Danner and de Meer *Energy Informatics* 2021, **4**(Suppl 3):14
https://doi.org/10.1186/s42162-021-00176-2

Energy Informatics

RESEARCH Open Access

Location and solar system parameter extraction from power measurement time series

Philipp Danner* and Hermann de Meer

From The 10th DACH+ Conference on Energy Informatics
Virtual. 13-17 September 2021

*Correspondence:
philipp.danner@uni-passau.de
University of Passau, Innstraße 41, 94032 Passau, Germany

Abstract

Photovoltaic (PV) systems are considered an important pillar in the energy transition because they are usually located near the consumers. In order to provide accurate PV system models, e.g. for microgrid simulation or hybrid-physical forecast models, it is of high importance to know the underlying PV system parameters, such as location, panel orientation and peak power. In most open PV generation databases, these parameters are missing or are inaccurate.

In this paper, we present a framework based on particle swarm optimisation and the PVWatts model to estimate PV system parameters using only power feed-in measurements and satellite-based ERA5 climate reanalysis data. Our sensitivity analysis points out the most relevant PV system parameters, which are panel and inverter peak power, panel orientation, system location and a small but not negligible influence of ambient temperature and albedo. The detailed evaluation on one exemplary PV system shows an acceptable accuracy in panel azimuth and tilt for the use in microgrid PV system simulation. The extracted location has less than 25 km of positioning error in the best case, which is more than satisfying with respect to the underlying data resolution of the ERA5 dataset. Similar results are observed for 10 systems in Europe and the USA.

Keywords: PV model, Parameter estimation, Particle swarm optimization

Introduction

Due to the transition of the power grid towards clean energy, an increasing penetration of distributed renewable energy sources, mainly Photovoltaic (PV) systems at rooftops, have been observed. On the one hand, such a distributed power generation puts additional pressure on the power grid when it comes to grid management (grid limits, voltage stability or other ancillary services), but on the other hand, the reduced distance between power generation and consumption provides huge potential for regional consumption optimisation.

The trend of emerging energy communities and microgrids, as well as forming of virtual power plants for participating in ancillary service markets, such as primary frequency response, requires accurate power generation models for both simulation of different scenarios and improved forecasting with physical-hybrid models. More particular, PV generation models typically require essential system parameters, such as the panel azimuth, panel tilt, installation location, as well as PV cell and inverter electrical behaviour (e.g., influence of ambient temperature, efficiency and rated power limits).

Some of those PV system parameters are available in power plant databases such as PVOutput.org or national registration databases. However, the data quality is mixed because these values are sometimes user generated and thus suffer from incorrect reporting. Haghdadi et al. have shown that the panel tilt for 10% out of 5000 samples from pvoutput.org is missing and demonstrated to be wrong for 32% of the cases (Haghdadi et al. 2017). In addition, these databases often suffer from inaccurate value resolution (e.g., 45° interval at panel azimuth by defining only the compass direction) or have been identified as default values (e.g. 0° or 1° tilt angles) (Killinger et al. 2018). However, the generated power output of a PV system is usually measured for remuneration purpose using smart meters or is monitored via values from the inverters. This data could be used to automatically determine or validate given PV system parameters and gain a more accurate PV model for microgrid simulation or forecasting. Locating energy data might also cause privacy issues as shown by Chen et al. (2016); Chen and Irwin (2017a).

In order to analyse the feasibility of such an automatic PV system parameter estimation tool, we consider the following research question: *What are the most relevant parameters of an entire PV system and how accurate can these parameters be estimated by only considering historical time series feed-in power measurements and globally available satellite weather data?*

In order to answer this question, we contribute with the following points:

- Performing a variance-based global sensitivity analysis on the National Renewable Energy Laboratory NREL PVWatts model in order to reduce the search space dimension and extract the most relevant model parameters, which are: peak power, panel azimuth and panel tilt, as well as location.
- Describing and implementing a Particle Swarm Optimization (PSO)-based framework to estimate PV system parameters based on the NREL PVWatts model (Dobos 2014) and satellite-based ERA5 climate reanalysis data (Hersbach et al. 2020), which is available for most parts of the globe.
- Evaluating the applicability of the presented approach with a detailed numerical experiment on one exemplary PV system showing reasonable accuracy in peak power, location, as well as panel azimuth and tilt. Tests are repeated on 10 systems in Europe and the USA.

In the following "Related work" section, an overview on related literature about PV parameter estimation is provided. In the "Methodology" section, an optimisation problem of finding relevant PV system parameters based on the PVWatts model using ERA5 reanalysis data and different error metrics is defined. In addition, PSO as meta-heuristic solver is discussed. In "Experiments and discussion" section, the proposed method is tested on measured data and the accuracy of the estimated parameters is presented.

Finally, the work is concluded and an outlook to potential future work is provided in the "Conclusion" section.

Related work

In literature, PV parameter estimation has been performed on *cell/panel level* and on *system level*. A structured summary of related work is listed in Table 1.

On *cell level*, most authors focus on equivalent electric circuit models such as the Single Diode Model (SDM) (da Costa et al. 2010; Soon and Low 2012; Ma et al. 2013; Silva et al.

Table 1 Overview on PV parameter estimation in literature

Ref.	Model	Evaluation Data	Estimation Method	Fitness metric
da Costa et al. (2010); Soon and Low (2012); Ma et al. (2013); Silva et al. (2016); Kang et al. (2018); Jadli et al. (2018); Dali et al. (2015); Mughal et al. (2017); Oliva et al. (2017); Oliva et al. (2017); Chen et al. (2019)	SDM, DDM	mainly IV curve measurements under STC or varying conditions; sometimes extracted from data sheets	meta-heuristics like variants of PSO (Soon and Low 2012; Dali et al. 2015; Mughal et al. 2017), GA (Dali et al. 2015), differential evolution algorithm (da Costa et al. 2010), Cuckoo Search (Ma et al. 2013; Kang et al. 2018), SCA (Chen et al. 2019), Artificial Bee Colonies (Oliva et al. 2017; Oliva et al. 2017) or SA (Mughal et al. 2017; Jadli et al. 2018)	MSE, (normalised) RMSE, MPE, MAE, IAE
(Ruelle et al. 2016)	SAPM and SDM	Generation data from over 40,000 systems	direct search method	normalized MAE
(Haghdadi et al. 2017)	NREL PVWatts	extracted clear sky data	least squares method	-
(Saint-Drenan et al. 2015)	LUT for irradiance and power	power measurements of two PV systems	multiple steps	problem specific
(Williams et al. 2012)	astronomical approach	data of 135 PV systems	calculating longitude from biased peak production	-
(Mason et al. 2020)	DNN	evaluated on simulated data only	DNN	MAE
(Meng et al. 2020)	normalised POA irradiance	simulated data and 13 PV system power measurements	curve fitting on best clear sky day per month	RMSE
(Chen et al. 2016)	sun position algorithm	14 PV systems with almost optimal orientation and tilt	binary search using pre-processed sunrise/sunset and daylength	specific tightest bound fit
(Chen and Irwin 2017b)	sun position algorithm and weather data at a specific location	100 buildings with net load measurements	iterative, multi-step binary search	-
(Chen and Irwin 2017a)	sun position algorithm and weather signature	10 solar sites	(1) correlation clustering to find rough location, (2) weighted midpoint within the cluster	Pearson Correlation
this work	NREL PVWatts	ERA5 weather data/clear sky irradiance model, power generation time series	PSO	RMSE, MAE, MAD and IQR filtered RMSE and MAE

2016; Kang et al. 2018; Jadli et al. 2018), Double Diode Model (DDM) (Dali et al. 2015) or they compare both (Mughal et al. 2017; Oliva et al. 2017; Oliva et al. 2017; Chen et al. 2019). The SDM and the more complex DDM are used to model the current-voltage (IV) output characteristic of single PV cells or panels with relation to external influence of irradiation and temperature. This is achieved by stating an equivalent electric circuit of the cell with at least a series and a shunt resistance, diode reverse saturation current and a diode ideality factor (Gray 2011). Reasonable results, based on IV-curves under Standard Test Condition (STC) extracted from data sheets or measurement series under varying environment conditions, emphasise the usability of different meta-heuristics to solve the non-linear optimisation problem for parameter estimation. Most authors use Root Mean Square Error (RMSE) or Mean Square Error (MSE), as well as, Mean Absolute Error (MAE) metrics to evaluate the fitness of a sample solution. Equivalent electrical circuit models are limited to the characteristics of the PV cell or panel and require precise IV measurements for parameterisation. As, on the one hand, cell IV measurements are usually not available in smart meter data, and on the other hand modeling the detailed characteristics of PV cells is only one part of a whole PV system (location, panel tilt and orientation missing), pure equivalent electrical circuit models are not considered further.

