Forschungszentrum Karlsruhe
Technik und Umwelt

Wissenschaftliche Berichte
FZKA 6352

A Model Independent Method for Determination of Muon Density Fluctuations in EAS

M. Giller, T. Pytlos
Institut für Kernphysik

August 1999

FORSCHUNGSZENTRUM KARLSRUHE
Technik und Umwelt

Wissenschaftliche Berichte
FZKA 6352

A Model Independent Method for Determination of Muon Density Fluctuations in EAS

Maria Giller[1] and Tadeusz Pytlos[2]
Institut für Kernphysik

[1] University of Lodz, Polen

[2] On leave from Physics Department of University of Lodz, Polen

Forschungszentrum Karlsruhe GmbH, Karlsruhe
1999

Als Manuskript gedruckt

Forschungszentrum Karlsruhe GmbH
Postfach 3640, 76021 Karlsruhe

Mitglied der Hermann von Helmholtz-Gemeinschaft
Deutscher Forschungszentren (HGF)

ISSN 0947-8620

Abstract

It is shown how to determine fluctuations of the muon density in EAS at a given distance from the core, for showers with a fixed size. The method does not make use of any pre-assumed lateral muon distribution and uses only information whether muon detector has been hit by at least one muon. Moreover, it enables to check the homogeneity of the Array and/or the shower size determination.

Eine modellunabhängige Methode zur Bestimmung der Fluktuationen in der Myondichte von ausgedehnten Luftschauern

Es wird eine Methode aufgezeigt, wie Fluktuationen der Myondichte in Luftschauern bei gegebenem Abstand vom Schauerkern und fester Schauergröße bestimmt werden können. Die Methode vermeidet a-priori Annahmen über die Lateralverteilung der Myondichte und zieht ausschließlich die Informationen heran, ob ein Myondetektor von mindestens einem Myon getroffen wurde. Sie ermöglicht die Überprüfung der Nachweiseigenschaften des Detektor-Arrays und der Schauergrößen-Bestimmung.

1. Introduction

The main aim of the KASCADE experiment is to determine chemical composition of the primary cosmic ray flux at the energy 10^{15}-10^{16} eV. To achieve this the extensive air shower characteristics have to be measured possibly without systematic errors and conclusions from the measurements should be drawn without minimal number of a priori assumptions about EAS. It is well known that the muon component in a shower is an important indication of the primary mass. Thus, its study in a possibly model independent way would be worthwhile.

A shower initiated by a higher mass particle has, on average, more muons than that corresponding to a light particle of the same primary energy. This difference becomes even bigger when one compares showers with the same total number of electrons N_e. Thus, grouping showers with N_e = const makes muon component most sensitive to the primary mass providing that the fraction of heavy nuclei in the primary flux is not too small. In particular, muon density fluctuations at a given distance from the core would have smaller average value but larger width (dispersion) for proton initiated showers than that for iron ones. As it has been demonstrated by simulations by Haungs et al (1) for 2 GeV muons, their density fluctuations are sensitive to the primary mass only if showers are grouped with fixed N_e, as opposed to fixed truncated number of muons N_μ^{tr}, when no such sensitivity in seen. In this paper we present how to determine experimentally the muon density probability distribution at a given core distance R for showers with fixed shower size N_e. In our method the muon density ρ_μ at a given R is not being determined for each **individual shower** but has been reconstructed from a **sample** of showers with fixed N_e. There are two serious reasons for that.

First, to determine muon density at a given core distance for an individual shower, one needs to make an a priori assumption about the shape of its lateral distribution. Even if we knew the average lateral distribution well, we do not know what it is for a particular shower we consider, as it fluctuates from shower to shower quite significantly.

Secondly, assuming that we can, however, find the „true" lateral distribution for an individual shower by fitting one or two free parameters in the function describing it, we make the fit by the maximum likelihood method. In can be shown, however, that such a procedure, applied to each shower from a sample, gives generally wrong **distributions** of the muon density at a given distance for the sample (although the average value is reconstructed correctly). Our calculations show (2) that it can overestimate the width of the muon density fluctuations by factor of two.

2. The idea

Our aim is to determine probability density distribution *f(N)* of the number of muons *N* falling on a fixed distance ring, for a sample of showers from a fixed N_e bin. We shall use here the information from the muon Array detectors, with E_{th}>0.3 GeV. Fluctuations of *N* from shower to shower are caused by fluctuations in shower development in the atmosphere and by the distribution of the primary particle mass. As we shall show, distribution *f(N)* influences the shape of the histogram of the number of hit muon detectors, and, as we know precisely the effects of other reasons on this histogram, we can extract *f(N)* from it.

The shower sample suitable for this analysis has to be free of any conditions imposed on muon detectors. Thus, the main data set, corresponding to the so called „Selection 4" was not the proper one to study muon fluctuations as it rejected showers with relatively low number of muons by the condition log N_μ>4.745-0.212·logN_e. For our analysis we have used another, smaller sample, without the above condition, containing over $3 \cdot 10^5$ showers. After choosing almost vertical showers only, with zenith angle < 18^0 (for the sake of radial symmetry), their number has been reduced to about $5 \cdot 10^4$. Our sample has been further divided into rather narrow bins of shower size Δ log N_e=0.1. It contains showers in the range 3.5<log N_e<6 but, as we shall show, the lower part of this (log $N_e \leq 4.2$) is not an unbiased one. For log N_e >5.6, however, the number of events is too small, so the N_e range suitable for analysis turned out to be about 1 order of magnitude smaller than that above. The core distance has been divided into bins of $\Delta R = 10m$.

Let us first choose showers with a fixed number of muon detectors m at a given distance ring. If their number is $n(m)$ then the average number of showers $<F(k;m)>$ with k (out of m) muon detectors being hit by at least one muon, should be

$$< F(k;m) >= n(m)\int_0^\infty \binom{m}{k} (1-e^{-\alpha N})^k e^{-\alpha N(m-k)} \cdot f(N)dN \qquad (1)$$

where $\alpha = S_{det}/S_{ring}$ is the ratio of the area of a muon detector to that of the whole distance ring.

In the formula (1) it is assumed that showers have radial symmetry and that the N muons fall on the whole ring (R, $R+\Delta R$) independently and randomly.

To determine $F(k;m)$ experimentally we need a criterion for a muon detector to be hit by at least one muon. First, we make our analysis only for distances $R>40$ m, where the punch through effect can be almost neglected (at least for smaller showers). Next, after looking at the energy deposit ΔE distributions for single muons from many muon detectors we have chosen ΔE>3.5 MeV as our condition, that a detector has been hit by muon(s). We would like to underline here that for our purpose we do not have to worry about how many muons have hit a detector (which, for a particular event, is not always possible with a good accuracy).

The actually observed number of showers $F(k;m)$ with k hit detectors fluctuates according to the Poissonian distribution around its expected value, given by (1) and is of course the better representation of its mean $<F(k;m)>$, the bigger is the number of showers $n(m)$. The KASCADE experiment has a big advantage of having many muon detectors (192 in the Array), allowing the number m of available detectors in a given ring to reach values even above twenty (being around 10 most frequently).

Thus, in principle we can measure many values $F(k;m)$ as $0 \leq k \leq m$, and for many m as well. Our sample, however, was not big enough for all experimental $F(k;m)$ to represent their expected values $<F(k;m)>$ with a good accuracy. So, to determine $f(N)$ (for any N_e and R bin) we have summed our $F(k;m)$ histograms over all m (over all positions of the shower core), obtaining

$$F(k)=\int_0^\infty\left[\sum_m n(m)\binom{m}{k}(1-e^{-\alpha N})^k e^{-\alpha N(m-k)}\right]f(N)dN \qquad (2)$$

for k=0,1,...,m_{max}

By summing up over *m* we lose some information contained in the *k* distributions for each individual *m*. We gain, however, by getting smaller statistical relative uncertainties of *F(k)* and by simplifying evaluation of *f(N)*.

3. Factorial moments of the distribution of *k*

From formula (1) are can easily calculate moments of the probability distribution of *k*: $<k>$, $<k^2>$ and so on. It turns out, however, that in this case it is the factorial moments of *k* which are in a simpler way connected with the muon number distribution *f(N)*. We have, for any fixed number *m* of the available detectors:

$$<k>=m\int_0^\infty(1-e^{-\alpha N})f(N)dN \qquad (3a)$$

$$<k(k-1)>=m(m-1)\int_0^\infty(1-e^{-\alpha N})^2 f(N)dN \qquad (3b)$$

$$<k(k-1)(k-2)>=m(m-1)(m-2)\int_0^\infty(1-e^{-\alpha N})^3 f(N)dN \qquad (3c)$$

and so on for higher moments.

