

## Report on Tests and Measurements of Hadronic Interaction Properties with Air Showers

H.P. Dembinski<sup>1,\*</sup>, J.C. Arteaga-Velázquez<sup>2</sup>, L. Cazon<sup>3</sup>, R. Conceição<sup>3</sup>, J. Gonzalez<sup>4</sup>, Y. Itow<sup>5</sup>, D. Ivanov<sup>6</sup>, N.N. Kalmykov<sup>7</sup>, I. Karpikov<sup>8</sup>, S. Müller<sup>9</sup>, T. Pierog<sup>9</sup>, F. Riehn<sup>3</sup>, M. Roth<sup>9</sup>, T. Sako<sup>10</sup>, D. Soldin<sup>4</sup>, R. Takeishi<sup>11</sup>, G. Thompson<sup>6</sup>, S. Troitsky<sup>8</sup>, I. Yashin<sup>12</sup>, E. Zadeba<sup>12</sup>, and Y. Zhezher<sup>7</sup> for the EAS-MSU<sup>13</sup>, IceCube<sup>14</sup>, KASCADE-Grande<sup>15</sup>, NEVOD-DECOR<sup>16</sup>, Pierre Auger<sup>17</sup>, SUGAR<sup>18</sup>, Telescope Array<sup>19</sup>, and Yakutsk EAS Array<sup>20</sup> collaborations

<sup>1</sup>Max Planck Institute for Nuclear Physics, 69117 Heidelberg, Germany

<sup>2</sup>Institute of Physics and Mathematics, Universidad Michoacana, C.P. 58040 Morelia, Michoacan, Mexico

<sup>3</sup>LIP, Lisbon, Portugal

<sup>4</sup>Bartol Institute, University of Delaware, Delaware, USA

<sup>5</sup>Nagoya University, Nagoya, Japan

<sup>6</sup>High Energy Astrophysics Institute and Department of Physics and Astronomy, University of Utah, Salt Lake City, Utah, USA

<sup>7</sup>D.V. Skobeltsyn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University, Moscow 119991, Russia

<sup>8</sup>Institute for Nuclear Research of the Russian Academy of Sciences, 60th October Anniversary Prospect 7a, Moscow 117312, Russia

<sup>9</sup>Karlsruhe Institute of Technology, Eggenstein-Leopoldshafen, Germany

<sup>10</sup>Institute for Cosmic Ray Research, University of Tokyo, Tokyo, Japan

<sup>11</sup>Department of Physics, Sungkyunkwan University, Jang-an-gu, Suwon, Korea

<sup>12</sup>National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Moscow 115409, Russia

<sup>13</sup>EAS-MSU experiment, Moscow, Russia

<sup>14</sup>IceCube Neutrino Observatory, Madison WI, USA

Full author list: <https://icecube.wisc.edu/collaboration/authors/icecube>

<sup>15</sup>KASCADE-Grande experiment, Karlsruhe, Germany

<sup>16</sup>NEVOD-DECOR experiment, Moscow, Russia

<sup>17</sup>Pierre Auger Observatory, Malargüe, Argentina

Full author list: [http://www.auger.org/archive/authors\\_2018\\_10.html](http://www.auger.org/archive/authors_2018_10.html)

<sup>18</sup>SUGAR Array, Sidney, Australia

<sup>19</sup>Telescope Array Project, Salt Lake City UT, USA

Full author list: <http://www.telescopearray.org/research/collaborators>

<sup>20</sup>Yakutsk EAS Array, Yakutsk, Russia

Full author list: <https://ikfia.ysn.ru/en/theyakutskarrayteam>

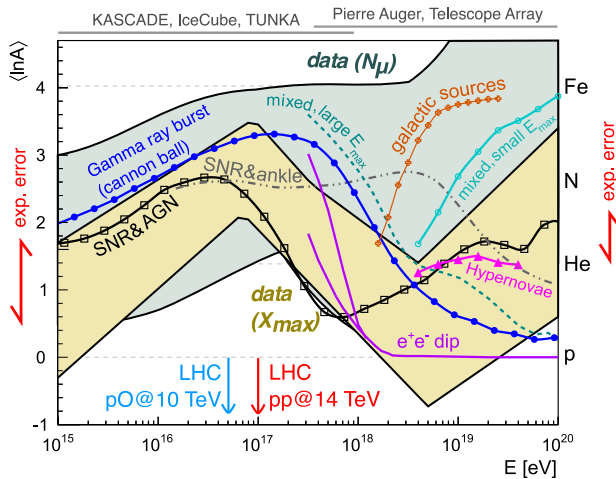
**Abstract.** We present a summary of recent tests and measurements of hadronic interaction properties with air showers. This report has a special focus on muon density measurements. Several experiments reported deviations between simulated and recorded muon densities in extensive air showers, while others reported no discrepancies. We combine data from eight leading air shower experiments to cover shower energies from PeV to tens of EeV. Data are combined using the  $z$ -scale, a unified reference scale based on simulated air showers. Energy-scales of experiments are cross-calibrated. Above 10 PeV, we find a muon deficit in simulated air showers for each of the six considered hadronic interaction models. The deficit is increasing with shower energy. For the models EPOS-LHC and QGSJet-II.04, the slope is found significant at 8 sigma.

### 1 Introduction

Cosmic rays with energies larger than  $10^{15}$  eV can only be indirectly observed via extensive air showers. A detailed understanding of the hadronic physics in an air shower is needed to infer the energy and mass of the cosmic ray from air shower measurements. The fluorescence technique has reduced the model-dependence for the energy measurement by tracking the longitudinal shower development, but inferring the mass accurately is still a challenge.

The energy-dependent mass composition of cosmic rays carries a unique imprint from the origin and propagation of cosmic rays. Attempts to measure the mass composition are complementary to direct searches for cosmic-ray sources via coincident observation in a multi-messenger approach or statistical correlation with potential sources. In Fig. 1, predictions (lines and markers) are shown for the mean-logarithmic-mass  $\langle \ln A \rangle$  from different theories. A measurement to an accuracy of 10 % of the proton-iron difference is technically possible, but uncertainties in hadronic interactions prevent the full exploitation of exist-

\*e-mail: [hdembins@mpi-hd.mpg.de](mailto:hdembins@mpi-hd.mpg.de)



**Figure 1.** Mass composition of cosmic rays quantified by  $\langle \ln A \rangle$  as a function of cosmic-ray energy  $E$ . Model predictions (markers and lines) are compared to data bands, all taken from the review by Kampert and Unger [1]. Vertical arrows at the sides indicate the instrumental error achieved by leading experiments at low and high energies. The figure is discussed in the text.

ing measurements. The two leading observables to infer  $\langle \ln A \rangle$  are the depth  $X_{max}$  of the shower maximum in the atmosphere (yellow band in Fig. 1), and the number  $N_\mu$  of muons produced in the shower (green band in Fig. 1). The width of those bands has two main contributions: the experimental uncertainties, and the uncertainties inherent in converting the air shower observables into  $\langle \ln A \rangle$ , which requires air shower simulations with hadronic interaction models. Leading models are EPOS-LHC [2], QGSJet-II.04 [3], and SIBYLL-2.3c [4]. Their variation is dominating the uncertainty. Moreover, the mass composition obtained with  $X_{max}$  and  $N_\mu$  is not consistent, which implies that the models are not correctly describing all aspects of hadronic physics in air showers.