Ruelle et al. (2016) propose to estimate PV system parameters (panel orientation and peak power) using a direct search method to minimise the normalised MAE between PV system simulation data and measured data. In order to avoid a local minimum, the initial estimate is settled to the best variable set out of 100 initial samples. Their simulation model is based on a SDM as part of the Sandia Array Performance Model (SAPM) (thus including Plane of Array (POA) irradiance calculations), which is parameterised with assumed electrical variables. The impact of these variables have not been further analysed and the location of the system is known. Performance improvements are achieved by filtering out cloudy days and shaded hours.

Saint-Drenan et al. (2015) developed an algorithm to estimate the panel orientation by using PV power and meteorological data measurements from a known location. Their PV model has three variables: panel tilt, azimuth and angular loss coefficient, with which the effective irradiance is calculated and the simulated power output is fetched from a Look-up Table (LUT) that takes irradiance and air temperature as inputs. The PV system variables with the maximum likelihood that simulated power matches measured power are accepted as best estimation.

Mason et al. (2020) present a Deep Neural Network (DNN) approach, which extracts PV panel tilt and azimuth from net load metering data. For that, they identified relevant features that have a relation with PV panel tilt and azimuth. Their model is trained on simulated PV data at one specific location that has been combined with customer load profiles. Their approach has been evaluated at one known location on synthetic data.

Meng et al. (2020) propose a data-driven method to parameterize panel azimuth and tilt based on the normalised shape of one clear sky day per month. The best fitting parameter sets (lowest RMSE between the POA irradiance and normalised measurements) are overlapped in order to infer the final estimate. They validated their method using simulated and real PV power measurements. Their curve fitting method requires Global Horizontal Irradiation (GHI) data, which has been taken from ground measurement stations nearby or satellites with increased estimation error. The method has not been designed to find the location of the PV system, but worked comparable well on 15 min, 30 min and 60 min data resolution.

Haghdadi et al. (2017) presented a two-step estimation of location and panel orientation. First, the longitude, which is stated to be independent from the other variables, is extracted from the position of the solar noon. A similar approach is also described in Williams et al. (2012). Second, latitude, panel azimuth and tilt are estimated using least square method to fit the variables of a simulation model (using NREL PVWatts model) to the measured power. This has been performed on clear sky days only, which have been identified by power output fluctuations and by fitting a 3-dimensional surface to the output data. Three extensive case studies provided good results for panel tilt and azimuth (MAE of 2.75° and 5.85°), however the location mismatch is quite high on latitude (4.08° $\sim$225 km), which might be improved using weather data.

Chen et al. (2016) describe a method to infer latitude and longitude independently. First, peak power production per day is fitted on an equation of time to compensate the difference between apparent solar time (actual sun movement) and mean solar time (solar noon 24hours apart). Longitude is then inferred from the extracted solar noon using binary search on a non-reversible sunset/sunrise calculation algorithm. Second, latitude - as function of daylength - is determined by extracting the average daylength within a year. Their approach requires high-resolution solar power measurement data (smallest possible area for minute resolution data is 28 km radius) and their prototypical evaluation focuses on almost south facing systems in the Northern Hemisphere. However, the impact of panel tilt and orientation has been shown as high influence factor for finding system's location but has been shifted to future work.

In a further work, Chen et al. 2017b iteratively apply a multi-step binary search in order to fit panel sizing (first step), orientation (second step) and tilt (third step) of a clear sky generation model to the daily maximum power generation of pre-processed, hourly smart meter net load measurements. The starting parameters is set to be the optimal panel orientation. The focus of that work is on disaggregating net load measurements into consumption and solar generation at a known location only.

For locating different types of energy data (consumption, wind and solar), Chen et al. use a weather signature, based on temperature, wind speed and cloud cover from ground weather stations of different locations (Chen and Irwin 2017a). In order to reduce the search space, daily correlation for initial filtering (k-means clustering) of the big weather database is used, before extracting the weighted (correlation) midpoint of locations in the cluster, based on a hourly analysis. To interpret and compare the weather signature with the solar generation data, a physical model with roughly estimated parameters using a combined approach from Chen et al. (2016) and Chen and Irwin (2017b) is used. Unfortunately, the accuracy of the system parameters (panel peak power, tilt and azimuth) have not been commented and the granularity, as well as the distribution of the underlying ground weather stations does not get clear. Satellite data might provide a more uniform coverage despite a potential lower spatial resolution.

Comparing the different approaches in literature, it can be seen that most related work does not consider PV system location, panel orientation (azimuth and tilt) and PV system component sizing (panel and inverter peak power) as unknown parameters simultaneously (exceptions in Haghdadi et al. (2017); Chen and Irwin (2017a)) and thus are limited to their specific use case. For search space reduction and avoidance of local optima, physically ideal starting parameters (Chen and Irwin 2017b), the best of set of initial samples (Ruelle et al. 2016) and filtering with lower resolution data (Chen and Irwin 2017a) has been used. As all parameters effect the power generation collectively, we consider estimating all parameter at once. We thus propose a simulation based PSO that uses ERA5 reanalysis data in order to improve location estimation. The usage of PSO is motivated further in "Particle swarm optimisation" section.

Methodology

For the proposed PV system parameter estimation framework, first the PV model that calculates the power output from relevant input variables is detailed. Afterwards, the objective function using different error metrics for comparison and the solving method is explained.

PV model

Unlike the equivalent electric circuit models, the PVWatts model directly estimates the power output of the PV panel. For evaluation, the more commonly available PV panel peak power can be compared to the estimated parameter from the PVWatts DC model; Measured IV curves under various environment conditions are not required. The PVWatts model still encompasses basic physical relation of input and output by relying on meaningful parameters, whereas other models such as the SAPM are mainly fitted with empirical measurements (King et al. 2004). We thus consider a PV system model chain that is mainly based on the PVWatts model and detailed in the following.

The model in this work is limited to the commonly used monofacial PV panels, also PV panel axis tracker are excluded. The focus is on one-sided PV panel orientation, however, east-west panel combinations are working as well, as shown later.

Inverter model

The PVWatts model includes multiple sub models. One of these is the inverter model that integrates the inverter efficiency η by defining the conversion from DC power P_{dc} to AC power P_{ac} and limiting to the inverter nominal power rating P_{ac0}, as shown in (1).

$$P_{ac} = min(\eta \cdot P_{dc}, P_{ac0}) \tag{1}$$

$$\eta = \frac{\eta_{nom}}{\eta_{ref}}(-0.0162\zeta - \frac{-0.0059}{\zeta} + 0.9858) \tag{2}$$

$$\text{where } \zeta = \frac{P_{dc}}{P_{dc\_limit}} \text{ and } P_{dc\_limit} = \frac{P_{ac0}}{\eta_{nom}}$$

The constant values in Eq. (2) have been extracted from an analysis of the California Energy Commission (CEC) inverter performance database and are part of the PVWatts model. The reference inverter efficiency η_{ref} from the actual most typical inverter is 0.9637, the default nominal efficiency η_{nom} is set to the proposed value of 0.96 (Dobos 2014). These assumptions represent a typical inverter efficiency, however the overall power output is mainly influenced by panel and inverter power (depending on sizing and irradiance) and panel orientation as shown in the sensitivity analysis later.

