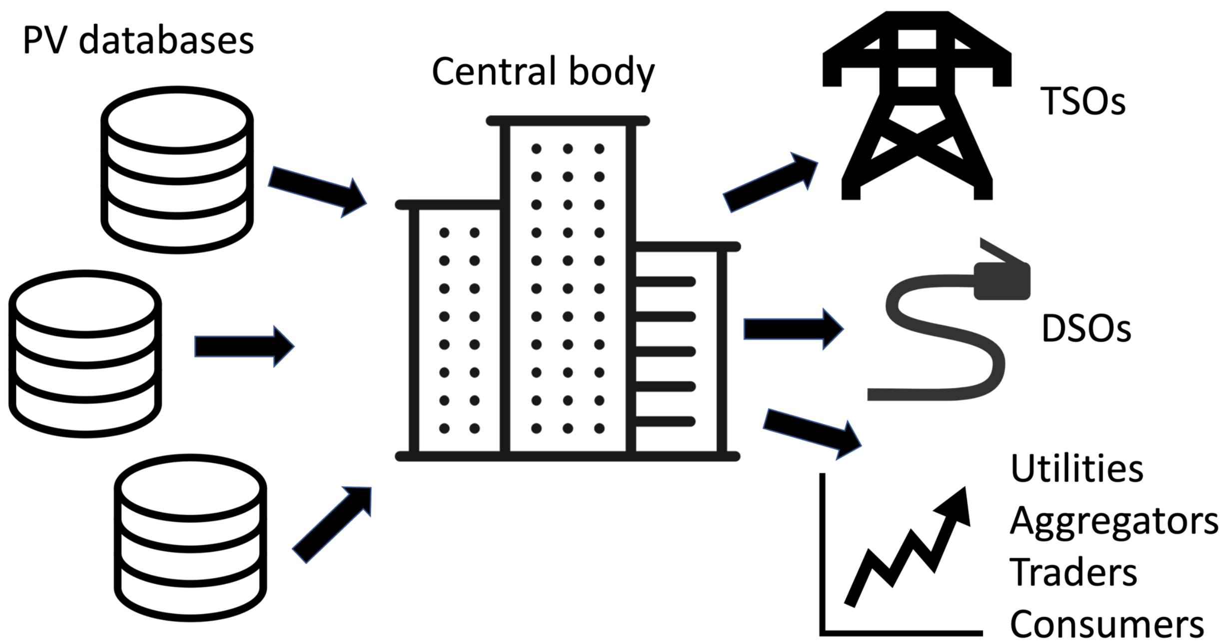
Photovoltaics Observatory

PVOBS

Final report



Public report

Project details

Projectnumber:	TEUE 1821102
Project title:	Photovoltaics Observatory
Coordinator:	Universiteit Utrecht
Partners:	SunData, CBS
Project period:	1 November 2019 – 31 October 2023

Contact details

For questions regarding the project, result, and follow-up projects you can contact:

Universiteit Utrecht
Wilfried van Sark
w.g.j.h.m.vansark@uu.nl

Authors

Wilfried van Sark (UU), Lex Schiebaan (SunData), Ben Laevens, Sabine Krieg, Hui Shan, Bob van den Berg (CBS), Anton Driesse (PV Performance Labs)

Subsidy

The project was supported with a subsidy from the Dutch Ministry of Economic Affairs, National EZ subsidies, Topsector Energie, performed by the Rijksdienst voor Ondernemend Nederland (RVO).

Acknowledgements

We would like to explicitly like to thank Rolf Heijnen of Dutch New Energy and Anton Boonstra for supplying data.

Cover

Summary

With the continuous increase in the number of PV systems in the Netherlands it is highly relevant for maintaining electricity grid balance and security of supply to know ***where*** these systems are located, ***what*** their capacity is and ***how much*** energy they generate. Such an overview is missing at present.

The objective of the PV Observatory project was to investigate how to best set-up a system in which performance data of all Dutch PV systems is collected. Such a national data base, termed the PV Observatory, would include spatially resolved system meta data (capacity, orientation, tilt) and power/energy data.

A set of requirements for both static data describing PV systems such as location and capacity as well as dynamic data such as PV power and time resolution has been defined. It has been investigated how existing data bases could be coupled, especially the CERES database managed by CBS, which contains static data, and others, such as from SunData and PVOutput.org. It was found that merging is not possible on a per-system basis. Further work is needed to find ways of sharing aggregated data.

Analysis of performance of PV systems has shown that energy yield in kWh/kWp varies over the years, due to solar irradiance differences. System degradation rates seems small, but additional data should be analyzed to corroborate that. The results can be used as input to the discussion of updating the Dutch specific yield value of 875 kWh/kWp, which is in use by CBS since 2014.

Finally, a new method has been developed and validated which can replace the use of aggregated global horizontal irradiance by one that uses aggregated global tilted irradiance.

Preface

This final report describes the work performed in the project PVOBS (Photovoltaics Observatory) as carried out within the framework of the Nationale regelingen EZ-subsidies, Topsector Energie, executed by the Rijksdienst voor Ondernemend Nederland. The report addresses the results obtained. In addition, several project changes, mostly due to the Covid pandemic, are described.

Contents

1. Introduction

With the continuous increase in the amount of distributed generation in the form of PV systems it is highly relevant for maintaining electricity grid balance and security of supply to know ***where*** these systems are located, ***what*** their capacity is and ***how much*** energy they will generate. This information is essential for grid operators and other stakeholders active in electricity generation and distribution. The level of detail needed, for example time resolution, is differs for the various stakeholders. For example, for statistical purposes reporting capacity and energy yield on a monthly and annual basis and on a country level is required by Eurostat, and CBS is responsible for that. An energy supplier would require energy generation data with 15-minute time resolution in line with present time resolution in the energy market. For the grid operator a high spatial resolution is required for operational reasons.

In the Netherlands, with 4.2 GWp [1] installed at the end of 2018, and 18.8 GWp at the end of 2022 [2] (see Figure 1) the contribution to electricity demand increased from 3% to 13.8% [2]. Based on the plans described in the Roadmap PV systems and Applications a 10-fold increase is foreseen by 2050 [3], which will cover increased electricity demand due to electrified heating and transport. We note that updated statistics from CBS report that 4.6 GWp was installed at the end of 2018 [4] and 19.6 GWp at the end of 2022 [5]. At the end of 2023 22.4 GWp is installed [5].

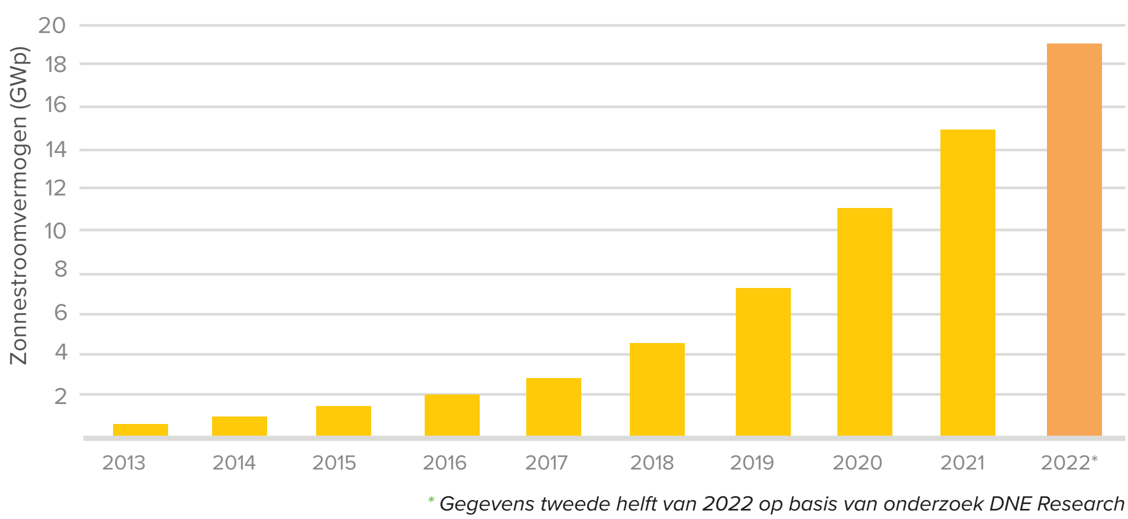


Figure 1. Increase in installed PV capacity in the Netherlands since 2013 (Source: DNE [2]).

1.1. National statistics

Reporting on the the amount of generated electricity by PV is described in the Protocol Monitoring Hernieuwbare Energie [6,7]. Since 2015, the amount of generated electricity is based on multiplication of the average cumulative installed PV capacity determined by CBS and a key figure ("kental") representing the amount of generated energy per installed unit of capacity. This has been determined at 875 kWh/kWp [8], while the uncertainty of annual generated electricity determined by this method is stated to be 20% [9].

For the determination of PV capacity several sources are used. The most important are 1) CERES (Centrale Registratie van Systeemelementen) from the grid operators for systems for small consumers and 2) data from VertiCer (previously CertiQ) for subsidized installations (SDE++, systems > 15 kWp), often large consumers. CBS checks this data through an extensive analysis and removes double counting within and between sources as well as possible.

For the determination of energy yield, for (large) systems listed in the VertiCer database (formally CertiQ), metered PV electricity generation data is used, as this information is the basis for the calculation of the SDE++ subsidy. For the other systems energy yields are calculated based on a bottom-up model. In this model calculation, the power per month per installation is multiplied by the global horizontal solar radiation in the most nearby KNMI weather station in that month [10] divided by the 30-year average annual radiation between 1981 and 2010 (368,378 J/cm^2, or 1023.3 kWh/m^2) and the key figure for production per unit of power in a normal year of 875 kWh/kWp. This key figure was determined by analyzing PV production data in 2012 and 2013, which showed 1-2% higher annual irradiation compared to the 30-year average. We note that the current 30-year average is taken for the years 1990-2020, which is 3% larger than the older average, mostly due to the sunnier years between 2010 and 2020.

Global Horizontal Irradiance (GHI) is defined as the solar energy (*irradiance*) coming from all directions (*global*) onto a *horizontal* surface. This can be split up into two parts: *diffuse*, coming from the sky hemisphere and *direct*, coming in a straight line form the sun. The *diffuse* portion is also measured on a horizontal surface by blocking the *direct* sun rays and is therefore referred to as DHI. The intensity of the *direct* portion is measured by pointing a sensor at the sun, so it represents the *irradiance* intensity on a plane that is geometrically *normal* to the sun rays, hence the name direct normal irradiance (DNI).

1.2. Specific energy yield

The Protocol Monitoring Hernieuwbare Energie [6,7] focuses on the annual statistics, which CBS is obliged to provide to Eurostat. The method is also used to possible to present monthly production, however the model needs refinement to include plane-of-array irradiance instead of global horizontal irradiance. Section 3.2.4. will describe a new model to calculate plane-of-array irradiance from global horizontal irradiance, direct normal irradiance, and global diffuse irradiance. In addition, the effect of temperature on the efficiency of the PV modules needs to be included, although this refinement is less important. The efficiency of a PV module is determined at a temperature of 25 °C, but it is temperature-dependent and decreases with about 0.4%/°C. Especially in summer months, depending on the installation manner (on roof-top or open field), PV modules may reach 80 °C, which thus leads to a reduction of module efficiency of 20-25%, compared to lab conditions, depending on the type of module. Hence a PV module rated at 400 Wp would then generate 300-320 Wp. Consequently, daily energy yields will be 10-15% lower than nominal. On an annual basis though, this temperature effect is around 5% for the Netherlands [11].

Moreover, the figure of 875 kWh/kWp is relatively old, as it was determined 10 years ago [8]. Since then, PV systems are increasingly installed in a non-optimum way, with orientations varying from East to West, and tilt angles varying from 10° to 90°. While it is actually useful to avoid PV generation peaks at noon, it lowers the annual yield, see Figure 2. CBS is therefore investigating how this relates to recent measurement data from VertiCer's systems and, if possible, other sources with measured data [12]. The difficulty is that malfunctions in PV systems are not always adequately detected nor repaired, and lower than expected energy yields will be found than used in statistical reporting. Also, orientation and tilt are often not known. In addition, reporting of capacity of the PV systems based on the installed components is not always correct, as either DC (total capacity of PV modules) or AC capacity is reported. Often the AC capacity of the inverters is lower than DC PV capacity, even down to 50% of PV capacity for operational reasons or due to grid congesting mitigation requirements for larger PV parks. All these types of uncertainties that are difficult to resolve limit the accuracy in reported capacity and energy yields that can be achieved.

