Norwegian University of Life Sciences

Master's Thesis 2021 300 ECTS
REALTEK

Optimizing production time of electrolyzers powered by off-grid PV-arrays

Sebastian Kihle
Environmental physics and renewable energy

Preface

My five years here at the Norwegian University of Life Sciences will conclude with this thesis. The thesis represents my knowledge and insight within a small field of my degree in Environmental physics and renewable energy. It started when Nel was in contact with the university, wanting to learn about ways to build an off-grid electrolyzer without batteries. Seeing the rapid change in power from PV when shaded, proved that it would be difficult to have a stable hydrogen production without batteries. This is why the main task of the thesis was to predict when cloud shadows would arrive at the PV-array. It has been challenging to develop a method to track the velocity of cloud shadows, meeting technical problems and waiting for the right weather conditions. In the end, it all worked out thanks to my skilled supervisors. I would like to thank Heidi S. Nygård for her assistance in the structure, content and feedback of the thesis. Writing well is difficult, but Heidi made it easier. Her course also made me understand power systems on a higher level, which definitely made it easier to solve some of the problems I faced. Espen Olsen's insight into PV technology made him a reliable source of knowledge, and his course taught me most of what I know about PV. Espen did a great job giving feedback and doing quality control of much of my writing. Mareile A. Wolff helped me understand clouds and the weather, making it possible to make accurate assumptions. She also contributed with data for making the some some of the figures in the thesis, which I would not have found myself. Furthermore, I would like to thank Marko Jakovljevic for his contribution in building the triangulation point. Marko built all the sensors, more or less from scratch, in addition to data logging. Being a great technician solved many problems along the way, and made all the results come together. Lastly, I would like to thank Lars Wikström, my contact person in Nel. We had many interesting conversations, and he came up with many good ideas which have become an essential part of the thesis. This thesis would not be possible without him. An important part of my studies have been my friends. They have helped me when I needed it, and made every day more fun. Without them as a source of mutual motivation, I would probably not have made it through the degree. I also thank my parents for nurturing my interest in natural sciences when growing up, and being there for me during both the good and hard parts of life.

Abstract

This thesis takes a look at how to maximize the production time and power of an electrolyzer powered by a PV-array in a off-grid system without any batteries. Since an alkaline electrolyzer can not be ramped down instantaneously and there are no backup batteries, the electrolyzer needs to be ramped down in advance of a cloud shadow hitting the PV-array. To detect and predict when a cloud will hit the PV-array, three sensors configured in an equilateral triangle are used to triangulate the cloud shadow direction and speed. The triangulation point, consisting of the three sensors, will be placed a distance from the PV-array to be able to estimate when a cloud shadow will hit the PV-array. To test the system, the triangulation point was built on a roof in Ås, Norway in an equilateral triangle with sides of 15m, with a fourth sensor placed down range of the cloud with a distance of about 1km for reference during measurements. In order to triangulate, a time difference between each sensor in the triangulation point is needed. The time difference was found by inspecting the power gradient. Yet, to avoid noise, the power gradient was calculated from a third order polynomial regression of the power curve with a high R^2-score. Three time difference methods were tested, but only one was used. The time difference method that proved most reliable, consists of inspecting a relative change in the power gradient on an interval. The intervals were chosen based on a rapid change in power over 20 seconds or less. In total, measurements were done over three days, with slightly varying conditions. The triangulation gave promising results on average, but struggles under some circumstances. With some adjustments to the triangulation points and the algorithms used, it is believed that the accuracy of the model can increase. Using multiple triangulation points around the PV-array seems necessary in order to detect cloud shadows from all directions. Multiple triangulation points also have the possibility of communicating with each other, which can increase the accuracy of the predictions. There is also a possibility of increasing the efficiency of the system by bypassing the DC/DC converter between the PV-array and the electrolyzer. This requires the electrolyzers current characteristics to match the PV-array's MPP during the optimal irradiation interval. The optimal irradiated power interval for bypassing the DC/DC converter, at 200W intervals, is 800W-1000W. When the DC/DC converter is bypassed at this irradiation interval, there is an increase in delivered energy of 0.849kWh per year per square meter of installed PV.

Sammendrag

Denne gradsoppgaven ser på hvordan man kan maksimere produksjonstiden til en elektrolysør drevet av en PV-array i et off-grid-system uten batterier. Ettersom elektrolysøren ikke kan skrus av øyeblikkelig og det ikke er noen backup-batterier, trenger elektrolysøren å skrus av i forkant av en sky som vil skyggelegge PV-arrayen. For å detektere og forutsi når en sky vil skyggelegge PV-arrayen, benyttes tre sensorer i en likesidet trekant til å triangulere skyskyggens fart og retning. Trianguleringspunktet, bestående av tre sensorer, plasseres et stykke unna PV-arrayen for å estimere når skyskyggen vil treffe PV-arrayen. For å teste systemet ble trianguleringspunktet plassert på et tak i Ås som en likesidet trekant med sider på 15m, med en fjerde sensor som ble plassert i ca. 1km i skyretningen som referanse under målingene. For å triangulere, trenger man å vite tidsdifferansen mellom hver sensor. Tidsdifferansen ble funnet ved å inspisere den deriverte av effekten til hver sensor. For å unngå støyen i dataene, ble tredjegrads polynomregresjon benyttet på effektkurvene, med høy R^2-verdi. Tre tidsdifferansemetoder ble prøvd, men bare én ble benyttet. Den tidsdifferansemetoden som var mest pålitelig baserer seg på en relativ endring i den deriverte effekten fra sensorene på et intervall. Intervallene ble valgt basert på en rask endring i effekt på 20 sekunder eller mindre. Det ble totalt gjort målinger for tre dager, med litt varierende forhold. Trianguleringen gav lovende resultater gjennomsnittlig, men sliter under noen omstendigheter. Med noen justeringer i trianguleringspunktene og algoritmene, kan sannsynligvis presisjonen til modellen øke. Det virker nødvendig å ha flere trianguleringspunkter rundt PV-arrayen for å kunne detektere skyskygger fra alle retninger. Dersom man har flere trianguleringspunkter, har de muligheten til å kommunisere med hverandre og samarbeide for å øke presisjonen på predikasjonene. I tillegg er det en mulighet til å øke effektiviteten til systemet ved å forbigå DC/DC omformeren mellom PV-arrayen og elektrolysøren. Dette krever at strømkarakteristikken til elektrolysøren passer med PV-arrayens MPP på intervallet med optimal innstråling. Det optimale innstrålingsintervallet for å forbigå DC/DC omformeren, på 200W intervaller, er 800W-1000W. Når DC/DC omformeren forbigås på dette intervallet, øker man den leverte energien med 0,849kWh per år per kvadratmeter installert PV.

Abbreviations

- PV – Photovoltaics
- HFCV – Hydrogen Fuel Cell Vehicle
- Si – Silicon
- c-Si – Crystalline Silicon
- B – Boron
- P – Phosphorus
- SRH – Shockley-Read-Hall
- V – Volt
- MPP – Maximum Power Point
- MPPT – Maximum Power Point Tracker
- DC – Dicrect Current
- PWM – Pulse-Width Modulation
- H – Hydrogen
- O – Oxygen
- K – Potassium
- KOH – Lye
- atm – Atmospheric pressure
- kWh – kilo Watt hour
- N m^3 – Normal cubic meter
- TP – Triangulation Point
- CS – Cloud Shadow
- CNN – Convolutional Neural Network
- NN – Neural Network
- AI – Artificial Intelligence
- GPS – Global Positioning System

- NMBU – Norwegian University of Life Sciences
- m – meter
- cm – centimeter
- PVGIS – Photovoltaic Geographical Information System
- KDE – Kernel Density Estimation
- WiFi – Wireless Fidelity
- NaN – Not-a-Number
- $\frac{\mathrm{m}}{\mathrm{s}}$ – Meters per second
- W – Watt
- s – second
- N – North
- E - East
- S – South
- W – West
- km – kilometer
- mAh – milliampere hours

List of symbols

- e^- – electron
- E_γ – photon energy
- E_g – bandgap energy
- E_v – energy of valence band
- E_c – energy of conduction band
- P_{loss} – resistive power loss
- R – resistance
- I – current
- I_{sun} – current produced in a PV-cell when there is sun
- V_{sun} – voltage produced in a PV-cell when there is sun
- E_s – high-voltage source
- E_0 – low-voltage source
- e_L – induced voltage in inductor
- I_a – maximum current
- T_a – switch *on* time
- T – period of one cycle
- D – relationship between the time a switch is open versus closed
- I_b – inductor discharge current
- I_0 – average current of a pulse
- I_S – average current of a cycle
- E_L – voltage over switch 2 in figure 2.3.2 (b)
- S1 – switch 1
- S2 – switch 2
- I_L – current through the inductor
- E_H – high voltage source

- p – western PV-module
- q – northern PV-module
- r – eastern PV-module
- a – distance between p and q
- b – distance between p and r
- c – distance between q and r
- α – angle between cloud shadow direction and a
- β – angle between cloud shadow direction and b
- θ – angle between a and b
- v – cloud speed
- t_b – time the cloud's shadow uses moving over b
- t_r – time when r shuts off
- t_p – time when p shuts off
- t_a – time the cloud's shadow uses moving over a
- σ – correction angle relative to north
- ϕ – cloud's shadow direction relative to north
- E_y – energy irradiated on an interval on average over a year
- τ – fraction of a year where the irradiation lies on the interval
- κ – number of samples in an interval
- H – number of samples in all of the data
- K_1 – lower limit for a power interval
- K_2 – upper limit for a power interval
- η_{DC} – efficiency of the DC/DC converter
- η_{PV} – efficiency of the PV-array with cable losses
- E_{save} – energy saved by bypassing the DC/DC converter on a power interval
- d_{PV} – distance between the sensors in the triangulation point
- dP – change in power

- dt – change in time
- G – Power threshold value for making an interval
- M – power drop or power rise interval
- B – data for a day
- $f(t)$ – third order polynomial approximation of power
- $Prod_{inc}$ – increase in hydrogen production per year by bypassing the DC/DC converter
- A_{PV} – total area of installed PV in the grid

Contents

1 Introduction

As the world shifts its focus towards climate change, alternatives to carbon based power production are gaining attention. This has resulted in increased energy production from renewable sources [1]. One of the renewable technologies that has seen the most improvement in recent years is photovoltaics (PV), more commonly known as solar cells [2]. Hydrogen gas based technology such as fuel cells has also advanced in recent years [3]. Unlike PV, which harnesses energy from the sun, hydrogen stored in pressurized containers and fuel cells are energy carriers. This means that energy can be stored in the form of hydrogen in a similar way to electric batteries. Assuming that hydrogen follows the "supply and demand" principle of economics, the demand for hydrogen will increase as the global consumption rises. According to Park, Kim and Lee, the amount of Hydrogen Fuel Cell Vehicles (HFCV) will increase over the next 18 years in Korea, and the next 30 years in the US, which likely will contribute to a large increase in hydrogen consumption [3]. However, hydrogen has more potential than just as fuel in cars. Because pressurized hydrogen containers do not lose energy over time like a battery does, they can store energy more efficiently over longer time than batteries. This gives way to transporting large quantities of energy, in the form of hydrogen, on ships (preferably running on hydrogen or electricity) or pipelines as an alternative to power cables. In addition, hydrogen is not dangerous for the environment in itself, so a leak would not impact the local climate like a fossil gas or oil. Hydrogen is explosive however, so leaks could be dangerous to the surroundings, especially if ignited. If the world wants to become more renewable and sustainable, hydrogen has the potential to make a positive impact, if produced in a renewable way.

One way of producing hydrogen in a renewable way is to use electrolyzers powered by PV-arrays. Electrolyzers are machines that run a chemical process where water is split into hydrogen gas and oxygen gas by using electricity. If the PV-array is big enough to power the electrolyzers, the electrolyzers do not need to be connected to the power grid. An off-grid electrolyzer powered by PV-arrays has the potential to be built anywhere in the world, as long as there is access to the sun. However, being disconnected from the power grid has its challenges. When the sun stops shining, the electrolyzers either need to be powered down, or get their power from an alternative source. For off-grid systems, batteries often serve as the backup power source. Still, in context of stand-alone systems that require high power, Von Meier explains the problem well: "Besides being expensive, batteries are toxic, corrosive, potentially explosive, and bulky; also, their performance is sensitive to proper treatment and maintenance. If a convenient and affordable alternative existed, it would no doubt revolutionize the field." [4].

This thesis will take a look at how to reduce or eliminate the need for batteries in an off-grid, PV-powered hydrogen plant. When a PV-cell is exposed to shading, it stops producing power almost instantaneously. Since the electrolyzer cannot be turned on or off instantaneously, it would be useful to know in advance when a cloud will cover the PV-arrays. By using triangulation and a measuring point, the speed and direction of a cloud can be determined,

which in turn indicates how much time will pass before the cloud hits the PV-arrays. The triangulation will be done by building a triangulation point a distance from the PV-array, consisting of three sensors to detect cloud shadow velocity. To check the accuracy of the triangulation point, the actual time the cloud shadow uses from the triangulation point to the PV-array will be measured. Being able to predict when power is lost is a useful tool in order to turn the electrolyzer on or off at just the right time to minimize production loss.

Additionally, a way to increase the system efficiency will be discussed. This can be done by bypassing the DC/DC converter. This thesis will try to find out which circumstances yields the maximum gain in hydrogen production when bypassing the DC/DC converter. This also includes the amount of energy saved and hydrogen produced from the DC/DC converter bypass.

2 Theory

The two main components in the off-grid system for hydrogen production powered by PV are the PV-cells and the electrolyzer. PV-cells are relatively easy to install and use. Nevertheless one should understand their working principles in order to utilize them in the most efficient way.

2.1 PV-cells

2.1.1 Semiconductors

PV-cells mainly consist of semiconductors. Semiconductors are a group of elements whose conductive properties can easily be manipulated. Their four *valence electrons* is what gives semiconductors their properties. Valence electrons are the electrons in the outermost shell of an atom [5]. In PV, the most common element to use as a semiconductor is silicon (Si) in crystalline form (c-Si). Since each Si atom has four valence electrons, it can form four covalent bonds with four other Si atoms because atoms are most stable when they have eight valence electrons.

2.1.2 Doping

To generate electricity, electrons need to move. An electron is a negatively charged particle that is a fundamental part of atoms. This is why we consider the lack of an electron, a *hole*, as a positive charge in an atom. To manipulate the electron concentration and the hole concentration, we use *doping*. Doping means to replace some atoms in a crystal lattice with other atoms that have more or fewer valence electrons. The most common elements to use in doping of c-Si are boron (B) and phosphorus (P) [5]. Consider phosphorus which has five valence electrons. When one Si atom is replaced by a P atom, four of the five valence electrons bond with the surrounding Si atoms, leaving one electron weakly bound to the P atom. Normal ambient room temperature gives off enough thermal energy to free the last electron from the P atom. Thus, the P atom "donates" a free electron and is thereby called a *donor* [5]. Donor atoms become positively ionized and a donor doped semiconductor is called an *n-type* material.

When considering doping with the B atom, which has three valence electrons, all three electrons bond to the neighboring Si atoms. However, one electron is missing in the B atom. The thermal energy from the surroundings will free an electron from a nearby Si atom, which will be "accepted" by the B atom. The freed electron will leave a hole, which acts as a positive charge. This makes the B atoms contribute to the concentration of holes and these atoms are called *acceptors* [5]. Acceptor atoms become negatively ionized and acceptor doped semiconductors are called *p-type* materials.

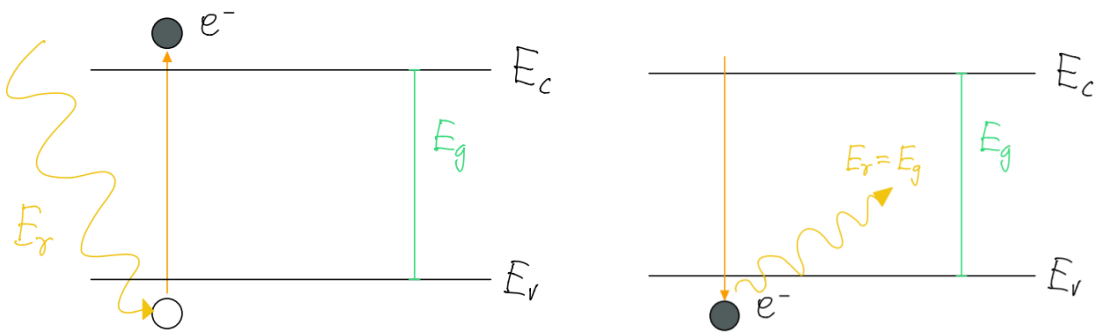
2.1.3 Charge transportation

There are two mechanisms for transporting charge carriers, *drift* and *diffusion*. "*Drift* is charged particle motion in response to an electric field." [5]. In an electric field, positively charged

particles, such as holes, move in the positive direction of the electric field, while negative particles, such as electrons, move in the negative direction of an electric field. This means that it is possible to separate holes from electrons using electrical fields. Diffusion is motion caused by a gradient in the particle concentration [5]. This means particles move from regions with high concentration to regions with lower concentration.

2.1.4 Generation and recombination

When light (photons) hit the semiconductor, an electron (e^-) can be freed using the energy from the photon. This requires the photon energy (E_γ) to be higher than the *bandgap energy* (E_g). The bandgap energy represents the energy difference between the bound state of an electron in the valence band (E_v) and the free state of the electron in the conduction band (E_c). As illustrated in Figure 2.1.1 (a), the excited electron leaves a hole in its previous position (the white dot).



(a) *Generation of an electron-hole pair caused by a photon.*

(b) *Recombination of an electron-hole pair.*

Figure 2.1.1: *Figures illustrating generation and recombination of charge carriers in the bandgap.*

The result is a free electron and a new hole, an electron-hole pair, which is a result of what is called *generation*. Generation does not exclusively occur by excitation from photons. It can also occur from thermal energy, yet most generation in a PV-cell occurs from the photons in sunlight.

Recombination is when an electron-hole pair meets and binds the electron (see Figure 2.1.1(b)). This reduces both the number of free electrons and holes, and generates a photon. There are four recombination mechanisms:

1. **Direct recombination**: Intrinsic recombination caused by the material. Also affected by temperature.
2. **Shockley-Read-Hall (SRH) recombination**: Recombination due to impurities and imperfections in the crystal lattice.
3. **Auger recombination**: Recombination that excites another electron to a higher energy

state, instead of generating a photon. The excited electron then thermalizes and recombines. This is a three-particle process.

4. **Surface recombination**: Recombination due to "dangling bonds" which are unsaturated bonds due contact with a different medium in one or more directions.

Recombination is largely undesired as it reduces the number of charge carriers in the semiconductor. Since SRH recombination makes up most of the recombination in indirect bandgap semiconductors like silicone, it is important to use extremely pure materials with few imperfections.