Cell model

The DC power P_{dc} of a PV panel is calculated with the PVWatts DC model as shown in Eq. (3). In this model, the panel efficiency is assumed to decrease at a linear rate with increasing temperature. This is governed by the temperature coefficient τ, which depends on the module type.

$$P_{dc} = \frac{I_{tr}}{I_{tr0}} P_{dc0}(1 + \tau(T_{cell} - T_{ref})) \tag{3}$$

The parameters are defined as following:

- I_{tr} represents the effectively transmitted plane of array irradiance on the PV cell in units of W/m^2. The angle of incidence losses need to be applied beforehand (detailed in "Irradiance" section).
- I_{tr0} is the reference irradiation, which is 1000 W/m^2.
- T_{cell} is the calculated PV cell temperature in °C.
- P_{dc0} is the nominal DC power of the PV module at reference irradiation I_{tr0} and cell reference temperature T_{ref}.
- τ represents the temperature coefficient in units of $1/°C$. This value is typically between -0.002 and -0.005 per °C.
- T_{ref} is the cell reference temperature, which is defined to be $25°C$.

From the DC model parameters, the temperature coefficient τ and the nominal DC power P_{dc} remain as variables in the optimization problem. The other parameters are calculated as defined in the following.

Temperature model

Instead of the temperature model from Fuentes 1987, that has been developed in the 1980s and is used in PVWatts, we calculate the cell temperature with the SAPM (King et al. 2004). This is because the early model has proven to be unnecessarily complex and thus is leading to integration issues of new module technologies. In addition, the SAPM uses less parameters by providing a temperature accuracy of $\pm 5°C$ resulting in an uncertainty of less than 3% of the power output, according to King et al. (2004).

The cell temperature is calculated in the SAPM by using the ambient dry bulb temperature T_a in $°C$, the plane of array effective irradiance I_{tr} in W/m^2 and the wind speed WS in m/s at a height of 10 meters, in order to include heating effects from the sun and cooling effects from the wind. The cell temperature is calculated in (4) and the back-surface module temperature T_m is defined in (5).

$$T_{cell} = T_m + \frac{I_{tr}}{I_{tr0}} \Delta T \tag{4}$$

$$T_m = I_{tr} \cdot e^{a+b \cdot WS} + T_{amb} \tag{5}$$

The ambient temperature T_{amb}, as well as the wind speed at 10m height are extracted from the ERA5 reanalysis data in this work and they depend on the PV system location. The parameter sets of a (coefficient for module temperature upper limit at low wind speeds and high solar irradiance), b (coefficient for the rate at which module temperature drops as wind speed increases) and ΔT represent the thermodynamics of the panels and their installation. Some empirically determined examples are shown in Table 2.

Irradiance

The effectively transmitted POA irradiance I_{tr} is a linear combination of the direct POA irradiance I_{beam}, the sky diffuse irradiance $I_{diffuse}$ and the ground reflected irradiance $I_{reflected}$, defined in (6). The calculation of these parts, Eqs. (7) and (10), is based on GHI,

Table 2 Typical temperature model parameter sets for different panel types and installations (King et al. 2004)

Module	Mounting	*a*	*b*	ΔT
glass/glass	open rack	-3.47	-0.0594	3
glass/glass	close roof	-2.98	-0.0471	1
glass/polymer	open rack	-3.56	-0.075	3
glass/polymer	insulated back	-2.81	-0.0455	0

Diffuse Horizontal Irradiation (DHI) and Direct Normal Irradiation (DNI). This weather-depending irradiance values can be gathered from the ERA5 climate reanalysis data, which requires location and time as input. Because the ERA5 data only provides GHI and DHI, the missing DNI can be calculated with Eq. (8). The angle of incidence α for a given, fixed PV panel with a surface tilt β, a surface azimuth γ is calculated from the solar azimuth γ_{sun}, and solar zenith θ_{sun} angles, as described in Eq. (9).

$$I_{tr} = I_{beam} + I_{diffuse} + I_{reflected} \tag{6}$$

$$I_{beam} = DNI \cdot cos(\alpha) \tag{7}$$

$$DNI = \frac{GHI - DHI}{cos(\theta_{sun})} \tag{8}$$

$$\alpha = cos^{-1}(sin(\theta_{sun}) \cdot cos(\gamma - \gamma_{sun}) \cdot sin(\beta) + cos(\theta_{sun}) \cdot cos(\beta)) \tag{9}$$

Panel surface tilt β and panel surface azimuth γ (the panel orientation) remain as problem variables, whereas the solar azimuth γ_{sun}, and solar zenith θ_{sun} are calculated from the position of the sun at a certain time using the NREL Solar Position Algorithm (SPA) (Reda and Andreas 2004; 2007). The additional parameters of the SPA are location (latitude and longitude), elevation (=altitude; can be derived from latitude and longitude with an elevation map), as well as the yearly average air temperature (assumed to be 12°C) and pressure (calculated from altitude) for atmospheric refraction correction.

The angle of incidence correction within PVWatts V5 to adjust the direct beam irradiance in order to account for reflection losses in the glass surface of the PV panel is not used in this work. This is because the difference in power output for standard glass modules is negligible according to Dobos (2014). The additional parameters would complicate the model by providing only minor impact.

For calculating the diffuse irradiance, multiple methods have been proposed. Loutzenhiser et al. evaluated seven models with experimental data on vertical building facades and found out that the Perez (1990) formulation provides the most accurate results for their building heating energy scenario (Loutzenhiser et al. 2007). For this work, however, we use the Hay-Davies model (Hay and Davies 1980) due to the following reasons: On the one hand, the Perez model is more complex and is based on empirically derived coefficients (Perez et al. 1990). On the other hand, the accuracy of the Hay-Davies model still has acceptable accuracy in irradiance on vertical plane (1.1% mean error compared to 0.5% mean error of Perez model at peak times (Loutzenhiser et al. 2007)). In addition, the impact of diffuse irradiance on the overall POA irradiance is even lower in non-vertical scenarios, such as it is the case with roof-top PV systems, which are usually oriented towards the sun and thus are dominated by the direct beam irradiation. The Hay-Davies model is composed of an isotropic and circumsolar component, and horizon brightening is neglected.

$$I_{diffuse} = DHI \cdot (A \cdot R_b + (1 - A) \cdot (\frac{1 + \cos\beta}{2}))$$
$$\text{where the anisotropy index } A = \frac{DNI}{I_{ET}} \text{and } R_b = \frac{cos(\alpha)}{cos(\theta_{sun})} \tag{10}$$

The radiation on the earth's atmosphere varies slightly over the year, thus this extraterrestrial radiation I_{ET} is calculated with a yearly varying term in order to account for the eccentricity of the Earth's orbit around the sun. We use the Spencer model that is defined through Fourier series (Spencer 1971) with x as the day angle for the earth's orbit around the sun in Eq. (11).

$$\begin{aligned} I_{ET} = I_{SC} \cdot [\, 1.00011 + 0.034221 \cdot cos(x) + 0.00128 \cdot sin(x) \\ +0.000719 \cdot cos(2x) + 7.7e - 05 \cdot sin(2x)] \\ \text{and the solar constant } I_{SC} = 1366.1 W/m^2 \end{aligned} \tag{11}$$

The ground reflected irradiance $I_{reflected}$ represents the reflected irradiance, which usually distinguishes between different types of ground by using the *albedo factor*. Although the albedo depends on the location and changes with seasonal effects such as snow or rain, an albedo of 0.25 is assumed in this paper. This value is a compromise of typical reflection for onshore surfaces (0.1 - 0.4) and the average albedo from Earth (0.34), which roughly represents the reflection of grass (McEvoy et al. 2012). The effect of different albedo factors on the total irradiance with an economically optimised PV system in central Europe (panel tilt of 36°) is less than 2% (excluding snow condition) and can thus be neglected for the purpose of this work. The reflected irradiance is calculated as defined in Eq. (12) extracted from Loutzenhiser et al. (2007).

$$I_{reflected} = GHI \cdot albedo \cdot (\frac{1 + \cos\beta}{2}) \tag{12}$$

Model parameter discussion

The overall PV system model is derived by combining the individual models defined in Eqs. (1) to (12) and the NREL SPA for calculating the position of the sun (sun azimuth and sun zenith).

