



## 1 Toegang tot HiSPARC gegevens

De data opslag van HiSPARC meetgegevens gebeurt op het Nikhef en bestaat uit een paar databases. Als eerst is er de ruwe dataopslag. Daarin worden metingen opgeslagen zodra ze naar het Nikhef worden gestuurd. Daarnaast is er nog een afgeleide database. In de afgeleide database zijn al een aantal analyses op de meetgegevens uitgevoerd. Met name de aankomst tijden van de shower in verschillende detectoren en het aantal deeltjes in een detector.

Toegang tot HiSPARC meetgegevens is vrij voor iedereen. Het downloaden van de ruwe data is niet heel eenvoudig, bovendien raden we dat ook niet meer aan nu de afgeleide database beschikbaar is. Data uit de afgeleide database wordt als tsv-bestand (tab gescheiden kolommen) aangeboden voor download via deze website: <http://data.hisparc.nl/data/download>.

Zodra data gedownload is kan deze bijvoorbeeld in Excel geïmporteerd worden om grafieken te maken en analyses uit te voeren. Omdat Excel niet altijd even eenvoudig werkt hebben we zelf een programma gemaakt dat in webbrowsers werkt. Het download data en kan daar direct grafieken mee maken.

## 2 jSparc bibliotheek

jSparc is de JavaScript bibliotheek die het makkelijker maakt om met HiSPARC data te werken. Zo biedt de bibliotheek een eenvoudige functie om gegevens op te halen van de HiSPARC server en deze gelijk om te zetten tot een formaat dat begrijpelijk is voor JavaScript. De broncode is hier te vinden: <https://www.github.com/hisparc/jsparc/>.

## 3 Data retrieval

Dit is een beschrijving van de pagina te bereiken via [http://data.hisparc.nl/media/jsparc/data\\_retrieval.html](http://data.hisparc.nl/media/jsparc/data_retrieval.html). Hiermee kan data opgehaald en bestudeerd worden. Aan het HiSPARC logo (rechts boven) is te herkennen of de pagina data van de HiSPARC server aan het ophalen is, dan is het logo namelijk geanimeerd. Als eerst haalt de pagina een up-to-date lijst van HiSPARC stations op, dit gaat zo snel dat het logo maar heel kort geanimeerd is. Onder het logo is een knop om bij de documentatie van de pagina te komen: <http://doc.hisparc.nl/jsparc/>.

### 3.1 Downloaden van data

Het Download data formulier, zie Figuur 3.1, staat toe een HiSPARC station, een start- en einddatum, en het data type te selecteren. Door te drukken op *Get Data!* wordt de data dan gedownload. Zodra de nieuwe data is geladen, verschijnt er een nieuwe sectie op de pagina die een overzicht weergeeft van alle datasets die geladen zijn, zie Figuur 3.2. Het is mogelijk meerdere datasets te

### Download data

Get data from the HiSPARC server.

Data type: ☒ Events ☐ Weather

Station:

Start date:

End date:

### Load local file

Import a downloaded .csv file (tab-separated values).

**Figuur 3.1** – Gedeelte van de website waar data mee gedownload of ingeladen kan worden.

Data type: ☒ Events ☐ Weather

### Select datasets to use

Choice 1	Choice 2	Station	Type	Start date	End date	Entries	Preview	Download	Remove
<input type="radio"/>	<input type="radio"/>	501	barometer	2012-06-27 00:00	1 day later	25980	<a href="#">show</a>	<a href="#">get csv</a>	<a href="#">x</a>
<input type="radio"/>	<input type="radio"/>	2010	events	2013-12-10 00:00	2013-12-11 00:00	24097	<a href="#">show</a>	<a href="#">get csv</a>	<a href="#">x</a>

**Figuur 3.2** – Overzicht van de ingeladen datasets.

laden door het Download data formulier opnieuw te gebruiken met andere instellingen. Met het rechter formulier; *Load local file* kan eigen of eerder gedownloade .tsv bestanden (tab gescheiden) inladen. Deze verschijnen dan ook in het overzicht.