2.1.5 The p-n junction

When a p-type and an n-type semiconductor meets, a large difference between the electron concentration in the materials cause a diffusion current of electrons from the n-type material to the p-type material. Like with the electrons, the difference between the hole concentrations in the two materials cause a diffusion current of holes from the p-type material to the n-type material. Because of this diffusion process, most of the mobile charge carriers, such as electrons and holes, are gone from the vicinity of the junction between the p- and n-type materials. However, the ionized doping atoms are not mobile and cannot move, and thus give rise to a potential difference. The potential difference sets up an electrical field over the junction which keeps the holes and electrons separated in each of the materials. The region with the electrical field is called the *space charge region* [5]. When the forces from diffusion and the electrical field are equal, the system is stabilized and enters equilibrium.

When exposed to sunlight, the photons will excite and break the covalent bonds, freeing electrons in the p-type material. The newly freed electron will be forced over to the n-type material over the junction by the electrical field. Now there is a higher concentration of electrons in the n-type material than in equilibrium, but the electrons cannot move to the p-type material against the electrical field. By connecting to an external circuit between the two materials, electrons can move from the n-type material to the p-type material. Thus electricity is produced from the PV-cells.

2.2 PV-modules

The PV-cells usually have thin silver lines over their surface known as "fingers" to pick up the electrons efficiently (see Figure 2.2.1). On top of that, there is an anti-reflective cover to both protect the cells and absorb as much sunlight as possible. It is common to use several PV-cells to increase their power output. This is done by connecting them in series and/or parallel on a panel known as a PV-module.

When two or more PV-modules are connected together, they make up a PV-array. PV-modules are more frequently connected in series (strings) than in parallel, because when connected in

series, the voltage produced by each cell adds up, while in parallel the current of each cell adds up. A schematic of a module with PV-cells connected in series is given in Figure 2.2.2. In the case of direct current, like PV-modules produce, the resistive losses are given by Equation 2.2.1 [5].

$$P_{loss} = RI^2 \tag{2.2.1}$$

Here, P_{loss} represents the power lost by heating the conductor, R is the resistance in the conductor, and I is the current. The loss is proportional to the current squared, which is why parallel connections tend to be avoided.

In Figure 2.2.1, there is a distinct difference between the PV-cells in the two middle PV-modules compared to the others. This is because the two middle PV-modules consist of PV-cells that are *mono crystalline*, which means that they are made from a single silicon crystal. Because silicon crystals are grown as cylinders, the shape of mono crystalline PV-cells are octagonal instead of completely square to utilize the surface area of the cylindrical crystal efficiently. The other PV-modules consist of *multi crystalline* silicon, which are square and generally less efficient compared to mono crystalline silicon.

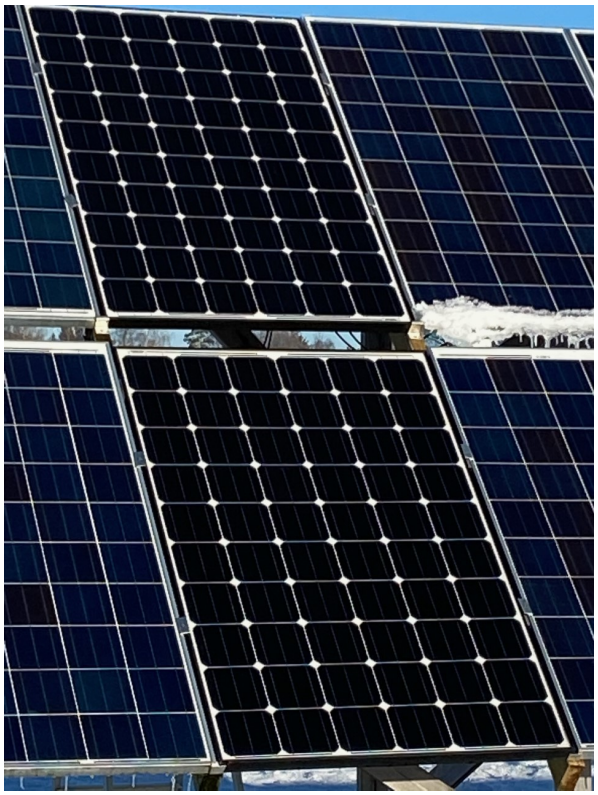


Figure 2.2.1: *Image of PV-cells (the small black squares) in a PV-module connected in a PV-array.*

2.2.1 Shading

One of the disadvantages of having all PV-modules connected in series (see Figure 2.2.2) is that if one stops producing power, they all do, in a similar way to old Christmas tree lights. Shading is the usual cause of local power disruptions in PV which is why it is important to place PV-modules and arrays without any obstacles nearby. When a cell becomes shaded, its current output is drastically reduced. Because current in a series connection is limited by the lowest current produced, all the other cells in the series will also have their current output reduced, as Figure 2.2.3 shows. Still, the voltage produced by the unshaded cells are the same, and will be lost over the shaded cell. This deposits all the generated energy in the string on the shaded cell, which heats it up and can break it [6]. Imagine if you have 20 PV-cells in series, and a shade covers half of one of the cells. The current produced in the cell that is partially covered will be halved, which means that the current in the 19 unshaded cells also will be halved. Now all of the 20 cells produce half power, while half of the power the 19 unshaded cells produce is lost as heat in the partially shaded cell. The same will happen if there are 200 cells in a series: if just half of one of the cells are shaded, all 200 cells will produce half power. The same applies if one considers the PV-cells as modules.

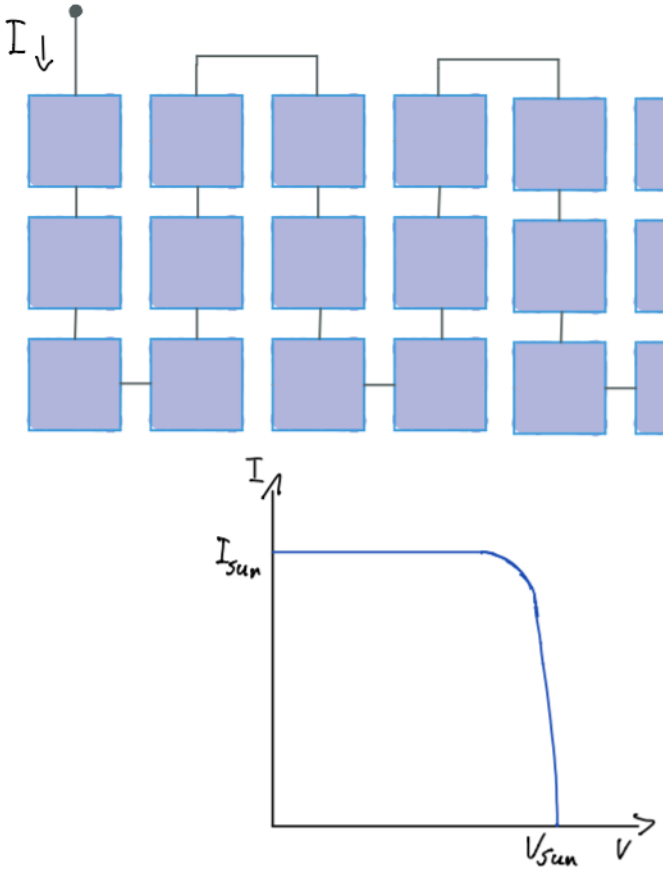


Figure 2.2.2: *Schematic of a PV-module where all the PV-cells are connected in series. Each blue square represents a PV-cell. The graph shows that the module produces a current of I_{sun} and a voltage of V_{sun}.*

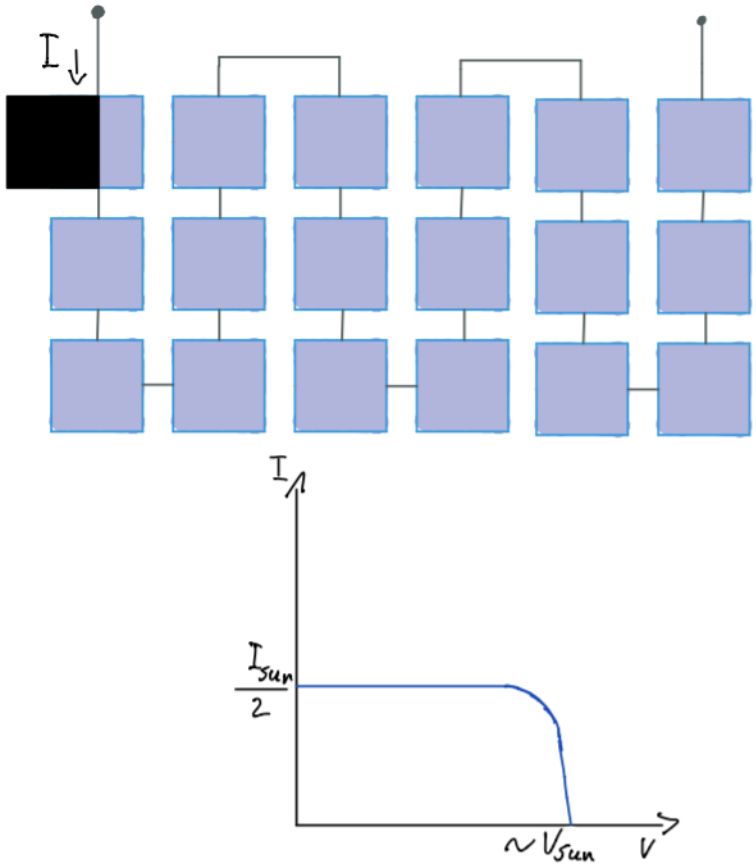


Figure 2.2.3: *Schematic of the same PV-module as in Figure 2.2.2, but here, half of one cell is shaded. The graph shows that the current in the module drops to half of* I_{sun}.

To reduce the power loss and potential damage from shading, *bypass diodes* are commonly used. They are usually placed between strings, in case one string is shaded. If only one string is shaded, the current in the series connection will run through the bypass diode instead of through the shaded string, thereby effectively disconnecting the shaded string, as Figure 2.2.4 illustrates. This makes sure that the current in the rest of the series connection is maintained while only a small amount of voltage is lost over the bypass diode. If a cell is completely shaded, a PV-module without bypass diodes would lose all power, while a PV-module with bypass diodes would only partially lose power. It would be ideal to have one bypass diode for each cell, but that would raise the cost of the PV-module drastically.

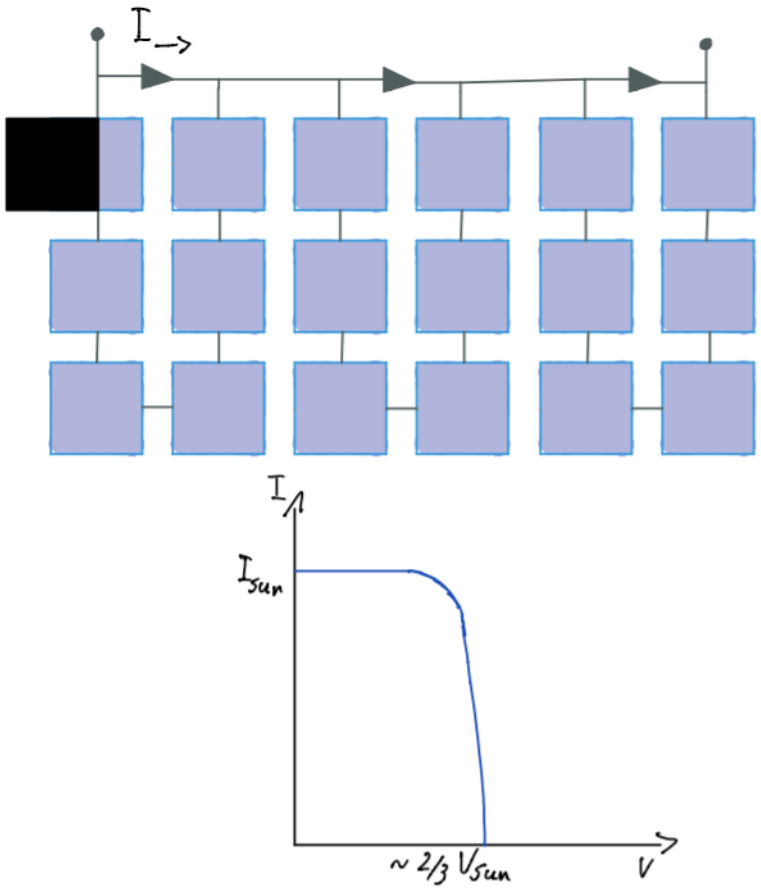


Figure 2.2.4: *Schematic of a PV-module with bypass diodes and a shaded cell. The graph shows that the voltage drops to* $\frac{2}{3}V_{sun}$*, but the current remains the same.*

Alternatively, PV-cells and PV-arrays can be connected in parallel. However, a parallel connection increases the current in the system, which increases the power loss as mentioned in Section 2.2.

2.3 Converters

The output power of a PV-module varies depending on the current [5]. To find the *maximum power point* (MPP), a *maximum power point tracker* (MPPT) is used. In Figure 2.3.1, the relation between voltage and current is clear. As the voltage on the x-axis increases, the current on the y-axis slowly drops until it rapidly drops to zero.

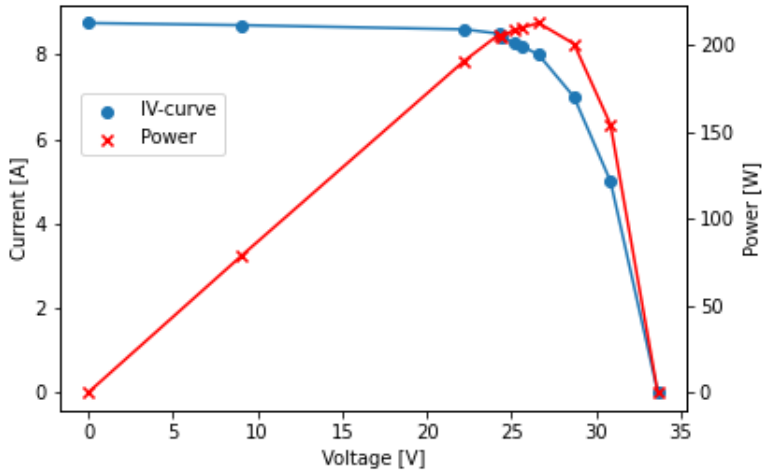
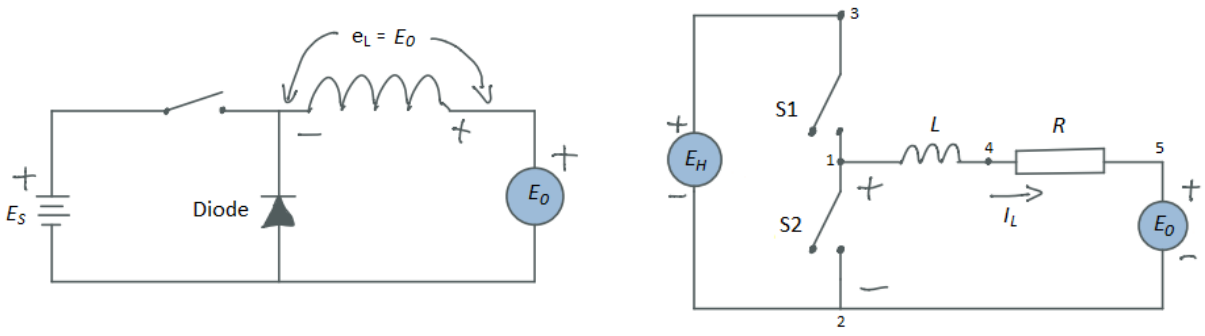


Figure 2.3.1: *Example of an IV-curve and power curve. Figure is based on measured data of an ITS MODUL EcoPlus Poly 250. Figure is reprinted with permission from [7].*

The blue line in Figure 2.3.1 represents the "IV-curve" of the PV-module. The current on the y-axis rapidly at about 27V, and it is at the same point the MPP lies, as expressed by the power graph. The task of the MPPT is to make sure that the current is adjusted so that the PV-module outputs maximum power. The MPPT changes the direct current (DC) by signaling a *DC-DC converter* which can alter the load voltage or current seen by the PV-module or array.

2.3.1 DC-DC converters

The following paragraphs simply explains how two different DC-DC converters operate according to Wildi [8]. In the case where power has to be transformed from a high-voltage source E_s to a lower-voltage source E_0, one can use a *chopper* (see Figure 2.3.2 (a)).



(a) *Schematic of a simple DC-DC chopper.* **(b)** *Schematic of a 2-quadrant DC-DC converter.*

Figure 2.3.2: *Figures illustrating two different simple DC-DC converters.*

A chopper connects an inductor between the source and the load, and has a switch that opens

and closes the circuit repeatedly. When the switch is closed, energy is stored in the inductor as the current builds up due to the applied voltage. When it opens, a voltage e_L is induced as the power transferred to the load must now come from the energy stored in the inductor. The diode prevents voltage loss in the form of an arc across the switch, and provides a freewheeling path for the load current when the switch is open. If the voltage drop over the diode is neglected, e_L will be equal to E_0. When the inductor is charged, it will contain the maximum current I_a which decreases as the inductor discharges.

By altering the relationship D between the time the switch is on T_a and the period of one cycle T (see Equation 2.3.1), it is possible to make the current oscillate between the maximum current and a lower current I_b, as illustrated in Figure 2.3.3.

$$D = \frac{T_a}{T} \tag{2.3.1}$$

The average current I_0 of each pulse can thus be calculated by Equation 2.3.2.

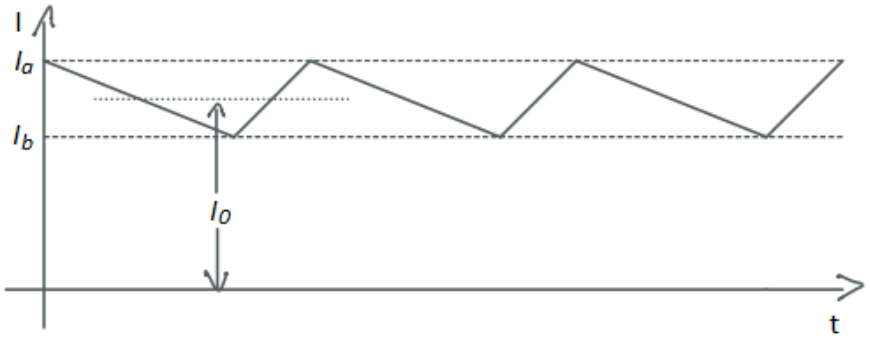


Figure 2.3.3: *Figure displaying how the current changes between I_a and I_b as the inductor charges and discharges, resulting in a mean current I_0.*

$$I_0 = \frac{(I_a + I_b)}{2} \tag{2.3.2}$$

The value of I_b can be lowered by letting the switch remain open for a longer period of time, letting the inductor discharge its energy for a longer time. This will also lower I_0 making it possible to adjust how the current will be transformed. For one cycle however, the average current I_S is calculated by Equation 2.3.3.

$$I_S = I_0 D \tag{2.3.3}$$

Yet, simple chopper converters can only transform a high current to a lower current, while a *2-quadrant DC-DC converter* can transform current both up and down.

