The influence of weather effects on the reconstruction of extensive air showers at the Pierre Auger Observatory

Alan Coleman[*a] **for the Pierre Auger Collaboration**[b]

[a]*Penn State Physics Department, State College, USA*
[b]*Observatorio Pierre Auger, Av. San Martín Norte 304, 5613 Malargüe, Argentina*
E-mail: auger_spokespersons@fnal.gov
Full author list: http://www.auger.org/archive/authors_icrc_2017.html

The extensive air showers created by highly energetic cosmic rays are measured at the Pierre Auger Observatory. The development of these air showers in the dynamic medium of the Earth's atmosphere affects the reconstruction and ultimately the determination of the energy for the primary particle. We present an analysis using data from the two surface detector arrays (with 750 m and 1500 m spacing) which studies the modulation of the reconstructed energy estimator due to changes in atmospheric pressure and density. These dependencies on local weather are expected to affect the cosmic ray energy measurements by $\sim$0.5% on average. While this is a rather small effect, not accounting for weather effects can introduce biases in event rates over a sidereal day and must be corrected to perform cleaner anisotropy analyses, in particular at large angular scales. Finally, a method by which these modulations can be corrected is detailed for the 750 m and 1500 m arrays.

---

[*]Speaker.

1. Introduction

The development of cosmic ray air showers occurs in the dynamic medium of Earth's atmosphere. The atmospheric properties affect this development which can be seen via changes in the measured signal due to secondary particles on the ground. Due to the daily and yearly cycles in Earth's weather, the changes in atmospheric conditions can produce biases in the measured air shower signals. These effects must be corrected when attempting to measure cosmic ray anisotropy and determine the energy flux spectrum.

We present an update to a previous work [1] which developed a method by which the small changes in signal due to atmospheric fluctuations can be corrected. This method identifies biases in the measured signal as a function of air density, ρ, and pressure, P. This update includes seven more years of Pierre Auger data from the 1500 m surface detector (SD) as well as measurements using the smaller 750 m SD array. The method is also improved by including a timing delay in the model to account for the difference in atmospheric conditions in the last two radiation lengths above the detector.

The Pierre Auger Observatory and data sets used in this study are described in section 2. The method to develop the updated weather correction is detailed in section 3. Finally we explain how the atmospheric model will affect the SD reconstruction of air showers in section 4.

2. Observatory and data set

The Pierre Auger Observatory [2] is a hybrid cosmic ray detector located in the Mendoza province of Argentina. The Observatory, situated around 1400 m above the sea level, employs a fluorescence detector (FD) which overlooks two nested surface detector (SD) arrays. Together, the SD and FD are able to observe the fluorescence light given off by developing air showers as well as their particle signature on the ground.

Each SD consists of a hexagonal lattice of water Cherenkov detectors which measure the energy given off by secondary particles entering the water volume. Physically, the two SD arrays differ only in the lattice spacing, 1500 m (750 m), between stations (since many values for the two SD arrays will be presented in parallel, this notation will be used for the rest of the paper). The separation between stations defines the characteristics of cosmic rays that can be observed. See Table 1 for a comparison of the physics capabilities of the two SD arrays.

2.1 Energy measurements

Estimating the energy of an air shower using an SD begins with a measurement of the signal, S, deposited in the SD stations on the ground. The first-order energy estimator, $S(r_{opt})$, is found by fitting the lateral distribution of signal amplitudes from the shower axis. The value of $S(r_{opt})$ is given by the expected signal amplitude at an optimal distance, r_{opt}. This distance has been specifically chosen to minimize the impact of the empirically chosen lateral distribution model on the measurement of $S(r_{opt})$. For the 1500 m (750 m) array, r_{opt} is 1000 m (450 m).