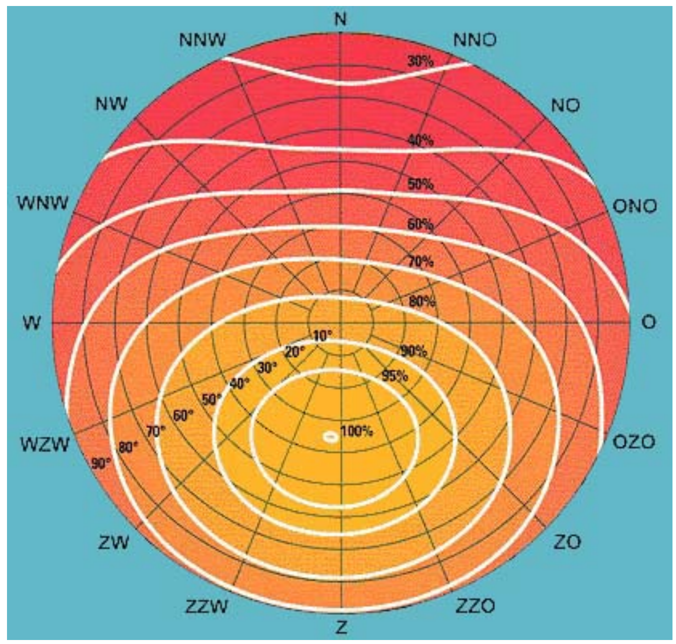


Fig. 2. Annual yield in % of maximum for all orientations and tilt values. Maximum yield for the Netherlands is obtained for a South facing PV system at 37° tilt (source: AgentschapNL [13]).

At the end of 2022, CBS reports a total amount of 2,294,154 PV installations with a total capacity of 19.6 GWp [14], or about 8.54 kWp per installation. We note that about half of the systems is in residential areas with capacity up to 3 kWp (or total capacity of about 3.4 GWp), for which actual monitored energy data is not available. CBS collects spatially resolved data, which allows to visualize total capacity per municipality, or, for example, the change from 2022 to 2021, as shown in Figure 3.

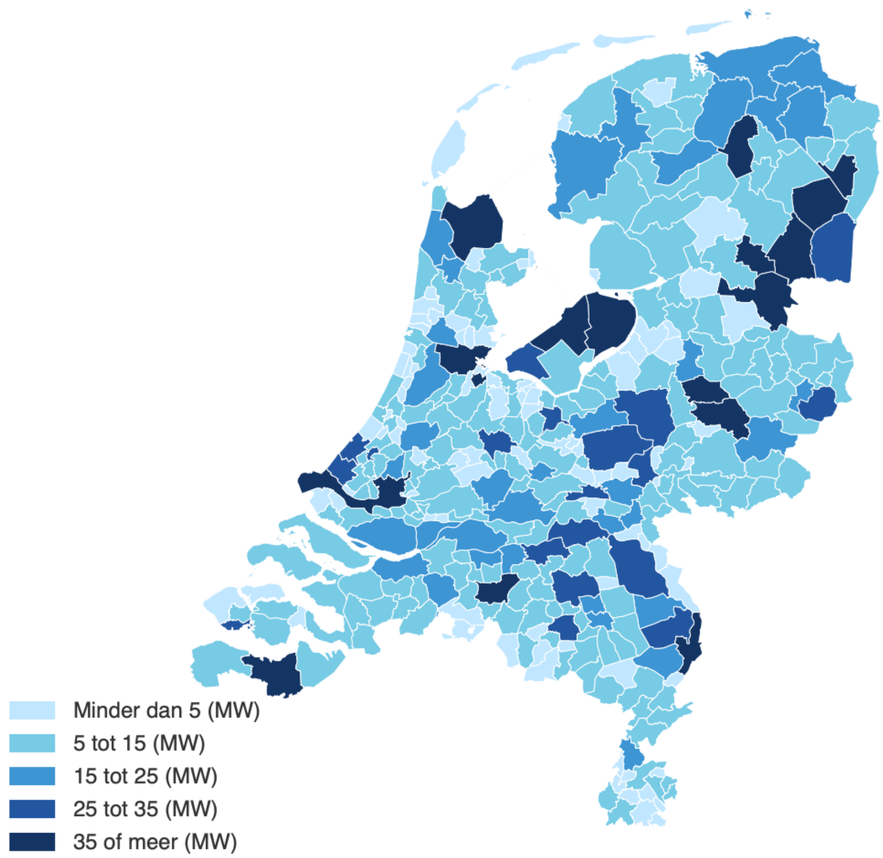


Figure 3. Additional PV capacity per municipality, illustrating PV increase between 2021 and 2022 (source: CBS [16]).

Annual generated energy in 2022 is reported to be 17.079 TWh [15]. From these data, a specific yield of 871 kWh/kWp can be inferred. This seems in line with the 875 kWh/kWp figure. However, the 2022 horizontal annual irradiation sum was 1211.7 kWh/m^2 compared to the 1051.7 kWh/m^2 30-year average (1991-2020), or 15% larger than normal [10]. Accounting for that, the actual specific yield would be 15% lower, or about 744 kWh/kWp. Compared to the 1981-2010 average irradiance which is 18% higher, the specific yield consequently is 18% lower, or 718 kWh/kWp. For the determination of energy yield of the small residential systems, the 875 kWh/kWp number is used, corrected for the irradiance sum, and for a total capacity of about 3.4 GWp. For the other systems, actual measured data is used from VertiCer, which is not corrected for the annual irradiance sum.

An important assumption in these corrections is that PV energy yield is linearly related to annual horizontal irradiance sum. The latter is measured on horizontal surfaces and is known more formally as global horizontal irradiance (GHI). PV modules are rarely mounted horizontally but tilted and oriented to a certain direction so that they capture more of the available solar energy. Energy yields are linearly approximated much better by using so-called plane-of-array (POA) irradiance, or the actual incident irradiation on the surface of the PV module. Only at high irradiance values, this linearity is broken due to temperature effects.

1.3. Data for operational support

In addition to the requirement for CBS to collect and disseminate statistical data on PV, information on PV capacity, actual or potential power flows is needed by distribution system operators (DSOs), such as Stedin, Enexis, Alliander, as well as the transmission system operator (TSO, TenneT). This information is required to properly manage the balance of supply and demand at the electricity grid at every level of the grid. Not knowing what and where systems are connected may be problematic certainly when PV capacity is increasing as fast as in recent years. Besides that, (local) congestion may limit the connection of new capacity to the grid, and at certain times with high feed-in of PV power, grid voltage may increase beyond permissible limits leading to shut-down of inverters. This passive curtailment reduces the amount of energy fed back into the grid. In addition, energy utilities such as Eneco, Essent, Vattenfall, need to know expected power generation in order to properly manage their assets as well as to optimize energy trading. For PV plant owners, from small (kWp) to large (multi MWp) size, energy yield directly couples to revenues, which is of special importance for SDE++ subsidy schemes. For systems larger than 1 MWp, monitoring solutions are usually deployed already to guarantee proper functioning, for smaller systems this is rare, and only a few SMEs are stepping into this market (such as SunData and SolarCare).

1.4. Proposed solution and challenges

This project proposed to investigate how to best set-up a system in which performance data of all PV systems is collected and that will be publicly available while privacy concerns issues (General Data Protection Regulation, GDPR) are respected. The project aimed to set-up requirements for a national data base, the PV Observatory, that includes spatially resolved system meta data (capacity, orientation, tilt, etc.) and power/energy data at a certain time resolution and ideally in real-time.

While much (meta-) data is available, it is scattered and/or not publicly available (such as VertiCer data). The main challenge of this project therefore is to bring together all data on a common platform. This requires automated data collection, for example directly from Internet-connected inverters, after properly addressing privacy issues, as well as using other data sources from installers and/or metering companies. All that data must be harmonized to fit in a uniform format for the data base. Irradiation data from KNMI and/or satellite sources are then added to be able to calculate specific yield values and performance ratios.

The data in the PV Observatory is intended to be used differently by different stakeholders based on their business model (Figure 4) [17]. For example, network operators can better plan grid extensions, while PV system owners and monitoring service companies can use performance data as a benchmark of their systems. This will undoubtedly lead to increased energy yields, as malfunctioning systems will be identified much sooner, as well as lower O&M cost. Visualizations on a country, provincial or municipal level can be informative for the general public, and the aggregated data will serve various purposes at CBS. Also, the Netherlands Enterprise Agency (RVO) will be able to provide annual statistics more easily and more accurately to the PV power systems program of the International Energy Agency (IEA-PVPS).

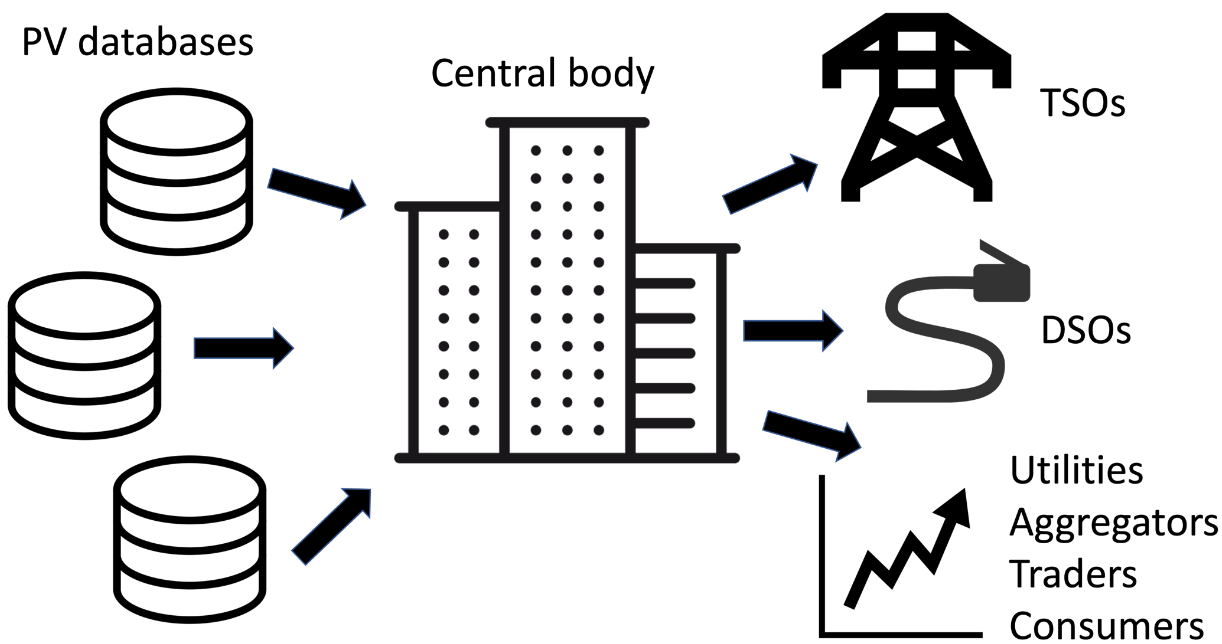


Figure 4. Proposed framework of PV data platform.

As over 2 million PV systems are installed in the Netherlands (in 2022) inclusion of all of them in the database is very and perhaps overly ambitious. Therefore, a step-by-step approach was used: the first step focused on selecting at least 5000 representative PV systems across the country from which an overall impression of PV performance can be obtained. The second step focused on increasing this number to 25000, and will also define automated ways to increase this further to reach near-100% coverage. The size of the data base is linked to the number of systems but also to the time resolution, which may vary from monthly data down to 5-min interval data.

This project was ambitious, with an important identified risk being that the amount of data collected will be too small or not representative for proper analysis or further development, due to, e.g., privacy issues. While clear progress had been made towards the realization of a data base, in the end an **open** data base could not be realized.

2. Results

2.1. Database set-up

The CERES database contains location (postal code) and system capacity. CERES is not publicly available, while it should be noted that CBS is disseminating data per municipality (see, e.g., Figure 3), and data is also disclosed on the klimaatmonitor website [18**]**. At present, listed capacity is both DC (direct current) and AC (alternating current) capacity, i.e. total PV module capacity and inverter capacity. In earlier years, this distinction was not made.