Tests and measurements of hadronic interaction properties with air showers address these issues. They can reduce uncertainties in modeling the air shower development, so that accurate estimates of the cosmic-ray mass can be obtained. They also offer opportunities to test the standard model of particle physics under extreme conditions. The cms-energy in the first interaction of an air shower initiated by a  $10^{20}$  eV cosmic ray is 432 TeV in the nucleon-nucleon system, 33 times higher than what is currently accessibly at the LHC.

In this report, we will review recent tests and measurements of hadronic interaction properties, starting with the measurements of the electromagnetic shower component, but then focusing on muon density measurements. There is a long-standing problem with the correct simulation of muons in air showers. The HiRes/MIA collaboration already reported a discrepancy in simulated and measured air showers between  $10^{17}$  to  $10^{18}$  eV in the year 2000 [5]. The NEVOD-DECOR experiment reported an increase of muon density relative to simulation from  $10^{15}$  to  $10^{18}$  eV in 2010 [6]. Above  $10^{17}$  eV, an excess of multi-muon events

at energies around  $10^{18}$  eV over the expectation was observed, and a muon deficit in simulations was suggested as an explanation. The experiments KASCADE-Grande [7] and EAS-MSU [8] reported no muon discrepancy in this energy range when the latest hadronic interaction models tuned to LHC-data are used to simulate air showers, while the SUGAR array [9] reported a muon deficit even for these models. The Pierre Auger Observatory [10, 11] and Telescope Array [12] also observed a muon deficit in  $10^{19}$  eV showers simulated with the latest models. These measurements, preliminary data from IceCube [13] and AMIGA [14], and unpublished data from Yaktusk [15] are systematically compared in this report.

## 2 Measurements of electromagnetic shower component

Most measurements of the electromagnetic component of air showers show good or acceptable agreement with simulations, especially when the post-LHC generation of hadronic interaction models is used. We list recent measurements briefly and focus on deviations.

**Proton-air cross-section** The proton-air cross-section has been measured based on the slope of the tail of the  $X_{max}$ -distribution [16, 17] by the Pierre Auger Observatory and Telescope Array. Conceptually, this is the most direct measurement of a hadronic interaction property with air showers, the dependence of the analysis on uncertainties in the mass-composition and the shower development is reduced. The measurements start to constrain hadronic interaction models.

**Moments of  $X_{max}$ -distribution** The first two moments of the  $X_{max}$  distribution measured by the Pierre Auger Observatory have been mapped to the first two moments of the  $\ln A$ -distribution with parameters from simulations [18]. Some unphysical second moments were found for QGSJet-II.04.

**Longitudinal shape** The average longitudinal shapes of air showers recorded by the Pierre Auger Observatory have been compared to simulations, more specifically the width of the profile around the maximum and the asymmetry of the rising and falling edge [19, 20]. The measurements are compatible with simulations using post-LHC models.

**Lateral density profile** The slope of the lateral density profile of electrons and photons is sensitive to the cosmic-ray mass. At the IceCube Neutrino Observatory, measurements of the slope with the surface detector array in 2.8 km altitude a.s.l. were compared to measurements of muon bundles below a kilometer of ice, which are also sensitive to the mass [21]. SIBYLL-2.1 and EPOS-LHC are inconsistent with the combined measurements.

**Attenuation with zenith angle** The attenuation of the signals measured at ground with increasing zenith angle have been compared by Telescope Array up to  $45^\circ$  [22], and up to  $40^\circ$  by KASCADE-Grande [7]. No deviations to simulations were found.

### 3 Measurements of muonic shower component

Most measurements of the muonic component of air showers show disagreement to simulations. The deviations are difficult to reduce to a single cause, further research on muons in air showers is needed.

**Lateral density** A large body of data on lateral density measurements is available and has been converted into a comparable format. The comparison of lateral density measurements will be discussed in the next section.

**Production depth or height** The muon production height has been inferred from signal arrival times in ground detectors of the Pierre Auger Observatory [23], and via a muon tracking detector within KASCADE-Grande [24]. The Pierre Auger Observatory converts the measured heights into equivalent slant depths, which is recommended, since hadronic cascades develop along slant depth. Disagreement to simulations are found for QGSJet-II.02 [25] at PeV energies, and for QGSJet-II.04 and EPOS-LHC at EeV energies.

**Attenuation with zenith angle** The attenuation of the muon lateral profile with growing zenith angle has been measured by KASCADE-Grande up to  $40^\circ$  [7]. Simulations show larger attenuation factors than data for the models QGSJet-II.02, QGSJet-II.04, SIBYLL-2.1 and EPOS-LHC. The attenuation has a dependence on the lateral distance to the shower axis, which is not reproduced by simulations.

**Multiplicity in muon bundles** The ALICE experiment has recently measured the multiplicity in muon bundles [26]. The relative abundance of high and low multiplicity events is well reproduced by QGSJet-II.04.

**Atmospheric flux of TeV muons** The IceCube Neutrino Observatory has measured the atmospheric flux of TeV muons [27, 28]. The total flux is well reproduced by SIBYLL-2.1. Some tension is found in the zenith-angle dependent flux, which could potentially be fixed by adding charmed mesons and their decays to the simulations. Contributions of charmed mesons are negligible for most air shower measurements, but contribute significantly to multi-TeV muons and neutrinos.

**Lateral separation of TeV muons** The IceCube Neutrino Observatory has observed events with laterally separated muons from muon bundles [29]. The events are interpreted as single muons with high transverse momentum, which dominantly originate from the first interaction in an air shower. The statistical distribution of the lateral separation is related to the transverse momentum distribution of hadrons produced in the first interaction. Partial agreement is found for SIBYLL-2.1 and 2.3, disagreement for EPOS-LHC and QGSJet-II.04.

**Rise-time of shower front** The Pierre Auger Observatory has published mass-composition measurements based on the normalized rise-time [30] and the rise-time asymmetry [31]. The rise-time is the time interval in which the

collected surface detector charge rises from 10 % to 50 % of the final value. Muons in an air shower tend to arrive before electrons and photons, so the rise-time is a tracer of the muon content of the shower. Both analyses show disagreement between data and simulations for QGSJet-II.04 and EPOS-LHC. At lateral distances larger than 1000 m from the shower axis, however, the rise-time asymmetry is compatible with QGSJet-II.04 predictions.

### 4 Measurements of muon lateral density

We are analyzing the following measurements of the lateral muon density from eight cosmic-ray experiments:

- EAS-MSU [8]
- IceCube Neutrino Observatory [13]
- KASCADE-Grande [7]
- NEVOD-DECOR [6, 32]
- Pierre Auger Observatory & AMIGA [10, 11, 14]
- SUGAR [9]
- Telescope Array [12]
- Yakutsk, based on preliminary unpublished data [15].

We further show HiRes-MIA data [5] for comparison in several plots, but exclude them from the final results. The HiRes-MIA result systematically differs from all more recent measurements and we did not succeed in contacting one of the authors to better understand the differences.