A correction is then applied to remove zenith angle biases using a constant intensity cut (CIC) [3]. This correction has the geometric motivation that an increase in zenith angle leads to increase in average slant depth for showers at ground level. Thus, highly inclined showers reach the detector

	1500 m Array	750 m Array
Total Stations	1600	61
Lattice Cells	1380	42
Detector Area	$3000\,\mathrm{km}^2$	$24\,\mathrm{km}^2$
Fully Efficient Energy Threshold	3 EeV	0.3 EeV
Maximum Zenith Angle Considered	$60°$	$55°$
Signal Optimal Distance (r_{opt})	1000 m	450 m
Dataset Events	$1,146,481$	$570,123$
Dataset Energy Threshold	1.0 EeV	0.1 EeV
Dataset Median Energy	1.5 EeV	0.15 EeV
Dates Used	$01.01.2005 - 31.12.2015$	$01.01.2011 - 31.12.2015$

Table 1: The Pierre Auger Observatory includes two SD arrays to measure cosmic ray air showers. The top section of the table lists differences in physical properties and physics capabilities of the two arrays. The bottom section details each array's data set used in this work.

with more attenuation which produces a zenith angle dependent bias when considering events above a given $S(r_{\mathrm{opt}})$. This bias is removed by scaling $S(r_{\mathrm{opt}})$ to $\hat{S}$, the value that would have been measured for a shower arriving at a reference zenith angle of $38°$ ($35°$). See [3] for a more detailed explanation.

The corrected energy estimator, $\hat{S}$, is then converted to a shower energy using a calibration between the FD and the SD arrays. The FD has the benefit of almost calorimeterically measuring the electromagnetic energy of the shower. Thus, the calibration between $\hat{S}$ and the shower energy is done by using the common set of high quality events that can be independently reconstructed by the FD and the SD arrays. For each respective SD array, $\hat{S}$ and the shower energy, E, are related via a power law $E = A\hat{S}^B$ (see [3] for more information).

The FD calibration also highlights one of the motivations for a weather correction. As will be shown later in this section, the atmospheric density and pressure have daily cycles. The FD only records data at night and thus the events used in the calibration of the SD are inherently biased. The weather correction will unbias the data before the FD calibration and will produce more accurate SD energy measurements.

2.2 The event set

The data set for this work was chosen to start when the construction of the respective SD array was approximately complete. Since construction of the 750 m array began much later than for the 1500 m array, the data periods cover different intervals of 10 (5) years.

To ensure a high quality measurement of the $S(r_{\mathrm{opt}})$ values, the data set uses a fiducial cut which requires that the station with the largest signal has all six nearest neighbors active at the time of the event. This cut effectively demands that the shower core is contained within the array and that the lateral distribution of particles is well sampled. A further cut in zenith angle was applied to
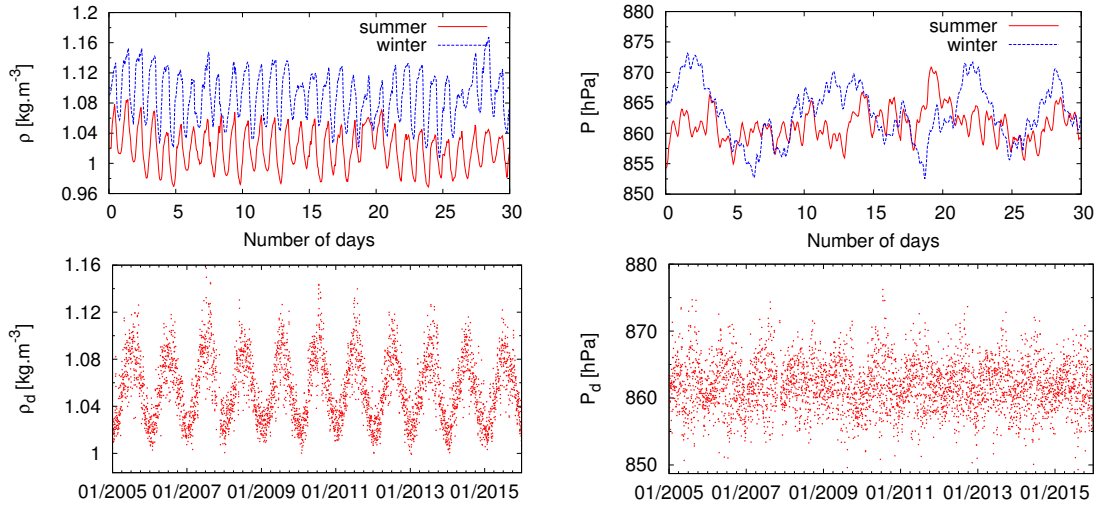
Figure 1: The weather data used in this study includes measurements of air density (left) and pressure (right). The top two plots show the hourly measurements over the course of one summer (solid red) and winter (dashed blue) month. Daily and yearly cycles can be see most clearly in the density measurement with amplitudes of 3% and 6%, respectively. The air pressure does not have such strong yearly modulations though daily cycles are apparent.

