



Contents lists available at ScienceDirect

Nuclear Inst. and Methods in Physics Research, A

journal homepage: www.elsevier.com/locate/nima

The HiSPARC experiment

K. van Dam^{a,*}, B. van Eijk^{a,b,*}, D.B.R.A. Fokkema^c, J.W. van Holten^{a,d}, A.P.L.S. de Laat^{a,1}, N.G. Schultheiss^{a,e}, J.J.M. Steijger^a, J.C. Verkooijen^a^a Nikhef National Institute for Subatomic Physics, Science Park 105, 1098 XG Amsterdam, The Netherlands^b Faculty of Science and Technology, University of Twente, 7500 AE Enschede, The Netherlands^c Faculty of Sciences, VU University Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands^d Lorentz Institute, Leiden University, Niels Bohrweg 2A, 2333 CA Leiden, The Netherlands^e Zaanlands Lyceum, Vincent van Goghweg 42, 1506 JD Zaandam, The Netherlands

ARTICLE INFO

Keywords:

HiSPARC
Cosmic rays
Extensive air shower detector
Scintillation detector
High school
Outreach

ABSTRACT

The High School Project on Astrophysics Research with Cosmics (HiSPARC) is a large extensive air shower (EAS) array with detection stations throughout the Netherlands, United Kingdom, Denmark and Namibia. HiSPARC is a collaboration of universities, scientific institutes and high schools. The majority of detection stations is hosted by high schools. A HiSPARC station consists of two or four scintillators placed inside roof boxes on top of a building. The measured response of a detector to single incoming muons agrees well with GEANT4 simulations. The response of a station to EASs agrees with simulations as well. A four-scintillator station was integrated in the KASCADE experiment and was used to determine the accuracy of the shower direction reconstruction. Using simulations, the trigger efficiency of a station to detect a shower as function of both distance to the shower core and zenith angle was determined. The HiSPARC experiment is taking data since 2003. The number of stations (~140 in 2019) still increases. The project demonstrates that its approach is viable for educational purposes and that scientific data can be obtained in a collaboration with high school students and teachers.

1. Introduction

Cosmic rays are energetic particles from space that hit the Earth's atmosphere at a rate of about 1000 per square meter per second [1]. They mainly consist of protons (~90%) and α -particles (~9%). A very small fraction contains heavier, ionized nuclei [1]. Cosmic rays with energies up to about 10^{10} eV are predominantly produced by the Sun (solar wind [2] and solar energetic particles [3]). Cosmic rays with energies between 10^{10} eV and 10^{18} eV are considered to be of galactic origin [4]. These galactic cosmic rays are believed to predominantly originate from supernova remnants. Particles with energies beyond 10^{18} eV up to the extreme energy of 10^{20} eV stem from extra-galactic sources [5]. However, little is known about these sources and acceleration mechanisms. The flux of galactic cosmic rays decreases rapidly with energy ($\sim E^{-2.7}$ above 10^{10} eV and drops to $\sim E^{-3.1}$ beyond 3×10^{15} eV). Thus, solar cosmic rays are many orders of magnitude more abundant than galactic cosmic rays. The flux of these galactic cosmic rays is in turn many times higher than that of extra-galactic cosmic rays. The cosmic ray rate above 10^{15} eV quickly drops to single events per square meter per year. At 10^{18} eV this rate drops to single events per

square kilometer per year. This implies that space-based experiments focusing on these energy ranges suffer from low statistics.

When an energetic cosmic ray hits the Earth's atmosphere it will most likely collide with a nitrogen or oxygen nucleus. A large number of (energetic) secondary particles may be produced. These secondary particles will interact with other atmospheric nuclei. The multiplication process continues until the energy becomes insufficient for further particle production. The result of this mechanism is called an 'Extensive Air Shower' (EAS) [6].

The size of the footprint of an EAS at the surface of the Earth rises with the energy of the primary cosmic ray. Nevertheless, statistical fluctuations lead to differences in the total number of particles reaching ground level up to factors of ten. An EAS consists mainly of gamma rays, electrons (positrons) and, to a lesser extent, muons and hadrons. Below $\sim 10^{13}$ eV, the shower leaves no footprint and only some remnants reach the ground. For cosmic ray energies below $\sim 10^{11}$ eV only one or two muons reach the Earth's surface and no electrons, positrons or gamma rays are left. Because of the high flux of these low energy cosmic rays there is a large number of isolated minimum ionizing muons. We refer to these muons as the single muon background as

* Corresponding authors.

E-mail addresses: kaspervd@nikhef.nl (K. van Dam), vaneijk@nikhef.nl (B. van Eijk).¹ Now in industry.

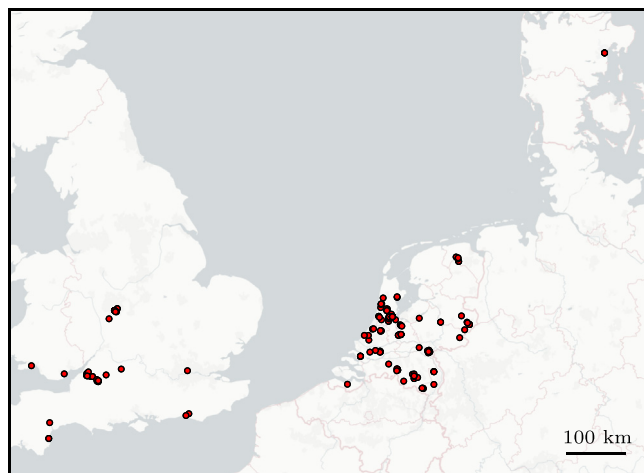


Fig. 1. Layout (early 2019) of the HiSPARC array. Each red dot represents one or more stations.

they do not stem from EAS. The footprint of an EAS ranges from meters to several kilometers in diameter for perpendicular incident primaries. Since it is difficult to cover large footprints with a single detector, EASs are usually sampled by relatively small detectors arranged in arrays. By reconstructing the EAS, properties of the original cosmic ray can be derived. Examples of sampling experiments are KASCADE (13 m detector separation, $200 \times 200 \text{ m}^2$) [7], KASCADE-Grande (137 m detector separation, $700 \times 700 \text{ m}^2$) [8], AGASA (1 km detector separation, 100 km^2) [9], Telescope array (1.2 km detector separation, 762 km^2) [10] and the Pierre Auger Observatory (1.5 km detector separation, 3000 km^2) [11].

The High School Project on Astrophysics Research with Cosmics (HiSPARC) [12,13] has approximately 140 detection stations distributed throughout the Netherlands, United Kingdom and Denmark (Fig. 1), and Namibia. HiSPARC is a collaboration of universities, scientific institutes and high schools each hosting their own detection station(s). The majority of the stations is located at high schools. HiSPARC has a strong outreach component. Stations are maintained by high school teachers, their students and university staff. Data are stored at Nikhef, the Dutch National Institute for Subatomic Physics [14]. A number of HiSPARC stations also employs a weather station. Their data are stored in the central database at Nikhef as well. Weather stations are used to analyze atmospheric conditions affecting EAS development.

Irregular arrays of cosmic ray detectors are not new, e.g. LAAS [15], SEASA [16], CHICOS [17], and recently EEE [18]. The latter three projects have a strong education component involving high schools as well. In 2001, the Nijmegen Area High School Array (NAHSA [19]) was founded. In 2003 Nikhef initiated the expansion of NAHSA into a nationwide network: HiSPARC. The geometry of the array is determined by the location of the collaborating institutions. This has led to an irregular grid of detection stations.

HiSPARC detection stations are relatively robust, cheap (5000 €/10,000 €, configuration dependent), small and straightforward to assemble. Single (remote) stations are mainly used for educational purposes and to provide local measurements of the single muon flux, EAS flux and EAS directions. Clusters of stations can be used for more advanced EAS reconstruction. Typically, high schools have limited resources for both building, and maintaining a station. In this paper the HiSPARC experiment, its hardware infrastructure, data acquisition, analysis tools and performance are described.

2. Scintillation detector

The detection philosophy of HiSPARC is to sample EAS footprints using scintillation detectors. The light output of a scintillator

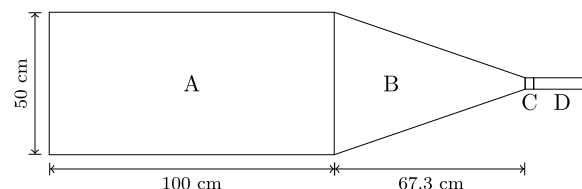


Fig. 2. Sketch of the HiSPARC detector. The scintillator and light-guide are denoted by the letters A and B resp. The light-guide adaptor piece (C) enables the cylindrical PMT (D) to be mounted to the square end of the light-guide.



Fig. 3. The HiSPARC detector inside a roof box.

is proportional to the number of charged EAS particles traversing the detector.

A scintillator ($100 \text{ cm} \times 50 \text{ cm} \times 2 \text{ cm}$), see A in Fig. 2, is glued to a slightly thicker triangular light-guide (base 50 cm, top 2.5 cm, height 67.5 cm, B) which is connected to a photomultiplier tube (PMT, D). A small adaptor light-guide (C) connects to the cylindrical PMT. Both light-guide (polymethylmethacrylate or PMMA) and scintillator have comparable refractive indices (1.49 and 1.58 resp.). The scintillation material [20] has a light attenuation length of 380 cm. The wavelength of maximum emission is 425 nm. The surfaces of the scintillator and light-guides are diamond polished to achieve a high surface reflectivity. The detector is wrapped in aluminum foil (thickness $30 \mu\text{m}$) and is made light-tight with black pond liner (thickness 0.45 mm). The assembly is placed inside a roof box (Fig. 3).

2.1. Expected energy loss

When a charged particle traverses the scintillator it loses energy. Apart from radiative energy loss, the particle primarily loses energy due to inelastic collisions with the atomic electrons of the scintillation material. Multiple molecules will be excited. These molecules will quickly return to their ground state and emit photons. The wavelength of the scintillation photons matches the wavelength characteristics of the PMT [21,22]. The photons may scatter several times inside the detector and only a fraction will reach the PMT. Fig. 4 shows a typical PMT output signal. The analog signal is sampled by an ADC in 2.5 ns bins. The area under the curve, or pulse integral, is a measure for the number of scintillation photons that have reached the PMT. The pulse integral is then calculated by summing the ADC values for each bin exceeding -10 mV (dotted line in Fig. 4).

The majority of muons and electrons (positrons) in an EAS carry high energy and can be considered minimum ionizing (MIP). For electron and muon MIPs the energy loss distribution is similar and is described by Landau theory [23]. The Landau distribution has a pronounced tail towards higher energy losses. A vertically incident

minimum ionizing particle has a most probable energy loss in 2 cm scintillation material of 3.51 MeV (\equiv 1 MIP).

Only a small fraction of the gamma rays in an EAS interacts with the scintillator via Compton scattering and less frequently, depending on their energy, via pair creation. These Compton electrons (positrons) are again detected via scintillation. Both direction and energy of the Compton electron depend on the scattering angle. The interaction depth differs for each gamma. Therefore, Compton electrons have different energies and travel different distances. The energy loss distribution due to gamma rays is a continuously decreasing function with energy. Pair creation is suppressed as high energy gamma rays are not abundant.

2.2. PMT

HiSPARC deploys PMTs with a cathode diameter of 25 mm. The 12 cm long glass tube is enclosed by mu-metal, providing shielding against external magnetic fields. The PMT-base is supplied with a DC voltage of -12 V which is converted into a DC voltage ranging from -300 to -1500 V. The quantum efficiency at 425 nm is typically 25%. Two different types of PMT bases are used. A commercial one [24] and an in-house developed version [25]. The Nikhef base provides a highly linear response over a large dynamical range, allowing to generate signals well in excess of -5 V (Fig. 5, red crosses). The rise-time is almost independent of the output signal amplitude. The impedance of the line driver, cable and subsequent readout electronics have been matched. The dynamical range of the commercial base is limited (Fig. 5, blue dots). The response curves were measured with a device containing 24 LEDs (light-emitting diodes). The LEDs were connected to the PMT using optical fibers. The pulsed light-output of each single LED was measured with the same PMT. Higher light intensities were obtained by bundling optical fibers. Fig. 5 shows that for small intensities both PMT assemblies behave linearly but at higher intensities the output flattens for the commercial base. The response of the PMT assemblies can be parametrized with a single function:

$$f(x) = \frac{ax}{(x^b + c)^{\frac{1}{b}}} + dx \quad (1)$$

with $a = 0.237$, $b = 13.5$, $c = 9.34 \cdot 10^4$, $d = 0.918$ for the Nikhef base and $a = 1.42$, $b = 2.74$, $c = 4.13$, $d = 0.150$ for the commercial base. A lab test on each PMT is not required. The PMT response function is derived from experimental data using the pulse integral (Fig. 4). Single minimum ionizing particles generate a peak in the pulse height and pulse integral distributions. The MIP-peak value in the pulse height domain (MIP_{ph}) is divided by the MIP-peak value in the pulse integral domain (MIP_{pi}). The pulse integrals (pi) are subsequently multiplied by this ratio to obtain inferred pulse heights:

$$ph_{inferred} = \frac{MIP_{ph}}{MIP_{pi}} \cdot pi \quad (2)$$

The measured pulse heights are expressed as function of the inferred pulse heights and the parametrization in Eq. (1) is fitted to obtain the PMT response function.

For perpendicular incident 10^{15} eV proton showers, electron densities (simulation) larger than $50 m^{-2}$ occur within a ~ 8 m radius. The high voltage on a PMT is chosen such that the MIP response distribution peaks at ~ 150 mV. The low and high thresholds are set to -30 mV (0.2 MIP) and -70 mV (~ 0.5 MIP) resp. These settings have been chosen to increase the sensitivity for gamma rays and low energy electrons while maintaining a decent dynamic range (ADC, ~ 2.3 V or ~ 15 MIPs). For larger pulses the read-out electronics are equipped with adjustable comparators in order to still be able to get an estimate of the pulse shape. For details see Section 3.1.