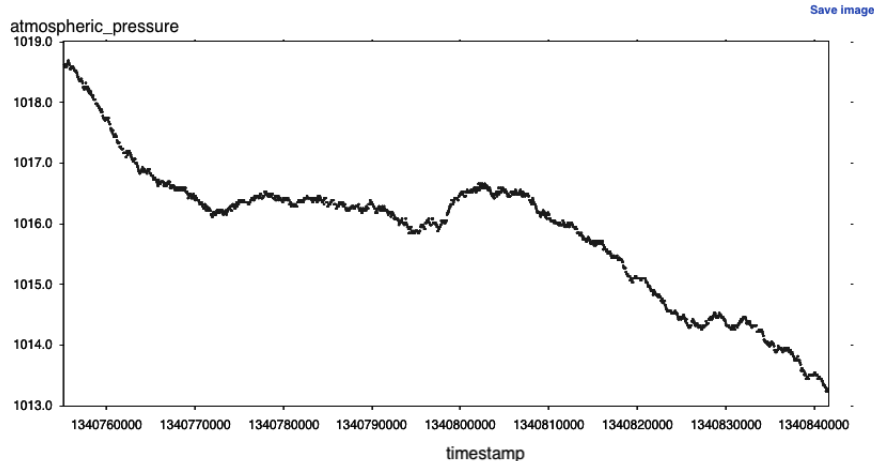
Met de geladen datasets kunnen verschillende acties uitgevoerd worden. Door in de *Choice* kolommen een dataset te kiezen verschijnen de plot opties en een lijst van de variabelen in die dataset. In sectie 3.2 is uitgelegd wat alle opties betekenen. Daarnaast kan er gekozen worden om de waarden in de dataset te bekijken door op *show* te klikken in de *Preview* kolom. Met de *tsv* link onder *Download* wordt de data als .tsv bestand gedownload. Met de *x* in de *Remove* kolom kan de dataset uit het geheugen van de browser gewist worden.

## 3.2 Grafieken maken

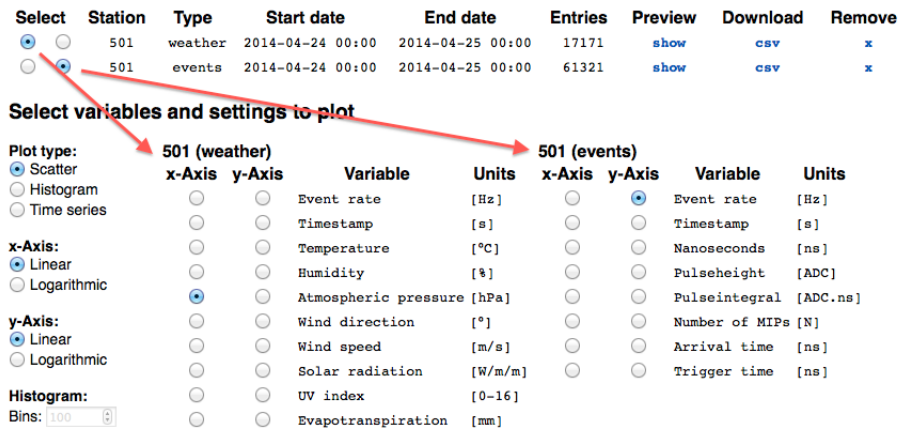
Zodra een dataset gekozen is kunnen er plots (grafieken) mee gemaakt worden. Er zijn 3 opties voor soorten grafieken. Bij het type *scatter* worden er twee variabelen voor iedere meting tegen elkaar uitgezet, een op de x-as en de ander op de y-as. Bij het type *histogram* kan maar één variabele gekozen worden (x-as). Bij een histogram wordt het hele bereik tussen de minimale en maximale waarde van die variabele opgedeeld in een aantal (normaal 100) evengrote delen. Dan wordt voor elke individuele waarde gekeken in welk bereik die past, het aantal waarden in een bereik komt dan op de y-as te staan. Als laatste is er nog de *time series*. Hierbij is de x-as de tijd (en datum) en kan een variabele voor de y-as gekozen worden.

Als de keuzes gemaakt zijn kan de plot gemaakt worden door op *Create Plot* te drukken, zie Figuur 3.3. Als de opties aangepast worden kan een nieuwe plot gemaakt worden door weer op die knop te drukken.