the two arrays to ensure full efficiency, zenith angle $< 60°$ ($55°$). Further, only data above $1\,\text{EeV}$ ($0.1\,\text{EeV}$) were used for this analysis. This event selection results in $1,146,481$ ($570,123$) events.

2.3 Atmospheric Data Set

The atmospheric measurements for this work were recorded using the Observatory's weather monitoring stations. The primary atmospheric monitoring station used in this study is located roughly in the center of the 1500 m array near the central laser facility (CLF). The CLF weather station records measurements of the atmospheric temperature and pressure every 5 minutes. Occasionally, there are gaps in these measurements. When there are missing data for intervals between 10 minutes and 3 hours, the values are interpolated. For longer intervals, the atmospheric data from the weather stations at the FD sites are used[1] after correcting for the difference in altitude. The weather stations' measurements of temperature, T, and pressure, P, are used to calculate the air density, ρ, using the dry-air relation $\rho = (M/R)(P/T)$ where $M/R = 0.3484 \frac{\text{kg}}{\text{m}^3}\frac{\text{K}}{\text{hPa}}$.

The top plots in Figure 1 show measurements of density (left) and pressure (right). They include measurements taken every five minutes over the course of one summer (solid red) and winter (blue dashed) month. The bottom two plots show the density and pressure daily-averages over the course of the 10 years. The strong yearly and daily modulations in density can be clearly seen with an amplitude of around 6% and 3%, respectively. Such modulations can also be seen in the pressure over the course of a year though the spread from day to day is not as large in amplitude. We also note that the fluctuations in pressure are much higher in winter months than they are in summer months.

---

[1] FD weather measurements make up $\sim$7% of the data set

3. Correcting for atmospheric conditions

We have shown in a previous work [1] that it is the atmospheric conditions in the last two radiation lengths above ground which most strongly correlate to changes in the measured signal. At these heights, up to one km above the detector, temperature fluctuations can be two to three times smaller than at the ground level. Thus, the ideal measurements to be used in a weather correction would be taken high over the SD array.

Instead of having to make measurements at these heights we note that there is an approximate 2 hour delay in the atmospheric conditions one km above the detector and at ground level. Thus, we propose a model which is based on the current pressure, P, the average density ± 12 hours of the event, ρ_d, and the density two hours previous $(\widetilde{\rho})$

$$S(r_{\text{opt}}) = S_0 \left[1 + \alpha_P(P - P_0) + \alpha_\rho(\rho_d - \rho_0) + \beta_\rho(\widetilde{\rho} - \rho_d)\right]. \tag{3.1}$$

Here $S(r_{\text{opt}})$ is the measured reference signal (see section 2.1) and S_0 is the signal that would have been measured at the reference weather conditions, P_0 and ρ_0. These values are defined to be the data set's yearly averages, $P_0 = 862\,\text{hPa}$ and $\rho_0 = 1.06\,\text{kg m}^{-3}$.

