MÄLARDALENS HÖGSKOLA
ESKILSTUNA VÄSTERÅS

PHOTOVOLTAIC SYSTEM YIELD EVALUATION IN SWEDEN

A performance review of PV systems in Sweden 2017-2018

ERIC SCHELIN

Akademin för hållbar samhälls- och teknikutveckling

Course: *Degree project in Energy system engineering*
Course code: ERA403
Subject: Energy engineering
Credits: 30 HP
Program: Master of Science in Energy System Engineering

Supervisor: Bengt Stridh
Examiner: Valentina Zaccaria
Date: 2019-06-13
Email:
Esn12006@student.mdh.se

ABSTRACT

The goal of this study is to evaluate Swedish photovoltaic systems regarding energy production from two different years and compare the gathered data with results from a model simulating optimal conditions. This is done to investigate how the energy production differs between each year, why there are differences, and also to evaluate the simulation tools compared to the real production data. A good way to measure performance is to calculate the specific yield, that is the energy produced per unit of installed power (kWh/kWp). In order to complete this study, a literature study was made to investigate reasons for potential variations in PV system yield. Besides that, the production data from 2373 PV systems in Sweden were collected from different databases, and the data were sorted and compiled in order to calculate specific yield (kWh/kWp). The total number of PV systems after sorting was 828 for the 2017-2018 data and 1380 systems for the 2018 data. Data from real PV system production was compared with calculations performed in two simulation tools, PVGIS and PVsyst. Differences in calculation methods were investigated for performance evaluations between the two programs, and also for comparison with the real plant data. The results showed that the average specific yield for Sweden as a whole, to be 798 kWh/kWp for 2017. For 2018 with the results where 890 kWh/kWp when looking at the exact same plants as for 2017. This is an increase of 11,5%. For the simulation tools the results where 974 kWh/kWp for PVGIS, and 978 for PVsyst for an optimized system. Larger variations in specific yield occurs between every of the 21 counties in Sweden. The solar irradiations show significant correlations to the variations of the 2017 and 2018 specific yield data. Differences between the production data from the two years and the simulation tools were investigated further. Reasons for this was discussed to be because of orientations of the panels and shading of the panels. Real PV systems differ in orientation and the amount of shadowing from the simulated calculations.

Keywords: PV, Photovoltaics, Specific yield, Simulations, Shadowing, PVGIS, PVsyst, Solar irradiation

PREFACE

This 30 HP degree project was made at Mälardalens Högskola in the spring of 2019 and is the final work of my masters in Energy System Engineering. The project focuses on Photovoltaic power performance evaluations made on Swedish PV plants during the year of 2017 and 2018. The work came to be because of my interest in future energy recourses, especially in solar power, and the topic was suggested by Senior Lecturer Bengt Stridh at MDH. Special thanks to Bengt Stridh for excellent supervision during the time this project was made. Also big thanks to my girlfriend Petra who helped me cope with everyday life during the time this work was done. I would not have had the strength and determination to complete this degree project in time if it wasn't for here.

Västerås June 2019

....

Eric Schelin

MÄLARDALENS HÖGSKOLA
ESKILSTUNA VÄSTERÅS

SUMMARY

The concept of using sunshine for energy conversion into electricity has been around for many years, but it is not until now the technology has catch up so that it clearly can compete with other sources of energy, and even beat it on multiple planes. PV, or Photovoltaic cells are devices that converts the sunshine directly into electricity. The growth of solar power has increased dramatically in the recent years, and the total installed solar power in the world reached 400 GW in the end of 2017, with over 100 of those gigawatts installed in 2017 alone (International Energy Agency, 2018). In Sweden, the total installed power in 2018 was 411,1 MW, from a total number of 25 486 plants. The Swedish government has a goal of making Swedish electricity production 100% renewable by 2040. A proposition for this to work was made by the Swedish Energy Agency and they state that 5-10 % of the total electricity production needs to come from solar.

This study was made to evaluate Swedish PV system production by comparing real PV plant data to optimized simulations done in different PV modeling tools. The aim was to calculate the specific yield (kWh/kWp) for every county in Sweden for 2017 and 2018 and compare the real values to optimized simulated values. Beyond that, reasons for variations in yield between the years, and between the simulations was investigated.

First, a literature study was made to investigate other work done on PV plant yield. This was done to get an insight into what could affect the yield for PV plants, and how big the variations could be due to these eventual causes. Then the data from Swedish PV plants was collected from different databases and sorted in Excel. This was done to calculate an average specific yield for each of the 21 counties in Sweden for 2017 and 2018, from the statistics of the databases. Then a simulation of the specific yield for each county was done in two different simulation tools, PVGIS and PVsyst. This was done to investigate if, and how much the calculations would differ to the real plant data.

The results show that the average specific yield for Sweden was 798 kWh/kWp for 2017, and 890 kWh/kWp for 2018. By adjusting the specific yield to an average solar irradiation for Sweden, those numbers were 801 kWh/kWp for 2017, and 790 kWh/kWp for 2018. The difference in yield from 2017 to 2018 from the actual data was 11,5 %, with regional differences varying from 2-22% in specific yield from 2017 to 2018. The solar irradiation shows close correlation to variations in specific yield between the years.

Reasons for differences between simulations tools most likely lies in the different ways they calculate the solar irradiation and how the effects of shadowing are accounted for. Depending in the number of PV systems in the study, the statistics will be better or worse. Some counties with low number of systems might not represent reality in the best way possible. Regarding the impact of shadowing, it is hard to say just from PV system production data how big the impact will be.

Differences from optimized simulated values was around 20 percent higher than the real PV system data. Reasons for this might be the optimal input values of the simulation tools, that most likely is not found in the real data. Literature review show that up to 7% of losses in global radiation can be due to shadowing, when comparing annual shading losses on roofs in Uppsala. A close correlation to specific yield and solar radiation can be found for Sweden as a whole, and this is to be expected. When comparing specific yield adjusted to average solar radiation the differences in yield for 2017 and 2018 was only 1.5%. Regional differences larger than that occur in some counties in Sweden.

TABLE OF CONTENTS

TABLE OF FIGURES

DESIGNATIONS

Name	Sign	Unit
Power	P	W
Energy	E	J, Ws, Wh,Wy
Global Irradiance	Gp	W/m²
Direct Irradiance (Beam)	I	W/m²
Diffuse Irradiance	D	W/m²
Global Irradiation	Ge	Wh/m²

ABBREVIATIONS AND CONCEPTS

Solar Irradiance	Solar power per unit area
Solar Irradiation	Solar energy per unit area
PV	Photovoltaic cell – other word for solar cell
Tilt	The vertical angle of the PV panel relative to the horizon
Azimuth	The position angle of panel compared to the south
TWy	Tera Watt year – Energy amount during a year
kWp	Kilo Watt peak - Max panel power output in kilowatts at STC
Specific Yield	Energy produced per unit of power
PVGIS	Photovoltaic simulation software
PVsyst	Photovoltaic simulation software
County	The 21 regions in Sweden, called “Län”
STC	Standard Test Conditions, meaning T=25°C, Gp = 1000 W/m², Air mass =1.5 (Sino Voltaics, 2011)
WMO	World Meteorological Organization

1 INTRODUCTION

This degree project aims to evaluate Swedish photovoltaic system production. This will be done by conducting a literature study showcasing the primary causes that will affect the electric yield of PV plants. When an understanding of why differences will occur during production, an analysis of PV system production data from two different years will be conducted to see how the yield will be affected by different variations, and later a comparison of a simulated system will be done to compare different simulation tools with the production data. All this is done to get an understanding on the potential of PV production in Sweden, and how solar can be a feasible solution for a renewable energy mix in the future. The key to good investments in renewable energy, and political legislations that will favour solar, is to have good understanding on how much energy that can be expected to be produced during a year, how large both seasonal and daily variations will be, and how much said production will cost.

1.1 Background

This section contains a brief background to the project, made for the reader to get a brief understanding on PV cells, how they work, and why this project is of interest. The background will finish with a problem definition that explains what this study will try to answer.

1.1.1 PV cells background

The world consumed around 18,3 TWy in the year of 2014, in the same time the planet received around 23 000 TWy of energy from the sun (Perez and Perez, 2015). Using sunshine for energy conversion into electricity has been around for many years, but it is not until now that the technology has gone into the state where it can clearly compete with other sources of energy, and even beat it on multiple levels. It is this wider approach of positive incitements that is needed for a change into better options when it comes to energy production for the masses. PV cells for electricity production have been around for decades, but it is when these multiple benefits, such as environmental reasons, reliability, efficiency and economic benefits all align, that the masses and society as a whole will adapt and transform their energy production into more sustainable alternatives, such as PV for power generation. This time is now.

PV cells, or Photovoltaic cells, are devices that are converting sunlight directly into electricity (Knier,2019). This report will only cover PV cell technology, even though other types of solar power devices exists, such as solar thermal energy or concentrated solar power to name some of them. The growth of PV power has seen a dramatic increase in the recent years, bringing the total installed power in the world up to over 400 GW in the end of 2017, with 100 of these gigawatts installed in 2017 alone (International Energy Agency [IEA], 2018). The reasons behind this recent growth probably depends on a multitude of reasons, but one of the biggest factors might be that the cost for PV panels has decreased rapidly in the recent years. When looking at module prices in Sweden over recent years, prices in Sweden for the end consumer has decreased from 70 SEK/Wp to 5.5 SEK/Wp, excluding VAT, in 2017 (IEA, 2017). Other factors that might influence the recent growth in PV power could be a desire to be more independent, with more personal control over electricity production and costs.

1.1.2 Installed PV system power and total yield in Sweden

The total installed PV power in Sweden was, including off grid, 322,4 MWp in the end of 2017, with 117,6 MWp installed in the year of 2017 (IEA,2017). An update to these numbers has been published by the Swedish Energy Agency (2019) that add the total grid connected installed capacity in the end of 2018 to **411,1 MW**, with a total number of **25486** plants that exists in Sweden.

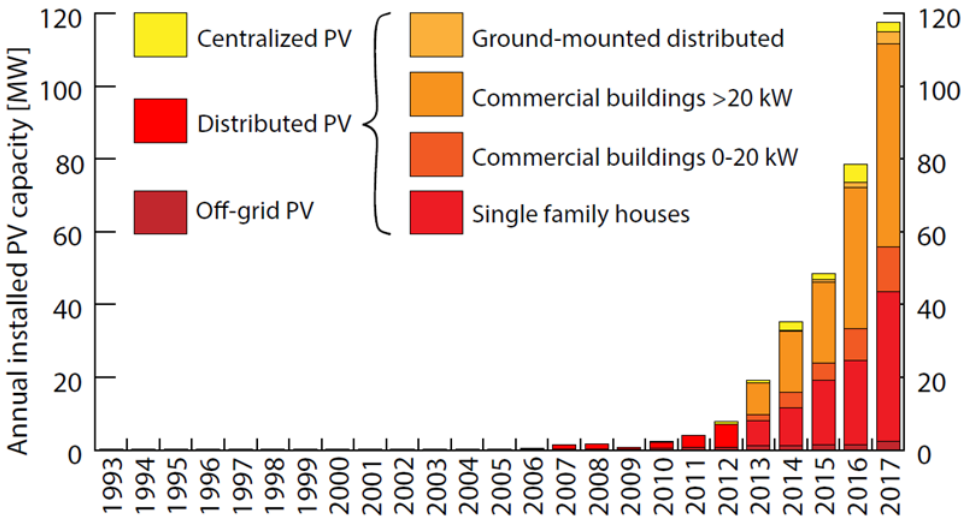


Figure 1 Annual installed PV capacity in Sweden (Source: IEA, 2017, approved by author).

The Swedish government has a goal of making 100% of the Swedish electricity production renewable by the year of 2040. The Swedish Energy Agency has suggested that, in order to meet that goal, 5-10% of the Swedish electrical production needs to come from solar. This means that 7-14 TWh per year needs to be produced by solar. An assumption was made for this that the amount of electricity production is the same as today by 2040. (Swedish Energy Agency, 2016). .

1.1.3 Thesis background

When looking at PV production in Sweden, both actual yield from production data, and that from simulations that uses solar radiation data for specific locations in the simulations, there is most certainly differences in electrical production. Actual yield data will have a lot of differences when looking at both different plants in the same region, as well as differences when looking at similar plants in different regions. When simulating a system for a specific location, there is often optimal placement of the PV panels, with perfect solar radiation for that specific location. The reality is different however, and not every solar plant owner has a roof pointing in optimal direction, or with the perfect angle of the roof itself. There will most certainly be shadowing that reduces output, as well as large differences in solar radiation on different locations on different years. This is why it is interesting to look at as many plants as possible, for all different areas in Sweden, and study the energy production of each plant, in order to get the best statistical data. Each PV panel, or system of panels will probably have a performance sheet telling the customer how much electricity can be expected from that system, and this is often the case in the best-case scenario. By looking at production data from a large number of systems, the specific energy production (kWh/kWp) will hopefully result in a number that reflects reality in a better way, for all different regions in Sweden. Another interesting thing to look at is how this specific energy production varies from simulations of optimized systems in all these different areas of Sweden, how big the variations can be, and why these differences occur. A great understanding of the performance of PV systems, and the variations in performances that can occur, are all key for making solar power a feasible solution for the future energy mix in Sweden.

1.1.4 Problem definition

When calculating investments for PV systems, the specific energy production in kWh per kW installed power is an important and interesting parameter to consider. There are multiple factors that affects how much electricity a certain plant produces per installed kW. Factors that play a big role in the variations are such as shading and variations in solar radiation for a certain location. That is why it is interesting to see how the production data looks for a larger number of plants, in different regions in Sweden and compare this production data to calculated values made by simulations on a typical PV plant. If variations occur, understanding on how big these variations in production can be are essential in order to make good viable investments into solar power, both for home owners and for political implementations.

1.2 Purpose

This study aims to evaluate a larger number of Swedish PV systems production data from 2017 and 2018. This data was sorted and compared with results from a simulated model in different simulation tools. The differences was compared and analyzed, and the differences within the simulation tools was investigated on underlying theory. This was done in order to evaluate the performance and methodology used for the simulation tools.

1.3 Research questions

- How much energy does Swedish PV plants produce on average (kWh/kWp)?
- How much does the actual yield differ from ideal, simulated values on specific yield, and what could be the cause of these variations?
- How well does the total yield in 2017 and 2018 correlates to the solar radiation differences between these 2 years?

1.4 Delimitation

- Production data gathered from 3 separate databases only accounts for 5% of total installed capacity in Sweden in the end of 2018.
- The years that were studied in terms of yield data are 2017 and 2018
- Gathered data on production with zero values in months, except December and January, has been sorted out of the study to get rid of inferior data sets.
- The studied PV technique is crystalline silicon, both mono-and poly-crystalline cells, but in the simulations only poly-crystalline cells was used due to them being more

common in Sweden at this moment (Solkollen, 2019). There is no way of knowing from the databases on production what kind of PV cell technique they are.

- From production databases, no known tilt or azimuth angles was described, hence when simulating, a set tilt of 31° and also an optimal tilt was chosen, and the azimuth angle was set to zero degrees (directed south)
- The values on the current production and installed capacity in Sweden are based on 2018 data when possible, and some on 2017 data. Installed capacity are rapidly growing so these values might not represent the real world by the time this report is read.
- The simulations are based on a single location, or town for each county (Län), so the actual values simulated might differ from location to location in a county, due to its size and meteorological differences.
- The solar radiation data are based on both statistics from SMHI solar radiation databases for 2017 and 2018, but also other tools to calculate irradiation on the locations that was not included in the SMHI data. SMHI stored Irradiation data for a number of years, but these are only based on data from 11 of the 21 counties in Sweden. For the counties that doesn't have real irradiation measurements, a modelled called STRÅNG was used to estimate solar irradiation for these locations.

2 METHOD

The method part is divided into 3 parts, where the first part of the study was the literature study of PV system yield and reasons for variations in PV production was investigated. This was done together with a research on the underlying energy calculations used in the two simulation tools PVGIS and PVsyst. Next part of the thesis focused on data gathering from databases storing annual production data from Swedish PV systems, and the sorting of this data to showcase specific yield. The sorting and gathering of production data were a big part of this study, at least time consuming wise. The third and last part of the study was to use simulation tools to compare the simulated specific yield to the real-life yield from the collected production data.

2.1 Literature study:

The literature study was done in order to get an understanding on the background to solar power and the technical aspects of how PV cells work. In order to understand why the specific production will vary during different years and with plants under different working conditions, reasons for these differences need to be investigated. The first thing to research was other reports that focus on PV system yield in other countries in order to investigate how they handle the data collection and what they found to be the contributing parameters effecting the yield in PV plants, such as shading, differences in solar radiation etc.
Another big part of the literature review was to investigate the underlying theory behind the simulation tools used in this study. This was done to get and understanding on the technical energy calculations that each simulation tool used to calculate the power output, and what the differences and accuracy these will have. The underlying theory from the PVGIS tool was obtained from their website, where a report covering all the theory used in the software. This consisted of both how the solar radiation data were determined, and how losses were addressed for example. The theory behind PVSYST is a little bit more complex, due to the higher complexity of the software compared to PVGIS. On the website PVSYST Help a explanation of the software can be found in a tab called “Physical models used”, and under “Project design-Shadings”.

2.2 Data gathering and sorting:

In order to get good statistics on PV system yield, data from a large number of plants needed to be collected. Different types of databases that store production data and size of each plant in Sweden has been used. Data has been collected both from a company called Check Watt that shared production data from 2373 plants, and from the Solar Edge open database. The plan was to collect data from the whole year of 2017 and 2018 in order to see variations during the months of a year, and total yield variations during these 2 years. Some months was seen to have zero production, and if that occurred in the middle of the summer, it was dismissed from the statistics. This could be because of either something wrong in the data transmission from the PV system, and in that case, it should be dismissed from the statistics, but if the system shows zero values because its broken, it should be included. A broken system shows the reality, and is therefore good for statistics, but because it is really hard to distinguish between a fault in the data collection and the actual PV system, all zero values in the summer was excluded. The more plants that can be gathered, the better the statistics will be. The goal was to at least find 1000-1500 plants in Sweden, and in every county in order to have good data to compare with simulations for different places in Sweden. The overall aim was to find as many plants as possible with full monthly data for 2018 to evaluate the yield in that year, but also find as many plants as possible that had full monthly data for both 2017 and 2018 in order to evaluate performance differences between these two years. This meant collecting data from plants that only had data for both 2017 and 2018. Because of the exponential growth of PV plant installations during the last years, there is much more data that could be obtained from 2018 than 2017, resulting in much less plants that could be used for the 17/18 comparison.

Much of the work in sorting the data consisted of using the built-in tools in Excel for sorting and formatting the data in a way that only the data that was interested could be obtained. Other than the plant yield data, solar irradiation data was obtained from SMHI for different locations that has measurements on global radiation. On the locations that doesn't have SMHI ground measurements, the model called STRÅNG was used. This model calculates irradiation on an arbitrary location anywhere in Sweden from coordinates. This data on solar radiation was then used to compare the production data for 2017 and 2018 with the solar radiation received for the two years, to see how big the correlations are.

2.3 Simulations on a designed model

In order for this study to be complete, a comparison of the gathered production data and calculations done in different simulation tools needed to be done. When investing in PV systems there will most likely be a calculation done by the manufacturer or supplier of the panels that will tell the customer how much energy can be expected to be produced on a monthly or annual basis. This is key for investments and economical calculations that needs to be done in order to fully understand the potential of PV power. The reasons for doing these simulations are to compare these to the actual production data for different regions in Sweden, in order to see how big the variations can be between real production and simulated values. The simulations have been performed in the public access program PVGIS and the commercial software PVSYST, and a review and comparison between the calculations on power output between these two programs have been made. In order to compare these two programs between themselves and the gathered production data, the input data and the working conditions needs to be the same when simulating the production. This means a set azimuth to the south, a fixed tilt on the panels, and to put the plants in close proximity on the maps. For example, if the yield should be compared from plant metadata in Jämtland and the 2 simulation tools, the plant should not be put in a mountain valley in PVGIS, and a flat field in PVSYST. The comparison that is most key to compare is the specific yield, that is the energy in kWh per year for each kW power. This means that the easiest way of simulating this is to calculate for a plant that is 1 kW, meaning that the total annual production will be in kWh/kW. This can easily be done in PVGIS, but not in PVsyst due to the fact that a specific number of panels will need to be built in the model, and this will most likely not add up to exactly one kW peak power.