## Plot



**Figuur 3.3** – Voorbeeld van een plot. Hier is de luchtdruk op de y-as uitgezet tegen de tijdstempel op de x-as.



**Figuur 3.4** – Meerdere datasets kiezen uit het overzicht.

In sommige gevallen zullen meerdere kleuren te zien zijn in de grafiek, dat komt doordat bepaalde variabelen door meerdere sensoren of detectoren gemeten worden. Met kleur codes wordt dan aangegeven om welke detector het gaat. In het geval van HiSPARC metingen is zwart detector 1, rood detector 2, groen detector 3 en blauw detector 4. Bij weermetingen met meerdere sensoren is zwart de binnen sensor en rood de buiten sensor.

### 3.3 Interpolatie

Als meerdere datasets zijn opgehaald met overlappende tijdperiodes is het mogelijk om deze tegen elkaar te plotten. Kies hiervoor een dataset in kolom *Choice 1* en de ander in *Choice 2*. De variabelen verschijnen dan naast elkaar, zie Figuur 3.4. Nu kan uit beide een keuze gemaakt kan worden voor de x- en y-assen. De interpolatie vindt plaats op basis van de tijdstempels van de gegevens.

Raw data in the dataset

#	date	time	timestamp	nanoseconds	pulseheights	integral	number_of_mips
1	2013-12-10	00:00:04	1386633604	316356284	267 142 -1 -1	2574 2171 -1 -1	1.0548 -999 -1
2	2013-12-10	00:00:04	1386633604	514687429	66 764 -1 -1	383 6941 -1 -1	0.1569 -999 -1
3	2013-12-10	00:00:10	1386633610	2041815	310 67 -1 -1	3316 419 -1 -1	1.3588 -999 -1
4	2013-12-10	00:00:11	1386633611	34179469	416 1104 -1 -1	6315 8643 -1 -1	2.5877 -999 -1
5	2013-12-10	00:00:11	1386633611	140911485	341 203 -1 -1	3755 2707 -1 -1	1.5387 -999 -1
6	2013-12-10	00:00:13	1386633613	835327243	185 324 -1 -1	2751 3776 -1 -1	1.1273 -999 -1
7	2013-12-10	00:00:16	1386633616	197540866	226 353 -1 -1	2859 1970 -1 -1	1.1715 -999 -1
8	2013-12-10	00:00:17	1386633617	572549186	80 293 -1 -1	634 3211 -1 -1	0.2598 -999 -1
9	2013-12-10	00:00:19	1386633619	325857969	90 213 -1 -1	708 4357 -1 -1	0.2901 -999 -1
10	2013-12-10	00:00:24	1386633624	476804673	236 241 -1 -1	2117 2764 -1 -1	0.8675 -999 -1
11	2013-12-10	00:00:25	1386633625	48580650	60 925 -1 -1	292 8169 -1 -1	0.1197 -999 -1
12	2013-12-10	00:00:30	1386633630	411534830	64 289 -1 -1	300 3687 -1 -1	0.1229 -999 -1
13	2013-12-10	00:00:33	1386633633	258047273	432 621 -1 -1	5754 8412 -1 -1	2.3579 -999 -1
14	2013-12-10	00:00:45	1386633645	453753695	298 117 -1 -1	3076 1119 -1 -1	1.2605 -999 -1
... truncated table (click to show more)							
24083	2013-12-10	23:59:11	1386719951	549731025	254 75 -1 -1	3007 930 -1 -1	1.2322 -999 -1
24084	2013-12-10	23:59:14	1386719954	639980112	778 109 -1 -1	9546 1727 -1 -1	3.9117 -999 -1
24085	2013-12-10	23:59:20	1386719960	930306758	128 79 -1 -1	939 1109 -1 -1	0.3848 -999 -1
24086	2013-12-10	23:59:22	1386719962	330236234	2026 1107 -1 -1	39474 20515 -1 -1	16.1755 -999 -1
24087	2013-12-10	23:59:22	1386719962	576740175	237 137 -1 -1	2132 1845 -1 -1	0.8736 -999 -1

Figuur 3.5 – Overzicht van de meetwaarden in een dataset.