As atmospheric conditions change, the temperature of the PMT assembly can vary between -30°C on cold winter nights and $+60^\circ\text{C}$ when the Sun heats the air in the roof box. Temperature differences affect the height of the signal pulse. For the commercial base a higher

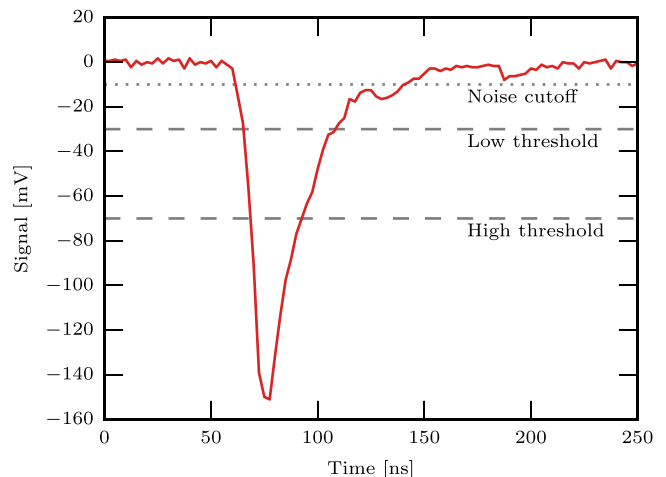


Fig. 4. Example of a typical signal with a pulse height of 150 mV. The FWHM is ~ 25 ns. The pulse integral is calculated by summing all values in the bins where the signal exceeds -10 mV (dotted gray line). The default trigger thresholds are -30 and -70 mV (dashed gray lines).

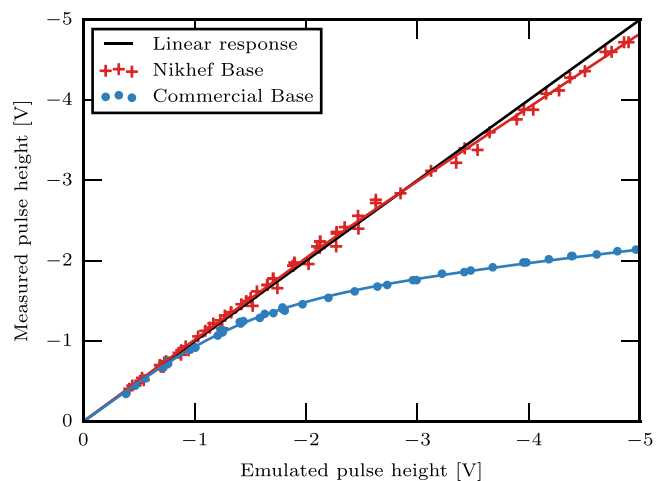


Fig. 5. Response of two HiSPARC PMT assemblies with different bases. On the horizontal axis the emulated pulse height (the combined pulse intensity of multiple LEDs) is shown. The vertical axis shows the measured pulse height (mV). The blue dots indicate the response for the commercial base whereas the red crosses show the pulse heights for the Nikhef base. Both response curves can be well described by the expression in Eq. (1). The black line indicates an ideal linear response. Up to about -0.7 V (~ 5 MIPs) both assemblies give a linear response.

temperature results in a lower gain. The Nikhef base shows a higher gain at higher temperatures. To arrive at a proper measure for the pulse height, temperature variations need to be taken into account.

The value of the pulse height corresponding to 1 MIP is derived from EAS data as a function of temperature. The height of the MIP-peak is averaged over short periods in time (4 h) and is used to calibrate the PMT output signal. Fig. 6 gives an example of the correlation between temperature (measured inside the roof box with a scale uncertainty of 0.5°C) and MIP-peak value for a commercial assembly. Of order 1000 events are collected within 4 h. A smaller time window results in a less accurate determination of the MIP-peak value. Alternatively, a method can be applied using a running average.

2.3. Transmission of scintillation light

When a (shower) particle propagates through the scintillator, the energy that is lost is converted into excitation photons. The yield

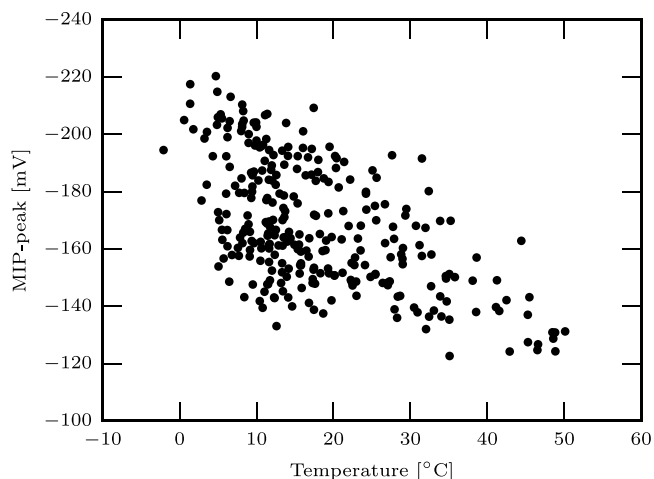


Fig. 6. Correlation between temperature and MIP-peak value for a PMT with a commercial base. The temperature was measured inside the roof box with a scale uncertainty of 0.5°C. The temperature and MIP-peak value were determined by averaging over 4-hour periods. There is a clear trend between the MIP-peak value and the local temperature.

of scintillation photons has been analyzed using GEANT4 [26] by exposing the detector (scintillator and light-guides) to perpendicular incident ionizing muons. As the scintillator has to be light tight, the detector is wrapped in black pond liner. To regain reflectivity, the scintillator is first packed in aluminum foil. The simulation assumes that the aluminum foil tightly encloses the detector. In reality, air pockets between scintillator and aluminum foil result in internal reflections. These reflections generate a higher yield of photons at the PMT than for a scintillator with a perfectly fitting aluminum envelope. In the simulation program this lack of air pockets is corrected for by increasing the aluminum reflectivity from 0.88 to a value of 0.93 to match the experimental data. This only scales the number of photons.

The photons have a different probability to reach the PMT depending on the location at which they are released in the scintillator. The plot on the right in Fig. 7 shows the distribution of the number of scintillation photons arriving at the PMT as a function of position. The muon flux was kept constant over the full area of the detector. The left hand figures show the distribution of the number of photons that reach the PMT from locations A (top) and B (bottom) in the scintillator. The distribution of photons follows Landau theory (black curves). The discrepancy between simulation and Landau curves is believed to be due to local differences in attenuation and reflection.

Fig. 7 also shows that the maximum light output is obtained when the muon hits a corner of the scintillator near the light-guide (top). Surprisingly, within the area that is represented by the mirror image of the light guide in the scintillator, the photon yield is relatively small.

To verify the simulation a table-top experiment was designed [13]. A 1.5 cm × 1.5 cm scintillator is connected to a small PMT. This small scintillator is positioned on top of the HiSPARC detector. A 3 × 5 grid (Fig. 8, 1–15) defines the locations. In the vicinity of for instance point 10, the simulation predicts a large light yield gradient. In this region four additional measurements (16–19) were carried out. The readout is triggered when both scintillators signal a MIP. In Fig. 9 the pulse integral distributions obtained at positions 16 and 17 are compared. Although the locations are very near, there is a sizeable difference in the photon yield reflected in the shift of the peak value. Position 17 has the peak at ~4400 mVns whereas the yield at position 16 (~3600 mVns) is by ~20% smaller. Both distributions can accurately be described by a Landau distribution convoluted with a Gaussian. PMT response characteristics are not included in this simulation. Moreover, simulations only account for perpendicular incident muons whereas the experiment is susceptible to muons from all directions where both scintillators generate a sufficiently large signal. This results in Gaussian smearing.

Fig. 10 shows the comparison between the light-output of the table-top experiment and detector simulation (scintillator and light-guide). Both experimentally measured and simulated MIP-peak values have comparable statistical uncertainties. The simulation reproduces the experimental data rather well. A more detailed analysis including photon arrival times is presented in a separate paper [27].

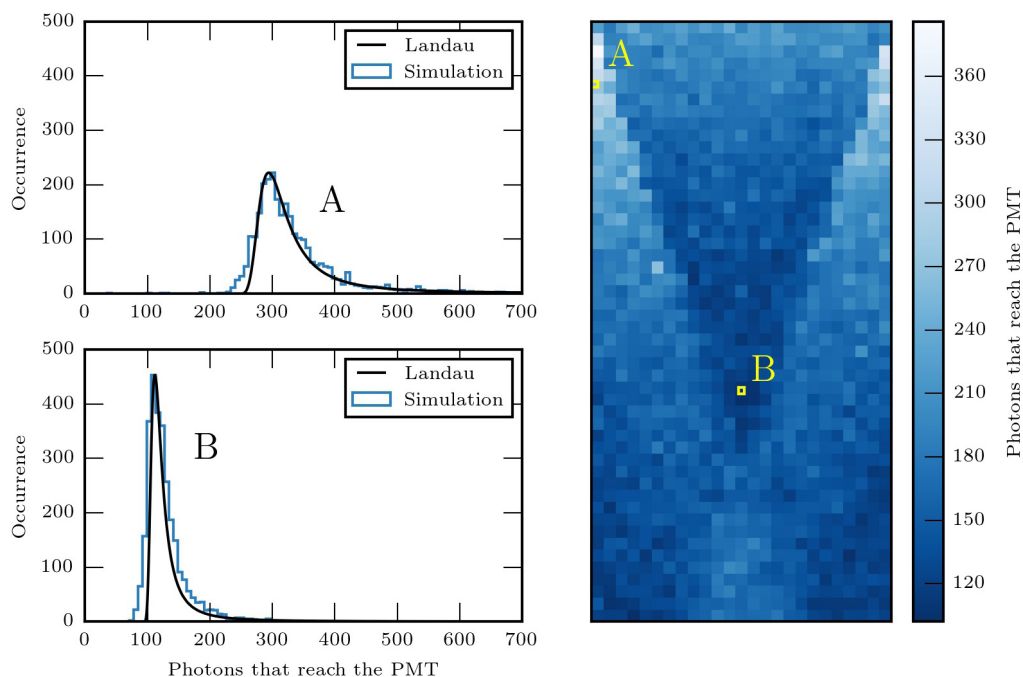


Fig. 7. At the right the average number of muon-induced scintillation photons arriving at the PMT as a function of the position in the detector is shown. In the simulation the detector was exposed to a large number of perpendicular incident relativistic muons. The muon flux was kept constant over the full area of the detector. The light-guide is connected at the top. The two histograms on the left show the fluctuation in the number of photons (in blue) released at locations A and B resp. Both distributions follow the Landau description (in black). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

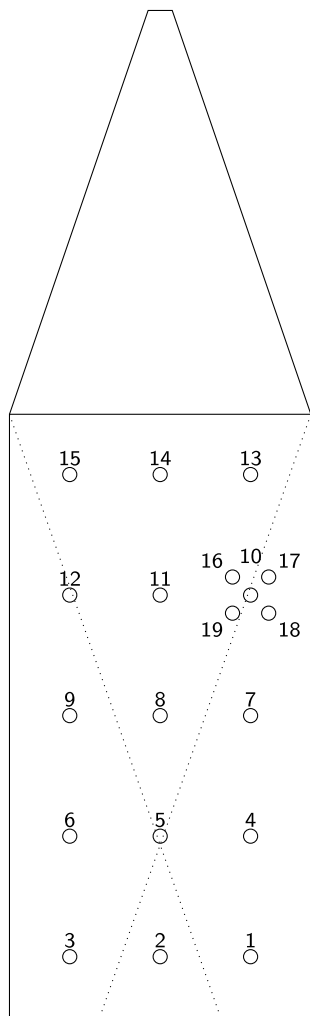


Fig. 8. The scintillator (including light-guide) light transmission is measured at 19 locations in the detector. The dotted lines show the position at which the largest gradients are observed (Fig. 7). The circles indicate the positions at which the efficiency is measured. Fifteen positions are defined on a grid, with four additional measurements performed around point 10.

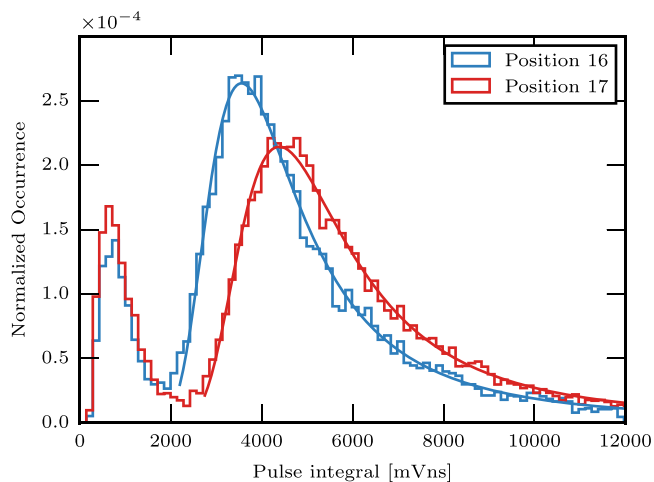


Fig. 9. Pulse integral distributions for events collected at position 16 and 17. Despite the small distance between the two positions there is a large difference in photon yield. Both distributions can accurately be described by a Landau convoluted with a Gaussian. The lower end of the spectrum results from gamma rays, low energy particles, and PMT noise.

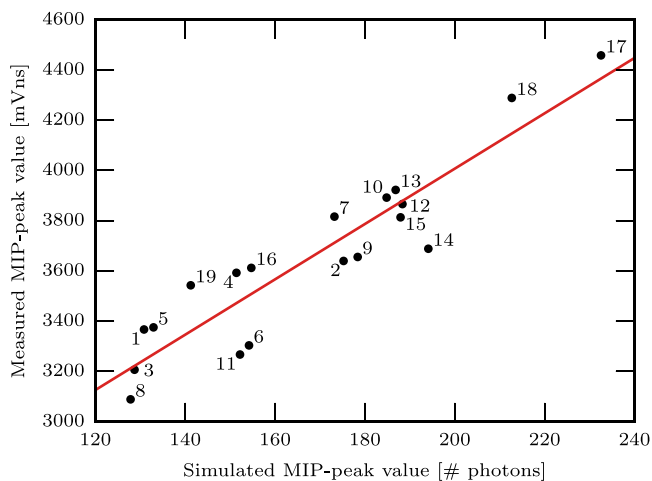


Fig. 10. Comparison between measured and simulated light yield of a HiSPARC detector (scintillator and light-guide). The numbers in the plot correspond to the locations indicated in Fig. 8.