3 THEORETICAL FRAMEWORK

This section will look at the research of previous work on PV system energy yield, as well as a background to how a PV system works, in order to understand why there will be differences in yield in different locations and different installations during different years. An understanding on how big the variations in energy yield are, as well as how production data should be investigated and sorted is key to get the best performance numbers, that reflect the reality in the best possible way.

3.1 Background to solar power

Photovoltaic cells are devices that converts sunlight directly into electricity. PVs are just one of many techniques sunlight can be converted into energy. A PV system for power generation consists of different types of photovoltaic cells in interconnection and capsuled to form a photovoltaic module. These modules can then be mounted in different ways, and depending on application, be connected to an inverter for electrical grid connection, or/and a battery for electrical storage. (*IEA,2018)*

There are different types of PV cells, generally single and multi-crystalline silicon, thin film or organic variants. The most common and produced type in 2018, are the crystalline silicon, accounting for over 97% of the overall production of PV cells. Mono crystalline silicon cells can reach a commercial efficiency of 16-25%, whilst the cheaper to produce, multi-crystalline silicon cells can reach 14-18% efficiency (IEA, 2018).The efficiency of a PV module can be described as how much of the sunlight energy that can be converted into electricity (What Factors Determine Solar Panel Efficiency?, 2013). Figure 2 show the efficiency for different types of PV cell technology in a best-case scenario.

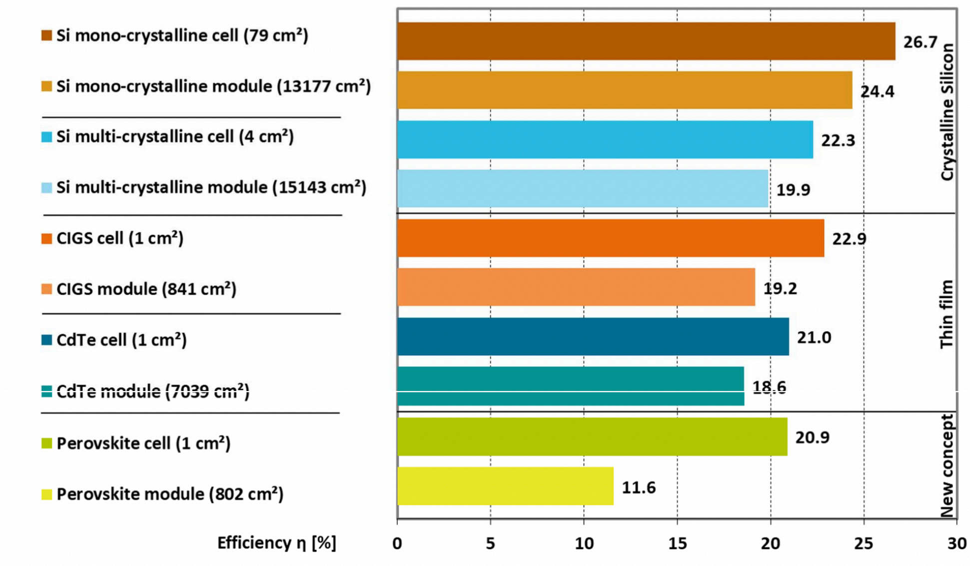


Figure 2 Efficiency best case scenario of different types of PV technologies (source: ISE, 2019)

3.1.1 Solar Radiation in Sweden

The source of all the energy that the panel receives, is of course the sun. This energy is almost constant when looking outside the earth's atmosphere. It will vary due to that the distance between the earth and the sun, and this distance varies during the year with approximately 1.5%, but the mean value of energy is 1366 W/m² outside the earth's atmosphere

(SMHI,2017). When the sunlight goes through the atmosphere, and eventually ending up at the earth's surface, some energy will have been lost due to absorption, dispersion or scattering to surrounding matter in the atmosphere (*PVGIS, 2017*).
The radiation from the sun that are being received at ground level is called global radiation, and this radiation is the sum of 3 components, called direct radiation, diffuse radiation and reflected radiation. The direct radiation is the radiation that is directly from the sun, seized at the ground totally unrestrained from the above atmosphere and others. The diffuse radiation is the component that has been scattered by the atmospheric matters, such as clouds and dust. The reflected radiation is radiation in the form of the previous to types that has been reflected on earth surfaces (Chalkias et al., 2013).

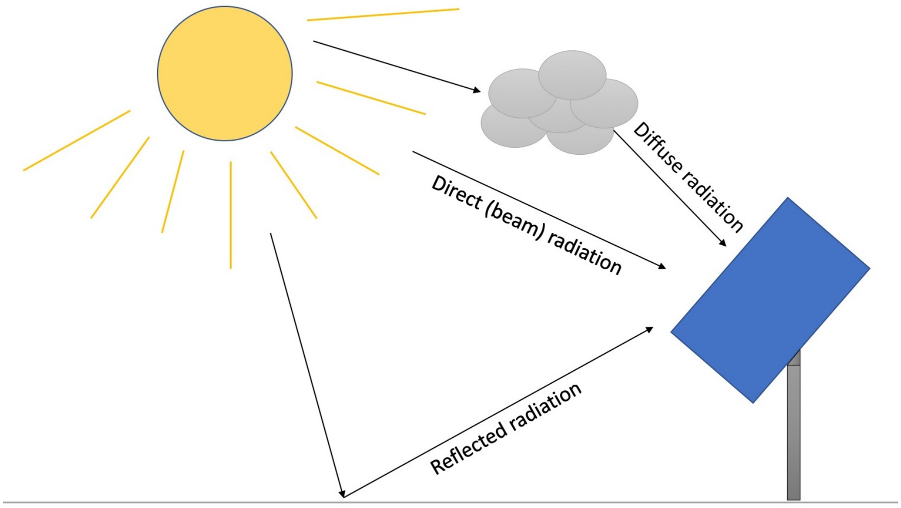


Figure 3 Global radiation components.

Figure 3 show the three different types of radiation that the global radiation consists of. Figure 4 shows the normal annual global radiation in Sweden, and the variations in intensity, and that it varies from around 750 kWh/m² in the very north, to around 1050 kWh/m² as a maximum. This is normal values during a year based on WMO data from 1961 to 1990 (SMHI,2017a). Figure 24 in the results part of this report will show higher values in solar irradiation based on newer irradiation data and a model called STRÅNG.

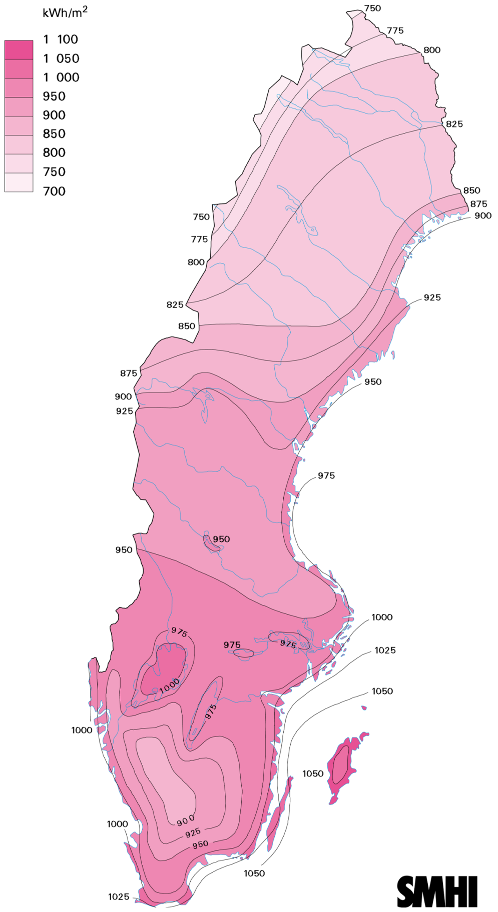


Figure 4 Annual global radiation (1961-1990) on a horizontal surface in kWh/m² (source: SMHI 2017a)

The energy in the form of solar radiation that reaches the surface of the earth varies during the day and the year. This is due to both variations of the position of the sun, and the amounts of clouds in the sky. SMHI measures the solar radiation reaching the earth at a horizontal surface and from all directions in the sky. The global radiation on a horizontal surface can be described as:

$G = I * \sin(h) + D$
$I = direct\ radiation\ (beam)$
$h = sun\ height$ [° above horizon]
$D = diffuse\ radiation$
(SMHI, 2015).

3.1.2 Swedish solar radiation in 2017 and 2018

SMHI will have measured data for some specific locations in Sweden when it comes to global radiation. In order to compare the differences in production from 2017 and 2018, a comparison of the global radiation during these two years will be performed. The data that can be found on the website contains measured global radiation (shortwave radiation) at automated radiation stations on the ground measured continuously and averaged as hourly values . The measurements are taken at the horizontal plane in the unit of Wh/m². The data contains the date, time in hour, value and quality of the measurement. The quality of the value is either green, that is a controlled and approved value, or yellow, suspicious value (SMHI,2019). More on the data gathering of solar radiation will be described in subsection 4- "Description of the study".

3.2 PV system yield

When looking at how much electricity is being produced in different countries, the total production in kWh per year is interesting, but also the specific production in kWh/kW installed is of great interest. This will showcase how much electricity from solar is being produced in other parts of the world, in relation to that locations total electrical production, but also take a look at the specific production to see if there are differences in production between different areas of Sweden.

The performance of PV systems will differ during different years and different locations. In order to review the performance of a plant, not only does the total electrical output need to be investigated, but this needs to be set in relation to what that specific location during that specific year really could produce. Variations in specific power output in a plant will vary due to several reasons, and a great understanding on what will cause variations in performance are important. As well as to understand how big these variations can be in order to make photovoltaics viable, both when investing and making other decisions of implementing solar power into the society.

3.2.1 PV in Sweden

When looking at installed data, the definition of installed data needs to be given. In the report by IEA (2017) the definition is that installation data are defined as all nationally installed ground PV applications of 40 W DC or more. A PV system are defined as modules, inverters, installation and control components. The statistics are based on sales, and therefore data on commissioning taking place during the specific year are not accounted for, but the accuracy of the data is still considered high due to the fact that the number of commissioned plants is low compared to how much are installed. The lifespan of a PV plant is high, and a majority of the plants in Sweden have been installed during the last 5 years, so commissioned plants are left out in these statistics due to their low number. Therefore, the total installed PV power should really be considered total accumulated installed power, rather than total PV capacity, even though these are almost the same. (IEA,2017)

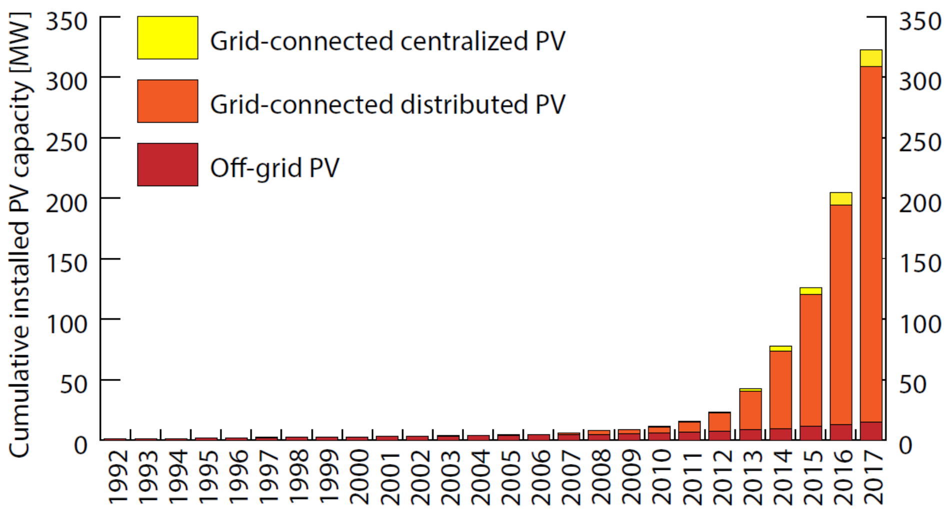


Figure 5 Total installed PV capacity in Sweden (Source: IEA, 2017, approved by author)

The total cumulative installed PV power was 322,4 MWp by the end of 2017. These amount of installed and sold PV accounts for both grid connected as well as off grid applications (IEA,2017). This means that this much have been sold and installed until the end of year 2017. When looking at the graph, one sees that a majority of the total cumulative power have been installed from around 2010 and forward, hence, most of the plants are still in use.

During the time this report was written, update to these numbers has been published by the Swedish Energy Agency (2019) that add the total installed capacity in the end of 2018 to **411,1 MW**, with a total number of **25 486** plants that exists in Sweden. This data has not been published in a report, but a database at the Swedish Energy Agency website will hold statistics that can be accessed and viewed.

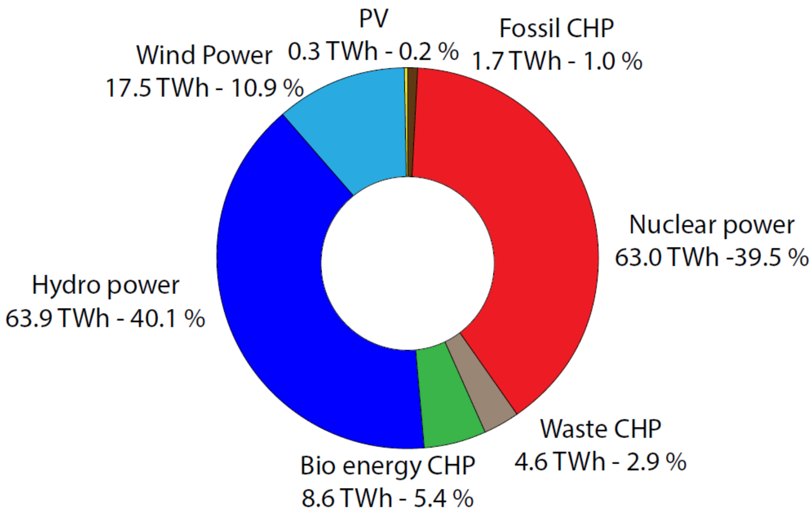


Figure 6 Total electricity production in Sweden 2017 (IEA,2017, approved by author)

Figure 6 shows the energy sector in Sweden in 2017. PV accounts for a total of 0,2 % of the total energy mix in Sweden. Incentives that will affect the future of PV power in Sweden is of course the political policies. The Swedish energy commission has agreed on a goal that Sweden will have 100% renewable power generation by 2040, and this will help push renewable energy to grow. The so-called Green electricity certificate will be extended from 2020 to 2030, and this prolongation will certainly help push the PV market. Beyond that, politicians plan to make it easier for small scale energy production, all things that will help the PV market in the future. (Swedish Energy Agency, 2017).

3.2.2 PV system yield analysis

There are key parameters that will affect the power output, and in order to review the performance of a system or geographic location for performance, the underlying reasons for performance variations need to be studied. A report covering Belgian PV system performance by Leloux, Narvarte and Trebosc, (2011) has studied 993 installations from performance data in Belgium and found that the orientation of the PV plants to cause 6 % losses as a whole, compared to an optimally orientated plant. The study showed that almost 70% of the plants lose less than 5% of energy due to orientation, and that the overall performance index, here described as the ratio between actual performance and performance from a very high quality optimally orientated plant, to be 85%. Meaning that the actual performance was 15% less than an optimal plant. The report also finds that the net annual energy for the year of 2009 was 902 kWh/kWp, and by adjusting these 2009 results by the ratio of solar radiation that year, to a solar radiation average during the last decade, a value of 836 kWh/kWp was found. Other than non-optimal orientation, losses due to soiling was said to be 3%, and shading losses expected to be 2%.

Another report by Killinger et.al, (2018) goes deep into PV power characteristics and performance, investigating data from over 2,8 million PV plants around the world. They found that the average specific annual yield to vary on a country basis from 786 kWh/kWp

(1,5 % of total capacity in 2017) in Denmark, to 1426 kWh/kWp (32% of total capacity in 2017) in the south of USA. The report focuses on investigating what they found to be the key parameters effecting the yield: - tilt angle, azimuth angle, installed capacity and the efficiency in the form of specific annual yield.

A report by van Sark et.al, (2014) investigates Dutch PV systems for an update on the specific yield for analyzing the PV contribution to renewable energy. By collecting PV system data from several data sources from the years of 2012 and 2013 they calculated an annual specific yield for the different regions in the Netherlands. They investigated 2,4 MW (6,5 %) of the total accumulated installed capacity in 2012, and 11,6 MW (1,6 %) of the total capacity in 2013. They found a specific yield of 877 kWh/kWp for 2012, and 878 kWh/kWp for 2013 when looking at an average for the country. Data on global irradiation for the two years was 1036 kWh/m^2 for 2012 and 1045 kWh/m^2 for 2013, an 0,9% increase. The report also concludes that regional differences in the Netherlands are large for a country of that size, up to 16 % difference in average specific yield between regions.

Another report by Leloux et.al, (2015) investigates performance of more than 31 000 PV systems in Europe installed between 2006 and 2014. In the report a focus was on 4 reference countries containing the majority of the PV systems. These were France, UK, Belgium and Spain. In France, 17 672 systems were investigated, Belgium 7648, UK 5835 and Spain 29 PV systems. The results from the study show that the mean annual specific yield in Belgium and the UK are similar at around 900 kWh/kWp. For France the number was 1115 kWh/kWp and total average for Spain was 1900 kWh/kWp. The number of investigated systems in Spain are only 29, but it has the most installed in terms of power. This means that they have bigger systems, or plants, and also big sun tracking systems. The average for the static systems in Spain was 1450 kWh/kWp, and for the sun tracking systems 2100 kWh/kWp. Other than this specific yield results, observations were made that PV system performance generally increases with increasing peak power. Also differences in total yield from losses in different types of inverters up to 5% was observed. Differences in yield depending on module manufacturer was observed up to 6%.

3.2.3 Analysis of shading impacts

Other crucial parameter that was investigated in the report by Kilinger et.al, (2018) was the impact of shading. Shading is one of those parameters that cannot be investigated through metadata, so the report focuses on performing a shading analysis in order to fully evaluate the parameters that influence the yield of PV plants. They find that several articles on research regarding shading to be simplified, and that results from these researches could be improved by understanding the influence of shading in a better way. So, in order to tackle the problem of shading, the report focuses on how to consider the impact of shading. A method of doing this is done by performing shading analysis on 48 000 buildings in Uppsala, and building a model based on roof shapes and LiDAR data. The model takes into account the surrounding objects, such as trees and other buildings that may shade the roof in question and produces a map showing what parts of the sky are visible from a specific part of the roof, and in that way showing if there is objects that are blocking the direct solar beam. The model will calculate for smaller parts of a specific roof and produces a mean shading value for the whole roof, and that way accounts for if the roof is partially shaded. The results from the shading analysis produced 2 functions that can be used to model the impact of shading. One of the functions calculates the beam irradiance subcomponent, describing the shade impact as a function of the solar elevation angle. Another function describes the view factor as a function of the roof angle and can be used to understand the losses of diffuse and reflected irradiance. One of the results from the study also showed that losses from shading of the diffuse irradiances component are higher than the beam component of shading. This is on an annual basis for sites of similar meteorological conditions as Uppsala, where the data for building the model was based on. The average impact of shading when considering all the roof facets in Uppsala, was for global irradiance 7,3%, with the subcomponents at 3,6% for beam, 6,3% for diffuse and negative (no loss, but gain) of 2,7% for reflected irradiance. The report takes into consideration that all the roofs in Uppsala was considered, and that if only roofs with PV installations would be considers, the total losses would most likely be smaller.

Figure 7 show the losses due to shadowing on the beam and diffuse component. The losses are in percent on the vertical axis, and the number of roofs with that loss are showed on the horizontal axis. Due to reflective irradiation, the losses can be negative as shown by the yellow lines.