As $1-e^{-\alpha N}=p$, where *p* is the probability of hitting a detector once *N* muons have fallen on the ring, we see that the integrals in the right-hand sides of the formulae (3) represent the successive moments of the distribution of the random variable *p*. That is

$$<k>=m\int_0^1 pg(p)dp \qquad (3d)$$

$$<k(k-1)>=m(m-1)\int_0^1 p^2 g(p)dp \qquad (3e)$$

$$<k(k-1)(k-2)>=m(m-1)(m-2)\int_0^1 p^3 g(p)dp \qquad (3f)$$

and so on, where $g(p)dp=f(N)dN$.

It is known that having all moments of a probability distribution it is possible to reconstruct the latter. Thus, in principle, one could obtain the probability distribution of *p*, *g(p)*, and then $f(N)=g[p(N)]\cdot dp/dN$.

We notice, however, that the higher is the order of the factorial moment of k, the smaller is the part of the k distribution on which it depends. Thus, as the number of showers with higher k finally decreases, one would need very big statistics in order to determine higher order moments of k (and p) with a reasonable accuracy. So, in our analysis we shall use formulae (3) to determine muon fluctuations *f(N)* in an approximate way only (§5b).

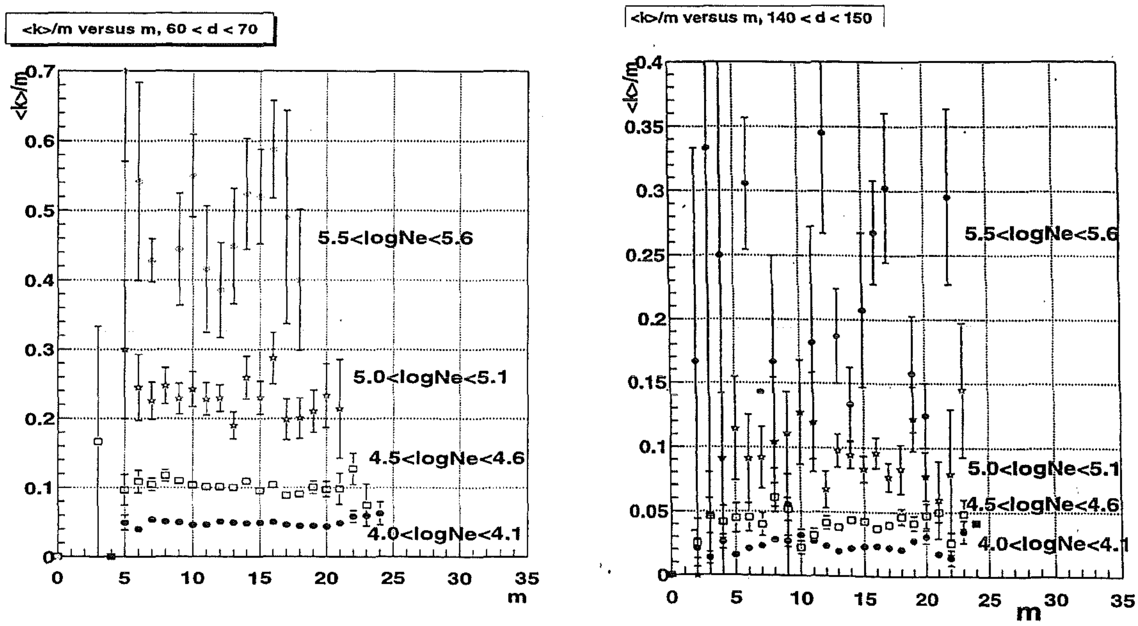


Fig.1. Average fraction of hit detectors $<k>/m$ as a function of their total number m in two distance ring: left – R=60-70 m and right – R=40-150 m, for several values of shower size N_e.

Formulae (3) can serve, however, as a check of the homogeneity of the detection conditions of the Array. From (3a) it follows that the fraction of hit detectors $\bar{k}/m$ should not, on average, depend on m, that is, on the position of the shower core. Fig.1 represents the experimentally obtained $\bar{k}/m$ as a function of m, for several shower size bins, for $60<R<70$ m (left) and for 140-150 m (right). It can be seen that within the statistical errors, the $\bar{k}/m$ values remain independent of m for almost any case.

The worst constancy is for small showers ($N_e \sim 10^4$) far from the core (~150 m). In this case $\bar{k}/m$ exhibits some systematic changes, much above statistical errors that are small. These can be caused by bigger biases in N_e determination (while it is small) correlated with the number of available detectors m. No biases are seen, however, for bigger showers (Fig.2, $N_e=10^{5.0}$-$10^{5.1}$), even up to distances ~180 m perhaps, because the muon density there is still about factor of 3 bigger than for $N_e \sim 10^4$ at 65 m.

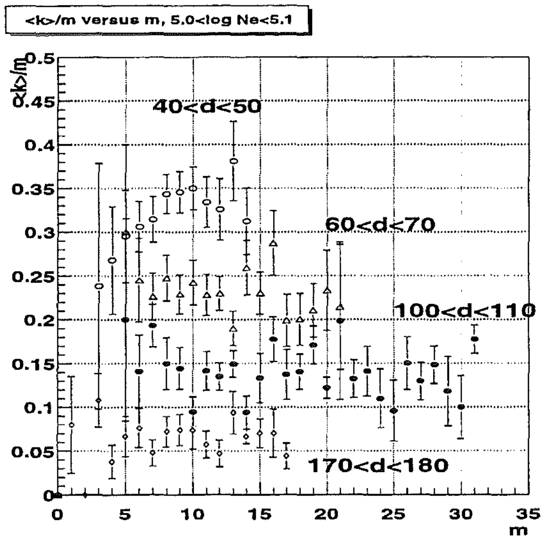


Fig.2. The same as in Fig.1 but for $N_e=10^5$-$10^{5.1}$ and several distance rings

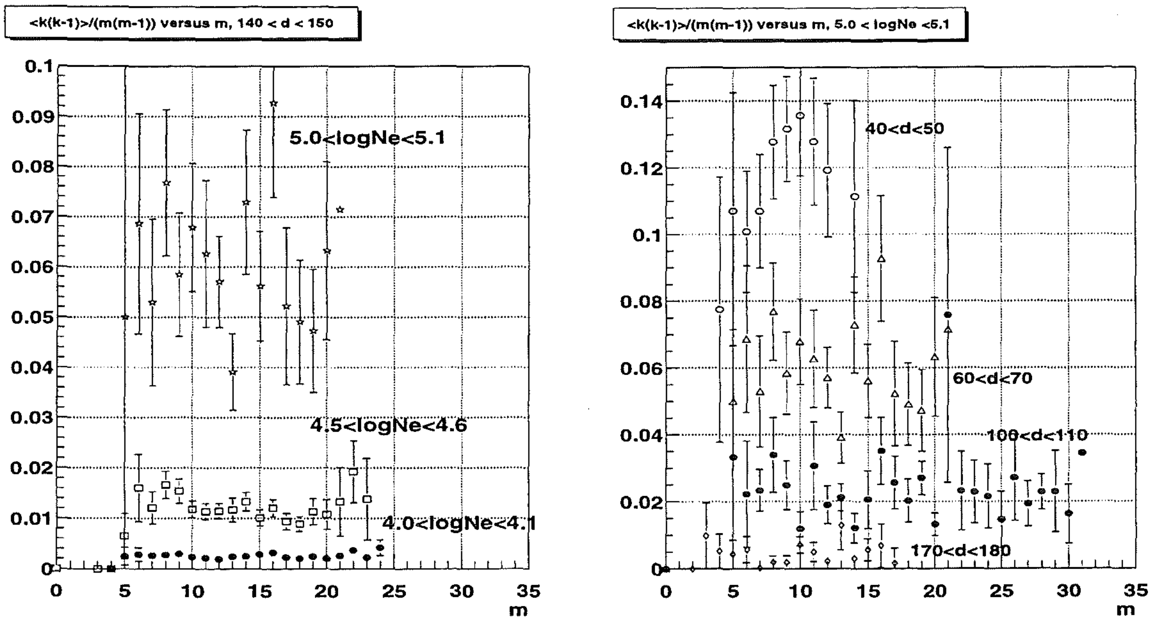


Fig.3. Second moment of p distribution, $<k(k-1)>/m(m-1)$ as a function of m. Left - for R=60-70 m and several N_e bins and right - for $N_e=10^5$-$10^{5.1}$ and several R bins.

