Starting Assignments

K. Schadenberg, G. Osinga & J.J. Velthuis

1 Introduction

Before we can do some serious data analysis with the HiSPARC data, we first need to know how we can access the data. In this assignment we will be using the data available via the HiSPARC website `http://www.hisparc.nl/`. There is more data available via the HiSPARC database but we leave this for a later assignment.

2 Downloading data

Obtaining data from a single HiSPARC detector is very easy. All data is available via `http://data.hisparc.nl/` . This text explains how to import the data into a spreadsheet application and perform a number of different analyses.

The HiSPARC data site shows all registered HiSPARC detectors, active detectors are shown as links. Clicking on a link shows the last day with available data from that station. In figure 2.1 the data from detector 13001, placed at the University of Bristol, from the 17th of June 2012 is shown as an example.

This web page of figure 2.1 shows the measurement data on the left and the detector settings on the right. Lets start at the top right of the page. The first thing we see is a small calendar. With this calendar we can browse to a specific date and look at the results from that day.[1] The link 'List' brings us back to the overview of detector stations.

Below the calendar is the position of the detector with a link to 'Openstreetmap' to see this location on a map. Note that this location is not always accurate because it is only updated very rarely.[2]

Below the location some hardware parameters are listed. Station 13001 has four detector plates and thus uses two electronics boxes, one master and one slave. Here you can see the hardware number, its firmware version, and the voltage supplied to the different photomultiplier tubes.

On the left hand side of the page the actual data is shown. There are three graphs:

Event Histogram shows the number of registered coincidences per hour.

Pulse height histogram shows the height of the registered coincidences.

[1] See how the address of the web page changes when you navigate to a different date. This might come in handy when you want to download large amounts of data for future assignments, or automate this process by writing your own software scripts.

[2] Detectors usually stay put so the long interval between updates is no problem. However, if the GPS antenna was not positioned or calibrated correctly it can take some time before the incorrect position is corrected on the website.

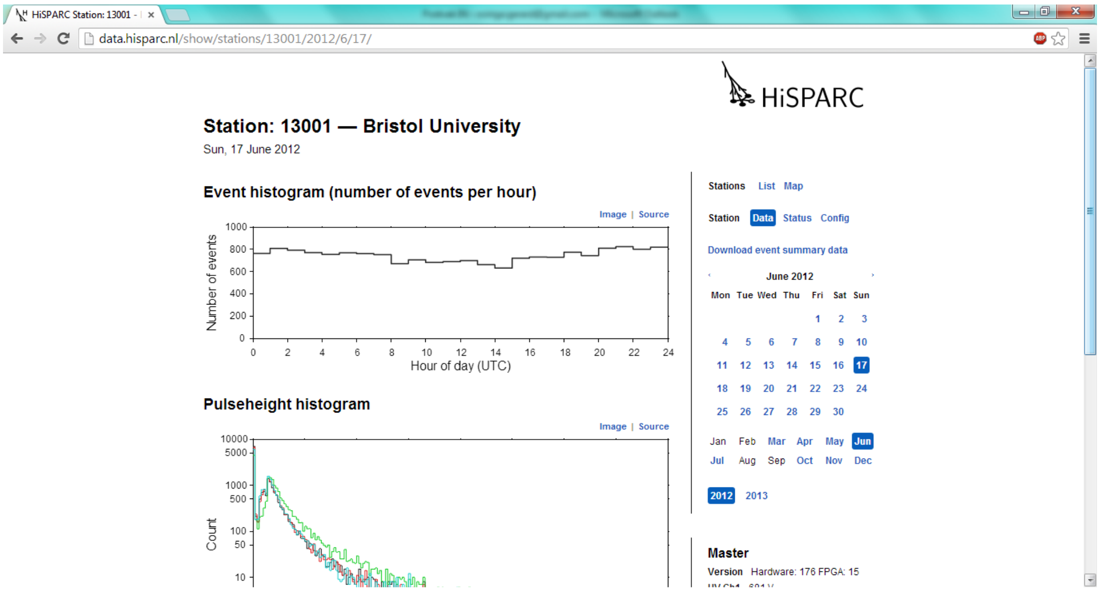

Figure 2.1: Information available via `http://data.hisparc.nl/` .

Pulse integral histogram shows the integral of the registered coincidences.

The data that is used to generate these graphs can be directly downloaded from this site. In the top right corner of every graph there is a link to the source data 'Source'. The data is saved in a comma separated value file (.csv). The contents of the file should look something like figure 2.2.[1] The top of the file gives a short description of the file contents. In the example of figure 2.2 the station registered 792 coincidences between 2 a.m. and 3 a.m.

3 Importing data

The process of importing data is described for two different spreadsheet applications; Microsoft Excel[2] and LibreOffice Calc[3].

3.1 Microsoft Excel

If the .csv file is correctly recognised by Windows the .csv file can be directly opened by Excel and the the bin number and measurement value will have been placed in separate columns. If this is not the case then the .csv data needs to be manually imported into Excel. To do this follow these steps:

1. Start Excel.

2. Select 'Data' in the Ribbon (top menu bar).

[1]Windows systems do not always correctly recognise the file settings of the .csv file. If this is the case you should use the second import method for Excel described in this module.

[2]This manual is written for a 'recent' version of Microsoft office using the Ribbon Interface.

[3]At the moment there are only small differences between OpenOffice and LibreOffice so this manual can be used for both.

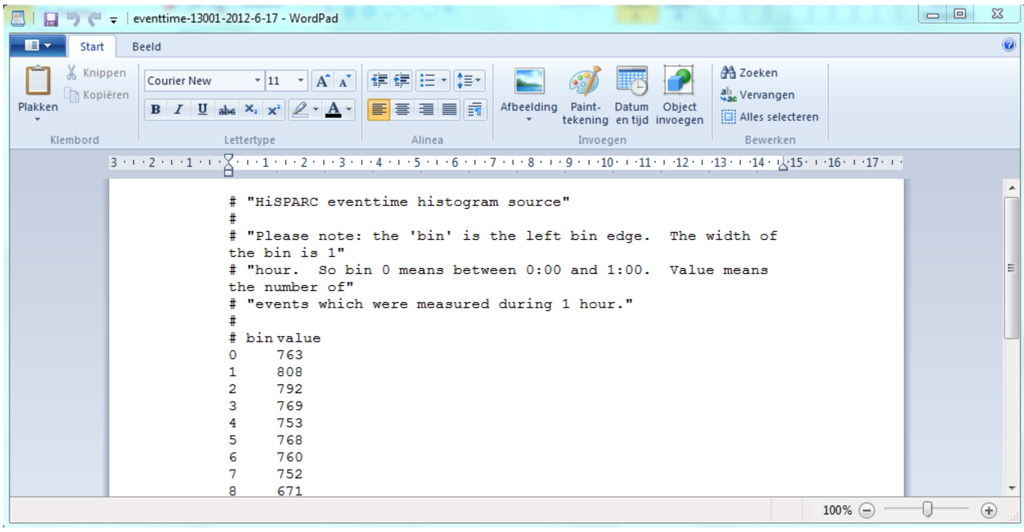

Figure 2.2: Part of the .csv file.

