О модификации массового состава при ультравысоких энергиях

(по данным AUGER Col., с привлечением формализма копулы)

А. Кириллов, И.Кириллов Дубна 2014

Статистический подход с целью ДОПОЛНИТЬ возможности эксперимента параметрами функций распределения

Введение (наиболее богатая статистика у коллаборации ОЖЕ)

Свободный доступ 1% событий

Событие, погрешности

Разделение на энергетические интервалы (12 интервалов)

Погрешности, Спектр, Хронология, Лодыжка

Пары идентифицирующих функций

Сравнительный анализ полученных распределений

а) Общий подход

(Напоминание о функции копула как характеристике внутренних связей)

б) Исследование внутренних связей

Эмпирическая и частотные копулы

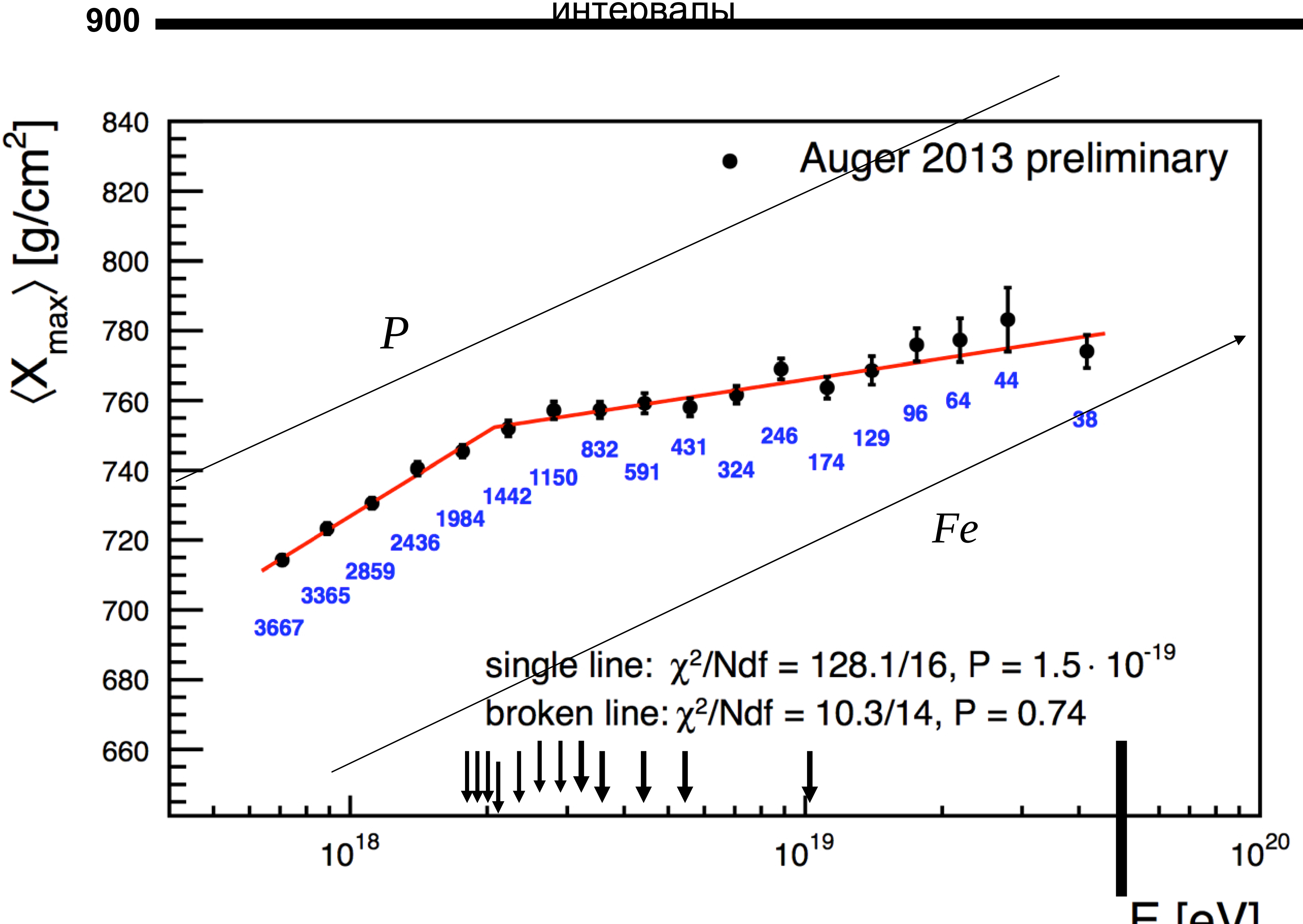
Сравнение крайних интервалов: копул и популярных мер для средних энерговыделений и расстояний от оси.