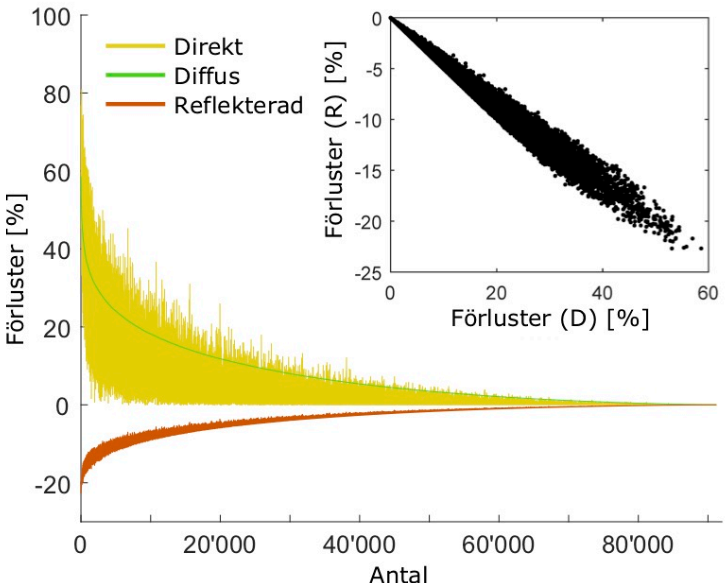


Figure 7 Beam and diffuse irradiance loss when considering shading (From report Killinger et.al, (2018), approved by author)

In a report by Bengtsson, Holm, Larsson, Karlsson, (2018), investigations on the shadowing effects was done. They mean that solar cells by themselves are really sensitive for shadowing, and to get the voltage up to useable levels, the cells will be connected in series. This results in a scenario where the cell receiving the least amount of solar radiation will decide the yield in all the cells in connected in that series. The consequences of shadowing can depend on different factors and will result in different impacts depending on the time of the day the shadowing occur, how the string of cells are placed, and how different types of technical components in the panel or system of panels are used to minimize the shadowing effects. Components like the bypass-diode, and different inverters that can optimize the power, will help to keep the effects of shadowing to a minimum. When a solar cell connected in series are exposed to shadows, it will generate less current, and this results in that the other cells has to handle a bigger current than the shadowed one. A so-called reverse voltage arises over the cell, and by using a bypass diode, that are mounted in the opposite direction will then receive a forward voltage and it will start to conduct current in addition to the shadowed cell.

Due to the fact that the global radiation is said to consist of 3 types of components, the impact of shadows from these three sub-components will vary. The diffuse solar radiation will normally account for around 50% of the global radiation during an average year in Sweden, but on a sunny day the that number drops to 10%, meaning that most solar radiation comes from the beam component. This means that the effects of shadowing if something is blocking the direct beam radiation will have big impacts of the total efficiency of the system. How much an object will cast shadow in a panel depends on the height and size of the object. The shadow casted from the sun are normally divided into 2 components, a core shadow and a half shadow. The ratio of this depends on how close the shadowing object is to the solar

panel, and the closer it is the darker the shadow will be. The core shadow will reduce the incoming radiation by around 60-80%, and the half shadow around half that.
The effects of seasonal and temporary shadowing, such as dirt, debris or snow are said to have low impact of the annual yield in Sweden. The losses due to these shadowing can be minimized by mounting the panels in an angle that will result in higher speeds on the cleaning rainwater falling on the panel. The best effects of this was said to be around an 30° angle. Snow accumulation and efficiency problems due to snow are worst during the winter, but this is also when the solar radiation is low, resulting in small effects of the annual yield. But effects of snow can still be considerable, if for example snow accumulate in piles because of snow fences on the roofs that are mounted to prevent snow falling down below, and these piles cast shadows on the panels. (Bengtsson, et.al, 2018)

By using smart monitoring of PV systems, a report by IEA-PVPS (2017) goes deep into investigations on improving efficiency using statistics on PV system performance. Some key parameters for the analysis to work was location, PV module type, inverter type, orientation, tilt and string configuration. On top of that, other monitoring hardware to measure current, voltage, frequency, active and reactive energy, are installed. By monitoring these parameters and comparing the predicted hourly production with actual measured production, the health of the system can be seen. By analyzing production data and patterns, the power generation can be measured over time to determine when a system is shaded. For doing this, statistical methods are used to calculate for shading losses. This is done by calculating deviations from expected PV system production, and if there are observations of constant deviations in the production, this loss is assumed to be because of shading. Other types of shading losses that was investigated was that of soiling of the panels. This could be predicted by using the statistics to see a low drop of performance over time, and then a sharp spike in performance after a rainfall. This would mean that the panels get cleaned, and the performance drop could be calculated and soiling losses measured.

3.3 Simulation software PVsyst

In order to conduct the best possible simulations that reflects the reality in a good way, the input data for all the simulations has to be good and thoughtful. In order for this to happen, input data such as panel slope has to be estimated, and this was done by researching the most common roof angles in Sweden. A report by (Kamp,2013) researched that the mean tilt of Swedish domestic houses to be 31°.

All the below subsection of PVsyst have information taken from PVsyst, (2018).
PVsyst is a software for modeling and study for sizing and production data of PV systems. The program handles both grid connected, stand alone, pumping and DC grid connected systems. It includes comprehensive databases for meteorological data and PV system components. In this report, the so-called Project design tool was used to simulate a system in different locations in Sweden. In the Project design the user can extensively specify different parameters to compare and analyse the effects of thermal behaviour, angle losses and shading losses from both far shading objects such as mountains, or closer objects such as buildings or trees. An economic evolution can also be done by using actual real component prices and investment costs. Databases for meteorological data and PV system components can be managed and created from scratch. There are real meteorological data from predefined sources and also new geographical sites that can be created, and corresponding generation of hourly meteorological data can be made to these new sites created. More on this in below subsection. (PVsyst, 2018)

3.3.1 Solar radiation, data retrieving and validation

PVsyst uses a software and database called Meteonorm that provides monthly meteorological global horizontal irradiance, diffuse horizontal irradiance, wind speed and temperature. It will generate synthetic hourly data from the monthly values using stochastic models. The Meteonorm database contains over 8000 weather stations worldwide, and if the simulated site is close to such a station, data on global radiation, temperature and wind data will be taken from these stations and other databases. Other than ground station measurements, satellite data will be used for locations that are not close to any ground measurement site. In Europe the distance for satellite data to be used solely is more than 50 km radius from a ground station site. Within that distance, down to 10 km, a mixture of satellite and ground measurements will be used, and within 10 km of a ground site, only ground measurement data will be used. The process of deriving solar radiation data from satellites uses different methods described more in depth in other reports, but put simply, satellite pictures from different geostationary satellites are used and processed to calculate daily means of global radiation and are summed up to monthly values. (PVsyst, 2018)

3.3.2 Global irradiance, beam and diffuse component

The global irradiance hourly values are generated from monthly values by using so called stochastic models. The model used first generates daily values from the monthly, then hourly values from the daily, using so-called Markov transition matrices. The matrices produced are made for the hourly values on irradiance, and the statistical properties and distributions are corresponding to real meteorological data that has been measured on ground sites.
The diffuse part of global irradiation can either be measured on ground stations, but when this is not available, the diffuse part has to be estimated from horizontal global irradiance. This is made using models, two of them mentioned here are the Liu and Jordan’s correlation, that uses experimental methods to express a ratio of diffuse and global irradiance based on a variable called clearness index. Another one is called the Perez model, that is more complex that uses hourly data for defining the diffuse part. The Perez model is not used in the current PVsyst model due to it requires very well measured data on global irradiation, and the Liu Jordan’s model is proven to provide good results. (PVsyst, 2018)

3.3.3 Shadowing

There are two types of shading that will affect the yield of a PV plant. Far shading, or the horizon line, such as mountains or valleys far away, making that type of shading acting on the PV plant in a global way, that is, the sun is either visible or it is not. Near shading are the type of shading that is produced closer to the PV plant such as trees or buildings that will cast visible shades on the panel. The near shading part is much harder to calculate and relies on performing a detailed three-dimensional description of the PV system and the surrounding area. The simulation on the near shading requires hourly calculations, and are calculated differently on the beam, diffuse and the so-called albedo component, that is the reflected component of the global irradiation. By default, the far shading part from mountains and valleys will not be included in the calculations, unless the user implements a horizon profile in the project area of PV syst. Also, the near shading objects needs to be implemented by the user. (PVsyst, 2018)

3.3.4 Calculation of PV power output and losses

The ambient temperature and wind will strongly affect the electrical performance of the PV system. The thermal profile of the panel is determined by and energy balance based on the ambient temperature and cell temperature that is affected by solar irradiance making the panel heat up. Therefore, the ambient temperature data are key for good performance calculations on the system. A general model for synthetic temperature data does not exist in PVsyst, and the temperature data generation are only adjusted on Swiss meteorological data, and not proven for any other site in the world. The wind speed will affect the calculation of PV module temperature profile, when estimating the so-called Array loss factor, and are taken as a default value or if possible, from the meteorological site data, but due to poor reliability of the wind data, it is recommended not to use this for simulations. (PVsyst, 2018)

The efficiency of the inverter is defined as the ratio of the output power to the input power. The efficiency is mainly a function of the power, but also the input voltage. In PVsyst there are a number of different ways that the efficiency is expressed. The most common way of describing the efficiency are by a Single efficiency profile, where the output power is a function of a linearly interpolated function of the input power. Another, more accurate way of calculating this is by using three different efficiency profiles for three different input voltages, and in the simulation software a so-called quadratic interpolation is performed between these three power curves that are a function of the actual input voltage. (PVsyst, 2018)

3.4 Simulation software PVGIS

PVGIS is a free to use online calculating tool for PV potential, and is made by Joint Research Centre from European Commission. It is simple to use and work without installation on both computers and other devices, and anyone with minimal basic PV knowledge can use it to perform PV output calculations. The program uses data on solar radiation and calculates the PV system performance based on several other inputs. (Tarai, Kale, 2016)

3.4.1 Solar radiation intensity, data retrieving and validation

The best way to measure solar radiation is to use good quality sensors on the ground, but the sensors need to fulfil a number of criteria in order to be really useful. The number of sensors on the ground that can do this are generally low, and spaced far apart, making it difficult to use them for input at specific locations, unless the sensors is right there. Therefore, it has become more common to use satellite data, mostly from geostationary meteorological satellites as this data are available for every location that the satellite image covers. There are some disadvantages of using satellite data for ground level solar radiation, one of them being that that it involves complicated algorithms that also has to use data for atmospheric water vapour, dust, particles and ozone, and some conditions can make the models loose accuracy, for example snow that can be interpreted as clouds. Other than that, the accuracy of satellite based solar radiation calculations are said to be very good, and PVGIS uses most of the solar radiation data from the satellite algorithms. How it works is that the satellite image estimates the effect that the clouds have on the solar radiation on the ground by looking at the reflection of the incoming sunlight on the clouds. Reflectivity of the clouds are calculated by looking at satellite pictures and focusing on a single pixel at the same time every day for a month. Then the darkest pixel during the month are assumed to be the one being equivalent to a clear sky, and then the cloud reflectivity is calculated relative to the darkest (clear day) pixel. This method works well in most cases but doesn’t work very well when the ground is covered in snow, which can be interpreted as clouds, and calculating irradiance values to low. The solar radiation data has been validated against measurements made on ground level sensors, and there is variance in differences depending on the location in the world. The most northern data in this report was 58 degrees north in Estonia, having a difference between ground station and measurement of radiation of 4-5 % depending on what radiation database was used. Differences of up to 14 % can be seen in one location, but generally the difference span only a few percent. (PVGIS, 2017)

3.4.2 Inclined planes

The solar radiation calculations used in PVGIS uses global and beam radiance on horizontal surfaces. PV panels are normally not horizontal, hence calculations of irradiance on inclined surfaces need to be done. An important thing to do is to estimate the values of beam and diffuse radiation that hits the tilted panel surface. The beam component is no problem to calculate if the position of the sun in the sky is known relative to the plane surface, but the diffuse component is not so easy, due to the fact that the diffuse radiation has been scattered by atmospheric components, and as such, can be described as coming from everywhere in the sky. The diffuse component is almost never uniform over the whole sky dome, due to changing cloud covers and different brightness in different parts of the sky, so more complex estimations models is needed. In PVGIS the model is a so called two component anisotropic model, that can distinguish between clear and overcast sky, and sunlit and shaded surfaces, resulting in a model that has been proved to have the best performance in a study by ESRA. (PVGIS, 2017)

3.4.3 Shadowing

PVGIS uses calculations on shadows on panels based on the terrain around it. The terrain can have huge effects on the output when near hills or mountains and the sun is behind the them. PVGIS uses information about the elevation of the terrain around it, with a resolution of 90 meters, meaning that for every 90 meter there is a value for ground elevation. This data is then used to calculate the time during the day that the sun is behind hills, and during this time, only the diffuse component of radiation is used. Due to the relatively “low” resolution of 90 meters, everything smaller than that, such as trees and houses, won’t be accounted for in the calculations. (PVGIS, 2017)

3.4.4 Calculation of PV power output beyond radiation

The factor that affects the output on a PV system the most is the amount of solar radiation that hits the panels, but there are other factors of importance that will affect the output. PVGIS will make corrections based on real operation condition effects, such as shallow angle reflection, changes in solar spectrum, module temperatures other system losses. The shallow angle reflection will vary depending on the angle of the light hitting the module, generally causing a loss of around 2-4% due to reflection in the panel, before the lights even reaches the cell. Changes in solar spectrum will vary during the time of day and meteorological conditions, and PV modules are sensitive to what wavelength of light it can use. The PV power output will vary depending on the spectrum of sunlight. In PVGIS solar radiation data from the satellites have been calculated for different spectral wavelengths to identify spectral changes and their effect on energy output in the PV system. Panel efficiency depends on two factors, solar irradiance and panel temperature. The efficiency normally decreases with increasing temperatures, but for most module types the efficiency is almost constant between 400W/m^2 to 1000 W/m^2 (at constants module temperature).
When the sun shines on a panel, the temperature will normally rise, well beyond the surrounding air temperature. The module temperature will depend on air temperature (T_a), irradiance (G) and wind speed (W). (PVGIS, 2017)

Before the energy can be converted into useful electricity, some other losses occur, such as inverter losses when converting the power into AC for grid connection, and cable losses, but these losses are not calculated in PVGIS, but users can add this input by them self. Other losses are panel degrading with age, and PVGIS suggests a number of 0,5 % of power per year. These system losses are not calculated individually by the simulation tool, but PVGIS recommend a total system loss of 14% that the user can give as input before simulating a system.
Other losses that are not accounted for in PVGIS simulation are for example snow covered panels. Even if only a part of the panel is covered by snow, the power output is generally low, and the losses due to snow are difficult to model, because it depends on how much snow, the melting process, the incline of the panel etc. Additional factors that causes loss are dusted and dirty panels, and partial shadowing of a panel, that may reduce the panel output strongly. (PVGIS, 2017)

The power output in PVGIS is calculated based on the following formula:

$$P = \frac{G_p}{1000} * A * \eta\ (G,\ T_m)$$

Where G_p is the irradiance, A is module area and η is the efficiency that is dependent on irradiance and module temperature. The efficiency is then calculated using more complex formulas with module temperature, irradiance and different coefficients depending on type of PV technology. (PVGIS, 2018)

3.5 Simulation on a designed model:

In order for this study to be complete, a comparison of the gathered production data and calculations done in different simulation tools needed to be done. When investing in PV systems there will most likely be a calculation done by the company or supplier of the panels that will tell the customer how much energy can be expected to be produced on a monthly or annual basis. This is key for investments and economical calculations that needs to be done in order to fully understand the potential of PV power. The reasons for doing these simulations are to compare these to the actual production data for different regions in Sweden, in order to see how big the variations can be between real production and simulated values.

The simulations have been performed in the public access program PVGIS and the commercial software PVSYST, and a review and comparison between the calculations on power output between these two programs have been made. In order to compare these two programs between themselves and the gathered production data, the input data and the working conditions needs to when simulating the production. This means a set azimuth to the south, a fixed tilt on the panels, and to put the plants in close proximity on the maps. For example, if the yield should be compared from plant metadata in Jämtland and the two simulation tools, the plant should not be put in a mountain valley in PVGIS, and a flat field in PVSYST. The comparison that is most key to compare is the specific yield, that is the energy in kWh per year for each kW power. This means that the easiest way of simulating this is to calculate for a plant that is 1 kW, meaning that the total annual production will be in kWh/kW. This can easily be done in PVGIS, but not in PVsyst due to the fact that a specific number of panels will need to be built in the model, and this will most likely not add up to exactly one kW peak power. Things that will result in differences between the input data between the two simulation tools are that they will use different databases for solar irradiation.

4 PV SYSTEM YIELD STUDY

This part of the report focuses on how the study, besides the literature study, was conducted. A thorough insight in how the data collection for metadata was done, and later sorted will be explained. Also, the calculations that was done in order to calculate the yield from this collected data will be described. The last part of the study will explain how the simulations in the two different simulation tools was set up and performed.

4.1 Data collection and sorting

A major part of this study is to compare data from real PV plants from two different years, in this case 2017 and 2018, to investigate any differences in production yield between the two years, and also between different parts of Sweden. The production data was taken from two different sources, Check Watt, and Solar Edge open database. In addition to these two databases, a small amount of data was collected from Sunny portal for the county of Blekinge, due to that Blekinge only had two PV plants in the other two databases, and that number was deemed too low for any statistical relevance for this thesis work.

4.1.1 Data from Check Watt

This section will focus on how the data of the PV system yield in Sweden was collected, and how the data was sorted and used for evaluation of the total yield for all the different regions in Sweden. The first data that was collected was from Check Watt with help from Bengt Stridh, senior lecturer at MDH. Check Watt is a Swedish company that offers products in energy production and consumption, with products such as electricity meters, electrical certificate metering and more. The data from Check Watt was bundled together into two files that was imported to Excel. The first file described the plants as in Table 1:

Table 1 Example of plant data table

Index	Zip code	City	Peak AC [kW]	Peak DC [kW]
3	73850	Norberg	25	26,88
5	74972	Fj‰rdhundra	20	20
7	83246	Frˆsˆn	10,22	10,22

An index for each plant, a zip code, city name, Peak AC power and Peak DC power for every plant. The other file contained the monthly production of each plant with data for every month of each year that the plant had data for in the following way:

Table 2 Example non sorted production data

Index	Year	Month	Energy Production [Wh]	Missing hours
3	2017	9	42386	704
3	2017	10	904397	135
3	2017	11	331146	0
3	2017	12	91145	0
5	2017	1	187436	17

The file with monthly data had all the data for each month labeled from 1-12 for every year on separate rows, meaning that each plant had at least 24 rows for 2017 and 2018 if that plant had data for every month of these two years. Many plants had data for more years that the desired years of 2017 and 2018. The plan was to get the data sorted in a way that would be easy to survey for each plant, so in this case that meant to sort each plant (index) in a single row, with plant data and monthly production data on the same row with columns for every month. The following layout pictured below, was the desired one in order to easily survey production data for each month together with plant information data. The way that this was conceived will be described below in a data sorting section.

Table 3 *Example of sorted combined plant info and production plant size*

Index 2	Zipcode	Location	Peak AC [kW]	Peak DC [kW]	Jan |Wh]	Feb |Wh]
3	73850	Norberg	25	26,88	111974	5537
5	74972	Fj‰rdhundra	20	20	127464	474926
7	83246	Frˆsˆn	10,22	10,22	0	403

The data sorting strategy is explained more in detail in APPENDIX 1, but the overall goal was to sort the data for every PV system with a total monthly electricity yield for 2017 and 2018 as seen in the Table 3 layout.