3.1 Cosmic ray arrival rate

The atmospheric conditions' impact on the arriving cosmic rays can be seen when considering the event rate above a given signal threshold, S_{cut}. Due to the modulation of measured signal by the changes in weather, some events will migrate above or below S_{cut}. Thus, the modulation can be found by looking at the event rate of observed air showers with $S(r_{\text{opt}}) > S_{\text{cut}}$ as a function of atmospheric conditions. The differential rate per unit area is given by

$$\frac{dR}{d\theta} = 2\pi \sin\theta \cos\theta \int_{S_{\text{cut}}}^{\infty} dS\, P_{\text{tr}}(S, \theta) \frac{d\Phi_{\text{CR}}}{dE_t} \frac{dE_t}{dS} \tag{3.2}$$

where P_{tr} is the function describing the trigger efficiency. The energy flux spectrum is described by a power law $d\Phi_{\text{CR}}/dE_t \propto E_t^{-\gamma}$ where E_t is the true cosmic ray energy. Using the FD calibration equation relating measured signal to this true energy, $E_t \simeq AS^B$, this can be rewritten to leading order in the weather corrections as

$$\frac{dR}{d\sin^2\theta} \propto \left[1 + a_P(P - P_0) + a_\rho(\rho_d - \rho_0) + b_\rho(\widetilde{\rho} - \rho_d)\right] \int_{S_{\text{cut}}}^{\infty} dS\, P_{\text{tr}}(S, \theta) S^{-B\gamma + B - 1}. \tag{3.3}$$

Here the weather coefficients have the relation $a_P = B(\gamma - 1)\alpha_P$ (and equivalently for α_ρ and β_ρ). Using the FD calibration constant $B = 1.023 \pm 0.006$ and the measured spectral index of the cosmic ray flux $\gamma = 3.29$ [4], we simplify the relationship to $a_P \simeq 2.3\alpha_P$.

3.2 Determination of atmospheric coefficients

To find the value of the atmospheric constants $\{a_P, a_\rho, b_\rho\}$, the data was split into hourly time bins. Using equation (3.3) for the arrival rate, the expected number of counts in bin, i, can be written as

$$\mu_i = R_0 A_i \left[1 + a_P(P_i - P_0) + a_\rho(\rho_{d,i} - \rho_0) + b_\rho(\widetilde{\rho}_i - \rho_d)\right] \tag{3.4}$$

Array	a_P [hPa^{-1}]	a_ρ [kg^{-1}m^3]	b_ρ [kg^{-1}m^3]	χ^2/dof
1500 m	$(-3.2 \pm 0.3) \times 10^{-3}$	-1.72 ± 0.04	-0.53 ± 0.04	1.013
750 m	$(-4.9 \pm 0.4) \times 10^{-3}$	-1.07 ± 0.06	-0.37 ± 0.06	0.998

Table 2: This table gives the values of the weather correction variables corresponding to equation (3.4) for the 750 m and 1500 m arrays.
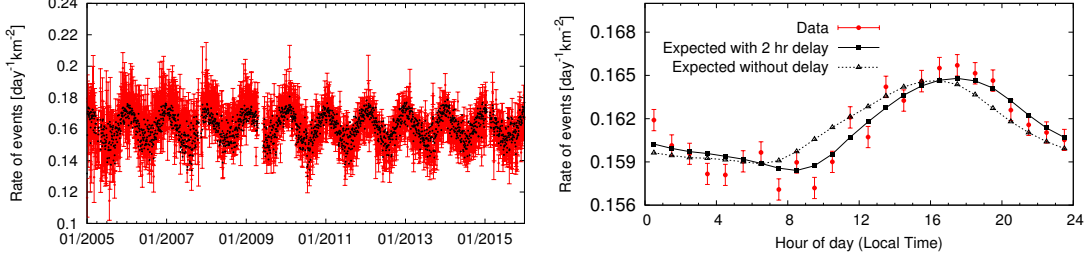

Figure 2: The measured and expected event rates for the 1500 m array are shown here. The left plot includes the measured (red) and expected (black) daily event rate over the 10 year data period. In the right plot, the event rates over the course of one day are shown. The measured rate (red) is shown along with the expected rate when using (black squares) and not using (grey triangles) the two hour delay in air density.

where R_0 is the global rate per area that would be observed at the reference atmospheric conditions and A_i is the combined area covered by active cells in the i-th time bin[2]. The optimal atmospheric constants are thus the values that maximize the likelihood function

$$\mathscr{L} = \prod_i \frac{\mu_i^{n_i} e^{-\mu_i}}{n_i!} \tag{3.5}$$

where n_i is the actual number of observed events in bin i compared to the expected number μ_i given in equation (3.4).