3. Click on 'From text' and select the file you want to import.
4. Excel will now start a Wizard to import the file.
5. Most probably Excel will recognise the correct settings. In step one of the wizard choose 'Delimited' as file type. In step two make sure that only the option 'Tab' is selected. In step three the 'Column data format' should be 'General'.
6. After clicking 'Finish' you can specify where you want to place the imported data.

3.2 LibreOffice Calc

When trying to open the .csv file with Calc the Text Import Wizard should automatically start. Make sure to select 'Tab' under 'Seperated by' and deselect 'Space' (this should be the default setting). Under 'Other options' the first option 'Quoted field as text' should be selected. Click 'OK' an you are done.

4 Understanding and using data

In this section we will try to recreate one of the graphs from the website. These graphs are of a special type and use different types of axes. A short explanation is given.

Exercise 1 : Download the coincidence data (number of events per hour) from a detector station and data of your choice and import it into you spreadsheet program. Be sure to choose a detector that was correctly working on the date from which you download the data. Save the data in the standard format of your program so you can easily edit and save it later on. This will also prevent you from making any changes to the source file.[1]

Try to recreate the graph from the HiSPARC data site. Remember: a graph always has a title and clear descriptions on the axis.

The Event Histogram is as the name suggest a histogram type graph. During the day a lot of possible showers created by cosmic rays are recorded by every detector station. All recorded events are put into 'bins'. A bin is like a bucket, all events from the first hour of the day are put into the first bucket, all events from the second hour into the second bucket, etc. At the end of the day the number of events in every bucket is counted and plotted into the graph.

Exercise 2 : When you look at the Event Histogram you see that the graph is not continuous or smooth, there are clear 'steps' in the plotted line. Explain the shape of the graph.

Exercise 3 : Explain in your own words the Pulse height histogram. What is plotted on the x- and y-axis? How are the 'buckets' filled? Take a look at the source data from the graph to find (some of) the answers.

Exercise 4 : The axis of the Event Histogram and Puls height histogram are different. One uses a log scale. Explain in your own words the difference between a 'normal' (lineair) scale and the log scale. Why is the pulse height displayed using a log scale?

[1] As a student, you can never edit data on the website so do not be afraid to break anything. You might however change your downloaded data and mistake if for the original data.

5 Simple data processing

Now that we know how to obtain data we can look at patterns and relationships. We could for instance examine if there are more coincidences during the day time, or if the moon influences our measurements. Let's start simple with the determination of the average number of coincidences during one week.

This section is written as a short introduction into the use of a spreadsheet program. If you are already familiar with this type of program you might want to consider not doing the (first few) exercises of this section.

Exercise 5 : Download and import the data of an entire week from a single station, make sure the detector was up and running correctly the entire week.

Exercise 6 : Calculate the average number of events per hour per day and place it below the recorded events per hour. Give this row of numbers a clear description in the first column so you can remember which number is which. Are the averages the same for every day? Do you have an explanation for the similarity or differences?

Exercise 7 : Calculate the average of every hour and place these numbers in a separate column. Again give this column a clear description. Make a graph of the averages. Are there differences between this graph and the previous graph you made? Can you see a pattern?

Exercise 8 : How would you compare the data from two different detector stations? What information about the different detectors is available that could help you explain any differences?

Exercise 9 : If you want to look at possible effects of the moon, what is the minimum amount of time you would have to take data from?

6 More advanced data processing

The next assignment requires a closer look at the data from the website. Important questions are ‘Is this difference a real (significant, meaningful, important, ...) difference?’, ‘Is the data reliable?’, and ‘How much data do I need to answer my question?’. After this assignment you can chose one of the more elaborate assignments from the next section.

An important - and still unanswered - question in the field of high energy cosmic radiation is the location of the sources of the radiation. Lets see if we can locate a possible source of cosmic radiation, or at least the direction in which we should be looking.

In theory the HiSPARC detectors could detect particles coming from any direction. However, because of the geometry of the detectors they are most sensitive to particles coming straight from above. Luckily this is also where most of the showers come from (see the Route Net module ‘Airshowers’ for more information). Lets assume that all particles we measure at 12 o’clock come from directly above our detector, that would be the direction of the Sun. An hour later our detector has turned away from the Sun and is now looking at a slightly different part of the Sky.

The Earth rotates both around its own axis and the around the Sun. This means that, although at 12 o’clock we will always look at the Sun, the stars will have a slightly different position. To take this into account in our calculations we need to use a different time system; celestial or sidereal time (see the Route Net module ‘The Sky’ for a bit of background information).

Every day the Earth completes one revolution around its axis of rotation and every year it makes one revolution around the Sun. If we take the stars as reference point however, it seems as though the Earth makes one extra revolution around its axis every year. That is why the *rotation period* as used by astronomers differs from the solar day, it excludes that part of the rotation that is needed to accommodate the portion of the *orbital period* during one day.

Exercise 10 : Calculate the duration of a sidereal day in solar hours.

Using this new time system we can try to figure out where cosmic radiation comes from. To do this we will divide the sky into four quadrants and calculate the amount of radiation coming from each of the quadrants using vector decomposition, see figure 6.1.

Time represents a certain angle with respect to the stars. We can switch between time and angle using the following equation:

$$\alpha = 2\pi \frac{t}{T} \tag{6.1}$$

t is the time we want to convert into an angle and T is the duration of an entire revolution. Try to figure out for yourself what time we need to use in this equation, solar or sidereal.