Физическая интерпретация

Заключение

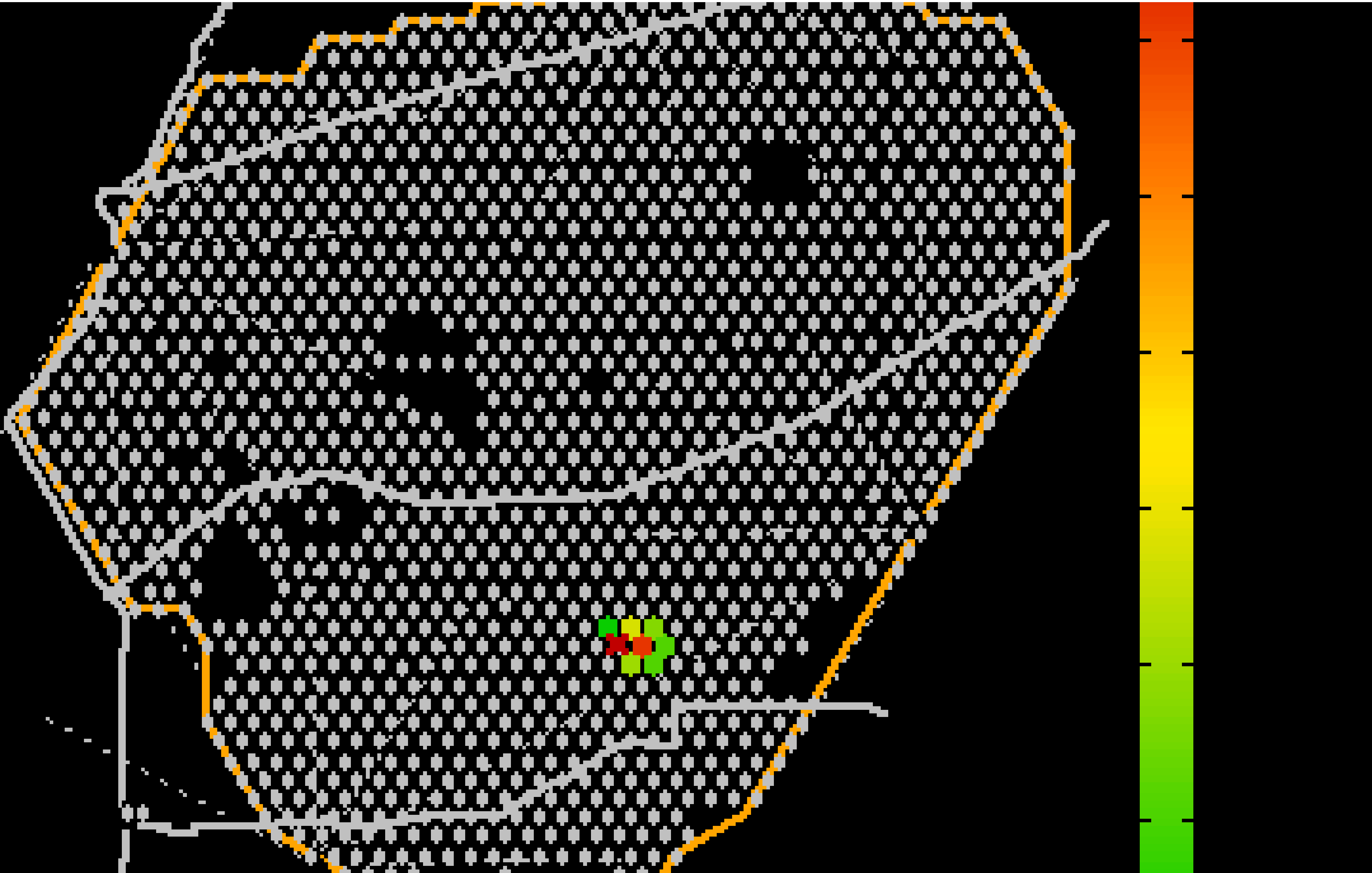
Только качественный уровень

Уровень наблюдения ОЖЕ, энергетическая область, рассматриваемые интервалы



Аргентина Общий вид Событие Исходные данные

Pierre Auger Observatory Public Event Explorer http://auger.colostate.edu/ED/

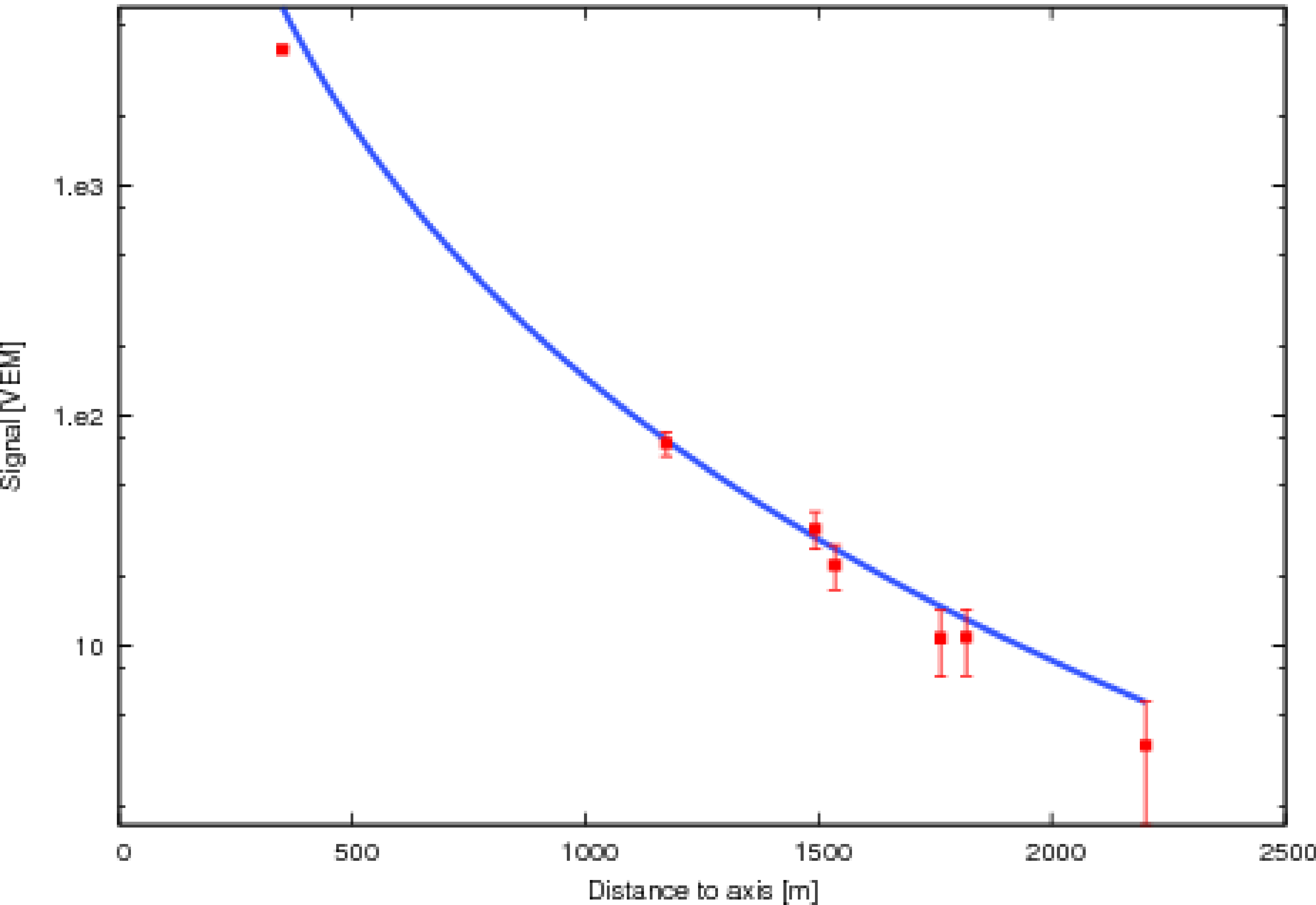


- 08042100: 25.93 EeV, 7 stations, 15.6 deg
- Generic Information
- Id / Date 8042100 / Wed Jul 22 20:27:14 2009
- Nb. of stations 7
- Energy 25.9 ± 2.2 EeV
- Theta 15.6 ± 0.3 deg
- Phi -78.5 ± 1.7 deg
- Curvature 8.3 ± 0.5 km
- Core Easting 480315 ± 29 m
- Core Northing 6086814 ± 33 m

