CellPress

Joule
Commentary

Photovoltaics performance monitoring is essential in a 100% renewables-based society

Wilfried van Sark[1],*

Wilfried van Sark is full professor of "Integration of Photovoltaics" at the Copernicus Institute of Sustainable Development. He is an experimental physicist by training (MSc/PhD) and has over 35 years experience in the field of PVs. He worked on various material systems such as crystalline and thin-film silicon and III-V solar cells, both experimentally and theoretically. His current activities focus on employing spectrum conversion to increase solar cell conversion efficiency for next-generation PV energy converters, and performance analysis of building-integrated and standard PV systems. This in particular links to integration of PV systems in smart grids in the built environment.

Photovoltaic (PV) solar energy installed capacity globally has reached the terawatt level in early 2022.[1] Although this is based on estimations as to the exact date or amount, as reporting of installed capacities is notoriously difficult, we can be pretty sure that 1 TWp installed capacity was reached before summer 2022. Annual markets are now around or approaching 200 GWp, and growth will continue.[2] This is deemed necessary to reach a (net) zero-emissions society, or a 100% renewables-powered society, by mid-century, which would require some 15–75 TWp of PV power[3] in combination with other renewable sources—in particular wind and storage.

When relying solely on renewables such as solar and wind for which the amount of generated energy depends on the weather, security of supply, defined as the uninterrupted availability of energy sources at an affordable price, may be at stake.[4] Energy security will no longer be depending on finite fossil fuel resource,s but mostly on infinite photon flows. Accurate weather forecasts in combination with optimized PV, wind, and storage capacity, as well as interconnection capacity between countries, are required. The least-cost optimization of a combination of PV, wind, and storage capacity at which renewable energy is providing security of supply has been proposed using the concept of "firm power". This requires increasing the capacity of PVs, wind, and storage to a certain extent by overbuilding capacity in combination with curtailment of power generation. This comes at a loss of energy and revenue, leading to increased levelized cost of electricity (LCOE). The ratio of the LCOE associated with firm power and the one for unconstrained PVs is defined as "firm kWh premium". The optimum, i.e., minimum value, of this firm kWh premium has been reported to range from 3 to 5 in several case studies for PV/wind curtailment of about 40%–60%,[5] while it is larger than 25 for zero curtailment due to much higher storage capacity requirements and associated costs.

Transitioning from a fossil-fuel-dominated energy system to a renewables-based one requires adequate monitoring of power generation of all generators, from small-sized PV modules to GW-sized PV and wind parks. Today, power output of existing power plants (centralized fossil-based, nuclear) is well monitored in the European Union. A database is available as open data from the Joint Research Center (JRC) Open Power Plants Database. Power generation of about 7,000 power plants is monitored and collected at the ENTSO-E (European Network of Transmission System Operators for Electricity) level and is represented in their transparency dashboard for most European countries.[6]

There data also form the basis for carbon emission intensity expressed in gCO_2eq/kWh per country in Europe as shown in Figure 1, a map made from data provided by Electricity Maps for March 1, 2022. Emission levels are color coded, and one can easily spot the low emissions in France, which are a result of high nuclear capacity, as opposed to the high emissions in Poland, due to the high coal-based running capacity.

Regarding wind energy, generation data are available, as wind turbines are typically installed in wind parks of normally 5 to >100 MW size, and proper Supervisory Control and Data Acquisition (SCADA) systems are in place as park owners trade electricity on the various markets.