Assigning the data from every hour to the quadrant of figure 6.1 can be done using the following equations:

$$Data_y = Data \cdot \sin(\alpha) \tag{6.2}$$

$$Data_x = Data \cdot \cos(\alpha) \tag{6.3}$$

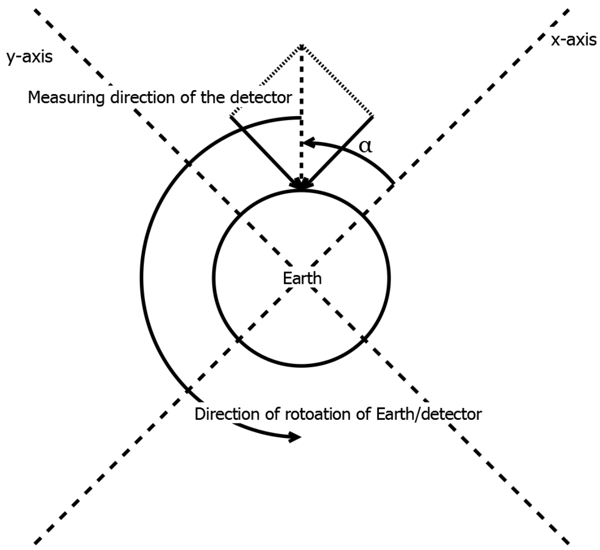

Figure 6.1: The sky is divided into quadrants and the data is assigned to the quadrant according to the measurement angle α

Or, if we combine these equations with equation 6.1:

$$Data_y = Data \cdot \sin\left(2\pi \frac{t}{T}\right) \tag{6.4}$$

$$Data_x = Data \cdot \cos\left(2\pi \frac{t}{T}\right) \tag{6.5}$$

In these equations $Data$ is the total amount of coincidences recorded during a solar hour. $Data_x$ and $Data_y$ represent the data assigned to the quadrants, there is a positive and negative direction making four quadrants.

If the measured cosmic radiation is uniformly distributed over all directions then the sum of $Data_x$ and $Data_y$ of one entire period will be zero. If however there is more radiation coming from a particular direction than at least one of the sums will not be zero. This enables us to point towards a direction of a possible source of cosmic radiation.

This analysis involves a lot of calculations, luckily the spreadsheet programs can do these calculations for us. But first we need to enter the correct formula and enough data. The calculations described above only work with a whole number of periods of data.

Exercise 11 : How many solar hours data are needed to obtain a whole number of sidereal rotation periods?

As the last question showed we need a lot of data to make the calculations work. Not only does this involve importing large amount of data (a lot of work) but our detectors have to be up and running without problems for the same period of time. To make life a bit easier we will use a shorter period of time. To negate the problems in our analysis caused by the incomplete number of periods used we will assign weighting factors to the data. Data at the start and end of our data set will count less towards the final answer while the middle part will be treated normally.

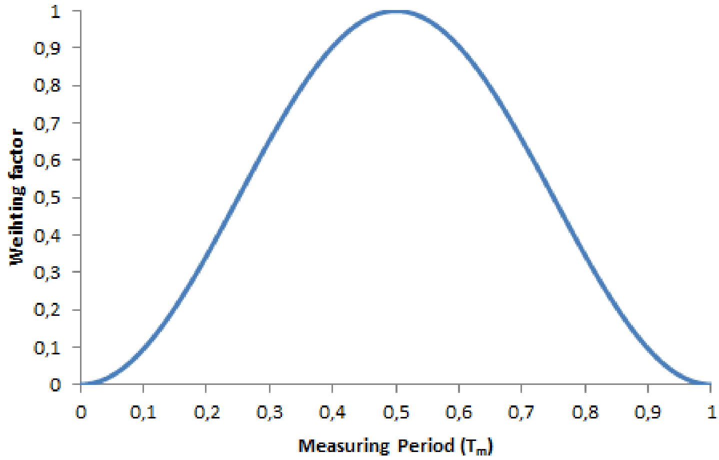

Figure 6.2: The weighting curve used on the data, a raised cosine

The weighting function should be a pure Gauss-curve. However this curve is infinitely long and we already had the problem of insufficient data. Therefore we will use an approximation which can be easily calculated using a spreadsheet program, a raised cosine:

$$\text{weighting factor} = \frac{1}{2}\left(1 - \cos\left(2\pi\frac{t}{T_m}\right)\right) \tag{6.6}$$

The entire procedure now becomes:

1. Collect the data from a single detector station.
2. Check the validity of the data: Was the detector always on? Are there no strange peaks? Etc.
3. Calculate the duration of a sidereal day.
4. Multiply the data with the raised cosine function to apply weighting.
5. Calculate $Data_x$ and $Data_y$.
6. Determine the sum of $Data_x$ and the sum of $Data_y$.

Are the sums 0? Then there is no direction from where more radiation is detected. Is the sum not 0, then we can take our analysis further. We can use our sums to reconstruct the angle from which the excess of radiation was detected. Lets say we obtained the following sums: $\sum Data_x = -525$ and $\sum Data_y = 260$. We can represent this as a line in a coordinate system such as figure 6.3.

Does this nice line mean their is actually something from the direction of the line? We cannot be sure because charged particles are deflected by magnetic fields. But before we start looking up at the sky, let us first look at the amount of excess radiation. To do

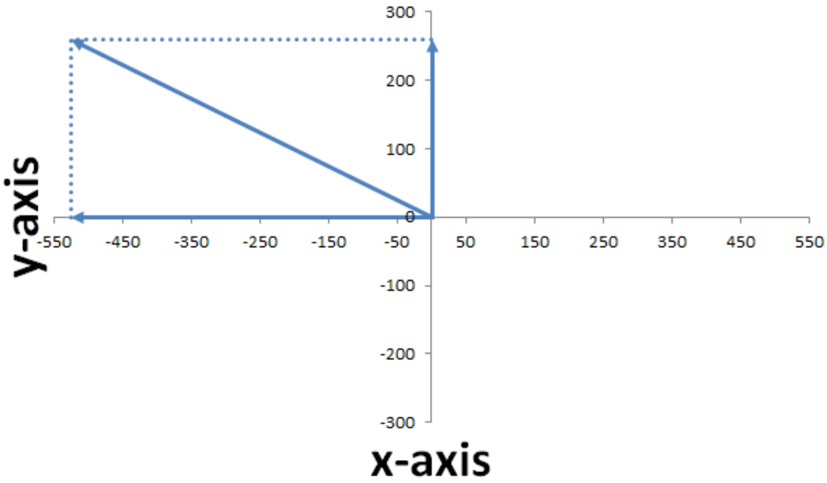

Figure 6.3: The result of the data analysis, the direction of excess radiation represented as a line

this we will compare the amount of radiation coming from the found direction with the total amount of radiation using:

$$\text{deviation} = \frac{\sqrt{\sum Data_x^2 + \sum Data_y^2}}{\sum Data} \cdot 100\% \tag{6.7}$$

Exercise 12 : Use the analysis described above to analyse one week's worth of data from a detector station of your choice. Do you see a region of the sky from which more cosmic radiation is hitting the Earth. Is the excess amount of radiation 'significant'?

Extra: You could use a program such as Stellarium or the Sky View option of Google Earth to look at the region of space from which you detected the excess amount of radiation. Is there something there? Be careful in answering this last question, (low energy) cosmic radiation is easily deflected by the different magnetic fields present in the universe.

7 Seasonal effects

In the following assignment you will do an analysis be yourself. The exercises will guide you through the process of doing research. You will need the basic techniques to download, import and process data that you have learned in the previous chapters. Once you have completed this analysis you can compare your analysis with the example given at the end of this chapter. Thereafter, you are ready to do the research assignments in the following chapter all by yourself.