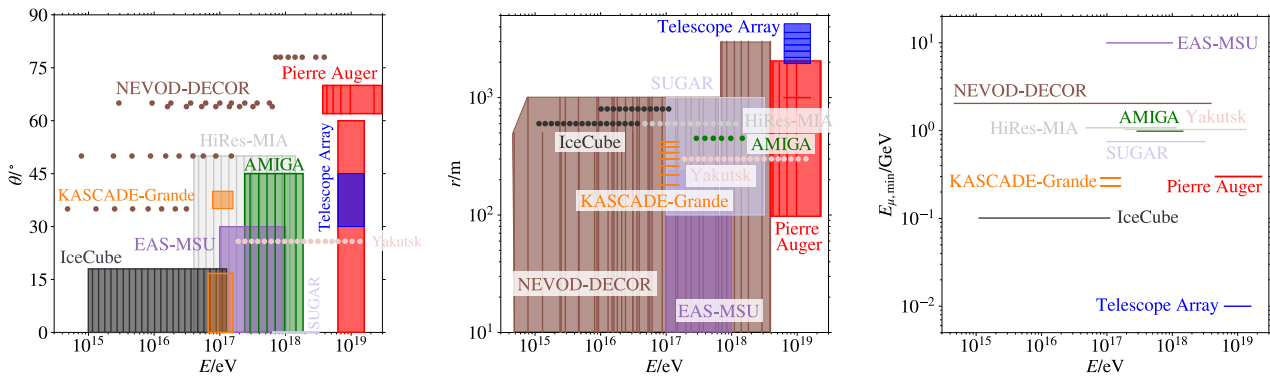
A direct comparison of muon measurements is not possible, since the muon measurements are performed under very different conditions and using different techniques. The muon density at the ground depends on many parameters which differ from experiment to experiment:

- Cosmic-ray energy  $E$ ,
- Zenith angle  $\theta$ ,
- Shower age (depends on altitude of the experiment, local atmosphere, and zenith angle of the shower),
- Lateral distance  $r$  from shower axis,
- Energy threshold  $E_{\mu,\min}$  of the detectors for muons.

Since direct comparisons of the measured muon density are unfeasible, each experiment usually compares to air shower simulations. The data/MC ratio is comparable between different analyses and different experiments. In a way, air shower simulations provide a universal reference. The caveat of this approach is that two measurements are only comparable, if simulations with the same hadronic interaction model are available for both.

How the measurements cover the space of parameters is shown in Fig. 2. The measurements as a whole cover most of the parameter space. This is very valuable, since it allows us to detect a possible dependence of the data/MC ratio along all dimensions of the parameter space.

Comparing data/MC ratios instead of just the data introduces a complication. The value of the ratio depends on how the corresponding air shower simulations are selected. While most of the shower parameters can be easily



**Figure 2.** Air shower experiments have measured the muon density at ground under various conditions, which are shown here. Points and lines indicate a measurement in a narrow bin of the parameter, while boxes indicate integration over a parameter range. *Left:* Zenith angle of air showers versus shower energy. *Middle:* Lateral distance of the muon density measurement versus shower energy. *Right:* Energy threshold for the muons that are counted in the experiment. Some experiments measure muons below a shielding, which increases the muon energy threshold.

matched in simulation and experiment, the cosmic-ray energy  $E$  is difficult to match. This has a large impact. According to the Matthews-Heitler model of air showers [33], the muon number depends on the energy  $E$  and mass  $A$  of the cosmic ray in the following way

$$N_{\mu} = A \left( \frac{E}{AC} \right)^{\beta} = A^{1-\beta} \left( \frac{E}{C} \right)^{\beta}, \quad (1)$$

with power-law index  $\beta \simeq 0.9$  and energy constant  $C$ . The muon number scales almost linearly with the cosmic-ray energy, and with a small power of the mass. For an easier discussion, we take the logarithm on both sides of Eq. 1 and compute the mean, which gives a linear equation

$$\langle \ln N_{\mu} \rangle = (1 - \beta) \langle \ln A \rangle + \beta \langle \ln(E/C) \rangle. \quad (2)$$

Air shower experiments usually have independently calibrated energy scales with systematic uncertainties of 10 % to 20 %. Two otherwise identical experiments with an energy-scale offset of 20 % would find a 18 % offset in the data/MC ratios, simply because equivalent measurements are compared to air showers simulated at different (apparent) energies. Cross-calibrating the energy-scales removes these offsets.

When  $\langle \ln N_{\mu} \rangle$  is directly compared to simulations, the value of  $\langle \ln A \rangle$  matters. The situation is better when comparing data from different experiments. The value of  $\langle \ln A \rangle$  is a function of the energy  $E$  only, so the effect on two experiments is the same, if both experiments compute the muon density from an unbiased sample of air showers, meaning that cosmic rays are included in the sample with a probability independent of their mass  $A$  [34]. If the sample has a different mass composition  $\langle \ln A \rangle$ , an offset  $(1 - \beta)(\langle \ln A \rangle - \langle \ln A \rangle)$  is introduced in a comparison to an unbiased sample. A poor detector resolution or using wide energy bins can also introduce biases [35].

Based on this discussion, we can classify the experiments by the measured observables into three groups.

**Shower energy and muon density** The Pierre Auger Observatory & AMIGA, Telescope Array, and the Yakutsk

experiment are all capable of measuring the shower energy  $E_{\text{cal}}$  using Cherenkov or fluorescence light, which can be converted into the primary energy  $E$  with a low model-dependence. These experiments come close to measuring the primary cosmic-ray energy  $E$  independently of the muon density at the ground, and can compute the data/MC ratio for showers with the same energy.

The IceCube Neutrino Observatory also falls into this category, although it does not observe the air shower optically. It can measure the shower energy with low model-dependence thanks to its high altitude, which places the detector close to the average depth of the shower maximum [36].

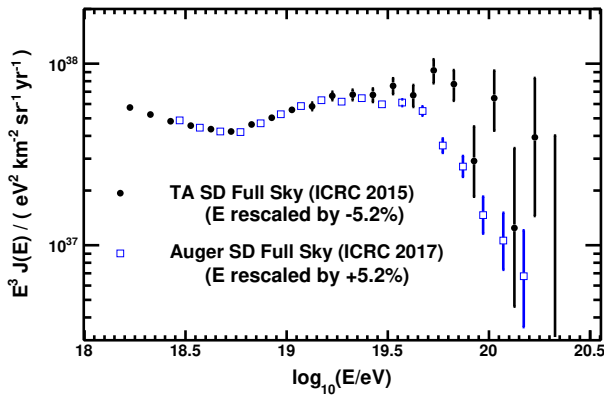
**Muon and electron density** KASCADE-Grande and EAS-MSU measure signals from electrons and muons separately. These experiments compute the data/MC ratio for showers in the same electron-density interval, not for showers in the same cosmic-ray energy interval. The electron density is correlated to the cosmic-ray energy, but also to the muon density [1]. Computing the data/MC ratio for showers in the same electron density interval is conceptually different from the previous case, since Eq. 1 and 2 do not apply. The data/MC ratios computed by the first and second class are not directly comparable.

**Muon density only** NEVOD-DECOR and SUGAR are pure muon detectors, without a separate energy estimator. The data/MC ratios are again computed differently. The flux of showers is measured in intervals of an event-wise muon density estimate. The measured flux is then compared with a simulated flux in the event-wise muon density estimate, computed from an external model of the cosmic-ray all-particle flux using a mass-composition assumption (proton or iron) and air shower simulations. If the fluxes differ, one can infer the data/MC ratio  $R$  based



**Table 1.** Energy-scale adjustment factors obtained from cross-calibration. The cross-calibration is explained in the text.