Another check of the Array can be done on the basis of equation (3b). Ratios $<k(k-1)>/m(m-1)$ are represented in Fig.3 as a function of m for R=60-70 m for different N_e, and for $N_e=10^{5.0}$-$10^{5.1}$ for different R. Here again it can be seen that the second moment of p distribution, measured for different m, gives values which are independent of it, as one should expect. Only for small N_e, as in the previous case, there are some systematic changes of the second moment. For N_e~10^5 the statistical errors become already big enough, so we have not checked the behaviour of the higher order moments, as their statistical uncertainties would be much bigger.

4. Variance σ_k^2 of the number *k* of hit detectors

It is instructive to derive variance σ_k^2 of the distribution *F(k)* given by formula (2). It can be represented in such a way as to show explicitly the contributions of the individual random processes. A particularly useful from of σ_k^2 is the following:

$$\sigma_k^2 = \overline{m}(\overline{m}-1)\sigma_p^2 + \overline{p}^2\sigma_m^2 + \sigma_{0,k}^2 + \sigma_m^2 \cdot \sigma_p^2 \qquad (4)$$

where σ_p^2 is the variance of $p=1-e^{-\alpha N}$, σ_m^2 is the variance of the total number of detectors *m* and $\sigma_{0,k}^2$ is the variance of *k* for $m=\overline{m}$ and $p=\overline{p}$.

It can be seen that fluctuations of *k* in terms of its variance can be represented by variances of the three variables. The first term is the one which is responsible for the intrinsic fluctuations of the muon number $N=-ln(1-p)/\alpha$. The second term describes the fluctuations in the total number of detectors in the ring. It has been introduced by us when summing up the distributions *F(k;m)* (eq.(1)) over *m* to obtain a bigger statistics. The third term describes pure Bernoulli fluctuations of *k* (due to Poisson fluctuations of the number of muons on a detector), as it is equal to its variance for fixed *m* and *N*, that is $\sigma_{0,k}^2 = \overline{m}\,\overline{p}(1-\overline{p})$.The fourth term is usually the smallest one as it is of the second order for small σ_p and σ_m.

From formula (4) it can be seen that a determination of the intrinsic muon fluctuations σ_p is possible only if the uncertainty of $\sigma_{0,k}^2$ is smaller than the first term. (The uncertainty of σ_m^2 can always be reduced by taking showers with fixed *m*). This uncertainty depends on the number of showers as $n_{sh}^{-1/2}$. In this analysis (as we shall show later) the statistics becomes too small for $N_e>10^5$.

5. Methods of determining muon density fluctuations *f(N)*

As we have already explained, the basis for determining *f(N)* (for any fixed N_e and *R* bin) is a set of equations (2) for $k = 0,1.....,m_{max}$. To find *f(N)* we can either calculate *f(N)* numerically or assume its analytical form depending on some parameters which can be determined from values *F(k)*.

Here, we have applied the three following methods:

5a. Numerical fit

The integral in the right-hand side of (2) was approximated by a sum of the integrand for 10 values of *N*. The ten unknowns $f(N_i)$ were then found by a maximum likelihood method allowing for the statistical fluctuations of *F(k)*. It is the CERN program MINUIT which was used to find these best fitting values $f(N_i)$. The only condition imposed on $f(N_i)$ was that they should not be negative.

5b. Method using three moments of N distribution

This method is the better, the smaller are values of N, *i. e.* the better is fulfilled the condition $\alpha N<<1$. Expanding $e^{-\alpha N}$ and keeping only three terms we have

$$1-e^{-\alpha N} \cong \alpha N - \frac{(\alpha N)^2}{2!} + \frac{(\alpha N)^3}{3!} \qquad (5)$$

Applying (5) to the formulae (3a,b,c) and averaging them over m, we get

$$\frac{<k>}{<m>} = \alpha\overline{N} - \frac{\alpha^2\overline{N^2}}{2!} + \frac{\alpha^3\overline{N^3}}{3!}$$

$$\frac{<k(k-1)>}{<m(m-1)>} \doteq \alpha^2\overline{N^2} - \alpha^3\overline{N^3} \qquad (6)$$

$$\frac{<k(k-1)(k-2)>}{<m(m-1)(m-2)>} = \alpha^3\overline{N^3}$$

The left hand sides are known from the experiment, so it is seen that from this set of equations it is very easy to find moments of N: $\overline{N}$,$\overline{N^2}$,$\overline{N^3}$. We have limited ourselves to calculate the three moments of N, but in principle the expansion in (5) could contain more higher order terms. We would have got more equations similar to (6) then, but solving them would not be more complicated (as the last equation would contain only one term on the right-hand side and the set would be triangular as it is in (6)). The only reason that we do not want to take into account higher moments is the increasing statistical uncertainty in determining the left-hand sides of (6).

Next, we assume that $f(N)$ has a shape of a gamma function

$$f(N) = \frac{q^p}{\Gamma(p)} N^{p-1} e^{-qN} \qquad (7)$$

We calculate the parameters p and q from the conditions that the first two moments of the gamma function in (7) be equal to the first two moments $\overline{N}$ and $\overline{N^2}$, determined from the experiment by use of equation (6). Thus, we get

$$p = \frac{\overline{N}^2}{\overline{N^2} - \overline{N}^2} \quad ; \quad q = \frac{\overline{N}}{\overline{N^2} - \overline{N}^2} \qquad (8)$$

The third moment $\overline{N^3}$ determined from (6), would not, in general, be equal to that calculated from (7) and (8) (although the information contained in it has been used in calculating $\overline{N}$ and $\overline{N^2}$ from (6)).

5c. Method adopting *f(N)* as a gamma function

Although in the previous method we did finally assume that *f(N)* is a gamma function, the main idea of that method is not based on this assumption. It shows how to calculate many (in principle) moments of the *N* distribution *f(N)*. The more moments we know, the better is the determination of the function.

Here, we assume from the very beginning that *f(N)* can be described by a gamma function. Inserting this to equations (3a) and (3b) and averaging over m we obtain

$$\frac{<k>}{<m>} = 1 - \left(\frac{q}{q+\alpha}\right)^p$$

$$\frac{<k(k-1)>}{<m(m-1)>} = 1 - 2\left(\frac{q}{q+\alpha}\right)^p + \left(\frac{q}{q+2\alpha}\right)^p \qquad (9)$$

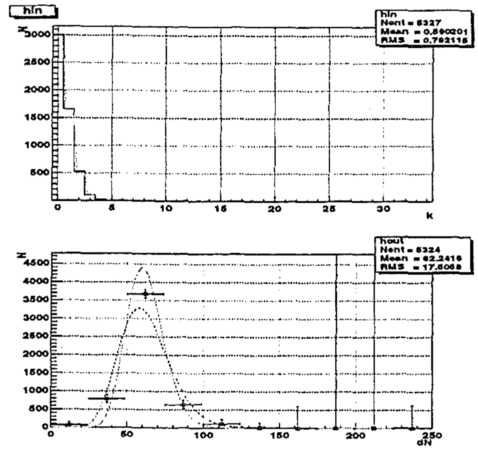
The parameters p and q can be found from equations (9) by solving them numerically. This method does not require small muon densities as in the case 4b.