In a 2-quadrant converter (see Figure 2.3.2 (b)), there is a constant DC voltage E_L between terminals 1 and 2. It consists of two switches, S1 and S2, instead of one as in a simple chopper. If E_0 is less than E_L, a DC current I_L will flow from terminal 1 to 5, and its magnitude will be given by Equation 2.3.4.

$$I_L = \frac{(E_L - E_0)}{R} \tag{2.3.4}$$

Under such circumstances, the converter works like a chopper discussed earlier, as power flows from the higher voltage E_H to the lower voltage E_0. If E_0 is greater than E_L however, the current I_L will change direction, and flow from terminal 5 to 1. In this case, I_L is calculated by Equation 2.3.5.

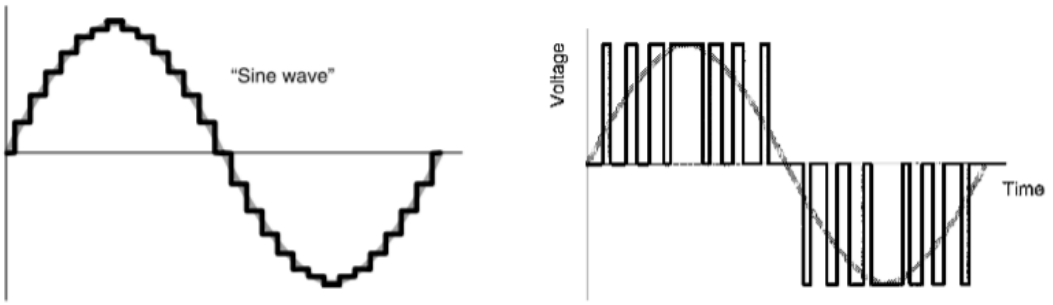
$$I_L = \frac{(E_0 - E_L)}{R} \tag{2.3.5}$$

In other words, if E_0 is greater than E_L, power flows from the from the low voltage side E_0 to the higher voltage side E_H.

Because the 2-quadrant converter requires a specific polarity from the load, it is preferred to use a *4-quadrant converter*. Put simply, the 4-quadrant converter uses two identical 2-quadrant converters to bypass the polarity problem, which is why 4-quadrant converters are the most flexible DC-DC converters.

2.3.2 Inverters

For machines and devices that use alternating current, it is usual to convert the direct current from PV-modules to alternating current. This is done by installing an *inverter* in the system. Figure 2.3.4 (a) shows how an inverter using a series of square waves called *quasi sine waves* imitates a sine curve.



(a) *Output from an inverter using a quasi sine wave approach.*

(b) *Result of current pulses from an inverter that approximates a sine wave using PWM.*

Figure 2.3.4: *Figures illustrating the operation of two different inverters. The figures are reprinted with permission from Von Meier [4].*

The current and voltage are kept at zero for a brief moment between each reversal to approximate the sine curve more accurately [4]. However, most modern inverters use a different approach, by sending small current pulses of varying length. This is called *pulse-width modulation* (PWM), and utilizes the short current pulses to create a an approximation of the sine curve as shown in Figure 2.3.4 (b) [4]. A low-pass filter consisting of inductors and capacitors smooths the pulses to an almost pure sine wave current.

Electrolyzers run on DC, which means that an inverter is unnecessary when the electrolyzer is connected to PV. This eliminates the need for both an inverter and a rectifier, which is beneficial because efficiency is lost at each conversion step.

2.4 Electrolyzers

An electrolyzer uses energy to split water molecules (H_2O) to hydrogen (H_2) and oxygen (O_2). An alkaline electrolyzer consists of two electrodes, an anode and a cathode, and an electrolyte (see Figure 2.4.1). In the anode, the solution is oxidized, while in the cathode, the solution is reduced. The electrolyte makes sure that the electrons move between the cathode and the anode. In theory, all you need is water as an electrolyte and electricity, but because of the high resistivity of pure water, aqueous lye (KOH) is more commonly used as electrolyte [9].

At the anode, the solution is oxidized, and the chemical reaction that occurs is given by Equation 2.4.1.

$$\text{Anode: } 2\,OH^-(aq) \longrightarrow \frac{1}{2}\,O_2(g) + H_2O(l) + 2\,e^- \tag{2.4.1}$$

While for the cathode, the solution is reduced, and the reaction is given by Equation 2.4.2.

$$\text{Cathode: } 2\,H_2O(l) + 2\,e^- \longrightarrow H_2(g) + 2\,OH^-(aq) \tag{2.4.2}$$

When the reactions are combined, the total reaction becomes Equation 2.4.3.

$$\text{Total: } H_2O(l) + \text{electrical energy} \longrightarrow H_2(g) + \frac{1}{2}\,O_2(g) \tag{2.4.3}$$

Both the anode and the cathode are illustrated in Figure 2.4.1.

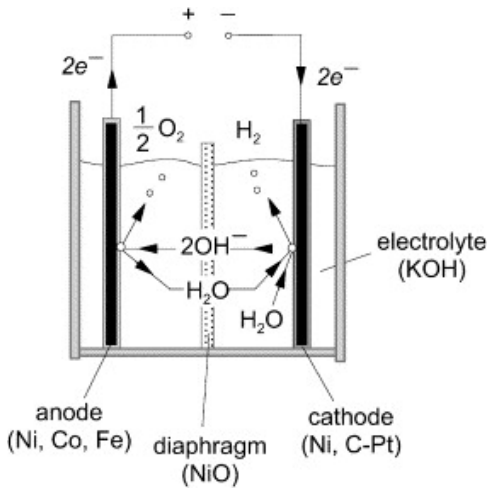


Figure 2.4.1: *Figure showing the principle of a alkaline water electrolysis. The figure is reprinted with permission from Ulleberg [10].*

The KOH makes sure that the ions are transported between the two electrodes in the form of $K^+$and OH^-. As shown in Figure 2.4.1, there is a diaphragm separating the anode from the cathode. The purpose of the diaphragm is to separate the two produced gases, oxygen and hydrogen, yet the ions need to pass between the diaphragm to maintain the reaction. Figure 2.4.2 shows a modern electrolyzer for reference.

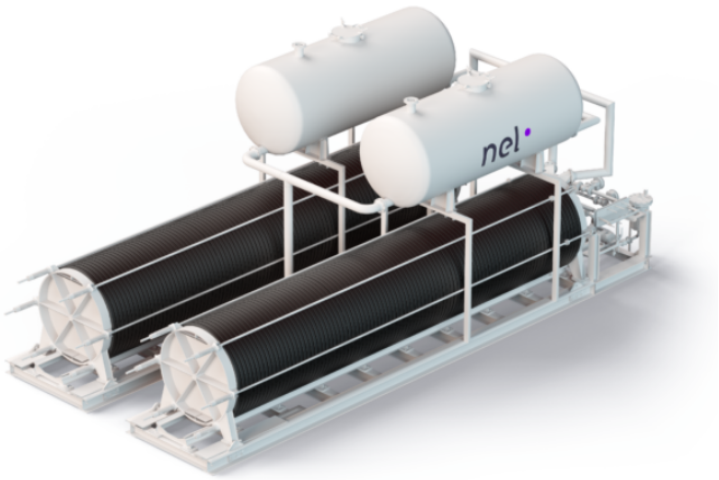


Figure 2.4.2: *A modern electrolyzer produced by Nel [11].*

In Figure 2.4.2, the horizontal black cylinders consist of many layers of electrolysis cells stacked next to each other. Figure 2.4.1 represents one of these electrolysis cells.

At standard conditions (298K and 1 atmospheric pressure (atm)), the total chemical reaction requires 1.226V. As a result of this, the minimum amount of energy required to produce one N

m^3 of H_2 is 2.94kWh [9]. Here, the N stands for "normal condition" which means the volume at 0℃ and 1 atm pressure. The reaction does in reality require more energy than 2.94kWh per N m^3 H_2 because there is resistivity in the electrolyte, in the diaphragm, and it requires an *overvoltage*. Overvoltage can be thought of as a "resistance" in the chemical reaction rate [9]. In reality, electrolyzers in industry (per 2003) require 4.3 - 5.3kWh per N m^3 H_2, and run on DC [9]. Recently however, this energy requirement has been reduced to 3.8kWh per N m^3 H_2 under certain conditions, though the normal efficiency is 4.4kWh per N m^3 H_2[11].

At the bottom of the electrolysis cells, there is a channel to transport lye through each electrolysis cell. There are also two channels at the top of each electrolysis cell, one for oxygen and one for hydrogen. During normal operation, there is a mixture of lye and gas in the two upper channels. When the power is cut off, the lye will not refill the electrolysis cell rapidly enough, which means that there will only be oxygen gas and hydrogen gas in the upper channel. This could lead to undesired pressure changes in the electrolyzer. The lack of lye will dry up the diaphragm between the two upper channels, which could lead to gas leakage through the diaphragm between the two channels. This would mix the oxygen gas and the hydrogen gas, which could be dangerous. To circumvent this, the electrolyzer has a gas quality sensor, which means that if the quality of the gases drop (by mixing), the electrolyzer would do an emergency stop with nitrogen purging. This is why the electrolyzer cannot be turned on or off instantaneously [12].

2.5 Clouds

Most clouds form when air rises, cools down and then condenses. The most common mechanisms that drives air upwards are surface heating, convection caused by topography (often mountains), weather fronts and convergence of surface air [13]. There are several different kinds of clouds which Luke Howard (1772-1864) divided into categories: *cumulus* which are clouds in a "pile", *stratus* which means "layer", *cirrus* which means hair strand, and *nimbus* which meant clouds carrying rain (see Figure 2.5.1) [14]. In addition, there are several sub-categories of clouds. Cirrus clouds however are not a cloud type of the same order as stratus or cumulus, as it is often used as a prefix to describe other cloud types such as cirrocumulus and cirrostratus.

The classification system is based on heights, which gives indications to what altitude a cloud lies at by looking at its type. Because clouds can have heights differentiating between a few hundred meters to well above five kilometers, it is hard to decide what speed they have, and what speed their shadows will have. This is why deciding a cloud's shadow speed is more reliably done by measuring the shadow itself. The wind speed, which moves the clouds, is dependent of height, which means that wind measured at ground level might not indicate the cloud direction. Higher clouds cast a shadow with a less distinct edge, making it more difficult distinguish the transition between sun and shadow. This also applies to clouds with less distinct edges, like cirrus clouds, which also are high clouds [15].

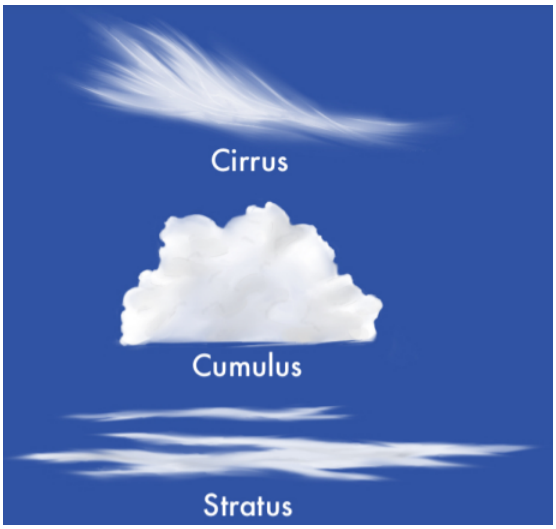


Figure 2.5.1: *Illustration of the three main cloud types. Illustration is by Abdi [16].*

The general wind speed and direction near ground level in Ås, where the test took place, can be represented by a wind rose as shown in Figure 2.5.2.

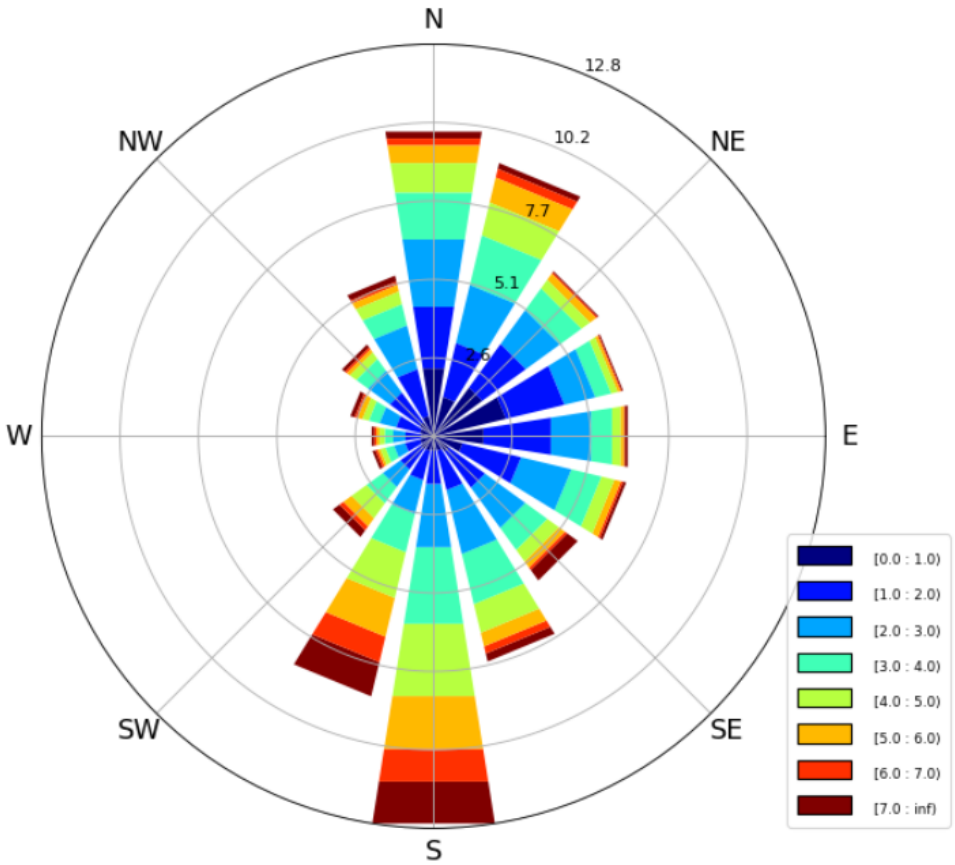


Figure 2.5.2: *Wind rose for the weather station in Ås based on wind measurements between 2014 and 2019. Measurements are means of 10-minute intervals. The values in the legend represent wind speed in meters per second.*

This wind rose is based on wind measurements from the weather station in Ås over five years. As the wind rose illustrates, the surface wind generally moves northeast or southwest, which is most likely due to a channeling effect by the nearby Oslo fjord. Yet, according to observations from *Ventusky.com*, the clouds in Ås appear to generally move southeast or southwest [17].

2.6 Triangulation

According to *The American heritage dictionary of the English language*, triangulation is: "A surveying technique in which a region is divided into a series of triangular elements based on a line of known length so that accurate measurements of distances and directions may be made by the application of trigonometry." [18]. In this case, three PV-modules placed in a triangle will be used to triangulate. For easier referencing, the PV-module triangle is called the *triangulation point* (TP).

When a cloud moves over the TP, its shadow will cover the PV-modules one by one, turning them off. By knowing the distances between each PV-module and looking at the time difference between when each PV-module turns off, it is possible to decide both the direction and speed of the cloud's shadow (CS). Figure 2.6.1 illustrates how a triangle can be used to triangulate.

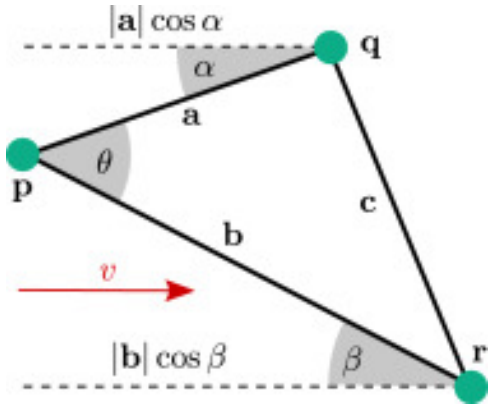


Figure 2.6.1: *Illustration of a triangle used in triangulation. Image is reprinted with permission from Cantwell [19].*

In this case, the v vector represents the CS, under the assumption that the CS front is linear. p, q and r represents the three PV-modules. The distances between points pq, pr and qr are known as a, b and c respectively. α denotes the angle between the CS direction and side a, while β denotes the angle between the CS direction and side b. θ is the angle between sides a and b.

The speed of the cloud can be calculated using Equation 2.6.1 [19].

$$v = \frac{|b| \cos \beta}{t_b} \tag{2.6.1}$$

where

$$t_b = \frac{t_r - t_p}{b} \tag{2.6.2}$$

Here, t_b denotes the time the CS uses over b. t_r and t_p are the points in time when p and r shuts off respectively.

α will be used to express the direction of the CS, which Equation 2.6.3 describes how to find [19].

$$\alpha = \arctan\left(\frac{t_a|a| - t_b|b|\cos\theta}{t_a|b|\sin\theta}\right) \tag{2.6.3}$$

Similarly to t_b, t_a is calculated by Equation 2.6.2 and represents the time the CS uses to pass distance a. However, α does not give the angle relative to north, which is the desired way of referring to the cloud direction. To adjust for this, a correction angle σ, is added to get the cloud's direction relative to north ϕ, as shown in Equation 2.6.4.

$$\phi = \alpha + \sigma \tag{2.6.4}$$

The σ is based on the orientation of the triangle, and differs based on which point is activated first, and what order the other points are activated. Values for σ are further used in Section 3.1.

2.7 An alternative way to predict clouds using neural networks

One alternative to using triangulation of CS is using a *convolutional neural network* (CNN). By using a camera, it is possible to analyze images of clouds for predicting their direction and speed. CNNs is a type of *neural network* (NN) that processes images by recognizing local patterns that are otherwise invisible [20]. These patterns can be used to extract information from images in a similar manner to the way humans do. CNNs are commonly used in fields like facial recognition, detecting cancerous tumors and reading handwriting [20]. NNs use artificial neurons (see Figure 2.7.1) that either "fire" or not, just like the neurons in a brain. By connecting many neurons (see Figure 2.7.1) in a network, they cooperate when processing data in a similar way to the human brain. NNs are a part of *deep learning* which is a subcategory of *machine learning* or *artificial intelligence* (AI).

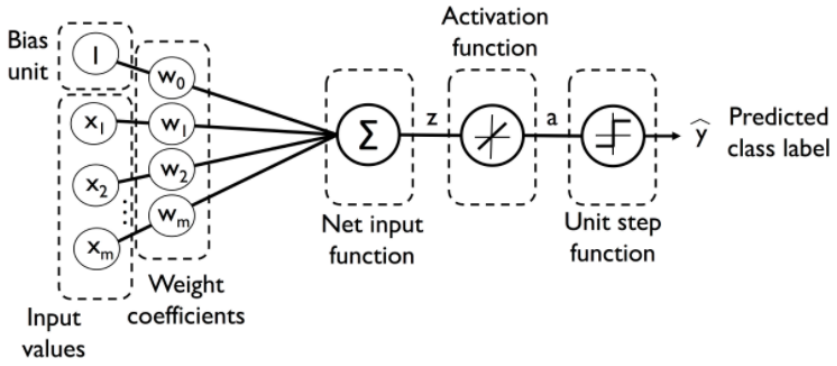


Figure 2.7.1: *Detailed representation of an artificial neuron. Image is reprinted with permission from Raschka [20].*

Because the output of each neuron is based upon mathematics and calculations, the capacity and parameters of the CNN model can be tailored to each problem in order to maximize the precision of the model. In Figure 2.7.1, each "box" represents either values or functions that work together in order to learn from the input data. When a model has finished training, it can use new data to make predictions. This however would require a "training" data set that the model could learn from in order to be used. To create such a data set would require a large number of images and data processing which is beyond the scope of this thesis.