4.1.2 Data collection from Solar Edge

The data collected from Solar Edge had to be done manually from the open database on Solar Edge homepage (Solar Edge, 2019), due to being unable to collect any files directly from Solar Edge as in the case with Check Watt. On the Solar Edge homepage there is an open database storing plant information that can be accessed by everyone. On April of 2019 when the data was collected, the open database contained 525 plants in Sweden. The strategy to extract the data from the portal was a manual approach, hence it being deemed to easiest, but most time-consuming alternative. On the Solar Edge portal, a search for every plant in Sweden was made, and the approach was to start from the top, and work manually through each of the 525 plants. The first thing that was done was to click on the plant name and look for plants that had full data for 2017 and 2018, that could be seen on a graph, otherwise that plant was dismissed. Then, if the plant met the criteria of full monthly production in 2017 and 2018, the plant essential data, such as zip code and power (kWp) was inserted into an Excel file. Then the production data was downloaded as a csv file and copied into the Excel file containing plant data. The actual process is described below in APPENDIX 2:

4.1.3 Solar irradiation data

The solar radiation data was taken from SMHI, that has a database consisting of ground measurement data of global radiation. The location of the measurement site is sparse, meaning that not every county in Sweden are represented. The radiation data could be downloaded from SMHI as a .csv file, that could be open in Excel. The data represents ground station values that could be found in 11 of the 21 counties in Sweden. The data was represented as shown below, with hourly irradiance data shown in Watts per square meter. The quality means that the data is either verified or not.

Table 4 *Global Irradiance SMHI data example*

Date	Time (UTC)	Global Irradiance [W/m2]	Quality
2018-03-21	12:00:00	486.22	G
2018-03-21	13:00:00	505.92	G
2018-03-21	14:00:00	432.14	G

The data was downloaded and it had hourly data for every hour of each year from 2008 to the beginning of 2019, so all the rows containing data that was not 2017 and 2018 was sorted out, and the sum of every month for 2017 and 2018 was calculated to get an overview of the solar irradiance for each month in kWh per square meter.

For all the counties that does not have sites with ground measurements, an online model called STRÅNG was used. The data are from SMHI and produced in collaboration with the Swedish Radiation Protection Agency and Swedish Environmental Agency. The model calculates a radiation quantity based on interpolated values from field data by using a bilinear method. The error for yearly model data is less than 10 % when comparing to a ground observation measurement on global irradiation. The mean value calculation on global irradiation was based on data from 1999 to 2015 (STRÅNG,2019).

4.2 Method of calculations

Most of the calculations in this study was done in the simulation software's, hence not much difficult calculations were done manually. Most of the calculations done otherwise revolved around mean value calculations and yield performance under average solar irradiation as shown below.

4.2.1 Average yield 2017 and 2018 adjusted to average solar radiation

In order to compare and analyse how the solar radiation correlates to the specific yield for the year of 2017 and 2018, it had to be adjusted to an average solar radiation for either the county or all of Sweden if that is to be investigated.

$$Adjusted\ specific\ yield = \frac{Specific\ yield(Year\ x)}{Solar\ irradiation(Year\ x)} * Average\ Solar\ irradiation$$

The solar irradiation for each specific year was gathered from hourly data from SMHI, as shown below in section 4.2.2. But in order to compare to an average solar irradiation, an average of more years need to be compared. This can be done by downloading a file from SMHI that stores monthly values of irradiation from 1983 to 2014. An average annual value of those years was calculated and used for average solar irradiation.

4.2.2 Solar radiation data calculation

The solar radiation data was taken from SMHI hourly radiation datasets that can be downloaded for a few specific places in Sweden. This data does not account for every county in Sweden, and therefore all county's without data on solar radiation has been calculated using a model called STRÅNG, as mentioned in section 4.1.3.

4.3 Simulations

The simulations were done in order to compare simulated values on specific yield to the actual yield data taken from the databases. Simulations were made on both the online tool PVGIS and the software PVSYST that is a more complex software that is not free. These two tools were chosen because of their differences in complexity and target users. The PVGIS tool is easy to use, even with very limited knowledge about PV systems and modeling, and it is free to use, meaning that for example home owners thinking about investing in solar can make easy calculations on production values. The PVSYST software is more complex, targeted more at professional users and companies, and it is interesting to see if this software will vary to the PVGIS simulations, and how close or for away from the reality of the actual yield data these simulations really are. In the literature review, there are a number of parameters that are discussed to have an effect to the power output of a PV plant, and these are the panel tilt, azimuth, irradiation, installed power and shading (Killinger et.al, 2018). In order to simulate PV systems that will reflect the real world in the best possible way, the panel tilt was simulated both as a fixed and an optimal tilt angle. The fixed tilt angle was based on the average roof tilt in Sweden, that was found to be 31. The azimuth was set to the south (0°) for all simulations. This because that is the best position for optimal yield, and information about the average azimuth angle for Swedish plants are difficult to obtain.

4.3.1 PVGIS Simulations

The simulations done in PVGIS was made after some initial input data was obtained. In order to get the best conditions possible that reflects the reality in a good way, information about radiation data and roof tilts had to be obtained. There are not many things the user can change on a roof mounted panel, but these are important to get right, for simulation accuracy. These are the parameters that can be changed for roof mounted panels:

- Radiation database
- PV technology type
- Installed peak power
- Estimated losses
- Slope
- Azimuth

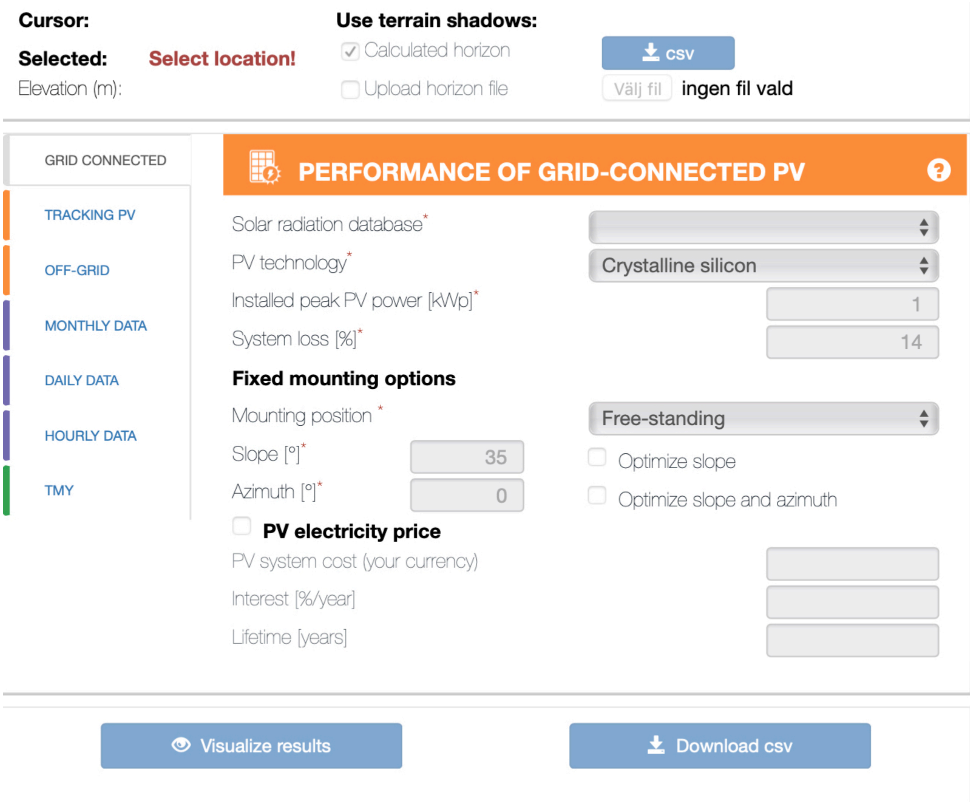


Figure 8 PVGIS simulation tool, input variable screen

The PV technology used for the simulations are the most common crystalline silicon. Installed peak power in this case is 1 kWp, because that is what we want to compare (kWh/kWp). The estimated losses are according to PVGIS developers recommended to be set at 14%. This number is can be compared to some of the numbers discussed in the literature review where for example the report by Leloux et.al (2011) found that on average the real PV systems yield was 15% lower than optimal PV systems. They also found that soiling and shading accounted for about 5% of losses. The other report by Kilinger et.al (2018) found irradiance losses of 7,3 % due to shading on buildings in Uppsala. Other losses not investigated fully is the losses due to the inverters, but overall the 14% losses suggested by PVGIS seems reasonable.
The Solar radiation database used was the PVGIS-ERA5 database, that because of the fact that there is no other option above 60° north. The slope of the roof was set at a value that would represent the most amounts of roof in Sweden, meaning finding the average roof tilt of Swedish roofs, and this value was found to be 31° (Kamp, 2013). Both the average slope on the tilt and the optimal slope was simulated.

The input values were as follows:

- Solar radiation database = PVGIS ERA-5
- PV technology = Crystalline Silicon
- Installed peak PV Power = 1 kWp
- System losses = 14% ´
- Mounting Position = 31° for one simulation, and optimal slope for another simulation
- Azimuth = 0°

Each simulation was done for every county (Län) in Sweden. The actual location in every county was chosen to be the biggest city of the county, because that is where most people live, in an intent to reflect the reality in the best way possible. Different locations in the same county was found to have minimal differences in yield, exceptions to this is of course big county's such as Norrbotten or mountain areas with high valleys etc.

4.3.2 PVSYST Simulations

The PVsyst simulation procedure started by choosing the project design part of the software, then choosing to model a grid connected system. The goal here was to choose to do a simulation that would have the same kind prerequisite's as PVGIS, that means to do the simulation at the same locations with the same power. But the way you set up the simulation in these two programs is different. In PVGIS the user can choose to calculate for example a 1 kW system, and that will in this case result in the required specific yield kWh/kWp, but that is harder to do in PVsyst. In that program the user will choose exactly the type of panel and inverter that will be simulated, from a list of actual panels and inverters. In order to do a simulation of systems that actually exists in real life, a peak power of the system of 5 kW was chosen, and then the produced electricity was divided by the installed power to get the specific yield. Due to the fact that the same type of system would be simulated for both a set tilt angle and an optimal, a way of storing the same system, but choosing different locations and tilt angels was done. A detailed process description is showed below, based on explanations in the project design window shown below.

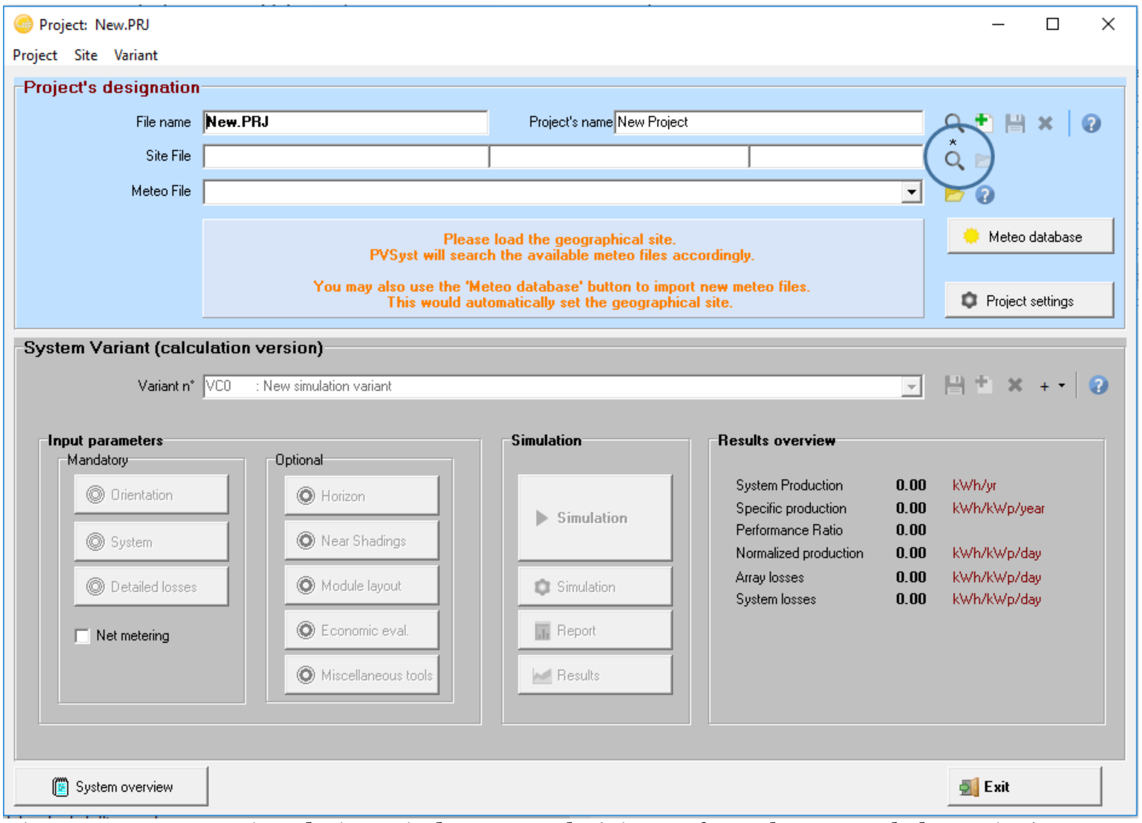


Figure 9 PVsyst Project design window example (Picture from the PVsyst help section)

- In the Project name, the current location for the simulation was written
- In the Meteo database, the location for the simulation was searched for. The database has some predefined locations based on ground station data, but as shown in the literature study of PVsyst, it does not cover location in every county in Sweden, so in that case they have to be generated. This was done in the first page of the PVsyst tab, where a tab called Databases can be clicked.

- In the Databases, Geographical sites is clicked. Here, the coordinates for the locations can be filled in. The location was the same town that was used for PVGIS for every county.
- When the coordinates have been filled in, meteorological data are imported by clicking Meteo data import. Then the file is saved for every location that does not have ground station data by default in the tool. The data is generated and saved hourly.
- Now the input parameters can be done, and this starts with the Mandatory part as seen in the above picture, Orientation and System. The orientation was set at a fix tilt at 31°, as well as an optimum tilt angle that could be shown graphically in the Orientation tab. The azimuth was set at zero degrees (south).
- The next part is the System, that is where the PV panels and inverters and the overall layout of the system is defined.
- The plan was to use a system of normal private home size, not to small, not to big, in order to reflect reality in a good way. A system size of 5 kWp was chosen. Modules of 250 Wp was chosen, and the exact model was Jinko JKM 250P-60.
- The panels needed to match the 5 kWp was 20 pieces, and the layout was 10 panels in each row. This meant an inverter that could handle two strings of panels needed to be chosen.
- The chosen inverter was SMA Sunny Tripower 5000 TL-20, and this needed to be set up to handle two sub-arrays of 10 panels each in series.
- After the system was built, the simulation can be started. There are optional parameters such as horizon far shading that can be implemented, as well as near shading objects, but in this simulation, none of this was implemented due to unknown shading profiles for every location.
- The simulation started with simulating each variant, first the set tilt angle at 31°, then the optimum angle for every location. The result for every simulation was saved by exporting the result tables into an Excel file.
- This was done for all the 21 counties in Sweden. The system profile could be saved so that the panel and inverter, and the layout of the panels did not have to be made for each project simulation.

5 RESULTS

The result section will display the result of the data gathering to showcase the specific yield of the 2017 and 2018 data. This data will be compared to results from simulations on specific yield. In order to display reasons for differences in specific yield, global irradiation data are shown and compared to yield data. For understanding and validation of the compared data, the compared number of systems and installed power of these systems are shown.

5.1 Specific yield 2017 and 2018 from real plant data

The comparison below are based on production data from 3 different databases, all data from systems that doesn't contain data from both years have been sorted out, meaning that only systems that were active in 2017 and 2018 have been used to produce the specific yield data shown below. But this also means that data has been sorted out from 2018 due installations that has been made in 2017 that will only show in 2018.

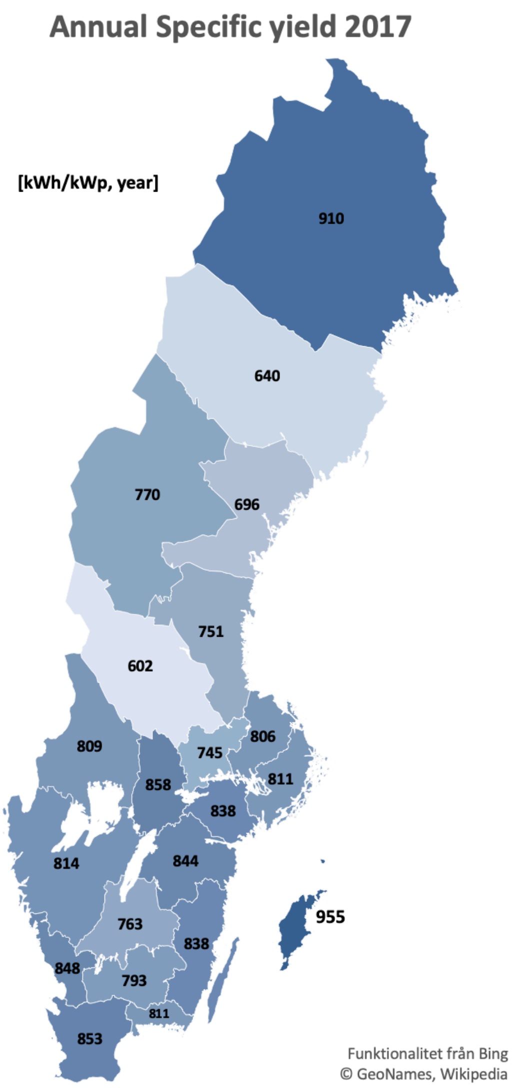


Figure 11 Specific yield real plant data 2017

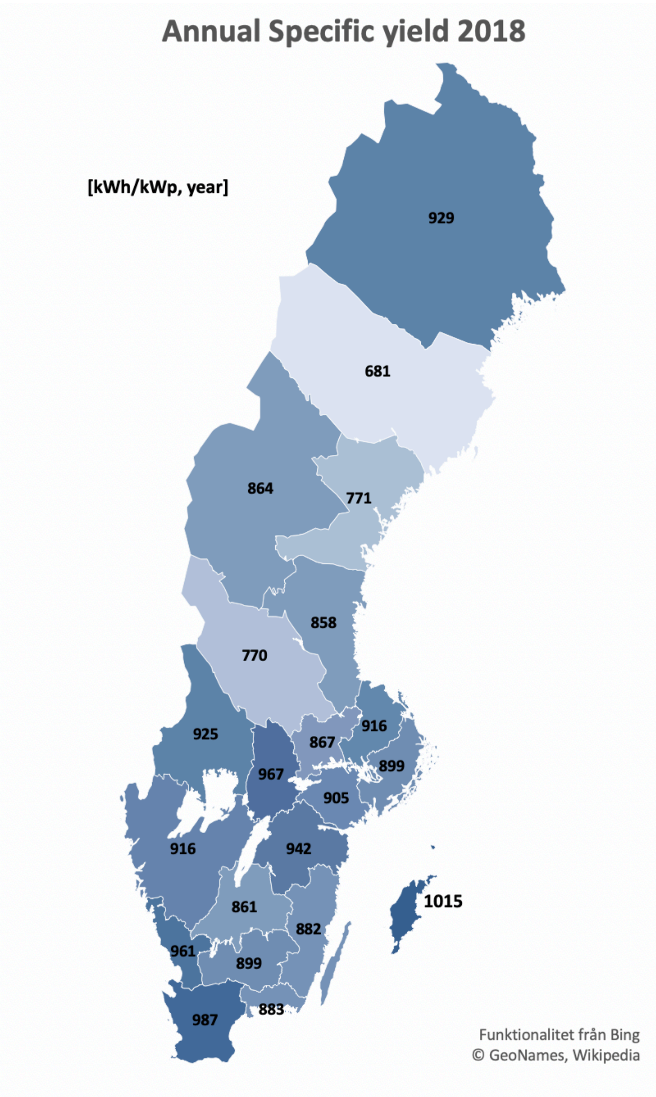


Figure 10 Specific yield real plant data 2018

The figure 10 and 11 show the specific yield in kWh per installed kW of power. They show data based on real system performance from 2017 and 2018, to showcase the differences between each year. What stands out is the high yield in Norrbotten at over 900 kWh/kWp. This is high compared to every other county in the north. This county has yield data based on just a few plants, with every plant showcasing high yield numbers. The maximum value for 2017 can be found in Gotland at 955 kWh/kWp, and the lowest in Västerbotten at 640 kWh/kWp. For 2018, the maximum is in Gotland again at 1015 kWh/kWp, and the minimum in Västerbotten at 681 kWh/kWp.