Experiment	$E_{\text{data}}/E_{\text{ref}}$
EAS-MSU	unknown
IceCube Neutrino Observatory	1.19
KASCADE-Grande	unknown
NEVOD-DECOR	1.08
Pierre Auger Observatory & AMIGA	0.948
SUGAR	0.948
Telescope Array	1.052
Yakutsk EAS Array	1.24



**Figure 3.** Relative cross-calibration of the energy scales of the Pierre Auger Observatory and Telescope Array by matching flux measurements, as presented by the *Spectrum Working Group*. Shown are the adjusted fluxes.

on this equation:

$$J_{\text{data}}(N_{\mu\text{data}}) = J_{\text{sim}}(N_{\mu\text{sim}}) \frac{dN_{\mu\text{sim}}}{dN_{\mu\text{data}}} = J_{\text{sim}} \left( \frac{N_{\mu\text{data}}}{R} \right) \frac{1}{R}. \quad (3)$$

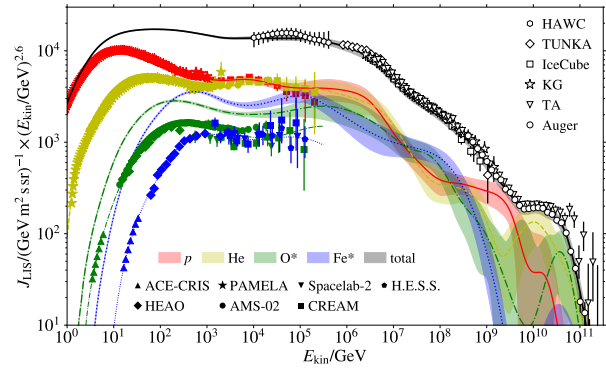
NEVOD-DECOR uses an average cosmic-ray flux computed from multiple experiments, while SUGAR uses the flux measured by the Pierre Auger Observatory. An energy is assigned to  $R$  based on the muon density and  $E \propto N_{\mu}^{1/\beta}$ .

#### 4.1 Energy scale offsets and cross-calibration

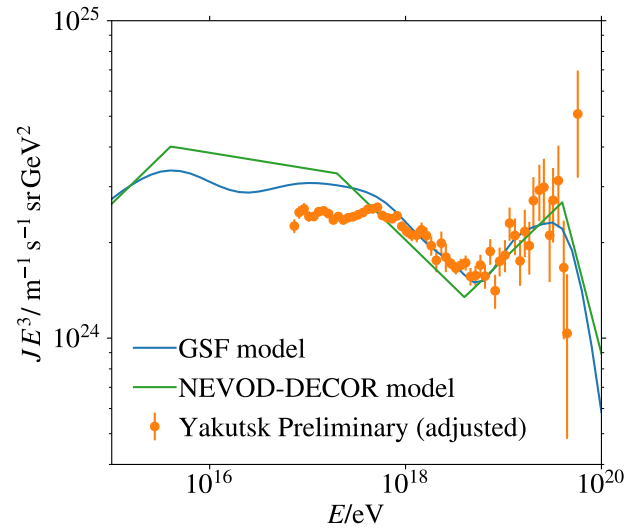
If an experiment uses the same energy proxy for measurements of the cosmic-ray flux and the muon density, relative offsets in data/MC ratios can be corrected that arise purely from the different energy scales.

The cross-calibration uses that the cosmic-ray flux is very isotropic up to  $10^{19.2}$  eV and can serve as a universal reference. If we assume that all deviations in measured fluxes between different experiments arise from energy-scale offsets, then a relative energy-scale ratio  $E_{\text{data}}/E$  can be found for each experiment so that the all-particle fluxes overlap, based on an equation analog to Eq. 3. This approach is well-known and has been used successfully in other works [37, 38].

The cross-calibration factors used in this report are given in Table 1. The *Spectrum Working Group* formed by the Pierre Auger and Telescope Array collaborations



**Figure 4.** Global spline fit (GSF) of cosmic-ray data (lines) and energy-scale adjusted data (points) [38]. The GSF treats energy-scale offsets as soft-constrained free parameters and produces energy-scale adjustment factors as a fit result.



**Figure 5.** All-particle flux from GSF [38] and NEVOD-DECOR [32] models, and a preliminary update of the Yakutsk spectrum (see Pravdin et al. for the last publication [39]), which was adjusted by an energy-scale factor 1.15 to match the GSF.

found a relative energy-scale shift of 10.4 % between the two experiments, see Fig. 3. The reference energy scale  $E_{\text{ref}}$  is placed between the two experiments.

The scaling factor for IceCube was taken from an updated Global Spline Fit (GSF) model [38], see Fig. 4. The GSF also uses cross-calibration internally, with a reference energy scale  $E_{\text{ref,GSF}}/E_{\text{ref}} = 0.948/0.88 \approx 1.08$ . The scaling factor for IceCube is  $E_{\text{IceCube}}/E_{\text{ref,GSF}} \times E_{\text{ref,GSF}}/E_{\text{ref}} = 1.10 \times 1.08 \approx 1.19$ .

The factor for Yakutsk is obtained by matching a preliminary update of their all-particle flux against the GSF model, see Fig. 5, to obtain a ratio  $E_{\text{Yakutsk}}/E_{\text{ref,GSF}} \times E_{\text{ref,GSF}}/E_{\text{ref}} = 1.15 \times 1.08 \approx 1.25$ . The factor for NEVOD-DECOR is obtained by comparing their custom all-particle flux parameterization with GSF (see same figure). The two models differ locally, which could be translated into energy-scale offsets within  $\pm 2\%$ , but no global offset is apparent. The energy scale of NEVOD-DECOR is there-

fore taken to be the same as GSF,  $E_{\text{NEVOD-DECOR}}/E_{\text{ref,GSF}} \times E_{\text{ref,GSF}}/E_{\text{ref}} = 1 \times 1.08 = 1.08$ .

No cross-calibration factor can be given for KASCADE-Grande, since the KASCADE-Grande flux is computed using a different energy estimator. For EAS-MSU, no all-particle flux is available for cross-calibration. SUGAR uses the flux from the Pierre Auger Observatory in its computation of the data/MC ratio and therefore has the same energy-scale adjustment factor.

We emphasize that the cross-calibration cannot eliminate a global offset of all experiments to the true energy scale, with corresponding shifts in the data/MC ratios. The energy scales of leading experiments have uncertainties in the order of 10 to 20%, we assume that the reference energy-scale has an uncertainty of at least 10%.

## 4.2 Combined measurements

Eq. 2 displays a simple relationship between the measured muon density,  $\langle \ln A \rangle$  and logarithmic shower energy. To compare all the measurements, we introduce the  $z$ -scale, which is inspired by Eq. 2,

$$z = \frac{\ln(N_{\mu}^{\text{det}}) - \ln(N_{\mu_p}^{\text{det}})}{\ln(N_{\mu_{\text{Fe}}}^{\text{det}}) - \ln(N_{\mu_p}^{\text{det}})}, \quad (4)$$

where  $N_{\mu}^{\text{det}}$  is the muon density estimate as seen in the detector, while  $N_{\mu_p}^{\text{det}}$  and  $N_{\mu_{\text{Fe}}}^{\text{det}}$  are the simulated muon density estimates for proton and iron showers after full detector simulation. The  $z$ -scale, while being rather abstract, has advantages over other choices that were proposed. The energy-dependence of  $N_{\mu}$  is removed and the expected range is from 0 (pure proton showers) and 1 (pure iron showers), if there is no discrepancy between real and simulated air showers. This is convenient. Furthermore, biases of the form  $\ln N_{\mu}^{\text{det}} = A + B \ln N_{\mu}$  in the measured muon density estimate  $N_{\mu}^{\text{det}}$  with respect to the true muon density  $N_{\mu}$  cancel in  $z$ .