2.8 DC/DC converter energy saving

Since the electrolyzer runs on direct current, the PV-array does not require an inverter. Avoiding the use of an inverter is beneficial because it increases the efficiency of the system. A DC/DC converter is however advantageous, since it allows the PV panels to operate at MPP and makes sure that the electrolyzer receives the right current and voltage. If the PV-array powering the electrolyzer is scaled with the electrolyzer in mind, the nominal voltage and current produced by the PV-array will be similar to the voltage and current consumed by the electrolyzer. When the power produced by the PV-array is close to the power consumed by the electrolyzer, the current does not need to be converted by a DC/DC converter. This means that the DC/DC converter can be bypassed, leading the power from the PV-array directly into the electrolyzer. Bypassing the DC/DC converter will increase the efficiency of the system by skipping a power conversion step. When the irradiated power is outside the bypass interval, the current will run through the DC/DC converter as normal.

To maximize production, the electrolyzer can be customized to draw a current close to the MPP current when it is desired to bypass the DC/DC converter. This is because the electrolyzer will work as a stabilizing force on the grid, making it possible to operate as an MPPT on a small interval [12]. This current must be chosen based upon the power from the PV-array in a way that utilizes both the power produced, and how frequent the power is produced. To explain, here is an example:

Imagine if the DC/DC converter is bypassed at nominal power production. Nominal power production is only achieved during short periods of the day, which means that even though the benefit from bypassing the DC/DC is largest during nominal production, the time period is short. If the DC/DC converter is bypassed at the average power produced during a day, the time is maximized, but the increase in production is not that large. Therefore, there is a need to find a balance between power gained and duration of the bypass activation.

When looking at how much energy that can be saved from the bypass, a closer inspection of the power interval where the DC/DC converter is bypassed is needed. To calculate how much energy is irradiated over an interval over the course of one year, Equation 2.8.1 can be used.

$$E_y = 365\text{d} \times 24\frac{\text{h}}{\text{d}} \times \tau \times \frac{K_1 + K_2}{2} \tag{2.8.1}$$

Here, τ is the fraction of the year where the irradiation lies on the interval $[K_1, K_2\rangle$, given by Equation 2.8.2. In Equation 2.8.1, E_y is the average energy irradiated on the interval $[K_1, K_2\rangle$ over a year in Wh.

$$\tau = \frac{\kappa}{H} \tag{2.8.2}$$

In Equation 2.8.2, κ is the number of samples in a power interval given by $[K_1, K_2\rangle$. K_1 is the lower power limit of the interval, while K_2 is the upper limit of the power interval. H represents the number of samples in all of the data. The yearly energy savings over the power interval $[K_1, K_2\rangle$ by bypassing the DC/DC converter is given by Equation 2.8.3.

$$E_{save} = (E_y - E_y\eta_{DC}) \times \eta_{PV} \tag{2.8.3}$$

The η_{DC} represents the efficiency of the DC/DC converter, while η_{PV} is the efficiency of the PV-array, including cable losses. E_{save} is the average energy saved, in Wh per m^2 of PV-array, over a year by bypassing the DC/DC converter.

3 Method

When a cloud's shadow hits a PV-array, most of the power in each string will be lost immediately because of the series connection. For an off-grid system, this means that all systems powered by the PV-array will shut down unless there are batteries available. Because, in this case, it is undesirable to use batteries in the system, all components must power down. However, the electrolyzer cannot simply be shut off instantly (see Section 2.4). An emergency shutdown would require technicians to service and reset the electrolyzer before it can operate again. Therefore, the electrolyzer's production should be gradually lowered until it shuts down or is sufficiently reduced in advance of the power outage. To predict when the power output from the PV-array will be drastically reduced, it is necessary to predict when shadows from clouds will hit the PV-array.

3.1 Predicting cloud velocity with triangulation

The TP consists of three PV-modules placed in an equilateral triangle with sides of length d_{PV} and three microcontrollers. Microcontrollers are used to log current values for each of the PV-modules, which will be uploaded to a database.

Figure 3.1.1 illustrates a simplification of the PV powered electrolyzer with the TP for cloud prediction. Here, the CS moves south-east towards the PV-array, as indicated by the red vector $\vec{v}$. In the TP, the shadow will cover PV-module q first, then p and r for this situation. Using the time difference between each the PV-modules shuts off, the speed and direction is calculated using Equations 2.6.1 and 2.6.3 respectively. The angle relative to north is calculated using Equation 2.6.4 using values for σ based on where the triangulation starts. Values for σ are given in Table 3.1.1.

Starts in point:	σ
q	30° if $t_r \leq t_p$, else 330°
r	210° if $t_q \leq t_p$, else 270°
p	150° if $t_q \leq t_r$, else 90°

Table 3.1.1: *Values for σ used in Equation 2.6.4.*

The calculations are done by accessing the logged data, filtering it and running a python script [21]. Since the CS in this case would cover the PV-array that supplies the electrolyzer, the measurements from the TP will be used to calculate the time until the PV-array is hit by the shadow. The time before the CS hits the PV-array, the *reaction time*, is based on the speed of the CS and the distance d.

The required reaction time depends on the time it takes for the electrolyzer to ramp down, and increases proportionally to d. However, the longer distance between the PV-array and the TP,

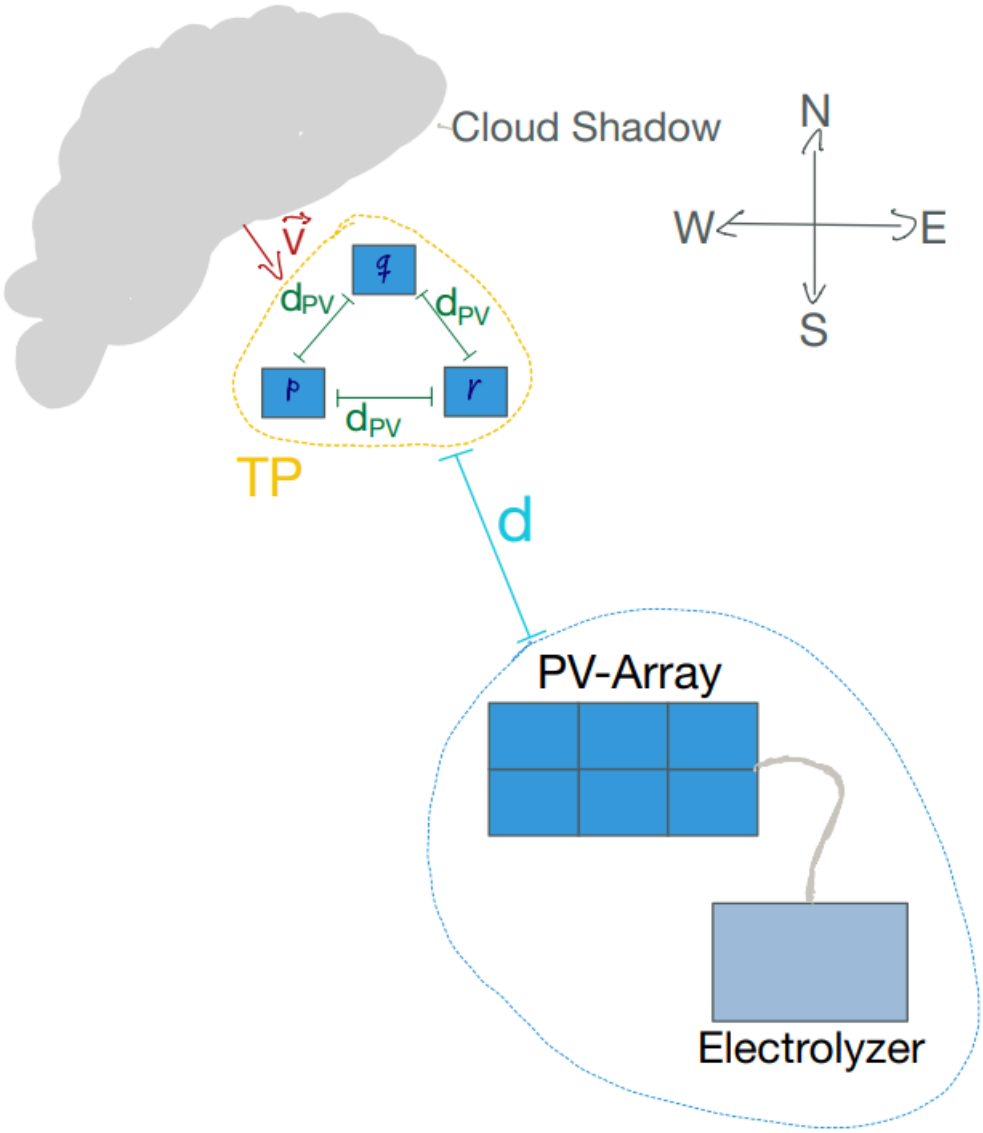


Figure 3.1.1: *Simple illustration of the electrolyzer powered by a PV-array with the triangulation point.*

the more time the cloud will have to change its shape, which can impact the reaction time and how the CS hits the PV-array. To simplify, it is therefore assumed that the cloud's shape does not change while moving from the TP to the PV-array. To test the accuracy of the TP, a 4th sensor will be stationed in the CS direction relative to the TP. The reaction time will be calculated at a point after the cloud has passed, making it possible to compare the theoretical results and the actual results.

When the cloud has passed the TP completely, all three PV-modules will turn on again. Assuming that the CS speed is constant between the TP and sensor 4, the time until the electrolyzer can turn back on will be the same as the reaction time. Knowing this can be utilized to maximize the production time of the electrolyzer.

In order to find the time difference (phase shift) between the sensors caused by the CS, one should look at the changes in power. This is because the sensors measures different power values because of the slightly different surroundings, yet the changes in power should be more or less the same for all of them. To find the rapid changes in power, the data is divided into intervals. These intervals M are chosen based on changes in power dP over a time period dt. The intervals are chosen from cases where the dP is higher than G and dt is less than 20 seconds. G represents a threshold value, and is chosen based upon how many intervals it generates. The intervals M can be written on mathematical form as in Equation 3.1.1.

$$M = [G \leq dP],\ [dt \leq 20s] \qquad dP, dt \in B \tag{3.1.1}$$

Here, B represents all the data for a day.

The easiest way to inspect the changes in power is to look at the power's gradients $\frac{d(\mathrm{W/m^2})}{dt}$. This can be done by calculating the difference in power between measurements, however, due to the rapid measurements, this approach leads to inconsistent values as illustrated in Figure 3.1.2.

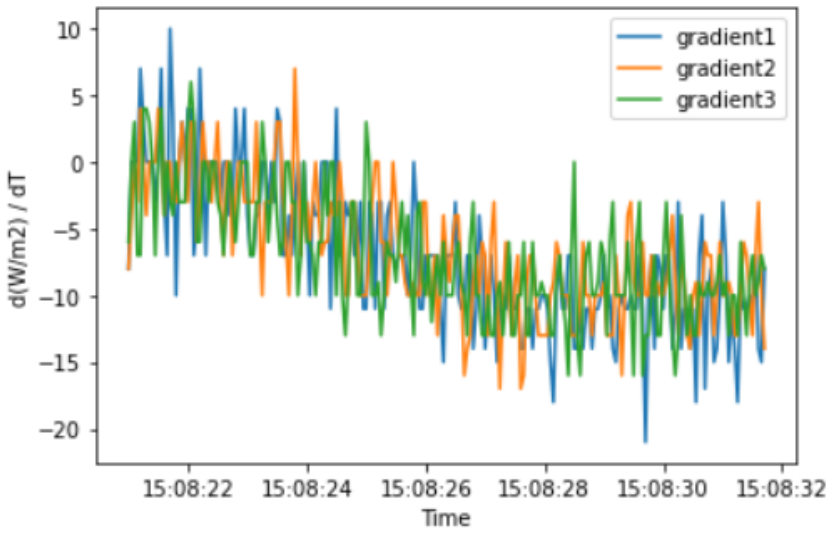


Figure 3.1.2: *Gradients of sensors 1-3 based on local changes in power.*

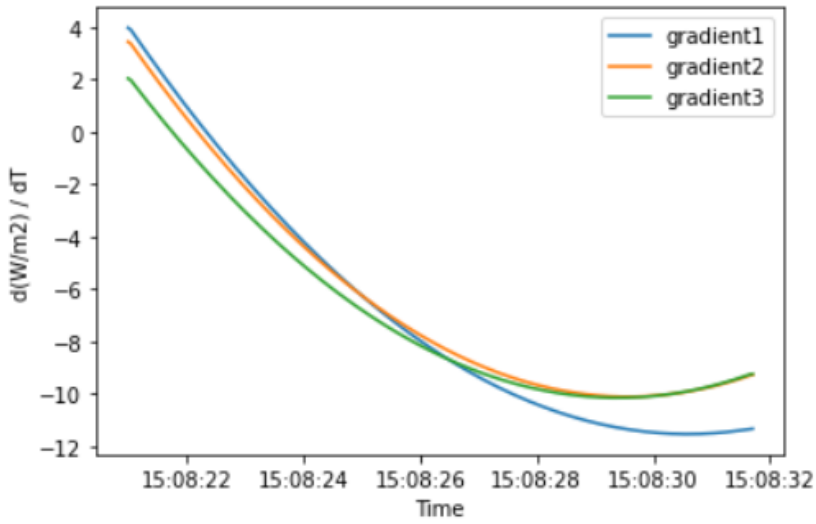


Figure 3.1.3: *Gradients of a third order polynomial approximating sensors 1-3.*

To avoid the inconsistent gradient values from looking at local changes, one can smooth out the curve by approximating it with third order polynomial regression. The regression converts the points to a continuous function, whose local derivatives are much more consistent. Applying the same approach, with local changes in power, now yields a smooth curve as shown in Figure 3.1.3. Using regression to approximate graphs can affect the data, however the regression done on the short intervals where the power increases or decreases on average yield a R^2-score higher than 0.995. A R^2-score close to 1 like this indicates that the third order polynomial regression is a good fit to the data, as Figure 3.1.4 illustrates.

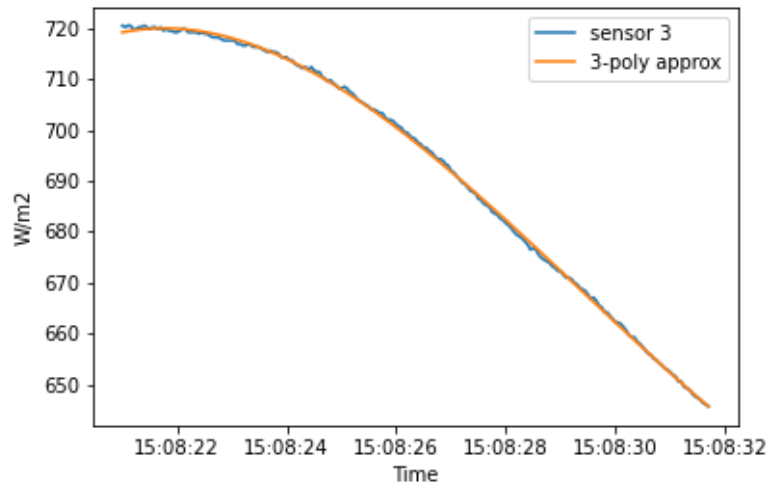


Figure 3.1.4: *Power values for a sensor, with the third order polynomial regression graph on top. These graphs are from the same interval as Figure 3.1.3. The R^2-score for this regression is higher than 0.999.*

The triangulation requires a time difference, or a phase shift, between the sensors in order to triangulate. To find this time difference, three methods were tested. The first method was to look at the top or bottom point of a curve, i.e. where the derivative of the function is zero, according to Function 3.1.2.

$$\frac{df(t)}{d(t)} = 0, \quad t \in M \tag{3.1.2}$$

Here, $f(t)$ is the third order polynomial approximation of the graph on a interval. These points should be slightly different from each other, and the points can be used as a reference to find the time difference between the sensors.

The second method consists of using the double derivative of the polynomial approximation of the graph. This is done by finding the point where the double derivative equals zero as reference point. Equation 3.1.3 shows the calculation.

$$\frac{d^2f(t)}{dt^2} = 0, \quad t \in M \tag{3.1.3}$$

The last method tested to find the time difference between the sensors, considers a relative change in the gradient. This method finds the point in time where the gradient curve has decreased by e.g. 40% and increased by 60%. This calculation is shown in Equation 3.1.4.

$$\frac{df(t)}{d(t)} = (\frac{df(t)}{d(t)})_{min} + ((\frac{df(t)}{d(t)})_{max} - (\frac{df(t)}{d(t)})_{min}) \times 0.6, \quad t \in M \tag{3.1.4}$$

3.1.1 Building the triangulation point

The triangulation point consists of three sensors in the form of small sensor boxes. There is also a reference pyranometer which is used to calibrate the sensors. The three sensor boxes are for measuring changes in irradiation and not necessarily the exact irradiation level. For easier referencing, the sensors are labeled 1, 2 and 3 which represents points q, r and p respectively in Figure 3.1.1. The sensors are pictured in Figure 3.1.5.

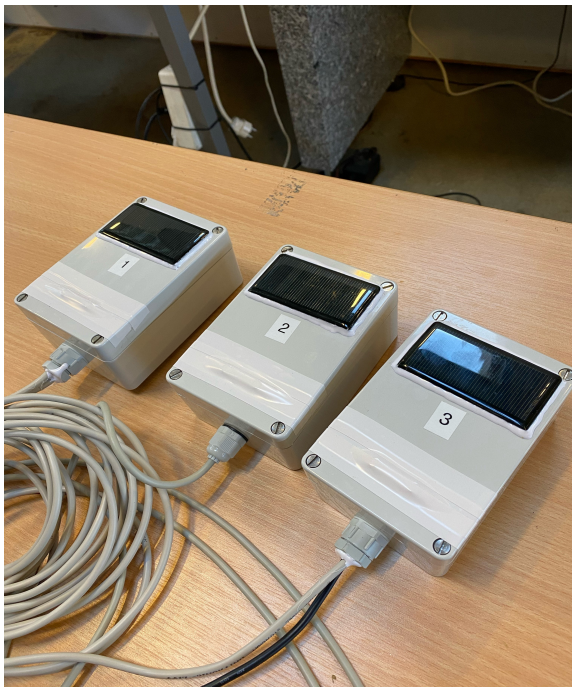


Figure 3.1.5: *Picture of the three sensors that are used in the triangulation point.*

All of the three sensor boxes are identical, consisting of a small PV-module on top of the cover, with a microcontroller and a GPS-tracker within the box. The microcontroller is of type *PHOTONH* produced by *Particle* [22]. The GPS module is a *746 GPS Module* by *ADAFRUIT INDUSTRIES*, with a *CR1220 Lithium coin* battery by *RS* [23][24]. The bread board was of type *RE994-S2* by *Roth Elektronik* [25]. There is also a shunt resistor of type *UPF25B50RV* by *TE Connectivity* [26]. A *5V 60mA 68x37mm Micro Mini Power Solar Cell* by *AMX3d* was used [27]. The box was a plastic box with 50x80x120mm dimensions. The circuit diagram in Figure 3.1.6 shows how all the parts are connected.