Table 5 Specific yield average for Sweden 2017 and 2018

	Specific yield average for Sweden [kWh/kWp]	Number of plants	Installed power [MW]
2017	798	828	13,8
2018	891	828	13,8

5.1.1 Total yield 2018 with full data

Figure 12 contains all the available data from every system that could be found containing full data for 2018, unlike the 2018 figure above that has system data sorted out for systems that doesn't contain 2017 data. This means that the data from the 2018 will be more statistically accurate due to almost double the amount of data.

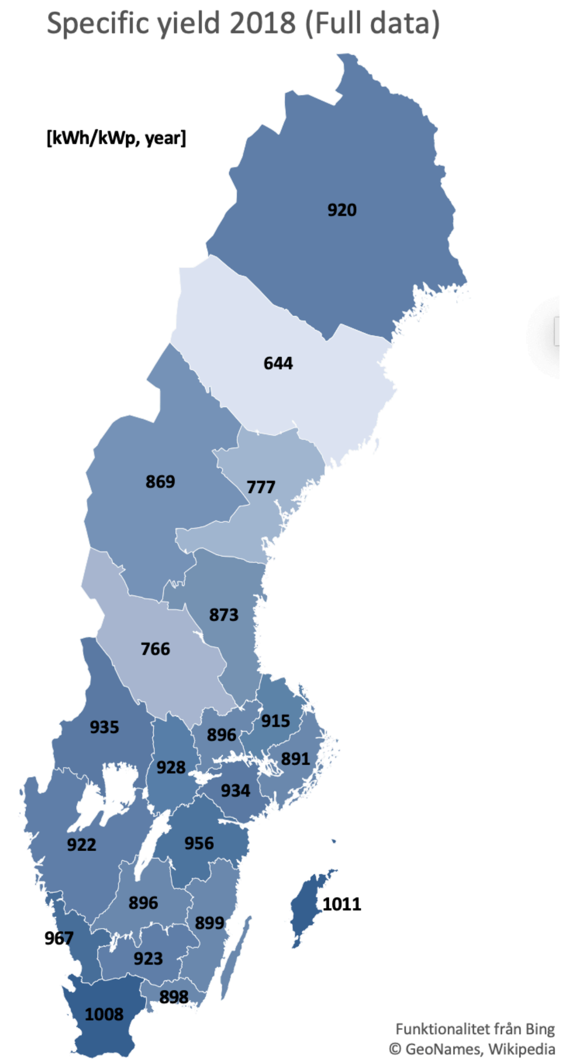


Figure 12 Specific yield 2018 with all available system data

Table 6 *Specific yield average Sweden (2018 full data)*

	Specific yield average for Sweden [kWh/kWp,year]	Number of systems	Installed power [MW]
2018 full data	897	1381	23,9

The average yield for Sweden when comparing the 2018 data with the less amount of systems and this 2018 full data are only 897 kWh/kWp compared to 890 kWp/kWp. But this will vary a lot on between the counties. A bigger part of the total amount of PV systems in the compared data will result in better statistics.

5.2 PVsyst simulations

The specific yield for the PVsyst simulations is shown in figure 13 and 14, with figure 13 showing the annual yield by simulating the systems with an optimized yield for every location. This optimal yield was found by an optimization graph built in into PVsyst, where a graph shows losses due to panel tilt. The exact panel tilt was hard to find due to the size of the graph and the manual approach, meaning in this case that the optimal tilt could differ 1-2 degrees and still show optimal tilt losses. Therefore, the optimal tilts shown below will not be accurate to the exact degree.

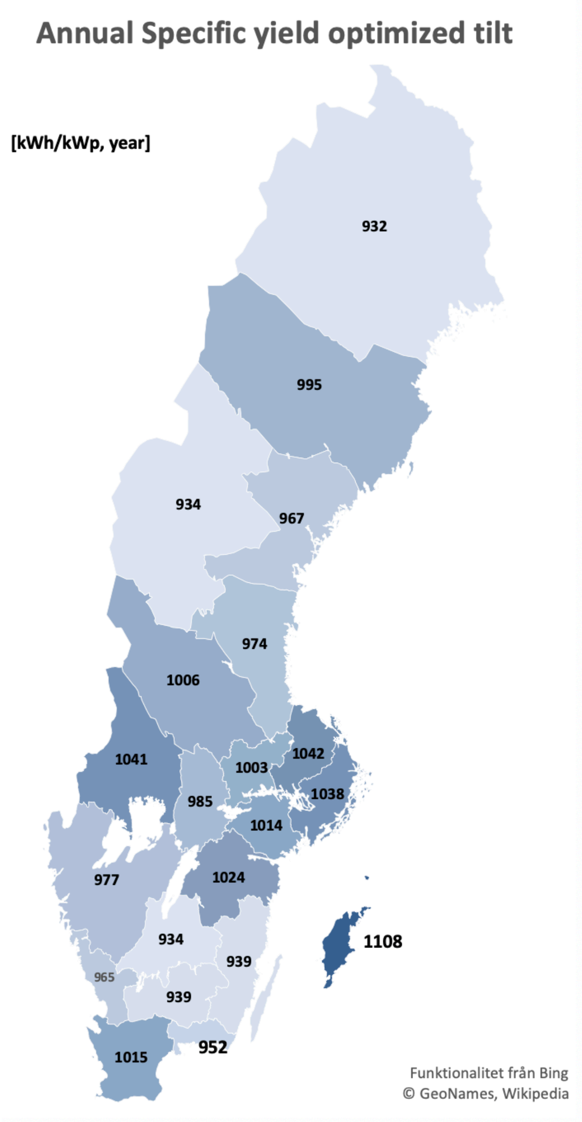


Figure 13 Specific yield PVsyst optimized tilt

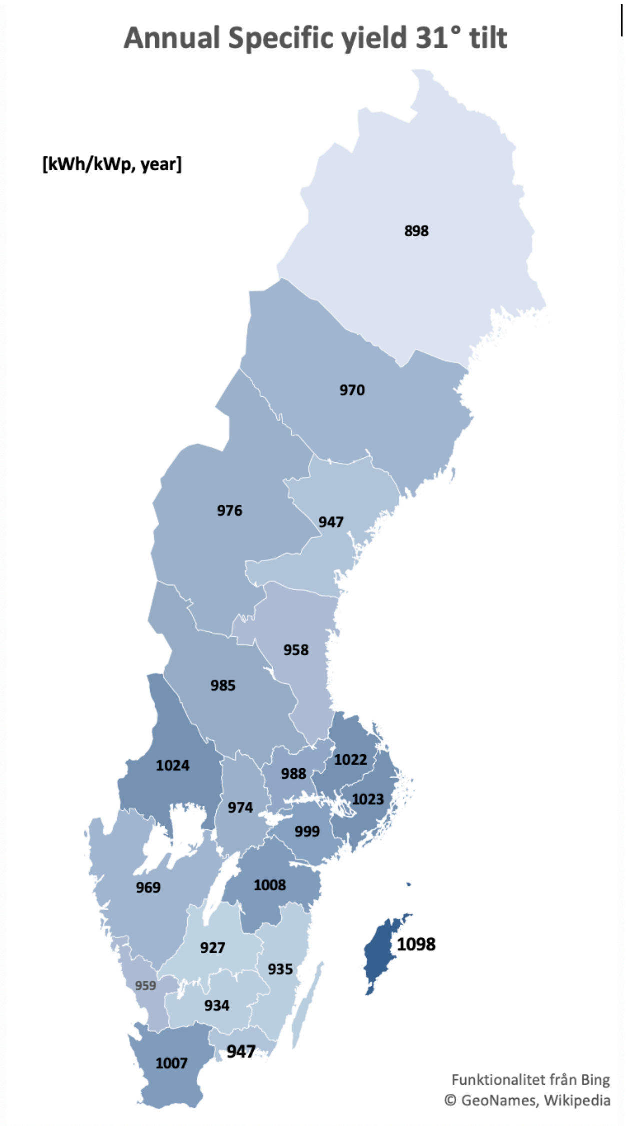


Figure 14 Specific yield PVsyst 31° tilt

Figure 15 show the optimal tilt angle according to PVsyst simulations. In PVsyst a graph will showcase the optimal tilt, but the optimal tilt will be optimal for several degrees, and hence these numbers will be good even at +- 1-2 degrees. Optimized tilt angles vary from 40 degrees in Blekinge, to 49 in Jämtland and Västerbotten.

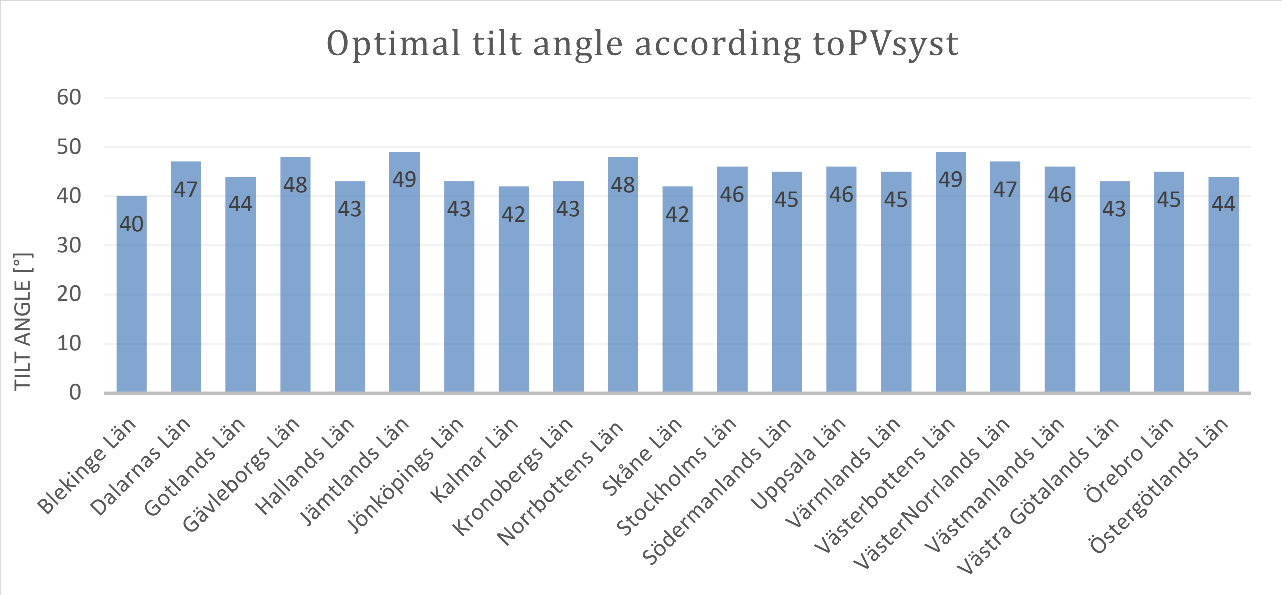


Figure 15 Optimized tilt angle from PVsyst simulation

5.3 PVGIS simulations

The figures above show the specific yield for the set tilt angle of 31°, as well as the yield for optimal tilt angles, for the simulations done in PVGIS. Maximum values for 31° tilt are found in Gotland at 1141 kWh/kWp, and minimum in Norrbotten at 786 kWh/kWp. For optimized tilt the same numbers are 1167 kWh/kWp for Gotland, and 816 kWh/kWp in Norrbotten.

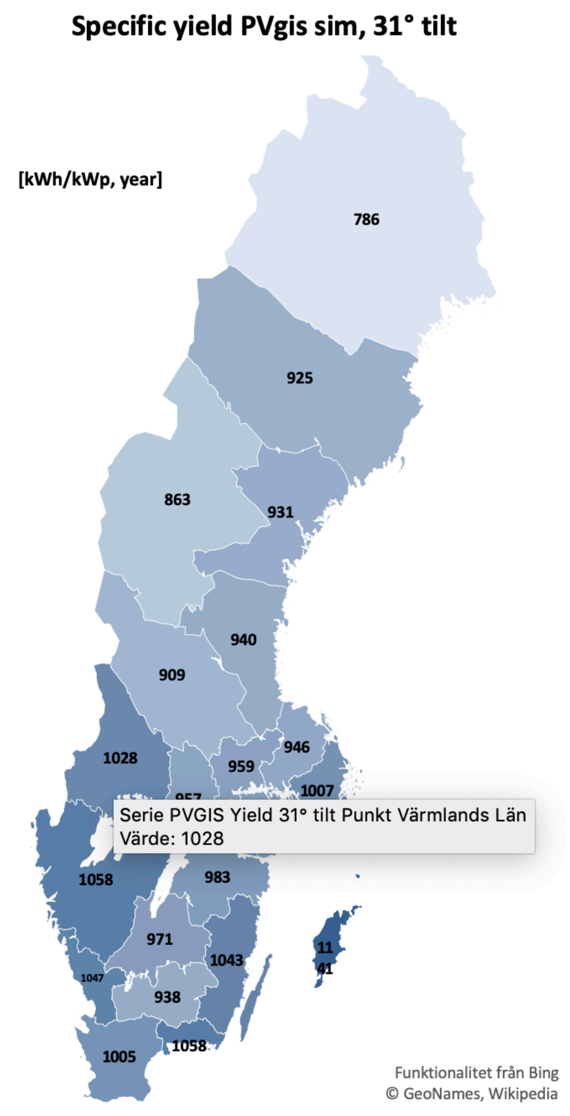


Figure 16 PVGIS 31° tilt simulation

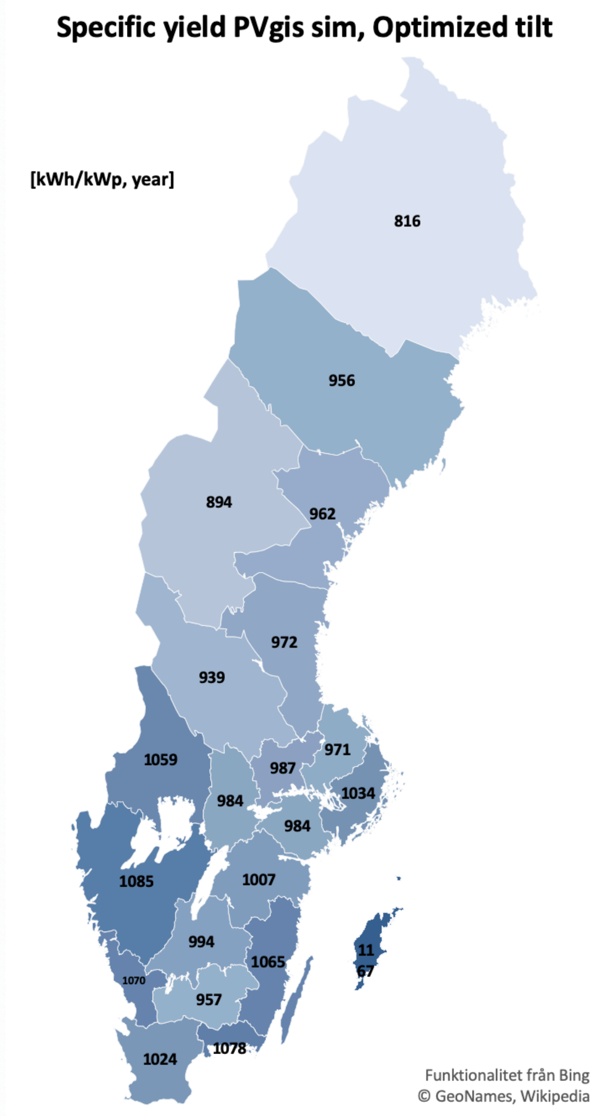


Figure 17 PVGIS Optimized tilt simulation

5.4 Global irradiation from SMHI and STRÅNG model

The global irradiation figures 19 and 20 show the global irradiation on a horizontal surface in kWh per square meter for a year. The data was based on measurements taken from ground stations that could be obtained from SMHI. An important thing to notice as explained in the "Description of the study" is that only 11 of the 21 counties will have irradiation values based on real measurements from SMHI ground station measurements. The other data are based on calculations by an irradiation model called STRÅNG. More on this is explained at Table 7.

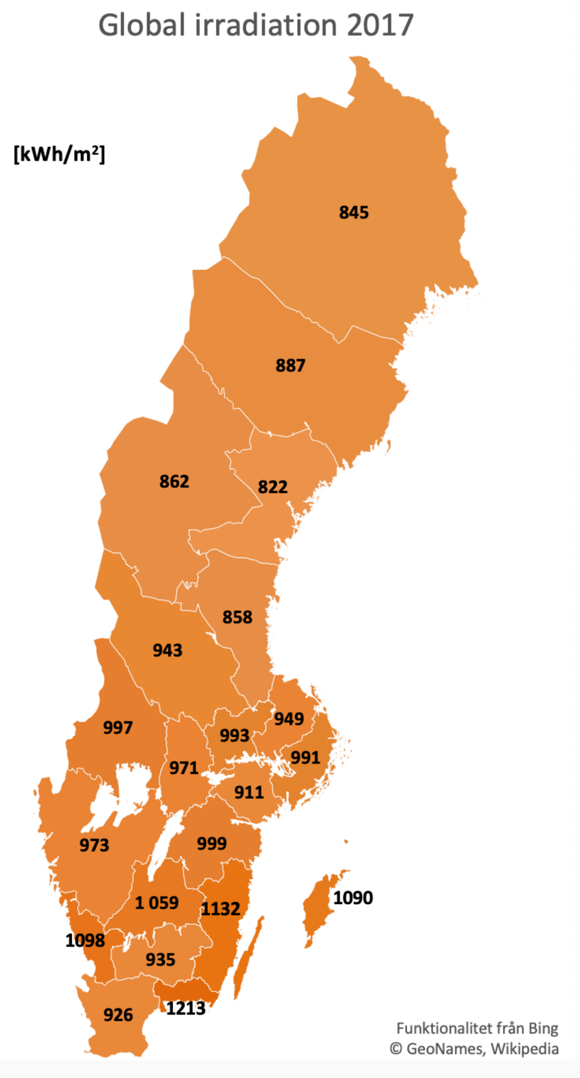


Figure 18 Global horizontal irradiation 2017

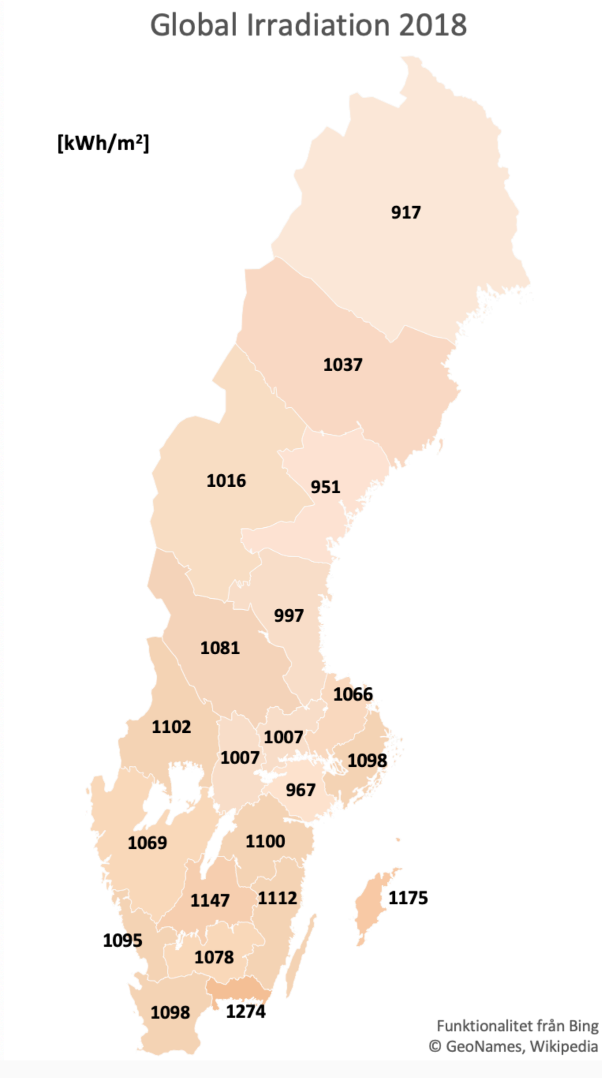


Figure 19 Global horizontal irradiation 2018

Table 7 show the irradiation data in kWh/m^2. The data was taken from SMHI irradiation database, that store hourly values. The data was added summed up into annual values for 2017 and 2018. Some data is based on ground station irradiation measurements, and all the **red** values are data that are based on the STRÅNG model. More in this in the 4.3.2 section on the "Method of calculations".