Shown in Fig. 6 are the converted measurements. The  $z$ -values are computed relative to simulations and therefore a different result is obtained for each hadronic interaction model although the same data are used. The conversion to  $z$  is only possible when  $N_{\mu_p}^{\text{det}}$  and  $N_{\mu_{\text{Fe}}}^{\text{det}}$  are available for that model. Therefore not all data points can be shown for all models. Overall, the data suggest an energy-dependent trend, but with a large scatter.

The scatter is drastically reduced after the cross-calibration, as shown in Fig. 7. The cross-calibration causes a shift in the simulated values  $N_{\mu_p}$  and  $N_{\mu_{\text{Fe}}}$ , which were computed for the energy  $E_{\text{data}}$ , but are needed for  $E_{\text{ref}}$ . Based on Eq. 2, we get  $N_{\mu_{\text{ref}}} = N_{\mu_{\text{data}}} - \beta \ln(E_{\text{data}}/E_{\text{ref}})$ . The shift is the same for proton and iron showers. It cancels in the denominator of Eq. 4, but enters with the opposite sign in the numerator. We get

$$z_{\text{ref}} = z_{\text{data}} + \frac{\beta \ln(E_{\text{data}}/E_{\text{ref}})}{\ln(N_{\mu_{\text{Fe}}}^{\text{det}}) - \ln(N_{\mu_p}^{\text{det}})} \quad (5)$$

with  $\beta = 1 - (\ln N_{\mu_{\text{Fe}}} - \ln N_{\mu_p}) / \ln 56$ , based on Eq. 2. The values of  $N_{\mu_{\text{Fe}}}$  and  $N_{\mu_p}$  are taken for each model from

CORSIKA simulations. The points also move horizontally by the relative amount  $(E_{\text{data}}/E_{\text{ref}})^{-1}$ , a minor effect.

As expected, the cross-calibration improves the agreement of data from different experiments. Before and after the cross-calibration, the  $z$ -values rise above the iron line beyond  $10^{19}$  eV. The interpretation at lower energies changes, however. In case of IceCube, the originally negative  $z$ -values suggested that the muon density in proton showers simulated with EPOS-LHC for shower energies below  $10^{16}$  eV was too high. After the correction, the  $z$ -values fall between proton and iron. In case of Yakutsk, the original data suggested very low muon densities with partly negative  $z$ -values. After the correction, the Yakutsk data is consistent with others within uncertainties. We emphasize again that the reference energy-scale after cross-calibration has a remaining uncertainty of at least 10%. This means that  $z$ -values in all plots can be collectively varied by about  $\pm 0.25$ .

To further refine the conclusions, we consider the effect of an energy-dependent mass composition. With Eq. 2 and Eq. 4 the expected value  $z_{\text{mass}}$  for a given mean-logarithmic-mass  $\langle \ln A \rangle$  is computed as  $z_{\text{mass}} = \frac{\langle \ln A \rangle}{\ln 56}$ . As mentioned in the introduction, the experimental value of  $\langle \ln A \rangle$  is uncertain. Shown in Fig. 7 is a band, an envelope over optical measurements of the depth  $X_{\text{max}}$  of shower maximum from several experiments, and converted to  $\langle \ln A \rangle$  based on air shower simulations with EPOS-LHC. We will use this as a rough estimate of the mass composition. The band is independent of the muon measurements here, and therefore can be used as a reference. The  $z_{\text{mass}}$  value computed from the GSF model is also shown, which is based on optical and muon measurements and averages over experiments and model interpretations of air shower data. The line mostly falls inside the envelope.

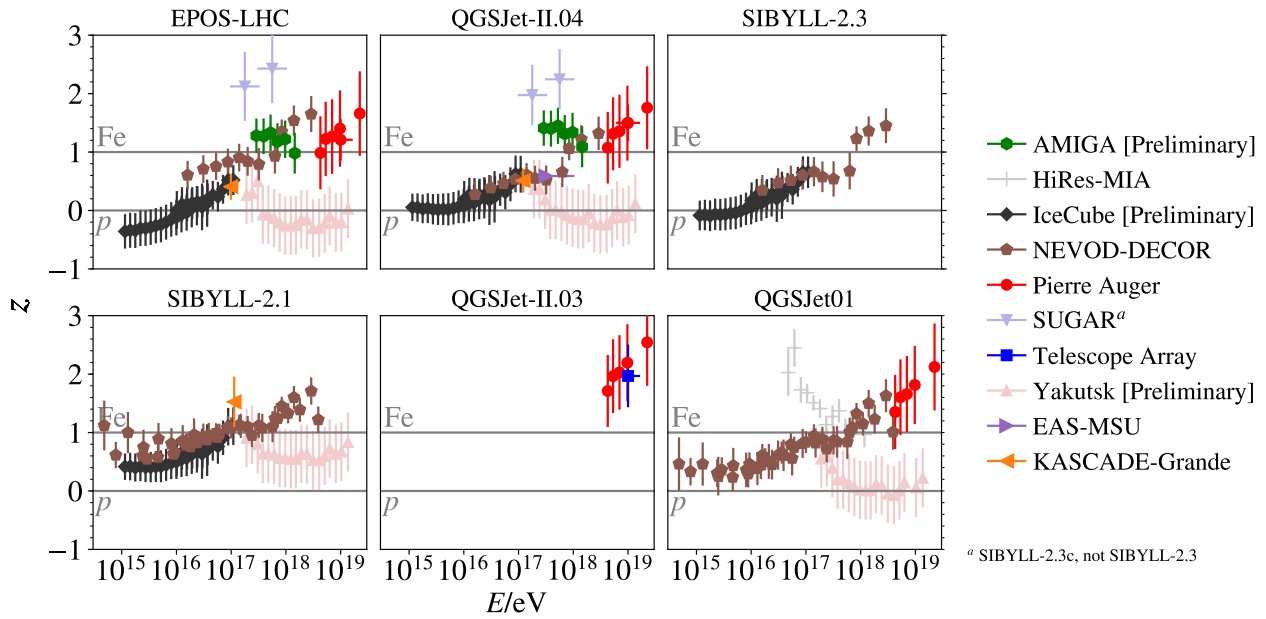
If the measured  $z$  values follow  $z_{\text{mass}}$ , the model describes the muon density at the ground consistently. This is overall not the case. The pre-LHC generation of hadronic interaction models, SIBYLL-2.1, QGSJet-II.03, and QGSJet01, show larger muon deficits than the models tuned to LHC data, EPOS-LHC, QGSJet-II.04, and SIBYLL-2.3. EPOS-LHC, QGSJet-II.04, SIBYLL-2.3, and QGSJet01 give a reasonable description of data up to a few  $10^{16}$  eV. At higher shower energies, a muon deficit in simulations is observed ( $z > z_{\text{mass}}$ ) in all models. Shown in Fig. 8 are zoomed plots for EPOS-LHC and QGSJet-II.04, the two latest-generation models with most data points. Shown in Fig. 9 is the difference  $\Delta z = z - z_{\text{mass}}$ . Subtracting  $z_{\text{mass}}$  is expected to remove the effect of the changing mass composition. An energy-dependent trend in  $\Delta z$  remains.