For PV the situation is different. Figure 2 shows the annual evolution of grid-connected PV systems in Europe in the past 10 years illustrating that over 50% of the PV systems are distributed systems, i.e., in residential areas, and typically of small size (<15 kWp).[2] Globally, just over 40%

---

[1] Utrecht University, Copernicus Institute of Sustainable Development, Princetonlaan 8a, 3584 CB Utrecht, the Netherlands

*Correspondence: W.G.J.H.M.vanSark@uu.nl

https://doi.org/10.1016/j.joule.2023.06.012

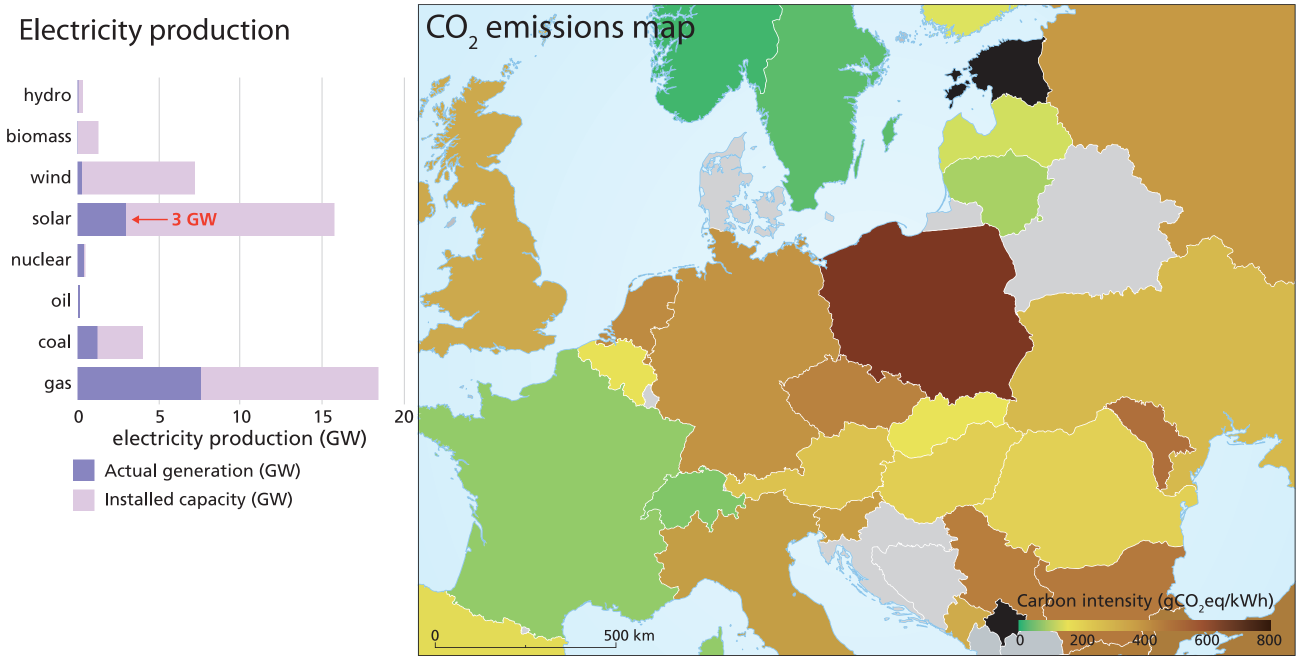

Figure 1. Carbon emission intensity in gCO$_2$eq/kWh per country for March 1, 2022

Emission levels are color coded for fast visual interpretation. Data used is from Electricity Maps (https://app.electricitymaps.com/map). The left panel provides details for the Netherlands, showing 3 GW of PVs being generated at 1 p.m. on March 1, 2022. Emission is 396 g CO$_2$eq/kWh based on the generation mix and emission factors per generator.

of the systems are distributed ones.[2] For these systems, (near) real-time power generation data are not accessible. Note, though, that power generation data are available locally on the inverter level, either on a display or—due to data uploads to data clouds of inverter manufacturers—on a personalized app.

The question now is how this distributed data can be used, without violating General Data Protection Regulations (GDPR), to determine the contribution of PVs to the electricity demand. As an example, one inverter manufacturer provides hourly data from its own cloud, in combination with assumed market penetration, to show hourly generation for 40+ regions in Germany (SMA, https://www.sma.de/en/company/pv-electricity-produced-in-germany). Similarly, the Australian PV Institute (APVI, https://pv-map.apvi.org.au) presents live (15 min time resolution) PV generation maps, taken from 6,000+ PV systems across Australia, supplied by the company Solar Analytics.