7.1 Your analysis

Research always starts with a research question. Sometimes you have to formulate it on your own. In this assignment, however, the research question is already given:

Does the amount of cosmic radiation measured by HiSPARC stations depend on the season?

In order to find out whether, and how, the seasons influence the amount of measured cosmic rays, we will need to know some more about the seasons.

Exercise 13 : Find out why the Earth experiences seasonal change throughout the year. What is the distance between the Sun and Earth during the different Seasons? Is this influencing the seasons? What are the differences between the northern and southern hemispheres? What is happening to the Earth during its path around the Sun? Try to answer these questions, but there may be more to it!

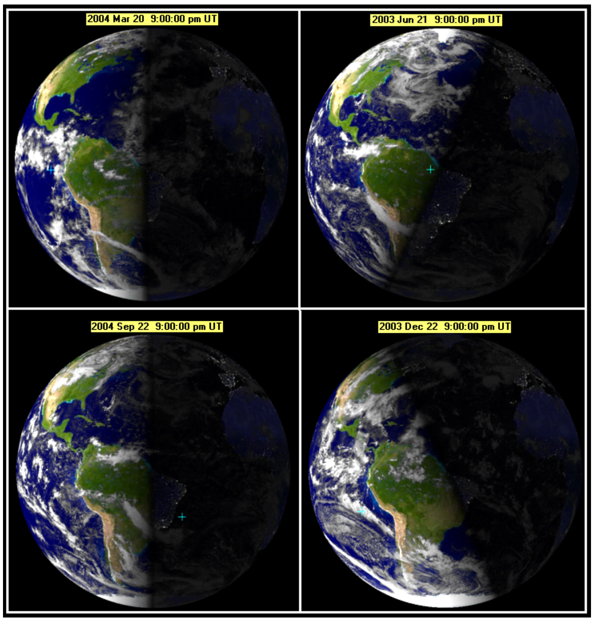

Figure 7.1: Illumination of the Earth at each season

As you probably found out, or already knew, the Earth is orbiting around the Sun. One trip around the Sun takes one year. Because of the elliptical shape of the orbit, the distance to the Sun differs during the year. However, it is not the orbital shape, but the axial tilt of the Earth that is mainly responsible for the four seasons. During the summer, the Sun is higher above the horizon causing the days to become longer and warmer than in the other seasons. When it is summer on the northern hemisphere, it is winter on the southern hemisphere.

At the northern hemisphere and following the astronomical calendar, spring starts on the day when the Sun reaches the point where it is standing exact above the equator (19th or 20th March). Summer starts on the longest day of the year (21st June). Autumn begins when the Sun is, again, standing right above the Equator (22nd or 23rd September). And winter begins when the Sun is at lowest on the horizon and thus the day is the shortest (20th or 21st December).

Meteorologists, however, use another calendar. For practical reasons, the Societas Meteorologica Palatina, the first international meteorological organisation, decided in 1780 that a season consist of three full calendar months. Therefore, spring starts on March 1st, Summer on June 1st, Autumn on September 1st and Winter on December 1st.

Now you know what mechanism is causing the seasons we experience. Lets find out whether this influences the amount of cosmic rays we measure with HiSPARC.

Exercise 14 : Try to think of several reasons why and how the seasons could be influencing the number of cosmic rays measured by HiSPARC and write them down. Keep in mind, your ideas cannot be wrong. It is what you are expecting. You will test your hypothesis with an analysis of HiSPARC data later.

You could also look on the internet for previous work considering cosmic ray intensity and seasons. You might be able to find some studies performed by researchers on this subject[1].

Before you can start testing your hypothesis lets think about the method of your analysis first. This method is the path you will follow to answer the main question of this assignment.

In Chapter 6 you already learned that importing large amounts of data involves a lot of work. Also, the HiSPARC detectors should have been up and running without problems during the whole period of time you are interested in. Recall the things you learned in previous chapters to negate these problems.

In the following exercises you will make several decisions. For each choice you make you should be able to explain why you made that decision. Remember, you design your method to test your hypothesis. Take care to make notes of all relevant parameters regarding station, time period or any other noteworthy peculiarities. This may be of use when comparing your results with your fellow researchers! Parameters regarding station hardware can be found on the data website.

[1]Note that the results of other scientific work does not make your analysis superfluous, your work may lead to new insights!

Exercise 15 : Decide how much data you want to download from the `http://www.data.hisparc.nl` website and for which period(s) of time. Also, think about whether you want to use the astronomical or meteorological calendar to define the seasons.

Exercise 16 : Decide which detector station you want to use to do your analysis. And if you want to compare the cosmic ray intensity with other parameters, for example temperature, make sure these data are available for the station of your choice. For weather data you might want to use the following website: `http://knmi.nl/climatology/daily_data/selection.cgi`

Exercise 17 : Decide which numbers you want to compare. You can compare daily, weekly or monthly averages in different seasons, compare seasons or whatever you think is best for testing your hypothesis. You also might want to decide whether you compare absolute or relative numbers.

Once you have finished the last three exercises you can continue to perform your analysis. Keep your your method in mind at all time. You already decided how you are going to test your hypothesis and thus what your next steps are going to be.

Exercise 18 : Download the data you want to use in you analysis and import the data to a spread-sheet application. Organise the data such that you can easily perform your analysis. You may want to recall your method.

Exercise 19 : Calculate all the numbers you need. Remember that you already decided in exercise 16 which numbers you want to compare. Make sure you give these new numbers a clear description. In this way you will remember what the numbers represent. Also, other people may want to use your data and will be able to see what the numbers mean.

Exercise 20 : Put all the interesting numbers on a single spread sheet (whithin the same file) . This will make it easier to compare the numbers. Again give the numbers clear descriptions. You want to know what you are comparing.

Exercise 21 : Make graphs of your numbers. Decide which numbers you want to put in a graph. Also decide what kind of graphs you want to use. Histograms, scatter plots or maybe another type of graph? Make sure the right quantity goes on the right axes. Give each plot a clear title. The axis should have a clear description too. Do not forget to mention units. When comparing you results with your fellow scientists, you might want to know what you are comparing!

Now that you have got your results, you can draw your conclusions from your results. You could either confirm or reject your hypothesis. It is also possible that your results neither confirm nor reject your hypothesis.

Exercise 22 : Recall your hypothesis and compare them with your results. Are they in agreement? And are your results in agreement with the studies performed by other researchers? And what about the results of your fellow researchers? Describe why it is in agreement.

It can be that there is a discrepancy. Try to explain why your results do not match your expectation. Maybe there were other parameters influencing the measurement? Was there something wrong with the detector station you chose?