The PV Observatory database aims to include spatially resolved system meta data (capacity, orientation, tilt, etc.) and power/energy data at a certain time resolution and ideally in real-time. This would entail extending the present CERES database by adding, per system, information about orientation and tilt, thus identifying the CERES database as a database with ***metadata*** only. Operational data, i.e., time-series of power/energy generation would require building a new operational database, for example CERES-Ops (CERES Operations). Depending on the required time resolution, and the number of systems, it would require substantial data acquisition and storage efforts.

Presently, Dutch grid operators are collaborating via a digital service provider for the energy market, i.e., Energy Data Services Netherlands (EDSN) [19], which is a collaborative platform. The operational database could be part of that platform.

The required data are identified to be:

Metadata

1. Total system DC capacity, *PDC*, in kWp
2. Location of system, *LocPV*, postal code-6 and streetnumer, e.g., 1234AB89
3. Orientation of modules, *Om*, in degrees (0-360^o, with South is 180^o)
4. Tilt of modules, *Tm*, in degrees (0-90^o)
5. Installation date, *DateIns*, as dd-mm-yyyy
6. Type of integration, *Itype*, e.g., in-roof or on roof or free standing
7. Number of PV modules, *Nm*
8. PV module capacity, *Pm*, in Wp
9. PV module brand, *PVbrand*, e.g., Sunpower
10. PV module type, *PVtype*, e.g., Maxeon 3, SPR-MAX3-410
11. Total inverter system AC capacity, *PAC*, in kWp

12. Number of inverters, *Ninv*
13. Inverter capacity, *Pinv*, in Wp
14. Inverter brand, *INVbrand*, e.g., SMA
15. Inverter type, *INVtype*, e.g., Sunny Tripower 5.0 Smart Energy
16. Inverter class, *INVclass*, e.g., string inverter, optimizer, microinverter
17. Comments, e.g., potential shade from trees, dormers etc.

Metadata #1 and #2 are now present in the CERES database. Metadata #5 can reasonably be assumed to be equal to the date of entry in the CERES system. The other metadata items are not available for small systems. For larger systems metadata is known to the system owner and/or installer, but not openly available. An effort is necessary to collect all this data. Note, that systems may also consist of PV modules at multiple orientation and tilt values, and/or various module and inverter types, etc.

<u>Operational data</u>

1. Date, *Date*, yyyymmdd, e.g., 20230713
2. Time, *Time,* hhmmss, e.g., 130415
3. Power (instantaneous) *Pi*, in W, only integers
4. Energy (cumulative per day), *Ed*, in kWh, one digit after decimal point
5. Module temperature, *Tmp*, in ^{o}C, e.g., 12.54 ^{o}C
6. Irradiance, horizontal, Gh, in W/m^2, one digit after decimal point
7. Irradiance, plane-of-array (tilt), Gt, in W/m^2, one digit after decimal point
8. Wind direction, *Wo*, in degrees (0-360^{o}, with South is 180^{o})
9. Wind speed, *Ws*, in m/s, one digits after decimal point

For small, residential systems, operational data is not available publicly. Power and energy data though can be inspected in apps connected to data clouds of inverter manufacturers. Larger systems, especially those that are entitle to receive SDE++ subsidy have a certified energy meter, and usually, if the system is large enough, some kind of monitoring system is in place which provides warnings when malfunc tions are detected. CBS through VertiCer has access to annual energy yield for all SDE++ systems. Meteorological data (operational data #6, #8, #9) can be derived from KNMI data series, or local sensors. Plane-of-array irradiance can be derived using metadata for orientation and tilt, but ideally are measured using local sensors (pyranometers or reference cells).

The required metadata and operational data listed above will be difficult to collect. Therefore, a mininum list of data is defined which would provide already good insights in PV performance across the country.

<u>List of minimally required data:</u>

<u>Metadata</u>

1. Total system DC capacity, *PDC*, in kWp
2. Location of system, *LocPV*, postal code-6 and streetnumer, e.g., 1234AB89
3. Installation date, *DateIns*, as dd-mm-yyyy
4. Total inverter system AC capacity, *PAC*, in kWp

<u>Operational data</u>

1. Date, *Date*, yyyymmdd, e.g., 20230713
2. Time, *Time,* hhmmss, e.g., 130415
3. Power (instantaneous) *Pi*, in W, only integers
4. Energy (cumulative per day), *Ed*, in kWh, one digit after decimal point

2.2. Combining existing databases

Sundata has metadata and performance data of numerous systems in the Netherlands, but a relative low coverage in Limburg, Friesland, Groningen and the north of North-Holland. In an attempt to investigate if the CERES database could be merged with the Sundata database, a small number (100) of systems in Amersfoort is selected. These systems have been realized with the identical PV modules of 280 Wp each, but with different system capacity. In this way, metadata from both databases could be compared. Location information (address) was modified to latitude and longitude information, rounded to two decimals after the decimal point, thus complying with privacy regulations, see also [20]. This led to data points on a 1x1 km^2 grid. Yield data, both actual and estimated, were made available.

From the selected 100 installations in Amersfoort, 98 are in the CERES database. This is a high number, as it has been found that on a national scale about 85% of the systems are in the CERES database [21], as owners have to register systems themselves. For the Sundata database, installers have provided the metadata.

There are 98 installations with an installation date in both the CERES and Sundata database. However, only in 6 cases, the dates are equal. In all other cases, the CERES date is earlier than the Sundata. The mean of the differences is 11.6 days, the largest difference is 58 days. In the comparison, the installation date provided in the Sundata database is used. As Sundata's database contains about 15000 systems, merging would be useful, as the systems cover a large part of the Netherlands.

Tables 1 and 2 show PV system rated power (DC) and the ratio of the inverter power (AC) and PV system rated power (DC). This ratio is sometimes smaller than 1. This is because DC rated power is rarely generated due to PV module temperature much higher than 25 ^{o}C. The three values of DC power are due to the use of 6, 8, or 10 panels of 280 Wp each.

Table 1: Distribution of DC power of PV systems in the set of 100 systems in Amersfoort.

DC power	1680	2240	2800
number of systems	9	21	68

Table 2: Distribution of the AC/DC power ratio in the set of 100 systems in Amersfoort.

AC/DC ratio	0.75	0.89	1	1.04	1.05	1.25
number of systems	1	1	92	2	1	1

In the CERES database, information on AC and DC power capacity is available for installations from 2021 onwards. For installations installed in 2020, about half of them have both AC and DC power capacity data. For older installations, the information is available only for a small percentage of installations. It is noted that often AC and DC power data are exactly the same, which is considered as unrealistic and not reliable. Figure 5 shows the distribution of the AC/DC power ratio from the CERES database for three years. Note, systems with a power ratio of exactly '1' are excluded from the figure. It is clear that the average AC/DC ratio is below one, and mostly between 0.8 and 0.9. A further distinction between systems of size below 10 kWp and larger than 10 kWp did not change this finding.

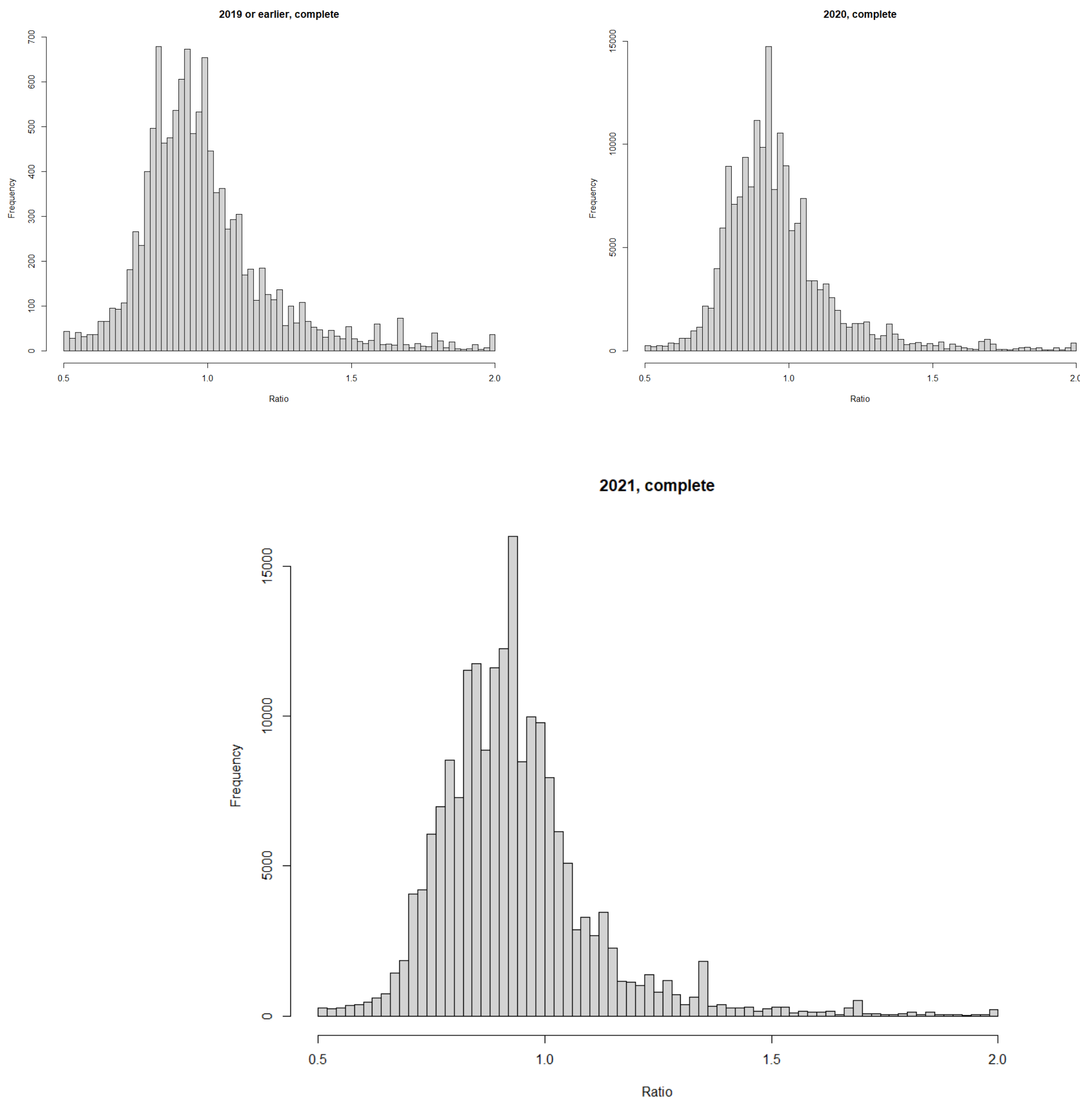


Figure 5. Distribution of the AC/DC power ratio from the CERES database for three years.

Another dataset with performance data is available from PVOUTPUT.org, at which today 87.5 MWp is registered in the Netherlands [22], with average system size of 6.6 kWp, or over 13000 systems scattered across the country. PV system owners voluntarily supply meta data as well as operational data to PVOUTPUT.org. Quite some metadata is supplied, such as orientation, tilt, panel and inverter brands and capacities. The regional distribution is not even over the whole country, see Figure 6. The province of Noord-Brabant accounts for 27% of the total installed capacity.

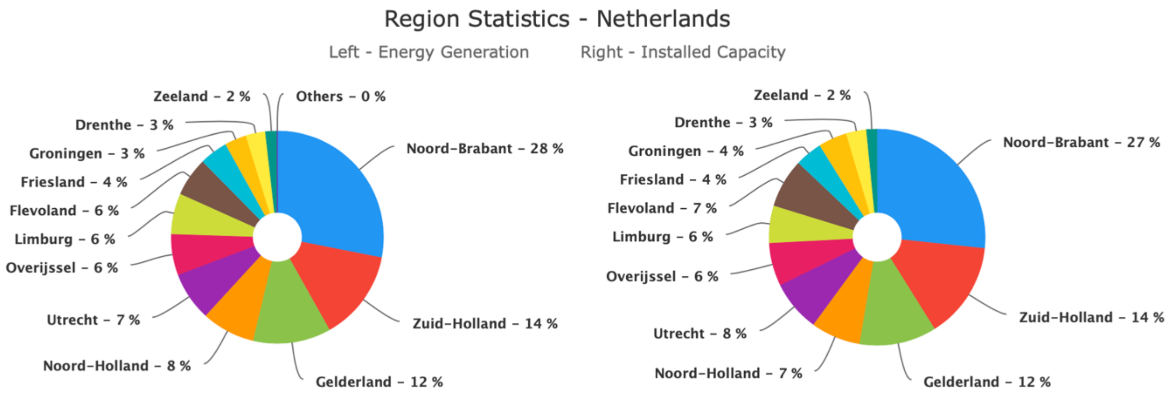


Figure 6. Regional distribution of PV systems (energy and capacity) in the Netherlands (Source: PVOUTPUT.org [22].