Irradiation (GHI, DHI) is obtained from the ERA5 reanalysis data, which requires the location of the PV system and the considered time as input. Environment condition, like the ambient temperature, and wind speed at 10 m height can be extracted from the ERA5 data as well. In order to further reduce the amount of parameters, altitude is calculated from SRTM 90m Digital Elevation Database v4.1 (Reuter et al. 2007; Jarvis et al. 2008), which is based on satellite data (by the NASA). Thus, the remaining PV system model parameters that are considered as decision variables are listed in Table 3.

Sensitivity analysis

In order to reduce the number of model parameters, a variance-based sensitivity analysis according to Sobol 2001, using the samples improvement by Saltelli (2002); Saltelli et al. (2010), is performed. Instead of the measured values from the ERA5 data set, GHI and DHI are calculated using the Ineichen clear sky model and thus depend on the PV system location. Wind speed (WS) at 10m height and ambient temperature (T_{amb}), as well as

Table 3 List of model parameters

Parameter	Description	Realistic range
P_{ac0}	Inverter nominal power	> 0.0
P_{dc}	Cell peak power	> 0.0
τ	Cell temperature coefficient	-0.002 .. -0.005
β	Panel tilt	0° (upright) .. 90° (horizon)
γ	Panel azimuth (0° facing north)	0° .. 360°
lat	Latitude of the PV system	90°N .. 0 .. 90°S
lon	Longitude of the PV system	-180° .. +180°
a	Coefficient for module temperature upper limit at low wind speed and high solar irradiance	-3.56 .. -2.81 (from Table 2)
b	Coefficient for module temperature drop at increasing wind speed	-0.075 .. -0.0455 (from Table 2)
ΔT	Module temperature difference for coefficient a and b	0 .. 3 (from Table 2)

albedo and inverter efficiency η_{nom} are added as parameters in the sensitivity analysis to measure their influence on the power output. For the model sensitivity analysis only, the time has been fixed to the 21th June, 12:00 UTC (day with most hours of daylight in Northern Hemisphere) in the year 2020 (relevant for the extraterrestrial radiation). Changing the date or time (excluding night times) does not change the sensitivity of the considered parameters significantly. The bounds of the considered parameters are listed in Table 4 (latitude and longitude roughly covering Europe) and a sample size of 10000 is used.

From the analysis of the first-order and total-order indices (compare Fig. 1), it can be concluded that the main parameters with the highest influence on the output power are, as expected, the inverter and panel peak power, the panel orientation (β and γ) as well as the location of the PV system. Ambient air temperature (T_{amb}) and the albedo factor still have a small but measurable impact on the output power.

The temperature coefficient τ, as well as the parameters for the PV module heating model a, b and ΔT and inverter efficiency eta_{nom} have negligible impact. These findings are inline with Hansen et al. 2013, who performed a detailed sensitivity analysis on the individual models. They observed a dominating contribution to the uncertainty in daily energy by the POA irradiance and the effective irradiance models, which depend on location, and panel orientation. Thus, we use the glass/glass close roof configuration from Table 2 and $\tau = -0.003$ in order to reduce the number of dimensions in the search space.

Table 4 Parameter bounds for sensitivity analysis

Parameter	Range	Parameter	Range
P_{ac0}	0 .. 20000 W	lat	35° N .. 60° N
P_{dc}	0 .. 20000 W	lon	-10° .. +30°
τ	-0.002 .. -0.005	β	0° .. 90°
a	-3.56 .. -2.81	γ	0° .. 360°
b	-0.075 .. -0.0455	T_{amb}	-20° .. 50° C
ΔT	0 .. 3	WS	0 .. 30 $\frac{m}{s}$
η_{nom}	0.92 .. 0.99	*albedo*	0.04 (asphalt) .. 0.8 (snow)

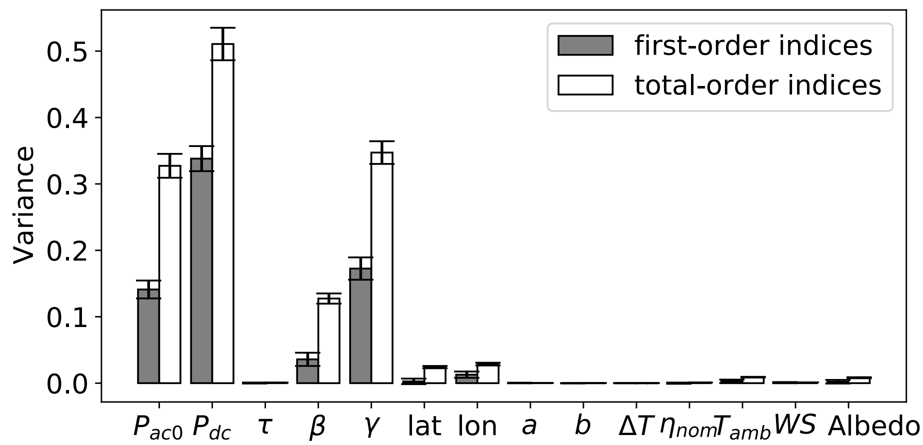

Fig. 1 First-order and total-order indices

Problem formulation

The discussed model parameters are subject to an optimisation problem to minimise the error between the calculated P^t_{ac}, according to the equations defined above, and the measured AC power $P^t_{measured}$ at each timestamp t. The error metric thus defines the objective function (sometimes called fitness function in the context of PSO).

We consider commonly used RMSE in Eq. (13), which tends to emphasise the effect of outliers, and MAE in Eq. (14), which is more robust to outliers and thus better represents the average characteristics of a potential solution. In addition, these metrics are compared to MAD in Eq. (15) and IQR filtered RMSE and MAE metrics. The latter three completely avoid the effect of outliers as only the better half of the error series is considered. According to Stein et al. 2010, satellite irradiance data provide a similar accuracy compared to ground measurements considering the mean error. However, the standard deviation is larger and thus filtering out outliers in these metrics seems to be a suitable option.

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(P^t_{ac} - P^t_{measured})^2} \tag{13}$$

$$MAE = \frac{1}{N}\sum_{t=1}^{N}\left|P^t_{ac} - P^t_{measured}\right| \tag{14}$$

$$MAD = median(\left|P^t_{ac} - P^t_{measured}\right|) \tag{15}$$

Particle swarm optimisation

The Particle Swarm Optimization (PSO) is an optimisation technique that emulates the social behaviour of biological organisms, such as bird or fish swarms. First, a set of particles, referred to as the swarm, is randomly initialised in the n-dimensional search-space (evenly distributed). Each particle represents one candidate solution. In order to find the optima, the particles then move around the search-space using historical position and velocity of themselves and their neighbours.

The original PSO algorithm is attributed to Kennedy and Eberhart (1995); Shi and Eberhart (1998) and has been developed to solve non-linear equations. Over time, many variations (e.g., different topology, search-space characteristics or constraints) have been used in research in order to solve a variety of problems. For this work, we use the classical star-topology, in which each particle is attracted by the best performing particle of the whole swarm, which is assumed to be near the global optimum.