Note that when your results do not match your expectation that your analysis is not necessarily wrong. Maybe nature is just not the way you thought. Also, there are a lot of factors influencing your measurements, just try to find out why your results are different.

At this point you have almost completed this research assignment. You have taken a good look at the theory of the seasons and how the seasons can influence the intensity of cosmic rays. Maybe, in your analysis, you found that the seasons do influence the cosmic-ray intensity, maybe you found it does not. Whatever your findings are, you must look critically at your own work. You can discuss every single decision you made during the whole process. Self-reflection is an important part of your work. You will see that it will increase the quality of your next work!

Exercise 23 : Look critically at your own work by discussing the choices you made and answering the following questions. How reliable are your findings? What would you do different if you were to do this same research again? What would you do the same? What would be a next step in future work?

The exercises you have done to complete this assignment are the basic steps you have to follow during a research assignment. It all starts with a research question. You first take a look into the theory of the subject you want to study. Look for useful websites or articles of other researcher who already have been looking into your problem. Based on the theory you found in the literature you formulate a hypothesis. You design a method with which you want to test your hypothesis and try to answer your research question. For HiSPARC related problems, you will mainly use data from the HiSPARC website and analyse them. You draw conclusion from the results to answer your research question and end by giving a discussion of your own work.

In the next chapter, ten research assignments are given. Whatever the assignment is, try to follow the steps you have been following in this assignment.

- Research question
- Theory
- Hypothesis

- Method
- Analysis
- Results
- Conclusion
- Discussion

Helpful websites
`http://scholar.google.co.uk`
`http://h2g2.com/approved_entry/A526673`
`http://astro.unl.edu/naap/motion1/animations/seasons_ecliptic.html`

7.2 Example analysis

In order to give you an idea of how such an analysis could look like, the following example is given. Again, note that if your analysis is looking different from this example it is not per definition wrong!

Research question:
Does the amount of cosmic radiation measured by HiSPARC stations depend on the season?

Theory:
The article of P. Adamson and colleagues, called "Observation of muon intensity variations by season with the MINOS far detector" which was written in 2012, states the following:
'When very high energy cosmic rays interact in the stratosphere, mesons are produced in the primary hadronic interaction. These mesons either interact and produce lower energy hadronic cascades, or decay into high-energy muons which can be observed deep underground. While the temperature of the troposphere varies considerably within the day, the temperature of the stratosphere remains nearly constant, usually changing on the time scale of seasons (...). An increase in temperature of the stratosphere causes a decrease in density. This reduces the chance of meson interaction, resulting in a larger fraction decaying to produce muons. This results in a higher muon rate observed deep underground.'

Hypothesis: Following the theory described above, in summer, when the temperature of stratosphere increases, the density decreases. Because of this decrease in density, less interaction will take place resulting in an increase in the number of muons reaching the ground. In the winter a decrease in temperature will result in an increase in density. More interactions will take place resulting in a lower number of muons observed at ground level. We assume that the number of muons reaching the ground is a measure of how many cosmic-ray particles we measure with a HiSPARC station.

Due to temperature changes in the stratosphere (and thus density changes) throughout the year, a change in the number of particles measured by HiSPARC is observerd.

- During the summer we will observe an above average amount of measured particles.
- During the winter we will observe a below average amount of measured particles.

Method:

In this analysisl the astronomical calendar will be used to define the seasons. Data will be takem from the middle of each season. This is because the assumption is made that it takes time before the stratosphere reaches its highest temperature. Data from the summers and winters of 2006, 2009, 2010, 2011 and 2012 will be downloaded (2007 and 2008 data was not available). For each season we will use one week data, resulting in a total of 70 days of data. That is for each year from February 3rd to February 9th and from August 3rd to August 9th. Unfortunately, we were unable to find any data considering the stratosphere temperature, hence we will use temperature data retrieved from the following website: `http://knmi.nl/climatology/daily_data/selection.cgi`. Here, it was chosen to do the analysis using the Nikhef-501 station. Most of the data required for the analysis are available. The weather data will be retrieved from the Schiphol weather station.

From the `http://data.hisparc.nl` website we can see that the number of events changed over the years. In 2009 there was an increase in the average of numbers of particles measured. From 2010 it decreased. We might need to use a correction for this trend behavior. There seems to have been some troubles with the detector. Especially in April 2011. It was down for maintenance on April 16th 2011. Assuming that the decline of the number of particles over the years is due to hardware issues, we will compare summer and winter averages with yearly averages. Thereafter we will compare the relative differences of several seasons. Furthermore we will compare the daily average temperature with the average number of measured cosmic-ray particles.

Analysis & Results:

We downloaded all data and imported them in Microsoft Excel. We put each season in a separate spread sheet (see figure 7.2) and calculated all numbers needed for the analysis (see figure 7.3). That is average number of measured particles per hour each day and per season. Because we are interested in the influence of stratosphere's temperature, data of all day (day & night) is averaged. Also the mean temperature in a season is calculated (see figure 7.4). In the screenshots you can see that we calculated a few more numbers. These numbers may be used for when we want to study other effects then described in the method. For now, they will not be used.

Now, all numbers of interest are put in a single work sheet (see figure 7.5). In this way, other interesting numbers can be calculated and compared, like the relative difference in the number of particles measured in summer and winter in comparison with the whole year. The other reason for making a summary of all interesting data is that it makes it a lot easier to select data to make histograms and scatter plots. Here, a histogram was made of the number of particles measured in each seasons relative to the average number of particles in that year (see figure 7.6). Also, because we made the assumption that the temperature influences the number of particles measured, a scatter plot was created with the temperature and the corresponding average number of particles per hour (see figure 7.7). This temperature however, is not the temperature of the stratosphere but

	A	B	C	D	E	F	G	H	I	J	K	L
1	Winter 2010											
2	Date:	February	3rd	4th	5th	6th	7th	8th	9th			
3		# bin	value	value	value	value	value	value	value	Average/bin		
4		0	3542	3505	3348	3409	3148	3291	3485	3390		
5		1	3474	3349	3449	3352	3309	3275	3412	3374		
6		2	3601	3432	3476	3389	3157	3348	3436	3406		
7		3	3561	3407	3495	3306	3234	3255	3355	3373		
8		4	3476	3418	3433	3378	3224	3459	3545	3419		
9		5	3489	3218	3605	3317	3169	3396	3493	3384		
10		6	3487	3283	3489	3280	3192	3373	3454	3365		
11		7	3458	3349	3506	3405	3085	3259	3411	3353		
12		8	3417	3312	3461	3289	3145	3517	3437	3368		
13		9	3348	3311	3550	3281	3222	3448	3454	3373		
14		10	3357	3297	3511	3265	3244	3363	3478	3359		
15		11	3327	3200	3537	3215	3160	3294	3395	3304		
16		12	3270	3297	3476	3404	3156	3365	3421	3341		
17		13	3340	3336	3558	3315	3108	3336	3266	3323		
18		14	3199	3203	3546	3288	3265	3381	3462	3335		
19		15	3379	3285	3545	3355	3182	3542	3443	3390		
20		16	3242	3295	3619	3296	3245	3562	3431	3384		
21		17	3244	3236	3507	3375	3274	3312	3380	3333		
22		18	3381	3386	3486	3227	3243	3387	3423	3362		
23		19	3326	3411	3381	3343	3245	3394	3498	3371		
24		20	3351	3414	3383	3280	3295	3351	3458	3362		
25		21	3335	3442	3423	3268	3266	3376	3520	3376		