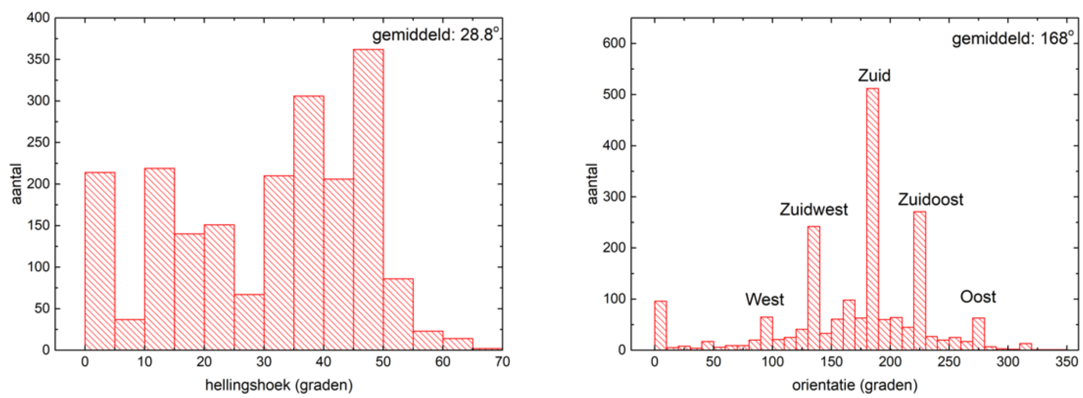


Figure 7. Distribution of PV system tilt and orientation for the Tweakers/PVOUTPUT data (Source: [23,24]).

For over 2000 systems data are shared and analyzed via tweakers.net, which is organized by Anton Boonstra [23]. Total capacity is 12.76 MWp., with average DC capacity of 6.09 kWp. Monthly performance analyses per province are shared on twitter/X [23], including comparison with KNMI irradiance data. From this dataset, tilt and orientation are shown in Figure 7. Average values are capacity of 6.09 kWp, orientation 168° (South), and 28.8° tilt [24]. While many of the systems are oriented South, also systems are oriented South-West and South-East, while systems that are oriented East and West may represent harmonica type systems, in which half of the panels are oriented East and half West at about 10° tilt, which are typically found on flat roofs. Note, for several systems both orientation and tilt are equal to '0', which may be erroneous data.

Figure 8. Regional distribution of PV systems for the Tweakers/PVOUTPUT data (Source: [23,24]).

The systems are not evenly distributed over the country, as shown in Figure 8. This prompts the question on representativity of metadata and performance data, which is necessary to ask when only a subset of all systems is used. This issue is also discussed by Laevens et al. [25] and may easily lead to 10-20% errors in estimation of national PV yields.

2.2. Data analysis

2.2.1. 100 Amersfoort systems

A performance analysis has been done for the selected 100 systems in Amersfoort by CBS. The yield of the systems is predicted using the specific yield of 875 kWh/kWp corrected for irradiance variations. The KNMI station of De Bilt is closed to the city of Amersfoort, i.e., about 15 km South-West of Amersfoort. The calculation is:

$$\text{Production} = \text{power [in kWp]} * 875\ \text{kWh/kWp} * \text{radiation [in J/cm}^2\text{]} / 368.378\ \text{J/cm}^2$$

In which the value 368.378 kJ/cm^2 (equivalent to 1023.27 kWh/m^2) is the mean annual irradiance sum (30-year average of 1981-2010). We note that the value of 875 kWh/kWp was established using data from 2012 and 2013 in which solar irradiance was similar to the 30-year average.

First, a few installations are selected, and actual power generated is compared to the theoretical production according to the above equation. The comparison is started at April 1, 2019, and daily and monthly production is compared.

The ratio of actual and calculated daily production for 1 installation is shown in Figure 9. It is clear that this ratio is around 1, with daily fluctuations. These fluctuations are smaller in summer compared to winter. In winter, the mean could be slightly larger than 1, in summer, slightly smaller. There are only a few days with (almost) no production.

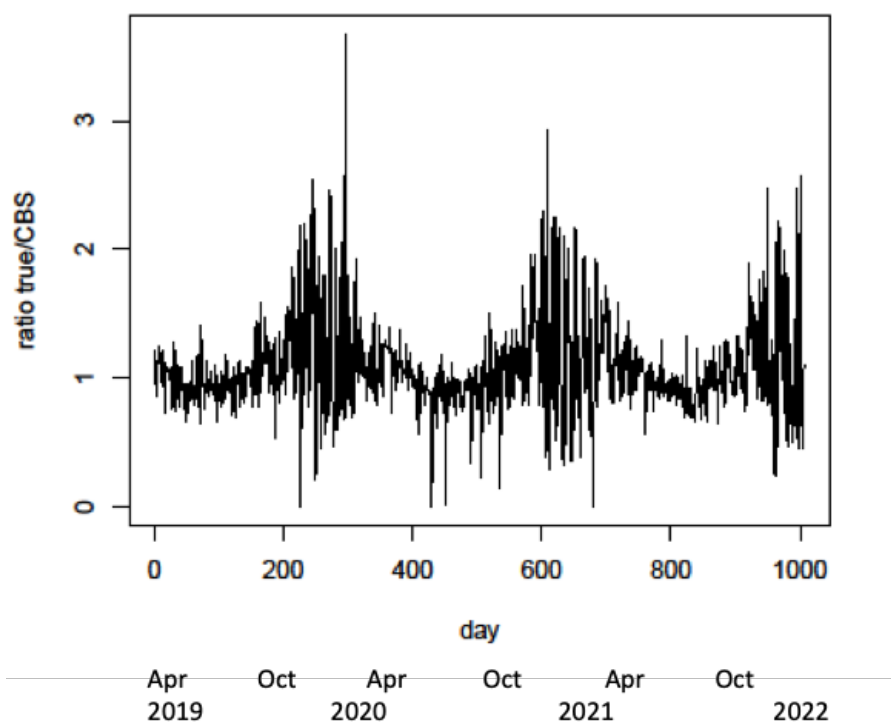


Figure 9. Ratio of actual ("true") and calculated ("CBS") PV production data.

The monthly production for three other installations is shown in Figure 10. In all figures the red line is the actual production and the black line the predicted one. There is a seasonal pattern, with smaller ratios (true production smaller than formula) in the summer. This can be attributed to the use of GHI instead of POA (see section 2.2.4). It is very clear for installations A and B (Fig. 10a,b and 10c.d), and it is not visibly for system C (Fig. 10e,f). For system C a decreasing ratio during the (almost) 3 years is observed, and the ratio is almost always smaller than 1. Presumably because the tilt is close to zero. Installation B has the highest monthly ratio (around 1.6). Tilt and orientation effects will be important in this.

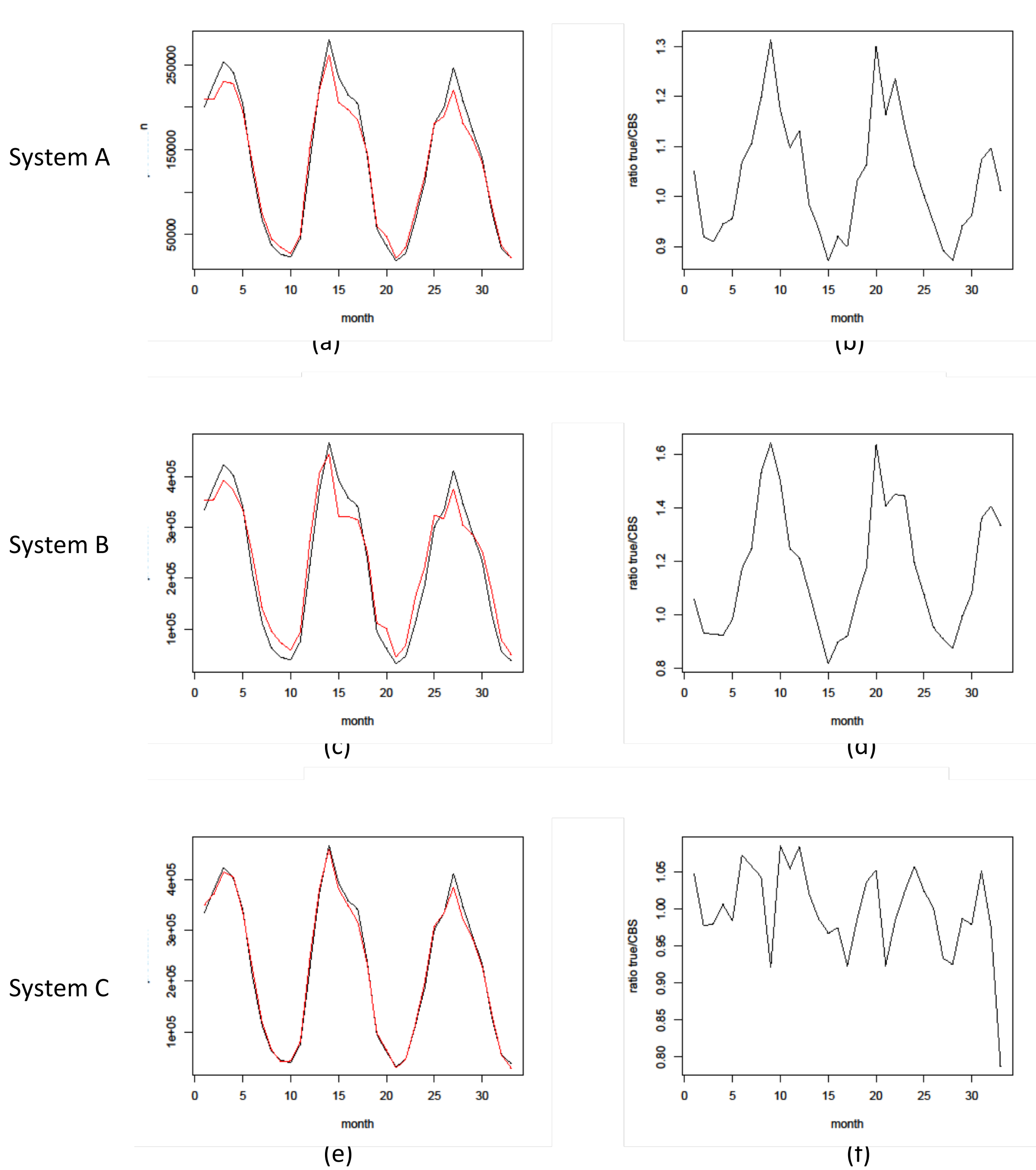


Figure 10. Comparison of monthly power of three PV systems (actual, red and predicted, black), and associated ratio. Clear summer and winter peaks are seen.

Further details of daily production are shown in Figure 11 for winter ("selection 3") and in Figure 12 for summer ("selection 2"). As shown in Figures 9 and 10, also for the daily power pattern, there is a good correlation of actual and predicted power, with some differences though. For example, installation B shows higher actual production on sunny days in winter compared to predicted power, while on cloudy days (lower irradiance) it is the other way around. This is presumably due to a system with high tilt angle. In summer actual production is mostly smaller probably due to daily POA being lower than daily GHI. Temperature effects on PV panel efficiency are secondary in affecting PV energy yield. Similar effects are seen for installation A and D. In contrast, for installation C actual production is always smaller than the predicted production (sunny and cloudy days) in winter.

Figure 13 shows the distribution of annual energy production for 2020 and 2021. The mean of the distribution for both years is 837 kWh/kWp, about 5% lower than the 875 kWh/kWp value. The PV systems with low annual energy production suffer from shading losses, and have mean value of about 725 kWh/kWp. Not taking these systems into account for the calculation of the mean, the annual energy production for unshaded systems is estimated to be about 875 kWh/kWp. The system with annual energy production of >1100 kWh/kWp likely is due to an erroneous DC capacity (increasing capacity by two panels brings the mean to 875 kWh/kWp).