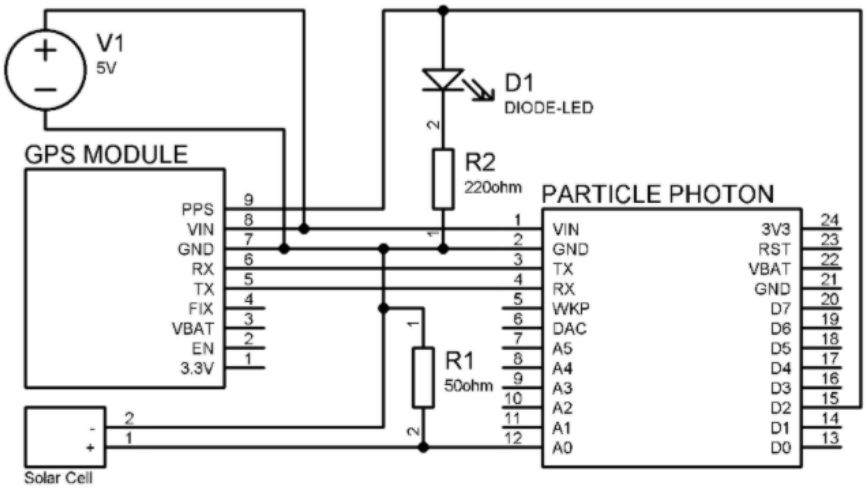


Figure 3.1.6: *A circuit diagram of the components that make up the sensors.*

The microcontroller measures the short circuit current of the PV-module, which is multiplied by a *calibration factor* in the microcontroller to estimate the irradiation. To get the best irradiance estimates, the calibration factor is calibrated using a series of lights focused through a square tunnel (see Figure 3.1.7) and using the reference pyranometer.

Figure 3.1.7: *Image of the light "tunnel" used to calibrate the sensors.*

It is important that the reference pyranometer and the small PV-modules are at the same distance from the light source during calibration because the irradiance varies depending on the distance from the light source. The microcontroller logs the data from each sensor at 20 times a second, and sends data at one second intervals to a database hosted by *Google Firebase* [28]. The purpose of the GPS-tracker is to synchronize the time in each of the three sensor boxes so that their measurements correspond to the same timestamps. To make the sensor boxes waterproof, there is a rubber gasket between the bottom of the box and the lid, while the PV-module is sealed around the edges by melted plastic. The three sensor boxes are placed in an equilateral triangle with a distance of 15 meters between each sensor box.

3.1.2 System placement

The placement of the TP relative to the PV-array is essential. If the PV-array is a large-scale power plant, it would make sense to have one TP in each of the four directions: north, south, east and west. Yet if the distance between the TPs is larger than the CS, it is theoretically possible that a cloud could pass between two TPs, making its way to the PV-array without being predicted. This would result in an undesired emergency shut down of the electrolyzer.

In this test, the TP is placed on the roof of a building at *The Norwegian University of Life Sciences* (NMBU) called "TF fløy 4". Its roof placement is due to the easy access to power and WiFi, and makes the TP less exposed to interference from people and shadows. Building an equilateral triangle on top of the roof proved more difficult than anticipated due to the uneven surface of the roof. Therefore, the triangle had to be rotated 180° so that it points southwards instead of northwards. The sensor points were measured up based on an image of the roof from Google Maps [29]. By measuring the width of the roof and measuring the image with a ruler, it was possible to make a *measurement coefficient*. Using the measurement coefficient, which was $2.08\frac{\text{m}}{\text{cm}}$, the sensor points were found by making a circle with a radius corresponding to 15m, with the south sensor as center. Furthermore, the distance between the two northern points were found by a 15m straight line touching the circle in both ends, with the center of the line going through a line from the center of the circle and the northern most part of the circle. A schematic of the TP placement relative to the roof is illustrated in the in Figure 3.1.8.

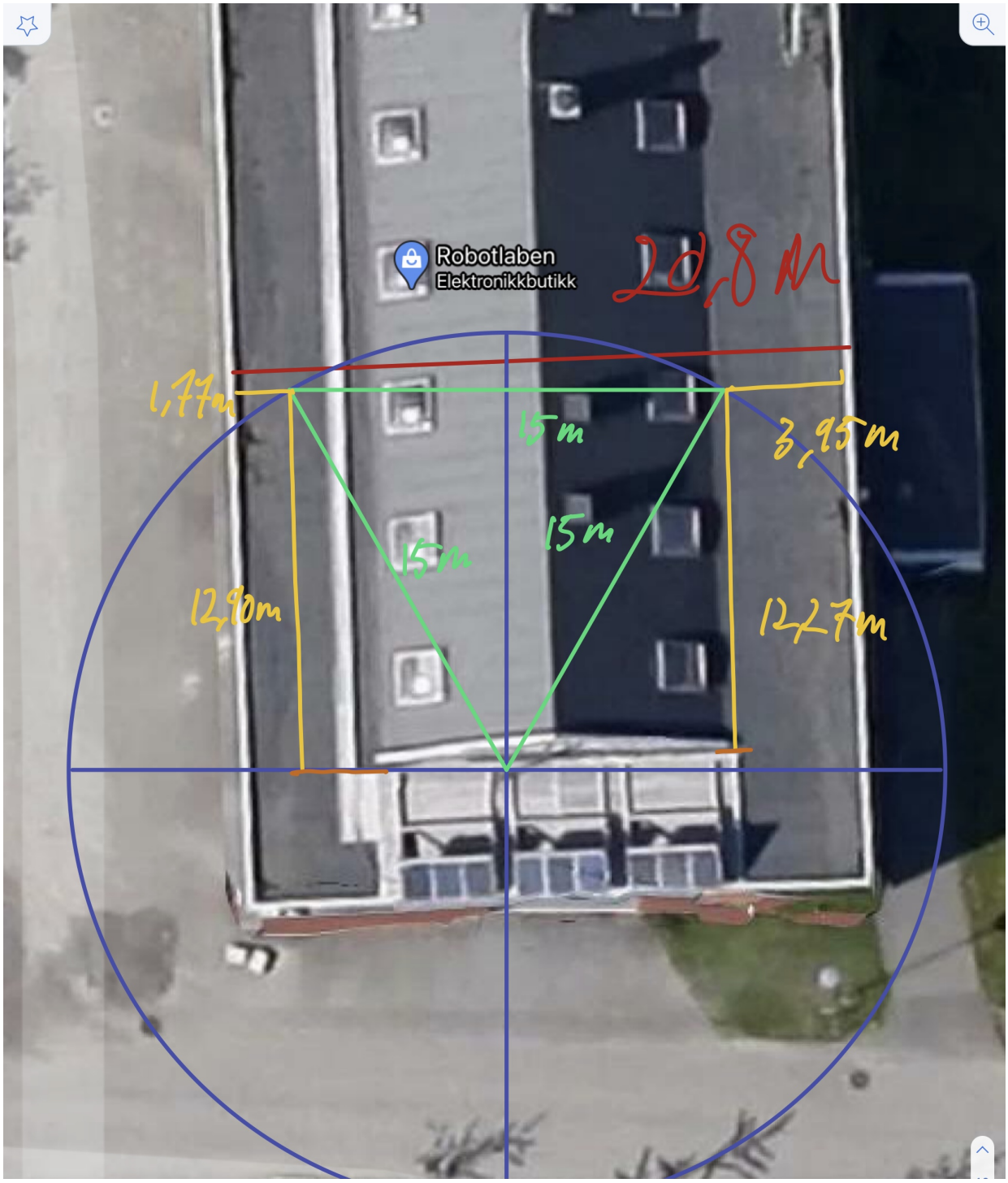


Figure 3.1.8: *Image of sensor placement on the roof of "TF fløy 4". The image is from Google Maps with distances drawn on top of the satellite image [29].*

Because the ruler used to measure lengths has an accuracy of 1mm and the measuring coefficient was $2.08\frac{\text{m}}{\text{cm}}$, the sensor placement points have an accuracy of ± 20.8cm. When the three sensor points had been found, the distance from the edges of the roof was measured. The measured

distances of the sensor positions were used to place the sensors. The accuracy of the measuring tape used to measure the distances was 1mm, yet because the measuring tape was only 5m long and the roof was uneven, a more realistic error margin is 1cm. There is also an error margin when placing the sensor on their position at about 1cm, which totals an error margin of ±22.8cm for the sensor placement. Figure 3.1.9 shows the sensors after placement.

Each of the sensor boxes is mounted on a wooden plank with an angle $50^\circ \pm 5^\circ$ relative to the horizontal plane (see Figure 3.1.9). For two of the sensors, the wooden plank is mounted on top of a 185cm ±0.5cm tall aluminum pole facing to the south end of the building. Each of the aluminum poles has three feet which are weighed down by a cinder block. For the last sensor, the wooden plank rests on top of a fuse box with the same elevation as the poles. The purpose of the elevation is to avoid the shadow from the higher parts of the roof. All three sensors are facing to the south end of the building with a margin of $\pm 5^\circ$. The building faces southwards, but not directly south, as is visible in Figure 3.1.8.

(a) *Image of sensor 1 (south sensor).*

(b) *Image of sensor 2 (east sensor).*

(c) *image of sensor 3 (west sensor).*

Figure 3.1.9: *Images of the three triangulation sensors after placement.*

The "PV-array dummy" called "sensor 4" will be moved to a location the CS will pass after hitting the TP, rather than being at a stationary point. This approach makes the system more flexible with respect to cloud direction, and is powered by a rechargeable power bank.

3.2 Bypassing the DC/DC converter

To find the optimal current the for the electrolyzer, it is easiest to inspect the expected irradiated power. The *Photovoltaic Geographical Information System* (PVGIS) has a database that covers 2005-2016 called "SARAH" [30]. Using the database in PVGIS, along with the optimal slope angle for a PV-module at 43° (in Ås), the expected irradiation can be simulated.

The expected radiation can be used to make a *Kernel Density Estimation* (KDE) plot, which gives information about both irradiation levels, and how frequently they occur.

The KDE plot can be used to find intervals where both time and power is maximized. A density plot, like a KDE plot, gives information of the distribution of values. The fraction an interval makes up of all of the data is given by the area under the curve in the interval, which can be calculated by Equation 2.8.2. By running iterations with intervals through the KDE plot and using Equations 2.8.1 and 2.8.3, it is possible to find the power intervals where it is most beneficial to bypass the DC/DC converter.

The optimal bypass interval length is based on the characteristics of the electrolyzer. By using a simplified load curve for the electrolyzer given by Nel, and simulating a PV-cell using methods described in "Solar energy" by Smets, the MPP and load relationship can be illustrated, as in Figure 3.2.1 [12][5].

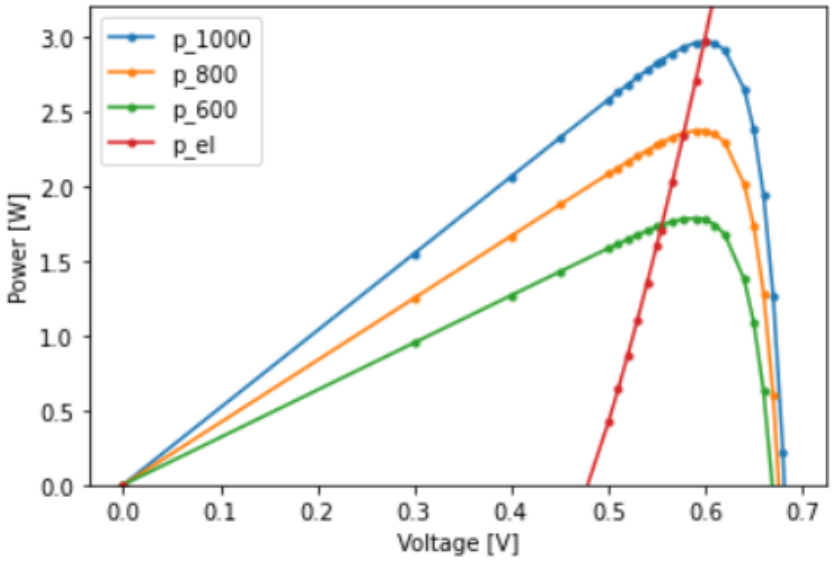


Figure 3.2.1: *Power curves for a PV-cell at different levels of irradiance with a load curve for a fraction of an electrolyzer.*

In Figure 3.2.1, the red load curve represents one millionth of an electrolyzer, while the power curves are for one individual PV-cell of type "Bistar TP6H60M" at different irradiated power levels [31]. The blue, yellow and green curves are the power graphs of the PV-cell at $1000\frac{\mathrm{W}}{\mathrm{m}^2}$, $800\frac{\mathrm{W}}{\mathrm{m}^2}$ and $600\frac{\mathrm{W}}{\mathrm{m}^2}$ respectively. The load curve of the electrolyzer follows the MPP reasonably well from $1000\frac{\mathrm{W}}{\mathrm{m}^2}$ to $800\frac{\mathrm{W}}{\mathrm{m}^2}$, which means that a bypass interval of length $200\frac{\mathrm{W}}{\mathrm{m}^2}$ is well suited.

It is assumed that the efficiency of the PV-array η_{PV} is 0.15, and that the efficiency of the DC/DC converter η_{DC} is 0.98 [32]. The amount of energy saved by bypassing the DC/DC converter also depends on the size of power interval where the bypass can be done. A bigger power interval saves more energy. It is assumed that the DC/DC converter bypass interval is

$200\frac{\mathrm{W}}{\mathrm{m}^2}$. When it is known what power is optimal for the DC/DC converter bypass, the electrolyzer should be customized so that it draws a current that matches the MPP for the PV-array when the DC/DC converter is bypassed. The optimal current also depends on the IV-characteristics of the PV-modules used in the PV-array. A simple circuit of the DC/DC converter is shown in Figure 3.2.2. The switch and related DC/DC converter activation is preferably controlled with hysteresis, so that it is not opened and closed too frequently, but opens quickly in response to clouds causing very sub-optimal operation points.

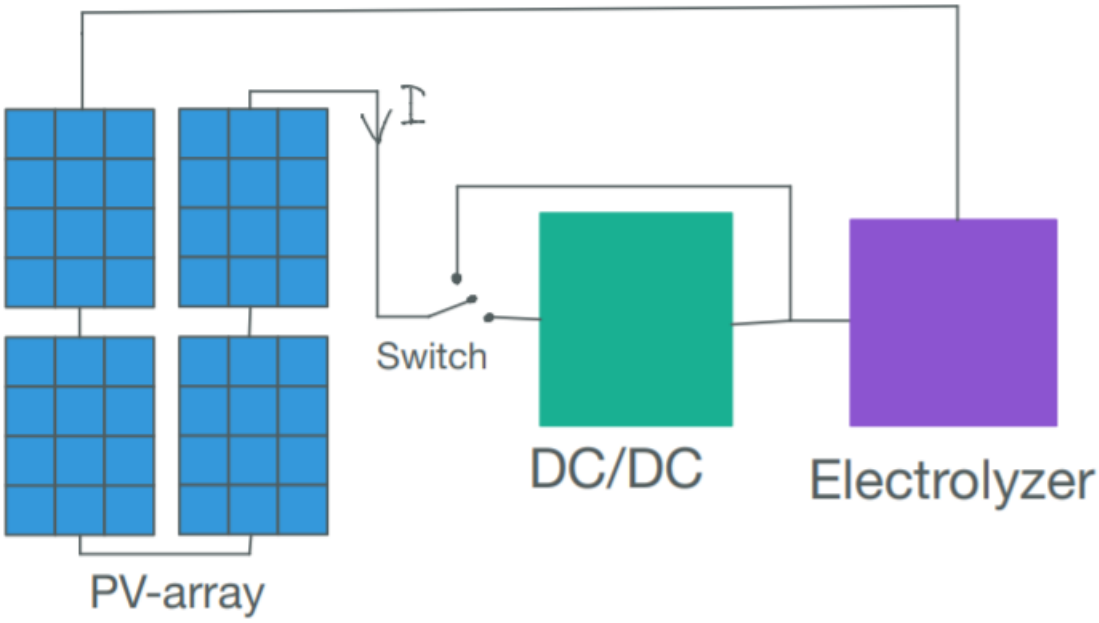


Figure 3.2.2: *A simple circuit diagram showing an external bypass of the DC/DC converter. The bypass is activated and deactivated by either en external or integrated switch in the DC/DC converter.*

4 Results

The results are based on measurements from three days. Information about these days are given in Table 4.0.1.

Date	Cloud conditions	Cloud direction	Sensor 4 distance
17/03 2021	Cirrus	South	1060m
22/03 2021	Cirrus	Southeast	1020m
13/04 2021	Cumulus	South	1120m

Table 4.0.1: *Table with information about days and conditions when measurements were done.*

4.1 Data and format

After downloading and unpacking the measurements for a day, the data is structured as a "Pandas dataframe" which is a data structure in python [33]. The dataframe consists of timestamps in two different formats and power values for all four sensors. An example of such a dataframe is illustrated in Figure 4.1.1. In addition to Pandas, NumPy was used during data processing [34].

	datetime	time	sensor 1	sensor 2	sensor 3	sensor 4
0	10:43:26	104326.00	1059.4	960.5	948.6	948.3
1	10:43:26.050000	104326.05	1059.1	960.5	948.2	949.0
2	10:43:26.100000	104326.10	1059.4	960.5	947.6	948.7
3	10:43:26.150000	104326.15	1059.4	960.5	947.9	949.0
4	10:43:26.200000	104326.20	1059.4	960.2	947.6	949.7
...	...	...	...	...	...	...
331175	15:19:24.750000	151924.75	744.5	678.0	672.3	643.3
331176	15:19:24.800000	151924.80	744.1	678.6	672.0	644.0
331177	15:19:24.850000	151924.85	745.2	677.3	672.0	643.6
331178	15:19:24.900000	151924.90	744.8	677.3	672.0	643.6
331179	15:19:24.950000	151924.95	744.8	678.6	672.0	643.3

331180 rows × 6 columns

Figure 4.1.1: *Image of the data after unpacking. The measurements in this figure are from 17/03/2021.*

This dataframe can be visualized as a graph by plotting the power value of each sensor to its corresponding point in time, as is shown in Figure 4.1.2. Notice that sensor 3 has a dip down to zero at about 13:45. This is because if the GPS signal or WiFi signal is lost during measurements, the values are not logged into the database. These values are recorded as

"not-a-number"s (NaN). NaNs are replaced with zeros in the dataframe. All graphs in the results are made in python, using the Matplotlib package [35].