6. Results

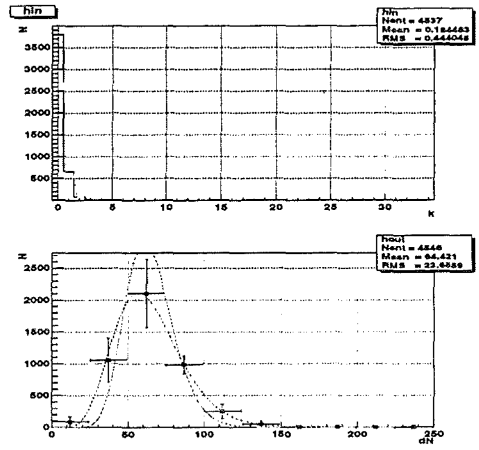
6a. Comparison of the results obtained by the three methods

Fig. 4a,b,...,f show several examples of the histograms *F(k)* (upper parts) and calculated from them distributions *f(N)*, multiplied by number of showers and ΔN (lower parts). *f(N)* are calculated by the three different methods, as described in § 5. The results obtained by MINUIT (method 5a) are represented by points, the dashed curves correspond to the method 4b (moments of N) and the dotted curves - to the method 4c (gamma function). The dotted histogram in the upper part is the prediction for the hit detectors distributions *F(k)*, for *f(N)* obtained from the MINUIT calculation. Thus comparing it with the solid histogram (experimental distribution), one can see how good the MINUIT fit is. In general, if the statistics is large enough the *f(N)* obtained by different methods are not very different.

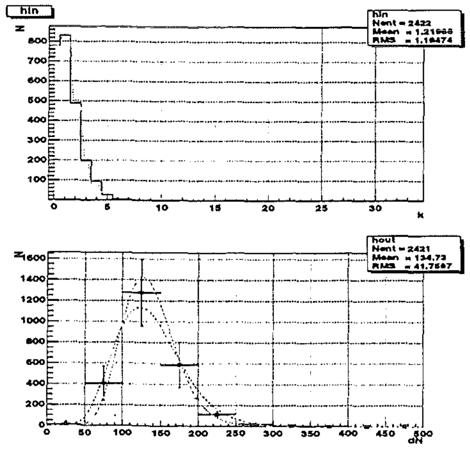
A more quantitative comparison of the three methods is contained in Tables 1 and 2. Table 1 shows average numbers of muons $\overline{N}$, calculated by the three methods and its standard deviations (in brackets) in different distance rings for $4.0<\log N_e<4.1$. Table 2 represents the same for $5.0<\log N_e<5.1$. It can be seen that the spread of the three values $\overline{N}$ is more or less contained within its dispersions. However, the dispersions of $\overline{N}$,$\sigma_{\overline{N}}$ (being the dispersions of *f(N)*, σ_N, divided by the square-root of the number of showers) show bigger differences. If the statistics is not large enough, fluctuations of the number of events in the histogram *F(k)* cause rather large changes in the higher moments, and as a result, in the parameters of the gamma functions calculated in the second and third methods. In some cases it was not possible to determine σ_N by the second or third method apparently because the distribution *F(k)* fluctuated in such a way that it was narrower than that for a single value of *N*. In general, the dispersions $\sigma_{\overline{N}}$ (and σ_N) are the largest for the MINUIT method and the smallest for the third method (gamma functions). We believe that the former reflect better the true dispersions, as *f(N)* has been fitted there with 10 free parameters (its values for 10 N_i,



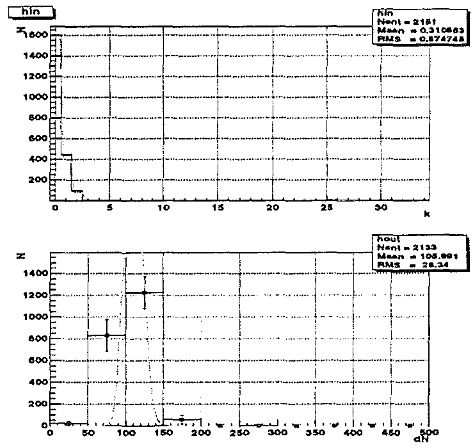
a) N_e=10^4-$10^{4.1}$, R=60-70 m



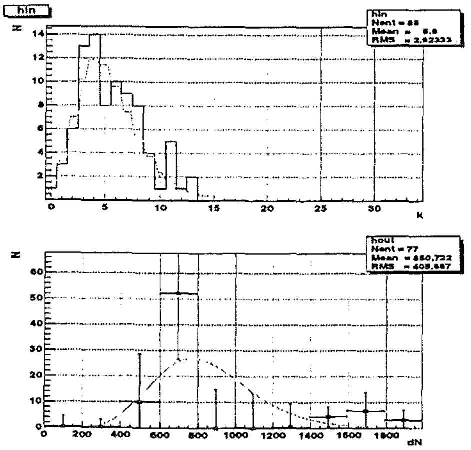
b) N_e=10^4-$10^{4.1}$, R=160-170 m



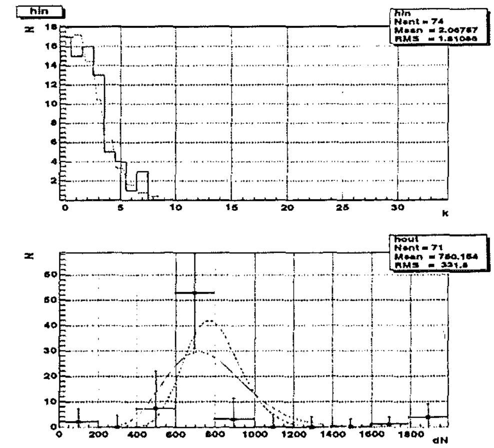
c) N_e=$10^{4.5}$-$10^{4.6}$, R=60-70 m



d) N_e=$10^{4.5}$-$10^{4.6}$, R=160-170 m



e) N_e=$10^{5.5}$-$10^{5.6}$, R=60-70 m



f) N_e=$10^{5.5}$-$10^{5.6}$, R=160-170 m

Fig. 4a,b,...,f. Observed distributions *F(k)* of the number of hit muon detectors *k* (upper histograms, solid line). Calculated distributions of the muon number *N* in the distance ring (lower graphs). Points - method 5.1; dashed line - method 5.2, dotted line - method 5.3. The dotted histogram (upper graphs) is the predicted *k* distribution by method 5.1 (MINUIT).

Table 1. Average number of muons $\overline{N}$, determined by the three methods (last three columns) in different distance rings (first column) for $N_e=10^4$-$10^{4.1}$. In brackets are their estimated uncertainties. Second column shows number of showers. Symbols ▼ and ▲ are when the method did not work (see text).

R(m)	n_{sh}	Method 5.1	Method 5.2	Method 5.3
40-50	5231	67.00(0.20)	67.21(0.13)	67.02(0.12)
50-60	5345	64.87(0.39)	64.57(0.19)	64.38(0.14)
60-70	5324	62.24(0.24)	62.41(0.23)	62.24(0.17)
70-80	5325	60.99(0.26)	60.94(0.16)	60.85(0.12)
80-90	5211	63.83(0.17)	64.09(▼)	64.14(0.11)
90-100	5263	61.26(0.22)	61.906(0.004)	61.92(0.11)
100-110	5346	61.40(0.27)	60.79(0.09)	60.78(0.11)
110-120	5309	61.71(0.35)	62.02(0.02)	62.04(0.11)
120-130	5325	63.00(0.36)	63.60(0.28)	63.49(0.20)
130-140	5337	63.64(0.30)	63.75(0.28)	63.65(0.21)
140-150	5303	65.34(0.23)	65.39(▼)	65.44(0.11)
150-160	5055	63.49(0.29)	63.26(0.21)	63.21(0.15)
160-170	4546	64.42(0.35)	64.61(0.33)	64.52(0.23)
170-180	3953	64.87(0.45)	64.58(0.20)	64.55(0.14)
180-190	3442	63.73(0.37)	65.15(0.36)	65.07(0.22)
190-200	2856	57.00(0.41)	58.99(▼)	59.50(0.14)

Table 2. The same as in Table 1, but for $N_e=10^5$-$10^{5.1}$.