The position of the particle x_i at the current step s is updated with the computed velocity v_i at $s+1$, as in Eq. (16). The velocity of a particle is calculated as a linear combination of: (1) its own damped previous velocity (parameter w for inertia), (2) its deviation to its p_i neighbourhood (parameter c_1 for cognitive behaviour), and (3) its deviation to the best particle of the swarm p_g (parameter c_2 for social behaviour), as defined in Eq. (17). The two parameters c_1 and c_2 define if the swarm is more explorative (following personal best) or exploitative (following swarm's global best). The independent random numbers r_1 and r_2 in the range of $[0, 1]$ introduce a certain randomness into the velocity (next iteration), more details can be found in Shi and Eberhart (1998).

$$x_i(s+1) = x_i(s) + v_i(s+1) \tag{16}$$

$$v_i(s+1) = w \cdot v_i(s) + c_1 r_1 [p_i - x_i(s)] + c_2 r_2 [p_g - x_i(s)] \tag{17}$$

One of the main advantages of the PSO algorithm is that it does not use the gradient of the function. Thus, it is not required to have an objective function that is differentiable. As we obtain irradiance from the ERA5 dataset based on the location parameters, our problem cannot be differentiated. In more general, PSO can be classified as metaheuristic as it makes few (in our case decision variable boundaries) or no assumptions about the underlying problem to be optimised. Compared to variants of the population based Genetic Algorithm, PSO provides the same quality of solution while reducing the computational effort (Hassan et al. 2005). As panel tilt β, cell peak power P_{dc} and efficiency parameters (roughly scaling the output power) have similar effects on the overall generation (Chen and Irwin 2017b), it appears that searching the entire parameter search space is required in order to avoid local optima. In addition, similar weather conditions in different areas of the ERA5 dataset could lead to local optima of the parameter set. PSO is a suitable method to avoid local optima, as lots of sample solutions, spread in the search space, are compared at each step and the overall solution is steadily directed towards the best known optima.

Experiments and discussion

In the following, the proposed method is evaluated with an exemplary PV system, for which all relevant parameters are known. First, the used data is described and second the results are presented and discussed.

PV system data and pre-processing

The exemplary PV system DC peak power is rated with 11.55 kW and the panels are connected to two inverters with each 4.6 kW nominal AC output (5.06 kW maximum). The inverter efficiency ($\eta_{nom} = 95.9\%$ extracted from the curve in the datasheet according to weighted CEC definition) roughly matches the typical values extracted from the CEC inverter database ($\eta_{nom} = 96\%$) quite well. The PV system is installed on a roof top with a roof pitch (panel tilt) of 23°; Panel azimuth is 195.45° (slight south west direction). The PV system is 13 years old, thus a degraded panel peak power is expected. There is a minor shadow effect in the morning, which makes the PV system a good candidate for a detailed analysis as perfect systems are rare.

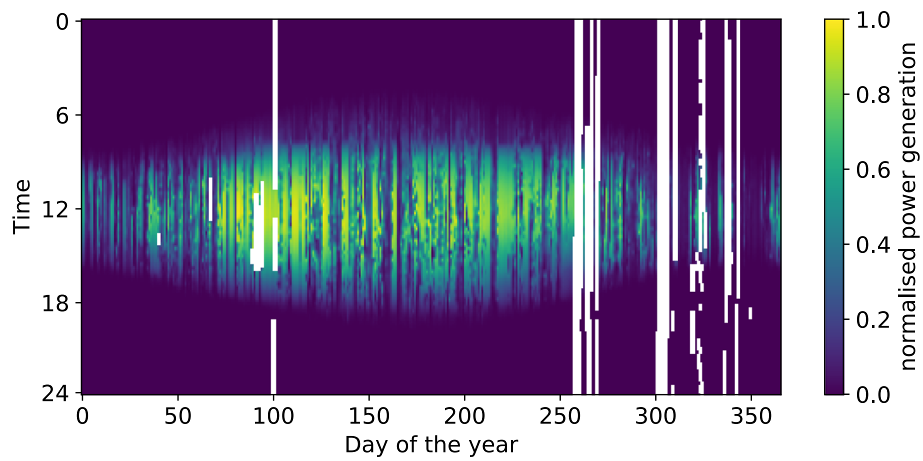

Fig. 2 Normalised power generation in 15 min (white represents missing data)

The power output of the tested PV system has been collected at the digital, calibrated energy meter and is averaged to 1 hour mean values for the year 2020 to match the temporal resolution of the ERA5 reanalysis data. Due to some measurement issues, 374 hours are missing and have been excluded from the optimisation. The measured power output and data gaps are visualised with a quarter hourly resolution in Fig. 2.

In order to focus the error metric on productive time and improve calculation performance, night conditions (measured power smaller than 100 W) have been filtered out. This is especially important for the metrics that focus on the better half of the error series (MAD, IQR filtered RMSE and MAE), which is the case at night condition when the error is almost to zero.

The considered period has been limited from beginning of April until end of October in order to avoid the influence of snow and tree shadows (due to low sun zenith) as good as possible. This assumption is backed by observations but can also be observed from the monthly Pearson correlation between GHI (from ERA5 weather data at the location of the system) and the power measurements (see Fig. 3). Winter and late autumn season seem to have a higher mismatch induced by snow covered PV panels and higher impact of shadows due to lower sun zenith.

Solar irradiance, wind speed and ambient temperature are extracted from the ERA5 reanalysis data based on the considered location. ERA5 data (reanalysis-era5-single-levels) has been prepared with a resolution of 0.25° for both latitude and longitude, which is around 16 - 20 km on latitude and 28 km on longitude in the tested part of Europe (same area as in the sensitivity analysis, Table 4). For a more precise calculation, the ERA5 data is linearly interpolated with the given location at each step.

When running the simulation model with ERA5 data using the actual system parameters for the whole year, a MAE of 567.42 W can be found. The deviation between the measured and simulated PV system can be explained with observed minor shadow effect in the morning, degraded panel peak power, model inaccuracy and mainly by the inaccuracy of the satellite weather data (compare sample period in Fig. 4).

Time series data (measured power and ERA5 data) is shifted by 30 minutes in order to calculate the sun position at the halftime of the corresponding mean period.

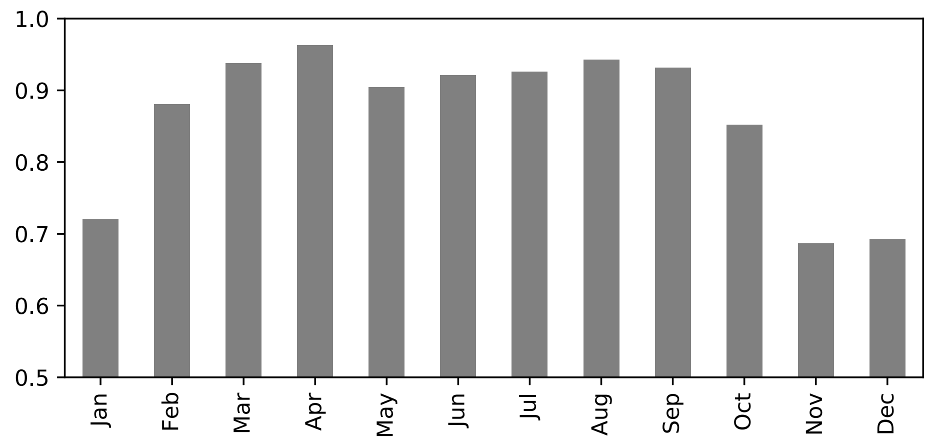

Fig. 3 Monthly Pearson correlation between ERA5 GHI at the system's location and measured power

Results

For the experiment a swarm size of 200 with 400 iterations and $c_1 = 0.7$, $c_2 = 0.3$ and $w = 0.9$ results in a stable solution. The swarm acts more explorative than exploitative and thus finds the global optimum in most of the times. After around 150–200 iterations, the particle velocity of all six parameters converges. This is visualised as velocity history graph, normalised to the parameter's search range, for MAE in Fig. 5. The velocity of the inverter nominal power P_{ac0} increases after a while and stabilises again at the boundary of the search space. Similar convergence behaviour is observed for all tested metrics. The calculations have been repeated 15 times in order to measure the impact of the random swarm initialisation.

For the exemplary system, RMSE and MAE find similar optimum parameter sets in all 15 repetitions as visualised in Fig. 6, whereas MAD and IQR filtered RMSE and MAE metrics converge in slightly different solutions. This is caused by only considering the better half of the error series, which changes in each iteration.