Table 7 Global horizontal irradiation comparison from 2017 to 2018

County	2017 [kWh/m2]	2018 [kWh/m2]	2018 to 2017 irradiation [%]
Blekinge Län	1 213	1 275	5,1%
Dalarnas Län	943	1 081	14,6%
Gotlands Län	1 090	1 175	7,8%
Gävleborgs Län	858	997	16,2%
Hallands Län	1 099	1 095	-0,3%
Jämtlands Län	862	1 016	17,9%
Jönköpings Län	1 059	1 147	8,3%
Kalmar Län	1 132	1 112	-1,7%
Kronobergs Län	935	1 078	15,3%
Norrbottens Län	845	917	8,5%
Skåne Län	926	1 098	18,5%
Stockholms Län	991	1 098	10,8%
Södermanlands Län	911	967	6,1%
Uppsala Län	949	1 066	12,3%
Värmlands Län	997	1 102	10,5%
Västerbottens Län	887	1 037	17,0%
VästerNorrlands Län	822	951	15,7%
Västmanlands Län	993	1 007	1,4%
Västra Götalands Län	973	1 069	9,9%
Örebro Län	971	1 007	3,7%
Östergötlands Län	999	1 100	10,1%

Figure 20 shows the irradiation data that was gathered from both PVsyst simulation tool, and SMHI annual data on irradiation. The PVsyst data are a mean value from several years based on the built-in databases for climatological data in the software. The 2017 and 2018 SMHI data are described more in detail above under Table 5, and these are based on ground measurements and STRÅNG model calculations. Similarities between PVsyst and the 2017 SMHI data can be seen, and that the 2018 solar irradiation are generally much larger.

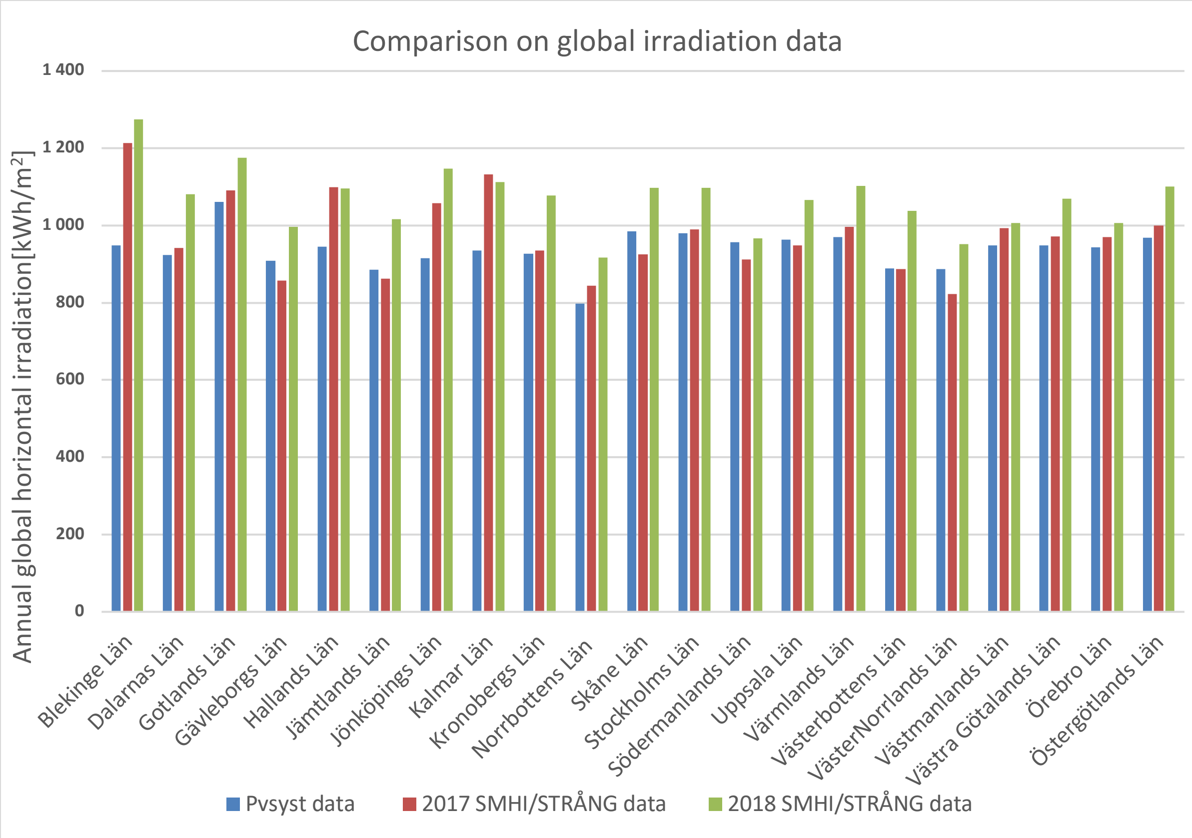
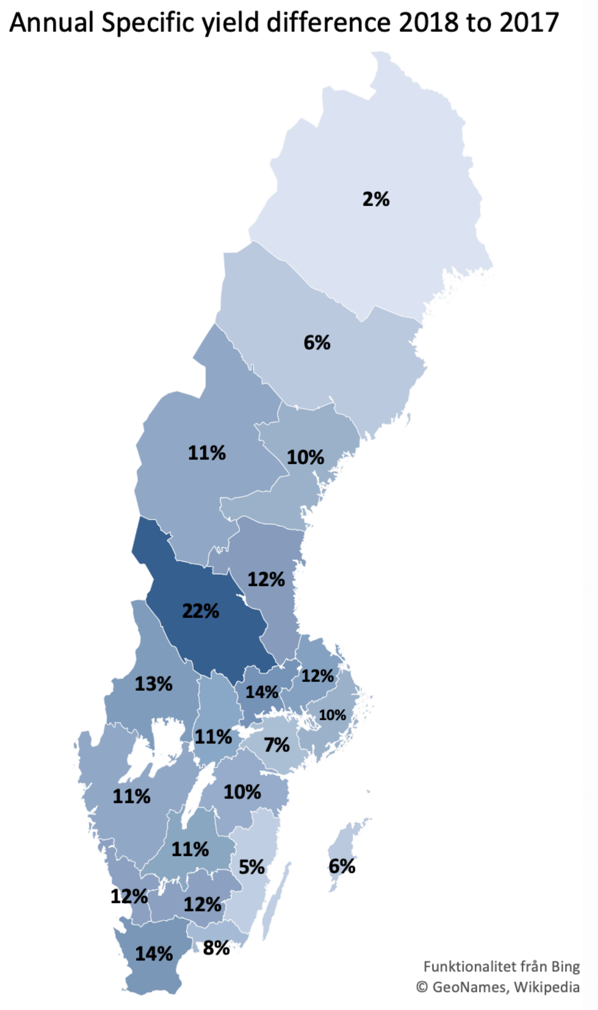


Figure 20 Comparison on irradiation data from SMHI 17/18 and PVsyst irradiation meteo

5.5 Comparison on collected data, solar irradiation and simulations

Figures 21 and 22 show the annual difference in specific yield from the real PV plant data for 2017 and 2018, with comparison of difference in solar irradiation for these years. As described in chapter 4.1.3, the solar radiation data collection, some of the irradiation values from some counties are based STRÅNG model data. This means that the actual values can differ slightly to real measured values if that location would have measurements.

Figure 21 and 22 show that overall the specific yield differences are correlating with the differences in solar irradiation for these 2 years. Some differences can be seen, especially in the north where the correlations it is not as strong as in some of the other counties.



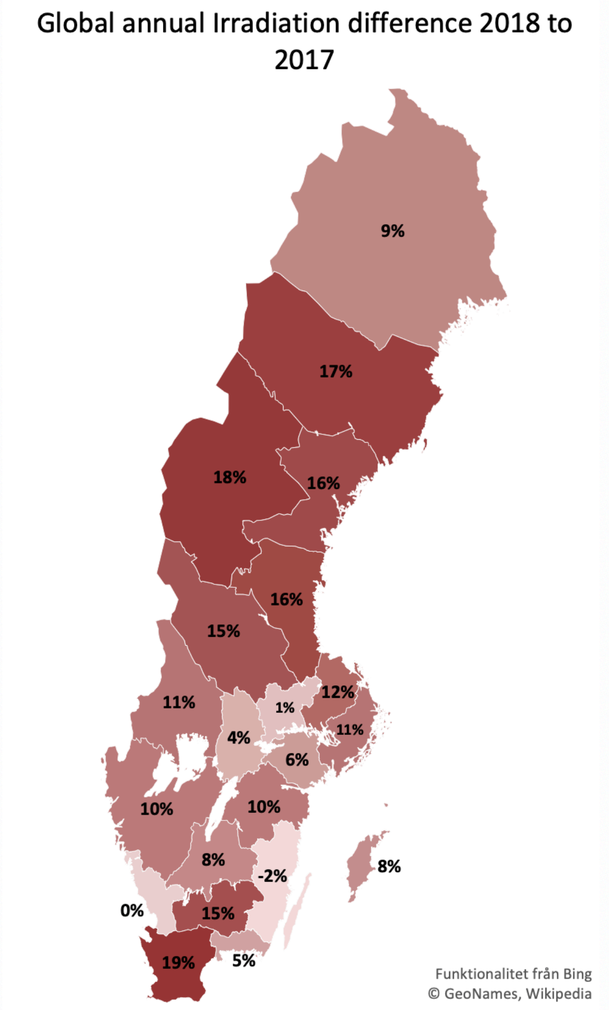


Figure 21 Global annual Specific yield comparison 2018 to 2017 data

Figure 22 Solar irradiation difference SMHI 2018 to 2017 data

Figure 23 will show the comparison of how much the specific yield differ from PVGIS values to PVsyst. The green line shows the difference in percentage between each simulated value for each county. A positive value on the green line means that the PVGIS specific yield value is bigger than for PVsyst.

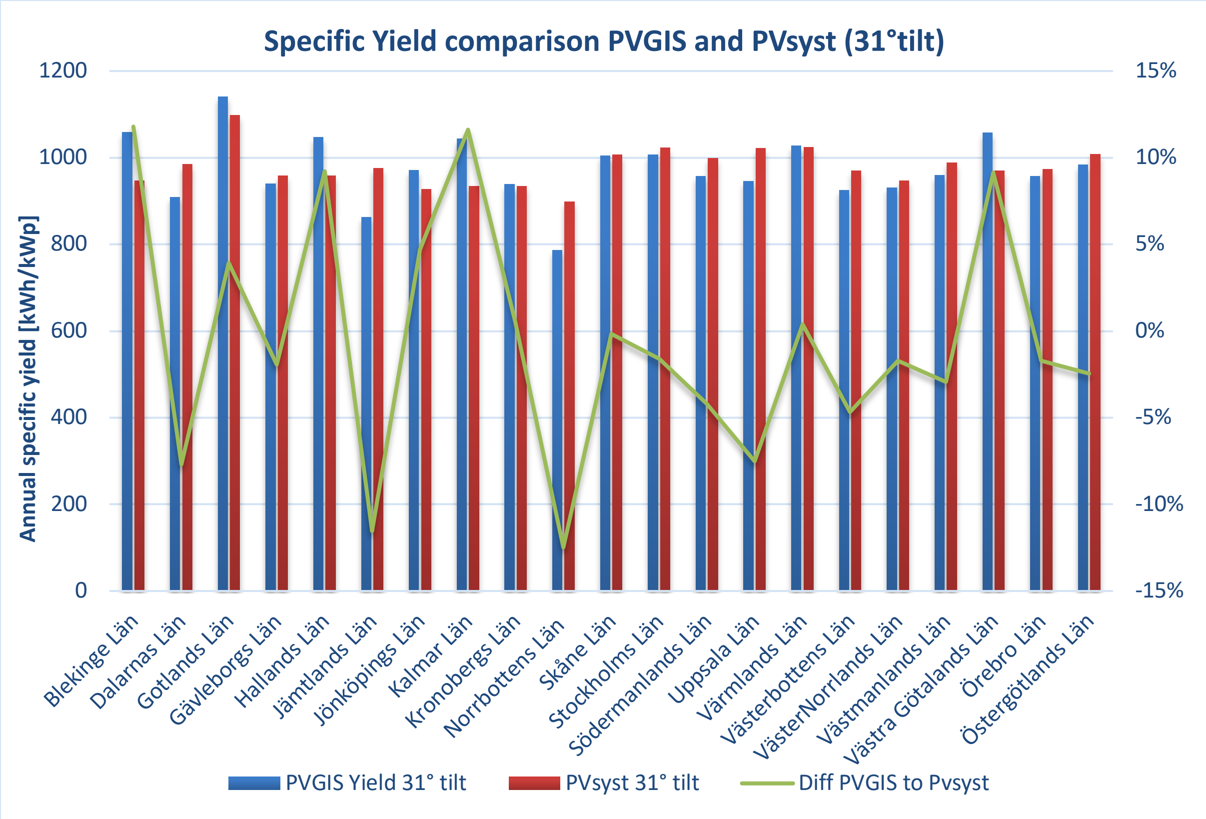


Figure 23 Comparison between PVGIS and PVsyst yield results

The table 8 will show a full comparison with annual specific yield values for both gathered real plant data and the simulations in PVGIS and PVsyst. The simulated values will have data for both optimal tilt angles and a set tilt at 31°.

Table 8 Specific yield comparison of all data and simulations

Annual Specific Yield Comparison [kWh/kWp]						
Län	**Plant data 2017**	**Plant data 2018**	**PVGIS 31° tilt**	**PVGIS optimal tilt**	**Pvsyst 31° tilt**	**Pvsyst Optimized tilt**
Blekinge	811	883	1058	1 078	947	952
Dalarna	602	770	909	939	985	1 006
Gotland	955	1 015	1141	1 167	1 098	1 108
Gävleborg	751	858	940	972	958	974
Halland	848	961	1047	1 070	959	965
Jämtland	770	864	863	894	976	934
Jönköping	763	861	971	994	927	934
Kalmar	838	882	1043	1 065	935	939
Kronoberg	793	899	938	957	934	939
Norrbotten	910	929	786	816	898	932
Skåne	853	987	1005	1 024	1 007	1 015
Stockholm	811	899	1007	1 034	1 023	1 038
Södermanland	838	905	957	984	999	1 014
Uppsala	806	916	946	971	1 022	1 042
Värmland	809	925	1028	1 059	1 024	1 041
Västerbotten	640	681	925	956	970	995
Västernorrland	696	771	931	962	947	967
Västmanland	745	867	959	987	988	1 003
Västra Götaland	814	916	1058	1 085	969	977
Örebro	858	967	957	984	974	985
Östergötland	844	942	983	1 007	1 008	1 024
Average for Sweden	**798**	**890**	**974**	**1 000**	**978**	**990**

5.6 Specific yield adjusted to average solar irradiation values

In order to look at how much the solar irradiation will affect the annual yield, a calculation was done in Excel to adjust the annual irradiation to average values. This meant that the total annual yield for both years was correlated with annual mean values from SMHI irradiation data, based on mean values of more than 30 years of ground measurements on solar irradiation.

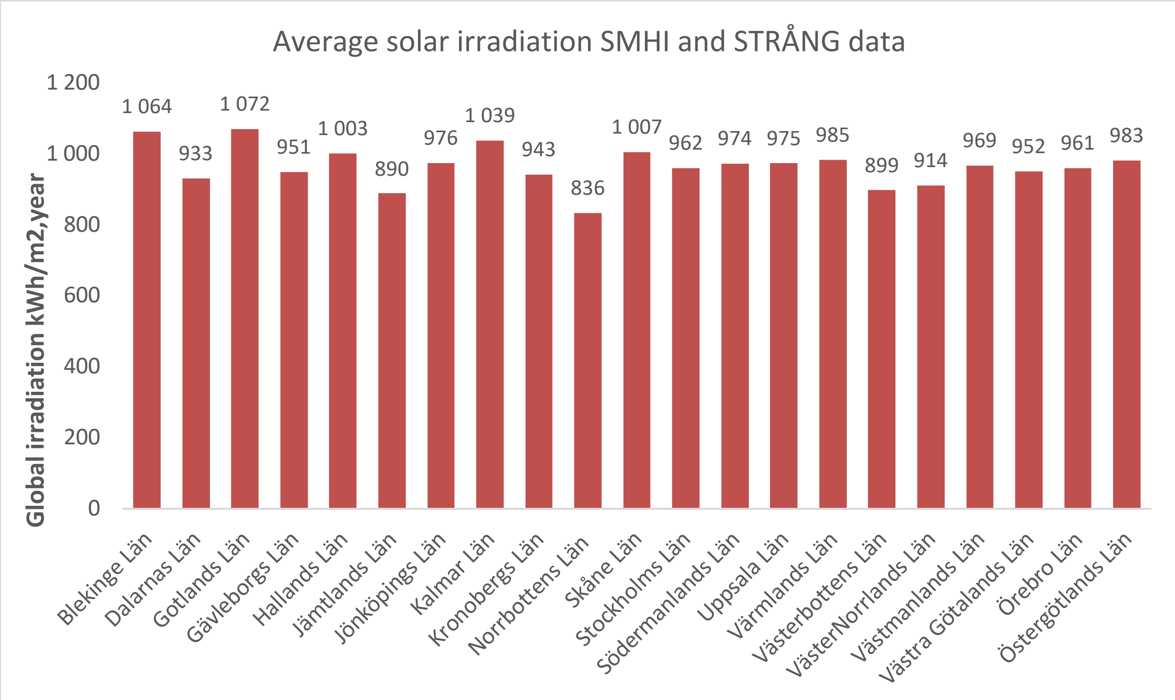


Figure 24 Average solar irradiation SMHI and STRÅNG data

Figure 25 show the 2017 and 2018 specific yield from collected data adjusted to a normal global irradiation. The mean value on irradiation are based on data from SMHI ground measurements from 1983 to 2015, and from 1999 to 2015 for the STRÅNG model. The figure 24 are meant to show how much the yield differs between the years if you remove the input of differences in solar irradiation. If the differences are small, one can say that the correlations of specific yield and solar irradiation are significant. If there are bigger differences, other local factors as non-optimal tilt and shadowing might be occurring. The specific yield values for 2017 was almost identical to an average year for Sweden as a whole, but regional differences occur.