R(m)	n_{sh}	Method 5.1	Method 5.2	Method 5.3
40-50	429	363.04(6.73)	378.90(▼)	358.09(5.74)
50-60	434	357.97(6.69)	368.31(▼)	352.40(5.06)
60-70	428	320.61(4.89)	334.34(▼)	325.84(3.68)
70-80	436	325.70(5.61)	315.52(2.06)	309.40(3.38)
80-90	428	321.97(6.82)	320.51(3.66)	313.99(4.01)
90-100	421	304.24(4.95)	310.70(1.90)	307.00(2.47)
100-110	437	299.36(6.37)	304.61(1.90)	298.43(4.27)
110-120	436	277.66(7.34)	272.73(5.14)	268.13(3.97)
120-130	437	279.00(5.86)	267.39(2.97)	265.40(2.78)
130-140	438	265.39(5.31)	263.76(4.00)	261.31(3.26)
140-150	439	285.21(8.70)	275.25(5.77)	271.30(4.33)
150-160	429	250.70(5.92)	256.36(7.33)	253.41(4.02)
160-170	388	222.87(4.70)	239.64(1.09)	239.10(▲)
170-180	372	229.45(4.53)	233.86(3.20)	232.78(1.97)
180-190	331	195.32(2.82)	200.59(▼)	202.27(▲)
190-200	278	226.02(13.69)	220.79(11.46)	215.51(8.84)

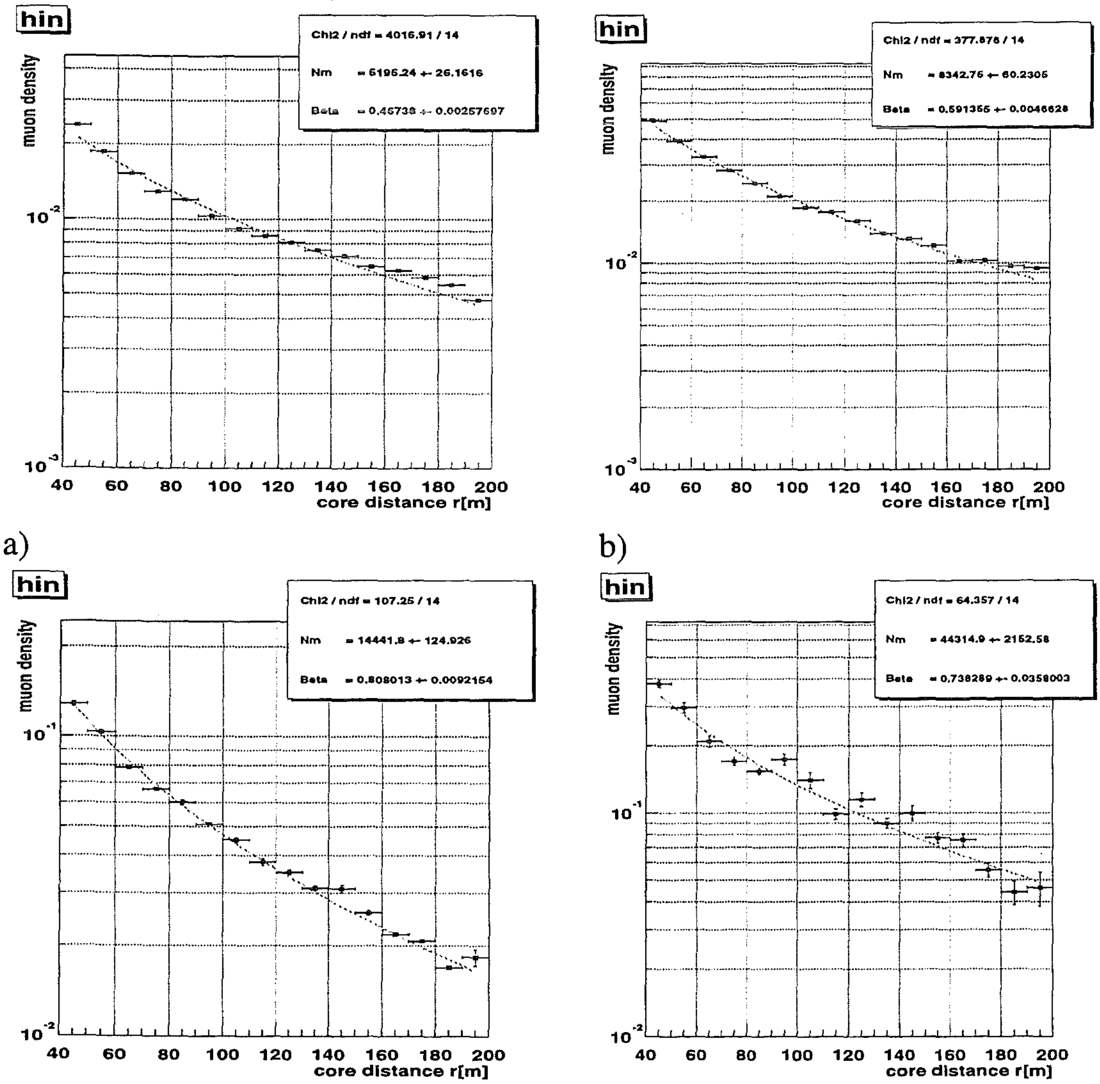
although in practice 6-7 values were obtained as being larger than zero). Assuming from the beginning that *f(N)* is a gamma function seems to be a too big simplification which underestimates the width of the muon fluctuations. In what follows we shall present results obtained by MINUIT.

6b. Muon lateral distributions

Dividing average number of muons $\overline{N}$ in the ring by its area, we can determine the average lateral distribution $\rho_\mu(R)$ for showers with fixed N_e. Fig. 5a,b,c,d, show the results for four bins of N_e. The dashed lines are the best fitted Greisen distributions

$$\rho_G(R) = \frac{N_\mu}{2\pi R_o^{\,2}} \cdot \frac{\Gamma(2.5)}{\Gamma((2-\beta)\,\Gamma(\beta+0.5)} \left(\frac{R}{R_o}\right)^{-\beta} \left(1+\frac{R}{R_o}\right)^{-2.5} \qquad (10)$$

where R_o=320 m and N_μ and β were fitted to the points.



a) b) c) d)

Fig.5a,b,c,d. Average lateral distribution of muons (in m^{-2}). The dashed curve is the best fitted Greissen distribution (see text).

a) $N_e=10^4$-$10^{4.1}$, b) $N_e=10^{4.5}$-$10^{4.6}$, c) $N_e=10^5$-$10^{5.1}$, d) $N_e=10^{5.5}$-$10^{5.6}$.

The Greisen function does not describe well the muon distribution at low N_e. The χ^2 values (for 14 degrees of freedom) are given in the inserts, as well as N_μ and β. (Their uncertainties have been automatically calculated by ROOT but, as χ^2 values are so big that the hypothesis about the fit would be rejected, they are rather meaningless.) We have also fitted the NKG function (with R_o=420m) but the obtained curves are practically non-distinguishable from the fitted Greisen distributions. Our aim here was not to determine lateral distributions of muons, so we have not been particularly careful with subtracting possible contamination by electrons or hadrons. As we have already mentioned, we have allowed only for the electron/photon punch-through by taking into analysis muon detectors with core distances larger than 40 m. If this is sufficient for small showers, it may not be the case for bigger ones.

We have compared, however, our lateral distributions with that obtained by a more sophisticated analysis with a much bigger statistics by Leibrock et al. (3), the actual $\rho_\mu(R)$ for fixed N_e have been given to us by A.Haungs. The comparison is shown in Fig.6 for three bins of N_e. As their N_e bins were larger ($\Delta \log N_e$=0.2) we have corrected their points by factor $10^{0.05 \cdot 0.85}$=1.1 (up or down), assuming that N_μ(40-200m)~$N_e^{0.85}$. It can be seen that the agreement is quite good although our distributions become steeper with increasing shower size N_e, whereas this effect is much weaker in the analysis by Leibrock et al.

In Fig.6 we have also shown results obtained by EAS-TOP experiment (4). The points (full circles) are for $5<\log N_e<5.3$ at the experimental level which is ~200g·cm^{-2} higher in the atmosphere than KASCADE. As the attenuation length of the electromagnetic component in showers is about 43% per 100 g·cm^{-2}, showers with size $N_e=10^{4.5}$ at the KASCADE level correspond to $N_e=10^{5.2}$ at the EAS-TOP level. Keeping in mind that the muon attenuation at 1 GeV (EAS-TOP threshold) is small and that there is ~20% less muons with E>1GeV than these with E>0.3 GeV (5) we conclude that the agreement is quite satisfactory.