The inverter power P_{ac0} is overestimated independently of the used metric (mostly at upper search space boundary at 20 kW). This is due to minor impact of higher inverter

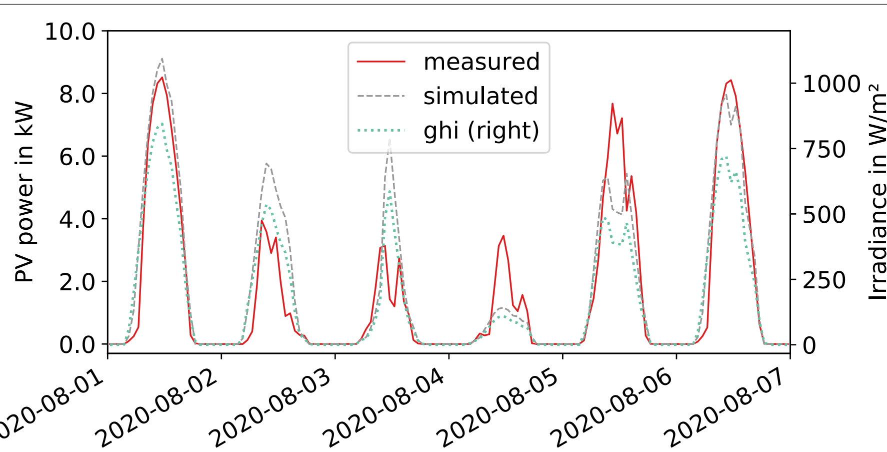

Fig. 4 Model accuracy in a sample period

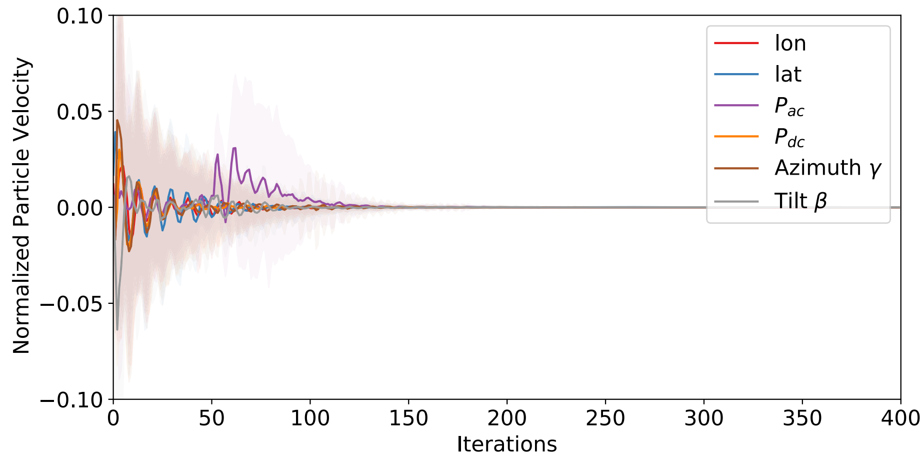

Fig. 5 Mean and standard deviation of velocity history sample run, normalised to the parameter's search range, using MAE metric

power compared to PV panel power (only efficiency). The initial under-sizing of the inverter (10.12 kW maximum inverter power versus 11.55 kW panel peak) has not been detected. This might stem from panel degradation and soiling, which reduced the panel peak power below the maximum inverter power.

The nominal peak power of the panels is underestimated in average between 0.8 and 1.55 kW with different metrics. This can be explained with the initial under-sizing of the inverter (1.43 kW), in addition to panel degradation and soiling effects. Nevertheless, these fitted parameters might better represent the current state of the system compared to the rated power at installation time.

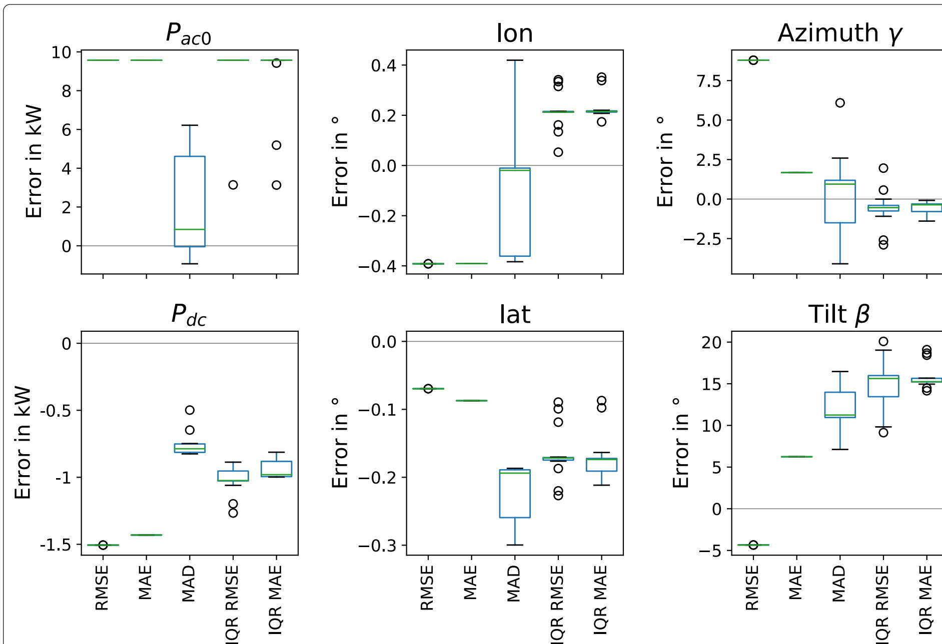

Fig. 6 Mismatch between the estimated and actual parameters (15 repetitions). Parameters search space according to Table 3. Inverter peak power P_{ac0} is mostly optimised to the search space boundary, the remaining parameters are within the search space

Regarding the location parameters latitude and longitude, RMSE and MAE provide a stable solution with a distance of 30 km from the actual location (mainly west, slightly south direction). This deviation could be explained with a regular minor shadow in the morning that shifted the longitude in general to the west (ignoring different weather). MAD and especially IQR filtered metrics halve longitude error, but double the latitude error, resulting in around 25 km distance error (southeast). Using these metrics, the impact of the minor shadow effect in the morning is reduced while the results are not stable in each run.

Panel azimuth γ estimation is comparable with all metrics except RMSE and ranges between +1.68° to - 0.59°. The over-estimation with RMSE (+8.8° error) is assumed to origin in the observed minor shading in the morning and the fact that the RMSE metric emphasise the impact of outliers. Panel tilt β error is quite high in all metrics, which might stem from its smaller influence on the system, as shown in the sensitivity analysis.

Discussion

The estimated PV system parameters should be considered as the best fitting parameters for the simulation model. The ERA5 data is useful for locating the system by incorporating a broad variety of weather situations, however it does not provide an accurate representation of the local situation due to its low spatial resolution.

Regarding the isolated impact of panel tilt and azimuth error in more detail, the calculated annual energy generation differs by in average of 0.35% per degree tilt and 0.08% per degree azimuth in the range of $\pm 15°$ around the actual orientation (compare Fig. 7). The absolute estimation error in azimuth is 1.68° for MAE metric and thus approximately 0.13% error in annual energy generation, which could be considered as negligible. However, the estimation error in tilt (6.24° for the MAE metric) results in a annual energy generation error of 2.184%. Comparing this error with the isolated annual energy generation error for location mismatch, which accounts for around 0.2% error of annual energy, it becomes clear that a better tilt estimation is required.

The method using MAE has also been tested on a east-west sided PV installation. Its location is estimated with a similar accuracy (23.79 km distance error) compared to the single-sided system in the same area. As the weather signature in that region mainly

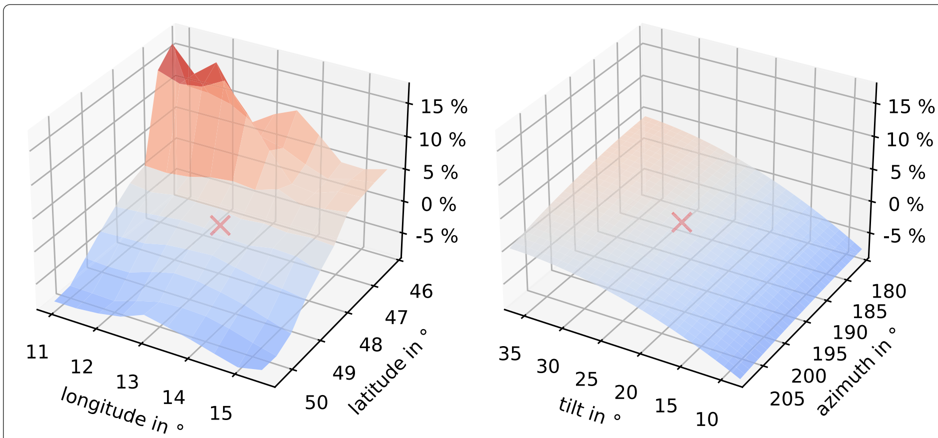

Fig. 7 Impact of azimuth and tilt error on yearly energy generation compared to actual orientation

influences the location estimation of the system, it was even possible to identify a two-sided setup. The angles of the panels (azimuth error of up to 22°) and their peak power (deviation of around 25%), however, are not very accurate.