Given the fact that, at present, about half of the PV systems are small-scale residential systems for which no generation data are known, we argue that the security of electricity supply is based on quicksand rather than on solid rock. With future scenarios of PV reaching 600 GWp in Europe by 2030,[7] this is simply unacceptable and seriously jeopardizing the security of supply. Direct access to PV generation data is essential to enable DSOs (distribution system operators), TSOs (transmission system operators) and other stakeholders to guarantee a reliable and sustainable electricity supply.

Here, we will argue that some kind of data monitoring and database system should be in place to guarantee security of supply in a PVs-/wind-dominated energy system. This goes beyond the recommendations from IEA-PVPS[8] that focus on setting up databases of installed capacity only, which is referred to as meta-data (or static data) of PV systems. The report also analyses the present data model for 13 countries, excluding the Netherlands. In this perspective, recommendations for operational (dynamic) data are given. In support of this, we analyze the PV landscape in the Netherlands as a case study example.

Case study: the Netherlands

Figure 1 also shows details for electricity production in the Netherlands. The actual generation power and maximum capacity per source are indicated. For example, a PV capacity of 16 GW[2] is installed, while 3 GW is generated at 1p.m. due to weather conditions. The 7 GW wind capacity did generate less than 1 GW, similar to nuclear. Gas-fired power plants generate the rest, i.e., about 8 GW, while maximum installed capacity is 18 GW. Hence, total demand was about 13 GW, which is an average power demand at midday. With assumed emission factors, this leads to an emission intensity of 396 g CO$_2$eq/kWh.

The total installed capacity of 16 GWp could generate about 12–13 GW on a sunny day, due to increased panel

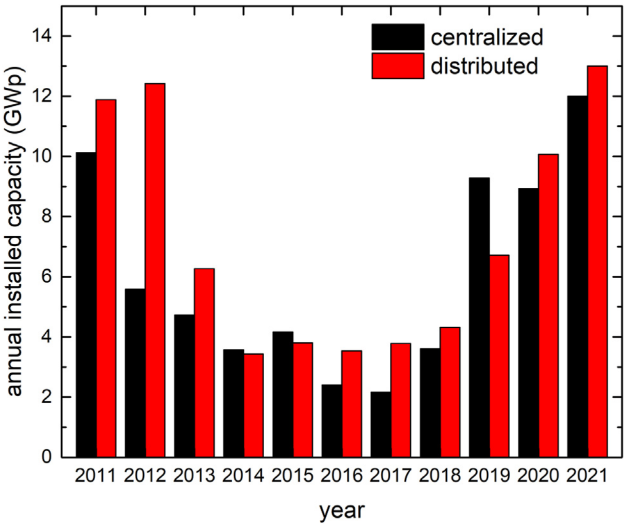

Figure 2. Annual evolution of grid-connected PV installations in Europe, per segment (utility-scale and distributed/residential)
(Data source: IEA-PVPS[2]).

temperatures that lead to lower conversion efficiencies. That would thus be providing for nearly all demand, and conventional power plants would need to be shut down or else would be exporting massive amounts of electricity to neighboring countries as long as interconnection transport capacity is not limiting that ability. This already leads to negative electricity market prices around noon. Also, defining relative overbuild capacity as the ratio of PV capacity and peak load capacity shows that this ratio is larger than one for the Netherlands already. This means that curtailment is necessary on sunny days in May, which are relatively cold. As has been reported, a 50% continuous curtailment leads to only 14% annual yield loss (in the Netherlands).[9] With the ongoing additions of PV capacity, curtailment as a cheap option will be competing with battery storage and hydrogen generation.