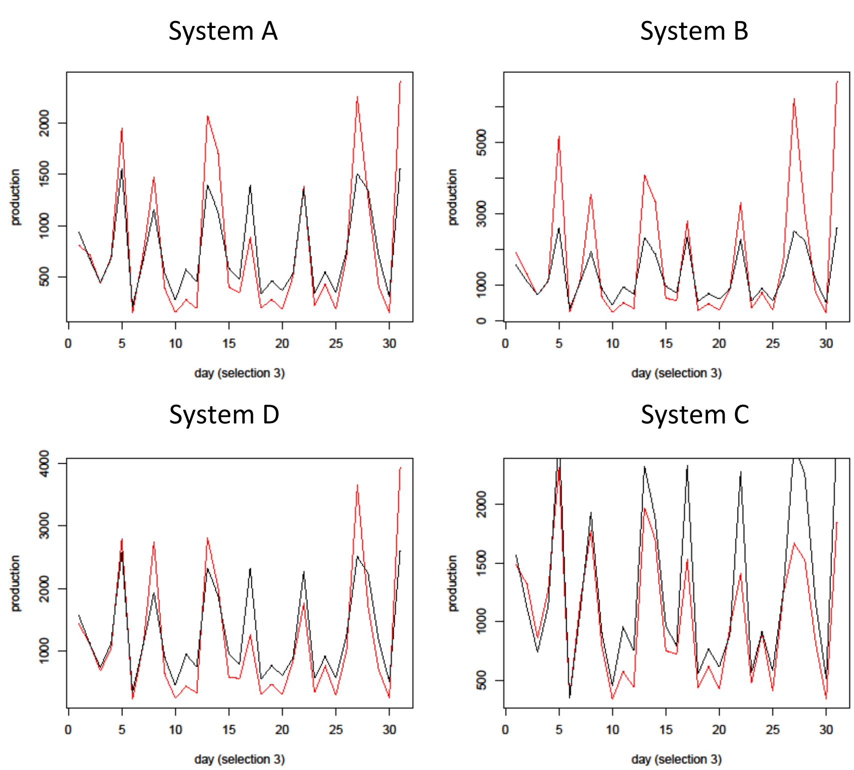


Figure 11. Comparison of daily power of four PV systems (actual and predicted) for days in winter.

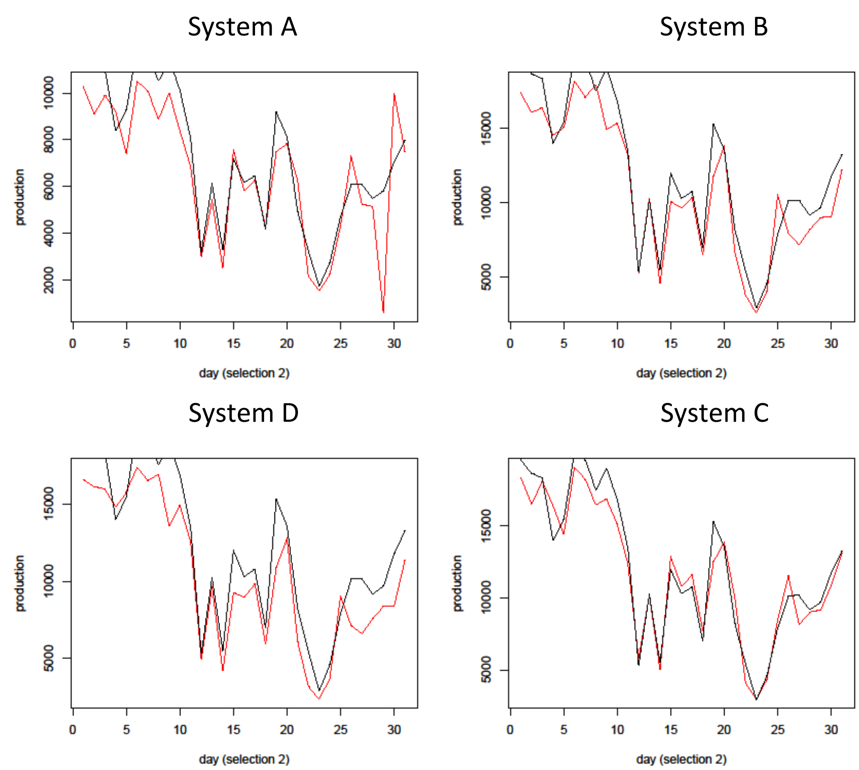


Figure 12. Comparison of daily power of four PV systems (actual and predicted) for days in summer.

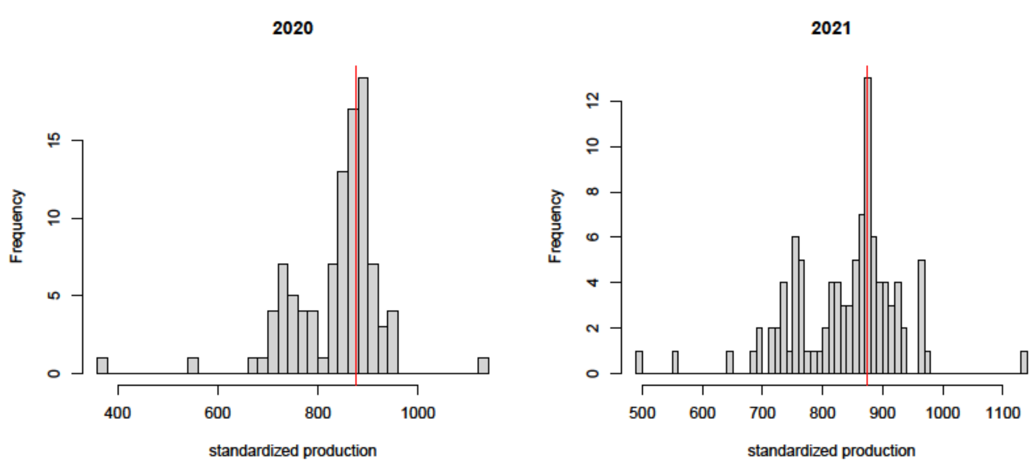


Figure 13. Distribution of actual annual specific energy yield of all 100 PV systems

All in all, from the above analysis it is shown that the annual yield calculation method is reasonably accurate for the 100 selected systems. This means that available metadata on PV systems in combination with weather data could in principle be used to determine actual PV energy yield on annual basis at reasonable accuracy.

2.2.2. Verticer/CertiQ data

Part of solar power data available at CBS comes from the VertiCer data source (previously CertiQ). In addition to power information for an installation, this source also records energy production per installation (for some of the installations, and up to a monthly frequency). This section is a shortened version of the report prepared by CBS "Onderzoek naar productiefactoren zonnestroom in 2022" [12].

For these installations the specific yield in kWh/kWp is calculated by CBS, hence based on the energy production and rated power and an irradiance correction factor (annual radiation/long-term average). Specific yield data have also been reported by others sources (Laevens et al. [25], SolarCare [26]). We compared specific yield data with each other for the years 2016-2021, see Figure 14. The calculated specific yield data are further broken down by capacity of the capacity, i.e., small (residential, P < 15 kWp), medium (15 kWp < P < 100 kWp), and large (P > 100 kWp).

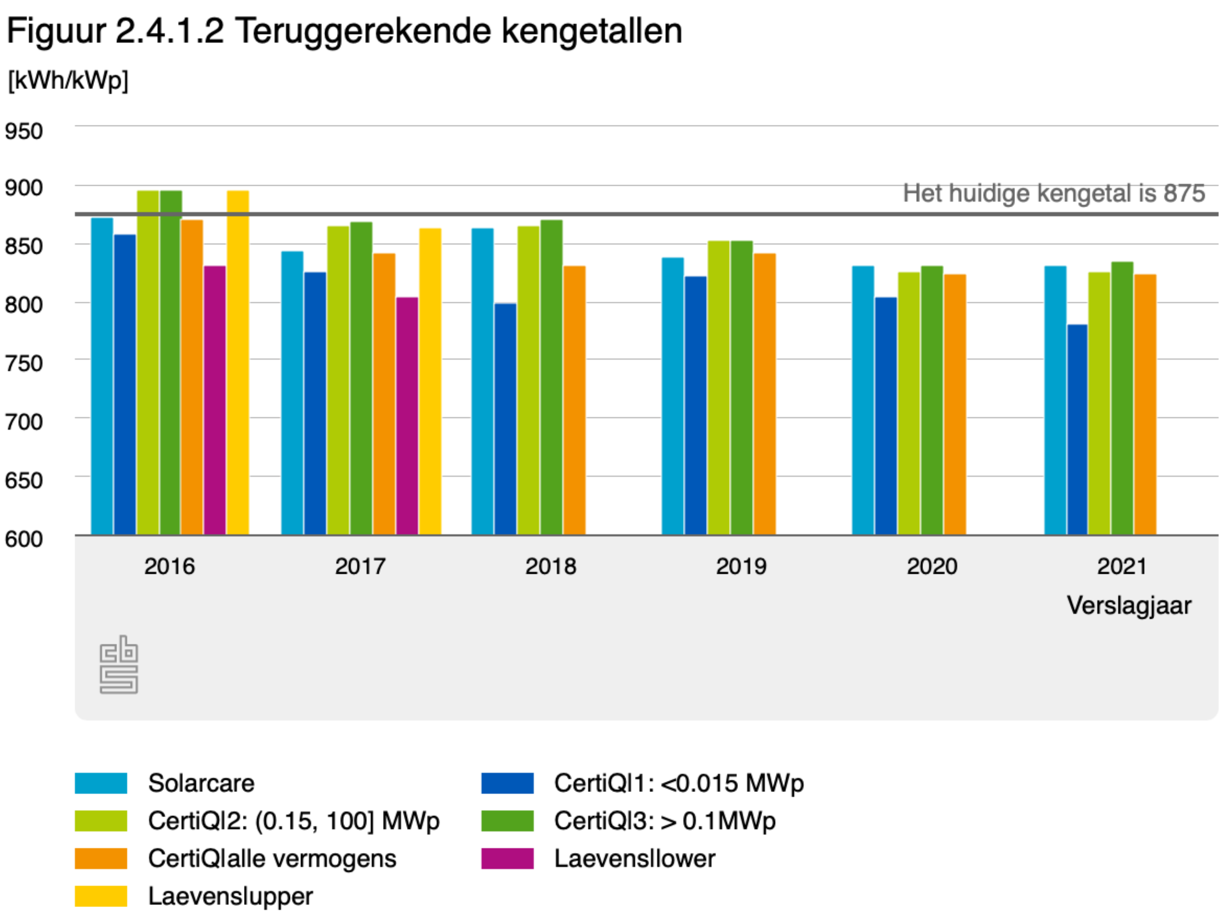


Figure 14. Comparison of specific yields corrected for annual irradiance differences for the period 2016-2021 per size category for data from SolarCare [26], CertiQ, and Laevens [25]. The black solid line indicates the 875 value of specific yield. Figure is taken from [12], with permission.