In addition, we tested 5 PV system installations in California/USA for which the given ZIP code covers the smallest area, which are mainly located in cities. It was possible to allocate the hourly resolution time series of the year 2016 with an error of between 35.74 km and 62.4 km to the centre point of the ZIP code area using the MAE metric. The panel angles are not very accurate, which can be accounted to the shading from nearby buildings in the urban area. This can also be observed when comparing the GHI of clear sky days at the estimated location with the power measurements, where the power is heavily reduced in the morning and in the evening. The tilt angle for one system with flat panels (0° tilt), however, could be identified reliably. Location estimation on 5 further PV system at different locations in Bavaria/Germany is working with a mixed accuracy ranging from 17-90 km. Panel orientation is not documented for these five systems.

When comparing the parameter estimation accuracy with related work - even when using the exact same dataset is not possible - Saint-Drenan et al. 2015 performed better on panel azimuth and tilt error (less than 2° in optimal cases) using satellite irradiance and temperature values. No accuracy on their location estimation was given. The data-driven approach by Meng et al. (2020) achieved an MAE in azimuth 4.5° and tilt 4.3°. When applying our approach with known location using RMSE metric, a tilt error of 0.55° and azimuth error of 7.22° is achieved. The azimuth error is supposed to originate from the minor morning shading of the observed PV system.

Williams et al. 2012 state a longitudinal error of less than 50 miles (around 80 km) using their astronomical approach and one month of data. The panel orientation deviation is found to be $\pm 7°$. Haghdadi et al. 2017 achieved a mean absolute deviation of 0.2° longitude, 4.08° latitude, 2.75° panel tilt and 5.85° azimuth working with clear sky data and a temporal resolution of 5 minutes. Our PSO approach with IQR filtered error metrics outperforms the location estimation (less than 0.3° for both longitude and latitude) using hourly temporal resolution (1.4° longitude MAE for hourly resolution have been achieved in Haghdadi et al. (2017)). Even the other tested 5 PV systems have been located better than 1° for both longitude and latitude. However, our approach lacks accuracy in panel orientation.

Chen et al. 2017a achieved a better accuracy in allocating their PV systems, however, the granularity, as well as distribution of the underlying ground weather stations does not get clear. Satellite data, as used in this work, provide a uniform coverage making our approach more generic and applicable equally almost all over the globe.

Conclusion

For advanced modeling of PV power generation in microgrid simulation scenarios or for improved forecasting with physical-hybrid models, the PV system parameters, such as location, orientation and nominal power are required but not available in all cases.

This paper presents a framework to estimate the most relevant PV system parameters by using power measurements and ERA5 reanalysis weather data in combination with a PVWatts based PV simulation model. The relevant parameters, more specifically longitude, latitude, panel tilt and azimuth, as well as inverter/panel peak power, have been identified with a global sensitivity analysis on the simulation model. As most of the parameters show a dependency on each other, all parameters are optimised at once. This is achieved by minimising the error between the measured and the simulated power output time series using PSO and different error metrics. We compared commonly used MAE and RMSE with median error and IQR filtered metrics, which only consider the lower half of absolute errors and thus ignore outliers. The latter perform slightly better for location estimation but lack accuracy in panel tilt.

We demonstrated with one exemplary PV system and measurements over one year that the location can be estimated with an error of less than 25 km using hourly measurement resolution. This estimation error roughly matches the spatial resolution of the underlying ERA5 data. Regarding the panel orientation, azimuth estimation is acceptable while the tilt angle, which also has a lower sensitivity, remains a point for improvement. The panel peak power is in the expected range of panel degradation of the analysed system. Similar observation can be found with the 10 PV system in Europe and the USA.

In contrast to related work, we can also apply our approach to dual-sided PV installation. As result, it is possible to distinguish between single and dual-sided systems; however, the panel orientation mismatch is higher than on single-sided systems only. The location accuracy is comparable to singled-sided systems.

The presented framework is limited, on the one hand, by accurate weather data (temporal and spatial) and, on the other hand, by an accurate PV model. Localisation using ERA5 weather data has been demonstrated to work quite well with regard to the available resolution of 0.25° on latitude and longitude. However, the panel orientation extraction might be improved by removing the uncertainty of irradiance, e.g., by more focusing on clear sky days. In addition, shadow and snow detection could be integrated to avoid or correct the measurements of shaded periods. The parameter estimation might be improved by finding the best trade-off between filtering inaccurate data (e.g., panel shading) and a suitable metric that devalues outliers, both without losing the general correlation to the ERA5 data for location estimation.

Acknowledgements
We want to thank the contributors to pvLib-python (F Holmgren et al. 2018), PySwarms (James V. Miranda 2018) and SALib (Herman and Usher 2017) for working on scientific open source tools.

About this supplement
This article has been published as part of *Energy Informatics Volume 4 Supplement 3, 2021: Proceedings of the 10th DACH+ Conference on Energy Informatics*. The full contents of the supplement are available online at https://energyinformatics.springeropen.com/articles/supplements/volume-4-supplement-3.

Authors' contributions
PD conducted the main research including related work, concept and implementation. HdM provided research direction, supervision, and helped to write the final version of the paper. Both authors read and approved the final manuscript.

Funding
This project has received partial funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 764090. Publication funding was provided by the German Federal Ministry for Economic Affairs and Energy.

Availability of data and materials
PV system time series data from California/USA with all parameter information have been downloaded from https://www.californiadgstats.ca.gov/downloads/. Additional PV power production time series from Germany have been used from https://openenergy-platform.org/dataedit/view/demand/emsig_energy_data_by_ems. The altitude map can be found at https://srtm.csi.cgiar.org/.
ERA5 data have been obtained form https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels.
Source code is available via request to the authors.

Declarations

Competing interests
The authors declare that they have no competing interests.

Published: 13 September 2021

References

Chen D, Irwin D (2017) Weatherman: Exposing weather-based privacy threats in big energy data. In: 2017 IEEE International Conference on Big Data (Big Data). IEEE. pp 1079–1086. http://ieeexplore.ieee.org/document/8258032/

Chen D, Irwin D (2017) SunDance: Black-box behind-the-meter solar disaggregation. In: Proceedings of the 8th International Conference on Future Energy Systems. ACM. pp 45–55

Chen D, Iyengar S, Irwin D, Shenoy P (2016) SunSpot. In: Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments. ACM, New York, USA. pp 85–94

Chen H, Jiao S, Heidari AA, Wang M, Chen X, Zhao X (2019) An opposition-based sine cosine approach with local search for parameter estimation of photovoltaic models. Energy Convers Manag 195:927–942

da Costa WT, Fardin JF, Simonetti DSL, de VBM Neto L (2010) Identification of photovoltaic model parameters by differential evolution. In: 2010 IEEE International Conference on Industrial Technology. IEEE, Via del Mar, Chile. pp 931–936. http://ieeexplore.ieee.org/document/5472557/

Dali A, Bouharchouche A, Diaf S (2015) Parameter identification of photovoltaic cell/module using genetic algorithm (GA) and particle swarm optimization (PSO). In: 2015 3rd International Conference on Control, Engineering & Information Technology (CEIT). IEEE, Tlemcen, Algeria. pp 1–6. http://ieeexplore.ieee.org/document/7233137/