## 4.3 Energy-dependent trend

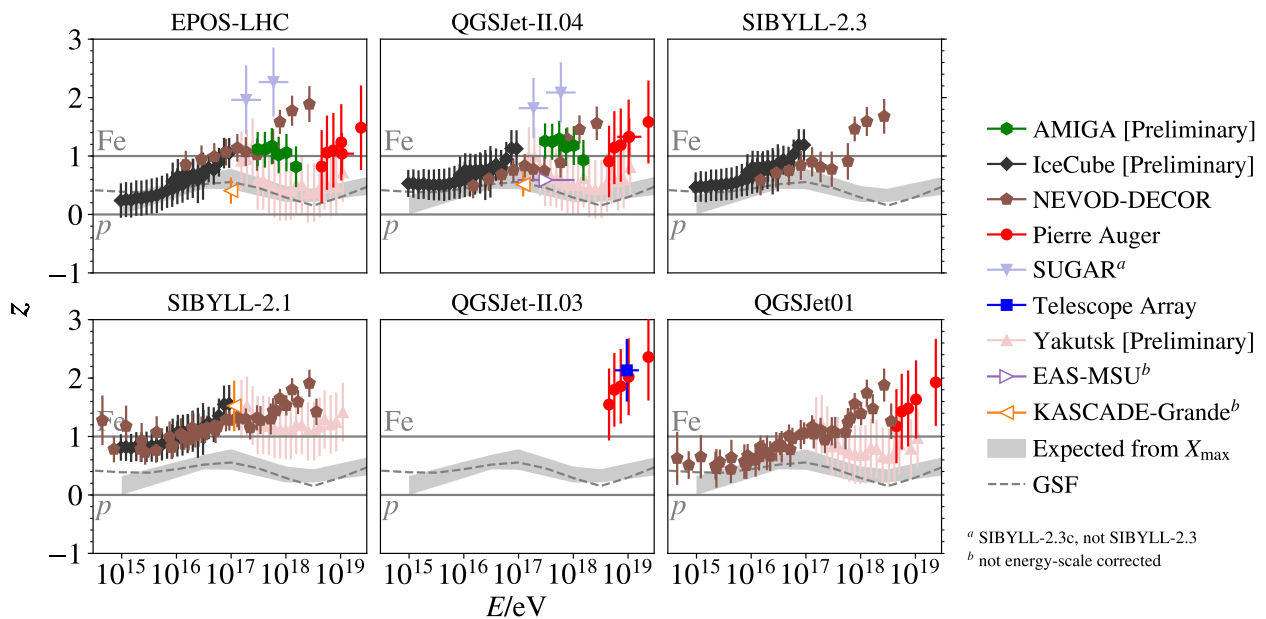
To quantify the observed trend in  $\Delta z$  as a function of energy, a line-model is fitted to the data shown in Fig. 9,

$$\Delta z = a + b (\log_{10}(E/\text{eV}) - 16), \quad (6)$$

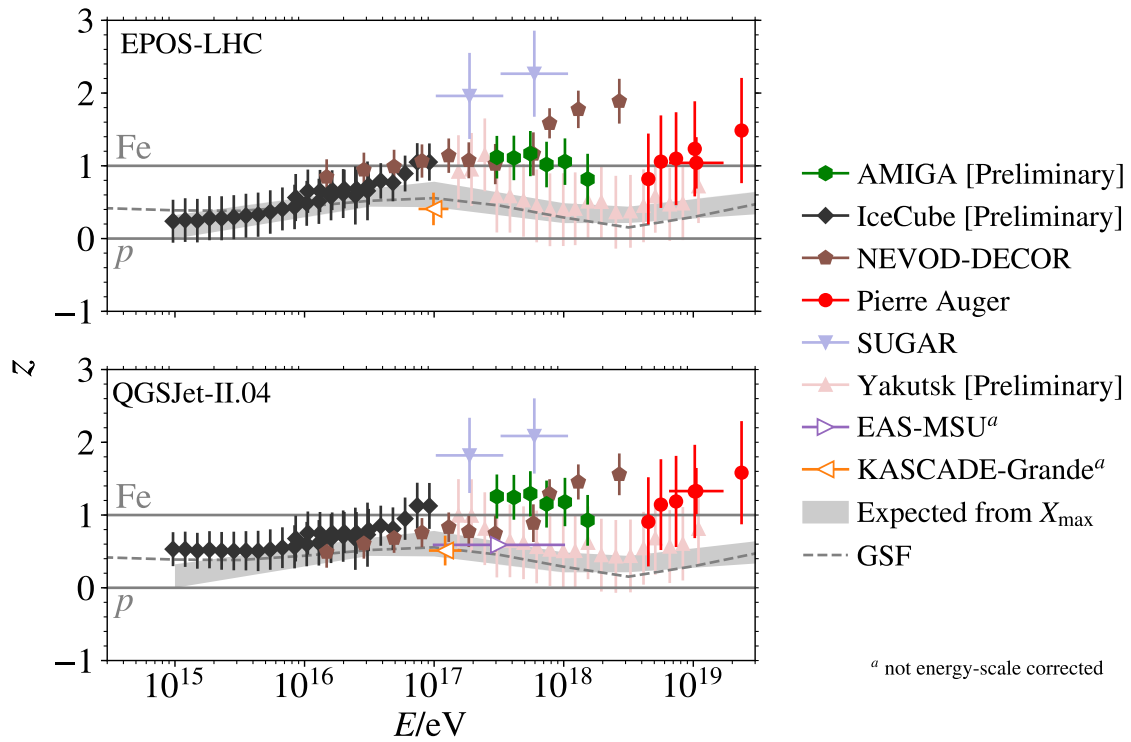
with free parameters  $a$  and  $b$ . The slope  $b$  is the increase in  $\Delta z$  per decade in energy. The  $z$ -values from KASCADE-Grande and EAS-MSU are not included in the fit, since



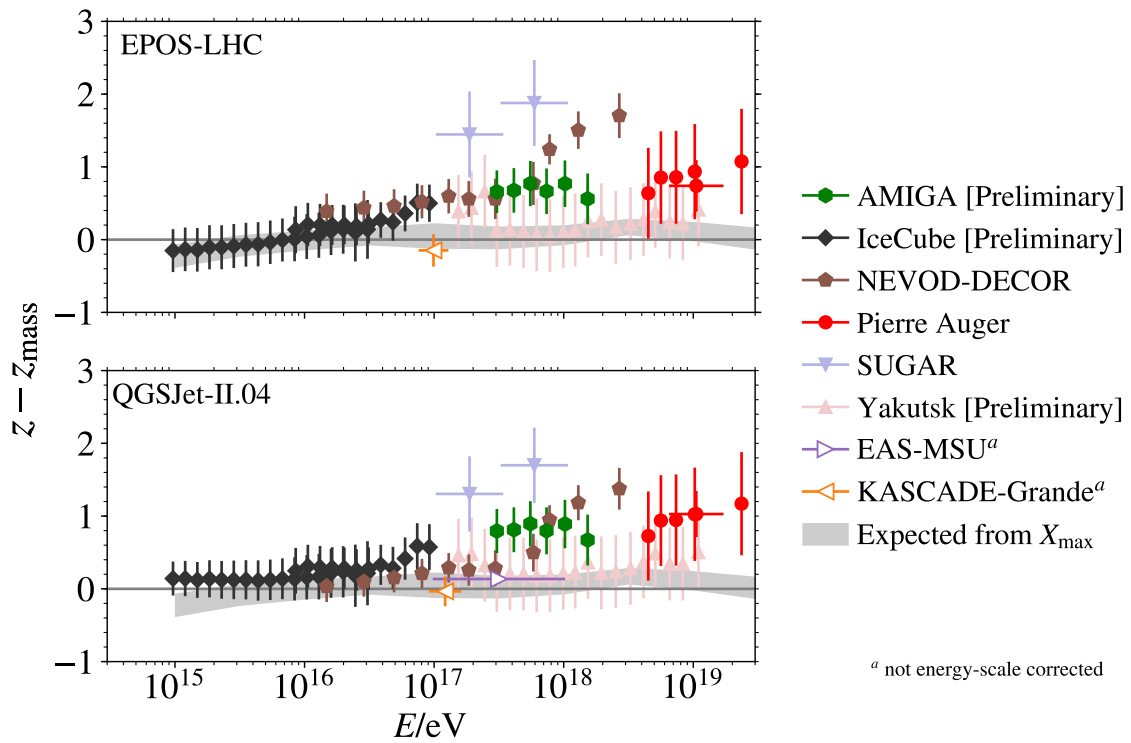
**Figure 6.** Muon density measurements converted to the  $z$ -scale, as described in the text. It depends on the hadronic interaction model, so the same data sets are repeatedly shown for different models. When corresponding simulations are missing for an experiment, no points can be shown. Error bars show statistical and systematic uncertainties added in quadrature (systematic uncertainties are dominant for nearly all measurements).