3.3 Results for the SD Arrays

The fit of the atmospheric constants for the two arrays are given in Table 2. A chi-square test shows that the model fits the data well for both arrays. The number of degrees of freedom for the fits was 88,126 (39,258). As an example, the measured and expected event rates for the 1500 m array are shown in Figure 2 (the 750 m array results look similar). The daily event rates (red) and the expected rates (black) spanning ten years are shown in the left plot. The right plot includes the average rates over the course of one day (red circles). The expected values have also been plotted when the two hour delay is both included (black squares) and neglected (empty triangles). The two hour delay seems to be justified both visually and via a reduced χ^2 goodness of fit test, changing from 4.2 to 1.9 when the air density is shifted by two hours.

---

[2]Each hexagonal cell contributes $\sqrt{3/4}d^2$ where d is the lattice spacing.

Array	Coefficient	c_0	c_1	c_2
1500 m	a_P [hPa^{-1}]	$(2.1 \pm 0.9) \times 10^{-3}$	$(-2.6 \pm 0.6) \times 10^{-2}$	$(2.6 \pm 0.7) \times 10^{-2}$
	a_ρ [kg^{-1}m^3]	-2.7 ± 0.1	1.5 ± 0.8	2.2 ± 1.0
	b_ρ [kg^{-1}m^3]	-1.0 ± 0.1	1.2 ± 0.8	0.0 ± 1.1
750 m	a_P [hPa^{-1}]	$(-2.5 \pm 0.8) \times 10^{-3}$	$(-0.8 \pm 0.2) \times 10^{-2}$	-
	a_ρ [kg^{-1}m^3]	-1.6 ± 0.1	1.8 ± 0.3	-
	b_ρ [kg^{-1}m^3]	-0.4 ± 0.1	0.1 ± 0.3	-

Table 3: The weather corrections were carried out in different zenith angle bins. The resulting coefficients were then fit to a polynomial as described in equation (3.6). Since the 750 m array was split into only three zenith angle bins, it was fit to a line.
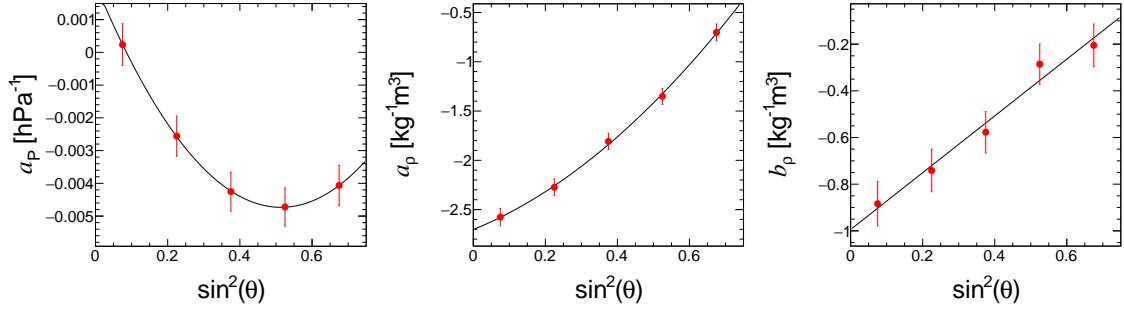

Figure 3: The weather correction was repeated in bins of equal $\sin^2 \theta$. The results below show the weather correction coefficients (see equation (3.1)) in red and the fitted curves whose values are given in Table 3.