However, actual PV power generation is estimated and not measured. This is because about half of installed PV capacity is in the form of small-size (<15 kWp) residential systems for which no central monitoring system is in place, similar to the average situation in Europe. In contrast, large-size (>15 kWp) PV parks have been realized up to 80 MWp, which was incentivized by national subsidy schemes (SDE+(+)[10] that provide a premium per kWh delivered to the electricity grid. For this a calibrated energy meter is required. Administration of the subsidy scheme is done by a certification body (CertiQ), which collects information on installed capacity (MWp) and annual energy generated (kWh). Hence, for utility scale, PV generation data is available; however, not on a sufficiently high time resolution. Assuming though that the installed PV capacity is known, meteorological data (solar irradiance, temperature, and wind data, from satellites and/or ground-based stations) can be used to calculate the power generated at the same time resolution (15 min) of the meteorological data, while orientation and tilt data usually are not known.[11]

For systems <15 kWp, net metering is in place but it will be abolished in 2031 and phase out will start in 2025. For the majority of these systems, smart meters are measuring electricity taken from and delivered to the grid, which is readable via separate displays. Note that electricity delivered to the grid is PV generation minus direct in-house consumption. Although this data is collected by DSOs, it is transferred to independent data operators and utilities are only allowed to access this data once every 2 months for billing purposes, except when given explicit permission by owners to access data every 15 min.

For PV power systems, a centralized database of the total installed capacity is available at Statistics Netherlands (CBS). It is mandatory to register any PV system in databases operated by the local DSOs, which then are collected nationally by Statistics Netherlands. The database contains about 2 million small (1–5 kWp) systems in residential areas as well as many multi-MWp systems in predominantly rural areas. This database contains system capacity (DC and AC in many cases) and location without information regarding orientation and tilt.[12] Now, as stated in the protocol on monitoring renewable energy (RVO[12]), estimates of the amount of energy generated can be provided per month only, which is based on the actual average weather—i.e., global horizontal irradiance and ambient temperature—from 30+ meteorological stations in the country.[12] The analysis method would allow us to calculate (sub-)hourly, daily, annual, national, and regional energy generation, and could in principle be used for forecasting PV power based on weather forecasts, such as provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). However, the PV system database is not publicly accessible.

No real-time power generation data is available in the database. A recent spatially resolved (per municipality) analysis has been performed based on open data available at https://pvoutput.org/.[13] Here, PV system owners upload generation data, manually or automatically, on a voluntary basis. Annual specific yields calculations in kWh/kWp have been performed for 2016 and 2017, using satellite-based irradiance data and assuming that all systems were oriented to the South.[11] It was found that actual measured annual yields are about 4% lower, which is because not all systems are oriented South. System tilts vary between 20° and 50°. From this analysis one could infer that calculating annual yields assuming south orientations will only lead to a 4% larger value. However, it should be noted that systems installed before 2015 were mostly installed at near-ideal orientation and tilts as optimum yields were required to reach economic payback times of 7–10 years. Nowadays, as PV systems are considerably cheaper, non-ideal orientations and tilts are used, e.g., systems combining rows of East- and West-facing modules with 10° tilt are becoming more common on flat roofs, as this optimizes generated kWh per m^2 roof surface. Thus, the method of calculating annual yields assuming south orientations will yield incorrect results.

What to do?

We propose the following actions to address the issues described above.

1. Organize a PV installation database
2. Perform yield calculations
3. Determine actual yield
    a. using representative set of PV systems
    b. using data from inverter manufacturers

1. Organize a PV installation database

As argued also by IEA-PVPS,[8] a database with PV installations, each having its own unique system identification (SID), should be organized by countries and national statistics institutes (NSIs) are most probably best equipped to host such a database, as well as to provide access to various stakeholders, assuring that GDPR is not violated. Procedures should be in place to allow researchers to have access to all data for research purposes, whereas aggregated data could be open or restricted based on different aggregation levels. Ideally, the database should contain information on location and owner, total installed capacity (AC and DC), and installation date. It is highly recommended to also include orientation and tilt. Further information, such as cell technology, module and inverter brands, installation type (building applied or integrated, floating, etc.), and cost would qualify as ''nice to have''. This information is necessary to report statistical data on development of PV capacity in countries and should be disclosed on a monthly basis.