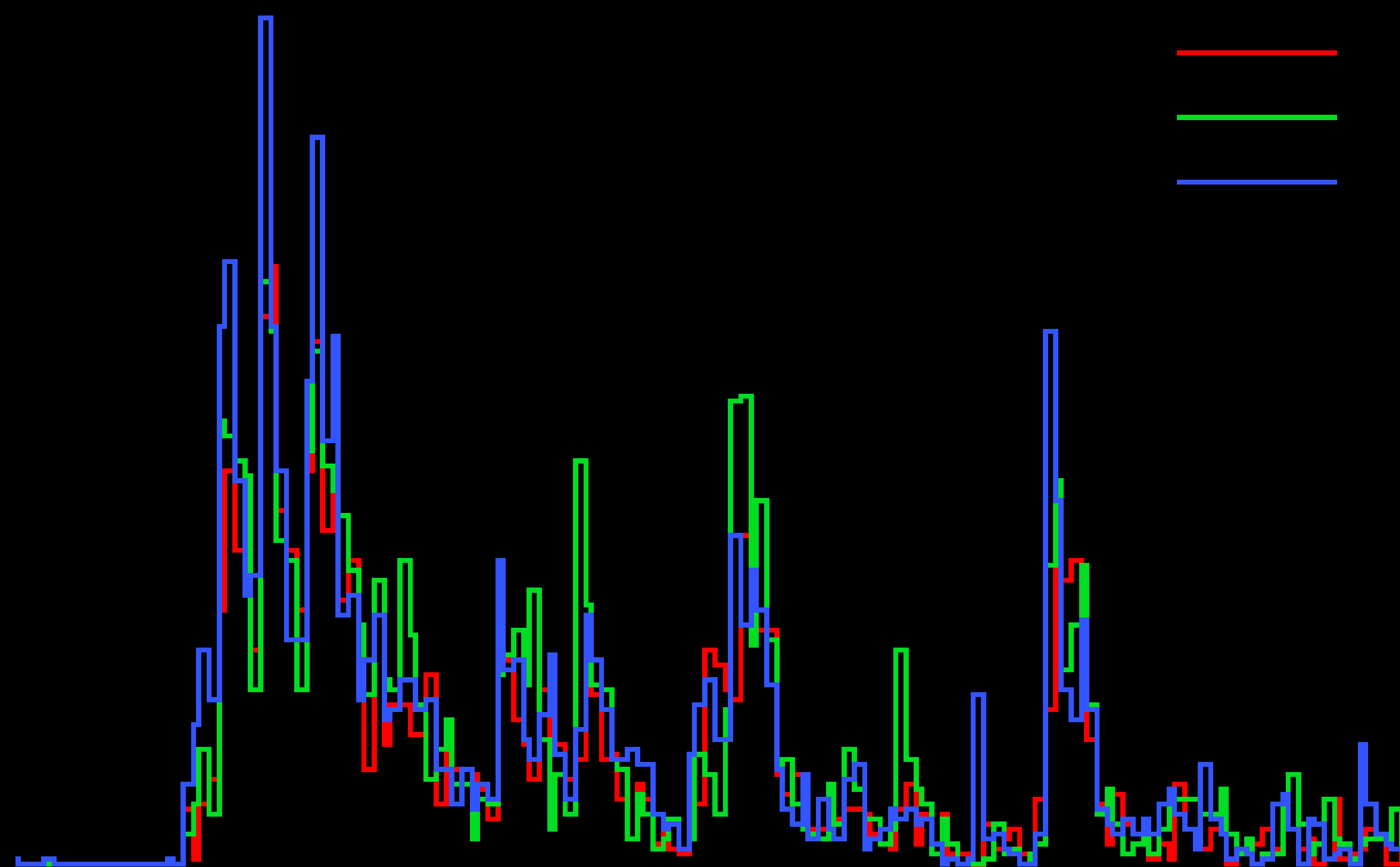
Событие   Индивидуальные данные станций

- # Id Signal(VEM) Time(sec) Time(ns) Easting(m) Northing(m) Altitude(m)
- 263 3927.3 1248294419 795219569 480628.13 6086648.24 1375.45
- 236 76.23 1248294419 795221082 479879.11 6087947.59 1377.94
- 256 32.4 1248294419 795218980 479877.51 6085343.99 1377.36
- 455 22.56 1248294419 795221005 481377.62 6087946.35 1373.93
- 498 10.96 1248294419 795219954 482127.55 6086642.45 1374.22
- 262 10.86 1248294419 795218928 481377.94 6085344.54 1374.39
- 221 3.72 1248294419 795222002 478384.56 6087940.35 1379.98

Событие Величины сигналов и пространственные расстояния до оси



Событие Структура сигнала индивидуальной станции



Рассматриваемые данные содержат ≈ 15% погрешности.
Влияние этих погрешностей на расчёт коэффициента корреляции иллюстрируется следующим рисунком:

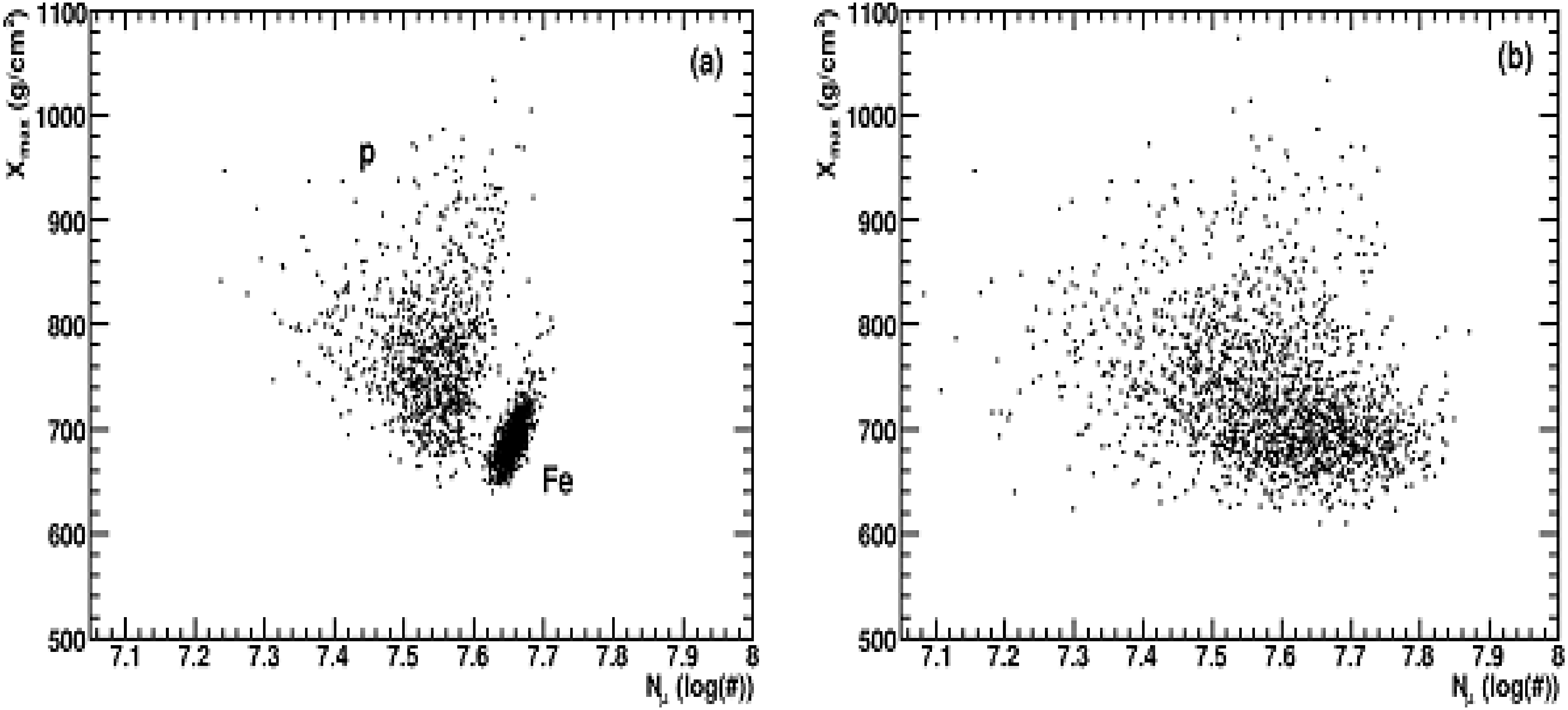
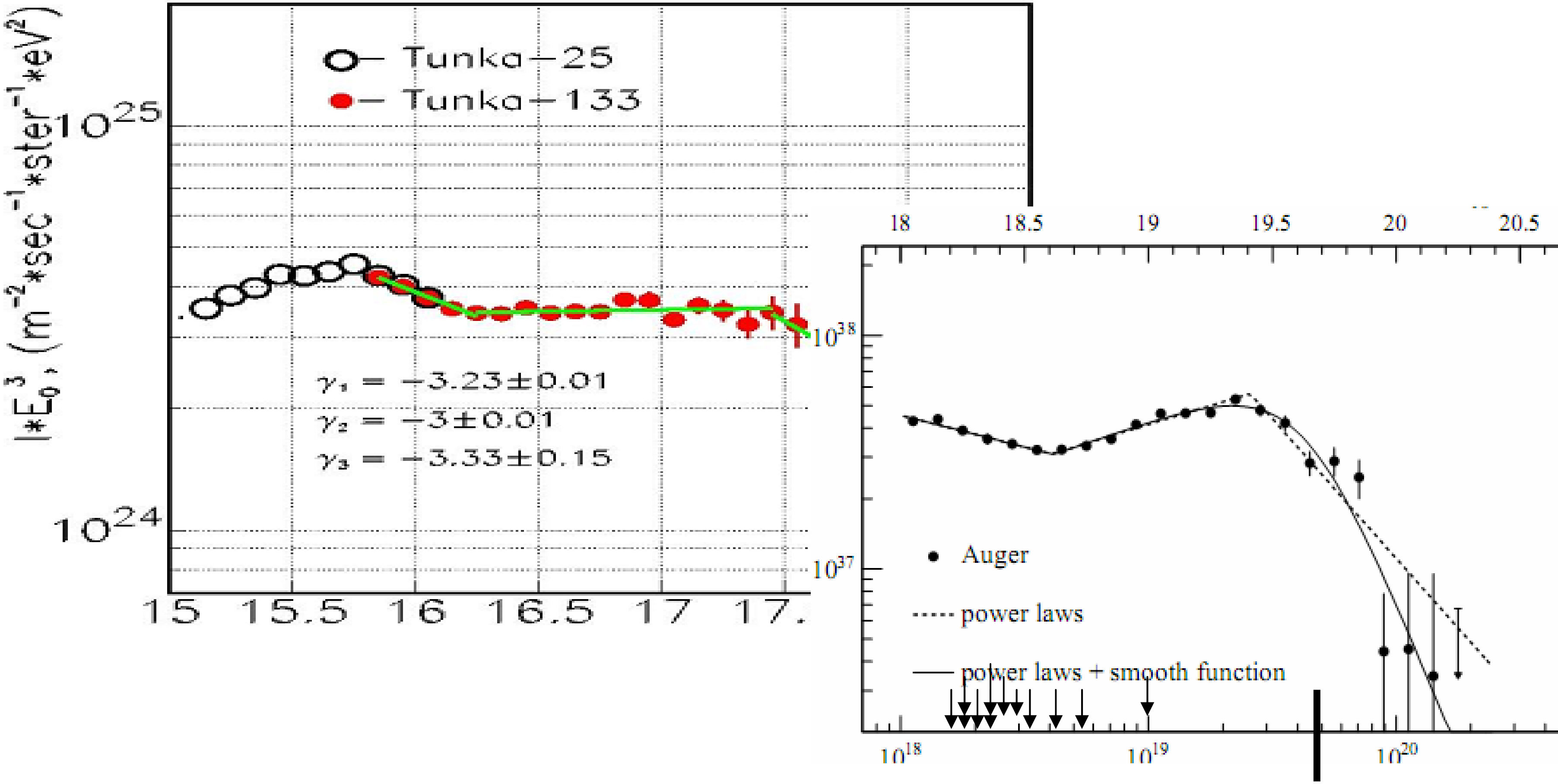