### 3.4 Gegevens inkijken

Door op de knop *show* te drukken in het overzicht van de datasets verschijnt een tabel met op elke rij in kolommen verdeelt de waardes van een meting, zie Figuur 3.5 voor een voorbeeld. Niet alle rijen worden direct getoond, het kost de browser namelijk te veel tijd om duizenden regels te tonen, dus eerst worden er zo'n 30 getoond, door op de middelste regel te klikken ('click to show more') kan er meer ingekeken worden. Hier is ook duidelijk te zien dat sommige kolommen meerdere waarden hebben, omdat die waarden door meerdere sensoren of detectoren gemeten zijn. Voor deze waarden gelden dezelfde kleur aanduidingen als bij de grafieken.

### 3.5 Traces bekijken

Van iedere detectie van een air shower zijn de signalen uit de PMTs te bekijken. Gebruik eerst de knop om de gegevens in te kijken zoals beschreven in sectie 3.4. Ga dan naar de kolom met de kop *Traces*, druk daar op de *show* knop van een event. Dan zullen de PMT signalen worden opgehaald van de server en worden getoond als grafiek, zie Figuur 3.6.

### 3.6 Fitten van data

Bij de plot opties kan ook gekozen worden om een fit te maken van de data geplot gaat worden. Als variabelen geselecteerd zijn, klik dan onder *Fit* om het type functie kiezen. Als functie-keuze verschijnen dan *No fit*; *Linear*; *Exponential*; *Logarithmic*; *Power*; *Polynomial*. Bij de optie *Polynomial* kan de macht van de polynoom reeks aangeven worden bij *Degree*. Door nu op *Create Plot* te klikken, wordt de plot van de variabelen gemaakt inclusief een fit van de datapunten. Onder het plotje verschijnt de vergelijking van de fit curve. Zie Figuur 3.7.

## Plot



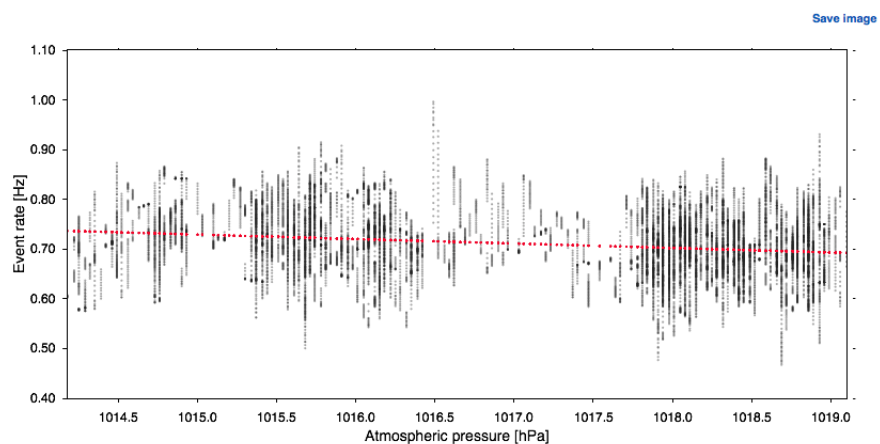
## Raw data in the dataset

pulseheights	integral	number_of_mips	arrival_times	trace
16 9 2 438	4466 0 0 6095	1.4755 0 0 1.7169	12.5 -999 -999 30	show
2 259 304 2	0 2749 2987 0	0 0.8001 0.8782 0	-999 12.5 12.5 -999	show

**Figuur 3.6** – De traces van een meting. De met rood gearceerde kolom bevat de knoppen om de traces voor een event op te halen.

**Histogram:** ☐ ☐ UV index [0-16]  
 Bins:  ☐ Evapotranspiration [mm]  
**Fit:** ☐ ☐ Rain rate [mm/h]  
 Type:  ☐ Heat index [°C]  
 Degree:  ☐ Dew point [°C]  
☐ Wind chill [°C]

## Plot



Fit formula:

Functie van de fit

$$y = -9.00 \cdot 10^{-3}x + 9.86 \cdot 10^0$$

**Figuur 3.7** – Instellen van een fit van twee variabelen. Hier bijvoorbeeld event rate tegen luchtdruk