Dobos AP (2014) PVWatts Version 5 Manual. Technical Report September, National Renewable Energy Laboratory (NREL), Denver West Parkway Golden, CO. http://www.nrel.gov/docs/fy14osti/62641.pdf

F Holmgren W, W Hansen C, A Mikofski M (2018) pvlib python: a python package for modeling solar energy systems. J Open Source Softw 3(29):884

Fuentes MK (1987) A simplified thermal model for flat-plate photovoltaic arrays. Technical report, Sandia National Laboratories, Albuquerque, New Mexico 87185 and Livermore, California 94550

Gray JL (2011) The Physics of the Solar Cell. In: Handbook of Photovoltaic Science and Engineering. John Wiley & Sons, Ltd, Chichester, UK. pp 82–129

Haghdadi N, Copper J, Bruce A, MacGill I (2017) A method to estimate the location and orientation of distributed photovoltaic systems from their generation output data. Renew Energy 108:390–400

Hansen CW, Pohl A, Jordan D (2013) Uncertainty and Sensitivity Analysis for Photovoltaic System Modeling. Technical Report, Sandia National Laboratories, Albuquerque, New Mexico 87185 and Livermore, California 94550

Hassan R, Cohanim B, de Weck O, Venter G (2005) A Comparison of Particle Swarm Optimization and the Genetic Algorithm. In: 46th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference. American Institute of Aeronautics and Astronautics, Reston, Virigina Vol. 13. pp 274–283

Hay JE, Davies JA (1980) Calculations of the solar radiation incident on an inclined surface. In: Proc of First Canadian Solar Radiation Data Workshop Canada: Ministry of Supply and Services. Minister of Supply and Services Canada, Toronto, Ontario, Canada. pp 32–58

Herman J, Usher W (2017) SALib: An open-source Python library for Sensitivity Analysis. J Open Source Softw 2(9):97

Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, Nicolas J, Peubey C, Radu R, Schepers D, Simmons A, Soci C, Abdalla S, Abellan X, Balsamo G, Bechtold P, Biavati G, Bidlot J, Bonavita M, Chiara G, Dahlgren P, Dee D, Diamantakis M, Dragani R, Flemming J, Forbes R, Fuentes M, Geer A, Haimberger L, Healy S, Hogan RJ, Hólm E, Janisková M, Keeley S, Laloyaux P, Lopez P, Lupu C, Radnoti G, Rosnay P, Rozum I, Vamborg F, Villaume S, Thépaut J (2020) The ERA5 global reanalysis. Q J R Meteorol Soc 146(730):1999–2049

Jadli U, Thakur P, Shukla RD (2018) A New Parameter Estimation Method of Solar Photovoltaic. IEEE J Photovolt 8(1):239–247

James V. Miranda L (2018) PySwarms: a research toolkit for Particle Swarm Optimization in Python. J Open Source Softw 3(21):433

Jarvis A, Reuter HI, Nelson A, Guevara E (2008) Hole-filled SRTM for the globe Version 4. Available from the CGIAR-CSI SRTM 90m Database (http://srtm.csi.cgiar.org)

Kang T, Yao J, Jin M, Yang S, Duong T (2018) A Novel Improved Cuckoo Search Algorithm for Parameter Estimation of Photovoltaic (PV) Models. Energies 11(5):1060

Kennedy J, Eberhart R (1995) Particle swarm optimization. In: Proceedings of ICNN'95 - International Conference on Neural Networks. IEEE, Perth, WA, Australia Vol. 4. pp 1942–1948. http://ieeexplore.ieee.org/document/488968/

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

King DL, Boyson WE, Kratochvil JA (2004) Photovoltaic array performance model, SANDIA Report SAND2004-3535. Sandia Report No. 2004-3535 8:1–19

Loutzenhiser PG, Manz H, Felsmann C, Strachan PA, Frank T, Maxwell GM (2007) Empirical validation of models to compute solar irradiance on inclined surfaces for building energy simulation. Sol Energy 81(2):254–267

Ma J, Ting TO, Man KL, Zhang N, Guan S-U, Wong PWH (2013) Parameter Estimation of Photovoltaic Models via Cuckoo Search. J Appl Math 2013:1–8

Mason K, Reno MJ, Blakely L, Vejdan S, Grijalva S (2020) A deep neural network approach for behind-the-meter residential PV size, tilt and azimuth estimation. Sol Energy 196:260–269

McEvoy A, Markvart T, Castañer L. (2012) Practical Handbook of Photovoltaics: Fundamentals and Applications. 2nd edn. Academic Press, Waltham

Meng B, Loonen RCGM, Hensen JLM (2020) Data-driven inference of unknown tilt and azimuth of distributed PV systems. Sol Energy 211:418–432

Mughal MA, Ma Q, Xiao C (2017) Photovoltaic Cell Parameter Estimation Using Hybrid Particle Swarm Optimization and Simulated Annealing. Energies 10(8):1213

Oliva D, Abd El Aziz M, Ella Hassanien A (2017) Parameter estimation of photovoltaic cells using an improved chaotic whale optimization algorithm. Appl Energy 200:141–154

Oliva D, Ewees AA, Aziz MAE, Hassanien AE, Peréz-Cisneros M (2017) A Chaotic Improved Artificial Bee Colony for Parameter Estimation of Photovoltaic Cells. Energies 10(7):865

Perez R, Ineichen P, Seals R, Michalsky J, Stewart R (1990) Modeling daylight availability and irradiance components from direct and global irradiance. Sol Energy 44(5):271–289

Reda I, Andreas A (2004) Solar position algorithm for solar radiation applications. Sol Energy 76(5):577–589

Reda I, Andreas A (2007) Corrigendum to "Solar position algorithm for solar radiation applications". Sol Energy 81(6):838

Reuter HI, Nelson A, Jarvis A (2007) An evaluation of void-filling interpolation methods for SRTM data. Int J Geogr Inf Sci 21(9):983–1008

Ruelle VD, Jeppesen M, Brear M (2016) Rooftop PV Model Technical Report. Technical Report July, University of Melbourne, Melbourne. https://aemo.com.au/-/media/files/electricity/nem/planning_and_forecasting/demand-forecasts/nefr/2016/uom-rooftop-pv-model-technical-report.pdf

Saint-Drenan YM, Bofinger S, Fritz R, Vogt S, Good GH, Dobschinski J (2015) An empirical approach to parameterizing photovoltaic plants for power forecasting and simulation. Sol Energy 120:479–493

Saltelli A (2002) Making best use of model evaluations to compute sensitivity indices. Comput Phys Commun 145(2):280–297

Saltelli A, Annoni P, Azzini I, Campolongo F, Ratto M, Tarantola S (2010) Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Comput Phys Commun 181(2):259–270

Shi Y, Eberhart R (1998) A modified particle swarm optimizer. In: 1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence (Cat. No.98TH8360). IEEE, Anchorage, AK, USA. pp 69–73. http://ieeexplore.ieee.org/document/699146/

Silva EA, Bradaschia F, Cavalcanti MC, Nascimento AJ (2016) Parameter Estimation Method to Improve the Accuracy of Photovoltaic Electrical Model. IEEE J Photovolt 6(1):278–285

Sobol IM (2001) Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Math Comput Simul 55(1-3):271–280

Soon JJ, Low K-S (2012) Photovoltaic Model Identification Using Particle Swarm Optimization With Inverse Barrier Constraint. IEEE Trans Power Electron 27(9):3975–3983

Spencer JW (1971) Fourier series reprensentation of the position of the sun. Search 2(5):172

Stein JS, Perez R, Parkins A (2010) Validation of PV performance models using satellite-based irradiance measurements: A case study. In: 39th ASES National Solar Conference 2010 Vol. 1. pp 265–290

Williams MK, Kerrigan SL, Thornton A (2012) Automatic detection of PV system configuration. In: Fellows C (ed). World Renewable Energy Forum. American Solar Energy Society, Denver, Colorado Vol. 3. pp 1933–1937

Publisher's Note