Monte Carlos to understand exactly how much information can be meaningfully extracted from a detected shower and how to best analyze sets of events. The present shower generator can be a useful tool in exploring this question.

## References


Bishop, C. M., 1997, Neural Networks for Pattern Recognition (Oxford: Oxford University Press)

Ding, L. K. et al., 1997, ApJ. 474, 490

Engel J. el al., Phys. Rev. D46, 5013

Fletcher, R. S. et al, 1994, Phys. Rev. D50, 5711

Gaisser, T. K. et al., 1993, Phys. Rev. D50, 1919

Gaisser, T. K., 1992, Cosmic Rays and Particle Physics (Cambridge: Cambridge University Press)// Gaisser, T. K., Hillas, A.M., 1977, Proc. 15th Int. Cosmic Ray Conf. 8, 353

Greisen, K., 1956, Prog. Cosmic Ray Physics 3, 1

Hillas, A. M., 1981, Proc. 17th Int. Cosmic Ray Conf. 8, 193

Klein, S. R., 1998, AIP Conference Proceedings 433, 132

Kopeliovich, B. Z., Nikolaev, N. N., Potashnikova, I. K., Marzo, C. N. et. al., 1998, AIP Conference Proceedings 433, 87

Sokolsky, P., 1989, Introduction to Ultrahigh Energy Cosmic Ray Physics (Redwood City: Addison-Wesley)

Streitmatter R. E. et. al., 1998, AIP Conference Proceedings 433, 95

1989, Phys. Rev. D39, 769

Westfall, G. D. et al., 1979, Phys. Rev. C19, 1309