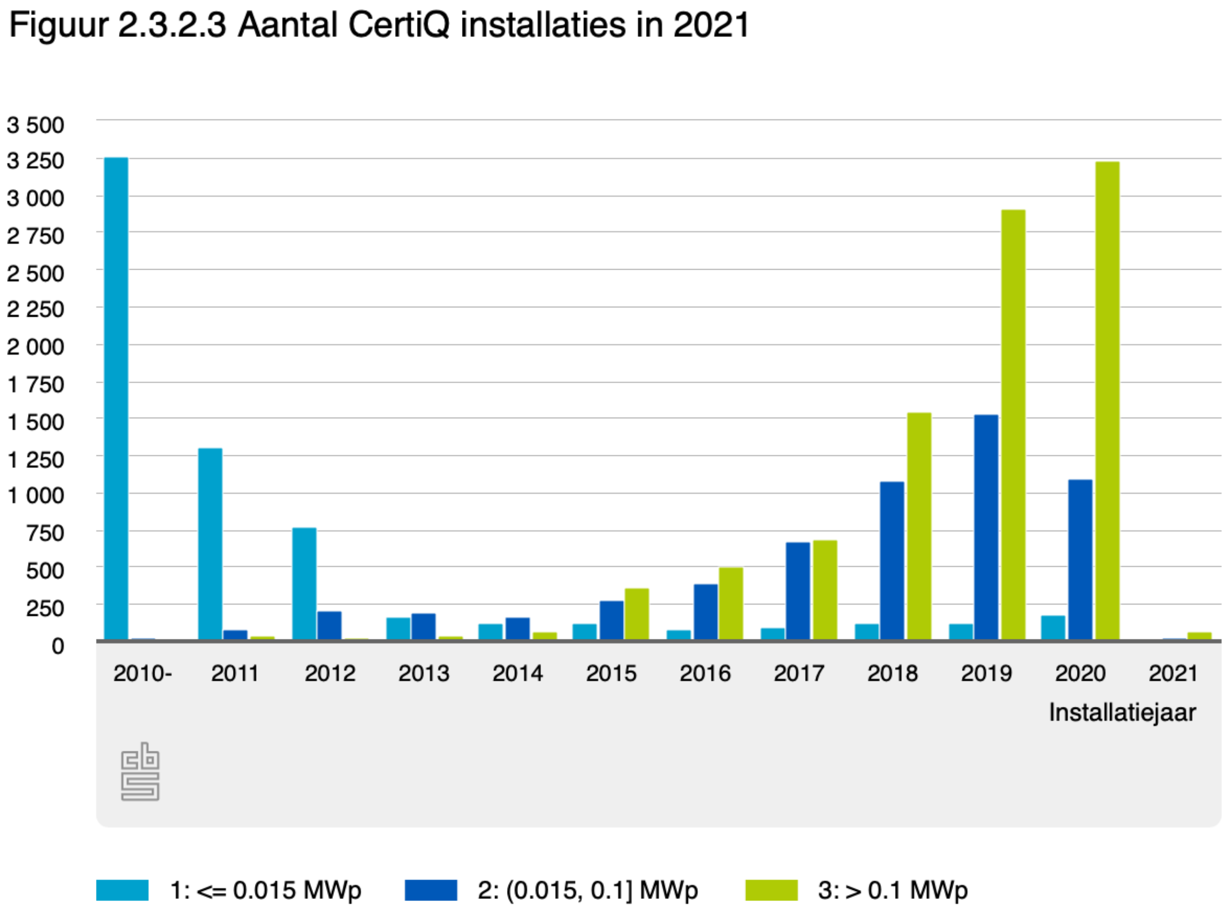


Figure 15. Number of installations subdivided by capacity class used in the analysis of specific yields of CertiQ data as a function of year of installation. Figure is taken from [12], with permission.

In order to further investigate the observed trends, a detailed analysis of this data is performed for the three different capacity classes (small, medium, large) as a function of year of installation. Figure 15 shows the number of systems analyzed subdivided by capacity class, clearly showing that the number of small-sized systems is decreasing in time, whereas the number of large-sized systems has increased. Figure 16 then shows average specific yield since 2010 for these classes. The specific yield of small-sized systems is decreasing from about 850 kWh/kWp to about 700 kWh/kWp, while for medium and large-sized systems the specific yield remains more or less the same. The decrease for small-sized systems may be due to PV systems being installed at less optimal conditions, such as East/West orientations at 10^{o} tilt. More details can be found in [12].

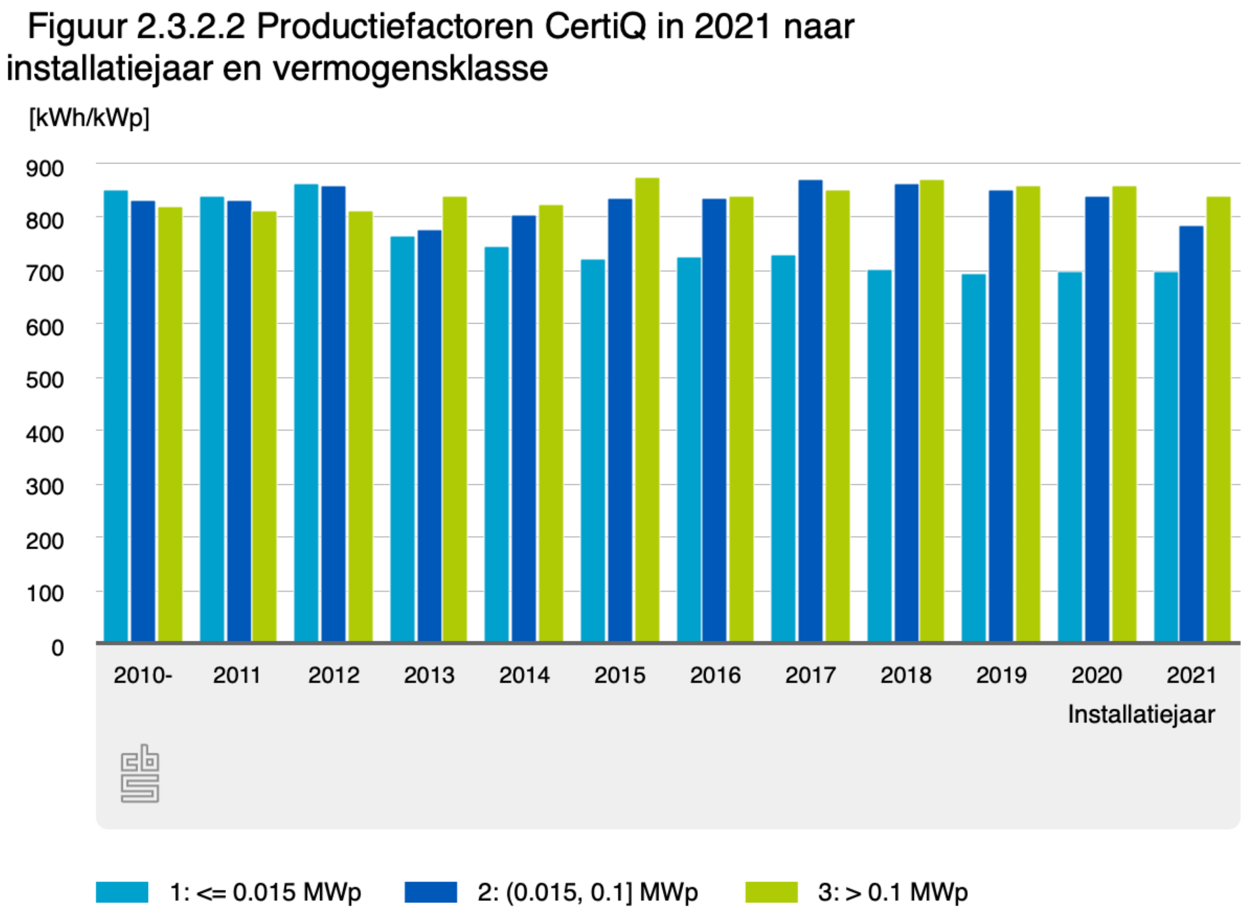


Figure 16. Specific yield corrected for irradiance for CertiQ data for three different size classes as a function of year of installation. Figure is taken from [12], with permission.

3.2.3. Tweakers/PVOUTPUT

Energy yield data at a monthly time resolution is available for the Tweakers/PVOUTPUT from January 2014 onwards. The analysis was carried out until December 2022. In total, this is a maximum of 108 months of data per system. However, Figure 17 shows that actually only for a small number of systems data is available for 108 consecutive months. To analyze energy yield data, it was therefore decided to only include those systems per year for which 12 months of data are available in that year.

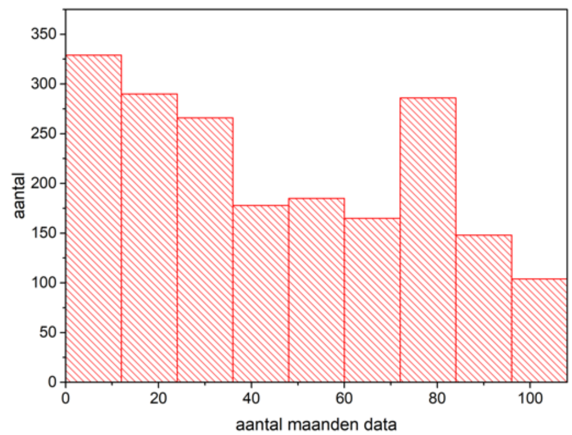
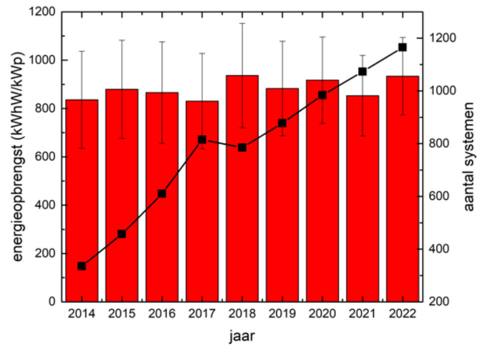


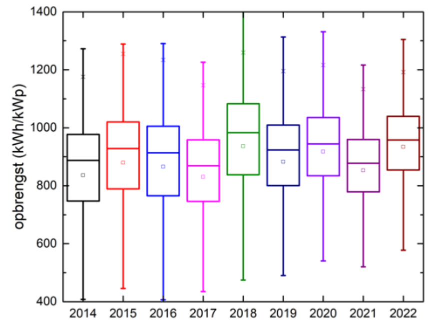
Figure 17. Number of months of performance data available per system.

Figure 18 shows yield data from systems for which 12 months of data is available in the relevant years 2014-2022. In 2014 there were 335 systems, but that number increased to 1165 in the year 2022. The average slope angle and orientation for those systems can also be seen in Figure 18. The figure shows that energy yields are similar across the years, however, with quite a wide range in values. Both orientation and tilt averages are about similar too.

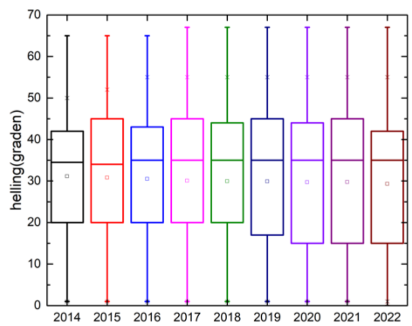
Figure 19a shows energy yield per year compared to the average annual horizontal solar irradiance in the Netherlands as measured by KNMI. The variation in yield correlates well with the variation in solar irradiance. This can be seen more clearly in Figure 19b which shows the ratio between yield and solar irradiation. The variation in the ratio seems to correlate reasonably well with the average ambient temperature. It is well known that a higher module temperature negatively affects the efficiency of a PV panel. A higher average ambient temperature in combination with a higher irradiance leads to relatively higher panel temperatures and as a consequence lower yields of PV systems. This thus results in a lower ratio, which is clearly visible in the figure.



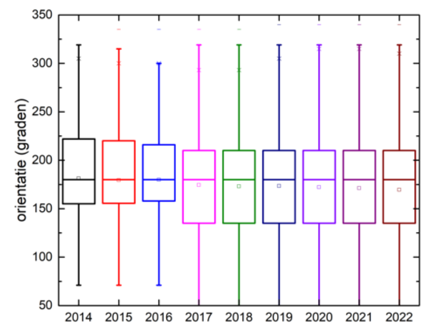
(a)



(b)

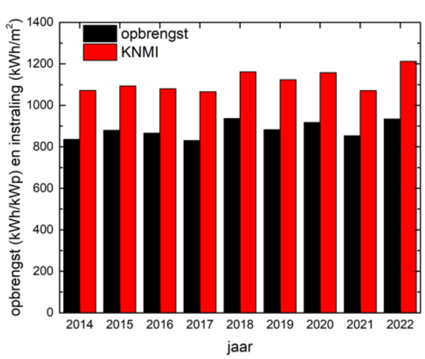


(c)

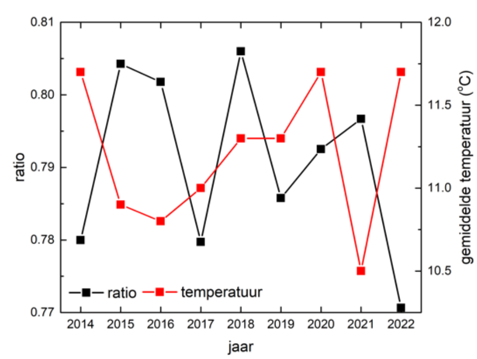


(d)

Figure 18. Mean and standard deviation of yield data from systems for which 12 months of data are available in the relevant years, box plots of (b) yield data, (c) mean tilt, and (d) orientations.