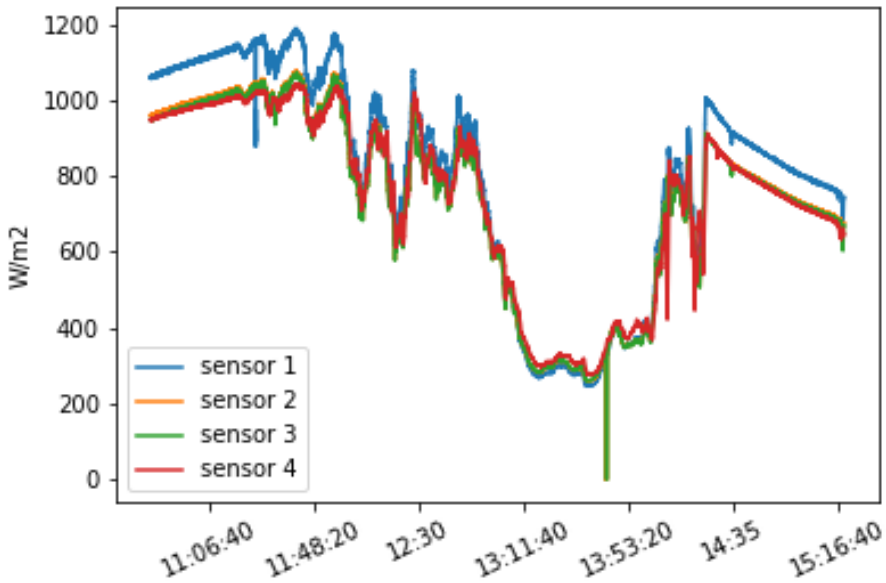
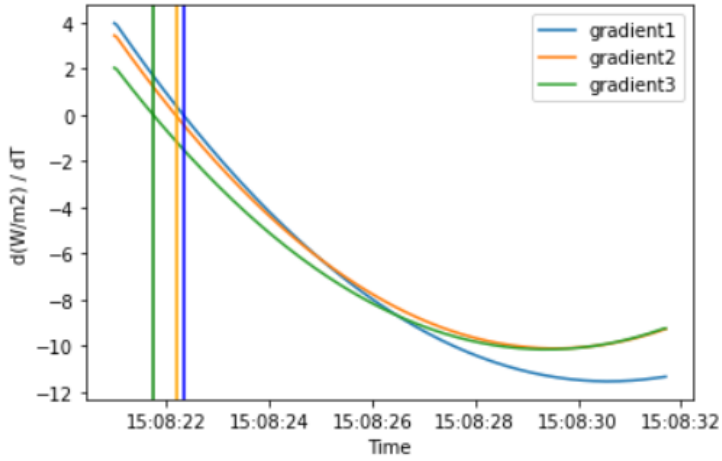


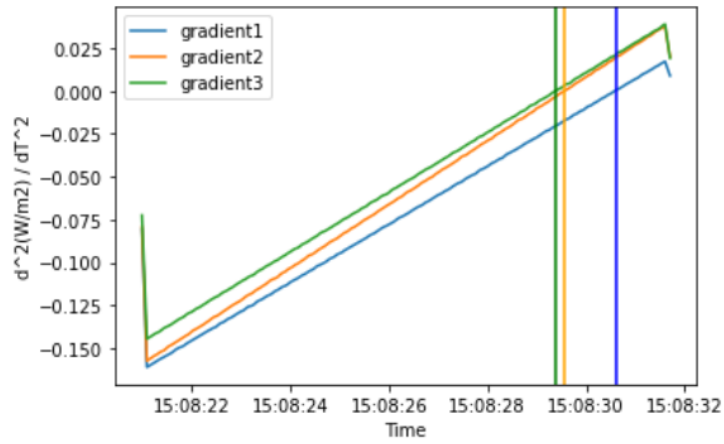
Figure 4.1.2: *Power measurements for all sensors from 10:43 to 15:19 17/03/2021.*

4.2 Time difference methods

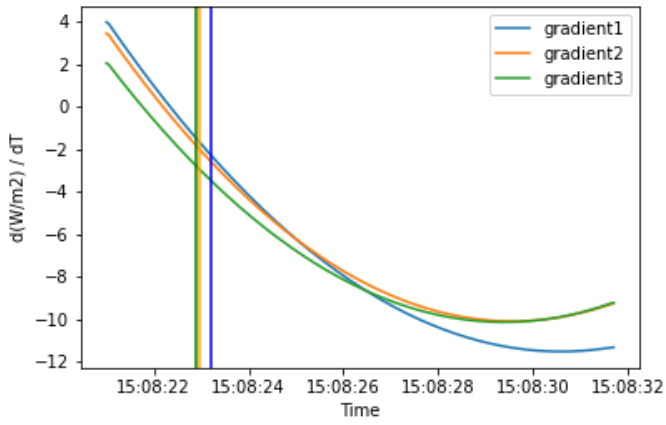
When looking at the results from the three different methods for finding the time difference (see Section 3.1), they yield similar results. The vertical lines indicate the points in time used as reference to find the time difference between the sensors. Figure 4.2.1 shows that the activation order is the same for all three methods, which is good because it indicates that these methods generally give similar results.



(a) *Reference points found by looking at where the gradient is zero (equation 3.1.2).*



(b) *Reference points found by looking at where the double derivative is zero (equation 3.1.3).*



(c) *Reference points found by looking at a relative drop in the gradient (equation 3.1.4).*

Figure 4.2.1: *Images of the three time difference methods on the same interval.*

However, the time difference between the sensors varies. Notice that in Figure 4.2.1 (a) the time difference between sensor 1 and 2 is higher and the time difference between sensor 2 and 3 is lower than for Figure 4.2.1 (b) and (c). The difference in time difference can impact how the velocity of the CS is triangulated. From testing, it appears that the methods depicted in Figure 4.2.1 (a) and (b) are more exposed to errors because the zero point is not always within the intervals defined by M (see Equation 3.1.1). Because of this, the favored and chosen method is the one depicted in Figure 4.2.1 (c) due to its flexible relative drop approach.

4.3 Estimated velocity

After importing the data and filtering out the intervals where there is a rapid change in power defined by M, the third order polynomial approximation to each sensor was calculated. Then the gradients of the polynomials were calculated and plotted. An example of this is illustrated in Figure 4.2.1 (c).

Table 4.3.1 shows the difference in time between the sensor activations. The table indicates that sensor 3 was activated first, followed by sensor 2 after 0.05s and lastly sensor 1, after 0.30s. The time difference is used in the triangulation to estimate both the speed and direction of the cloud. Figure 4.3.1 shows the output from the triangulation script using the values in Table 4.3.1.

Table 4.3.1: *Estimated relative time differences between sensor 1 to 3.*

Sensor	Relative time difference [s]
1	0.30
2	0.05
3	0.00

In Figure 4.3.1, the estimated speed was $46.7\frac{\text{m}}{\text{s}}$ with direction 171° north. The distance between the TP and sensor 4 was about 1060m and the cloud direction was observed to be southwards. With a speed of $46.7\frac{\text{m}}{\text{s}}$ it would take 22.7s for the cloud to reach sensor 4. To see how accurate the estimation is, it is best to use sensor 4 as a reference. Figure 4.3.2 shows the same interval as in Figure 4.2.1, which makes it easier to decide the actual time difference between the TP and sensor 4 at this interval.

```
triangulate(0.3, 0.05, 0)
```

```
C:0 | B:0.05 | A:0.3
Starts in C, angle:81.05, angle north:171.05, speed:46.66 m/s
```

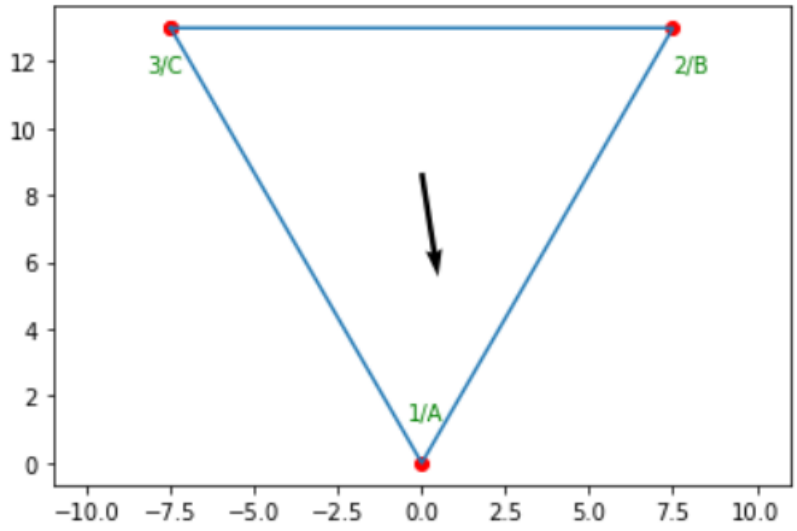


Figure 4.3.1: *Figure illustrating the output from the triangulation script. The figure shows both the estimated speed and direction of the cloud. The numbers and letters in the corners of the triangle represents the sensor number or letter.*

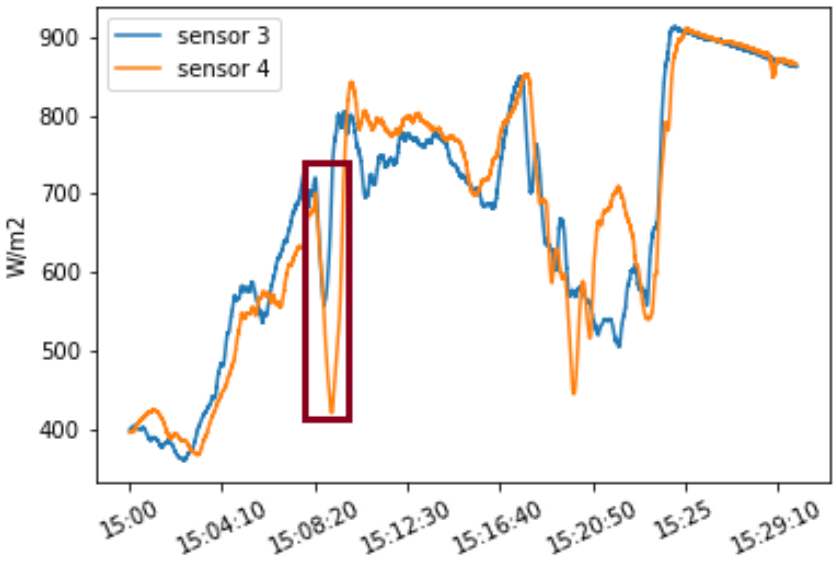


Figure 4.3.2: *Power curves for sensor 3 and 4 with a red box to indicate what interval that is inspected.*

Taking a closer look at this interval by zooming in, yields Figure 4.3.3.

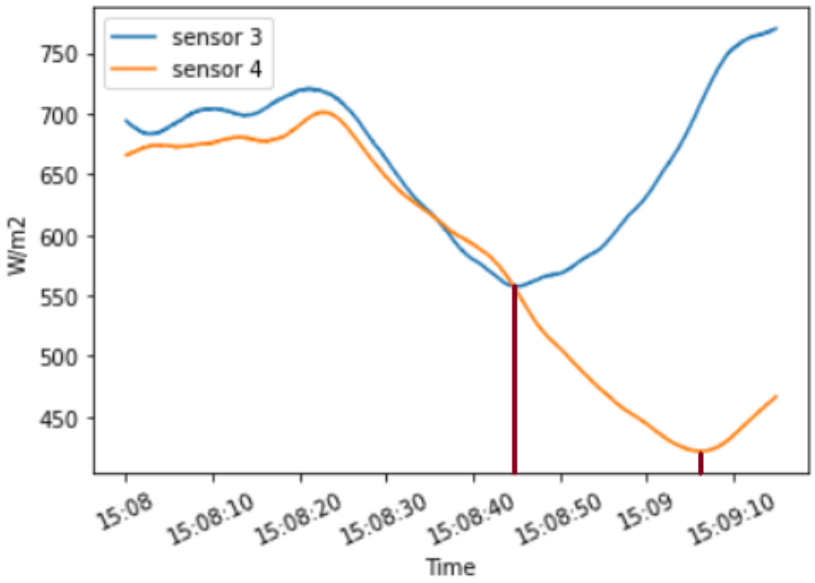


Figure 4.3.3: *A zoomed in image of the power curves with the minima of the curves marked with a red vertical line.*

By reading off the graphs, one can see that there is a time difference between sensor 3 and sensor 4 of approximately 20s. Here, the time difference between the sensors within the TP are neglected and it is assumed that the cloud moves in a straight line from the TP to sensor 4. Still, the model is usually not as accurate as in the previous example. Using the same approach at another interval in the same data set, the estimated time the CS uses to sensor 4 was 174 seconds, while in reality the time was about 35 seconds. A table describing the accuracy of each interval during the measurements of 17/03 2021 is given in Table 4.3.2. The Ω represents relative error in speed, meaning the estimated speed divided by the measured speed.

Table 4.3.2: *Estimated direction and speed the CS had from the TP to sensor 4, and the measured speed the CS had. Calculated using triangulation with measurements from 17/03 2021. The clouds were observed to move south. These intervals are defined by M using a G of 75W.*

Time interval	Measured speed $[\frac{m}{s}]$	Estimated speed $[\frac{m}{s}]$	Ω	Estimated direction
13:25:26 - 13:25:36	25	48	1.92	139°/SE
15:08:20 - 15:08:32	53	46	0.87	171°/S
15:08:46 - 15:09:02	53	13	0.25	120°/SE
15:09:00 - 15:09:06	39	59	1.51	263°/W
15:09:06 - 15:09:20	39	88	2.26	120°/SE
15:17:38 - 15:17:52	42	20	0.48	172°/S
15:19:32 - 15:19:44	35	71	2.03	106°/E
15:23:45 - 15:23:53	30	14	0.47	248°/SW
15:23:53 - 15:24:00	30	14	0.47	46°/NE
15:24:00 - 15:24:15	30	30	1.00	143°/SE
Average values	38	39	1.13	165°/SE
Standard deviation	9	27	0.71	66°

Table 4.3.3: *Estimated direction and speed the CS had from the TP to sensor 4, and the actual speed the CS had. Calculated using triangulation with measurements from 22/03 2021. The clouds were observed to move southeast. Intervals are defined by M using a G of 200W.*

Time interval	Measured speed $[\frac{m}{s}]$	Estimated speed $[\frac{m}{s}]$	Ω	Estimated direction
09:55:42 - 09:55:53	20	2	0.10	312°/NW
10:03:33 - 10:03:48	19	73	3.84	134°/SE
10:04:05 - 10:04:15	19	60	3.16	203°/SW
10:06:36 - 10:06:47	15	23	1.53	98°/E
10:06:50 - 10:07:06	15	10	0.67	174°/S
10:10:47 - 10:10:55	20	60	0.33	143°/SE
10:11:00 - 10:11:14	20	57	2.85	191°/S
11:32:34 - 11:32:53	17	5	0.29	179°/S
11:42:32 - 11:42:51	17	9	0.53	91°/E
13:22:10 - 13:22:29	19	18	0.95	110°/E
13:37:49 - 13:38:09	16	8	0.50	118°/SE
13:39:43 - 13:39:53	16	16	1.00	143°/SE
Average value	18	28	1.31	158°/SE
Standard deviation	2	25	1.21	58°

The measured speeds in Table 4.3.2 and 4.3.3 were calculated using the approach illustrated in Figure 4.3.3 to find the time with a distance of 1060m to sensor 4. The letters in the "estimated direction" column in table 4.3.2 represents the direction, where N is north, E is east, NE is northeast etcetera. Similarly to Table 4.3.2, the estimates for 22/03 2021 are given in Table

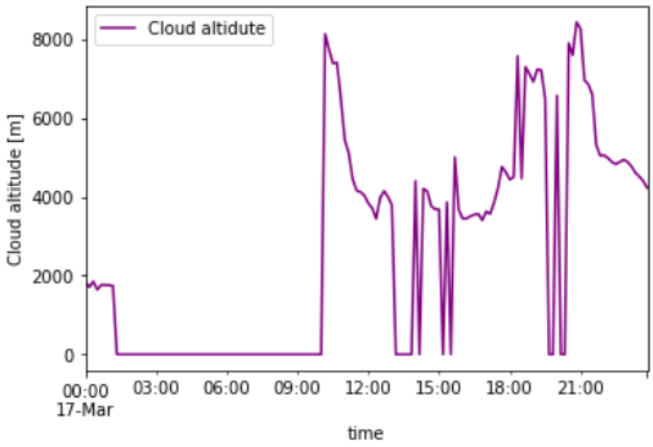
4.3.3, with a distance of 1020m to sensor 4.

The measurements from both 17th and 22nd of March 2021 were made during weather with a challenging kind of clouds, cirrus (see Section 2.5).

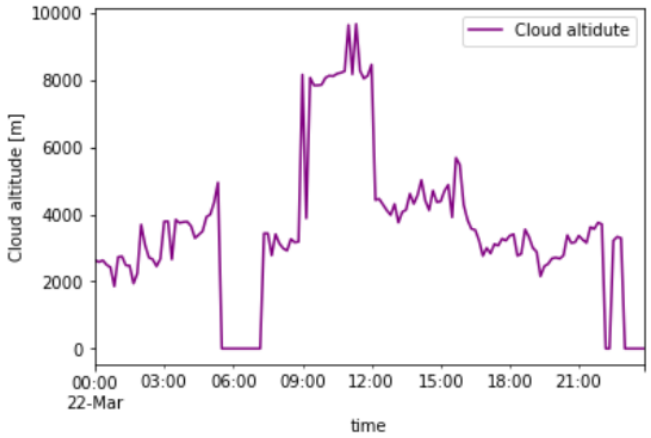
The 13th of April however, had more preferable weather, with cumulus clouds. An image of the clouds from 13/04 2021 are given in Figure 4.3.4. Measurements for 13/04 2021 are given in Table A.0.1 in the appendix with values from Table A.0.2 as reference.

Figure 4.3.4: *An image of clouds from 13/04 2021.*

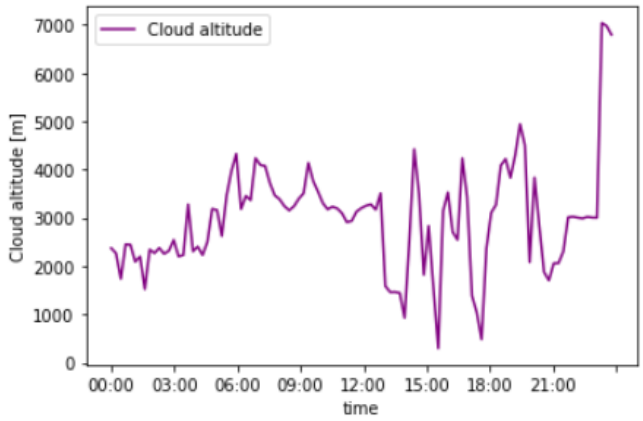
The Meteorological Institute in Oslo measures cloud altitude using a ceilometer, and using their data, the cloud altitude for the 17th and 22nd of March and 13th of April have been plotted in Figure 4.3.5.