2. Perform yield calculations

With the basic PV installation information, a *maximum* yield can be calculated assuming systems to be oriented south at optimal tilt, such as done in Laevens et al.[11] We note that optimal tilt differs per country and depends on the latitude of the system. The European Commission tool Photovoltaic Geographical Information System (PVGIS)[14] allows us to predict annual/monthly/daily yields of PV systems in all European countries, including finding the optimum tilt, based on historical solar irradiance data. If orientation and tilt are provided, actual system yield can be predicted, under the assumption that all PV systems are functioning properly without suffering from shading or technical malfunctions. Associated yield loss could well be some 10% annually. One can also include published degradation rates if the installation date is known.

TSOs and DSOs and other energy market stakeholders can use (aggregated) PV installation data to predict maximum PV yield based on weather forecasts (i.e., forecast of global horizontal irradiance [GHI]) in support of energy market trading and (local) congestion management, thus ensuring the security of supply. We note that with increasing PV penetration, PV system output will need to be actively managed, i.e., employing curtailment and storage, which will introduce new aggregator parties to the market.

3. Determine actual yield

Ideally, all operational data of every PV installation is available in a centralized (per country) database, is updated every 5 or 15 min, and is used in the ENTSO-E transparency platform in some aggregated form. One thus can treat all PV systems in a country as one large power plant of multi-GW size. Such operational databases can be managed by NSIs per country in collaboration with TSOs and DSOs. To arrive at that ideal situation from today's total absence of such a database, several intermediate approaches can be taken.

Representative set of PV systems

For every country, a representative set of PV systems should be found that should be monitored at 5-min time resolution, thus providing PV energy yields. These should include systems of different sizes and different types of locations (urban, sub-urban, rural, floating, etc.). If needed, non-disclosure agreements (NDAs) could be implemented. Although data acquisition may be a technical challenge, it should be noted that various companies offer services to PV system owners to monitor PV systems to optimize financial revenue, and hence organize fast responses to potential failures or underperformance. Data from inverters or from separate power sensors can be used for performance analysis. These companies run their operations based on fleet management principles, which will flag any operational issues.[15] Supplying (part) of these data to an NSI should be easy to organize as NSIs are bound by law not to disclose confidential information. In addition, the citizen-science-based approach followed by https://pvoutput.org// could also be used, where the NSI could acquire national data. We note that independent data operators (IDOs) could also take the role of operational data collector and distributor. The NSIs subsequently provide these data to parties such as TSOs, DSOs, and other energy market stakeholders (utilities, aggregators, energy traders) at various aggregation levels for use in their operations. Figure 3 summarizes this framework and data flows. These parties then can upscale the operational data with

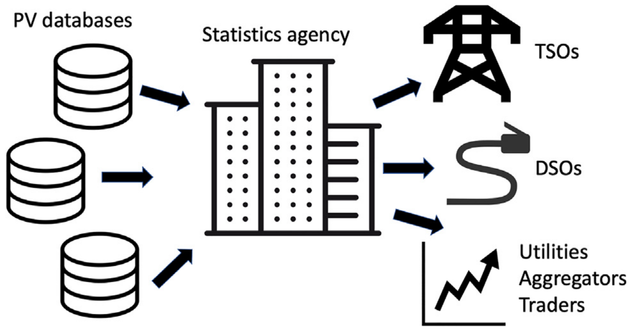

Figure 3. Organization of PV data
Data are sent from several companies to the national statistics institute, who provides it to TSOs, DSOs, and other energy market stakeholders.

installed capacity data, which is also provided by the NSI.

As actual PV yields are used, direct usage within households or companies is not accounted for. This self-consumption can be found from comparing the generation data with actual measured data at transformer stations (if they are available). Likewise, active curtailment of larger-size PV plants may not be apparent in the data. On a higher aggregation level actual TSO/DSO data from imbalance or other (local) congestion markets can be used to validate the upscaled data. It can be envisaged that machine learning approaches are suitable to perform these validations.