Figure 1: X_{max} - N_μ distribution for proton showers and iron showers with an energy of 10^{19} eV and a zenith angle of 45°. Figure (a) is for an ideal detector (i.e., zero measurement uncertainty). Figure (b) is for a realistic detector.



$6 \cdot 10^{15} < E < 2 \cdot 10^{16}$	$\gamma \approx 3.23$
$2 \cdot 10^{16} < E < 3 \cdot 10^{17}$	$\gamma \approx 3.0$
$3 \cdot 10^{17} < E < 4 \cdot 10^{18}$	$\gamma \approx 3.33 / 3.27$
$4 \cdot 10^{18} < E < 2.56 \cdot 10^{19}$	$\gamma \approx 2.68$
$2.56 \cdot 10^{19} < E$	$\gamma \approx 4.2$

О результатах моделирования ОЖЕ (ICRC 2013)

1 Результатов оч. много: в терминах X_{max}, N_u, S_{em} … при параметрах VeM, зен.уг., p, Fe для различн. моделей взаимодействия, учёт эксперим. погр.

2 Коэффициент корреляции используется как цель и инструмент исследов.

3 Универсальные св-ва ШАЛ в терминах X_{max}, N_u, S_{em}.

4 Разделение мюонной и электромагнитной компонент.

5 Построение пространственного распред. в терминах E_{VeM}, r, N_u, S_{em}

6 Вывод о достаточности описательных возможностей наземных детекторов для установления массового состава.

7 Предложено несколько методов а принципе полезных для решения проблемы массового состава.

8 К сожалению численные результаты только из анализа X_{max} и они модельно зависимы.

Эти результаты надо иметь в виду при интерпретации мер связи исследуемых распределений. В часности:

Доля фотонов энергии выше 1, 2, 3, 5,10 (ЕэВ): 0.4, 0.5, 1.0, 2.6, 8.9 (%)

X_{max}- X_{max}(мюон)=170г/см^2 r_{max}(Fe)-r_{max}(p)=100m.

Лёгкая компонента – узкие, тяжёлая компонента -- широкие ливни, ливни от тяжёлой компоненты развиваются быстрее.

<LnA> ОЖЕ

Heitler model of extensive air showers. In this context Xmax is a linear function of the logarithm of the shower energy per nucleon:

$$\langle X_{\max}\rangle = X_0 + D\,\log_{10}\left(\frac{E}{E_0 A}\right) ,$$

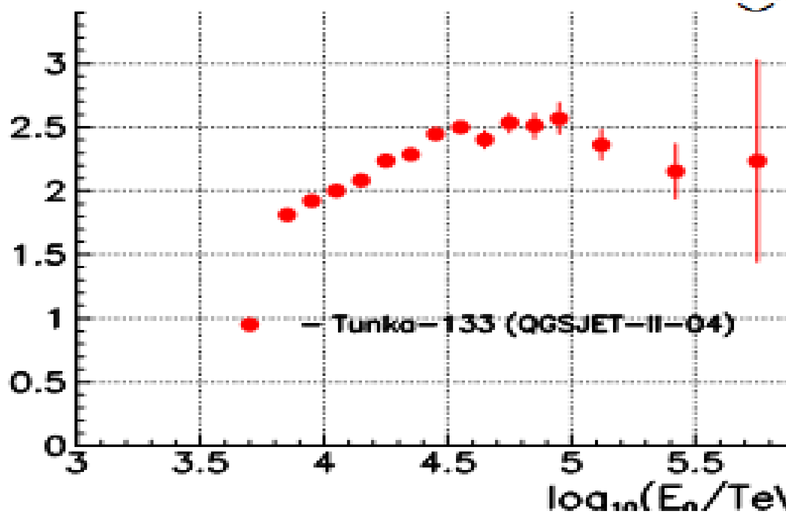
where X_0 is the mean depth of proton showers at energy E_0 and D is the elongation rate, i.e., the change of Xmax per decade of energy.

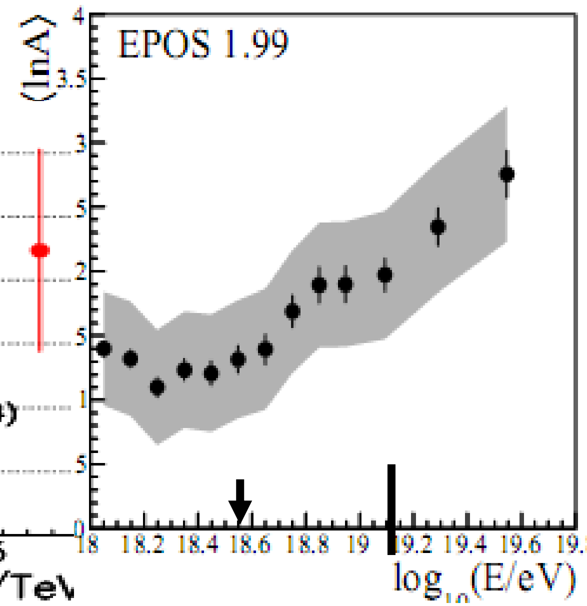
Простейшая модель флуктуаций (учёт дисперсии)

$$\langle X_{\max}\rangle = \langle X_{\max}\rangle_p + f_E \ln A$$

$$f_E = \xi - D/\ln 10 + \delta \log_{10}(E/E$$

parameter	EPOS 1.99
X_0	809.7 ± 0.3
D	62.2 ± 0.5
ξ	0.78 ± 0.24
δ	0.08 ± 0.21





При E>$10^{18.3}$ эВ <lnA> растёт от лёгких к средним массам, а дисперсия lnA убывает во всём диапазоне.