The change of the muon lateral distributions with shower size N_e is illustrated in Fig.7 where the dependence of β (from formula 10) on N_e is shown. Full squares correspond to our best fit with the statistical uncertainties of $\overline{\rho}(R)$ only. From Fig.5 it can be seen, however, that the dispersion of points from Greisen distributions is larger. To some extent this is simply due to the fact that the Greisen function does not have to describe well the true lateral distributions. But there must be some systematic errors as well which are probably mainly responsible for a rather large jump of the β value at $\log N_e$=5. In order to allow somehow for bigger than statistical errors of ρ_μ, we have recalculated the best fitted Greisen functions assuming that the relative dispersion of $\overline{\rho}_\mu$ is equal to the relative dispersion of the average number of hit detectors $\overline{k}$. This may serve as an upper limit to the statistical fluctuations of $\overline{\rho}_\mu$, as for most our cases $\overline{\rho}_\mu \sim \overline{k}/\overline{m}$. The new values of β are represented by open circles. We can see that the increase of β with growing N_e is smoother now and we believe that it reflects better the reality.

The solid lines show the predictions of β obtained by CORSIKA simulations (6), described by the formula

$$\beta(\theta,N_e)=0.41+(0.13-0.07/\cos\theta)\cdot\log N_e \qquad (11)$$

The two lines show β range corresponding to $0<\theta<18^0$. The simulations are for a mixed chemical composition. It can be seen that the experimentally obtained rate of the β increase with N_e is bigger than that obtained by simulations.

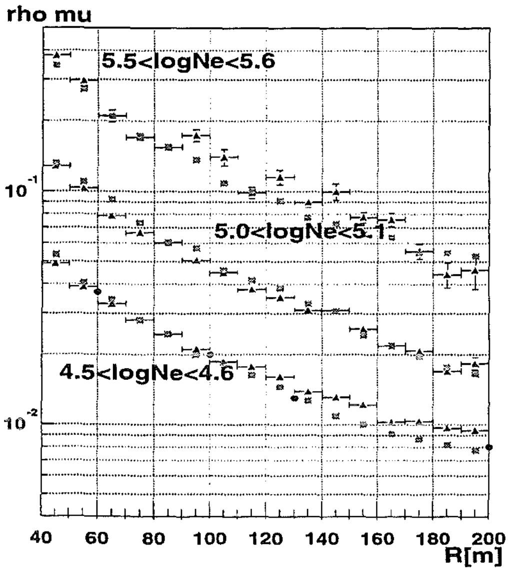


Fig.6. Comparison of the average lateral distributions of muons (ρ_μ in m^{-2}) obtained in this work (triangles) for three N_e bins with Leibrock et al. (squares). Circles show values obtained by EAS-TOP (4) for E_μ>1 GeV.

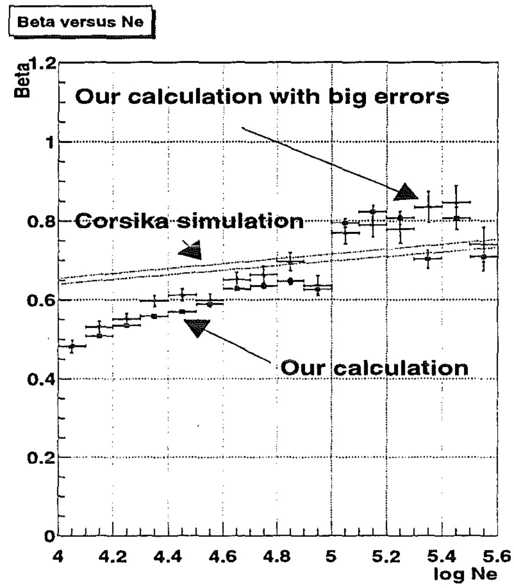


Fig.7. Parameter β from the Greisen function versus shower size N_e. See text for more details.

This point, however, needs further, more careful studies. In particular, as we have already mentioned, a bigger contribution of the punch-through of the electromagnetic component in bigger showers (not taken into account here) could be responsible for a larger increase rate of β with N_e. On the other hand we would like to stress here that the best fitted β values are quite sensitive to the uncertainties $\sigma_{\bar{\rho}}$ of muon densities $\bar{\rho}$. These in turn are sensitive to the choice of bins ΔN for the MINUIT program. Thus, a recalculation of $\sigma_{\bar{\rho}}$ with a bigger shower statistics is needed before any conclusions about β increase rate could be drawn.

Integrating our lateral distributions in the distance range $40<R<200$m we obtain average truncated number of muons N_μ^{tr}. Their dependence on N_e is represented in Fig.8. Also shown are the results obtained by another KASCADE analysis by Weber et al. (7). The agreement is not bad, although the discrepancy in N_μ^{tr} at $N_e=10^{4.5}$ is ~40%, being 12% or so at our highest N_e value, N_e~$10^{5.75}$. The power law dependence N_μ^{tr}~$N_e^{0.85}$ describes well our results for $N_e>10^{4.7}$.

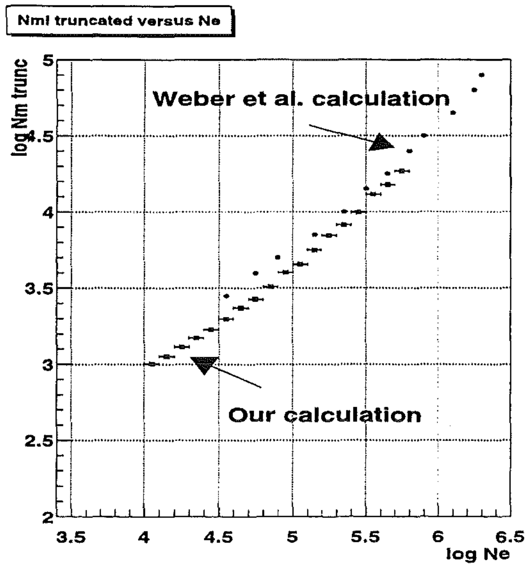


Fig.8. Truncated number of muons N_μ^{tr} as a function of N_e.

6c. Muon density fluctuations

The relative fluctuations $\sigma_\rho / \bar{\rho}_\mu$, as determined by the MINUIT method, are shown in Fig. 9a,b,c,d for several bins of N_e as a function of the core distance.

The uncertainties are practically only those of σ_ρ and they have been determined in the following way. We have assumed that the *F(k)* experimental histogram (for any N_e and *R*) is representative (average) for the true distribution of *k*. Then, for each value of *k* we have randomly chosen from the Poisson distribution (with the mean equal to *F(k)*) a new number of showers *F'(k)*. Applying our MINUIT method to *F'(k)* we obtain a new value of σ_ρ. Repeating this several times we get a sample of σ_ρ values, the standard deviation of which is our statistical uncertainty δ of a single σ_ρ. The values of δ divided by $\bar{\rho}_\mu$ are marked in Fig.9 as the error bars. They have been calculated from a sample of 10 values of σ_ρ (9 from randomly chosen *F'(k)*).

The uncertainties δ can also be derived in an analytical (although slightly approximate) way. The derivation is given in the Appendix. We have checked that the two values of δ, calculated by the two methods, give practically the same results.

Inspection of Fig.9 shows that it is difficult to draw conclusions about the dependence of $\sigma_\rho / \bar{\rho}$ on the core distance. Even for smaller showers ($4.4<\log N_e<4.5$) despite the fact that the best fit shows a trend of an increase, a hypothesis of constant $\sigma_\rho / \bar{\rho}$ would probably pass the χ^2 test. From this preliminary analysis we can only conclude that if there is any change of $\sigma_\rho / \bar{\rho}$ with *R* (for a fixed N_e) it is probably not larger than 30% within distance range $40<R<200$m.

It can be seen, however, that the fluctuations below N_e~10^5 are the order of 30-40%. Above N_e=10^5 the uncertainties δ become too large, sometimes they are comparable to the values of $\sigma_\rho / \bar{\rho}$ itself, so that values of the latter can not be treated seriously.