The key issue in this approach is the representativity of the PV operational dataset. Statistical analysis of the full dataset provides distributions of system size, tilt, and orientation. Based on these, PV system yields can be calculated using actual weather data. Comparison with operational data distributions of daily/weekly/monthly yields allows us to assess the representativity of the dataset and may lead to the use of correction factors. If data on orientation and/or tilt are not available, methods exist for the analysis of data for several sunny days, thus allowing us to infer apparent orientation and tilt, provided the time resolution is sufficiently high, i.e., 5–15 min.[16] The NSI as data collector should validate the representativeness of the subset of collected PV system data. This is an ongoing process, as many new PV systems will be installed throughout the year at most probably non-optimal orientations and tilts. A monthly update most probably will suffice.

Data from inverter manufacturers

Nowadays all inverter manufacturers operate a cloud-based data acquisition and visualization system, which allows the PV system owners to follow energy generation over the day. These data are available in the database of the inverter manufacturer. The inverter manufacturer can easily provide aggregated data for the whole country, region, or municipality.

If a country decides to design a policy for PV data sharing, in the interest of security of supply, inverter manufacturers would be obliged to share their PV yield data along with other data, such as owner information, location, and system size. It should be made clear, e.g., by the DSO, to the PV system owner that PV data will be aggregated to such levels that individual PV data never can be traced back to its owner, thus complying with GDPR. An NSI would be the data-receiving party, as any information onthe number of systems and market share of manufacturers cannot be disclosed by law. Also, in this way a cross-validation with the national installations database is possible, linking the national SID with the ID in the inverter database. As stated above, if data on orientation and/or tilt are not available, they can be added to the national database after yield data analysis on sunny days.

This approach will lead to the realization of a high time-resolution PV energy generation database, but the process to get there may not be easy and should be enforced by national governments developing the necessary policies and legislative procedures. If this is not possible, inverter manufacturers may still be urged to share data on a voluntary basis. Even if a few inverter manufacturers would supply data, as their inverters will be scattered over the country, scaling the aggregated power to the total installed power would be sufficient for network operation; however, representativeness of this subset of PV systems should be investigated.

For the recommendations to be followed, the main challenge is an organizational one. Data are available and collected, such as via inverter data clouds, and these should be shared. At the state level, ministries responsible for energy and/or relevant national authorities should design and implement legal measures to ensure PV data sharing in support of ensuring energy security. The national statistics agency is responsible for collecting vast amounts of various data, and PV data will be part of this collection process, for which existing procedures can be used. Data quality checks need to be in place, for which national statistics agencies are already well equipped.

Conclusion

We have discussed the importance of PV for a 100% renewables-based society and the necessity for the availability of operational data for grid operators. PV systems installed in residential areas are monitored, but performance data remains local and/or is available in data clouds of inverter manufacturers only. Grid operators do not have access to that data and cannot guarantee security of supply, especially in the future with massive amounts of PV systems feeding power into the electricity grid. Based on an analysis of the situation in the Netherlands, which is typical for many other countries as well—including countries with well-developed, as well as emerging, PV markets we provide recommendations to overcome this problem. These are based on three aspects, i.e., organization of a national PV installation database, performing yield calculations, and determining actual yields using either a representative set of PV systems and/or using data from inverter manufacturers. We believe it is urgent to address this issue so that grid operators can also guarantee security of supply in the future. Based on these databases, organizations such as IRENA (International Renewable Energy Agency) could demonstrate the contribution of PV to the world's electricity needs. Also, comparing performances from these databases in relation to climatic zone differences would elucidate climate-dependent performance and degradation differences.

ACKNOWLEDGMENTS

We are grateful to Bob van den Berg, Ben Laevens (CBS), Lex Schiebaan (Sundata), and Anne Gerd Imenes (University of Agder) for fruitful discussions as well as the COST Action Pearl-PV for allowing us to present first ideas during the training school "Potential of monitoring tools, advanced operation and maintenance practice for security and predictability of PV performance" organized 8–11 March 2022 at the University of Twente, Enschede, the Netherlands. We acknowledge financial support from the Dutch Ministry of Economic Affairs and Climate Policy through Topsector Energy funding for the TKI-Urban Energy project PV Observatory.