2.4. Light-guide

The light-guide reduces the scintillator light yield as demonstrated in Fig. 7. However, the light-guide also may add Cherenkov photons when a charged particle penetrates. Using GEANT4 the production and propagation of these Cherenkov photons have been investigated. The light-guide was exposed to perpendicular incident relativistic muons. Again, the muon flux was kept uniform over the full surface of the light-guide. Fig. 11 shows the spatial distribution at which the Cherenkov photons are released scaled with the number of photons that reach the PMT. The further away from the PMT the Cherenkov photons are generated, the smaller the probability they reach the PMT. Even for large Cherenkov photon yields close to the PMT, the number of photons is small compared to the number of photons generated in the scintillator.

Figs. 9 and 10 show the measured signal yield and transmission at selected grid points in the scintillator. To obtain the muon response over the full surface of the detector (scintillator and light-guide) taking into account the proper energy spectrum and angle of incidence of the muon, a second experiment was conducted. Two stacks of each two detectors are placed parallel at a distance of 6 m. When three or more detectors generate a signal exceeding the noise cut-off at -15 mV (Fig. 4), the event is recorded. In the analysis, events containing signals with a time difference between both stacks of less than 300 ns are discarded. This excludes the contribution from EASs in which particles arrive in a relatively small time window. If the time difference is larger than 300 ns, the event most likely stems from a random coincidence. These random coincidences are due to single muon background and “noise pulses” (PMT dark pulses, low energy electrons, gamma rays, etc.). When two detectors in the same stack generate a signal exceeding -15 mV, a MIP traversed both detectors. If only one detector of a stack generates a pulse, it is most likely caused by ‘background noise’. This noise generates a pulse-spectrum that peaks at (~ -50 mV [27]). The high rate of noise pulses in detectors of the same stack causes only a small number of random background coincidences. By choosing those events in which two detectors in one stack are triggered by a MIP (time difference with a signal in the other stack is larger than 300 ns), single muons are selected. The result is shown in Fig. 12 (red histogram). All four detectors generate a similar pulse height spectrum.

Next, the single muon response in a detector is investigated using the simulation procedure described before. In addition to the propagation of the photons in the scintillator and the light-guide an analytical

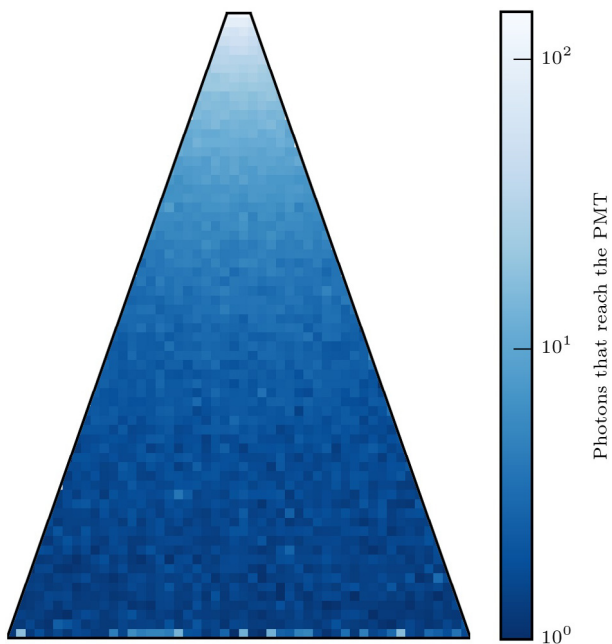


Fig. 11. Cherenkov light yield as function of position in the light-guide for photons reaching the PMT. In the GEANT4 simulation the light-guide was exposed to perpendicular incident relativistic muons. The muon flux was kept constant over the full area. The probability of measuring Cherenkov photons strongly decreases with increasing distance to the PMT.

description of a single photo-electron PMT response is added [28]. Muon direction and energy follow the distributions presented in [29]:

$$I(p, \theta) = \cos^3(\theta) I_V(p \cos \theta) \quad (3)$$

with p the muon momentum and θ the zenith angle. The expression connects the relative flux per unit zenith angle to the perpendicular incident muon flux $I_V(p)$ [$\text{cm}^{-2} \text{sr}^{-1} \text{s}^{-1} \text{GeV}^{-1}$]:

$$I_V(p) = c_1 p^{-(c_2 + c_3 \log(p) + c_4 \log^2(p) + c_5 \log^3(p))} \quad (4)$$

The parameters are $c_1 = 0.00253$, $c_2 = 0.2455$, $c_3 = 1.288$, $c_4 = -0.2555$ and $c_5 = 0.0209$ ($\log \equiv \log_{10}$). The energy and zenith-angle are sampled using the Metropolis–Hastings algorithm [30]. The muons are uniformly distributed across the detector surface. Below -100 mV, the simulation (blue histogram) in Fig. 12 compares well with the experimental data. The small peak at ~ -50 mV is due to the small number of randomly coinciding background pulses in two detectors of the same stack. The small pulse heights stem from Cherenkov radiation in the light-guide. The simulation tends to slightly overestimate the contribution from the Cherenkov photons.

2.5. Detection efficiency

The detector simulation describes the experimental data rather well and can be used to investigate the detection efficiency for electrons, muons and gamma rays. The efficiency depends not only on a combination of gain (i.e. precise value of the MIP-peak) and applied signal threshold (defined in terms of the fraction of the MIP-peak value), but also on the energy of the particle and its angle of incidence. A lower limit on the efficiency is obtained when only perpendicular incident particles are considered. The detection efficiency is then defined as the fraction of the pulse height distribution exceeding the threshold divided by the full range of pulse heights. Only the area covered by the scintillator is considered; the contribution from Cherenkov photons in the light-guide is ignored.

Fig. 13 shows the detection efficiency curves for the various particles. The majority of muons is considered to be minimum ionizing

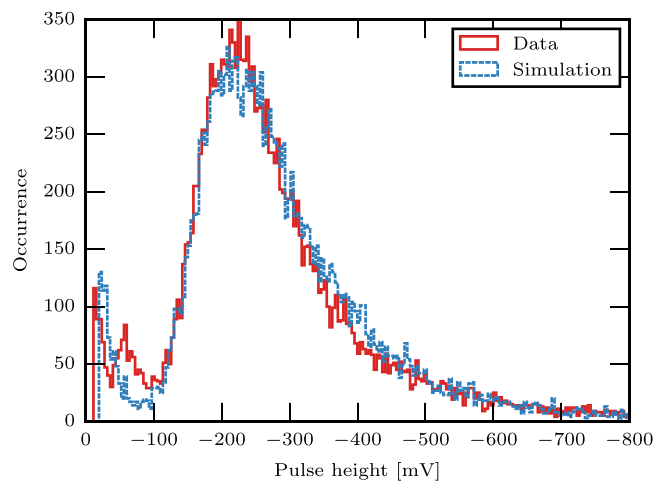


Fig. 12. Comparison between single muon pulse height distributions of simulated and experimental data. Below -100 mV the simulation represents the experimental data rather well. The small peak at ~ -50 mV is due to a small number of randomly coinciding background noise pulses. The contribution below -50 mV is dominated by Cherenkov radiation and is slightly overestimated in the simulation.

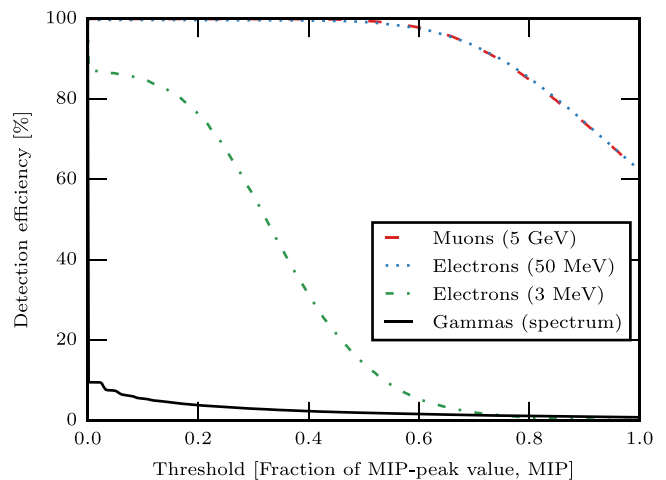


Fig. 13. Detection efficiency of a detector (only the area covered by the scintillator is considered) as a function of threshold (fraction of the MIP-peak value) for muons, electrons and gamma rays. The MIP-peak is defined as the most probable signal response of a single minimum ionizing particle. For high energy electrons (50 MeV, blue, dotted line) and energetic muons (5 GeV, red dashed line) the efficiency curves coincide. For low energy electrons (3 MeV, green dash-dotted line) the efficiency drops significantly. For gamma rays, the full energy range is considered. The energy spectrum is obtained from air shower simulations. The gamma detection efficiency turns out to be small (black solid line).

as they are produced at high altitudes. Electrons however, are also generated in electromagnetic showers closer to ground. Low energy electrons are therefore more abundant than low energy muons. For high energy muons (5 GeV) and high energy electrons (50 MeV) the detection efficiencies are the same (dashed red and dotted blue lines coincide). This is to be expected since at these energies both particles are minimum ionizing. For thresholds up to 0.5 MIP their detection efficiency is very close to 100%. For low energy electrons (e.g. 3 MeV), the detection efficiency is very different (green dashed–dotted line) and stays well below 100% over the full range of threshold values. Low energy particles have a high chance to be absorbed in the scintillator and will therefore produce only a limited number of excitation photons.

Gamma rays are abundant in EASs. Their energy spectrum is obtained from EAS simulations (CORSIKA [31]). The majority of gamma rays will not interact in the scintillator; their detection efficiency will

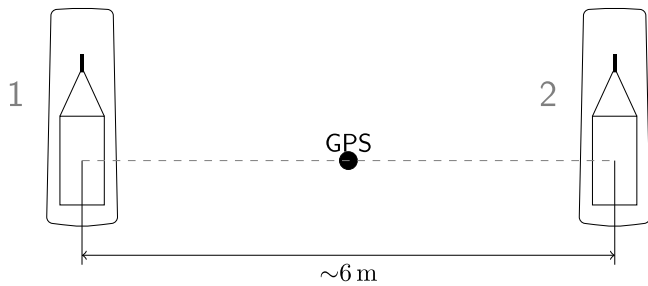


Fig. 14. Typical configuration of a two-detector station. A GPS antenna is located in between the two detectors and provides a signal to precisely timestamp the arrival time of the EAS particles.

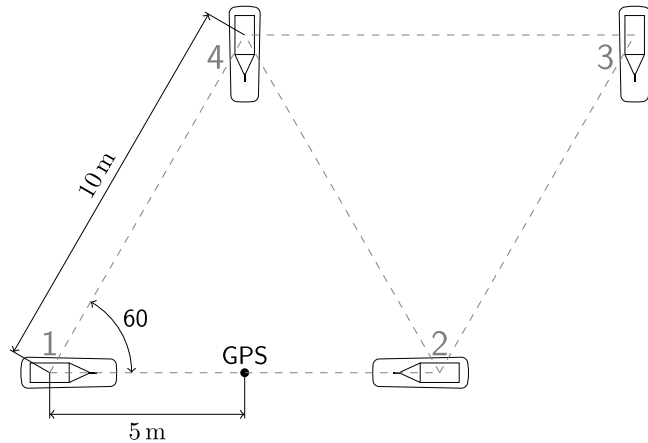


Fig. 15. Four detectors placed in a diamond formation.

therefore be small. For perpendicular incident gamma rays, a lower limit on their detection efficiency as function of threshold is also depicted in Fig. 13 (solid black line). Only a fraction (less than 10%) of the gamma rays will be detected.

3. HiSPARC station

A HiSPARC station combines two or four detectors with the aim to distinguish EASs from single background muons. Since the arrival times of particles in an EAS are highly correlated, this is achieved by demanding a response in two or more detectors within a small time frame.

The layout of a HiSPARC station with two detectors is shown in Fig. 14. The majority of high schools deploy a two-detector station. Four-detector stations explore two different layouts; a diamond formation (Fig. 15) and an equilateral triangle with one detector at the centroid of the triangle (Fig. 16). When at least three detectors in a four-detector station observe one or more particles of an EAS, the direction of the EAS (and thus the direction of the primary cosmic ray) can be obtained by triangulation. When only two detectors are hit, as for a two-detector station, the time difference between the two detectors only allows for the reconstruction of the arrival direction along the axis that connects the two detector centers.

Changes in atmospheric pressure are of importance; the higher the pressure, the larger the probability for a (low energy) shower particle to be absorbed before reaching the ground. Consequently, this will affect the size of the footprint. The pressure also influences at which height the first interaction occurs. Currently, 19 HiSPARC stations are equipped with a weather station [32]. Weather data are collected together with the cosmic ray data and stored in the Nikhef central database.

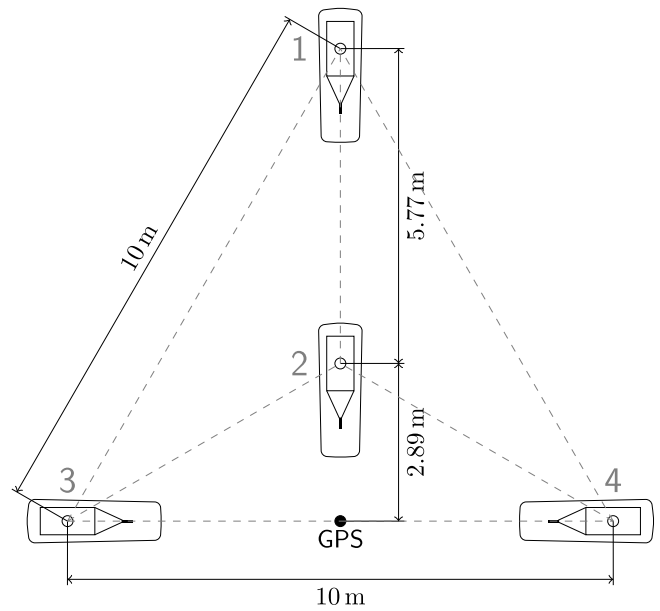


Fig. 16. Four detectors placed in a triangle formation with the fourth detector at the centroid.



Fig. 17. Front (top) and back (bottom) of the HiSPARC readout unit. Two PMTs are connected to the front of the unit and supplied with +12 V DC and a reference voltage. In the PMT-base, this reference voltage (+0.3 to +1.5 V) is converted into a high voltage (−300 to −1500 V). An orange LED next to the signal input flashes when the input signal (−10.3 to +0.3 V) exceeds the lower threshold. A white LED behind the air outlet (center) flashes when a trigger condition is met. At the back of the unit there are connectors for the 1.5 A 12 V DC power supply, GPS antenna cable and USB port for monitor and control of the GPS unit. At the far left there is a USB connector for data output. Two UTP ports facilitate the communication between Master and Slave units. There is also an input for an (+3 V) external trigger. The ADCs are set to a dynamical range of −2.2 V to +0.1 V. The maximum trigger rate is in excess of 30 events per second.