(a) *Cloud altitude during 17/03 2021.*



(b) *Cloud altitude during 22/03 2021.*



(c) *Cloud altitude during 13/04 2021.*

Figure 4.3.5: *Figure illustrating the altitude of the lowest cloud layer during the 17th (a) and 22nd (b) of March, and 13th of April 2021.*

In Figure 4.3.5, the parts where the graphs are at zero are due to there being no clouds, and

thus the values are set to zero. For the intervals where there is a registered cloud altitude, the average altitude was calculated for all days, being 4.8km the 17th, 4.1km the 22nd and 3.0km the 13th. Note that these measurements are done in Oslo, more than 25km north of the TP.

4.4 DC/DC converter bypass power saving

Based on the data from the 12 years in the SARAH database, with a PV-module with a 43° slope, a KDE plot for the simulated irradiation on the PV-module was made. The KDE plot is illustrated in Figure 4.4.1. The total yearly production is estimated to be about $230\frac{\text{kWh}}{\text{m}^2}$.

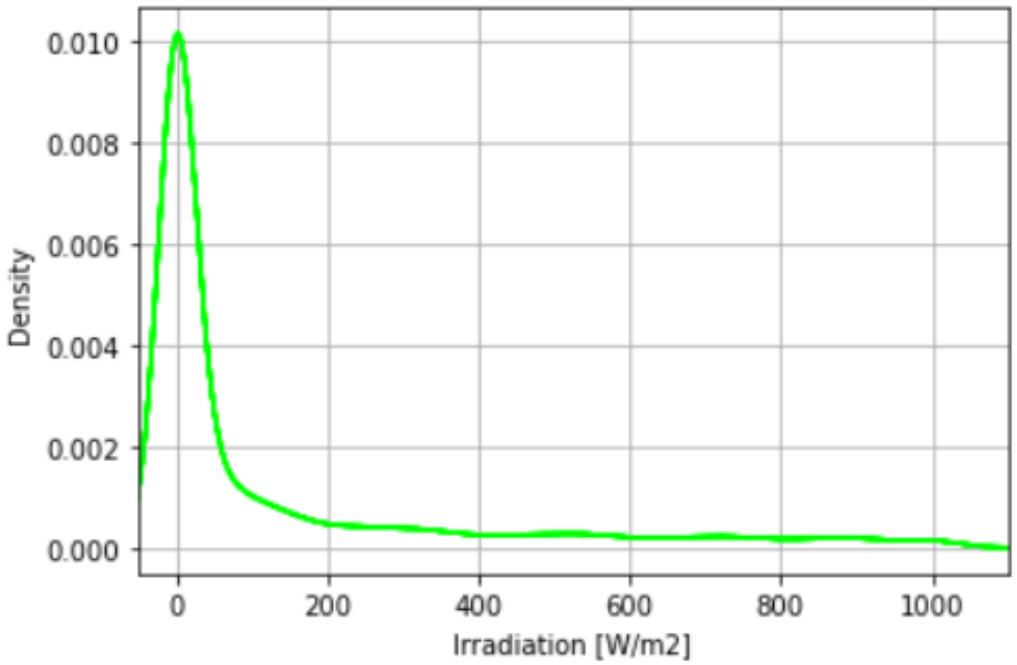


Figure 4.4.1: *A Kernel Density Estimation plot of the power distribution based on a PV-module simulation using 12 years of data.*

Using the data from SARAH and the KDE plot, power intervals from $0\frac{\text{W}}{\text{m}^2}$ - $1000\frac{\text{W}}{\text{m}^2}$ with lengths of $200\frac{\text{W}}{\text{m}^2}$ were iterated through. Each iteration calculated the power saved for each interval (Equation 2.8.3) and the area under the curve (Equation 2.8.2). The results of these calculations are given in Figure 4.4.2.

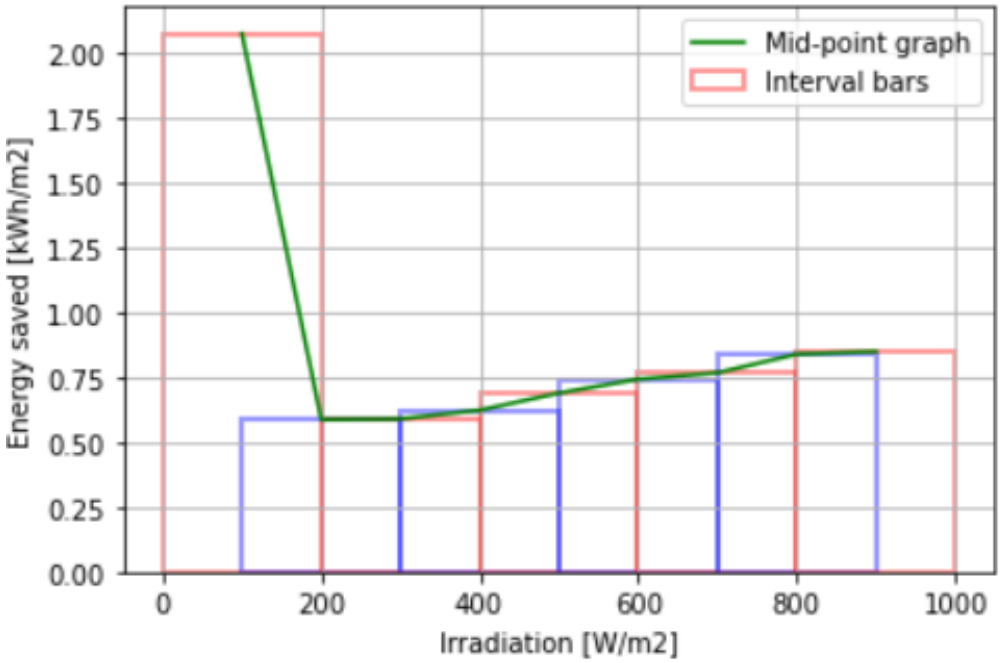


Figure 4.4.2: *Graph illustrating the potential saved energy by bypassing the DC/DC converter at different power intervals of length 200$\frac{\mathrm{W}}{\mathrm{m}^2}$, with a step size of 100$\frac{\mathrm{W}}{\mathrm{m}^2}$. The intervals are marked by bars that alternate between blue and red. The middle of each interval was used to plot the trend line, given by the green line.*

The three power intervals where most energy could be saved are given in Table 4.4.1.

Energy saved [$\frac{\mathrm{kWh}}{\mathrm{m}^2}$]	Interval [$\frac{\mathrm{W}}{\mathrm{m}^2}$]
2.074	0 - 200
0.849	800 - 1000
0.838	700 - 900

Table 4.4.1: *Energy saved per year per square meter of PV-array by bypassing the DC/DC converter at different intervals of 200$\frac{\mathrm{W}}{\mathrm{m}^2}$.*

5 Discussion

5.1 NaN treatment and sensors

The three sensors in the TP and sensor 4 all worked as intended for the most part. The instances where data were missing were replaced by zeros, which could lead to artificial drops detected by the interval algorithm for M. However, this was circumvented by excluding any intervals where the power was exactly equal to zero. An alternative to replacing the NaNs with zeros was to interpolate. Interpolation could lead to better results because the intervals of missing data usually are only one second long due to the one second data upload rate. Yet, the interpolation would interfere with the data if there was a rapid change in the power. Another reason that interpolation was left out is that because of the structure of the data, the interpolation would be computationally heavy due to iterations through all of the data. The ideal solution to this problem would be to have sensors that do not lose connection, by e.g. using a cable instead of wireless transmission or uploading data at a higher rate. This will be discussed in higher detail in Section 5.4.

From Figure 4.1.2 and several other figures that show the sensor power, it is clear that sensor 1 registers a higher power than sensor 2, 3 and 4. As all sensors were placed with the same angle, direction and height, and have the same sensitivity, they should all yield the same power when there is sun. This is not the case due to the placement of sensor 1. In Figure 3.1.9 (a) one can see that behind sensor 1 there is a vertical pole of metal. It is believed that the pole is the cause for the higher power picked up by sensor 1, caused by reflecting more light to the sensor from the pole. This was tested, and when moved further away from the pole, it registered a power value more similar to the other sensors. Unfortunately, there were no other place to place the sensor. The pole worked in the opposite way when there was less light outside, probably by blocking ambient light. To solve this problem, it would be best if all the sensors were placed in an open field without any structures to intervene. However, such an area was not available during the tests for this thesis.

5.2 Time difference methods problems

The three different methods used for calculating the time difference between the sensors in Figure 4.2.1 gave similar results. Because the zero point of the gradient is not always within the interval M, the methods in Figure 4.2.1 (a) and (b) were less favorable. Another value than zero could be used as a reference point, but there is no guarantee that any specific value is within M, which is why the relative gradient drop method was used. However, there are challenges with the relative gradient drop as well, as Figure 5.2.1 illustrates.

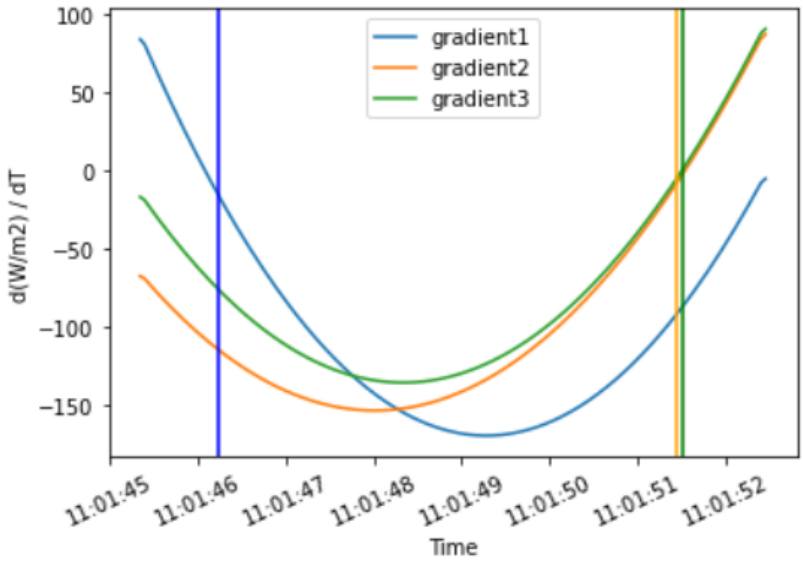


Figure 5.2.1: *Figure illustrating an instance where the relative gradient drop method struggles.*

In Figure 5.2.1, one can tell by looking at the minima of the graphs that the order the sensors were exposed to a change in power should be sensor 2 followed by sensor 3 then sensor 1, yet sensor 1 was registered as the first one. Intervals that are symmetric and have a clear minimum or maximum point make the reference points harder to locate because some power gradient values are located at two points on the y-axis instead of just one. To circumvent this, the algorithm was changed to find the first point where the power dropped by 40% or increased by 60%. In the rare case of Figure 5.2.1, sensors 2 and 3 only have one point at 60% of maximum value, while sensor 1 has two. The first point in sensor 1 is selected, however the same point for sensor 2 and 3 are outside the interval. This results in the algorithm setting the wrong reference point. This has major consequences for the results, as sensor 1 is the southern sensor. While the CS is moving south as is the case here, the wrong activation order of the sensors will predict that the cloud is moving north. This 180° error in direction pollutes the results, giving the model a worse accuracy at both speed and direction. It is hard to tell how often this problem occurs, but it is believed to be the the biggest prediction error factor. Luckily, it is possible to solve this problem by altering the algorithm for finding the sensor reference point, but the problem has yet to be solved.

5.3 System accuracy

The “measured speeds” in Table 4.3.2, 4.3.3 and A.0.2 are based on the assumption that the cloud moves in a straight line from the TP to sensor 4 at a constant speed. It is improbable that this assumption is accurate for an entire day, as sensor 4 is placed at the start of the measurements for a day and remains stationary. If the CS changes direction, it will impact the time difference between the TP and sensor 4 and thereby the measured time.

The results from Table 4.3.2 shows that there is a lot of variance in the estimations, as the standard deviations indicate. However, on average, both the speed and direction are fairly correct. The clouds were observed to move south, and the estimated direction was on average southeast. At some intervals, the estimates misses by a lot. From intervals at 15:09:06 - 15:09:20 and 15:19:32 - 15:19:44 the speeds were estimated to be $88\frac{\text{m}}{\text{s}}$ and $71\frac{\text{m}}{\text{s}}$ respectively, which is more than twice the measured speed. The estimated speed generally varies a lot, which is thought to be the result of the challenging cirrus clouds and "bugs" in the algorithm. The error margin from the sensor placement is relatively small, which is why it is thought to have little effect on the results.

Table 4.3.3 also have high standard deviations, but although the average speed is less correct than in Table 4.3.2, the direction is more accurate. This might be a result of the difference in CS speed or a difference in clouds. The cloud height could also influence the results, still, the average cloud altitude for both days were relatively similar. There is no obvious correlation between the cloud altitude and the estimated speed and direction for the respective days.

The 13th of April had "good clouds", meaning cumulus clouds which are dense and have distinct edges (see Figure 4.3.4). The cumulus clouds yields more rapid changes in power, which lead to more intervals than for 17/03 2021 and 22/03 2021 as seen in Table A.0.1. Most of the intervals had an estimated speed between 1 - $9\frac{\text{m}}{\text{s}}$, yet because of some outlier values, the average speed was $15\frac{\text{m}}{\text{s}}$. The highest speed in Table A.0.1 was at the interval from 12:32:15 - 12:32:35, where the speed was $130\frac{\text{m}}{\text{s}}$. This interval is depicted in Figure 5.3.1.

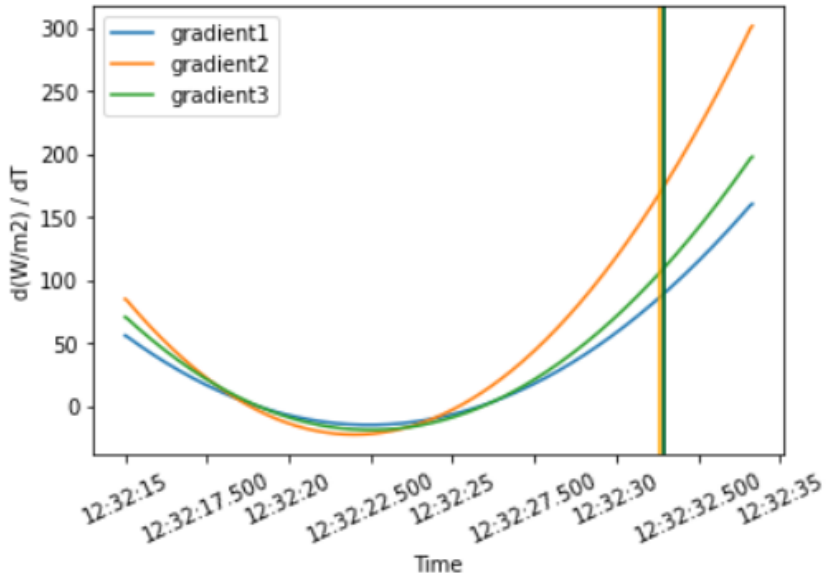


Figure 5.3.1: *Interval where the gradients look normal, but results in a high estimated speed.*

In Figure 5.3.1, the gradients look normal, but the activation point for each sensor are at almost

the same position in time. This results in that the model estimates a high speed because the time is short, and the distance is fixed. Unlike the problem in Section 5.2, which results in a low speed and is easy to see, this problem is more difficult to detect and yields very high speeds. An alternative time difference method could potentially solve this problem. Intervals which are similar to the one in Figure 5.3.1 are the main cause for outlier values with high speeds.

Although the average cloud altitude was lower for 13/04 2021 as Figure 4.3.5 indicates, the accuracy is roughly the same as for 17/03 and 22/03 2021. This might imply that the cloud altitude has little impact on the accuracy of the model, as using the CS itself factors in both altitude and speed.

5.4 Potential improvements

When building the TP, it would be preferable to use a large open field and increase the distance between the sensors. The open field would make it easier to have all sensors equally power sensitive and make them register close to equal power levels. This would circumvent the problem discussed in Section 5.1 where sensor 1 registers a higher power than sensor 2 and 3. Increasing the distance of the sensors would increase the time difference between the activations of the sensors, leading to more accurate speeds and directions. The increased distance between the sensors could also reduce the amount of intervals with problems such as the one described by Figure 5.3.1. To maximize precision, all distances should be measured by laser distancing sensor when building the TP. The extra distance between the sensors will make the activation point of each sensor further apart in time, which means that unless the length of time in M is increased, the activations of each sensor might not be included in the same interval. However, if the length of M increases, G also need to increase unless a large amount of intervals are desired. To increase the robustness of the TP, additional sensors could be added. This would require tweaks to the triangulation algorithm, but should reduce noise and give more robust information about the cloud shadow. The resolution, or, the sample rate of the data in this thesis are limited to 0.05 seconds. An alternative to increasing the distance between the sensors in the TP could be to increase the resolution. Yet, this would greatly increase the amount of data that needs to be logged and processed, and might not be as robust as simply increasing the distance between the sensors in the TP.

To avoid “holes” in the form of NaNs in the data, the data upload rate could be increased from one time per second. In addition, a stronger WiFi and GPS signal would be beneficial. This could be done by using a larger antenna or directly connecting the sensors by cable. The cable option might be more convenient in a stationary system. Still, the sensors are relatively energy effective. During the measurements, the sensors used about 2500mAh per hour when logging. This means that if the sensors are equipped with a decently sized battery and has the option to charge the battery when it is not logging, it could potentially work completely wireless.

The different time difference methods that are used have their strengths and weaknesses. To

increase robustness and possibly the accuracy of the estimations by the model, one could use an ensemble of time difference methods. Using an ensemble of the three time difference methods mentioned in Section 3.1, or more, could pick up more information. The ensemble could use a majority voting system to chose the direction and speed which is most probable. Using a majority voting system requires there to be an odd number of algorithms in the ensemble. Ensembles of algorithms are commonly used in machine learning to capture more information in addition to being more general.

When implementing the TP system for an actual PV-array, it would be safest to have one TP in each of the four directions: north, east, south, west. Alternatively, if the PV-array is small, the three TPs could be placed in an equilateral triangle around the PV-array. In certain places, like a valley, there is a possibility that the clouds generally only move in one or two directions. If one is certain of this, fewer TPs could be used. The spacing between each TP is important, because if it is too big, a CS could slip between two TPs and reach the PV-array without being detected. This could lead to undesired results. Using several TPs around the PV-array in different configurations could lead to some interesting advantages. If the TPs communicate with each other, the configuration of the TPs themselves could also be used to triangulate. In this way, the TPs can filter out internal movements within the cloud and rectify measurement deviations where the edge of the cloud is not perpendicular to the direction of travel, leading to more accurate results.

The triangulation algorithm considers shading of the PV-array as an "on/off" situation. However, if a CS covers the PV-array, not all power will be lost. Since the power only will be drastically reduced, the electrolyzer does not need to do a full shutdown. This means that if a CS causes a 60% power drop from the PV-array, the electrolyzer only need to reduce production by 60%. Because the sensors in the TP consists of small PV-cells, they can detect not only a reduction in irradiated power, but also how much. By tweaking the algorithms used in the triangulation, it should be possible for the TP to also predict the absolute drop in irradiated power, which can be used to further increase the efficiency of the system.