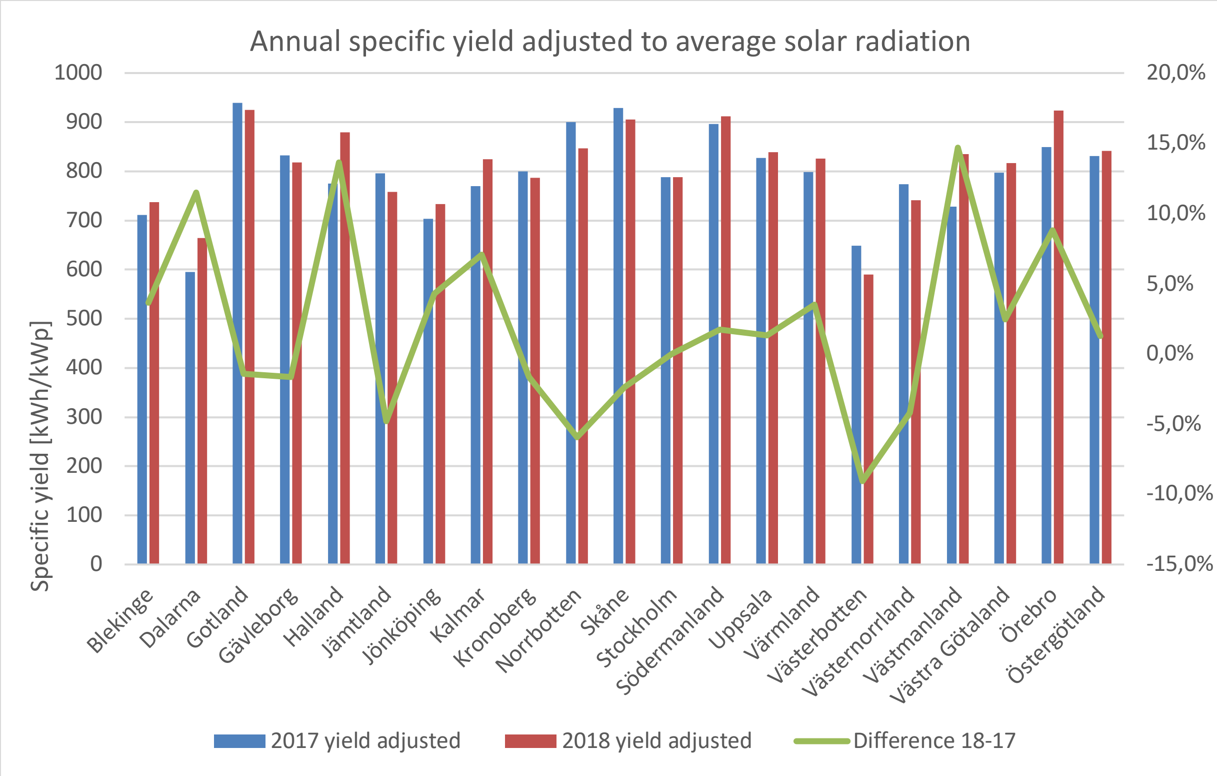


Figure 25 Specific yield adjusted to average solar irradiation

Table 9 Specific yield adjusted to average solar irradiation, average for Sweden

Specific yield adjusted to solar irradiation average	2017 yield adjusted	2018 yield adjusted
Average for Sweden [kWh/kWp,year]	795	809

5.7 Information and validation on 2017 and 2018 data

In order to understand the validity of the data presented in the 2017 and 2018 specific yield calculation results, the number of plants that represent the mean values in each county has to be shown. As seen in figure 26, the variations in the number of plants vary significantly. This will result in more or less accuracy, where bigger number of plants will represent a better statistical. The total number of systems in Sweden found in the 2017-2018 comparison are 828. For the 2018 full data, the number is 1380. In Västra Götaland the total number of PV systems are 115 for the 2017-2018 data, and 213 for the 2018 full data. For Västerbotten that number is 3 and 4 systems.

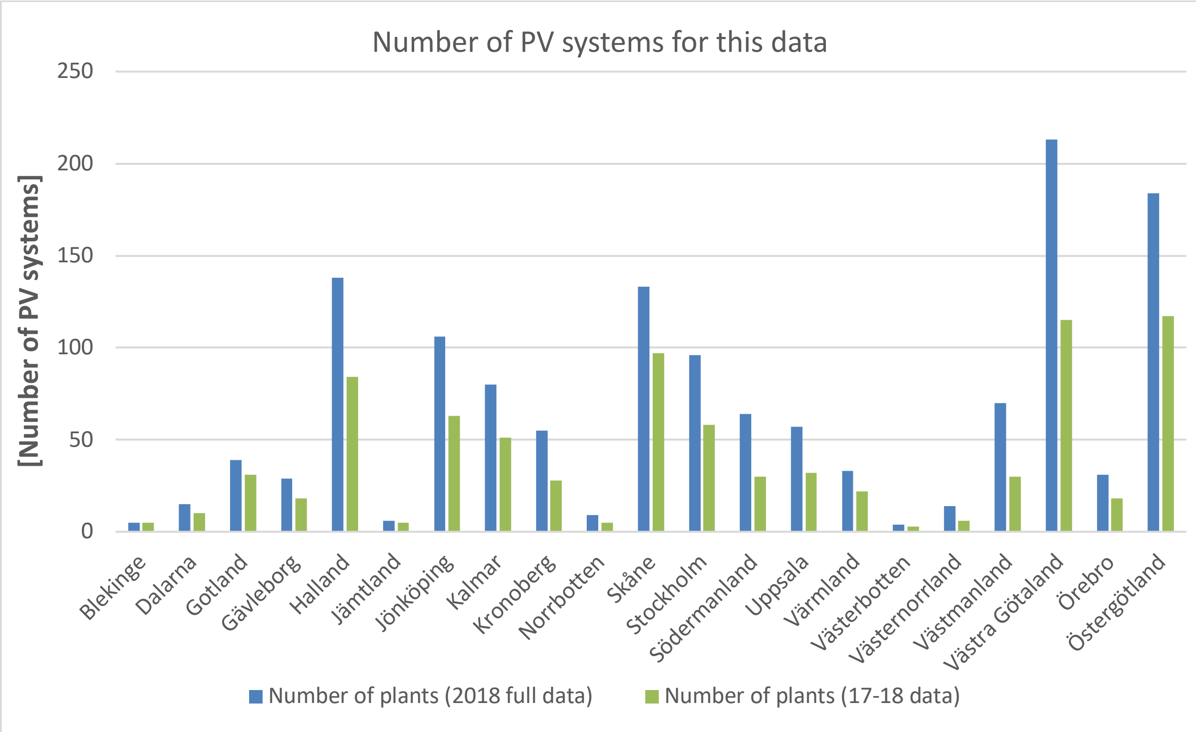


Figure 26 Number of PV systems in the sorted data

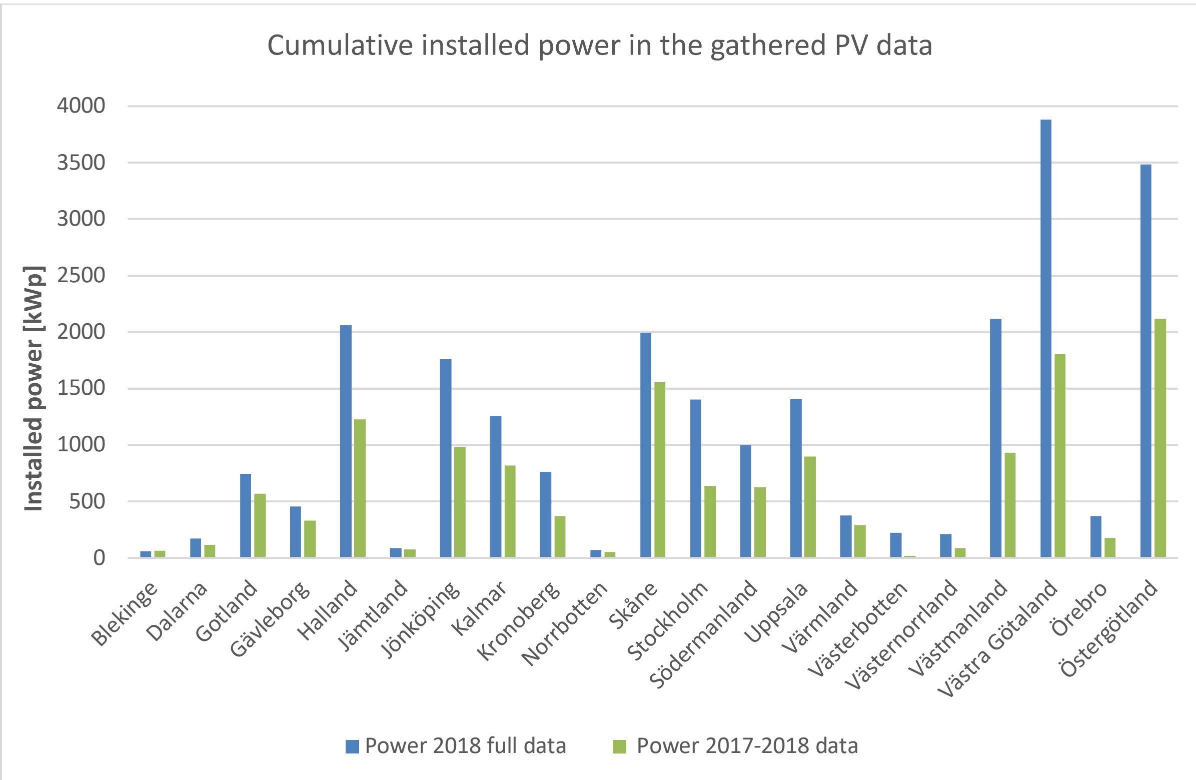


Figure 27 Installed power of the sorted data in

Another important factor to understand the significance of the results in specific yield is to look at the installed power that this sorted data consists of. This varies a lot from minimum values of 22 kWp in Västerbotten for 2017-2018 data, to maximum values of 3883 kWp in Västra Götaland for the 2018 full data.

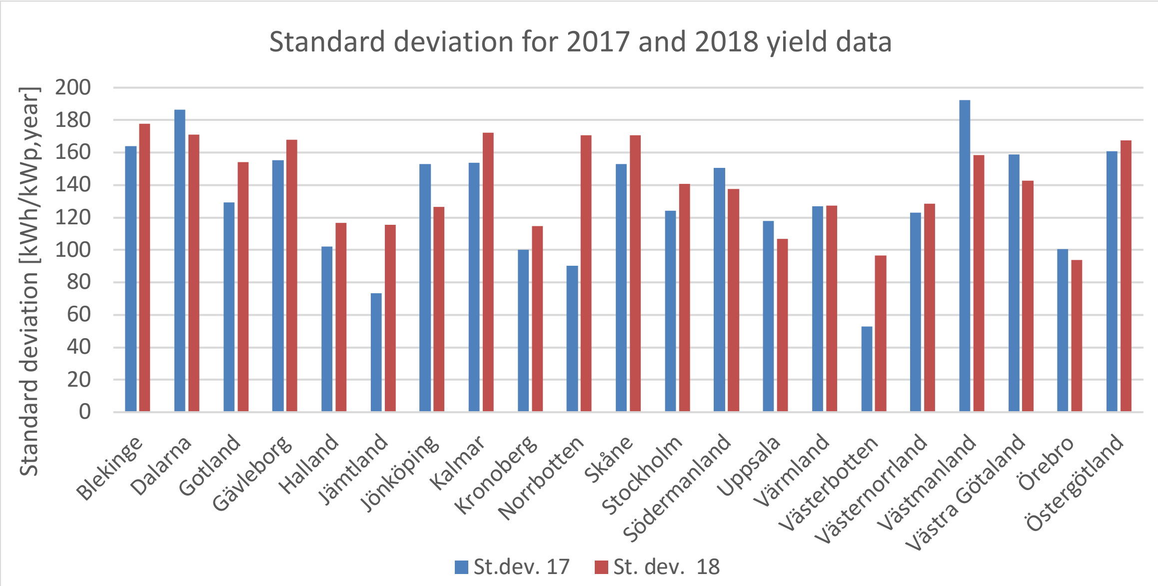


Figure 28 Standard deviation for each county, 2017-2018 dataset

The standard deviations are a statistical measurement of the variability around an average value. The variability or dispersion is the difference between the average value of a dataset and the actual value. Large dispersion of actual values from mean values means higher standard deviation. (EduPristine, 2017). In figure 28 the higher the number, the bigger the standard deviation are, meaning that there are bigger variations in specific yield from that data. The standard deviation is not the same as minimum or maximum values, but rather tell how big the variations usually are.

Table 10 Standard deviation for the specific yield Sweden total

Standard deviation (kWh/kWp]	2017 data	2018 (2017 plants)	2018 (allt plants)
	132	141	156

Figure 29 will show the standard deviation compared to the specific yield for the 2018 data. The 2018 data contains **1380** plants, and this means that these statistics will be better than the above statistics based on the 2017 and 2018 data, only containing 828 plants. The green line in the figure below show the standard deviation as a percentage of the specific yield for each county. The 2018 number of systems accounts for 5,4 % of the total number, and the 2017-2018 number 3,2 % of total number of installed PV systems in Sweden.

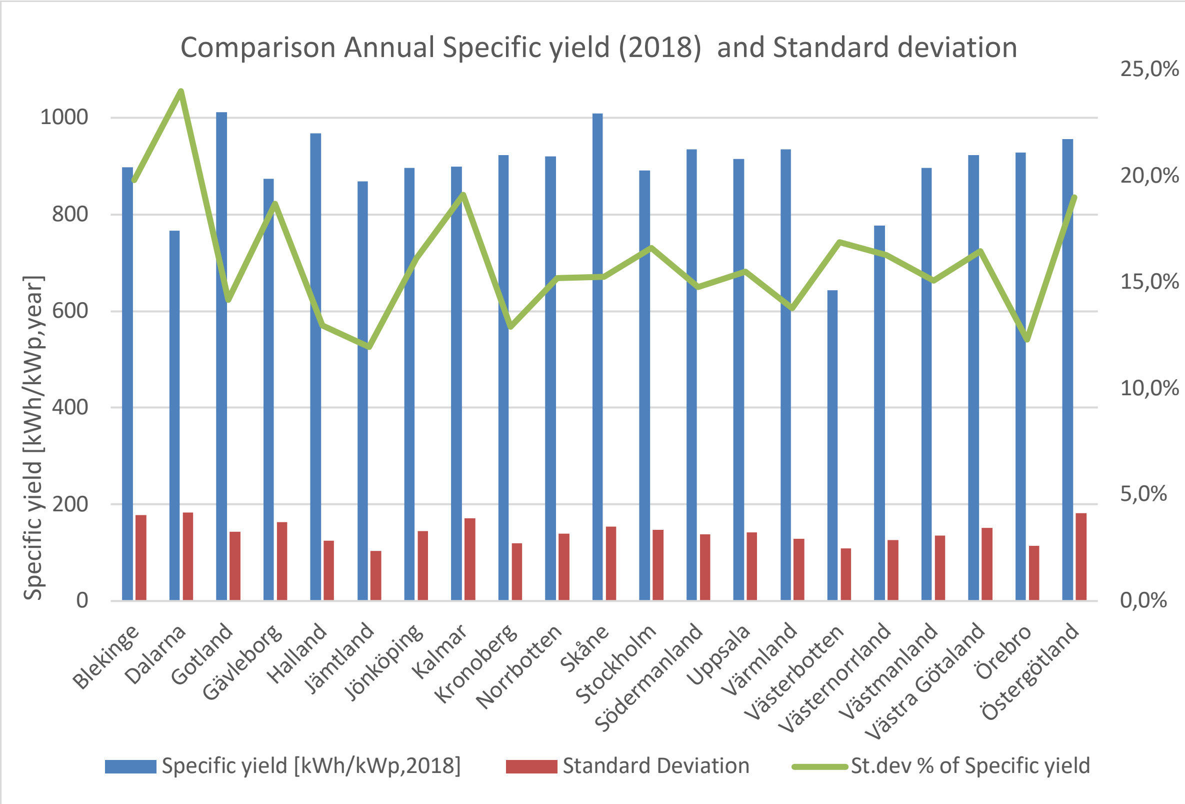


Figure 29 Comparison standard deviation and specific yield 2018 data

6 DISCUSSION

The discussion part will discuss the outcome of the results, how the main result can be interpreted and explanations on why the results are the way they are. Beyond that, the method and tools used to gain these results will be discussed and evaluated. The main reason for this study was to evaluate the PV system yield from plants in Sweden. The evaluation in this case means to understand the potential of PV power yield, and the reasons for variations in that potential. To use solar for power generation are key for transforming the power generation to sustainable alternatives. In order for this to happen, great understanding on the potential of the technique, how to optimize yield and to understand the variations of it will have to be understood in a good way.

6.1 Simulation software

When people invest in solar, there will most certainly be some evaluation on the potential of how much electricity will be produced on that site with a PV installment of a certain size. To simulate or model a system before installing or even thinking about investing in one, are key for evaluating if one should invest or not. That is one of the reasons why there will be calculations of performance when PV system are built. For a calculation to be any good they need in the best extent reflect how the system will perform in reality. If not at least give indications on how big the variations will be in performance, from modeled values to what a PV system will produce in reality. That are the main reasons why these simulations have been done, to evaluate how well they reflect the reality. Another reason why these two different simulation tools have been chosen are because how different they are. PVGIS is a simple simulation tool that everyone with an internet connection can use. It will with minimal knowledge about solar system give a number of how much electricity a certain number of installed watts will produce for a certain location. PVsyst on the other hand require a little more knowledge about PV systems, but in return it is expected to more accurately produce a result based on much more input variables.

We can see some difference between the simulation and calculation on specific yield between the two software's, where the biggest variations occur in Jämtland, Norrbotten, Dalarna and Kalmar. Possible reasons between the result will most likely lie in either shadowing or irradiation values. As seen in the literature study about calculation methods of the two software's, the solar irradiation calculation methods differ. In PVGIS the solar irradiation is based on satellite picture data. Whilst in PVsyst it is based on either ground station measurement data, satellite data or a mix of both. In both Jämtland, Norrbotten and Dalarna, SMHI has ground measurement data, and this will likely be used by PVsyst for calculation if the simulated region is in close proximity to a ground station. Other reasons for variation might be the way the two programs use information about shadows. In PVGIS the far shading from the horizon, in the form of elevation changes in locations, by mountains and valleys etc., will be accounted for automatically. This is not the case for PVsyst, where the user has to define a horizon profile for far shading. Simulations in this study does not consider the effect of changes in elevation. The higher yield for PVGIS simulations in Kalmar is harder to understand and can't really be explained by horizon changes. Possible reasons for this higher yield might lie in the irradiation database used. When simulating in PVGIS the user can chose what irradiation database to use as input, and the results can vary quite a lot for every database when performing simulation, especially as far north as in Sweden. When choosing other databases for simulations in Kalmar, the specific yield went from 1043 kWh/kWp down to around 1000 with other irradiation databases that can be chosen in the PVGIS interface. It is hard to say what database is best for a certain location without extensive testing and comparison with real world measurements.

6.2 Data gathering and evaluation

One of the goals of this study was to retrieve as much data as possible on Swedish PV system production. This meant to gather data from as many plants as possible, this in order to get good statistics. A greater number of PV systems in each county will produce better mean values on specific yield, than say only a few plants. A good example of that would be if you want to know the average height of a class of second graders, you don't just grab two of them and take the average height of those two. And then expect the rest of the class to be that average height. The same thing can be said of the specific yield results from some of the counties in this study. If there are only 3-5 plants and the specific yield has been calculated by a mean value of those few plants, chances are that result in real life would vary a bit more. An interesting insight though is that when looking at figure 23 and 24 in the results, most of the counties that has low number of plants, also have a low standard deviation. This means that the variations in specific yield in a particular system is low compared to the mean value. This is good, but because of the really small amounts of plants used for these statistics, this low standard deviation can also be sheer luck, meaning that only plants with similar specific yield has been collected from the databases. On the counties with higher amounts of plants in the statistics, there will obviously be mean values of specific yield that better reflect the reality, due to the higher number of plants used for the statistics. But some counties with many plants still have a relatively high standard deviation, meaning that there is large spread in specific yield values within different plants. This is to be expected due to the higher number of systems in the data.

A lot of plants have been removed from both the 2018 total comparison, and the 2017-2018 comparison. First in the 2017-18 comparison, all plants not containing full monthly values for either 2017 or 2018 have been removed, in order to compare specific yield between the years. Also, for the 2018 data, all plants not containing full monthly data in 2018 has been removed, except zero yield data in December to February, due the fact that snow and lack of sunshine might actually result in no production. But for some areas in Sweden, zero values in the winter will probably be improbable, so these zero values in the winter might actually represent faulty plants. In case of faulty plants, they should not be removed from the statistics,due to them reflecting the reality. In the case of zero values from faulty data transfers, they should be removed from the statistics, but there is hard to distinguish from a broken PV system or a fault in the data. In this study, a choice was made to remove all zero values in except from December to February. This means that some data from actual faulty plants might have made their way into the statistics, lowering the specific yield for that county.