3.1. Read-out electronics

Detectors are connected to read-out electronics by cables with a standard length of 30 m. The custom designed electronics control and read out two PMTs (Fig. 17). One electronics unit facilitates two detectors. For a four-detector station two units are connected in Master-Slave configuration. All four detectors are treated exactly the same. Their signals can be used to construct a matrix of trigger conditions.

The digitization of the analog input signal is carried out by two 12 bits, 200 MHz analog-to-digital converters (ADCs). One ADC is triggered at the rising edge of the clock, the other ADC at the falling edge, thus doubling the sampling frequency to 400 MHz, i.e. at 2.5 ns intervals. A calibration procedure (ADC ‘alignment’) ensures that both ADCs yield the same baseline, gain and dynamical range ($\sim +0.1$ V to ~ -2.2 V or 0–15 MIPs). For larger signals two comparators with adjustable threshold (default at -2.5 V (~ 17 MIPs) and -3 V (~ 20 MIPs) resp.) are added. The comparator data significantly improve the offline reconstruction of large signals that exceed the dynamical range of the ADCs [33].

The output of the ADCs is transferred into embedded memory in a field-programmable gate array (FPGA). The FPGA is clocked at 200 MHz as well. Per channel two thresholds can be defined: a low threshold and a high threshold. The FPGA raises flags when a signal exceeds a threshold. Combining flags (AND/OR) for up to four detector channels provides an extensive matrix of trigger conditions. For test purposes the unit has a separate external trigger input that can be combined with this trigger matrix. Fig. 18 gives an example on how a trigger condition for a two-detector station is generated and how an event is composed.

Assume the trigger condition is defined such that both detectors have to generate a signal that exceeds the low threshold within a limited amount of time. The time window has to be large enough to be fully efficient for detecting particles belonging to the same EAS (with different arrival times, covering all possible angles of incidence of the shower), while it has to be small enough to minimize the random number of coincidences between background muons. In the figure a first flag is raised when detector 1 generates a signal that exceeds the low threshold (channel 1) — the signal may also be large enough to raise the flag for exceeding the high threshold. At the same time a ‘coincidence’ time window is opened. The length of this time window is typically of order $1.5 \mu\text{s}$. If during the $1.5 \mu\text{s}$ the second detector also generates a signal that exceeds the low threshold, the trigger condition is met and an event is generated. An event consists of data taken just before the coincidence window was opened (the pre-trigger time window, typical length $1 \mu\text{s}$), the coincidence time frame ($1.5 \mu\text{s}$) and post-trigger period; the post-trigger time window ($3.5 \mu\text{s}$).

The maximum length of an event window is $10 \mu\text{s}$ (= 2000 12 bit ADC samples = 3 kB memory). Per detector channel, the maximum event size becomes 6 kB. The (embedded) memory may contain multiple events (maximum 3 for the time windows defined above) and acts as a de-randomizing buffer. Events are transferred from the FPGA’s embedded memory into a USB buffer from where they are read by the DAQ PC. If the internal memory is full and the USB buffer is not read out fast enough, new events are discarded until an event is transferred to the DAQ PC. In practice, the trigger rate remains below 1 Hz for all three station layouts. The probability that a fourth event immediately occurs after three consequent triggers ($18 \mu\text{s}$) is negligible.

A small GPS module [34] is mounted on top of the electronics mother board. It generates one second ‘tick marks’ that are sent to the FPGA that generates a separate message for the DAQ software. Each second a counter is started that counts until the trigger flag is raised or the next one second information arrives. In addition to the one second information (in ns) and value of the counter, the message also includes the number of GPS satellites and their signal strength. The DAQ software combines three one second messages, the counter values for the intervals and GPS quantization error to calculate the precise time stamp of an event. In a four-detector station a ‘Slave unit’ is connected to the ‘Master’. Both Master and Slave contain the same electronics and FPGA firmware. However, the Slave unit has the GPS module removed. By connecting the two units (Fig. 17), the trigger matrix is extended to four channels. The individual Master and Slave messages are combined by the DAQ software to produce the full event. Each second and for all channels, also the number of times a signal exceeds the high and/or low threshold is recorded.

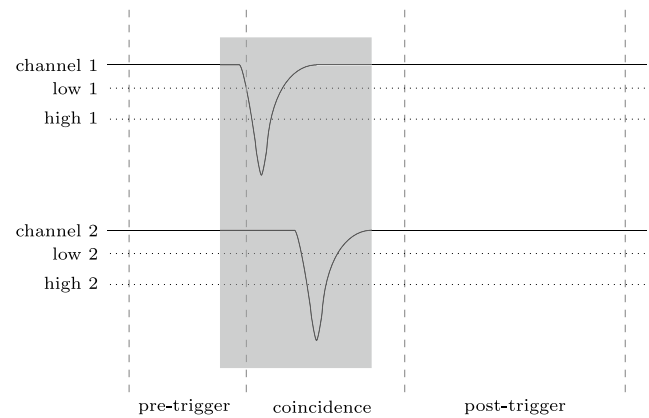


Fig. 18. A schematic representation of an event. Dashed vertical lines: the pre-trigger ($1 \mu\text{s}$), coincidence ($1.5 \mu\text{s}$) and post-trigger ($3.5 \mu\text{s}$) windows. Shaded area: the data-reduction window. Data outside this window will not be stored. The dotted lines indicate the low and high threshold values.

There exist two versions of the electronics. They have the same functionality. The HiSPARC II unit has an on-board memory chip that contains firmware that is loaded into the FPGA at power-up. The HiSPARC III box has an additional USB channel via which the DAQ software transfers the firmware to the FPGA at run-time.

3.2. DAQ software

To control the electronics unit and to monitor the data, a graphical user interface based on LabVIEW™ [35] has been developed that executes on a Windows™ PC. A screenshot of one of the interface panels is shown in Fig. 19. Typical pulse height distributions for a four-detector station are shown in the central graph. On the right the pulse integral distributions are displayed. With a typical pulse length of 30 to 50 ns (Fig. 4) a zero-suppression algorithm is applied to reduce event sizes.

The DAQ software performs a preliminary analysis of the data. The baseline (which is also used by the zero suppression algorithm) and the fluctuation of the baseline are recorded. Subsequently, the pulse height and pulse integral are calculated. The software also keeps track of the number of accessible GPS satellites and their signal strength over time.

The DAQ software stores the raw data from the electronics unit(s) combined with the first analysis results into a local MySQL [36] database. A Python [37] program monitors the number of written events and uploads the data at regular intervals to the data-store server at Nikhef.

The Nagios [38] software package monitors the status of both hardware (electronics and PC) and DAQ software for each station. In case a service fails an e-mail is generated and automatically forwarded to the person responsible for the station with the request to intervene. To avoid down time during weekends and long high school (summer) breaks, Nikhef operates an OpenVPN [39] network that gives remote access to each station (TightVNC [40]). This network is also used to install software updates on the stations.

3.3. Trigger efficiency

The default trigger condition for a two-detector station is straightforward. An event is selected if the signals in the two detectors exceed the low threshold within the coincidence time window ($1.5 \mu\text{s}$). The default trigger conditions for a four-detector station are: the signals of at least two detectors pass the high threshold or at least three detectors generate a signal that exceeds the low threshold.

Fig. 13 demonstrated that the detection efficiency for perpendicular incident MIPs traversing a detector is $T_L = 0.999$ (0.2 MIP) for the low threshold and $T_H = 0.997$ for the high threshold (~ 0.5 MIP). For a

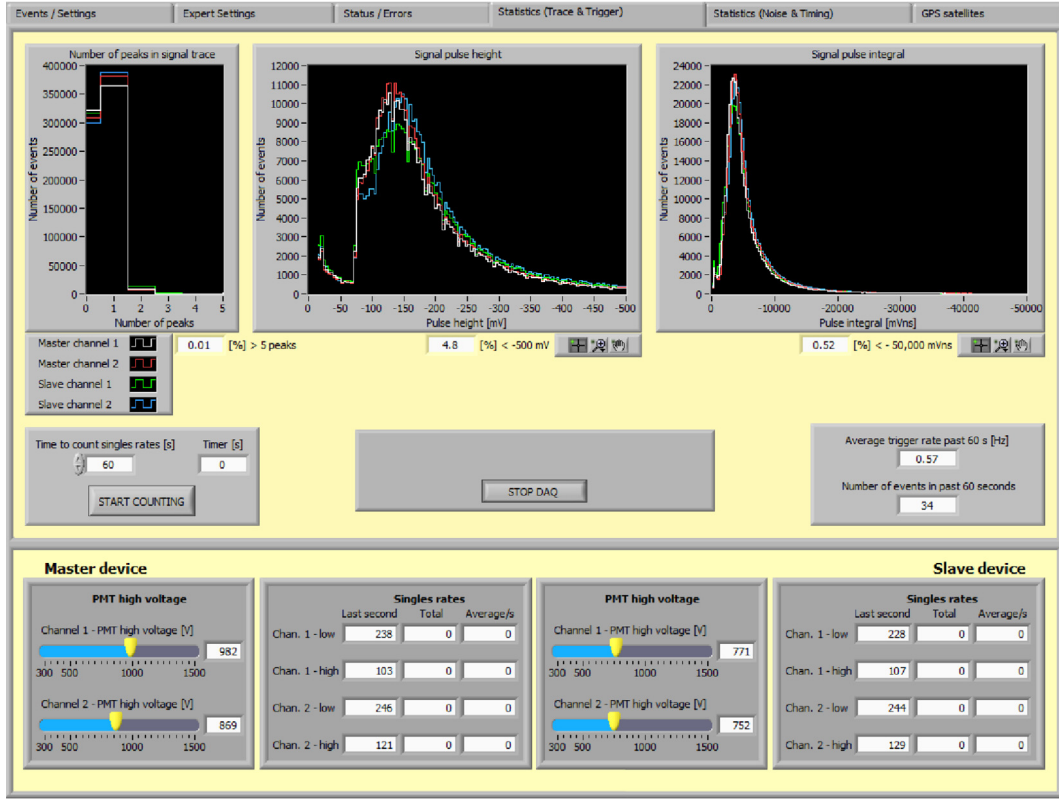


Fig. 19. Screenshot of a panel in the HiSPARC DAQ control and monitor user interface. The top three histograms show (from left to right) the number of signals above threshold in an event for each detector, the pulse heights and the pulse integral distributions resp. The bottom control panels set the PMT high voltage for Master (left) and Slave (2nd from right). The other panels display single rates for signals exceeding low and high thresholds. Additional tabs (visible at the top of the screenshot) include event settings, status and error messages, and display GPS data. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

two-detector station this leads to a trigger efficiency of $T_L^2 > 0.99$. For a four-detector station the trigger efficiency depends on the number of detectors that are hit. For two MIP events the trigger probability depends on the high threshold: $T_H^2 > 0.99$. For three MIP events, for which it is unknown which three detectors are hit, the station will still trigger even if one of the particles is not detected ('1 non'). If one of the MIPs does generate a signal that goes over the low threshold, the two other MIPs may still generate a signal that goes over the high threshold (2 high). This combination occurs three times:

$$\begin{aligned}
 P_{3\text{mip}}^{\text{four}} &= P(3 \text{ low}) \\
 &\quad + 3 \cdot P(2 \text{ high} | 1 \text{ non}) P(1 \text{ non}) \\
 &= T_L^3 + 3 \cdot T_H^2 \cdot (1 - T_L) \\
 &\gg 0.99
 \end{aligned} \tag{5}$$

Finally, for four MIPs crossing four different detectors the station is also very efficient:

$$\begin{aligned}
 P_{4\text{mip}}^{\text{four}} &= P(4 \text{ low}) \\
 &\quad + 4 \cdot P(3 \text{ low} | 1 \text{ non}) P(1 \text{ non}) \\
 &\quad + 6 \cdot P(2 \text{ high} | 2 \text{ non}) P(2 \text{ non}) \\
 &= T_L^4 + 4 \cdot T_L^3 \cdot (1 - T_L) \\
 &\quad + 6 \cdot T_H^2 \cdot (1 - T_L)^2 \\
 &\gg 0.99
 \end{aligned} \tag{6}$$

In all cases the trigger efficiency of HiSPARC stations for MIPs is well above 99%. In practice not only MIPs will be detected. Also gamma rays, low energy electrons and low energy muons are part of the EAS. The detection efficiency then strongly depends on their energy and nature.

3.4. Timing offsets

Assuming that the distribution of the angle of incidence of EAS is isotropic, each detector of a station has about equal probability to be hit first. Fig. 20 shows the (arrival) time differences between two detectors in a two-detector station (station #4). About 8×10^5 events were collected between the 5th of March and the 2nd of April 2018. The shift of the peak with respect to the dashed line at 0 ns shows an average timing offset between the two detectors of 12 ns. A plateau (blue horizontal line) is reached for offsets larger than ~ 300 ns. This is the result of random coincidences, i.e. two uncorrelated particles, a particle and a spontaneous emission in the PMT etc. The height of the plateau is obtained by making a fit for offsets between 300 ns and $1.5 \mu\text{s}$ (default coincidence time window). The expected number of random coincidences per second (N) can be calculated with $N = 2\tau r_1 r_2$. τ is the coincidence time window and r_1 and r_2 are the recorded single rates in the detectors. Integration over the full length of the plateau region yields 3.21×10^5 events. This is well in agreement with the estimated number of random coincidences: 3.17×10^5 . Note that the number of random coincidences within a time difference of 300 ns (blue crossed region) is small.

The difference in arrival times in detectors of a four-detector station are shown in Fig. 21. Station #501 (diamond formation, Fig. 15), recorded 3.1×10^5 events from the 1st of January till the 1st of March 2018. Only time differences in events for which all four detectors generated a signal exceeding the low threshold are considered. The random coincidence plateau is absent. The probability that four random signals coincide within $1.5 \mu\text{s}$ is thus very small. The timing offset in detector combinations 1–2 (blue) and 1–4 (green) show almost identical distributions. The distance between the detectors in the two pairs is the same. The distance between detectors 1 and 3 is larger as

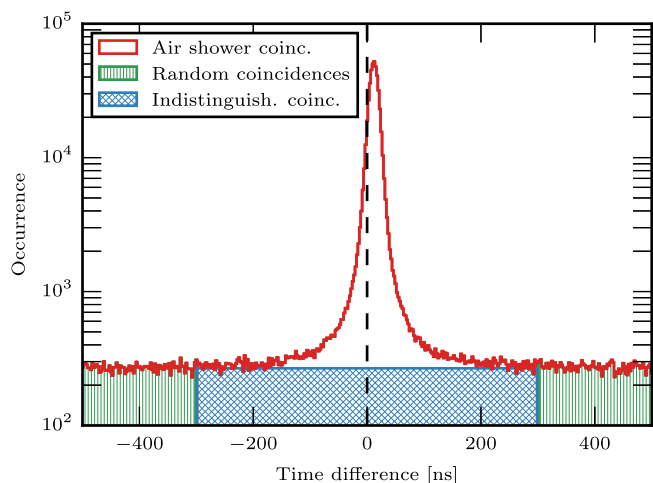


Fig. 20. Timing offset between two detectors in a two-detector station (in red). The plateau (blue horizontal line) is due to random coincidences (uncorrelated particles, PMT noise etc.). For time differences smaller than 300 ns (blue crossed region) the random coincidences are indistinguishable from air shower coincidences. Their contribution is however small. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

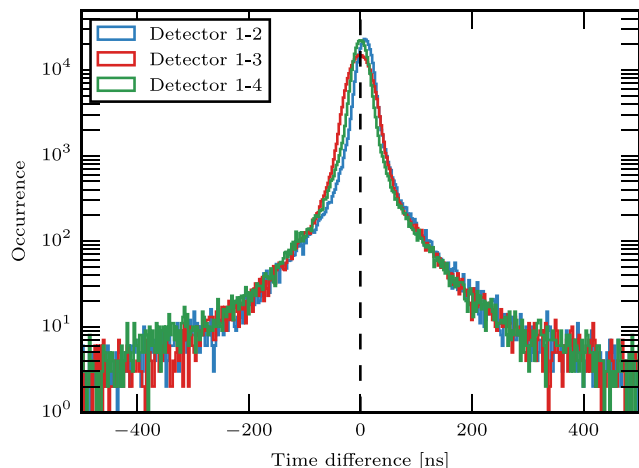


Fig. 21. Timing offsets between detectors in a four-detector station (color online). Only events in which all detectors generated a signal exceeding the low threshold are taken into account. There is no plateau due to the small probability that a random coincidence between the four detectors occur at the same time. The distance between detectors 1 and 2, and detectors 1 and 4, are equal and show an almost identical time offset distribution (in blue and green resp.). The distance between detectors 1 and 3 is larger; they lie along the diagonal of the diamond. The time difference distribution (in red) is therefore slightly broader. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

they lie along the diagonal of the diamond. The time offset distribution for this combination (in red) is therefore slightly broader.

For each station the timing offsets are calculated and stored on a daily basis. They are important parameters in for instance the direction reconstruction of an EAS.

4. Clusters of stations

A HISPARC cluster is a collection of stations. The surface that the cluster covers often depends on the number of stations that a high school hosts and/or how many high schools there are in the neighborhood. For extensive research a dense cluster at the Amsterdam Science Park is created (Fig. 22). The Science Park cluster contains thirteen four-detector stations. The distance between the stations varies from

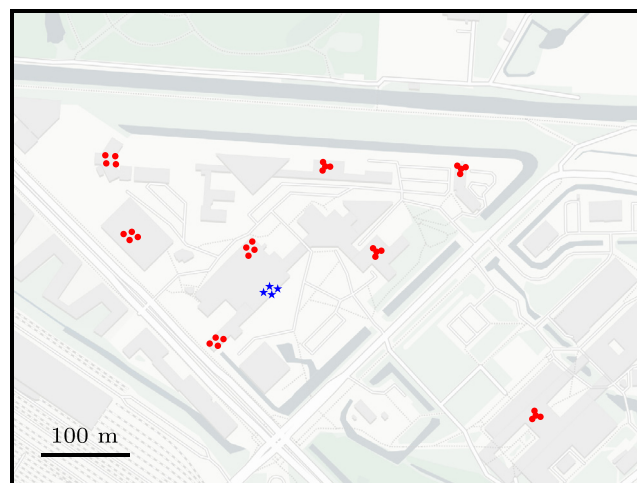


Fig. 22. Location of stations at the Amsterdam Science Park cluster. One of the stations is located inside the Nikhef building and is not on the map. Each red dot represents a detector and each combination of four dots forms a station. There are diamond and triangle shaped station configurations. At one location, four ‘diamond’ stations are essentially positioned ‘on top of each other’ (detectors are ~ 1 m apart) and only a single station is displayed (blue stars).

~ 1 m up to 280 m. Several stations may sample the same (large) EAS while each station generates its own GPS time stamp. To reconstruct the angle of incidence of the primary cosmic ray the arrival time of EAS particles in each station has to be precisely known. This implies that it is crucial to account for time offsets between GPSs.

4.1. GPS timing

Each GPS is operated in ‘overdetermined clock mode’ and is at a fixed location. Upon installation, the GPS receiver performs a 24 h self-survey to accurately determine its position which, in turn, also provides the absolute coordinates for the detectors of that station. The precision of the self-survey is investigated by performing multiple self-surveys using various combinations of GPS antennas, cables and GPS modules (of same type and make). These systematic studies show that 50% of the longitudinal differences stay within 0.73 m, 75% within 1.8 m and 95% within 4.1 m while for the latitude differences are 50% within 0.48 m, 75% within 1.1 m and 95% within 2.6 m. Analysis of the altitude data shows that 50% of the combinations stay within 1.5 m, 75% within 2.8 m and 95% within 6.1 m. The manufacturer of the GPS electronics quotes a 1σ accuracy of 15 ns [34]. When combining data from multiple stations in reconstructing an EAS, the precise timing offsets between all stations in a cluster have to be known. By replacing the detectors by stations, the same method can be applied with which the timing offsets between detectors within one station were obtained. Again, the angle of incidence of EASs is assumed to be isotropically distributed; stations have equal probability to be hit first. The timing offsets between ~ 100 station pairs have been examined. Combining their offsets, a Gaussian distribution with $\mu = 2.7$ ns and $\sigma = 18.9$ ns is obtained. As the distance between the majority of the stations is (much) larger than the distance between the detectors within a station, the rate at which coincidences between stations occur is much smaller than the event rate in a station. Consequently, GPS offsets can usually not be derived on a day-to-day basis due to lack of statistics.

4.2. Acceptance

Obviously, a station near the equator covers a different part of the sky when compared to a station near, for instance, the North Pole. Moreover, with increasing angle of inclination w.r.t. the zenith, the distance the shower particles have to travel through the atmosphere

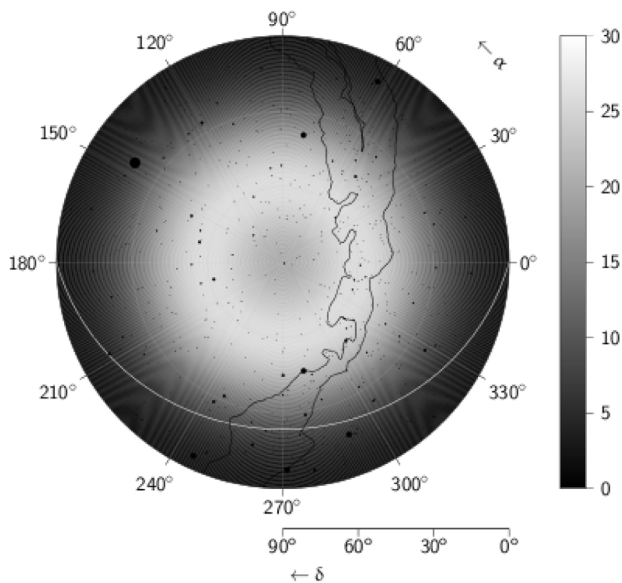


Fig. 23. The acceptance of an Amsterdam Science Park station integrated over a day. The sky map is represented in equatorial coordinates. Due to the Earth's rotation cosmic ray sources will only be visible part of the day. The integrated acceptance over a full day is shown in gray-scales from 0% (black) to a maximum of 30% (white). The boundaries of the Milky Way are shown as black curved lines while major stars are indicated by black dots. The ecliptic is shown in white. Note that only cosmic rays from the Northern Hemisphere can be observed.

increases. This implies that for the same energy of the primary particle the number of shower particles that reach the ground decreases with increasing zenith angle.

The zenith angle (θ) dependent rate distribution integrated over all energies can be approximated by: $dN/d\theta \propto 2\pi \sin\theta \cos^\beta \theta$ with $\beta = 6$ [41]. The rate for small zenith angles is therefore suppressed ($\theta \rightarrow 0^\circ$). The zenith angle dependent flux is obtained by dividing by the geometrical factor $2\pi \sin\theta$ (thus a high flux for small zenith angles). Next, the coordinate system in which the flux (integrated over azimuth) is defined, is converted into an equatorial coordinate system with α the right ascension and δ the declination. In addition, the longitude, latitude and altitude of the station, and the rotation of the Earth are taken into account. The acceptance is calculated in this coordinate system while the flux is integrated over 24 h.

We assume that an EAS generated by a source directly above Amsterdam Science Park ($\theta = 0^\circ$) has an acceptance of 100%. As the Earth rotates the zenith angle increases and the number of EAS particles that reach the ground decreases. At some point the source disappears behind the horizon. The overall acceptance integrated over 24 h becomes $\sim 30\%$ (light 'donut' shaped band in Fig. 23). Sources at $\delta = 0^\circ$ will be below the horizon most of the time. As a result, the acceptance for primaries stemming from the equatorial plane diminishes [42]. The maximum acceptance is obtained at equatorial directions that match small zenith angles in the local coordinate system of the station. Despite the fact that the area around Polaris is always visible from Amsterdam Science Park, and that the acceptance is relatively high, it never reaches 100% since the zenith angle is always larger than zero.

5. Data processing

All stations (~ 140) send their data to Nikhef where they are stored in HDF5[®] [43] format. Each night, a Django [44] web application preprocesses the raw data and generates an event summary database (ESD). The ESD contains information such as the pulse heights and timestamp of an event, detector position and timing offsets, etc. The same application supports direct web access to the ESD. It also provides

an application programming interface (API) through which also raw data are publicly accessible.

For data manipulation a Python based module has been developed. The SAPHIRE (Simulation and Analysis Program Package for HiSPARC Research and Education) [45] facilitates downloading HiSPARC data and performing analyses. It contains an extensive set of EAS reconstruction tools. In addition to the Python module a JavaScript library is developed called jSparc [46]. With jSparc students can work with (or create their own) web applications to explore the HiSPARC data.

The SAPHIRE (Python) library also forms the basis for a set of Jupyter Notebooks [47]. These notebooks are developed for use in the classroom and for (high school) student research projects.

6. EAS direction reconstruction

In Section 2 a detailed simulation of the single particle response of a HiSPARC detector was presented. Using this detector simulation, the response of a four-detector HiSPARC station to EASs can be investigated. With CORSIKA [31], proton initiated EASs are generated with energies ranging from 10^{13} to $10^{16.5}$ eV. Their relative abundance follows the cosmic ray energy spectrum. 'Thinning' [48] was not applied. For high energy hadronic interactions the QGSJET-II [49] model is selected. Interactions of hadrons with energies below 80 GeV are simulated using GHEISHA [50]. Electromagnetic interactions are described by the EGS4 [51] model. While the location of the station in the simulations remains fixed, the position of the EAS core is randomly chosen within a circle with a radius of 100 m centered at the station. Arrival directions are chosen isotropically. When one or more EAS particles hit a detector the full detector simulation is applied. For events satisfying the trigger conditions and having at least two MIPs in each detector, the direction of the shower is reconstructed assuming a flat shower front using the (triangulation) algorithm described in [52, Chapter 5].

Fig. 24 gives the 1σ uncertainty in the shower direction reconstruction for a four-detector triangle shaped station (Fig. 16) as a function of zenith angle (blue dots). The distribution is obtained by comparing the direction of the primary cosmic ray as set in the CORSIKA Monte Carlo program and the reconstructed direction after detector simulation. The average uncertainty ($\theta < 40^\circ$) is 7.7° .

The shower direction can be decomposed in terms of zenith angle and azimuth. Fig. 25 shows the 1σ uncertainty in the azimuthal angle (blue dots) and zenith angle (blue crosses) as a function of zenith angle. Again, the distribution is obtained by comparing the direction in which the shower developed as set in the CORSIKA Monte Carlo program and the reconstructed direction after detector simulation. With increasing zenith angle, the 1σ uncertainty on the reconstructed zenith angle slightly increases. The uncertainty on the reconstructed azimuth however, rapidly increases for smaller zenith angles; in the limit where the zenith angle goes to zero the azimuth becomes undefined.

In 2008 a four-detector HiSPARC station (triangle configuration) was integrated in the KASCADE experiment [7]. When KASCADE detected an EAS in the area where the HiSPARC station was located, the experiment generated a signal that triggered the DAQ of the HiSPARC station. Between July 1 and August 6 2008 more than 5×10^5 events were recorded [13]. The direction of these showers was reconstructed by applying the same algorithm as used in the reconstruction of the simulated EASs (again demanding at least 2 MIPs in each detector). KASCADE reconstructed the direction of these EASs with zenith angles between 0° and 40° with 0.3° accuracy [7]. Fig. 24 shows the 1σ uncertainty on the HiSPARC direction reconstruction (red dots). The average uncertainty ($\theta < 40^\circ$) is 6.1° . Fig. 25 gives the decomposition into the azimuth (red dots) and zenith (red crosses) angles as a function of the zenith angle.

When compared to KASCADE data it appears that the uncertainty obtained from the simulations (7.7°) slightly underestimates the real direction reconstruction performance of the HiSPARC station (6.1°). However, the simulated data only contain proton initiated showers with

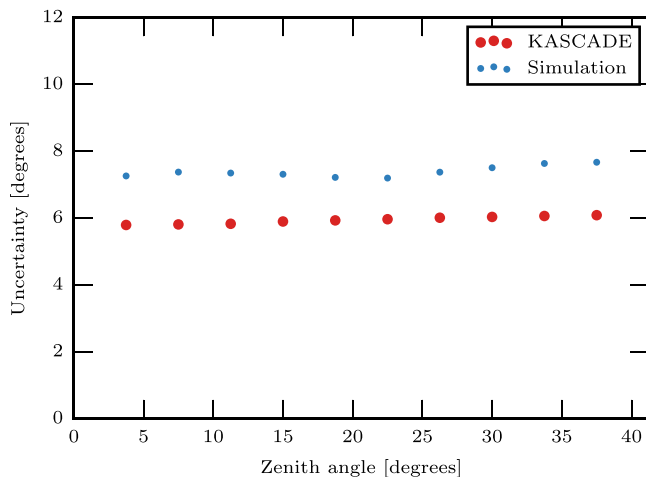


Fig. 24. The 1σ uncertainty in the shower direction reconstruction as a function of zenith angle (blue dots) is calculated by comparing the direction of CORSIKA generated showers and the direction reconstructed after full detector simulation in a four-detector, triangle shaped, HiSPARC station. A station with the same configuration was integrated in the KASCADE experiment. The reconstruction algorithm that was applied to the simulated EASs was used to obtain the shower directions from the HiSPARC data in the KASCADE setup. The algorithm only used events (satisfying the trigger conditions) with at least 2 MIPs in each detector. The comparison between the shower direction reconstructed by KASCADE (0.3° accuracy) and measured by the HiSPARC station is shown as a function of zenith angle (red dots).

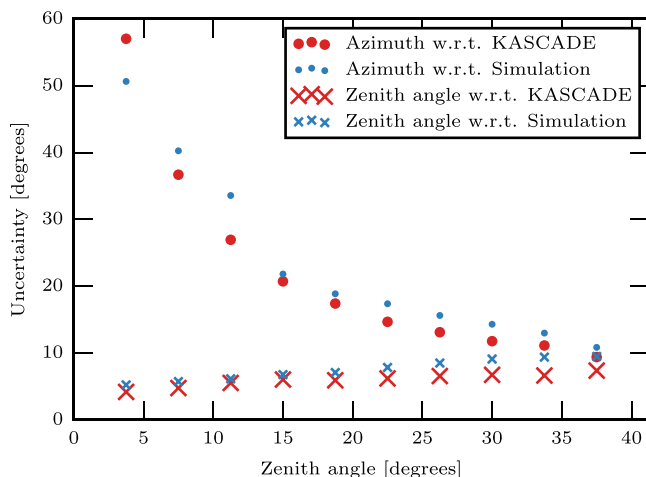


Fig. 25. The 1σ uncertainty in the reconstruction of azimuth (blue dots) and zenith (blue crosses) as a function of zenith angle are calculated by comparing the direction of CORSIKA generated showers and the direction reconstructed after full detector simulation in a four-detector, triangle shaped, HiSPARC station. A station with the same configuration was integrated in the KASCADE experiment. The reconstruction algorithm that was applied to the simulated EASs was used to obtain the shower directions from the HiSPARC data in the KASCADE setup. The reconstruction algorithm only used events (satisfying the trigger conditions) with at least 2 MIPs in each detector. The comparison between the shower direction reconstructed by KASCADE (0.3° accuracy) and measured by the HiSPARC station, is shown for azimuth (red dots) and zenith (red crosses) as a function of the zenith angle.

energies starting at 10^{13} eV. The KASCADE data contain all primary cosmic ray compositions while only showers with an energy in excess of 10^{14} eV are reconstructed. This results in a higher contribution from events with slightly higher particle densities. Moreover, in the simulations the shower core positions were evenly distributed in all directions up to a distance of 100 m from the station center. The nearest boundaries of the (square) KASCADE array are in two directions only ~ 55 m and ~ 70 m away from the HiSPARC station [13]. The contribution from showers with their core position close to or beyond

Table 1

The first column lists the pairs of stations for which the reconstruction of the shower direction was compared. The second column gives the 1σ difference between the two angles. For details see [53].

Stations	Difference [$^\circ$]
501-510	6.02 ± 0.04
501-512	6.14 ± 0.04
501-513	6.06 ± 0.04
510-512	6.35 ± 0.04
510-513	6.37 ± 0.04
512-513	5.93 ± 0.04

these boundaries are therefore suppressed. This will also reduce the number of lower multiplicity events observed by the HiSPARC station.

Recently, one month of data (April 2019) from four closely spaced four-detector diamond shaped stations (#501, #510, #512 and #513) was analyzed [53]. The relative distance between the centers of the stations ranges from ~ 1.5 m to ~ 5 m. The direction of showers was reconstructed provided all four stations triggered at the same time and all 16 detectors observed a signal corresponding to at least 2 MIPs. These conditions favor the high multiplicity region in air showers and resulted in 123800 events. In each station the direction of the shower was reconstructed. By pairwise comparing the directions 6 distributions were obtained. For each distribution the 1σ difference between the two measurements was calculated. The results are listed in Table 1. They agree well with the outcome of the KASCADE data analysis.

7. EAS energy reconstruction

Despite the evident limitations of a single HiSPARC station, i.e. the limited number of detectors, the relatively small distance between the detectors and the limited precision in the determination of the number of particles traversing a detector, it is possible to reconstruct the energy of the primary cosmic ray using a single four-detector station. The energy is determined by fitting a lateral density function (LDF), i.e. an analytical description of the particle densities in EAS footprints, to the measured particle densities. The number of particles traversing a detector is determined by dividing the measured pulse integral by the MIP-peak value. Since only the particle numbers at four detector positions can be determined, the maximum number of free parameters in the LDF is three; two for the core position and one for the energy. The core positions of low energy showers ($\sim 10^{14.5}$ eV to $\sim 10^{15.5}$ eV) are expected to lie within the station since EASs with their core outside the station are not likely to result in a trigger. For large showers there are two solutions to the LDF fitting procedure; one with the core inside and one outside the station area. Also, if all four detectors measure a similar number of particles, the data contain no information that can be used to accurately reconstruct the core position and thus the energy. Finally, in order to determine the energy flux, the effective surface area of an EAS footprint as function of energy needs to be known. Proton induced CORSIKA simulations in the energy range from $10^{14.5}$ eV to $10^{16.5}$ eV have been used to determine the effective surface area and to define the LDF.

Fig. 26 shows the simulated particle numbers as a function of distance to the shower core for multiple energies. For each shower the number of particles that traverse a 1 m by 0.5 m detector has been determined for various core positions. The markers display the averages of these particle numbers. For each energy a different marker is used. Only perpendicular incident proton simulations are shown. The fluctuation in the number of particles within showers of the same energy is large. The shaded regions show the 1σ spread. The lines show the LDF used for the energy reconstruction. Despite the limited number of free parameters, the particle numbers are still well described by the function. A modification of the Nishimura–Kamata–Greisen formula [1, and references therein] has been used as LDF. The number of free parameters in the original NKG formula has been reduced. The age

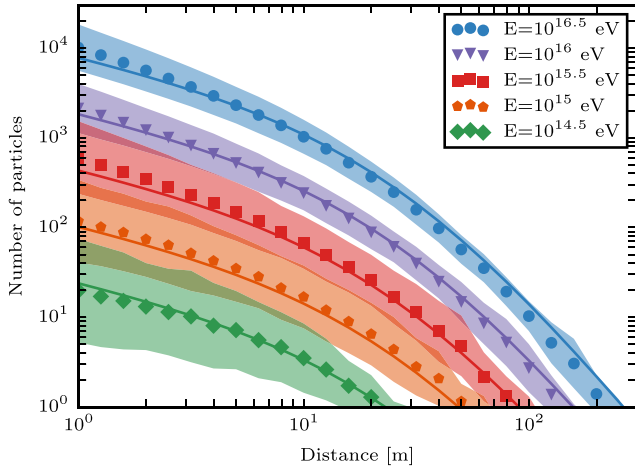


Fig. 26. The simulated average number of particles that fall inside a 0.5 by 1 m rectangle as a function of distance to the shower core at different energies. The markers show the average particle numbers. The shaded regions show the 1σ fluctuations. The lines show the modified NKG formula for each energy. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

parameter which is related to the shower maximum is effectively taken as a constant. In the simplified NKG formula the number of particles N at distance r is given by:

$$N(r) = A \left(\frac{r}{r_0} \right)^a \left(1 + \frac{r}{r_0} \right)^b \quad (7)$$

with $r_0 = 29.6$, $a = -0.566$, $b = -2.57$ and A the fit parameter related to the energy. For inclined showers the particle numbers are reduced due to the increase in path length through the atmosphere. This is corrected for by using:

$$A_{\perp} = A \cdot \exp \left(p \left(\frac{1}{\cos \theta} - 1 \right) \right) \quad (8)$$

with θ the zenith angle and $p = 6.937$. The energy of the primary cosmic ray is then calculated with:

$$\log(E) = c \cdot (\log(A_{\perp}) + d) \quad (9)$$

here $c = 0.797$ and $d = 17.62$. The modified NKG formula is circle symmetric and two parameters are required to determine the core position.

Three months of coincidences between stations 501 and 510 have been used to investigate the energy resolution obtained using the modified NKG method. Only events in which all 8 detectors of both stations measured a signal above 2 MIPs are considered. Furthermore, only events for which both directions deviate less than 15° are used. Also, the average zenith angle has to be less than 35° . Finally, if the spread in the number of MIPs in the four detectors is less than 2, the event is discarded. This selection avoids events with similar particle numbers that contain no information about the core position. Coincident events between station 501 and 510 provide two independent energy measurements. Per reconstruction two initial guesses for the parameters are used. One with the core position inside the four detectors and another one outside. The fit parameters of the solution with the lowest χ^2 value are used. If the best χ^2 value of one of the two stations is below 5, the event is discarded. Fig. 27 shows the distributions of reconstructed EAS energy differences between station 501 and 510. Four distributions are displayed each containing a selection of energies as determined by station 501 (within a $\log(E) = 0.5$ bin width). At higher energies a bimodal distribution appears reflecting the two solutions for the core position. Multiple stations at sufficiently large distance are required to resolve this problem.

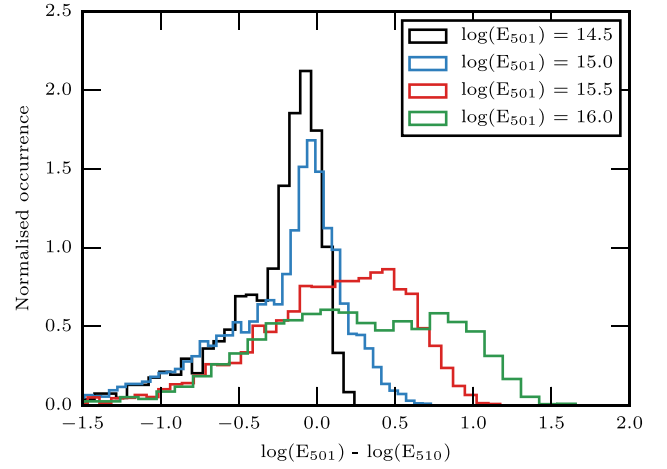


Fig. 27. Distributions of reconstructed EAS energy differences between station 501 and 510. Four distributions are displayed each containing a selection of energies as determined by station 501 (within a $\log(E) = 0.5$ bin width). At higher energies a bimodal distribution appears reflecting the two solutions for the core position (inside and outside the station). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The cosmic ray energy spectrum is obtained by dividing the distribution of energy measurements by the solid angle (zenith angle of 35°), the time span (3 months) and the effective surface area of an EAS footprint. CORSIKA simulations have been used to obtain the radius at which the particle number drops below 2 (Fig. 26). The effective surface area at an energy is determined as the area of a circle with that radius. The radii at intermediate energies have been estimated by fitting a second order polynomial to the radii of the simulated energies. Fig. 28 shows the energy spectrum obtained using stations 501 (red stars) and 510 (blue dots). Both stations yield very similar results. For energies between $10^{14.8}$ eV and $10^{15.5}$ eV (dotted lines) a slope (α) has been fitted to both distributions (solid lines). The values of these slopes (2.85 and 2.86) are similar and do not deviate much from the known value of 2.7 [1]. For higher energies the measured energy spectrum starts to soften whereas a steepening is expected. This follows from the bimodal distribution at higher energies (Fig. 27). If only solutions to the energy reconstruction problem with the core position inside the four detectors are taken into account, the spectrum steepens rapidly and no energies larger than 10^{16} are measured. The black dashed line shows the known energy spectrum with a steepening to a slope of 3.1 at $3 \cdot 10^{15}$ eV. The large offset between the HISPARC measurement and the spectrum is due to the detection efficiency and because of the numerous analysis cuts that are applied (e.g. all detectors have at least 2 MIPs).

A similar analysis using KASCADE data has been carried out. Unfortunately the number of KASCADE reconstructions was too limited for a decisive analysis. The uncertainty in the $\sim 10^{15}$ eV region seems to agree with the analysis using stations 501 and 510. For two-detector stations it is not possible to reconstruct the energy of individual EASs. However, by investigating the pulse height distribution it is possible to probe the cosmic ray energy spectrum. This will be discussed in a separate paper [27]. In several regions stations are positioned close enough that a combination of stations can be used for energy reconstruction. A preliminary study with the Science Park Cluster [52, Chapter 7] has been carried out. A more elaborate study that uses an AI algorithm in combination with properties derived from the pulse size and pulse shape (e.g. particle multiplicity and arrival time) is in progress [27].

8. Shower detection

To derive the EAS detection efficiency of a station a detailed comparison has to be made between the measured detector response curve

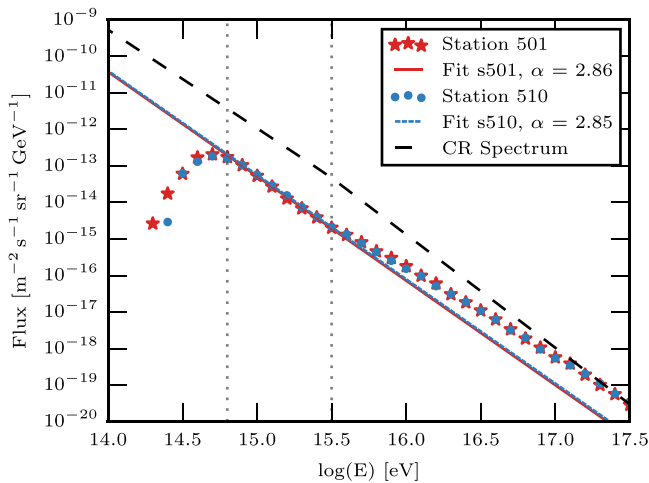


Fig. 28. The energy spectrum obtained using stations 501 (red stars) and 510 (blue dots). For energies between $10^{14.8}$ eV and $10^{15.5}$ eV (dotted lines) a slope (α) has been fitted to both distributions (solid lines). For higher energies the measured energy spectrum starts to soften whereas steepening is expected. This is due to the bimodal energy reconstruction for higher energies. The black dashed line shows the known energy spectrum. The large offset between the HiSPARC measurement and the spectrum is due to the detection efficiency and because of the numerous analysis cuts that are applied.

and the pulse height distribution obtained in the simulation. Showers are again generated in an energy range from 10^{13} to $10^{16.5}$ eV following the cosmic ray energy spectrum with randomly chosen directions and EAS core positions.

8.1. Pulse height distribution

Fig. 29 shows the pulse height distribution of a detector in a two-detector station compared to simulations. For pulses in excess of -400 mV the simulation (blue) agrees reasonably well with the data (red). Note that EASs with higher energies than $10^{16.5}$ eV are not included. For small pulses however, there appears to be a clear discrepancy. In the simulation however, an important contribution is absent. Only EASs with energies larger than 10^{13} eV are considered. Lower energy showers responsible for the single muon background are not taken into account. When an energetic muon decays prior to reaching the Earth it will produce an energetic electron. This electron initiates an electromagnetic ‘mini-shower’. These muon-induced mini-showers are indeed present in for instance 10^{10} eV proton induced CORSIKA showers. Adding this contribution resolves the discrepancy and is discussed in detail in [27].

Alternatively, contributions from muon-induced mini-showers can also be rejected. As the number density of these mini-showers is low, the probability of detecting a mini-shower rapidly decreases when three or more detectors are required to produce a signal in excess of -30 mV. **Fig. 30** shows the simulated pulse height distribution for a detector in a four-detector station (blue) and the measured pulse height spectrum (red). All four detectors generated a signal in excess of -30 mV. The discrepancy for small pulse heights completely disappears; demanding a signal in all four detectors completely removes contributions from single muon induced mini-showers. Contrary to what is observed in **Fig. 29**, the experimental data now show a pronounced excess for large pulses. Since EAS energies beyond $10^{16.5}$ eV are not included, indeed the simulations underestimate the contribution from EASs with higher particle densities [27].

The pulse height distributions in **Figs. 29** and **30** receive contributions from electrons, muons and gamma rays. This is demonstrated in **Fig. 31** where the contribution from a number of particles (and their combinations) to the simulated spectrum in **Fig. 29** (blue) is shown. The contributions from one electron (black), one gamma (red), one muon

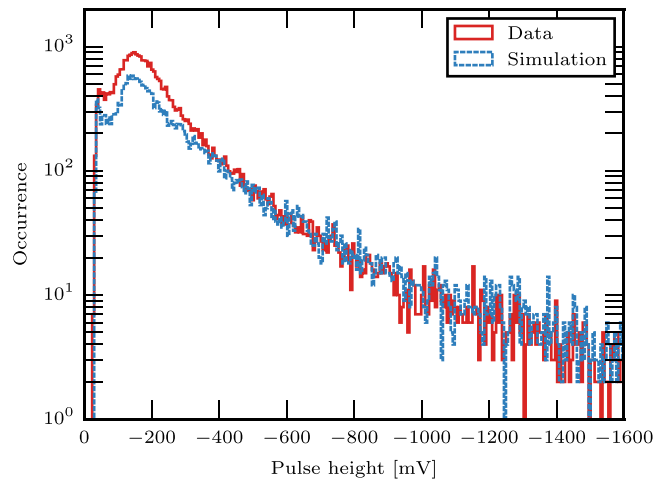


Fig. 29. Pulse height distribution of a detector in a two-detector station (red, solid) compared to simulations (blue, dashed). The contribution from random coincidences has been removed from the data while the MIP-peak value has been corrected for temperature fluctuations. For large pulses, data and simulations agree rather well. The discrepancy for small pulses is caused by the absence of muon induced ‘mini-showers’ (muons decaying into single electrons which in turn generate an electromagnetic shower) in the simulations. These events are caused by the large number of showers for which the primary particle carries an energy less than $\sim 10^{12.5}$ eV.

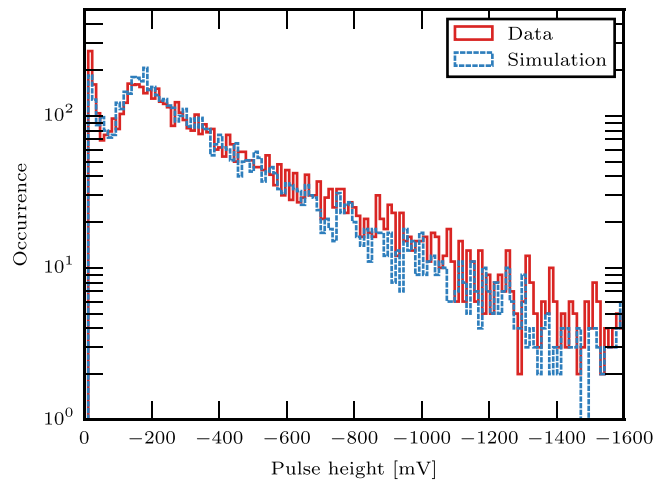


Fig. 30. Pulse height distribution of a detector in a four-detector station (red, solid) compared to simulations (blue, dashed). Only events that contain a signal in excess of -30 mV in each of the four detectors are selected. In contrast to **Fig. 29**, there is no discrepancy between data and simulation for small pulse heights as the size of the shower footprint has to be as large as or larger than the size of the station (i.e. EASs with an energy larger than $\sim 10^{13}$ eV). The discrepancy at large pulse heights becomes apparent; the simulations do not include EASs with energies larger than $10^{16.5}$ eV.

(light blue), two electrons (green) and electron plus gamma (purple) are shown separately. Small pulses are predominantly generated by single gamma rays and single low energy electrons. The blue histogram is the sum of all these contributions and also includes higher multiplicity combinations; it matches the simulated pulse height distribution in **Fig. 29**. **Fig. 31** also demonstrates that the pulse height is a measure for the number of particles that traverses a detector.

8.2. EAS detection efficiency

The probability for detecting an EAS with a single two-detector station as function of distance between the shower core and station center and zenith angle is investigated with CORSIKA shower simulations. The station was exposed to EASs with an energy of 10^{15} eV. The direction

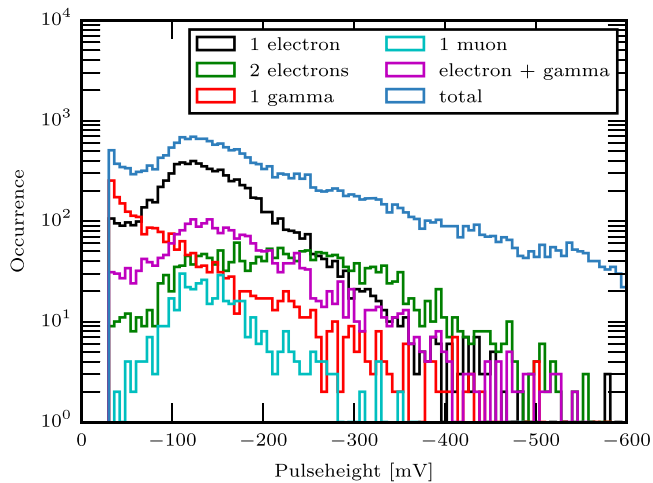


Fig. 31. Simulated response (blue histogram) of a detector in a two-detector station (color online). The distribution receives contributions from electrons (single electrons: black, 2 electrons: green), single muons (light blue), single gamma rays (red), single electron plus single gamma (purple) and higher multiplicities of various combinations (not shown separately). This simulation can be used to attribute a likelihood to finding the particle multiplicity in a detector at a specific pulse height. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of primary cosmic rays is chosen uniformly between 0° and 60° . The positions of the shower core with respect to the station center were homogeneously distributed within a circle with a radius of 150 m. Note that for increasing zenith angle the size of the shower footprint augments, while at the same time the chance of particle absorption in the atmosphere significantly increases. Both effects result in a lower particle density in the footprint. Also, the number of EASs that arrive at large core distances is larger than at small core distances because of the homogeneous exposure and the increasing effective area at larger (ring shaped) core distance bins. The same goes for the zenith angle. The solid angle Ω of the circular field of view subtended by a rotated zenith angle θ is given by $\Omega = 2\pi(1 - \cos(\theta))$. This implies that larger zenith angles result in larger solid angles.

Fig. 32 shows the EAS detection efficiency as function of core distance for zenith angles of 7.5° (blue dots), 30° (blue crosses) and 45° (blue stars). At 7.5° and for small core distances the efficiency is close to 100%. With increasing core distance the EAS detection efficiency rapidly decreases. **Fig. 33** shows the detection efficiency as function of zenith angle for core distances of 10 (blue dots), 25 (blue crosses) and 50 m (blue stars). If the shower core is close to the center of the station (e.g. 10 m in the figure) the EAS detection efficiency remains close to 100% for even relatively large zenith angles. If the shower core is further away from the station center the efficiency becomes much lower.

A 2D parametrization (combining the fits in **Figs. 32** and **33**) is derived that describes the detection efficiency as a function of the distance between station and shower core, and zenith angle. For small core distances, because of the high number density near the shower core, the EAS detection efficiency as function of core distance is expected to be 100%. At a distance r_m , at which the probability for EAS particles to miss a detector becomes substantial (e.g. for the 7.5° EAS detection efficiency in **Fig. 32** the value of r_m is ~ 20 m), the efficiency decreases. For core distances larger than r_m the radial dependency of the EAS detection efficiency can accurately be described by the formula of an exponentially modified Gaussian distribution. The zenith angle dependence is obtained by shifting r_m to smaller and eventually negative values depending on the zenith angle. This leads to the following parametrization:

$$p(r, \theta) = \begin{cases} f(r_m, \alpha, \mu(\theta, \chi, \rho), \sigma, \lambda) & \text{for } r < r_m \\ f(r, \alpha, \mu(\theta, \chi, \rho), \sigma, \lambda) & \text{for } r \geq r_m \end{cases} \quad (10)$$

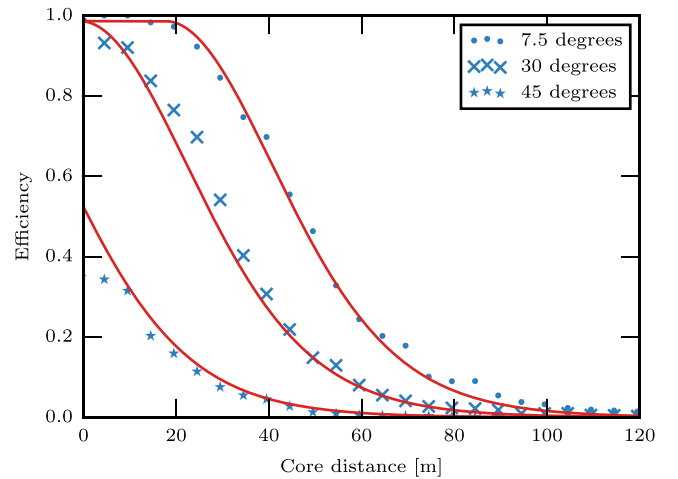


Fig. 32. EAS detection efficiency of 10^{15} eV showers as function of distance between the shower core and station center for zenith angles of 7.5° (dots), 30° (crosses) and 45° (stars). At 7.5° and for small core distances the efficiency is close to 1. For larger core distances the EAS detection efficiency decreases rapidly. The lines display the parametrization in Eq. (10).

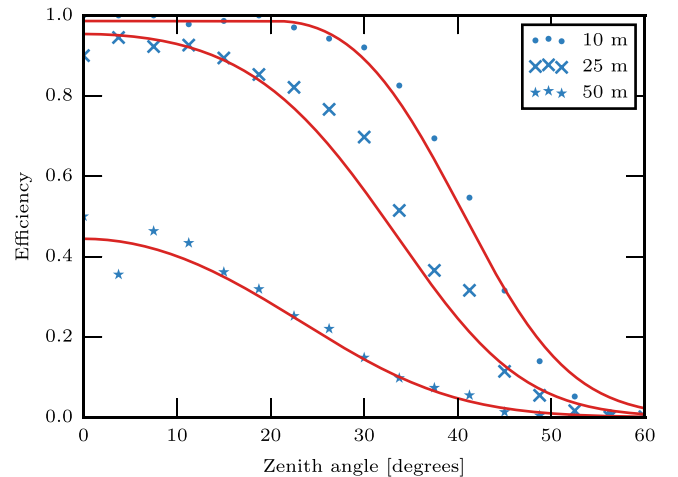


Fig. 33. EAS detection efficiency of 10^{15} eV showers as function of zenith angle for core distances of 10 (dots), 25 (crosses) and 50 m (stars). If the shower core is close to the center of the station (e.g. 10 m in the figure) the EAS detection efficiency stays close to 1 up to relatively large zenith angles. The lines display the parametrization in Eq. (10).

with the modified Gaussian distribution $f(r, \alpha, \mu, \sigma, \lambda)$ given by:

$$f(r, \alpha, \mu, \sigma, \lambda) = \alpha \exp\left[\frac{\lambda}{2}(2\mu + \lambda\sigma^2 - 2r)\right] \times \text{erfc}\left(\frac{\mu + \lambda\sigma^2 - r}{\sqrt{2}\sigma}\right) \quad (11)$$

μ and σ are the mean and standard deviation of the Gaussian part of the distribution and λ is the rate of the exponential distribution. The $\text{erfc}(x)$ factor is the complementary error function and is given by:

$$\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-y^2} dy \quad (12)$$

The value of r_m is the mode of $f(r, \alpha, \mu, \sigma, \lambda)$ and depends on μ , σ and λ . The function $f(r)$ is thus continuous at r_m . The shift of r_m as function of the zenith angle depends on μ . The zenith angle dependency is thus absorbed in μ :

$$\mu(\theta, \chi, \rho) = (\chi + \rho) \exp[-(\sec \theta - 1)] - \chi \quad (13)$$

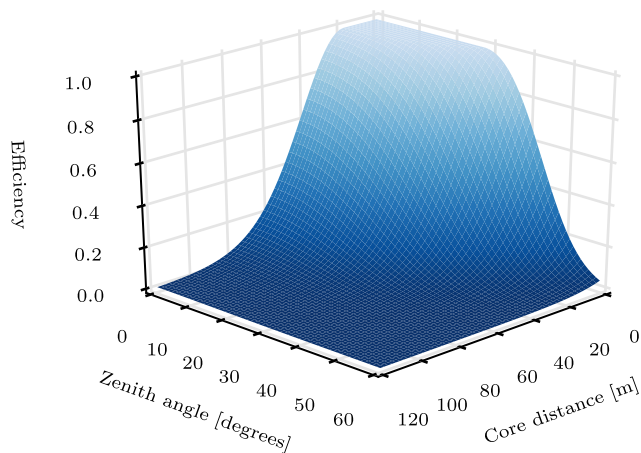


Fig. 34. A 3D display of the parametrized EAS detection efficiency of 10^{15} eV proton induced showers. For small zenith angles and core distances the efficiency is close to 1. For larger angles and distances the efficiency decreases but the surface area and solid angle increase substantially.

The exponential decrease of μ (and thus r_m) with the secant of the zenith angle can be thought of as an exponential diminishing of the shower particle density for larger travel distances through the atmosphere.

The lines in Figs. 32 and 33 show the parametrizations obtained by fitting α , σ , λ , ρ and χ to the simulated EAS detection efficiency. The obtained fit parameters are 2.15, 20.9, $7.22 \cdot 10^{-2}$, 7.84 and 129 resp. Fig. 34 shows a 3D plot of the parametrized EAS detection efficiency. With increasing EAS energy the ‘plateau’ at ~ 1 for small zenith angles and core distances increases. The shower detection efficiencies at other primary energies are discussed in [27].

9. Summary and conclusion

The High School Project on Astrophysics Research with Cosmics (HiSPARC) is a large air shower detector array with approximately 140 stations throughout the Netherlands, United Kingdom, Denmark and Namibia. HiSPARC is a collaboration of universities, scientific institutes and high schools each hosting one or more detection stations. The majority is managed by high schools.

The detection principle of HiSPARC is to sample EAS footprints using scintillation detectors. A HiSPARC detector consists of a scintillator glued to a light-guide which is connected to a PMT. The single muon detector response shows good agreement with simulations.

A detector is placed inside a roof box on the rooftop of a building. A detection station consists of two or four detectors. Custom made electronics read out and control two PMTs. For four-detector stations two units are used in a Master–Slave configuration. Multiple stations can be combined to sample an EAS footprint. To precisely correlate coincident samples accurate absolute timing is required. All stations are equipped with a GPS receiver that provides location and the timestamp for an event.

In 2008 a HiSPARC four-detector station was integrated in the KASCADE experiment. The shower direction reconstruction from a HiSPARC station (uncertainty $\sim 6^\circ$) was compared to KASCADE measurements. Similar results were obtained by comparing the direction reconstruction of four closely spaced four-detector station sampling the same shower. CORSIKA EAS simulations have been used to verify the uncertainty as well. Measurements and simulation yield comparable results. CORSIKA was also applied to determine the EAS detection efficiency of a station as function of distance to the shower core and zenith angle of the primary cosmic ray.

For small showers of which the core position falls within the area covered by the HiSPARC station, the energy can be estimated. Combining more stations at distances up to ~ 1 km, core reconstructions for larger showers improves considerably.

A subset of stations not too far from each other (distances smaller than the footprint size of the highest energy showers) forms a HiSPARC cluster. The Amsterdam Science Park Cluster is the densest cluster containing thirteen four-detector stations. Timing offsets between GPS antennas have been measured using EAS data. The field-of-view of the Science Park stations reaches beyond Polaris.

The large distance between stations in the network (up to 1000 km, excluding the station in Namibia) allows for searching coincident showers resulting from spallation of a primary cosmic ray (GZ-mechanism [54]). The HiSPARC array is also used to investigate the relation between cosmic rays and weather phenomena such as lightning [55].

HiSPARC is collecting data for already more than 15 years. A Dutch high school curriculum containing chapters on quantum mechanics, particle physics and special relativity in combination with a low cost (5000 €/10,000 €), robust HiSPARC station, easy accessible (pre-processed and raw) data and an extensive library with analyses tools, are the keys to success. HiSPARC brings science and research into the class room!

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

K. van Dam: Methodology, Validation, Formal analysis, Investigation, Writing - original draft, Visualization. **B. van Eijk:** Conceptualization, Methodology, Software, Resources, Writing - review & editing, Supervision, Project administration, Funding acquisition. **D.B.R.A. Fokkema:** Formal analysis, Investigation, Data curation, Software. **J.W. van Holten:** Writing - review & editing. **A.P.L.S. de Laet:** Formal analysis, Investigation, Data curation, Software. **N.G. Schultheiss:** Formal analysis, Investigation. **J.J.M. Steijger:** Writing - review & editing, Supervision. **J.C. Verkooijen:** Software.

Acknowledgments

The HiSPARC experiment could not have been realized without the interest of and contributions from the high schools, their physics teachers, technical assistants, ICT departments and, most importantly, their students. Schools managed to access public funds, but also received financial contributions from individuals and commercial businesses. The exhaustive list of persons, institutions and companies that (financially and/or in person) support the project is too long to give here.

In 2004 HiSPARC received the ‘Altran Foundation for Innovation’ award ‘Discovering, understanding and enjoying science through innovation’ [56,57]. We are indebted to Altran for their commitment and the exposure the award generated.

The Dutch National Institute for Subatomic Physics (Nikhef), The Netherlands has been indispensable for the development and realization of the project. The institute offered financial support, lab space, enabled the development of detector hardware and electronics, software and ICT infrastructure for data transfer and data storage. Without the generous help of Nikhef’s former director Frank Linde and Nikhef’s managing director Arjen van Rijn the realization of the project would not have been possible.

We are grateful that the KASCADE collaboration offered us the opportunity to install a test station at their site in Karlsruhe, and supplied us with the necessary infrastructure and KASCADE data.

References

- [1] T.K. Gaisser, R. Engel, E. Resconi, *Cosmic Rays and Particle Physics*, Cambridge University Press, 2016.
- [2] A.J. Hundhausen, *Coronal Expansion and Solar Wind*, Vol. 5, Springer Science & Business Media, 2012.
- [3] K.-L. Klein, S. Dalla, *Acceleration and propagation of solar energetic particles*, *Space Sci. Rev.* 212 (2017) 1107–1136.
- [4] P. Blasi, *The origin of galactic cosmic rays*, *Astron. Astrophys. Rev.* 21 (2013) 70.
- [5] A. Aab, P. Abreu, M. Aglietta, et al., *Observation of a large-scale anisotropy in the arrival directions of cosmic rays above 8×10^{18} eV*, *Science* 357 (2017) 1266–1270.
- [6] P.K.F. Grieder, *Extensive Air Showers: High Energy Phenomena and Astrophysical Aspects - A Tutorial, Reference Manual and Data Book*, Springer Science & Business Media, 2010.
- [7] T. Antoni, W.D. Apel, F. Badea, et al., *The cosmic-ray experiment KASCADE*, *Nucl. Instrum. Methods Phys. Res. A* 513 (2003) 490–510.
- [8] W.D. Apel, J.C. Arteaga, A.F. Badea, et al., *The KASCADE-grande experiment*, *Nucl. Instrum. Methods Phys. Res. A* 620 (2010) 202–216.
- [9] N. Chiba, K. Hashimoto, N. Hayashida, et al., *Akeno giant air shower array (AGASA) covering 100 km² area*, *Nucl. Instrum. Methods Phys. Res. A* 311 (1992) 338–349.
- [10] T. Abu-Zayyad, R. Aida, M. Allen, et al., *The surface detector array of the telescope array experiment*, *Nucl. Instrum. Methods Phys. Res. A* 689 (2012) 87–97.
- [11] Pierre Auger Collaboration, *The Pierre Auger cosmic ray observatory*, *Nucl. Instrum. Methods Phys. Res. A* 798 (2015) 172–213.
- [12] HiSPARC website, <https://www.hisparc.nl>.
- [13] D.B.R.A. Fokkema, *the HiSPARC Experiment* (Ph.D. thesis), University of Twente, ISBN: 978-90-365-3438-3, 2012, https://www.nikhef.nl/pub/services/biblio/theses_pdf/thesis_D.Fokkema.pdf.
- [14] Nikhef website, <https://www.nikhef.nl>.
- [15] T. Wada, N. Ochi, T. Kitamura, et al., *Observation of time correlation in cosmic air shower network*, *Nuclear Phys. B Proc. Suppl.* 75 (1999) 330–332.
- [16] P. Hofverberg, *Imaging the High Energy Cosmic Ray Sky* (Ph.D. thesis), KTH, 2006.
- [17] E. Brobeck, *Measurement of Ultra-High Energy Cosmic Rays with CHICOS* (Ph.D. thesis), California Institute of Technology, 2009.
- [18] M. Abbrescia, A. Agocs, S. Aiola, et al., *The EEE experiment project: status and first physics results*, *Eur. Phys. J. Plus* 128 (2013) 62.
- [19] C. Timmermans, J. Schotanus, B. van Eijk, et al., *Wetenschappelijk onderzoek voor scholieren*, *Ned. Tijdschr. Natuurkd.* 69 (2003) 230–234.
- [20] BC408 Data Sheet, Saint-Gobain Ceramics & Plastics, Inc., 2016.
- [21] 9107B series data sheet, ET Enterprises Ltd, 2010.
- [22] R6094 data sheet, Hamamatsu, 1996.
- [23] L. Landau, *On the energy loss of fast particles by ionization*, *J. Phys.* 8 (1944) 201–205.
- [24] HV3020 series data sheet, ET Enterprises Ltd, 2018.
- [25] P. Timmer, E. Heine, H. Peek, *Very low power, high voltage base for a photo multiplier tube for the KM3NeT deep sea neutrino telescope*, *J. Instrum.* 5 (2010) C12049.
- [26] J. Allison, K. Amako, J. Apostolakis, et al., *Recent developments in Geant4*, *Nucl. Instrum. Methods Phys. Res. A* 835 (2016) 186–225.
- [27] K. van Dam, et al., in preparation.
- [28] S. Jetter, D. Dwyer, J. Wen-Qi, et al., *PMT Waveform modeling at the Daya Bay experiment*, *Chin. Phys. C* 36 (2012) 733.
- [29] D. Reyna, *A simple parameterization of the cosmic-ray muon momentum spectra at the surface as a function of Zenith Angle*, 2006, arXiv preprint: [arXiv:hep-ph/0604145](https://arxiv.org/abs/0604145).
- [30] W.K. Hastings, *Monte Carlo Sampling methods using Markov chains and their applications*, *Biometrika* 57 (1970) 97–109.
- [31] D. Heck, J. Knapp, J.N. Capdevielle, et al., *CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers*, Forschungszentrum Karlsruhe GmbH, Karlsruhe (Germany), 1998.
- [32] Davis Vantage Pro, Davis Instruments, See <https://www.davisinstruments.com>.
- [33] T. Kokkeler, *Improvement of the Pulse Reconstruction and Arrival Direction Estimation at the HiSPARC Experiment* (Bachelor thesis), Twente University, 2019.
- [34] *Resolution T GPS embedded board user guide*, Trimble, 2009.
- [35] C. Elliott, V. Vijayakumar, W. Zink, et al., *National instruments LabVIEW: a programming environment for laboratory automation and measurement*, *JALA* 12 (2007) 17–24.
- [36] *MySQL Reference Manual*, Oracle Corporation, Available at <https://www.mysql.com>.
- [37] *Python Language Reference*, version 2.7, Python Software Foundation, Available at <https://www.python.org>.
- [38] *Nagios Core Documentation*, Nagios Enterprises, LLC, Available at <https://www.nagios.org>.
- [39] *OpenVPN Documentation*, OpenVPN Inc., Available at <https://www.openvpn.net>.
- [40] *TightVNC Documentation*, GlavSoft LLC, Available at <https://tightvnc.com>.
- [41] J.G. Wilson, *Progress in Cosmic Ray Physics*, Vol. 3, North-Holland Publishing Company, 1956.
- [42] N.G. Schultheiss, *The acceptance of the HiSPARC experiment*, 2016, arXiv preprint: [arXiv:1602.06799](https://arxiv.org/abs/1602.06799).
- [43] *The HDF5 library & file format*, The HDF Group, Available at <https://www.hdfgroup.org/solutions/hdf5>.
- [44] *Django Documentation*, Django Software Foundation, Available at <https://www.djangoproject.com>.
- [45] D.B.R.A. Fokkema, A.P.L.S. de Laat, T. Kooij, *SAPPHIRE*, 2012, <https://github.com/HiSPARC/sapphire>.
- [46] *jSparc website*, <https://docs.hisparc.nl/jsparc>.
- [47] *Jupyter Notebook Documentation*, Jupyter Team, Available at <https://www.jupyter.org>.
- [48] A.M. Hillas, in: *Proceedings International Cosmic Ray Conference*, vol. 1, 1981, pp. 193.
- [49] S. Ostapchenko, *QGSJET-II: towards reliable description of very high energy hadronic interactions*, *Nuclear Phys. B Proc. Suppl.* 151 (2006) 143–146.
- [50] H. Fesefeldt, *GHEISHA the Simulation of Hadronic Showers*, Tech. Rep., RWTH Aachen, 1985, PITHA-85/02.
- [51] W.R. Nelson, D.W.O. Rogers, H. Hirayama, *the EGS4 Code System*, Tech. rep., Stanford Linear Accelerator Center, Stanford, California, 1985.
- [52] J.M.C. Montanus, *The Observability of Jets in Cosmic Air Showers* (Ph.D. thesis), University of Amsterdam, ISBN: 978-94-028-0549-9, 2017, https://www.nikhef.nl/pub/services/biblio/theses_pdf/thesis_H_M.Montanus.pdf.
- [53] J.J.M. Steijger, *Het Kwartet*, 2019, Internal note.
- [54] N.M. Gerasimova, G.T. Zatspein, *Disintegration of cosmic ray nuclei by solar photons*, *J. Exp. Theor. Phys.* 38 (1960) 1245–1252.
- [55] R. Beekman, T. Aldus, J. Steijger, *Speelt kosotherse straling een rol bij initiatie bliksem?* *Ned. Tijdschr. Natuurkd.* 84 (2018) 18.
- [56] *Schools at 10^{20} eV and beyond*, *Nature*, 429 (2004) 685.
- [57] *2004 Altran award summary*, http://www.altran-foundation.org/fileadmin/medias/1.fondation/documents/Foundation_2004_Award.pdf.