**Figure 7.** Data from Fig. 6 after applying energy-scale cross-calibration. The points for KASCADE-Grande and EAS-MSU cannot be cross-calibrated and are only included for comparison. Shown for comparison are  $z$ -values expected for a mixed composition from optical measurements (band), based on an update of the review by Kampert and Unger [1] by the original authors of that paper, and from the GSF model (dashed line).

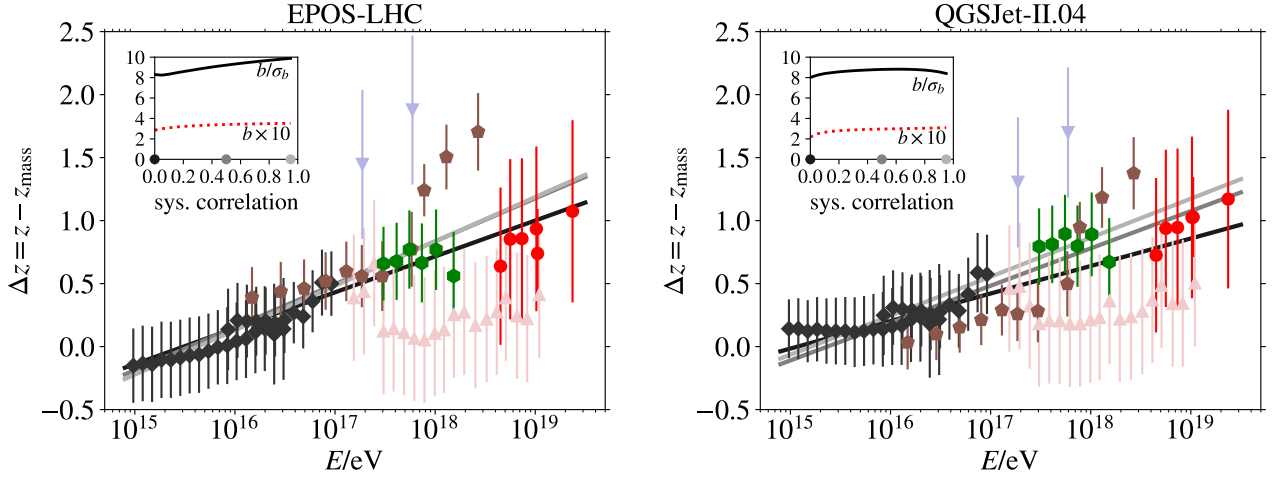


**Figure 8.** Zoom into Fig. 7 for EPOS-LHC and QGSJet-II.04. The points for KASCADE-Grande and EAS-MSU cannot be cross-calibrated and are only included for comparison.



**Figure 9.** Data from Fig. 8 after subtracting  $z_{\text{mass}}$ . The points for KASCADE-Grande and EAS-MSU cannot be cross-calibrated and are only included for comparison.





**Figure 10.** Fits of straight lines  $\Delta z = a + b(\log_{10}(E/\text{eV}) - 16)$  to the data points. Shown in the inset are the slope  $b$  and its deviation from zero in standard deviations for an assumed correlation of the point-wise uncertainties within each experiment. Examples of the fitted lines are shown for a correlation of 0, 0.5, and 0.95 in varying shades of gray.

they are not energy-scale corrected and not directly comparable to the other values.

The value of the slope  $b$  and its deviation from zero in standard deviations are of interest. The uncertainty of  $b$  scales with the uncertainties of the data. The error bars of most data points are dominated by systematic uncertainties, which are correlated for data points from a single set. Correlated uncertainties are the reason why the points in Fig. 9 do not scatter randomly as much as the error bars suggest. The exact amount of correlation is not known. We work around this problem by repeatedly fitting the data under different correlation assumptions.

We use the least-squares method for correlated uncertainties in the data. The score  $Q$  is minimized,

$$Q = (\Delta z - \Delta z^{\text{line}})^T C^{-1} (\Delta z - \Delta z^{\text{line}}), \quad (7)$$

where  $C^{-1}$  is the inverse of the covariance matrix  $C$  of the measurements. The matrix  $C$  is constructed as follows

$$C_{ij} = \sigma_{\Delta z,i} \sigma_{\Delta z,j} \times \begin{cases} 1 & \text{for } i = j \\ \alpha & \text{for } i, j \text{ from same data set} \\ 0 & \text{otherwise,} \end{cases} \quad (8)$$

where  $\alpha$  is the assumed correlation coefficient for the error bars within a data set. The matrix  $C$  has a block-diagonal form.

The fit is performed with MINUIT [40] and repeated for values of  $\alpha$  from 0 to 0.95. The HESSE algorithm is used to compute the standard deviation  $\sigma_b$  of  $b$ . To adjust for over- or underestimated uncertainties in the input, the raw result from MINUIT is corrected with the  $\chi^2$  value and the degrees of freedom  $n_{\text{dof}}$  of the fit,  $\sigma_b = \sigma_b^{\text{raw}} \sqrt{\chi^2/n_{\text{dof}}}$  [41].

The slope  $b$  and its deviation from zero in standard deviations as a function of the assumed correlation are shown in Fig. 10 for EPOS-LHC and QGSJet-II.04. The deviation from zero is always larger than 8 standard deviations, making the slope highly significant. The result is insensitive to the assumed correlation coefficient.

## 5 Summary and outlook

We presented a summary of recent tests and measurements of hadronic interaction properties in air showers with energies from PeV up to tens of EeV. Better agreement between simulation and experiment is found for the electromagnetic than for the muonic shower component.

We put a special focus on muon density measurements in this report. A comprehensive collection of muon measurements is presented. We developed the  $z$ -scale as a comparable measure of the muon density between different experiments and analyses. The  $z$ -scale uses air shower simulations as a reference to compare muon density measurements taken under different conditions. We demonstrated the importance of cross-calibrating energy-scales of experiments and apply it were possible, using the isotropic all-particle flux of cosmic rays as a reference.