(a)



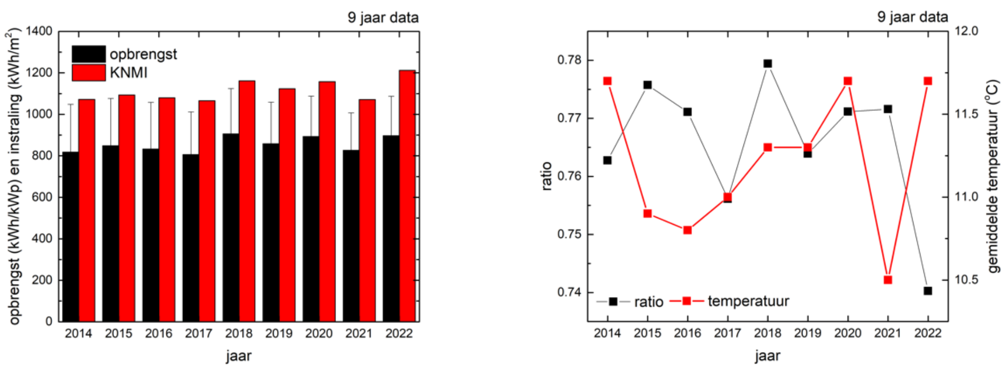
(b)

Figure 19. Yield data from systems for which 12 months of data is available compared to national average horizontally measured global horizontal solar irradiation by KNMI in the relevant years, (b) ratio of yield data and solar irradiation and average ambient temperature.

In order to assess if any degradation of yield is present, data was analyzed from systems with 108 months of data. Of the total of 143 PV systems, the average tilt is 31 ± 13 degrees and the average orientation is 183 ± 52 degrees.

Figure 20a shows the variation in yield and horizontal solar irradiation, and Figure 20b shows ratio between yield and solar irradiation. Here too, yield correlates well with the variation in solar irradiation, but the ratio values are approximately 4% lower. This may be due to technology differences between 2014 and later years.

Based on the variation of the yield/irradiation ratio, it cannot be properly determined if degradation of systems is occurring. It may even be absent.



(a) (b)

Figure 20. (a) Yield data from systems for which 108 months of data is available compared to national average horizontally measured solar irradiation by KNMI in the relevant years, (b) ratio of yield data to solar irradiation and average temperature.

2.2.4. New method: use of plane-of-array irradiance instead of global horizontal

2.2.4.1. Background

PV generation depends on multiple external conditions and system parameters, and the models to predict generation vary widely in complexity. Most important is the solar irradiance. In the data analyzed in previous sections, correction of specific yield is performed using the global horizontal solar irradiance (GHI). As most of the PV systems are installed at a non-zero tilt, the solar irradiance that PV systems receive is the so-called plane-of-array irradiance (G_{POA} or just POA).

In the simplest model, PV power output is just proportional to G_{POA} (not GHI):

$$P_{out} = G_{POA} \frac{P_{STC}}{G_{POA,STC}}, G_{POA,STC} = 1000\ W/m^2$$

where P_{STC} is the nominal or name-plate DC capacity of the modules or array. When this simple model is integrated over time we can simply substitute the integrals of P and G, which are E and H:

$$E_{out} = H_{POA} \frac{P_{STC}}{G_{POA,STC}}$$

In a slightly more realistic model, system losses are represented by a linear loss factor (LF), and the output remains proportional to G_{POA}:

$$P_{out} = G_{POA}(1 - LF) \frac{P_{STC}}{G_{POA,STC}}$$

In the equivalent energy model the same linear losses are more commonly expressed by a performance ratio (PR):

$$E_{out} = H_{POA} \cdot PR \cdot \frac{P_{STC}}{G_{POA,STC}}$$

If LF is constant over time—a simplifying assumption—then:

$$PR = 1 - LF$$

In Figure 21 the ratio between POA and GHI is shown using 11 years of observations. Figure 22 shows a further zoom in. POA is calculated using the Hay-Davies model [27] for a surface tilted at 37° toward the South, which has been reported to be the optimal angle for the Netherlands [13].

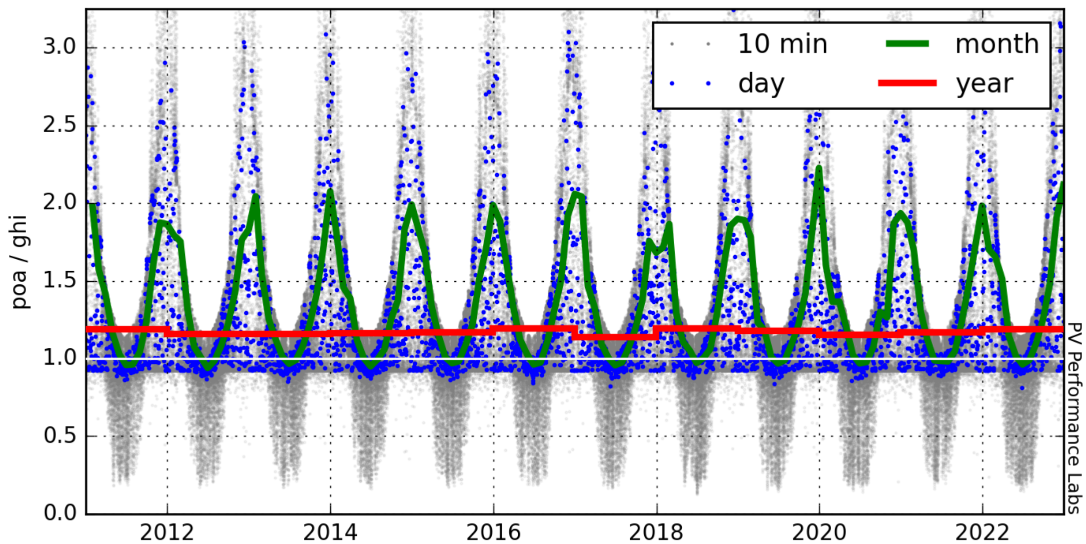


Figure 21. Comparison of POA and GHI by considering their ratio at different averaging intervals, from 10 min to 1 year. The plane of array is facing south and tilted 37°.

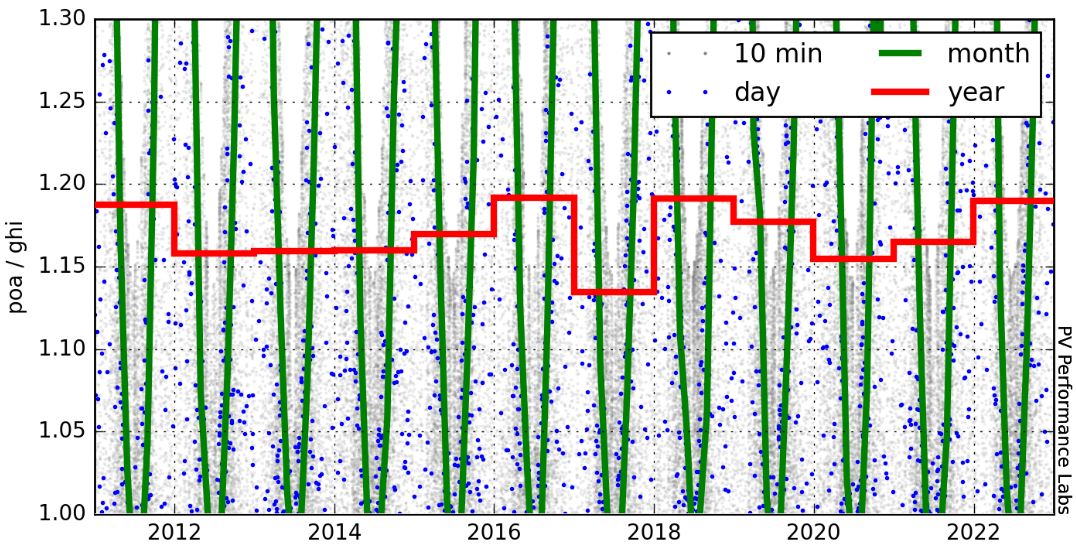


Figure 22. Comparison of POA and GHI by considering their ratio at different averaging intervals, zoom in of Figure 23 to focus on the annual average values.

The important observations from these graphs are:

- There are very large differences between POA and GHI on the shortest time scale of 10-minutes, but these gradually shrink with aggregation period.

- The monthly values show a strong seasonal pattern oscillating between POA/GHI ratios of ~0.9 and ~2.0. A similar pattern was observed by CBS (see section 2.2.1.) when comparing measured system output to estimated output based on GHI for systems tilted at 45° of variable orientation [12]. Thus it is clear that at the monthly level the distinction between POA and GHI is very relevant.
- Annual values of POA/GHI average around 1.17, meaning that there is about 17% more solar energy available on the 37° tilted plane than on the horizontal. However, the advantage fluctuates between 13% and 19% depending on the year. An overall bias would not pose a problem for the CBS adjustment method, but the year-to-year variations add uncertainty to the estimated annual values.
- Despite the annual fluctuations, there does not appear to be an overall increasing or decreasing trend in the annual ratio POA/GHI over this period. A longer period may be needed to reveal or refute such a trend conclusively.

For other tilt angles and orientations (not shown) the annual POA/GHI is generally lower and may also drop below 1.0, which means that the array receives less energy than a horizontal surface. The amplitude of the seasonal fluctuations also varies with tilt and orientation.

The insight that PV output correlates better with POA than GHI is certainly not new or original. The reason POA is not used by CBS for PV yield estimate is the fact that tilt and orientation data is not available. POA is different for each system and calculating it at high temporal resolution for many systems having different locations and orientations would indeed be cumbersome. In the following sections, we demonstrate a practical new method that can be used in place of such detailed calculations.

2.2.4.2. Model to estimate time-averaged POA irradiance

The purpose of tilting the plane of a PV array is to capture a larger portion of the available beam radiation than is available on a horizontal surface, while sacrificing a little bit of the sky diffuse radiation. For hourly and sub-hourly PV yield simulations it is standard practice to calculate the POA irradiance at each time step using the sun position and array orientation, plus knowledge of the direct and diffuse components of the GHI. This calculation—called *transposition*—is not possible using daily, monthly or annual irradiance because sun position cannot be aggregated the same way that irradiance can. We therefore developed a new transposition model that is specifically designed to work with aggregated irradiance values so that daily, but especially monthly and annual POA can be calculated more easily.

This model is tested for the Netherlands using a multi-year 10-minute data set of DNI, GHI and DHI at De Bilt [28,29]. For a variety of surface tilts and orientations POA irradiance was calculated at each time step using the Hay-Davies transposition model [27], and subsequently daily means of all irradiance components are calculated.

The daily POA calculated using the new model proved to correlate much better with daily POA (green points) than daily GHI (black points) as the example in Figure 23 shows. A strong improvement is also seen in the comparison of monthly aggregated irradiance values in Figure 24.

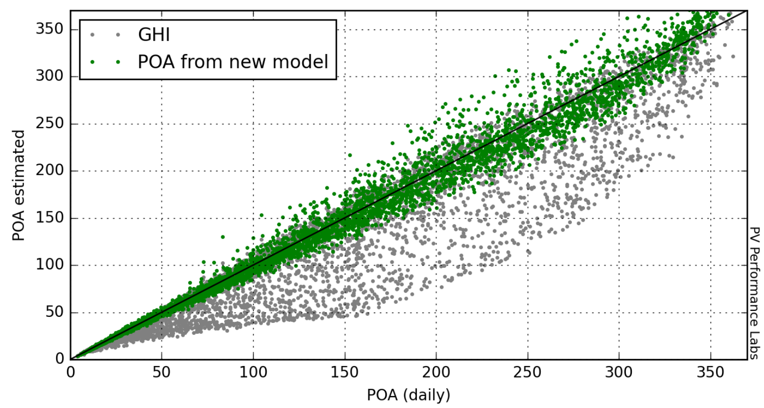


Figure 23. Estimation of <u>daily</u> average POA using the new model. The plane of array is facing south and tilted 37°.