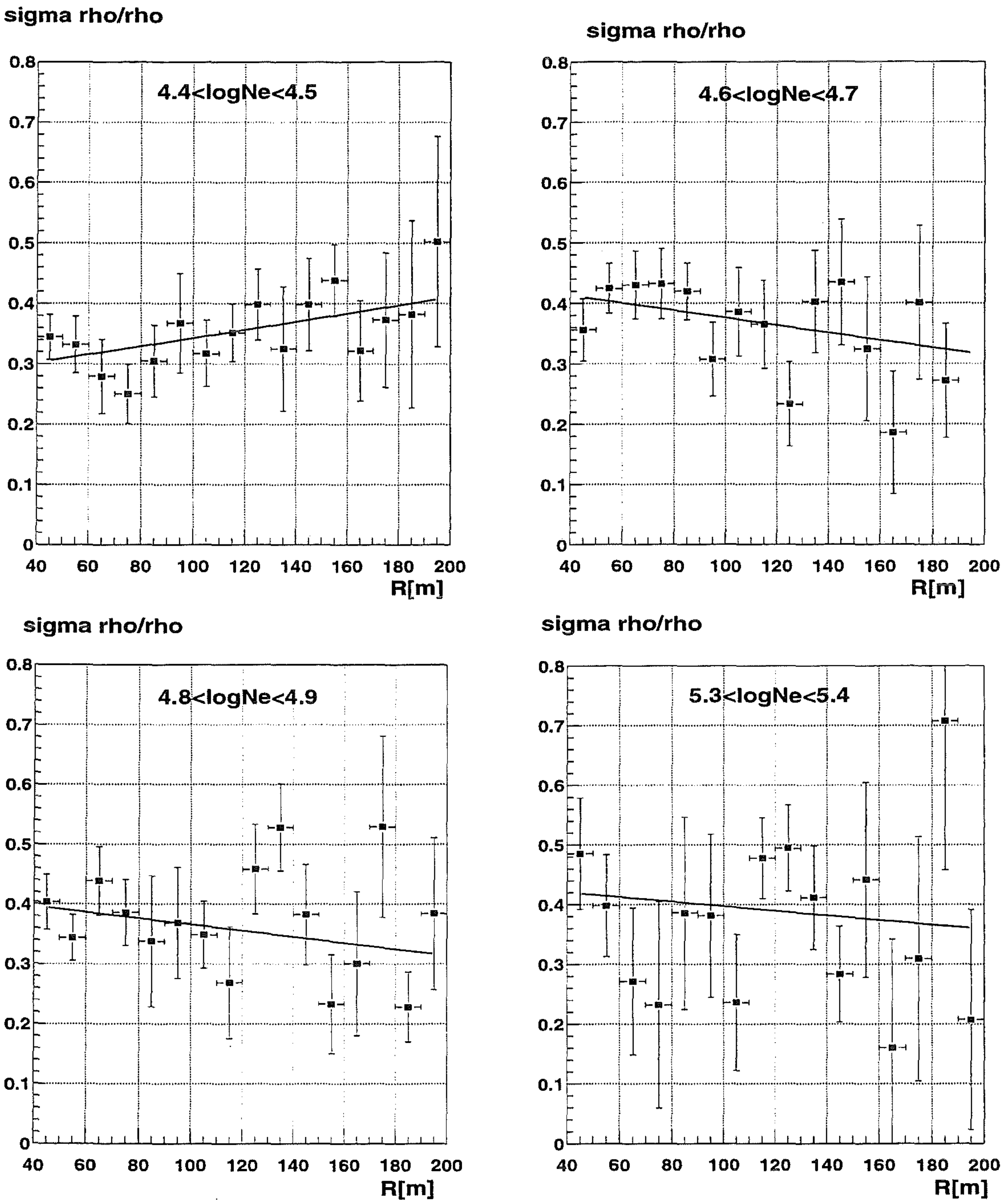


Fig.9. Relative dispersion $\sigma_p \, / \, \overline{\rho}_\mu$ of the muon density as a function of the core distance for several bins of N_e.

7. Summary and conclusions

Let us summarise the main results and conclusions obtained in this work:

- We point out that to determine muon density fluctuations (at a given distance from the core) it is necessary to do it for a sample of showers at a time. Calculating muon density for each individual shower by the maximum likelihood method and constructing then the density distribution leads, in general, to an incorrect dispersion of this distribution.
- We have derived formulas (3) relating factorial moments of the number of hit muon detectors with the moments of p, the probability of hitting a detector. With very big

statistics of showers these formulas would be sufficient to determine muon density probability distribution. Apart from that they can serve as a check of the homogeneity of the detection conditions of the Array and/or of the correctness of N_e determination in the KASCADE data tapes. Our analysis has shown that apparently there is a bias in the N_e determination for small showers (logN_e~4).

- Three different methods of deriving muon density fluctuations *f(N)* have been presented and analysed. The methods are „model independent" in the sense that: a/ no a priori assumptions are made about the shape of the muon lateral distribution (as it has to be done when analysing individual showers), b/ the only information needed about a muon detector is whether is has been hit by at least one muon or not. The correctness of such a decision is much better than a determination of the actual number of muons passing the detector.

  The method by using MINUIT turned out to be the most reliable and accurate. Reconstructing *f(N)* by fitting its two moments only or assuming that *f(N)* is a gamma function gives, in general, too narrow distributions. The average densities, however, determined by the three methods, agree very well.
- The lateral muon distributions obtained in this work agree pretty well with those determined by another KASCADE analysis (3). The truncated number of muons N_μ^{tr} determined here is slightly lower than that of Weber et al (7).
- The dispersions of the muon density fluctuations at a given distance from the core, for showers with fixed N_e, are of the order of 30-40%. Their dependence on the core distance and shower size is rather weak. Thus much bigger statistics are necessary to draw more quantitative conclusions about this dependence.
- An accurate determination of the shape of muon density fluctuations and its dependence on the shower size should reveal the chemical composition of the primaries (at least check the hypothesis that the fraction of heavy nuclei increases with energy).

The authors thanks prof. Gerd Schatz for valuable discussions. T.P. thanks Forschungszentrum Karlsruhe for the hospitality during his visit when part of this work was done.

This work was supported in part by the Polish Committee for Scientific Research (KBN) by the grant no. 2P03B16012.

References

(1) A.Haungs et al., Wissenschaftliche Berichte FZKA 6263, Forschungszentrum Karlsruhe, 1999
(2) T.Pytlos, PhD thesis, Lodz 1998 (in Polish)
(3) H.Leibrock, A.Haungs and H.Rebel, Unveröffentlicher Bericht, Forschungszentrum Karlsruhe, 1998 (in German)
(4) M.Aglietta et al., Proc.24th ICRC (Rome) **2** 664, 1995
(5) H.J.Simonis and H.Rebel, Unveröffentlicher Bericht, Forschungszentrum Karlsruhe, 1996
(6) A.Heiss, PhD thesis, Forschungszentrum Karlsruhe, 1996 (in German)
(7) J.H.Weber for the KASCADE Collaboration, Proc. 25th ICRC (Durban) **6**, 153, 1997

Appendix

Derivation of the statistical error of the muon density fluctuations

We shall show here how to calculate the statistical accuracy with which the width of the muon density distribution has been determined (Ne, R - fixed).
We shall start with the determination of the dispersion $\delta(\sigma_p^2)$ of the random variable $\sigma_p^2 = \overline{p^2} - \overline{p}^2$, being the experimentally determined variance of $p=1-e^{-\alpha N}$, uniquely related to N. (σ_p^2 is a random variable, each value of which corresponding to the whole set of showers taken for analysis).
We can see from Fig.4 that assuming that N has a Gaussian distribution is not a very crude approximation. Moreover, for most of our cases $\alpha N<1$, so $p \sim \alpha N$ and we may further assume that p has a Gaussian distribution as well. If this is the case then, according to the well known theorem, σ_p and $\overline{p}$ (determined from a sample) are independent random variables. Thus, σ_p^2 *and* $\overline{p}^2$ are independent as well. Then

$$\delta^2\left(\overline{p^2}\right) = \delta^2\left(\sigma_p^2\right) + \delta^2(\overline{p}^2) \qquad \text{A1}$$

where $\delta^2(x)$ denotes the variance of x. Thus the variance of the random variable σ_p^2 is equal to the difference of the variances of $\overline{p^2}$ and $\overline{p}^2$.
Let's start with deriving $\delta^2(\overline{p}^2)$. In our analysis we determine $\overline{p}$ with a quite small statistical error, so we can write that

$$\delta(\overline{p}^2) = 2\overline{p}\,\delta(\overline{p}) \qquad \text{A2}$$

and we have to find the dispersion of $\overline{p}$, $\delta(\overline{p})$. From (3a) averaged over m we have that $\overline{p} = \overline{k}/\overline{m}$. The dispersion of $\overline{p}$ has to be calculated taking into account that $\overline{k}$ and $\overline{m}$ are correlated random variables.
In general, if $z=f(x,y)$, where x and y are correlated random variables, then the variance of z, σ_z^2, depends on the variances σ_x^2 and σ_y^2 in the following way