Summary | Winter2010 | Summer2010 | Winter2011 | Summer2011 | Winter2012 | Summer2012

Figure 7.2: Screenshot of the seasons analysis. Data of the Nikhef-501 station is imported in excel. Each season is put into its own work sheet.

28	Hourly Average per day	3391	3341	3476	3313	3210	3384	3447	
29	Daily total	81390	80191	83428	79514	77039	81205	82723	
30	Hourly Average Winter 2010	3366							
31	total Winter 2010	565490							

Figure 7.3: Screenshot of the seasons analysis. All interesting numbers are calculated from the data retrieved from the data.hisparc.nl website.

Date	Tmean [°C]	TminGround [°C]	Hourly average per day
3-2-2010	2,3	-1,6	3391
4-2-2010	6	1,6	3341
5-2-2010	4,2	1,6	3476
6-2-2010	2,5	0,6	3313
7-2-2010	0,9	-0,1	3210
8-2-2010	-3,3	-4,2	3384
9-2-2010	-2,8	-4,4	3447
Mean	1,4	-0,9	3366

Figure 7.4: Screenshot of the seasons analysis. The temperature of 1 week is compared with the average amount of particles measured by the Nikhef-501 station.

	A	B	C	D	E
1	Bin average				
2					
3	Season	Total Count	Average p.h.	Mean T [°C]	% from yearly mean
4	Winter 2010	565490	3366	1	7%
5	Summer 2010	495468	2949	17	-7%
6	Winter 2011	470799	2802	7	4%
7	Summer 2011	434042	2584	18	-4%
8	Winter 2012	404874	2410	-2	-5%
9	Summer 2012	444086	2643	17	5%
10					
11	Year	Total count	Average p.h.		
12	2010	1060958	3158		
13	2011	904841	2693		
14	2012	848960	2527		
15					

Figure 7.5: All interesting numbers are put on to a single work sheet. Other numbers like percentage from yearly mean are calculated here.

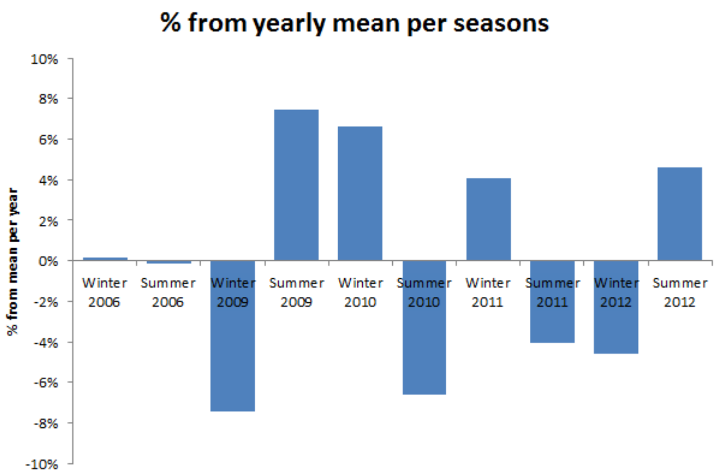

Figure 7.6: For each seasons was checked whether the amount of particles measured per hour was different from the mean of that year. In the years 2006, 2010 and 2011 the results look similar. In 2009 and 2012 however, the result is inversed.

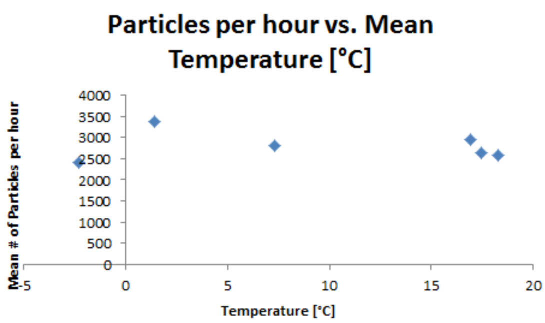

Figure 7.7: The mean temperature of a season is given with the corresponding mean number of particles measured per hour in that season. No trend is visible from the graph.

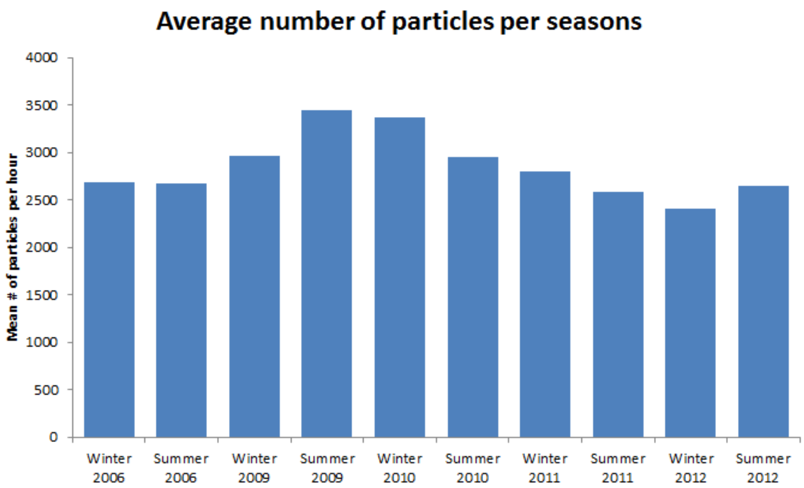

Figure 7.8: The average amount of particles measured per hour in each season over the years.

the temperature near the ground. This temperature can not be compared with that of the stratosphere.

Conclusion:

The hypothesis, the amount of cosmic-ray particles measured by HiSPARC is affected by the seasons, can neither be accepted nor rejected. The results of the seasons2006, 2010 and 2011 seem to be inversed with the results of the year 2006 and 2013. Also, the mean temperature does not seem to influence the number of particles measured. This is however, is the temperature near the ground. Nothing can be said of the influence of the temperature of the stratosphere on the number of cosmic-ray particles we measure with HiSPARC. As already stated in the method, a trend change of the average number of particles is observed in the data (see figure 7.8) over the years. From 2009 we see a decrease. This change is what we see in the results and it is not the seasonal influence. It is, however, unclear what is causing this decrease.