6.3 The impact of differences in solar irradiation

As seen in figure 19 and 20, the correlation between specific yield and solar irradiation can be seen quite well for most counties. Many counties show close correlation, while some are harder to draw any conclusion on. The biggest differences can be seen in Norrbotten, Västerbotten and Blekinge. Here the correlation between specific yield and irradiation is weak, and some reasons for this can probably be explained. One of the reasons for this can be the number of plants that the specific yield data are based on. All of these counties only have a few plants that contribute to the mean value of specific yield, so it might just be so that the data doesn't represent the mean value if more plants were involved in the same mean value calculation. Both Norrbotten and Västerbotten have solar irradiation data based on ground station measurements, but these are based on either single locations or as in the case of Norrbotten, an average of 2. Västerbotten is a really big county, and only one ground measurement on solar irradiation will contribute to these calculations on solar irradiation for one year. This means that some plants in this county most likely won't be in close proximity to a place with that actual irradiation.

When looking at the specific yield for 2017 and 2018 when adjusted to an average solar irradiation, we see close correlations. Figure 23 show some variations for some counties, mainly Dalarna, Västerbotten and Blekinge, with specific yield differences between 6-11% for the two years when adjusted to average irradiation. The overall for Sweden was a 1.5 % difference, showcasing that there is a strong correlation between specific yield and solar irradiation, and this is to be expected as seen in the literature study.

6.4 Shading impacts

It is difficult to say from only the gathered data what the influences of shadowing will be. But when we factor out the effects of variations in solar irradiation, the other key parameters are tilt, azimuth and shadowing. The tilt and azimuth will of course contribute to non-optimal yield, but a reason for variations when looking at the results when adjusted to average irradiation, could be the losses due to shadowing. If we would have information on the tilt and azimuth and the exact solar irradiation for a specific plant, the impact of the non-optimal placement of the panels could be factored out. This together with adjusting for solar irradiation mean values, means that a better understanding of shadowing impacts could be understood. This is not the case in this study, hence one could only speculate how much the shadowing impact the overall specific yield in the real data from 2017 and 2018.
When studying literature regarding shading impacts, especially the one by Killinger et.al (2018), the impact of shading will be obvious. Especially interesting for Sweden, is that the study of roofs in Uppsala revealed that the impact of shading from the diffuse component of the irradiation was bigger than the beam component, this for sites with similar meteorological conditions as Uppsala. This means that not only is it important to put PV panels out of the most obvious direct shading objects, such as trees blocking the direct sun beam. But also, to consider the objects that could potentially block the diffuse radiation, that will contribute to even higher losses from shading. As said in the literature study, research in the Killinger report found that the overall global irradiance shading losses, in relation to unshaded global irradiance, to be 7,3 percent. This is the mean value of shading losses for the whole building portfolio studied, a large number of roods, that most likely only have a few percent of actual PV panels on them. In reality PV installations will be made on the best side of a roof, in a tilt angle as good as possible. But the research in that report consider all roofs, not only optimal roofs, hence the losses could possibly be slightly exaggerated.
The literature study show that the partial shading can have different impact on the yield, depending on what type of shadows and type of panel it is. Normally in Sweden, the diffuse irradiation accounts for 50 % of the total irradiation. But on sunny days, the beam component will be major, and if something is blocking the direct beam, the drop in yield can be substantial. To cope with this, different types of techniques built in into the panels will work to minimize the effect of shadowing, but this will vary depending on the panel. Seasonal shadowing in the form of snow and dirt are said to have low impact of the total yield over a year.

6.5 Limitations and uncertainties

As discussed earlier, the number of plants will affect the statistics. The more plants in each county will contribute to a number of specific yield that better reflect the reality. Some counties only have a few plants, and this will most certainly affect the mean values of specific yield. Another result that is confusing the results for Norrbotten that shows really high

number in specific yields. Either this is due to the low number of plants, and that the gathered data only represent some really good plants. This county had a sun following system in the statistics, but this was removed due to the really high number in specific yield. This kind of system doesn't represent the norm of the kind of systems that this study investigates, due to the increased performance, but there is a chance that other sun tracking systems have made their way into the gathered data. Other factors that could make the yield look better than it really is, is that installed power might be higher than the statistics really tell. This can be due to a number of reasons. In this study a particular plant stood out of having a specific yield too good to be true, and the person owning this system could be reached by telephone. He told that the system had been expanded with more panels, but that the totals system power had not been updated. This meant that in the databases that this study was conducted on, would show a really good number of produced energy from a relatively low number of installed power. Resulting in a specific yield that looked really good, but in reality, was not. This could be the case for more plants in this study, but it is really hard to tell.
The standard deviation for the results of specific yield between each county, but also Sweden as a whole, will tell us that the results can vary. The standard deviation for the 2017 data was 132, for 2018 with the 2017 plants, 141 and for the 2018 total 156. A lower number in standard deviation in this case, means that more values of specific yield will be closer to the mean value calculated here. As seen in the figure 26 the standard deviation expressed as a percentage of the specific yield, varies between around 12-24% for the different counties. It is hard to draw any conclusions from this, due to the number of plants in each county. A low standard deviation normally tells us that most values lies close to the mean value, but if the standard deviation are based on just a few values, the low standard deviation can be sheer luck.

6.6 Comparison on PV plant data with simulations

The average specific yield values for Sweden when calculating the mean value for every county was 790 kWh/kWp for 2017, and 890 for 2018. The simulations showed values of 974 for PVGIS and 978 for PVsyst. When adjusting the yield with average solar irradiation values, the specific yield for the collected real plant data was around 800 kWh/kWp for both years. This means that the optimized specific yield for the simulation tools are around 20% higher than real values when looking at Sweden as a whole. The results will vary quite a lot within different counties, but overall the yield is substantially lower for the real plant data. Reasons for this could for example be either lack of data for statistics of the collected data, or overestimations for the simulation tools. The method chosen to collect the real PV system data meant that only a small fraction of all the plants in Sweden will be accounted for, with some counties only having a few single plants. If some of the counties with few plants really differs from the actual yield in that county, the overall average will be much lower, and this will affect the average value for Sweden as well. It would be really interesting to see how and if the average yield for both Sweden and all the counties would change if the total number of collected plants would increase drastically. After sorting out faulty data, only about 1200 plants remained from the about 2100 plants that was originally collected from the different databases.
In the simulations, both optimal placement in the form of tilt angles, and a set tilt angle of 31° was investigated. The azimuth was optimal and facing the south in all the simulations. In reality, this will most likely not be the case for panels. Non optimal placement of the panels will result in drops in specific yield. The literature showed that from a study performed in Belgium, that overall losses of around 6% due to orientation of the panels occurred. The literature study also showed that the overall losses for most of the plants was lower than 5%. This shows that non optimal placement of the panels will affect the yield, but this cannot solely explain the differences between the PV plant data and the simulations. This could mean that part of the reasons for differences could be due to shadowing.

7 CONCLUSION

This degree project was done in order to evaluate the PV system yield in Sweden. This was done by investigating the specific yield, that is the produced energy per installed unit of power in the PV plants. In order to do so, real plant data from thousands of Swedish PV plants was collected and analysed to compare yield data from 2017 and 2018. This was done to investigate reasons for differences in yield, and also to compare real plant data to simulated values. The simulations were performed in different simulation software and the results was used for comparison with the real annual yield data, and also to compare and analyse the simulation tools, and the calculation methods they use.

7.1 Conclusions on the research questions

This section on the conclusions will aim to answer the research questions that this degree project set out to find the answers to. A brief explanation on the results will be given, and for more explanations on the results, see the Discussion section.

7.1.1 How much energy does Swedish PV plants produce?

When looking at the specific yield from 2017 and 2018 data, the yield average for all of Sweden adds up to 798 kWh/kWp for 2017, and 890 kWh/kWp for 2018. The variations mean values in yield from the different counties (Län) in Sweden are from 602 kWh/kWp for Dalarna to 955 kWh/kWp for Gotland in 2017. For 2018 the mean values are from 681 kWh/kWp for Västerbotten, to 1015 kWh/kWp for Gotland. When looking at variations in yield when addressing the 2018 data that includes 1380 PV systems, the results look a little different. The lowest values on specific yield for full data of 2018 are 644 kWh/kWp for Västerbotten to maximum values of 1011 kWh/kWp in Gotland.

7.1.2 How much does the actual yield differ from ideal, calculated values on specific yield, and what could be the cause of this?

When adjusting the 2017 and 2018 specific yield to the average solar radiation, the differences between the years was small, as will be described in the section below. The differences between the collected data and the optimized systems in the simulations were substantially larger. When compared to each other the specific yield was around 20% higher for the simulations. In real numbers the adjusted collected data showed a specific yield around 800 kWh/kWp, and around 975 kWh/kWp for both the simulation tools. Possible reasons for the differences probably lie in the amount of data collected from the real PV systems. A larger fraction of the actual number of PV installations in Sweden would better represent the reality, especially in counties with small amount of systems for data collection in this study. Other factors that can cause variation from the optimal values from the simulations, are the placement of the panels. From the collected data, there isn't any information of panel tilts or orientation, meaning that variations between real values and simulations most likely will occur from these factors. Other factors that most likely can affect the yield as much as showed in this study are shadowing, both far shading from horizon and closer shading such as trees and buildings. From the literature review, a study studies revealed that shading losses on the global irradiance to be over 7% for a study done in Uppsala.

7.1.3 How well does the total yield in 2017 and 2018 correlates to the solar radiation differences between these 2 years?

There is a really close correlation between specific yield and solar irradiation. The overall difference in specific yield for the two years when adjusting for an average solar irradiation, was around 1.5 % for Sweden as a whole. Regional variance occurred for a few counties with differences around 6-11%. Most of these differences occur in counties with a low number of PV systems in the data, that could result in higher chance of non-representative data.
In real numbers the 2017 adjusted yield was 795 kWh/kWp, and for 2018, 809 kWh/kWp. When comparing these numbers to specific yield in other countries as investigated in the literature review, these numbers seems reasonable. Leloux et al. (2011) found numbers on specific yield for Belgium at 836 kWh/kWp. Kilinger et al. (2018) found specific yield in Denmark at 786 kWh/kWp, and Leloux et al. (2015) at 900 kWh/kWp for both Belgium and the UK.

8 SUGGESTIONS FOR FURTHER WORK

In order to get even better yield data for all the different regions in Sweden, more plants should to be incorporated in the study. It would be interesting to see how the yield would vary for the counties that already have good amounts of plants for evaluation, but even more so how the specific yield would change for the regions with really few plants. Large numbers of data from more plants so that a bigger portion of the Swedish PV plant would be represented are key for good statistics. Other things that would be really interesting to research are the mean tilt and azimuth angles for PV plants in Sweden, because this would mean that losses from non-optimal placement of the panels could be calculated. This would mean that other factors that would affect the yield, such as shading, could be better understood, if the tilt and azimuth “losses” could be subtracted from the equation. Other than that, more research on the impact of shading could be researched. This by performing experiments and simulations on real plants and compare the real-world shading losses to simulated or calculated values.

REFERENCES

Bengtsson, A., Holm, E., Larsson, D., Karlsson,B. (2017). Skuggningshandbok: Design för solcellssystem för minimerad inverkan av skuggning. *Energiforsk, 2017:385*.

Chalkias,C., Faka, A., & Kalogeropoulos, K. (2013). Assessment of the Direct Sun-Light on Rural Road Network through Solar Radiation Analysis Using GIS. *Open Journal of Applies Sciences, 2013, 3, 224-231*.

EduPristine. (2017, June). What is standard deviation?. Retrieved from https://www.edupristine.com/blog/what-is-standard-deviation

Fraunhofer Institute for Solar Energy Systems ISE. (2019). Photovoltaic Reports. Freiburg: ISE. Retrieved March 14, 2019, from https://www.ise.fraunhofer.de/content/dam/ise/de/documents/publications/studies/Photovoltaics-Report.pdf

International Energy Agency. (2018). Trends 2018 in Photovoltaic Applications: Survey Report of Selected IEA countries between 1992 and 2017. IEA PVPS. Retrieved February 4, 2019, from http://www.iea-pvps.org/fileadmin/dam/public/report/statistics/2018_iea-pvps_report_2018.pdf

IEA-PVPS. (2017). Improving Efficiency of PV Systems Using Statistical Performance Monitoring: *Report IEA-PVPS* (T13-07:2017)

Kamp, S. (2013). Sveriges potential för elproduktion från takmonterade solceller. Teoretisk, teknisk och ekonomisk analys. Examensarbete, Uppsala Universitet, UPTEC

Killinger, S., Lingfors, D,. Saint-Drenan, Y.M., Moraitis, P,. van Sark, W,. Taylor, J,. Et.al. (2018). On the search for representative characteristics of PV systems: Data collection and analysis of PV system azimuth, tilt, capacity, yield and shading. *Solar Energy, 173 (2018), 1087-1106*.

Knier, G. *(2019*). How do Photovoltaics Work?. Retrieved June 9, 2019, from https://science.nasa.gov/science-news/science-at-nasa/2002/solarcells

Leloux, J., Narvarte,L., & Trebosc,D. (2011). Review of the performance of residential PV systems in Belgium. *Renewable and Sustainable Energy Reviews, 16 (2012), 178-184*.

Leloux, J., Taylor, J., Moretón, R., Narvarte,L., Trebosc, D., Desportes, A., (2015). Monitoring 30 000 PV systems in Europe: Performance, Faults and State of the Art. *31st European Photovoltaic Solar Energy Conference and Exhibition (September 2015). 1-9*.

Perez, M., & Perez, R. (2015, November). UPDATE 2015 -- A FUNDAMENTAL LOOK AT SUPPLY SIDE ENERGY RESERVES FOR THE PLANET. *IEA-SHCP-Newsletter*, *62*. Retrieved February 5, 2019, from http://asrc.albany.edu/people/faculty/perez/2015/IEA.pdf

PVGIS. (2017). Overview of PVGIS data sources and calculation methods. European Commission. Retrieved March 20, 2019, from http://re.jrc.ec.europa.eu/pvg_static/methods.html

PVsyst. (2018). PVsyst 6 Help: Help system. Retrieved April, 2019, from http://files.pvsyst.com/help/general_descr.htm

Rakesh, K.T., Kale, P. (2016). Development of Rasterized Mapusing PVGIS for Assessment of Solar PV Energy Potentialof Odisha. *International Journal of Renewable Energy Research R.K. Tarai and P.Kale, Vol.6, No.1 (2016).*

Sino Voltaics. (2011). Standard Test Conditions (STC): definition and problems. Retrieved June 10, 2019, from https://sinovoltaics.com/learning-center/quality/standard-test-conditions-stc-definition-and-problems/

SMHI 2017a, Picture ”Normal Globalstrålning under ett år”. Retrieved (2019), from https://www.smhi.se/klimatdata/meteorologi/stralning/normal-globalstralning-under-ett-ar-1.2927

SMHI. (2019). Meteorologiska Observationer – Globalstrålning timvärde. Retrieved April 24, from https://www.smhi.se/klimatdata/meteorologi/ladda-ner-meteorologiska-observationer/#param=globalIrradians,stations=all,stationid=98735

SMHI. (2017b). Solstrålning. Retrieved March 21, 2019, from https://www.smhi.se/kunskapsbanken/meteorologi/solstralning-1.4186

SMHI. (2015). Solstrålning Sverige. Retrieved March 17, from https://www.smhi.se/kunskapsbanken/meteorologi/solstralning-i-sverige-1.89984,

Solar Edge. (2019).Monitoring Public. Retrieved Mayl 24, from https://monitoringpublic.solaredge.com/solaredge-web/p/home/public?locale=sv_SE

Solkollen. (2019). Solceller olika sorter for olika situationer. Retrieved June 10, 2019, from https://www.solkollen.nu/sv_SE/solceller-2/

STRÅNG. (2019). A mesoscale model for solar radiation. Retrieved May 26, from http://strang.smhi.se

Swedish Energy Agency. (2017). National Survey report of PV power applications in Sweden 2017: Photovoltaic Power System Programme. IEA PVPS. Retrieved February 4, 2019, from https://energiforskmedia.blob.core.windows.net/media/25135/national_survey_report_of_pv_power_application_in_sweden2017.pdf

Swedish Energy Agency. (2019). Solcellanläggningar,antal; Installerad Effekt (MW) [Nätanslutna Solcellsanläggningar]. Retrieved May 6, 2019, from http://pxexternal.energimyndigheten.se/pxweb/sv/Nätanslutna%20solcellsanläggningar/-/EN0123_2.px/?rxid=5e71cfb4-134c-4f1d-8fc5-15e530dd975c

Swedish Energy Agency. (2019). Förslag till strategi för ökad användning av solel. Eskilstuna: Swedish Energy Agency. Retrieved February 5, 2019, from http://www.energimyndigheten.se/globalassets/fornybart/solenergi/solen-i-samhallet/forslag-till-strategi-for-okad-anvandning-av-solel_webb.pdf

Van Sark, W., Bosselaar, L., Gerrissen, P., Esmeijer, K,. Moraitis, P., van den Donker, M., Et.al. (2014). Updata of the Dutch PV specific yield for determination of PV contribution to renewable energy production: 25% more energy*!. 29th European Photovoltaic Solar Energy Conference and exhibition (2014). 4095-4097.*

What Factors Determine Solar Panel Efficiency?. (2013, June). Retrieved from https://energyinformative.org/solar-panel-efficiency/

Appendix 1 Data sorting strategy in Excel

1. Original data inserted into a single Excel file
2. Use Excel sorting function to sort the monthly data after year, then add 2017 and 2018 data into a new column.
3. Sort 17/18 data after index (plants), then put 2017 and 2018 data info two separate sheets. Still having each month for each index row-wise. In order to get each index on a single row with monthly production data on columns, the function of Pivot tables was used.
4. Mark the table that is being transformed into a Pivot table, then Excel gives the choice where to put each parameter, so the parameters Month is drawn to columns, and Index to rows. This produces a new table with each row containing Index (plants) and each column a month (1-12)
5. Make the same procedure with pivot tables for both years, then copy the desired data into a new sheet for each year.
6. Mark the whole table for each year, click on Format – Conditional Formatting- Rules for marking a cell- More rules, then assign a colour to all numbered cells. This is done due to distinguish all the cells that does not have any data, that is all cells not containing production data for a whole year.
7. All cells containing any numbers now has a colour, then the sorting function is used, Custom sorting, sorting after cell colour, for each column (Jan-Dec)
8. This makes all the rows with full data in every month being sorted at the top of the table, making it easy to distinguish full rows with data for every month from those without. Then the rows with full data is copied into a new sheet.
9. The problem now is that the first table containing all the plant data does not match the number of now sorted production data, due to a lot of plants have being sorted out, so a strategy to match the table of plants data to production data was made.
10. To sort out plants not used in the statistics for production, another Conditional Formatting was made. The index (plant) column from the 2017 and 2018 sorted data was copied and inserted next to the full index (plant) data table, then marked both of these columns. Excel Conditional Formatting – Rules for marking cell – Double values. This is used to mark every value in the Plant data (Index) table with a certain colour if that value can be found in the Index column from the sorted production data.
11. The table with coloured index plant data is now sorted by colour, meaning that the two tables will match in terms of index numbers, with plant information and plant production will be on the same row, as seen in Table 3 above.
12. The two tables containing the full data for 2017 and 2018 was then sorted by Zip Code. The plan was to sort the data into every County (Län) in Sweden, but there is no easy way to catalogue Zip Codes into counties without having a database that stores Zip codes and county, and such databases could only be purchased.

Appendix 2 Solar Edge data collecting

1. If the plant in the database had full data for 2017 and 2018 the file was downloaded as a .csv file and open in an appropriate software. In the case of the Mac computer used in this case, the file was opened in the Numbers application, and the monthly data was copied for 2017 and 2018 and inserted into a Excel file that will contain all the data.
2. The data was inserted as a single column, but the plan was to sort every plant in a single row with every month in a separate column. In order to do this, the data had to be transposed, and in order to do this in Excel, the column containing the data needs to be copied again, and pasted using the “paste special” and inserted using the transposed function to insert the column into a single row for each monthly data.
3. The data was inserted from each of the Solar Edge plant files and the zip code and peak power was noted manually on a piece of paper and written into the Excel file for each plant together with the monthly data for 17/18.

MÄLARDALENS HÖGSKOLA
ESKILSTUNA VÄSTERÅS

Box 883, 721 23 Västerås **Tfn:** 021-10 13 00
Box 325, 631 05 Eskilstuna **Tfn:** 016-15 36 00
E-post: info@mdh.se **Webb:** www.mdh.se