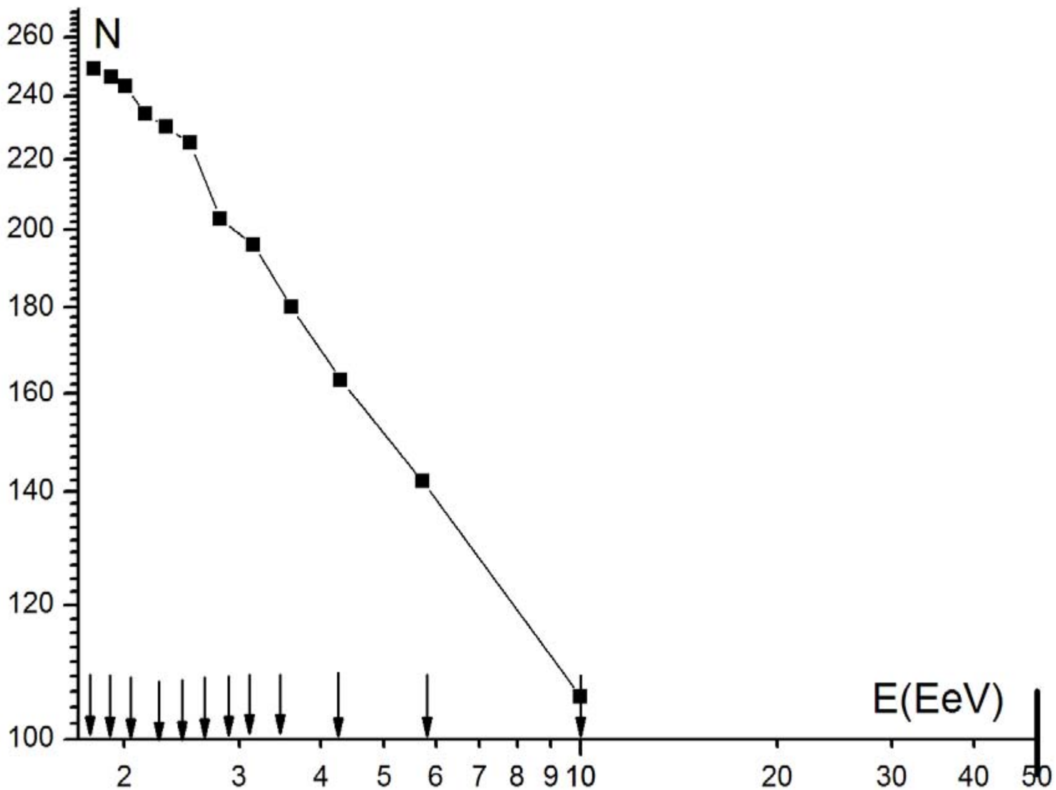
Энергетические интервалы и число ливней в каждом энергетическом интервале

Видно:

Степенной характер убывания.

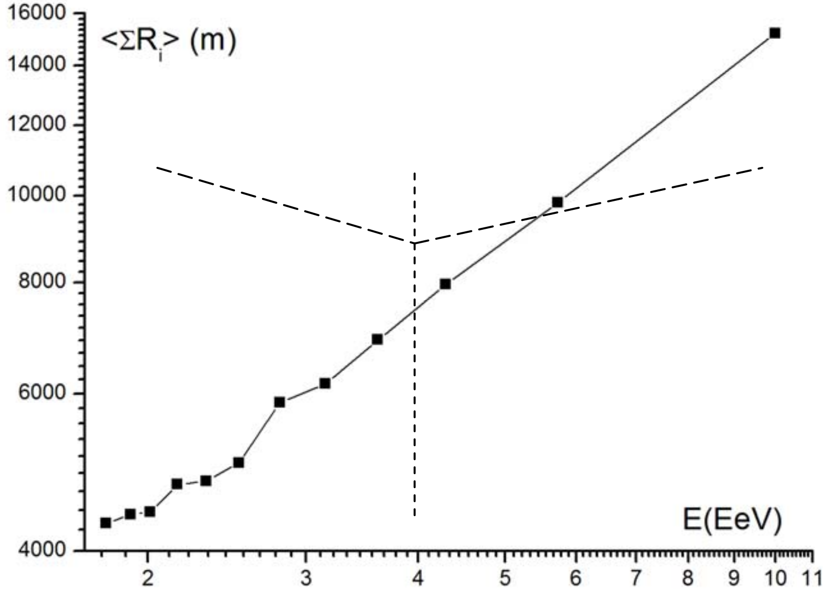
Излом при 2.7-3.1(ЕэВ).

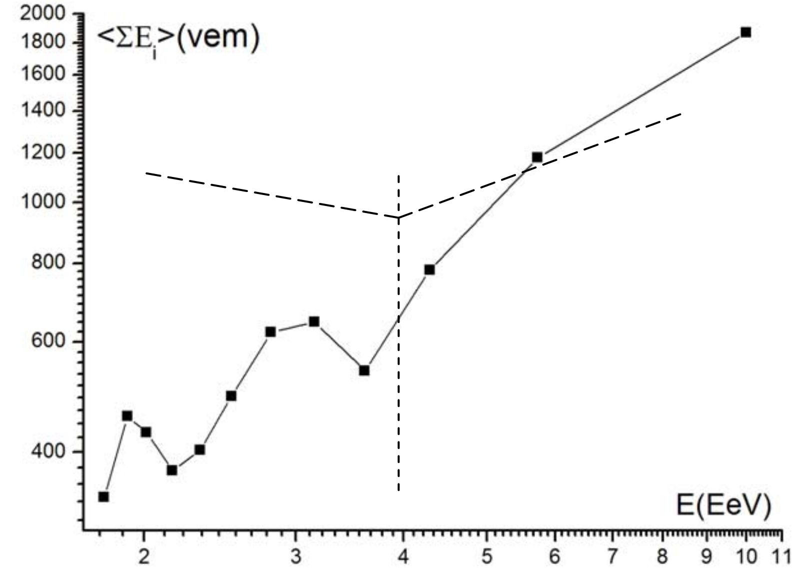
Рост наклона после Излома.



- Стрелки – нижние границы энергетических интервалов.
-
- Излом – излом спектра и массового состава в области «лодыжки».
- Изменение наклона – утяжеление массового состава.
- Это утяжеление может иметь различные причины: дефицит лёгкой компоненты или обогащение тяжёлой компоненты или утяжеление средней компоненты или различные комбинации этих основных причин.

Суммарные характеристики ливня каждого энергетического интервала





Суммарное расстояние датчиков от оси ливня .

Суммарная энергия показаний датчиков

Пунктиром схематически дана лодыжка.

Пары идентифицирующих функций

$$f_x = \frac{1}{n}\sum_{i=1}^{n} r_i \qquad f_y = \lg\left(\frac{1}{n}\sum_{i=1}^{n} e_i\right)$$

$$f_x = \frac{1}{n}\sum_{i=1}^{n} r_i \qquad f_y = \lg\left(\frac{1}{n}\sum_{i=1}^{n} e_i \cdot r_i\right)$$

$$f_x = \frac{1}{n}\sum_{i=1}^{n} r_i \qquad f_y = \lg\frac{1}{n}\sum_{i=1}^{n} \frac{e_i}{r_i})$$

$$f_x = r_{\max} - r_{\min} \qquad f_y = \lg(e(r_{\min}) - e(r_{\max}))$$

$$f_x = \frac{r_{\max} - r_{\min}}{\sum_{i=1}^{n} r_i} \qquad f_y = \frac{e(r_{\min}) - e(r_{\max})}{\sum_{i=1}^{n} e_i}$$

$$f_x = \sum_{i=1}^{n} r_i^2 \Big/ \sum_{i=1}^{n} r_i \qquad f_y = \sum_{i=1}^{n} e_i^2 \Big/ \sum_{i=1}^{n} e_i$$

.где: n – количество датчиков, зафиксировавших ливень

r_i - расстояние i - того датчика до оси ливня (m)

$r_{\min} \setminus r_{\max}$ - расстояние ближайшего к оси \ наиболее удалённого датчика

. e_i - энергия, показанная i - тым датчиком (vem)

. $e(r_{\min}) \setminus e(r_{\max})$ - энергия, показанная ближайшим \ наиболее удалённым датчиком

.

.

.

. **ДАВАЙТЕ РЕКОМЕНДАЦИИ О ДРУГИХ ФУНКЦИЯХ !!!**

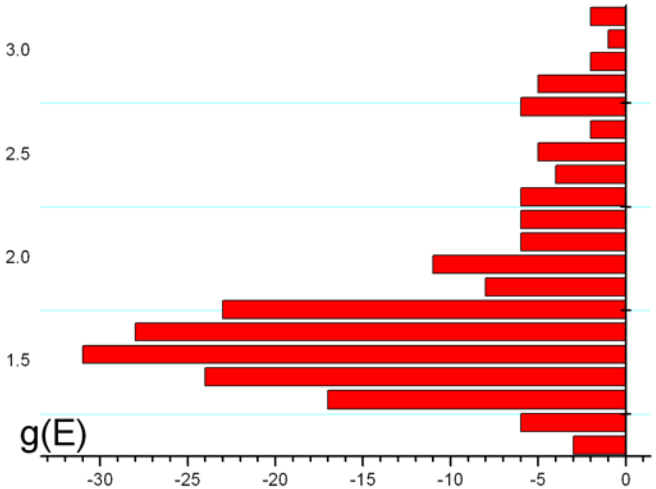
.

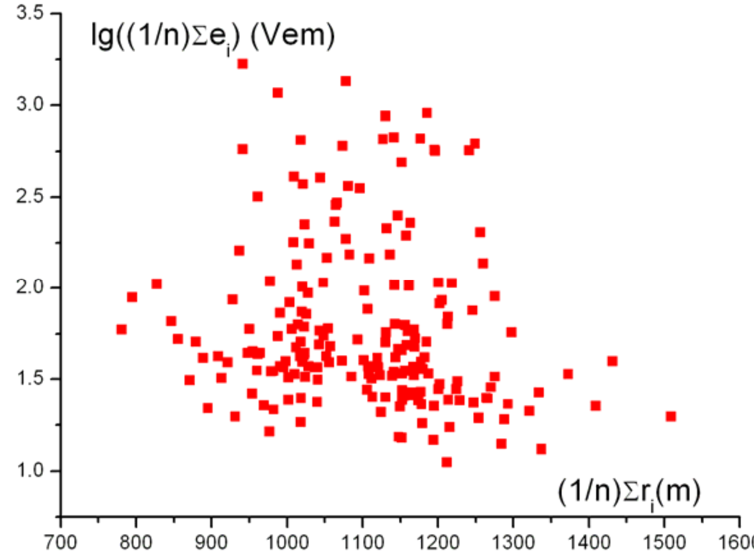
.

.

.

Интервал 8: 3.154 -- 3.609 (ЕэВ) маргиналы





Видны два кластера:

УСЛОВНО

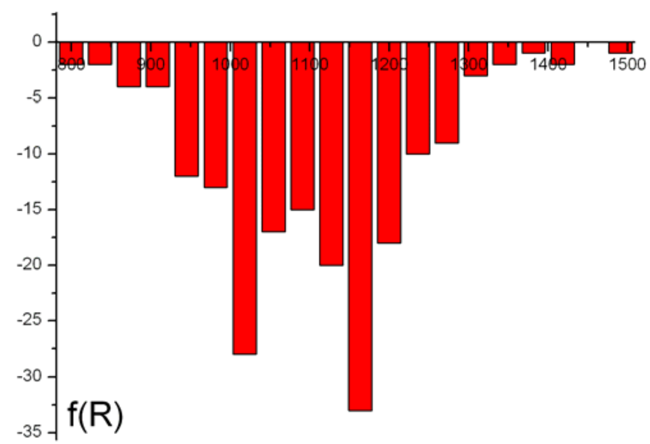
Левый «легкая» компонента

Правый «тяжёлая» компонента

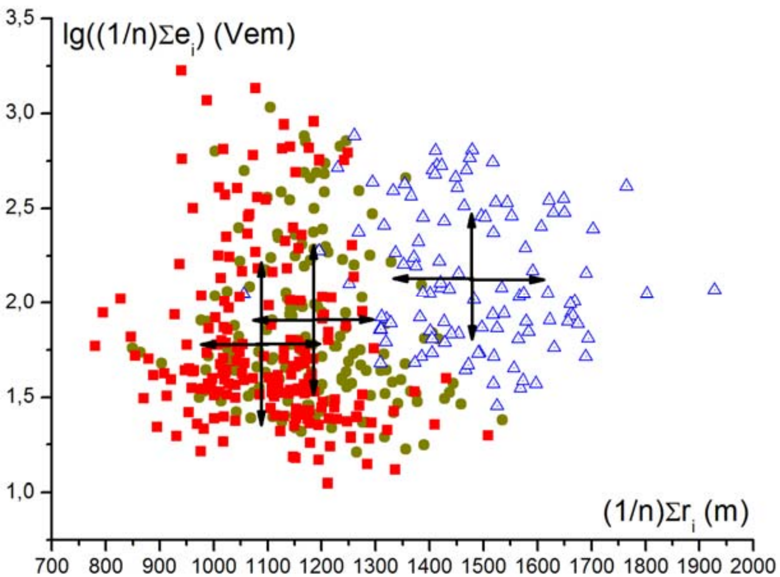
Видно влияние этих кластеров

на форму функций распределения

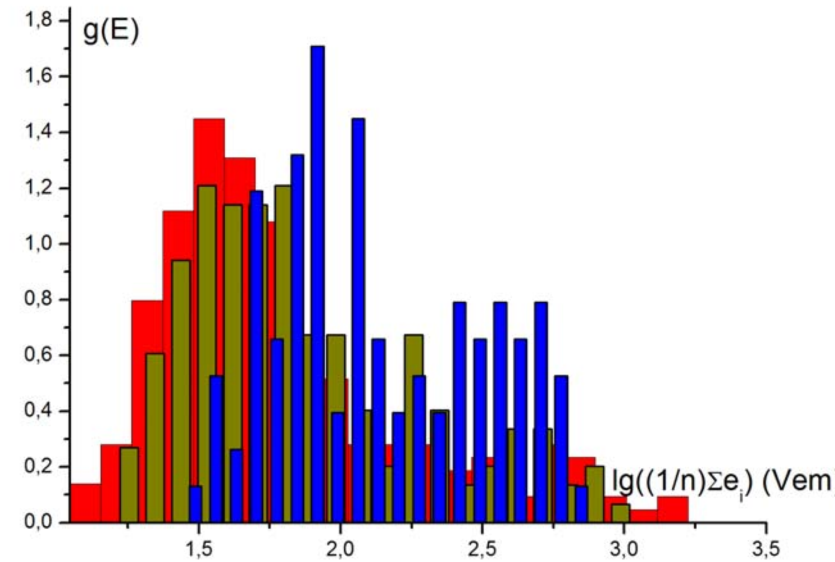
маргиналов .



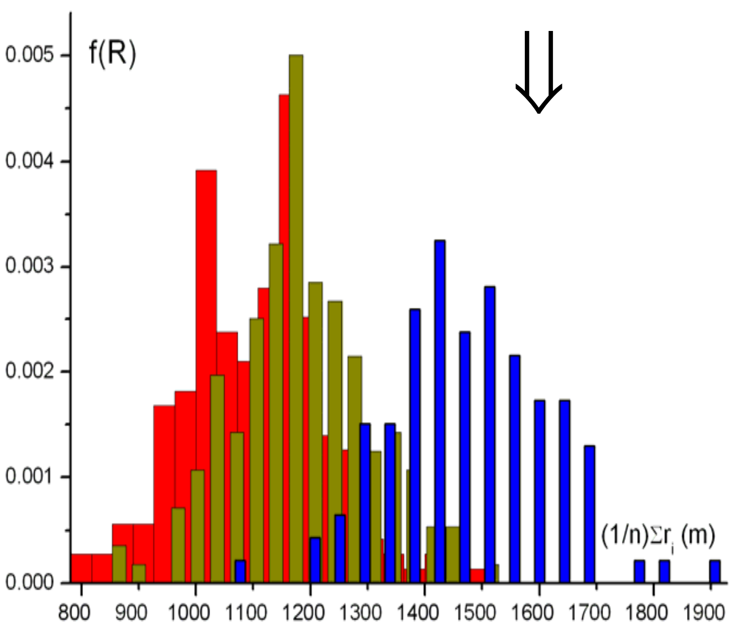
Общая картина эволюции при E>1EeV: (3.16 – 3.64; 4.92 – 7.29; 10.0 – 50.0) ЕэВ



⇒



⇓



кресты – центры тяжести распределении
с указанием дисперсий
Изменение формы распределений –
следствие изменения массового состава.

увелич. мин. расст. ⇒ обеднение лёгк. комп.
Увелич. средн энергии + ………………….
уменьш. максим энергии ⇒ утяжел. масс сост
Уширен. распределений ⇒ утяжел масс.сост
⇓
утяжеление массового состава .

Определение копулы (двумерной)

Двумерная копула **C(u,v)** -- функция $C: I^2 \to I$ такая, что

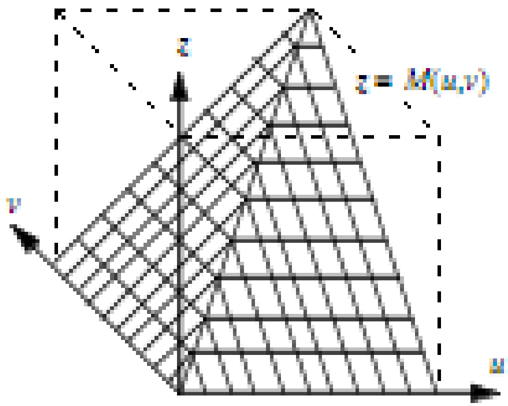
$C(0,t) = C(t,0) = 0$ *(граничное условие)*

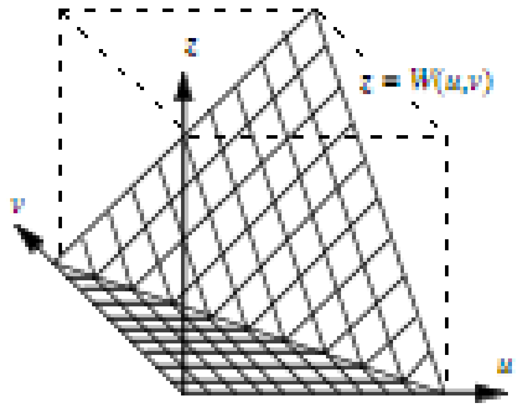
$C(1,t) = C(t,1) = t$ ***(маргиналы – равномерные)***

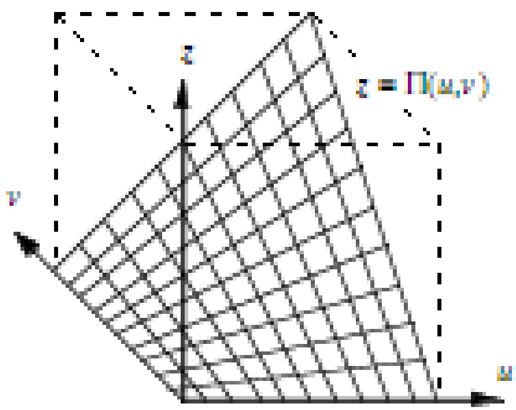
для всех $t \in I$ и

$C(u_2,v_2) - C(u_1,v_2) - C(u_2,v_1) + C(u_1,v_1) \geq 0$ (не убывание)

для всех $u_1, u_2, v_1, v_2 \in I$ и $u_1 \leq u_2 \quad v_1 \leq v_2$







Теорема Скляра

1959

Пусть $H(x,y)$ – совместная функция распределения с непрерывными маргиналами $F(x)$ и $G(y)$, то существует копула $C(F(x),G(y))$, такая, что

$$H(x,y) = C(F(x),G(y))$$

Если маргиналы непрерывны, то это представление единственно.

Верно и обратное:

$$C(u,v) = H(F^{-1}(u),G^{-1}(v))$$

Копула нормального распределения.

Плотности

стандартизованное

$$F(x,y)=\int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y} f(s,t)\,dt\,ds \qquad f(x,y)=\frac{1}{2\pi\sqrt{1-r^2}}\exp\left(-\frac{x^2-2rxy+y^2}{2(1-r^2)}\right)$$

$$C(u,v)=N_\rho(\Phi^{-1}(u),\Phi^{-1}(v))$$

$$=\frac{1}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{\Phi^{-1}(u)}\int_{-\infty}^{\Phi^{-1}(v)}\exp\left[\frac{-(s^2-2\rho st+t^2)}{2(1-\rho^2)}\right]ds\,dt$$

$$\underset{-1}{\overset{W}{\longleftarrow}}\!\!\!-\!\!\!-\!\!\!-\underset{0}{\overset{C}{\vert}}\!\!\!-\!\!\!-\!\!\!\underset{+1}{\overset{M}{\longrightarrow}} \qquad r$$

$$c(u_1,\dots,u_n,\dots,u_N)=\frac{\partial\,\mathbf{C}(u_1,\dots,u_n,\dots,u_N)}{\partial u_1\cdots\partial u_n\cdots\partial u_N}$$

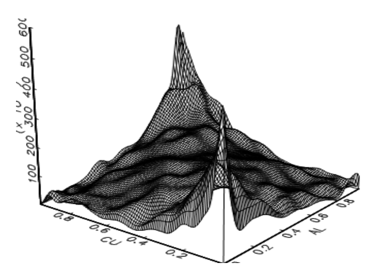
$$H(x,y)=C(F(x),G(y))$$

$$f(x_1,\dots,x_n,\dots,x_N)=c(\mathbf{F}_1(x_1),\dots,\mathbf{F}_n(x_n),\dots,\mathbf{F}_N(x_N))\prod_{n=1}^{N} f_n(x_n)$$

Двумерный случай **общий,**

не независимый

$$\frac{\partial^2 H(x,y)}{\partial x\partial y}=h(x,y)=c(x,y)\cdot f(x)g(y)$$



$$\neq 1$$

Продуктивность формализма копулы определяется тем, что известны способы измерения копул, т.е. измерение связи частных распределений конкретной многомерной функции сводится к измерениям копулы этой функци

Известна общая теория построения таких мер и можно строить нужные новые меры, определяемые спецификой задачи.

Мы воспользуемся известными наиболее популярными мерами:

Кендала, Спирмана, Гини.

Определения:

Кендал $\tau = \tau_{X,Y} = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]$

Здесь векторы (X_1,Y_1) и (X_2,Y_2) – независимые, распределённые по одной и той же совместной функции H(x,y).

Спирман $\rho_{X,Y} = 3\left(P[(X_1 - X_2)(Y_1 - Y_3) > 0] - P[(X_1 - X_2)(Y_1 - Y_3) < 0]\right)$

Здесь векторы (X_1,Y_1), (X_2,Y_2) и (X_3,Y_3) – независимые, распределённые по одной и той же совместной функции H(x,y). Заметим, что вектор (X_2,Y_3) получается распределённым по независимой функции $F(x)G(y)$.

Гини $g = 2E(|U + V - 1| - |U - V|)$,

где U и V – обратные функции маргиналов F(x) и G(y) распределения H(x,y),

Здесь и далее E(….) – математическое ожидание.

Выражения мер через копулу и интерпретация.

$$\tau = 4\iint_{I^2} C(u,v)dC(u,v) - 1 = 4E(C(u,v)) - 1 = 1 - 4\iint_{I^2} \frac{\partial}{\partial u}C(u,v)\frac{\partial}{\partial v}C(u,v)dudv$$

$$\rho = 12\iint_{I^2} uvdC(u,v) - 3 = 12E(u,v) - 3 = 12\iint_{I^2} (C(u,v) - uv)dudv$$

$$\gamma = 4(\int_0^1 C(u,1-u)du - \int_0^1 (u - C(u,u))du)$$

Использование этих мер аналогично использованию общеупотребительной меры линейной связи случайных векторов – коэффициента корреляции r -- меры Пирсона, интерпретации которой общеизвестны. Приведённые выше меры описывают более сложные связи. Интерпретации их, получаемые аналогично интерпретации меры Пирсона, также более сложны.

Например, из определения видно, что мера **Кендала** может интерпретироваться как <u>вероятность соответствия</u> больших (или меньших) значений одного вектора большим (или меньшим) значениям другого {Б $\Leftrightarrow$ Б } или {м $\Leftrightarrow$ м } <u>минус</u> <u>вероятность такого же несоответствия.</u>

.Аналогичная интерпретация меры **Спирмана** более сложна (но в некоторых случаях полезна). Здесь можно обратить внимание на то, что многомерное интегрирование по dC(u,v) есть математическое ожидание. В дан ном случае математическое ожидание такого обстоятельства, что в исследуемом процессе эти (фиксированные) маргиналы будут независимы.

Можно получить и другую интерпретацию этой же меры: на сколько данная

копула отличается от независимой копулы. Более точно: каков объём меж-

ду поверхностями данной копулы и независимой. (Нормировки на отрезок

[-1,+1] в этих интерпретациях – очевидны.)

Мера Гини может пониматься как мера асимметричности копулы.

$$C_n\left(\frac{i}{n},\frac{j}{n}\right)=\frac{\text{number of pairs } (x,y) \text{ in the sample with } x \le x_{(i)},\, y \le y_{(j)}}{n}.$$

where $x_{(i)}$ and $y_{(j)}$, $1 \le i,j \le n$, denote order statistics from the sampl

ь

The *empirical copula frequency* c_n is given by

$$c_n\left(\frac{i}{n},\frac{j}{n}\right)=\begin{cases}1/n, & \text{if } (x_{(i)},y_{(j)}) \text{ is an element of the sample,}\\ 0, & \text{otherwise.}\end{cases}$$

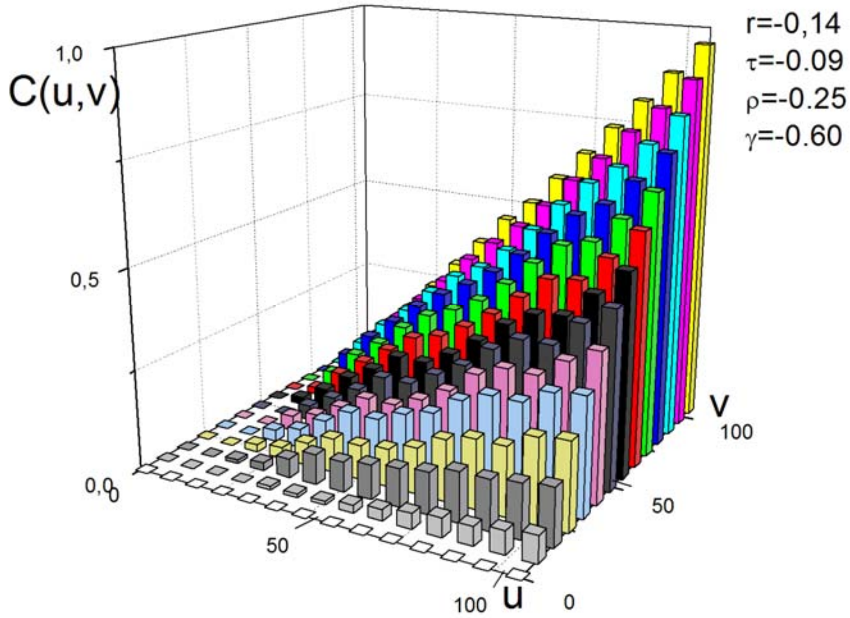
Приведём расчетные формулы для выбранных мер.

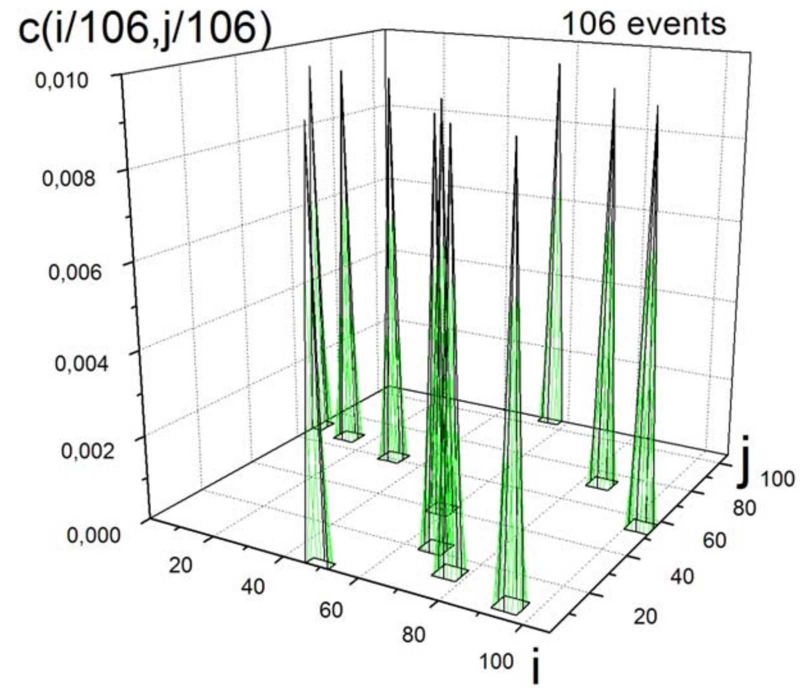
$$\tau = \frac{2n}{n-1}\sum_{i=2}^{n}\sum_{j=2}^{n}\sum_{p=1}^{i-1}\sum_{q=1}^{j-1}\left(c\left(\frac{i}{n},\frac{j}{n}\right)c\left(\frac{p}{n},\frac{q}{n}\right)-c\left(\frac{i}{n},\frac{q}{n}\right)c\left(\frac{p}{n},\frac{j}{n}\right)\right)$$

$$\rho = \frac{12}{n^2-1}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(C\left(\frac{i}{n},\frac{j}{n}\right)-\frac{i}{n}\frac{j}{n}\right)$$

$$\gamma = \frac{2n}{[n^2/2]}\left(\sum_{i=1}^{n-1}C\left(\frac{i}{n},1-\frac{i}{n}\right)-\sum_{i=1}^{n}\left(\frac{i}{n}-C\left(\frac{i}{n},\frac{i}{n}\right)\right)\right)$$

Копула и частотная копула наиболее энергичного интервала (10-50)ЕэВ





. Эмпирическая (слева) и частотная (справа) копулы функции распределения <ri> - средних значений расстояний сработавших детекторов и lg<ei> - логарифмов средних зарегистрированных ими энергий от ливней в интервале (10.0 – 50.0) ЕэВ

Из рисунка видно, что рассматриваемая копула несколько ($\rho = -0.25$) ниже четверти седла (независимой копулы), причём в области побочной диагонали ($u + v = 1$) наблюдается некоторое возвышение особенно заметное у концов диагонали. Количественно это подтверждается приведёнными значениями мер. Частотная копула явно показывает обеднение на главной и обогащение на побочной что подтверждается значением ($|\gamma| > 0.5$) меры Гини.

Заключение

- Скромная статистика (1% всех имеющихся у коллаборации ОЖЕ данных, 106 событий в самом высокоэнергичном интервале $(1 – 5)\cdot10^{19}$эВ) заставляет ограничиться только качественным уровнем исследования.
- Резких, существенных изменений массового состава имеющиеся данные (с учётом отмеченного выше о недостаточности доступной статистики при $E_0 > 2\cdot10^{19}$ (эВ)) в интервале $5\cdot10^{18}$ -- $5\cdot10^{19}$(эВ) не показывают, но указывают на утяжеление массового состава. Это утяжеление развивается постепенно с ростом энергии, имеет сложный характер, и происходит в основном за счёт обогащения более тяжёлой компонентой группы средних ядер.

Эксперимент $\Rightarrow$ функция распределения

Функция распределения (или её плотность) $\Rightarrow$ маргиналы,

но маргиналы $\not\Leftrightarrow$ функция распределения

Для определения функции распределения необходимо знание межкоор-динатных (структурных, внутренних) связей. (Коэффициент корреляции пример описания линейных внутренних связей.)

Полностью внутренние связи (и только они) описываются копулой – функцией, определяемой исходным распределением.

Функция распределения $\Leftrightarrow$ маргиналы + копула

- **Теорема Скляра (маргиналы – непрерывны)**

функция распределения $\Leftrightarrow$ **единственный набор маргиналов** $+$ **единственная копула**

«обратно»

любой набор n одномерных распределений (маргиналов) + любая копула

$\Downarrow$

единственная функция распределения.

Копула – тоже специфическая функция распределения той же размерно-сти, что и исходная, специфика в том, что все её маргиналы U_{01} -- равномерные распределения на [0,1].