5.5 Comparison to other work

This is not the first time somebody tries to predict a CS speed and direction. In an article called *Cloud shadow speed sensor*, a similar problem to the one in this thesis is described [36]. The system in the article uses nine sensors in a semi circle. The system discussed in the article is much more compact to the one discussed in this thesis and seems to yield more accurate results. Although the results appear to be better in the article, it is hard to compare the results in this thesis with the one in the article. This is because the measured speed and direction is estimated in this thesis, while in the article, an object is used to create a shadow for reference. To use a similar approach for reference in this thesis would be challenging due to the scale the of the object needed to cast a shadow over the TP. However, the system discussed in the article is

much more complex both in terms of hardware and software.

As mentioned in Section 2.7, an alternative to using illumination sensors could be using cameras and machine learning. In an article by Wieland, Li and Martinis, they use machine learning on spectral images from satellites to inspect clouds and cloud shadows [37]. Using a similar approach to the one in the article, with some adjustments, could be used to predict CS speed and direction. An advantage of a system consisting of a camera is that the system would be a lot more compact. Yet, such a system might require a lot more data storage depending on the image resolution and image capture rate. In addition, the camera might struggle to properly detect cloud velocity if there are several layers of clouds moving at different speeds.

5.6 Bypassing the DC/DC converter

Figure 4.4.1 gives information about how often different irradiation values occur over the course of a year. It is clear that most of the year, the irradiated power lies at 0-70 $\frac{W}{m^2}$. This is most likely because about half of all the power values are zero, as the data includes the entire year, of which about half is night. As expected, higher powers occur more scarcely, as great weather conditions are relatively rare.

The graph in Figure 4.4.2 illustrates how the potential energy saved varies with the irradiation. In Figure 4.4.2 and in Table 4.4.1, it can be seen that the highest amount of energy can be saved by bypassing the DC/DC converter when the irradiated power is between $0\frac{W}{m^2}$ and $200\frac{W}{m^2}$. The energy saved on this interval is 2.074kWh per square meter of PV per year. Yet, this interval is unrealistically highly weighted as it includes power values from nights as well, where the power is zero. This is why the interval of $0\frac{W}{m^2}$ - $200\frac{W}{m^2}$ should be ignored. The interval where realistically the most energy can be saved, is from $800\frac{W}{m^2}$ to $1000\frac{W}{m^2}$. At this interval, the potentially saved energy is 0.838kWh per square meter of PV per year. This corresponds to less than 0.4% of the total energy generated. However, for a large solar power plant with electrolyzers, this is still a considerable improvement of either production output or reduction of installation cost. The most efficient power interval of $800\frac{W}{m^2}$ to $1000\frac{W}{m^2}$ might indicate that the energy saved is more dependent on high irradiated power than the occurrence of high irradiated power. This claim is strengthened by how the little the graph in Figure 4.4.1 descends from $200\frac{W}{m^2}$ to $1000\frac{W}{m^2}$. The graph shows that there is little variance in how often higher irradiance values occur.

The optimal interval at $800\frac{W}{m^2}$-$1000\frac{W}{m^2}$ can be used to maximize production gain from the PV-powered electrolyzer. By looking at the IV-curve for the PV-module used in the PV-array, it is possible to find the MPP for the PV-module when the irradiated power is at about $900\frac{W}{m^2}$. When the MPP for the PV-module under these conditions is found, the current at the MPP I_{MPP} is the current the electrolyzer should be customized for. When the electrolyzer is customized with respect to I_{MPP}, the electrolyzer will work as as an MPPT when the irradiation is between $800\frac{W}{m^2}$ and $1000\frac{W}{m^2}$ and the DC/DC converter is bypassed. The increased yield in produced hydrogen depends on the amount of electrolyzers and the scale of the

PV-array, but can be calculated by Equation 5.6.1.

$$Prod_{inc} = 0.849 \frac{\text{kWh}}{\text{m}^2} \times A_{PV} \times \frac{1\text{N}\,\text{m}^3\,\text{H}_2}{4.4\text{kWh}} \tag{5.6.1}$$

Here, $Prod_{inc}$ is the increase in hydrogen production per year by bypassing the DC/DC converter when the irradiated power is between $800\frac{\text{W}}{\text{m}^2}$-$1000\frac{\text{W}}{\text{m}^2}$ and A_{PV} is the total area of installed PV (assuming an efficiency of 0.15). The last part of Equation 5.6.1 is the efficiency of the electrolyzer, which in this case is assumed to be the most efficient one from Nel (see Section 2.4). Using a larger power interval than $200\frac{\text{W}}{\text{m}^2}$ would increase the gain from the DC/DC converter bypass, but the size of the power interval depends on how stabilizing the electrolyzer current works on the power grid. Also, a too large interval would most likely not be beneficial because the MPP could vary so much that the gain from bypassing the DC/DC converter does not make up for it. Another challenge with the DC/DC bypass is that it will be less efficient if there are high variation in the temperature where the PV-array is. This is because the MPP moves down to a lower point if the temperature rises. This could impact the how well the electrolyzer works as an MPPT, even at short intervals.

6 Conclusion

The goal of this thesis was to find out if triangulation of cloud shadows can be used to predict cloud shadow speed and direction. In addition, a DC/DC converter bypass was looked into, to increase system efficiency. The triangulation was done by using three small PV-sensors positioned in an equilateral triangle, with sides of 15m. To detect rapid changes in power, the gradient of the power curves from each sensor was used on an interval. To avoid noise in the gradients, the curves were approximated using third order polynomial regression. The intervals were found by looking at a relatively high change in power over 20 seconds. Several time difference methods were tired, but only one was used for all days. The chosen time difference method consists of looking at a relative change in the gradients over the interval.

In general, the system yields fairly accurate results. For all three days measurements were done, the average speed and direction was more or less correct. However, due to some outlier values, the model struggles with some clouds based on different factors. There is room for improvement in the triangulation point. This can be done by placing the TP in an open field and increasing either the distance between the sensors or increasing the data sampling rate. Equipping the sensors with an antenna or connecting them by cable would increase their reliability and reduce data loss. The time difference algorithm used is not perfect, and could be improved by solving the problems discussed in Section 5.2, or using an ensemble of time difference algorithms. Using an ensemble of time difference algorithms should increase both the accuracy and robustness of the model.

With some minor adjustments to both the TP itself and the algorithms used, the system discussed in this thesis should be accurate enough to reliably predict the speed and direction of incoming cloud shadows. When implementing triangulation for an actual PV-array, it is safest to have several TPs around the PV-array. It is important that their spacing is small enough that no CS can pass between two sensors without being detected. Additionally, in theory, each TP should be able to connect to each other, cooperating to give more accurate results.

In order to maximize H_2 production, it is possible to bypass the DC/DC converter between the PV-array and the electrolyzer. Bypassing the DC/DC converter skips a conversion step, increasing the system's efficiency, and possibly the production with about 2%. This requires the electrolyzer to be customized so that it draws a current which makes the PV-array operate at MPP when the irradiation is optimal. Using intervals of 200W, the optimal irradiation interval for a DC/DC converter bypass is at $800\frac{\mathrm{W}}{\mathrm{m}^2}$ to $1000\frac{\mathrm{W}}{\mathrm{m}^2}$. Bypassing the DC/DC converter at this interval, increases the energy yield from the PV-array by 0.849kWh per square meter installed PV per year (assuming a PV efficiency of 0.15). A larger interval than $200\frac{\mathrm{W}}{\mathrm{m}^2}$ would save more energy, but the size of the interval depends on what the electrolyzer can handle. The yearly increase in hydrogen production by the bypass, is given in Equation 5.6.1.

References

[1] R. Haas, N.I Meyer, A. Held, D. Finon, A.Lorenzoni, R. Wiser, and K. Nishio. Promoting electricity from renewable energy sources – lessons learned from the eu, u.s. and japan. *Lawrence Berkeley National Laboratory*, 2008.

[2] V.V. Tyagi, Nurul A.A. Rahim, N.A. Rahim, and Jeyraj A./L. Selvaraj. Progress in solar pv technology: Research and achievement. *Renewable and Sustainable Energy Reviews*, 20:443 – 461, 2013.

[3] Sang Yong Park, Jong Wook Kim, and Duk Hee Lee. Development of a market penetration forecasting model for hydrogen fuel cell vehicles considering infrastructure and cost reduction effects. *Energy Policy*, 39(6):3307 – 3315, 2011.

[4] Alexandra Von Meier. Electric power systems. *A Conceptual Introduction*, 2006.

[5] Arno Smets, Klaus Jäger, Olindo Isabella, René Van Swaaij, and Miro Zeman. The physics and engineering of photovoltaic conversion, technologies and systems. *Solar Energy. Cambridge: UIT*, 2015.

[6] C.B.Honsberg and S.G.Bowden. Shading. *pveducation.org*, 2019, Read 18/01 2021. Available from: `https://www.pveducation.org/pvcdrom/modules-and-arrays/shading`.

[7] S. Kihle, A.S. Hoeimyr, D. Gulbrandsen, and S. Tallaksrud. Karakterisering av its modul ecoplus poly 250. *NMBU FYS375*, 2020.

[8] Theodore Wildi et al. *Electrical machines, drives, and power systems*. Pearson, 2014.

[9] Isao Abe. Energy carriers and conversion systems – vol. i - alkaline water electrolysis. *Encyclopedia of Life Support Systems*, 2008. Available from: `http://www.eolss.net/sample-chapters/c08/e3-13-03-02.pdf`.

[10] Øystein Ulleberg. Modeling of advanced alkaline electrolyzers: a system simulation approach. *International Journal of Hydrogen Energy*, 28(1):21 – 33, 2003.

[11] Nel. Atmospheric alkaline electrolyser. *Nel Hydrogen*, Read 07/01 2021. Available from: `https://nelhydrogen.com/product/atmospheric-alkaline-electrolyser-a-series/`.

[12] Lars Wikström. Correspondence with nel. *NEL Hydrogen*, April 2021.

[13] C Donald Ahrens. *Essentials of meteorology: an invitation to the atmosphere*. Nelson Education, 2011.

[14] William J. Burroughs, Bob Crowder, Ted Robertson, Eleanor Vallier-Talbot, and Richard Whitaker. *Været: En håndbok i meteorologi*. Spektrum, 2007.

[15] Andrew J Heymsfield, Martina Krämer, Anna Luebke, Phil Brown, Daniel J Cziczo, Charmaine Franklin, Paul Lawson, Ulrike Lohmann, Greg McFarquhar, Zbigniew Ulanowski, et al. Cirrus clouds. *Meteorological Monographs*, 58:2–1, 2017.

[16] Hassan Abdi. Cloud illustration. *Hassan's work*, 2021.

[17] InMeteo, Marek Mojzík, and Martin Prantl. Ventysky.com. *Ventusky.com*, 2021. Available from: `https://www.ventusky.com/?p=59.51;10.78;8&l=clouds-total&t=20210119/0000&w=dark`.

[18] Houghton Mifflin Harcourt. *American heritage dictionary of the English language*. Houghton Mifflin Harcourt, Read 11/01 2021. Available from: `https://www.ahdictionary.com/word/search.html?q=triangulation`.

[19] Chris D Cantwell, Caroline H Roney, Fu Siong Ng, Jennifer H Siggers, Spencer J Sherwin, and Nicholas S Peters. Techniques for automated local activation time annotation and conduction velocity estimation in cardiac mapping. *Computers in biology and medicine*, 65:229–242, 2015.

[20] Sebastian Raschka and Vahid Mirajalili. *Python machine learning*. Packt Publishing, 2019.

[21] Guido Van Rossum and Fred L Drake Jr. *Python tutorial*. Centrum voor Wiskunde en Informatica Amsterdam, The Netherlands, 1995.

[22] Particle. Particle photonh. *farnell.com*, 2021. Available from: `https://no.farnell.com/particle/photonh/dev-kit-wireless-connectivity/dp/2888494?st=particle`.

[23] ADAFRUIT INDUSTRIES. 746 gps module. *farnell.com*, 2021. Availible from: `https://no.rs-online.com/web/p/gnss-gps-modules/9054630/`.

[24] RS PRO. Rs pro cr1220 button battery. *rs-online.com*, 2021. Available from: `https://no.rs-online.com/web/p/button-batteries/8660653/`.

[25] Roth Elektronik. Re944-s2. *rs-online.com*, 2021. Available from: `https://no.rs-online.com/web/p/breadboards/8971663/`.

[26] TE Connectivity. Upf25b50rv. *rs-online.com*, 2021. Available from: `https://no.rs-online.com/web/p/through-hole-fixed-resistors/1649056/`.

[27] AMX3d. 60ma 68x37mm micro mini power solar cell. *amazon.com*, 2021. Available from: `https://www.amazon.com/AMX3d-68x37mm-Micro-Power-Panels/dp/B0721N5K34`.

[28] Google. Firebase. *Google*, Read 25/01 2021. Available from: `https://firebase.google.com/`.

[29] Google. Google maps. *Google*, Read 01/02 2021. Available from: `https://www.google.com/maps/@59.6661205,10.7779357,91m/data=!3m1!1e3`.

[30] European Comission. Photovoltaic geographical information system. *EU Science Hub*, Read 11/05 2021. Available from: `https://re.jrc.ec.europa.eu/pvg_tools/en/tools.html`.

[31] Talesun. *Bistar TP6H60M 120 half-cell*. Data sheet, 2020.

[32] Dynapower. Dps 500. *dynapower.com*, Read 29/04 2021. Available from: `https://www.dynapower.com/products/energy-storage-solutions/dc-coupled-utility-scale-solar-plus-storage/dps-500/#1476296633680-3b840381-89a8`.

[33] Wes McKinney et al. Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference*, volume 445, pages 51–56. Austin, TX, 2010.

[34] Charles R. Harris, K. Jarrod Millman, Stéfan J van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. Array programming with NumPy. *Nature*, 585:357–362, 2020.

[35] John D Hunter. Matplotlib: A 2d graphics environment. *Computing in science & engineering*, 9(3):90–95, 2007.

[36] V. Fung, J. L. Bosch, S. W. Roberts, and J. Kleissl. Cloud shadow speed sensor. *Atmospheric Measurement Techniques*, 7(6):1693–1700, 2014.

[37] Marc Wieland, Yu Li, and Sandro Martinis. Multi-sensor cloud and cloud shadow segmentation with a convolutional neural network. *Remote Sensing of Environment*, 230:111203, 2019.

A Appendix

Table A.0.1: *Estimated speed and direction of CS based on measurements from 13/04 2021. The distance from the TP to sensor 4 was 1120m, which was used to calculate the speed. Clouds were observed to move south. Intervals are defined by M using a G of 750W. Measured speed was sampled every 30 minutes, as given in Table A.0.2.*

Time interval	Measured speed $[\frac{m}{s}]$	Estimated speed $[\frac{m}{s}]$	Ω	Estimated direction
11:00:25 - 11:00.34	9	7	0.78	64°/NE
11:00:56 - 11:01:10	9	5	0.56	197°/S
11:01:30 - 11:01:43	9	37	4.11	142°/SE
11:01:45 - 11:01:52	9	2	0.22	0°/N
11:01:53 - 11:01:57	9	6	0.67	31°/NE
11:02:42 - 11:02:57	9	1	0.11	357°/N
11:23:43 - 11:23:55	9	3	0.33	185°/S
11:25:16 - 11:25:21	9	4	0.44	5°/N
11:25:45 - 11:25:57	9	40	4.44	172°/S
11:28:38 - 11:28:55	9	1	0.11	164°/S
11:54:54 - 11:55:06	9	3	0.33	189°/S
11:58:00 - 11:58:05	9	7	0.78	64°/NE
11:58:26 - 11:58:35	9	13	1.44	185°/S
11:58:58 - 11:59:12	9	3	0.33	182°/S
12:00:12 - 12:00:22	9	5	0.56	187°/S
12:03:48 - 10:03:58	9	3	0.33	186°/S
12:21:18 - 12:21:25	10	3	0.30	2°/N
12:22:04 - 12:22:13	10	4	0.40	188°/S
12:27:38 - 12:27:48	10	3	0.30	208°/SW
12:27:49 - 12:28:02	10	2	0.20	308°/NW
12:29:45 - 12:29:58	10	25	2.50	150°/SE
12:30:57 - 12:31:15	10	1	0.10	51°/NE
12:32:15 - 12:32:35	10	130	13.00	240°/SW
12:33:41 - 12:33:52	10	5	0.50	242°/SW
12:35:39 - 12:35:48	10	2	0.20	125°/SE
12:40:51 - 12:40:56	10	5	0.50	210°/SW
12:40:59 - 12:41:07	10	7	0.70	212°/SW
12:45:12 - 12:45:20	10	2	0.20	302°/NW
12:45:40 - 12:45:58	10	49	4.90	199°/S
12:47:59 - 12:48:07	10	2	0.20	299°/NW
12:55:27 - 12:55:31	10	7	0.70	313°/NW
12:58:03 - 12:58:10	10	3	0.30	55°/NE
12:58:11 - 12:58:15	10	15	1.50	186°/S
12:58:18 - 12:58:22	10	5	0.50	57°/NE
12:58:48 - 12:59:00	10	75	7.50	270°/W

Time interval	Measured speed $[\frac{m}{s}]$	Estimated speed $[\frac{m}{s}]$	Ω	Estimated direction
13:00:02 - 13:00:15	10	2	0.20	175°/S
13:00:17 - 13:00:27	10	49	4.90	199°/S
13:05:13 - 13:05:17	10	14	1.40	215°/SW
13:06:37 - 13:06:47	10	3	0.30	181°/S
13:08:19 - 13:08:23	10	8	0.80	329°/NW
13:08:23 - 13:08:31	10	25	2.50	199°/S
13:09:17 - 13:09:29	10	4	0.40	195°/S
13:09:30 - 13:09:37	10	9	0.90	191°/S
13:26:58 - 13:27:05	10	3	0.30	243°/SW
13:30:05 - 13:30:23	10	2	0.20	122°/SE
13:30:43 - 13:31:02	10	13	1.30	201°/S
13:31:21 - 13:31:27	10	5	0.50	177°/S
13:31:27 - 13:31:39	10	72	7.20	194°/S
13:35:19 - 13:35:30	10	49	4.90	341°/N
Average value	9.67	15	1.55	189°/SE
Standard deviation	0.47	24	2.46	86°

Table A.0.2: *Measured CS speeds the 13th of April 2021 found using method illustrated in Figure 4.3.3. The distance from the TP to sensor 4 was 1120m, which was used to calculate the speed. Clouds were observed to move south. Intervals are defined by M using a G of 750W.*

Time (±2 minutes)	Measured speed $[\frac{m}{s}]$
11:30	9
12:47	10
13:30	10
14:35	9
15:27	10
16:00	8

All the code used to make this thesis can be found on GitHub via this link:
`https://github.com/Sebbek8/master_thesis_2021`
It also contains most of the measurements.

Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway