

HiSPARC: big data voor scholieren

Het HiSPARC-experiment [1] is een samenwerking van universiteiten en middelbare scholen. Er worden deeltjes gemeten in kosmische lawines en er zijn inmiddels terabytes aan meetgegevens geproduceerd, dit is te beschouwen als big data. Voor het verwerken van deze grote hoeveelheid data zijn vaardigheden op het gebied van programmeren, software-engineering en data-analyse noodzakelijk. Dergelijke vaardigheden zijn ook buiten de natuurkunde steeds meer van belang. HiSPARC biedt een manier om middelbare scholieren en studenten kennis te laten maken met het verwerken van big data in een onderzoeksomgeving. Hiervoor zijn Python Notebooks [2] ontwikkeld. Deze notebooks zijn aanpasbare programmeervoorbeelden waarmee de scholieren data kunnen analyseren en leren programmeren. Tom Kooij, Niek Schultheiss, Kasper van Dam en Bob van Eijk

404

Het HiSPARC-experiment

Het HiSPARC-experiment (HiSPARC staat voor High School Project on Astrophysics Research with Cosmics) bestaat uit een uitgebreid netwerk van ongeveer 140 meetstations voor kosmische straling. Deze stations worden beheerd door universiteiten en middelbare scholen in Nederland, Denemarken en het Verenigd Koninkrijk [3]. De geometrie van het meetnetwerk wordt bepaald door de locatie van samenwerkende instellingen en is hierdoor bijzonder onregelmatig.

Het cluster op het Amsterdam Science Park [4] bestaat uit een tiental stations in een klein gebied. Elk station beschikt over vier scintillatoren. De gedetecteerde deeltjes – gamma's, elektronen en muonen – zijn onderdeel van een deeltjeslawine die door de botsing van een primair kosmisch deeltje hoog in de dampkring wordt veroorzaakt. Het aantal deeltjes in de lawine hangt af van het

soort primair deeltje – foton, proton of een samengestelde atoomkern – en zijn energie. De richting van de lawine is te bepalen aan de hand van de aankomsttijden van de deeltjes op het aardoppervlak.

Naast het cluster op het Amsterdam Science Park beschikt het netwerk over clusters op onderlinge afstanden tot maximaal duizend kilometer. Hiermee zijn de volgende onderzoeksvragen te formuleren:

- Hangt de hoeveelheid gemeten kosmische straling af van de plaats op Aarde en het tijdstip van de meting? Wat is bijvoorbeeld de invloed van het Aardmagnetisch veld?
- Is de verdeling van richtingen van de kosmische straling uniform? Komt er bijvoorbeeld meer straling uit het centrum van de Melkweg?
- Is er een correlatie tussen de gedetecteerde straling van de verschillende stations?
- Heeft de lokale toestand van de dampkring invloed op de

metingen?

De ruwe data worden gecontroleerd en verwerkt en zijn daarna via internet vrij beschikbaar voor iedereen. HiSPARC reconstrueert deeltjeslawines met geometrische informatie en meetgegevens. Hiervoor wordt de programmeertaal Python gebruikt [5]. De eenvoudige syntax van Python en de uitgebreide bibliotheek met functies vereenvoudigen het leren van deze taal. Bovendien is de taal geschikt om grote hoeveelheden data snel te verwerken en te analyseren.

Richtingsreconstructie

De richting van een deeltjeslawine, en daarmee het initiërende deeltje, is door triangulatie te reconstrueren als deeltjes bijna gelijktijdig in minimaal drie van de vier scintillatoren van een station worden gedetecteerd (figuur 2). De verschillen in aankomsttijden in de detectoren (gemeten met een nauwkeurigheid van enkele nanoseconden) worden gebruikt

om de richting van de deeltjeslawine te berekenen. Als alle aankomsttijden exact gelijk zijn, is er sprake van een verticale shower. De snelheden van de hoogenergetische deeltjes in de lawine zijn allemaal dicht bij de lichtsnelheid, en dus min of meer gelijk. Een verschil in aankomsttijd kan zo worden uitgedrukt in een afstand. Met drie detectoren en dus drie lengte-verschillen is het mogelijk de richting van de shower te bepalen. In plaats van scintillatoren in één enkel station kunnen ook meerdere stations, met een grotere onderlinge afstand, worden gebruikt voor een nauwkeurigere richtingsreconstructie van deeltjeslawines. De benodigde grotere voetafdruk van de lawine vereist een hogere energie.

Met een Python Notebook zijn histogrammen te maken van de richtingen van showers gebaseerd op de metingen met meerdere detectoren. In figuur 3 is links het histogram voor de zenithoek op de plaats van het station gegeven, dit is de hoek die de shower met de verticaal maakt. Rechts staat het azimut op de plaats van het station, de hoek met de oostelijke richting. Voor lage energieën heeft de verdeling van geladen deeltjes een oost-westafwijking door het magnetisch veld van de Aarde.

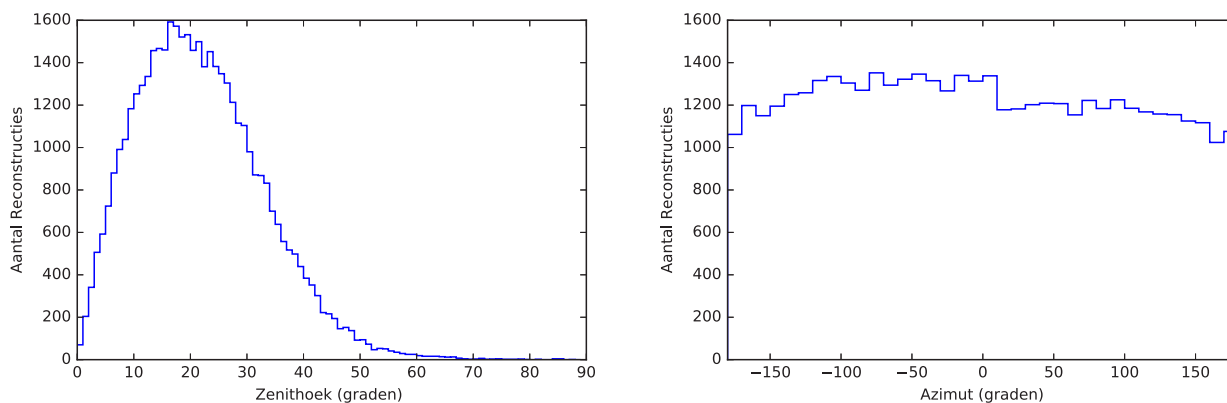
In het linker histogram van figuur 3 is te zien dat de verdeling van de zenithoek maximaal is rond twintig graden. Dit komt enerzijds doordat een HiSPARC-station minder gevoelig is voor grote zenithoeken en omdat niet-loodrecht invallende showers door meer atmosfeer moeten reizen waardoor minder deeltjes de grond bereiken. Anderzijds nadert het hemeloppervlak recht boven het station



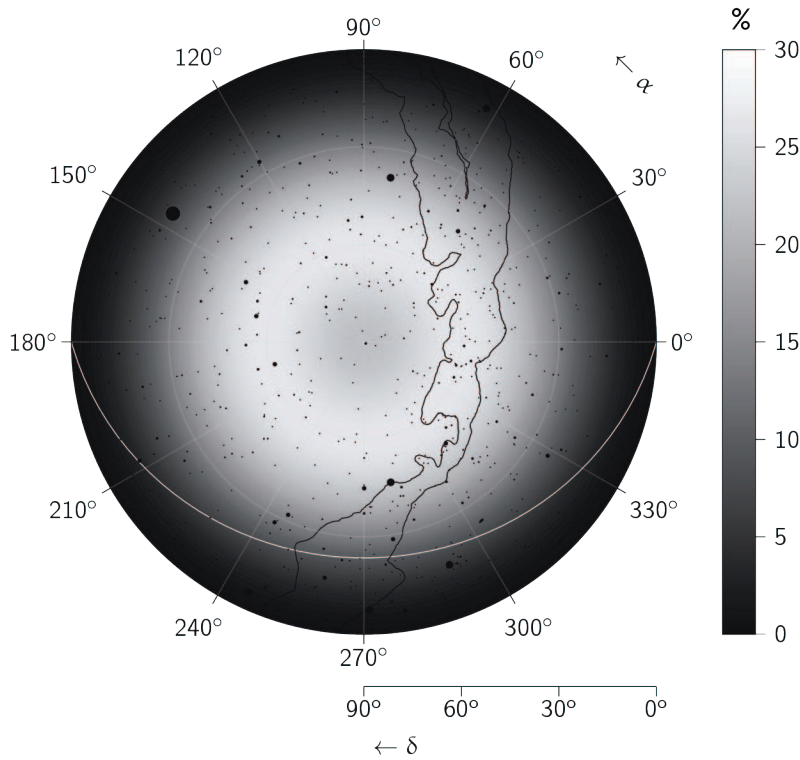
Figuur 1 Kaart met de stations van het HiSPARC-experiment. De getoonde stations bestaan uit twee of vier scintillatoren. Elk station bezit hierdoor een scintillatieoppervlak van 1 of 2 m². Op dit moment beschikt HiSPARC over ongeveer honderdveertig stations die gemiddeld iedere drie seconden minstens twee deeltjes van een kosmische lawine meten. Sinds de start van het experiment zijn inmiddels 6×10^9 metingen verzameld. De kaart is met een aangepast Python Notebook gegenereerd.



Figuur 2 Een skibox met daarin een scintillatiedetector van 100 cm x 50 cm met een dikte van 2 cm. De detector geeft een lichtpuls als een gamma-, elektrono of muondeeltje door de scintillator schiet. De hoeveelheid licht wordt als functie van de tijd gemeten.



Figuur 3 Verdeling van de gereconstrueerde zenithoek en azimut van twee maanden meetgegevens van drie stations op het Amsterdam Science Park. De gereconstrueerde zenithoek voor HiSPARC-stations heeft een maximum bij twintig graden. De verdeling van het azimut hangt af van de geometrie van de opstelling en heeft in het ideale geval een vlakke verdeling.



Figuur 4 Door de aardrotatie neemt het HiSPARC-experiment deeltjes uit een wisselend deel rond het zenit waar. De acceptantie van de detector is hierdoor gemiddeld maximaal 30%. In 70% van de gevallen komen de deeltjes niet uit het waarneembare gebied. De kans op detectie neemt af voor grotere of kleinere declinaties. Voor het Amsterdam Science Park varieert dit van 0 (zwart) tot 30% (wit). In het midden van het diagram is de Poolster met een declinatie van 90° te zien. Daaromheen zijn de andere sterren als zwarte punten afgebeeld. De witte kromme geeft de baan van de zon weer. De contouren van de Melkweg zijn in zwart aangegeven. [7]

406



Van links naar rechts zijn afgebeeld: Kasper van Dam (promovendus Nikhef), Bob van Eijk (professor hoge-energiefysica Nikhef/Universiteit Twente), Tom Kooij (onderzoeker Nikhef, docent Coornhert Gymnasium) en Niek Schultheiss (promovendus Nikhef en docent Zaanlands Lyceum). De laatste twee worden financieel ondersteund door de landelijke SNS-gelden voor outreachactiviteiten.

kaspervd@nikhef.nl

naar nul voor kleine zenithoeken. In het rechter histogram is de genoeg vlakke azimuthverdeling te zien, dit is een sterke indicatie voor een isotrope verdeling. De verdeling hangt ook af van de geometrie van de drie stations, in dit geval een gelijkzijdige driehoek. Om een idee te krijgen van de her-

komst van de kosmische straling is het inzichtelijk om de gereconstrueerde richtingen weer te geven in een equatoriaal coördinatensysteem. Hierin is de oorsprong het centrum van de aarde, de z -as de richting van de Noordpool en het xy -vlak het vlak van de evenaar. Coördinaten, of liever richtingen, worden aangeduid met

een rechte klimming en declinatie. De rechte klimming is de hoek ten opzichte van een gedefinieerd punt in het equatoriaalvlak (de vernal equinox om precies te zijn), de declinatie is de hoek loodrecht op dit vlak. De Poolster heeft dus een declinatie van bijna negentig graden (89,26°). Om de met de Aarde meedraaiende azimuth en zenithhoek om te rekenen naar de ten opzichte van de sterren vaste rechte klimming en declinatie, is de locatie van het HiSPARC-station op aarde en het tijdstip van de shower nodig. Deze complexe berekening kan stap voor stap overzichtelijk worden uitgevoerd met een computerprogramma. Met het equatoriaal coördinatensysteem kan de richting van de deeltjeslawine van meerdere kosmische deeltjes overzichtelijk weergegeven worden (figuur 4). Deze weergave is ook gebruikt om te onderzoeken of er een 'hotspot' aan de hemel bestaat van waaruit meer hoogenergetische deeltjes worden uitgezonden, dit is beschreven in een eerdere uitgave van dit tijdschrift [6].

HiSPARC-data verwerken met Python Notebooks

Python Notebooks zijn documenten waarin programmacode van Python gecombineerd wordt met tekst, formules, illustraties et cetera. Door deze elementen te combineren ontstaat een leesbaar document waarin de volledige workflow van data-analyse wordt gepresenteerd en waarmee resultaten kunnen worden gereproduceerd, zonder dat andere bronnen nodig zijn.

Voor het maken en gebruiken van een notebook is een computer met Python en Jupyter software [5,2] nodig. Notebooks werken via een (lokale) server [8] waarop berekeningen worden uitgevoerd. De server kan op iedere computer, maar ook op een centrale locatie, worden geïnstalleerd. Voor het uitvoeren van HiSPARC data-analyses is dan alleen een webbrowser nodig. Op dit moment wordt onderzocht of het mogelijk is om een centrale server aan te bieden.

Figuur 5 laat een deel van een notebook [9] zien waar de richting van kosmische deeltjes wordt weergegeven op een hemelkaart. Ieder blauw kruisje geeft de gereconstrueerde richting van een kosmisch deeltje aan. Hiervoor is een dataset van een

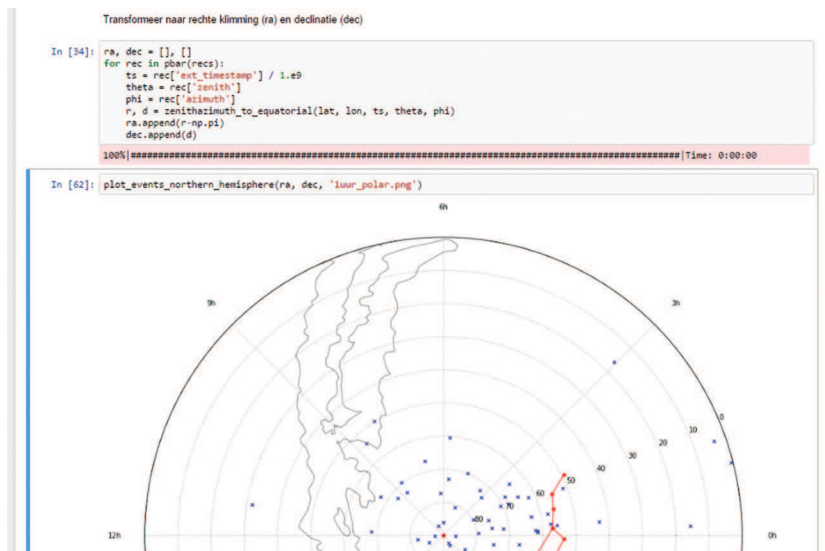
enkel station gebruikt, bijvoorbeeld afkomstig van de detector van de eigen school. Omdat hier bewust slechts één uur data is gebruikt, wordt maar aan een deel van de hemel gemeten (rond het zenit). Met dit notebook kan de leerling de richting van kosmische deeltjes van dit station of een ander station reconstitueren. Zowel de richtingsreconstructie als transformatie naar het equatoriale coördinatensysteem worden door de software automatisch uitgevoerd. Deze bewerkingen worden bewust als ‘zwarte doos’ geïntroduceerd, omdat interpreteren van de resultaten het doel is van dit notebook. De coördinatentransformatie kan echter ook een leerdoel zijn. In dit geval is het dan een programmeeropdracht om zenithoek, azimut en tijdstip om te rekenen naar rechte klimming en declinatie.