Discussion:

The article we used states that with an increase of temperature the number of muons measured below the earth increases because of the decrease of density. This indicates that the number of particles measured by HiSPARC should be higher in summer. There are however articles, like that of A. Duperier, that state exactly the opposite. Duperier wrote in his 1940 article *'The seasonal variations of cosmic-ray intensity and temperature of the atmosphere'* that with higher temperatures, the travel distance of cosmic rays and subsequent produced air showers through the atmosphere is longer which leads to a decrease of cosmic ray intensity measured at ground level.

For the analysis weather data was used from the KNMI website. Only the mean temperature just above ground level was measured at that station. However, the hypothesis was based on the temperature differences in the stratosphere. We have found no data on these temperatures.

The assumption that the decline of the averages is due to the hardware might not be right. Maybe we should have checked first whether this was the same for other stations. It could be that there is something else causing this decrease. It could for example be that it is the Sun. An article from 1954, written by S. Forbush, states that the amount of cosmic rays is inverse proportional with the number of sunspots. So, the more sunspots, the less cosmic rays we observe. From the website `http://spaceweather.com/` one can find that the number of Sun spots has been increasing over the years, which could explain the decrease observed in our data.

For future work we suggest that the same research is repeated for another HiSPARC station. It can be that the seasonal effect is measurable for other stations. It could also be that the same decrease is observed with other HiSPARC stations which would be an indication for the hypothesis that the solar cycle may be observed with the HiSPARC network. An interesting question might be how these sunspots influence the cosmic rays.

There may be other factors influencing the amount of cosmic rays during the seasons. For example, the distance from the Earth to the Sun might influence our measurements. The Sun is a source of cosmic rays and as it is standing closer to Earth during the summer, it might increase the amount of cosmic-ray particles we measure. One could test whether the sun is a significant source of observed particles in a day-night analysis of the data.

7.3 A critical look at the example analysis

Maybe you noticed that the example analysis was not perfect. There are a few good things to it, but if you look critically at this analysis you might see that some choices are not correct or complete. Let's take a look at the example analysis. For each exercise try to answer the questions, but you may add your own comments too.

Exercise 24 : First, take a look at the theory. What theory is chosen? Is this theory applicable in our situation? Explain why, or why not.

Exercise 25 : What can you say about the hypothesis? Where is it based on? Can it be that the hypothesis is wrong?

Exercise 26 : Now take a look at the designed metod. From which dates data were taken? Would you have picked the same? And what about the amount of data? Can the hypothesis be tested with the method? Can the research question be answered? Would you design the method differently.

The data used in this analysis shown a change in the number of particles over the years. It is assumed that this is a hardware problem. But the method does not state what hardware may cause the problem. Can you find out why? Hint: `http://data.hisparc.nl`

Exercise 27 : The analysis was performed after downloading and importing the data. What do you think about the way the data was sorted? And the usefulnes of the calculated numbers? Do the calculated numbers have clear descriptions?

Exercise 28 : The results are presented in several graphs. What do you think about the way the numbers are presented? Has each graph got an clear titel and description on the axes?

Exercise 29 : What do you think about the conclusion? Does it answer the research question? Can you draw any other conclusions based on the results?

Exercise 30 : The discussion is an important part of the research. Describe wich parts of the discussion you think is good and which parts could be better. Do you think that the discussion is complete?

Now you are not only able to look critical at your own work, but also that of others. The next chapter consists of several research assignments. When you want to compare your results with others, you are able to take a critical look at their work too. You will know whether their analysis is valid and if you are able to compare your results.

8 Research Assignments

Below is a short list of research assignments with brief descriptions. Choose one, compare and discuss your results with your fellow ‘researchers’.

Important: Not all detectors are working correctly 100% of the time. Also, some detectors are used to answer a specific research question and are therefore calibrated or set up differently. The detector in Karlsruhe for instance in placed inside a larger detector array KASKADE. It uses the triggers from this array to record its own data, not the normal HiSPARC triggers. A number of the Science Park detectors are also different. In most cases the name suggest as much (e.g. MPPC test station). For a list a special detectors and with remarks and comments take a look at Appendix A - List of special detector stations on page 25.

Always look at the data before downloading it and use only valid data[1]. Some of the assignments below have a specific list of station and/or cities to look at. When such a list is present only use these stations. You still need to look at different dates to see if the detector was working properly during the time frame in which you are interested.

Be prepared to answer the question why you chose a particular station on a particular day (or the reverse question: Why did you not use data from ...).

From all detectors a number of parameters are known. Via the website you can find the location of the station, the used hardware, the number of detector plates, and the voltage supplied to the PMTs. When needed use this information, perhaps already during your selection of detector stations to compare.

Latitude effect Does the amount of cosmic radiation hitting the atmosphere of the Earth depend on the location, more specifically the latitude?

The HiSPARC network has detectors placed at different locations across Europe. Some are clustered together at roughly the same latitude, but others are located far more South of North. Compare the event rate of two of the listed combination below or select one and add your own pair of detectors.

- Bristol - Groningen
- Bristol - Aarhus
- Bristol - Eindhoven
- Bristol - Tilburg
- Bristol - Leiden
- Eindhoven - Tilburg
- Groningen - Aarhus
- Eindhoven - Aarhus
- Enschede - Leiden

[1]Answering the question whether or not data is ‘valid’ can be very difficult, certainly with limited knowledge of the detectors. Therefore treat the ‘valid’ requirement as ‘looking similar to most stations’.

Is most cities there is more than one detector station. You are free to choose one (or perhaps combine the results of multiple stations). When choosing stations for your analysis, make sure your have selected stations with a similar configuration (two or four plates)!

Longitude effect Does the amount of cosmic radiation hitting the atmosphere of the Earth depend on the location, more specifically the longitude?

The HiSPARC network has detectors placed at different locations across Europe. Most are located within a small region of longitude with no more than a few degrees difference. Now with the addition of Bristol and cities nearby to the network we are able to look at possible differences caused by longitude.

Take a look at the different active detectors on the HiSPARC data site. Make (at least) two pairings of different detector sites with a large difference in longitude, but not in latitude. (Or perhaps make one pair with comparable longitude and latitude.) Test whether or not there is a longitude effect.

Be careful in your choice of detector stations, make sure your have selected stations with a similar configuration (two or four plates)!

Day-Night Variations Our Sun is a source a cosmic radiation. Because it is very close it has a large influence on life on Earth. Are the results from HiSPARC affected by the Sun? Make your own research plan to look for day/night variations. Should there be an influence from the Sun, is this a wanted or unwanted effect? Why?