After applying the cross-calibration, a remarkably consistent picture is obtained. Muon measurements seem to be consistent with simulations based on the latest generation of hadronic interaction models, EPOS-LHC, QGSJet-II.04, and SIBYLL-2.3, up to about  $10^{16}$  eV. At higher energies, a growing muon deficit in the simulations is observed, visible as an increase in  $z$  over the expectation. An analog trend is observed in older hadronic interaction models, with a more severe muon deficit. The slope of this increase in  $z$  per decade in energy is 0.22 to 0.35 for EPOS-LHC and QGSJet-II.04, with  $8\sigma$  significance.

We plan to further study the collected data, looking for other trends in the deviation between simulations and data. This will provide hints which aspects of hadronic interaction properties are the likely cause for these deviations.

## 6 Acknowledgements

We remember Mikhail Pravdin of the Yakutsk collaboration, who passed away on January 3rd, 2019. Dr. Pravdin had a leading role in cosmic ray physics. His kindness, skills, and expertise will be greatly missed.

## References

- [1] K.H. Kampert, M. Unger, *Astropart. Phys.* **35**, 660 (2012), 1201.0018
- [2] T. Pierog, I. Karpenko, J.M. Katzy, E. Yatsenko, K. Werner, *Phys. Rev.* **C92**, 034906 (2015), 1306.0121
- [3] S. Ostapchenko, *EPJ Web Conf.* **52**, 02001 (2013)
- [4] F. Riehn, H.P. Dembinski, R. Engel, A. Fedynitch, T.K. Gaisser, T. Stanev, *PoS ICRC2017*, 301 (2017), 1709.07227
- [5] T. Abu-Zayyad et al. (MIA, HiRes), *Phys. Rev. Lett.* **84**, 4276 (2000), astro-ph/9911144
- [6] A.G. Bogdanov, D.M. Gromushkin, R.P. Kokoulin, G. Mannocchi, A.A. Petrukhin, O. Saavedra, G. Trinchero, D.V. Chernov, V.V. Shutenko, I.I. Yashin, *Phys. Atom. Nucl.* **73**, 1852 (2010), [*Yad. Fiz.*73,1904(2010)]
- [7] W.D. Apel et al. (KASCADE-Grande), *Astropart. Phys.* **95**, 25 (2017), 1801.05513
- [8] Yu.A. Fomin, N.N. Kalmykov, I.S. Karpikov, G.V. Kulikov, M.Yu. Kuznetsov, G.I. Rubtsov, V.P. Sulakov, S.V. Troitsky, *Astropart. Phys.* **92**, 1 (2017), 1609.05764
- [9] J.A. Bellido, R.W. Clay, N.N. Kalmykov, I.S. Karpikov, G.I. Rubtsov, S.V. Troitsky, J. Ulrichs, *Phys. Rev.* **D98**, 023014 (2018), 1803.08662
- [10] A. Aab et al. (Pierre Auger), *Phys. Rev.* **D91**, 032003 (2015), 1408.1421
- [11] A. Aab et al. (Pierre Auger), *Phys. Rev. Lett.* **117**, 192001 (2016), 1610.08509
- [12] R.U. Abbasi et al. (Telescope Array), *Phys. Rev.* **D98**, 022002 (2018), 1804.03877
- [13] J. Gonzalez (IceCube), *Measuring the Muon Content of Air Showers with IceTop*, in *20th International Symposium on Very High Energy Cosmic Ray Interactions (ISVHECRI 2018) Nagoya, Japan* (2018)
- [14] S. Müller (Pierre Auger Observatory), *AMIGA preliminary measurement of lateral muon density (poster)* (2018), UHECR 2018, Paris, France
- [15] A.V. Glushkov, M.I. Pravdin, and A. Sabourov for the Yakutsk collaboration, private comm. (2018)
- [16] R. Ulrich (Pierre Auger), *PoS ICRC2015*, 401 (2016)
- [17] R.U. Abbasi et al. (Telescope Array), *Phys. Rev.* **D92**, 032007 (2015), 1505.01860
- [18] J. Bellido (Pierre Auger), *PoS ICRC2017*, 506 (2018), [40(2017)]
- [19] A. Aab et al. (Pierre Auger), Submitted to: *JCAP* (2018), 1811.04660
- [20] S. Andringa et al. (Pierre Auger), *Average shape of longitudinal shower profiles measured at the Pierre Auger Observatory*, these proceedings
- [21] S. De Ridder, E. Dvorak, T.K. Gaisser (IceCube), *PoS ICRC2017*, 319 (2018)
- [22] D. Ivanov (Telescope Array), *Telescope Array Experiment* (2018), TeVPA 2018, Berlin, Germany, <https://indico.desy.de/indico/event/18204/session/15/contribution/117>
- [23] A. Aab et al. (Pierre Auger), *Phys. Rev.* **D90**, 012012 (2014), 1407.5919
- [24] W.D. Apel et al., *Astropart. Phys.* **34**, 476 (2011)
- [25] S. Ostapchenko, *Phys. Rev.* **D74**, 014026 (2006), hep-ph/0505259
- [26] J. Adam et al. (ALICE), *JCAP* **1601**, 032 (2016), 1507.07577
- [27] M.G. Aartsen et al. (IceCube), *Astropart. Phys.* **78**, 1 (2016), 1506.07981
- [28] T. Fuchs (IceCube), *Development of a Machine Learning Based Analysis Chain for the Measurement of Atmospheric Muon Spectra with IceCube*, in *25th European Cosmic Ray Symposium (ECRS 2016) Turin, Italy, September 04-09, 2016* (2017), 1701.04067
- [29] D. Soldin (IceCube), *Atmospheric Muons Measured with IceCube*, in *20th International Symposium on Very High Energy Cosmic Ray Interactions (ISVHECRI 2018) Nagoya, Japan* (2018), 1811.03651
- [30] A. Aab et al. (Pierre Auger), *Phys. Rev.* **D96**, 122003 (2017), 1710.07249
- [31] A. Aab et al. (Pierre Auger), *Phys. Rev.* **D93**, 072006 (2016), 1604.00978
- [32] A.G. Bogdanov, R.P. Kokoulin, G. Mannocchi, A.A. Petrukhin, O. Saavedra, V.V. Shutenko, G. Trinchero, I.I. Yashin, *Astropart. Phys.* **98**, 13 (2018)
- [33] J. Matthews, *Astropart. Phys.* **22**, 387 (2005)
- [34] H.P. Dembinski, B. Kégl, I.C. Mariş, M. Roth, D. Veberič, *Astropart. Phys.* **73**, 44 (2016), 1503.09027
- [35] H.P. Dembinski, *Astropart. Phys.* **102**, 89 (2018), 1711.05737
- [36] M.G. Aartsen et al. (IceCube), *Phys. Rev.* **D88**, 042004 (2013), 1307.3795
- [37] J.R. Hoerandel, *Astropart. Phys.* **19**, 193 (2003), astro-ph/0210453
- [38] H.P. Dembinski, R. Engel, A. Fedynitch, T. Gaisser, F. Riehn, T. Stanev, *PoS ICRC2017*, 533 (2018), 1711.11432
- [39] M.I. Pravdin et al., *Bull. Russ. Acad. Sci. Phys.* **73**, 541 (2009), [*Izv. Ross. Akad. Nauk Ser. Fiz.*73,no.5,578(2009)]
- [40] F. James, M. Roos, *Comput. Phys. Commun.* **10**, 343 (1975)
- [41] F. James, *Statistical methods in experimental physics* (2006), ISBN 9789812567956