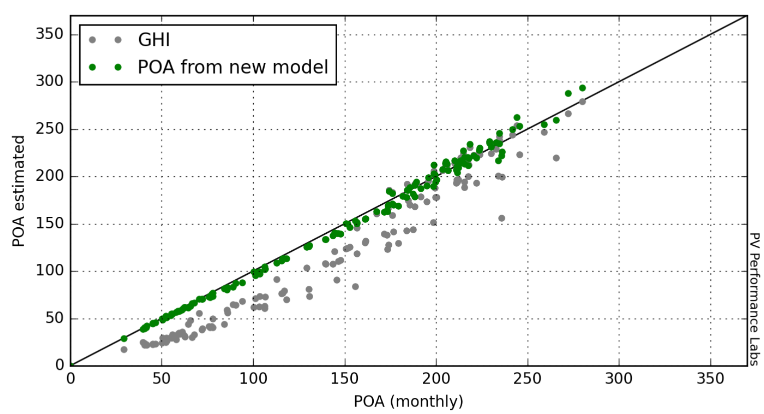


Figure 24. Estimation of <u>monthly</u> average POA using the new model. The plane of array is facing south and tilted 37°.

A final example is given for a vertical surface facing north. While this may be unlikely manner of installing PV it clearly shows the power of the newly proposed method.

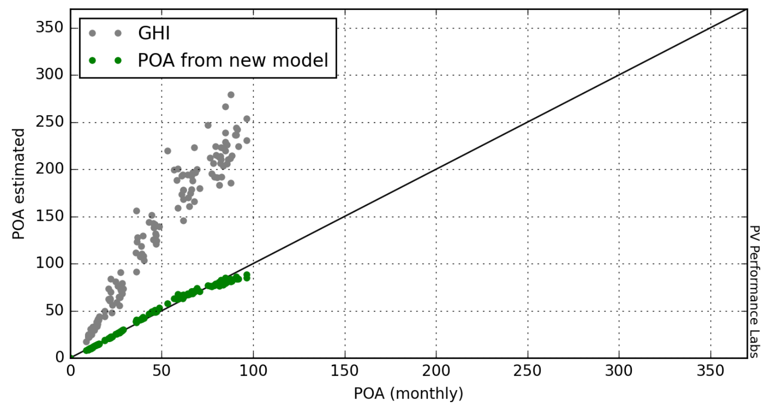


Figure 25. Estimation of monthly average POA using the new model. The plane of array is vertical and facing north.

The new method does need to be fine-tuned for a specific location—in particular its latitude, but to some extent also its weather patterns. It is likely that for the Netherlands the fine-tuning can be done for just one location and used throughout the country.

2.3. Privacy and legal issues

The CERES data base resides at CBS and CBS is legally not allowed to share data on individual systems. It is possible that (scientific) institutions can gain access to pseudonymised data at personal or company level under strict conditions, and only for statistical or scientific research [20]. In the case of third-party data sources, as it the case for the CERES data, permission always needs to be requested first, before such data can be made available for such research. CBS shares aggegrated data for statistical purposes (see, e.g., Figure 3). The analysis of the performance of the 100 systems in Amersfoort has been performed in compliance with privacy regulations and the CBS law [30].

2.4. Business plan

The activities in the project have shown that combination of data bases provides interesting insights in the actual perfornance of PV systems. Having one data base in which meta data and operational data is combined would be ideal. However, due to present regulations, this is not possible. Hence, a business case for a party that would exploit the data is not possible.

However, as shown in Figure 4, some central body, could collect performance data and could connect that to the meta data. Collection could be organized by approaching companies such as SolarCare, but also SunData, to supply performance data for individual systems. It should obviously comply with GDPR regulations. For larger systems, the VertiCer data could be used, if the time resolution would be improved to at least hourly and if access would be arranged. As tilt and orientation are presently not available in the CERES data base, a method such as developed by Laevens et al. [31] could be used to derive that information from operational data. In addition, inverter manufacturers could be asked to supply individual data GDPR compliant to the central body, or aggregated yield numbers.

For example, as CBS shares maps of the Netherlands showing PV capacity per municipality, inverter manufacturers could share maps of yields per municipality as well. This already provides insight in PV performance within the country.

3. Discussion

PVOBS was an ambitious project with innovative and challenging objectives given the available budget. The final goal of setting up one data base with meta data and operational data for all PV systems in the Netherlands has not been reached, while a potential way of organizing it is suggested, involving a central body.

The results of the analysis provided interesting insight in the evaluation of performance of PV systems in the Netherlands. More data at higher time resolution is needed to study aspects that affect performance such as degradation and/or curtailment, especially in regions that suffer from grid congestion or under conditions of negative electricity prices.

At present, global horizontal irradiance is used to correct annual energy yield. An improved method is suggested using calculation of the plane-of-array irradiance. This requires access to direct normal and diffuse horizontal irradiance information, which are available from KNMI, and can be derived and measured at ground stations and from satellite images.

4. Follow up activities

The PVOBS project has provided insights for the discussion of a potential renewal of the annual specific yield of 875 kWh/kWp that is used by CBS to determine the contribution of PV to the national electricity supply. It is clear that in the past 10 years many PV systems are not installed at optimum orientation and tilt. This leads to lower specific yields (see also Figure 2). It is therefore highly recommended to analyse the distribution of PV systems over orientation and tilt. Combining this with data from Figure 2 would provide a more accurate annual specific yield. Such distributions should be updated every year (or month, if monthly specific yields are to be reported). Orientation and tilt information could be derived from digital elevation maps as made by Kadaster. Initial work has been done by CBS and Kadaster [32].

5. Dissemination

Dissemination activities have aimed to promote non-confidential results obtained within the project as swiftly and effectively as possible for the benefit of the whole (scientific) community and to avoid duplication of R&D efforts. Besides papers and reports, participation in workshops is indicated as well.

Media:

- https://www.cbs.nl/nl-nl/over-ons/onderzoek-en-innovatie/project/innovatieve-methode-voor-het-schatten-van-de-opbrengst-van-zonnestroom
- https://solarmagazine.nl/nieuws-zonne-energie/i24209/cbs-presenteert-nieuwe-methode-om-opbrengst-van-zonnepanelen-te-bepalen

Published papers:

- Wilfried van Sark, *Photovoltaics performance monitoring is essential in a 100% renewables-based society*, Joule 7 (2023) 1388-1393. (doi:10.1016/j.joule.2023.06.012)
- Benjamin P. M. Laevens, Olav ten Bosch, Frank P. Pijpers, Wilfried G.J.H.M. van Sark, *An Observational method for determining daily and regional photovoltaic solar energy statistics,* Solar Energy 228 (2021) 12 – 26. (doi:10.1016/j.solener.2021.08.077)
- Benjamin P. M. Laevens, Frank P. Pijpers, Harm Jan Boonstra, Wilfried G.J.H.M. van Sark, Olav ten Bosch, *A Markov Chain Monte Carlo approach for the estimation of photovoltaic system parameters*, Solar Energy 265 (2023) 112132. (doi:10.1016/j.solener.2023.112132)

References

[1] Dutch New Energy Research, "Nationaal Solar Trend Rapport 2019," 2019.

[2] Dutch New Energy Research, "Nationaal Solar Trend Rapport 2023," 2023.

[3] W. Folkerts, W. van Sark, C. de Keizer, W. van Hooff, and M. van den Donker, "Roadmap PV systemen en toepassingen," 2017.

[4] CBS Statline, "Zonnestroom; vermogen bedrijven en woningen, regio (indeling 2019)", https://opendata.cbs.nl/statline/#/CBS/nl/dataset/84783NED/table?ts=1706000683647

[5] CBS Statline, "Zonnestroom; vermogen en vermogensklasse, bedrijven en woningen, regio", https://opendata.cbs.nl/statline/#/CBS/nl/dataset/85005NED/table?ts=1695713879046

[6] RVO, "Protocol Monitoring Hernieuwbare Energie", 2015.

[7] RVO, "Protocol Monitoring Hernieuwbare Energie – Herziening 2022", https://www.rvo.nl/sites/default/files/2022-07/Protocol-Monitoring-Hernieuwbare%20Energie.pdf , 2022.

[8] W. G. J. H. M. Van Sark, L. Bosselaar, P. Gerrissen, and K. Esmeijer, "Update of the Dutch PV specific yield for determination of PV contribution to renewable energy production: 25% more energy!," in *Proceedings 29th European Photovoltaic Solar Energy Conference and Exhibition*, 2014, pp. 4095–4097.

[9] CBS, "Slim zonnestroom in kaart brengen.", https://www.cbs.nl/nl-nl/onze-diensten/innovatie/project/slim-zonnestroom-in-kaart-brengen.

[10] KNMI, Maandoverzicht van het weer in Nederland (MOW), https://www.knmi.nl/nederland-nu/klimatologie/gegevens/mow.

[11] A. Louwen, R.E.I. Schropp, W.G.J.H.M. van Sark, A.P.C. Faaij, "Geospatial analysis of the energy yield and environmental footprint of different photovoltaic module technologies", Solar Energy 155 (2017) 1339–1353.

[12] Karen Brandenburg, Hui Shan Chan, Sabine Krieg, Manon van Middelkoop, Reinoud Segers, "Onderzoek naar productiefactoren zonnestroom in 2022", 2023. https://www.cbs.nl/nl-nl/longread/diversen/2023/onderzoek-naar-productiefactoren-zonnestroom-in-2022?onepage=true

[13] Agentschap NL, "Leidraad Zonnestroomprojecten", Publicatie-nr. 2DENB1003, 2010.

[14] CBS Statline, "Hernieuwbare elektriciteit; productie en vermogen" https://opendata.cbs.nl/statline/#/CBS/nl/dataset/82610ned/table

[15] CBS Statline, "Hernieuwbare energie; verbruik naar energiebron, techniek en toepassing", https://opendata.cbs.nl/statline/#/CBS/nl/dataset/84917NED/table?ts=1700658214274

[16] CBS, "46 procent meer energieproductie uit zonnestroom in 2022", https://www.cbs.nl/nl-

nl/nieuws/2023/24/46-procent-meer-energieproductie-uit-zonnestroom-in-2022

[17] Wilfried van Sark, *Photovoltaics performance monitoring is essential in a 100% renewables-based society*, Joule 7 (2023) 1388-1393.

[18] https://klimaatmonitor.databank.nl/jive

[19] EDSN, https://www.edsn.nl

[20] CBS, “Privacy”, https://www.cbs.nl/nl-nl/over-ons/dit-zijn-wij/onze-organisatie/privacy

[21] Sofie De Broe, Ralph Meijers, Olav ten Bosch, Bart Buelens, Alex Priem, Ben Laevens, Tim de Jong, Marco Puts, “From experimental to official statistics: The case of solar energy”, Statistical Journal of the IAOS 35 (2019) 371–385.

[22] https://pvoutput.org/region.jsp?country=165

[23] Anton Boonstra, https://twitter.com/antonboonstra

[24] Wilfried van Sark, “Analyse PV systemen beschikbaar via Tweakers.net”, november 2023.

[25] Benjamin P. M. Laevens, Olav ten Bosch, Frank P. Pijpers, Wilfried G.J.H.M. van Sark, “An Observational method for determining daily and regional photovoltaic solar energy statistics”, Solar Energy 228 (2021) 12 – 26.

[26] SolarCare, https://www.solarcare.eu/nl/nieuws/12-gemiddelde-zon-pv-opbrengst-nederland.html

[27] Hay, J.E., J.A. Davies, “Calculations of the Solar Radiation Incident on an Inclined Surface.” In *Proceedings of First Canadian Solar Radiation Data Workshop*. Ministry of Supply and Services, Canada, 1980.

[28] Copernicus Atmosphere Monitoring Service (CAMS) Atmosphere Data Store (ADS), CAMS solar radiation time-series, 2023. https://ads.atmosphere.copernicus.eu/cdsapp#!/dataset/cams-solar-radiation-timeseries?tab=overview

[29] Schroedter-Homscheidt, M., Azam, F., Betcke, J., Hanrieder, N., Lefèvre, M., Saboret, L., Saint-Drenan, Y. -M., “Surface solar irradiation retrieval from MSG/SEVIRI based on APOLLO Next Generation and HELIOSAT-4 methods”, Meteorologische Zeitschrift 31 (2022), 455 – 476.

[30] Overheid, Wet op het Centraal bureau voor de statistiek, https://wetten.overheid.nl/BWBR0015926/2022-03-02

[31] Benjamin P. M. Laevens, Frank P. Pijpers, Harm Jan Boonstra, Wilfried G.J.H.M. van Sark, Olav ten Bosch, “A Markov Chain Monte Carlo approach for the estimation of photovoltaic system parameters”, Solar Energy 265 (2023) 112132.

[32] CBS, Detectie zonnepanelen regionaal, https://www.cbs.nl/nl-nl/maatwerk/2022/36/detectie-zonnepanelen-regionaal