Extra: Take a look at the data from station 7101 (Hengelo - Bataafs Lyceum) between May 14 and May 28. Clearly there is something happening here. But the stations nearby (Enschede and Haaksbergen) report nothing out of the ordinary. Possibly there is something wrong with the detector in Hengelo which was fixed in the last week of May. What do you think was wrong with the detector?

Influence of the Moon The Moon is our nearest celestial neighbour. How does the Moon influence cosmic radiation on Earth? Does is shield the Earth from the rays, like is stops asteroids from hitting the Earth, or is the Moon too small to strongly influence cosmic rays in this way? Or could the Moon in some way attract cosmic rays towards the Earth?

Compose your own research protocol to search for possible influences of the Moon. Important questions to keep in mind are the amount of data you need and which detector station or stations you are going the use. (Hint: Take a second look at the text from section 6 - More advanced data processing, what important fact about the detector sensitivity is given?)

Height effect Most detectors are located on the roof of schools. But not all roofs are the same height. On the data page of each detector station, below the location of the detector the height is also given.[1] Use this information to test if the height of the detector has a (large) influence on the recorded rate (or for a more challenging

[1] Note that this height locations is not always accurate. For one it is dependent on the accuracy of the GPS antenna. An error of 10 metres in height is not uncommon. Also the location of the detector is only rarely updated. If the detector has been moved the listed location might not be correct any more.

investigation: on the pulse height spectrum). In your analysis of the results, relate the difference in height to the thickness of the atmosphere through which the radiation travels, remember to take into account the varying density of the air with height.

Detector placement Most detector are placed on a (flat) roof. Some however are placed inside, below a roof. Does this influence the recorded rate or the pulse height spectrum (more challenging question)? The following detectors are placed under a roof: 7001, 7002, 7003.

Weather effects - Pressure Does the weather affect cosmic radiation and airshowers?[1] There are a number of HiSPARC stations that are equipped with weather stations. Stations with a weather station show two extra graphs on the data page: Barometer data and Temperature data.

Does the ambient air pressure (pressure at the location of the detector station) influence the measured event rate? To answer this question you need to obtain a lot of data which is grouped per hour and determine the (average) pressure at the moment of measurement. Can you find a connection between the ambient pressure and the even rate? Can you make a graph showing the relationship?

The following station have a weather station (or had one in the past): 3, 501 (past), 509, and 8001. Can you use the weather data from these stations when looking at data from nearby stations? Why or why not?

Extra: The Royal Netherlands Meteorological Institute (KNMI) has made all its data about the weather in the Netherlands freely available and easily accessible via a website (also in English `http://www.knmi.nl/climatology/daily_data/selection.cgi`). If you want you can use this data to do you analysis. Unfortunately we have not found an easily accessible British counterpart.

Weather effects - Temperature Does the weather affect cosmic radiation and airshowers?[2] There are a number of HiSPARC stations that are equipped with weather stations. Stations with a weather station show two extra graphs on the data page: Barometer data and Temperature data.

Does the ambient temperature (temperature at the location of the detector station) influence the measured event rate? To answer this question you need to obtain a lot of data which is grouped per hour and determine the (temperature) pressure at the moment of measurement. Can you find a connection between the ambient pressure and the even rate? Can you make a graph showing the relationship?

Also try to answer the following question: Which temperature change are you looking at in this analysis? The temperature of the detector, the air, both, something else? Does this matter?

For more information about the weather data look at the description of the previous assignment.

Take a look at Appendix B - List of detectors with correct height on page 26 for a list of stations with a verified accurate listed height.

[1] Or is it the other way around?

[2] Or is it the other way around?

Weather effects The two previous assignments talk about possible effects of temperature and pressure. What if they both have an effect? What if they have an opposite effect?

Work together with the groups working on the previous assignments to compensate the pressure analysis for any temperature effects and/or vice versa. Do this analysis on your own, but work on the same data as your companion group.

Number of plates There are now three different detector station configurations present in the network. Most station have two plates placed 10 metres apart. A few have four plates laid out as an equilateral triangle with side of 10 metres with the fourth detector at the centroid of the triangle. Detector 13001 (University of Bristol) also has four plates but since the 25th of October they are placed on the points of a rhombus (diamond). The rhombus is constructed by laying two equilateral triangles with sides of 8 metres next to each other.

With four plates you have a greater change of 'catching' a shower. How much more showers are detected by a four plate detector (difference in recorded rate)?[1] Which showers are recorded by the four plate detector that are missed by a two plate set-up? Analyse the Pulse height spectra to answer this question.

[1] Assume the rate of showers is constant both in time and space.

A List of special detector stations

Some HiSPARC detectors are used to answer specific questions or to test new hardware or software. In most cases the data from these detectors cannot be used directly in other analyses.

Below is a (non exhaustive) list of detectors with different settings which should not be used for the assignments of section 8 - Research Assignments.

11001 HEP Group / Cavendish Laboratory Never recorded data intended to be used in the HiSPARC project.

12001 Prior Pursglove College Not calibrated to HiSPARC specifications.

701 Forschungszentrum Karlsruhe HiSPARC station placed inside the KASKADE project, using the triggers of the larger project.

20XX Nijmegen The detectors in the Nijmegen clusters use a different version of the HiSPARC hardware. These are being updated to the newest version. Comparison of data can be difficult.

93 DAQ III Test A special station used to test new hardware.

94 Tijd Test 2 (Swap) A special station used to test the timing of the hardware and possible differences between electronics boxes.

95 Tijd Test (Reference Used in conjunction with station 94 for testing timing differences.

96 Weather station test Contains only weather data, was used to write the weather station module of the HiSPARC DAQ software.

97 MPPC test station A special station that was used to research the possibility of replacing the now used photomultiplier tubes with silicon photomultipliers (Multi-pixel Photon Counter).

98 Used for various tests.

99 Used for various tests, amongst others different detector station geometries.

10001 University of Sheffield Only operated for a limited time to allow the University of Sheffield to test the detector.

8102 St. Odulphuslyceum Tilburg Not operating continuously.

2101 Raaijland College No recent activity, before that no usable results.

B List of detectors with correct height

Below is a list of detectors whose height is correctly displayed on the data website.

- 3
- 5
- 6
- 7
- 9
- 13
- 22
- 101
- 102
- 103
- 104
- 202
- 303
- 501
- 502
- 504
- 507
- 509
- 601
- 1006
- 1007
- 1008
- 2001
- 3001
- 3002
- 3104
- 3201
- 3301
- 3302
- 3303
- 3401
- 4001
- 7001
- 7101
- 7301
- 7401
- 8001
- 8002
- 8004
- 8005
- 8006
- 8007
- 8009
- 8101
- 8103
- 8105
- 8201
- 8303
- 13001
- 20002
- 20003