AUTHOR CONTRIBUTIONS

W.v.S. wrote and revised the initial draft.

DECLARATION OF INTERESTS

The author declares no competing interests.

REFERENCES

1. Weaver, J.F. (2022). World has installed 1TW of solar capacity. 15 March 2022. PV magazine. https://www.pv-magazine.com/2022/03/15/humans-have-installed-1-terawatt-of-solar-capacity/.
2. IEA-PVPS (2022). Trends in Photovoltaic Applications 2022. Report IEA PVPS T1-43:2022. https://iea-pvps.org/trends_reports/trends-2022/.
3. Haegel, N.M., Verlinden, P., Victoria, M., Altermatt, P., Atwater, H., Barnes, T., Breyer, C., Case, C., De Wolf, S., Deline, C., et al. (2023). Photovoltaics at multi-terawatt scale: Waiting is not an option. Science *380*, 39–42. https://doi.org/10.1126/science.adf6957.
4. Cergibozan, R. (2022). Renewable energy sources as a solution for energy security risk: Empirical evidence from OECD countries. Renew. Energy *183*, 617–626. https://doi.org/10.1016/j.renene.2021.11.056.
5. Perez, M., Perez, R., Rábago, K.R., and Putnam, M. (2019). Overbuilding & curtailment: The cost-effective enablers of firm PV generation. Sol. Energy *180*, 412–422. https://doi.org/10.1016/j.solener.2018.12.074.
6. ENTSO-E (2022). Transparency Platform. https://transparency.entsoe.eu/dashboard/show.
7. European Commission (2022). EU Solar Energy Strategy, Communication from the commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. COM/2022/221 final. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2022:221:FIN.
8. IEA-PVPS (2020). Data model for PV systems, Data Model and Data Acquisition for PV registration schemes and grid connection evaluations – Best Practice and Recommendations, report IEA-PVPS-T1/14-01:2020. https://iea-pvps.org/key-topics/data-model-and-data-acquisition-for-pv-registration-schemes-and-grid-connection-best-practice-and-recommendations/.
9. Braat, M., Tsafarakis, O., Lampropoulos, I., Besseling, J., and Van Sark, W.G.J.H.M. (2021). Cost-Effective Increase of Photovoltaic Electricity Feed-In on Congested Transmission Lines: A Case Study of The Netherlands. Energies *14*, 2868. https://doi.org/10.3390/en14102868.
10. RVO (2022). Stimulering Duurzame Energieproductie en Klimaattransitie (SDE++). https://www.rvo.nl/subsidies-financiering/sde.
11. Laevens, B.P., Ten Bosch, O., Pijpers, F.P., and van Sark, W.G. (2021). An Observational method for determining daily and regional photovoltaic solar energy statistics. Sol. Energy *228*, 12–26. https://doi.org/10.1016/j.solener.2021.08.077.
12. RVO (2022). Protocol Monitoring Duurzame Energie. Herziening. https://www.rvo.nl/sites/default/files/2022-07/Protocol-Monitoring-Hernieuwbare%20Energie.pdf.
13. PVoutput. https://pvoutput.org/help/overview.html#overview
14. European Commission. PVGIS Photovoltaic Geographical Information System, https://joint-research-centre.ec.europa.eu/pvgis-photovoltaic-geographical-information-system_en
15. Lindig, S., Gallmetzer, S., Herz, M., Louwen, A., Koumpli, E., Sebastian Enriquez Paez, P., and Moser, D. (2022). Towards the development of an optimized Decision Support System for the PV industry: A comprehensive statistical and economical assessment of over 35,000 O&M tickets. Progress in Photovoltaics. *30*. https://doi.org/10.1002/pip.3637.
16. Meng, B., Loonen, R., and Hensen, J. (2020). Data-driven inference of unknown tilt and azimuth of distributed PV systems. Sol. Energy *211*, 418–432. https://doi.org/10.1016/j.solener.2020.09.077.