$$\sigma_z^2 \cong \left(\frac{\partial f}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \sigma_y^2 + 2\frac{\partial f}{\partial x}\cdot\frac{\partial f}{\partial y}\Bigg|_{\overline{x},\overline{y}} \cdot \rho(x,y)\cdot\sigma_x\sigma_y \qquad \text{A3}$$

assuming that $\sigma_x/\overline{x} \ll 1$ and $\sigma_y/\overline{y} \ll 1$. Formula A3 can be easily derived by expanding $f(x,y)$ around the point ($\overline{x},\overline{y}$), averaging $(\Delta f)^2$ and keeping only terms of the smallest order. $\rho(x, y)$ is the correlation coefficient of the two variables, defined as

$$\rho(x,y) = \frac{E[(x-\bar{x})(y-\bar{y})]}{\sigma_x \cdot \sigma_y} \qquad \text{A4}$$

where the numerator represents the expected value of the product of the deflections of x and y from their mean values.
If $\bar{p} = z = y/x$, with $x = \bar{m}$ and $y = \bar{k}$, we can use the above formula to find $\delta^2(\bar{p}) = \sigma_z^2$, as the dispersions of $\bar{k}$ and $\bar{m}$ are really very small when compared with $\bar{k}$ and $\bar{m}$ respectively.

Now we need to find the correlation coefficient $\rho(\bar{m},\bar{k})$ of the two variables $\bar{m}$ and $\bar{k}$. First we notice that it is equal to the correlation coefficient of m and k (which can be shown from its definition A4).
If we neglect the difference between the expected values of k, m and $k \cdot m$ with their average values $\bar{k}, \bar{m}$ and $\overline{k \cdot m}$ respectively from our sample, then

$$\rho(m,k) = \frac{\overline{k \cdot m} - \bar{k} \cdot \bar{m}}{\sigma_k \cdot \sigma_m} \qquad \text{A5}$$

To calculate $\overline{k \cdot m}$ let us first do it for a fixed N. The probability $P(k,m)$ that for a fixed N we have a shower with a given value of m and k is equal

$$P(k,m) = w(m)\binom{m}{k} p^k (1-p)^{m-k} \qquad \text{A6}$$

where $w(m)$ is the probability of m.
Thus

$$\overline{k \cdot m} = \sum_m \sum_{k=0}^{m} k \cdot m \cdot P(k,m) = \overline{m^2} p \qquad \text{A7}$$

It is seen that allowing for the distribution of N is equivalent to averaging (A7) over $p = 1 - e^{-\alpha N}$, to obtain

$$\overline{k \cdot m} = \overline{m^2} \cdot \bar{p} \qquad \text{A8}$$

Substituting A8 to A5 and remembering that $\bar{k} = \bar{m} \cdot \bar{p}$ we finally get the correlation coefficient

$$\rho(m,k) = \frac{\overline{m^2}\bar{p} - \bar{m}^2 \cdot \bar{p}}{\sigma_k \cdot \sigma_m} = \frac{\sigma_m}{\sigma_k}\bar{p} = \frac{\sigma_m / \bar{m}}{\sigma_k / \bar{k}} \qquad \text{A9}$$

So $$\rho(\bar{m},\bar{k}) = \rho(m,k) = \frac{\sigma_{\bar{m}} / \bar{m}}{\sigma_{\bar{k}} / \bar{k}} \qquad \text{A10}$$

as both $\sigma_{\overline{m}}$ and $\sigma_{\overline{k}}$ are determined from the same number of showers.
Substituting A10 to A3, with $x = \overline{m}$, $y = \overline{k}$ and $z = \overline{p} = \overline{k} / \overline{m}$ we obtain that the dispersion of $\overline{p}$, $\delta(\overline{p})$ equals

$$\delta(\overline{p}) = \sqrt{\sigma_z^2} = \sqrt{\left(\sigma_{\overline{k}} / \overline{k}\right)^2 - \left(\sigma_{\overline{m}} / \overline{m}\right)^2} \cdot \frac{\overline{k}}{\overline{m}} \qquad \text{A11}$$

Together with A2 this gives us $\delta(\overline{p}^2)$ from A1.

We are now left with a calculation of the variance of $\overline{p^2}$, $\delta^2(\overline{p^2})$.
If we denote $\kappa = k(k-1)$ and $\mu = m(m-1)$, then

$$\overline{p^2} = \overline{k(k-1)} / \overline{m(m-1)} = \overline{\kappa} / \overline{\mu} \qquad \text{A12}$$

We see that now we are facing a similar problem as when calculating the variance of $\overline{k} / \overline{m}$. $\overline{\kappa}$ and $\overline{\mu}$ are correlated random variables, their correlation coefficient being the same as that for κ and μ. The latter equals (similarly to A5)

$$\rho(\mu, \kappa) = \frac{\overline{\kappa \cdot \mu} - \overline{\kappa} \cdot \overline{\mu}}{\sigma_\kappa \cdot \sigma_\mu} \qquad \text{A13}$$

Calculating $\overline{\kappa \cdot \mu}$, the average value of *k(k-1)m (m -1),* we obtain

$$\overline{\kappa \cdot \mu} = < \left[m(m-1)\right]^2 > \cdot \overline{p^2} \qquad \text{A14}$$

Taking into account that $\overline{k(k-1)} = \overline{m(m-1)} \cdot \overline{p^2}$ (formula (3b) averaged over *m*) and substituting this and A14 to A13 we get

$$\rho(\mu, \kappa) = \frac{\sigma_\mu / \overline{\mu}}{\sigma_\kappa / \overline{\kappa}} = \rho(\overline{\mu}, \overline{\kappa}) = \frac{\sigma_{\overline{\mu}} / \overline{\mu}}{\sigma_{\overline{\kappa}} / \overline{\kappa}} \qquad \text{A15}$$

As this formula has an identical form as (A10) and we want to find the variance of the ratio $\overline{\kappa} / \overline{\mu}$ of the two variables, similarly as before for $\overline{k}$ and $\overline{m}$, we obtain a formula as A11

$$\delta(\overline{p^2}) = \sqrt{\left(\sigma_{\overline{\kappa}} / \overline{\kappa}\right)^2 - \left(\sigma_{\overline{\mu}} / \overline{\mu}\right)^2} \cdot \frac{\overline{\kappa}}{\overline{\mu}} \qquad \text{A16}$$

Substituting A16, A11, and A2 to A1 we obtain

$$\delta^2(\sigma_p^2) = \left[(\sigma_{\bar{\kappa}} / \bar{\kappa})^2 - (\sigma_{\bar{\mu}} / \bar{\mu})^2\right]\left(\frac{\bar{\kappa}}{\bar{\mu}}\right)^2 - 4\left(\frac{\bar{k}}{\bar{m}}\right)^4\left[(\sigma_{\bar{k}} / \bar{k})^2 - (\sigma_{\bar{m}} / \bar{m})^2\right] \qquad \text{A17}$$

For big statistics, when σ_p^2 can be determined quite accurately, so that $\delta(\sigma_p^2)/\sigma_p^2 << 1$, we can calculate the dispersion of σ_p ,$\delta(\sigma_p)$, from the relation

$$\delta(\sigma_p) = \frac{\delta(\sigma_p^2)}{2\sigma_p} \qquad \text{A18}$$

To find finally the dispersion of σ_N, $\delta(\sigma_N)$, we need to have a relation between σ_N and σ_p. Strictly speaking, σ_N can be calculated only if the whole probability distribution of p is known, not only its moment σ_p. But it is probably not worthwhile to be so strict and we shall proceed in an approximate way. When $\alpha N << 1$, $p \cong \alpha N$ and then $\sigma_N = \frac{1}{\alpha}\sigma_p$. When $\sigma_p / \bar{p} << 1$ then $\sigma_N = \frac{dN}{dp}\sigma_p = \frac{\sigma_p}{\alpha(1-p)}$.

If none of the above two conditions is fulfilled we can estimate the statistical uncertainty in the determination of σ_N from the approximate formula

$$\delta(\sigma_N) \approx \frac{1}{2}\left(\left.\frac{dN}{dp}\right|_{\bar{p}-\sigma_p} \cdot \delta(\sigma_p) + \left.\frac{dN}{dp}\right|_{\bar{p}+\sigma_p} \cdot \delta(\sigma_p)\right)$$