Het notebook uit figuur 5 is door iedereen aan te passen. In figuur 6 is een tweede diagram te zien dat de hele hemelbol weergeeft. De meetgegevens zijn in dit geval afkomstig van zeven stations op het Amsterdam Science Park. Door het aanpassen van hetzelfde notebook zijn verschillende diagrammen te maken.

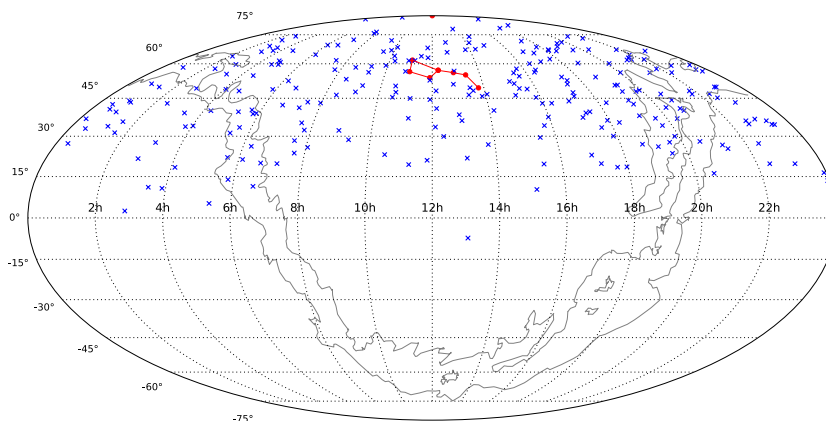
De flexibele opzet van de interactieve notebooks maakt verwerken van geometrische en meetgegevens mogelijk. Hiermee is bijvoorbeeld kosmische straling, gemeten door het gehele experiment – clusters en stations – te reconstrueren. HiSPARC combineert verschillende facetten van wetenschappelijk onderzoek. De combinatie van meetgegevens van een (eigen) meetstation met aangepaste notebooks maakt het gebruik van deze techniek in de les en voor profielwerkstukken interessant. Docenten in onderzoek bij HiSPARC maken al intensief gebruik van Python Notebooks [10]. Inmiddels maken ook de eerste leerlingen gebruik van notebooks voor hun profielwerkstuk.

Referenties

- 1 www.hisparc.nl.
- 2 jupyter.org.
- 3 <https://wikis.bris.ac.uk/display/HiSPARC/HiSPARC>.
- 4 http://data.hisparc.nl/show/stations_on_map/Netherlands/Amsterdam/Science%20Park.
- 5 www.python.org.
- 6 S.I. Beijen, *Een hotspot voor UHECR's*:



Figuur 5 Screenshot van een gedeelte van het in de tekst beschreven Python Notebook [9]. In dit voorbeeld staan (van boven naar beneden) tekst, een blok programmacode met resultaat (de rode statusbalk) en een tweede blok met het figuur dat door de code is gegenereerd. Dit figuur is een projectie van de noordelijke hemel. De blauwe kruisjes zijn aankomstrichtingen van deeltjeslwynes van een enkel HiSPARC-station binnen één uur. Ook de contouren van de Melkweg, de Poolster en het steelplan-asterisme van de Grote Beer zijn weergegeven.



Figuur 6 Een projectie van de gehele hemelbol. De blauwe kruisjes zijn aankomstrichtingen van deeltjeslwynes die door minstens zeven HiSPARC stations tegelijkertijd gemeten zijn. Als referentie zijn de contouren van de Melkweg, de Poolster en het steelplan-asterisme van de Grote Beer weergegeven.

zichtbaar met HiSPARC, NTvN **81-12**
(2015) 438 - 441.

- 7 N.G. Schultheiss, *The acceptance of the HiSPARC experiment*, arXiv:1602-06799.
- 8 http://docs.hisparc.nl/infopakket/pdf/notebook_installatie.pdf.
- 9 http://docs.hisparc.nl/infopakket/notebooks/10_sterrenkaart.ipynb.
- 10 <http://docs.hisparc.nl/infopakket>.