3.4 Zenith angle dependency

To investigate the dependency on zenith angle, the data sets for the two arrays were separated into 5 (3) equal bins in $\sin^2 \theta$. The same procedure described above was carried out on the data in each of the bins independently. The resulting atmospheric coefficients were then fit to a polynomial

$$f(x) = c_0 + c_1 x + c_2 x^2, \text{ where } x = \sin^2 \theta. \tag{3.6}$$

Here $f(x)$ can represent any of the atmospheric parameters, $\{a_P, a_\rho, b_\rho\}$. Note that due to the limited number of zenith angle bins used for the 750 m array, the data were instead fit to a line ($c_2 = 0$). The results are summarized in Table 3 and shown in Figure 3 for the 1500 m array.

The density coefficients (controlled by a_ρ and b_ρ) are negative and decreasing in magnitude with zenith angle. During times of higher than average air density, the lateral spread of electromagnetic particles is reduced. The upward trend is consistent with the electromagnetic component being more attenuated at higher zenith angles, decreasing this effect when measured at the ground. The values of the pressure coefficient (a_P) are generally negative showing that increasing pressure leads to a smaller signal due to an increase in traversed matter when the air shower reaches the ground. Further, this attenuation will be even more pronounced for highly inclined showers which is demonstrated by the increasingly negative value of a_P.

4. The impact of atmospheric effects on energy measurements

Using the values obtained for $\{a_P, a_\rho, b_\rho\}$ in the previous section, the measured signals can be corrected by converting back to $\{\alpha_P, \alpha_\rho, \beta_\rho\}$ and inserting them into equation (3.1). After being corrected for weather effects, the signals then undergo the CIC correction. Since the CIC procedure also attempts to account for the attenuation of showers in air, this process must be tuned after applying weather corrections. However, since the weather correction converts the signal to a reference value chosen to be the global average, the impact on the CIC is small.

As detailed in section 2.1, the CIC corrected signals, $\hat{S}$, are then converted into an energy via a calibration with the FD calorimetric measurements, $E = A\hat{S}^B$. Here we expect some small changes to the calibration constants, due to the FD events only being recorded at night. To estimate the effects of the weather corrections on the FD calibration, we first note that b_ρ has the largest impact on the signals measured during the night compared to the daily average. This difference translates to approximately a 0.5% shift in the energy assignment and would be observed as a 0.5% change in the A calibration constant. However, this change is much smaller than the overall systematic uncertainties which are around 14%.

5. Conclusions

We have presented here an update of a weather correction for air showers which are measured at the ground level. The signal produced on the ground is modulated by changes in the atmospheric conditions. A model describing the impact of weather conditions on the measured shower signal was tuned for the 1500 m and 750 m SD arrays at the Pierre Auger Observatory. This model was an update to a previous work which includes a two hour offset in the air density which accounts for the difference in the atmospheric conditions in the last two radiation lengths above the detector. The model was also tuned for data in different zenith angle ranges. The weather correction is expected to shift the energy scale of the SD arrays by around 0.5% due to a bias in the energy calibration method.

References

[1] J. Abraham *et al.* (The Pierre Auger Collaboration), *Atmospheric effects on extensive air showers observed with the surface detector of the Pierre Auger observatory*, Astropart. Phys. **32** (2009) 86 [astroph.IM/0906.5497v2].

[2] A. Aab *et al.* (The Pierre Auger Collaboration), *The Pierre Auger Cosmic Ray Observatory*, Nucl. Instrument. Meth. A **798** (2015) 172 [astroph.HE/1502.01323]

[3] J. Hersil *et al.*, *Observations of extensive air showers near the maximum of their longitudinal development*, Phys. Rev. Lett. **6** (1961) 22.

[4] I. Valiño for The Pierre Auger Collaboration, *The flux of ultra-high energy cosmic rays after ten years of operation of the Pierre Auger Observatory*, in proceedings of *34th Int. Cosmic Ray Conf.* POS(ICRC2015)271.