Applied Energy 361 (2024) 122923

Contents lists available at ScienceDirect

Applied Energy

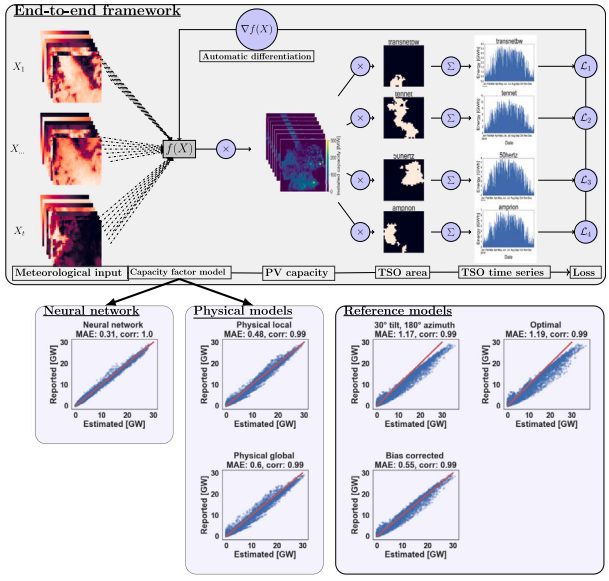
journal homepage: www.elsevier.com/locate/apenergy

End-to-end learning of representative PV capacity factors from aggregated PV feed-ins

Matthias Zech *, Lueder von Bremen

German Aerospace Center, Institute of Networked Energy Systems, Carl-von-Ossietzky-Str. 15, Oldenburg, 26129, Germany

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Energy meteorology
Solar energy
Automatic differentiation
Physics based deep learning

ABSTRACT

Energy system models rely on accurate weather information to capture the spatio-temporal characteristics of renewable energy generation. Whereas energy system models are often solved with high abstraction of the actual energy system, meteorological data from reanalysis or satellites provides rich gridded information of the weather. The mapping from meteorological data to renewable energy generation usually relies on major assumptions as for solar photovoltaic energy the photovoltaic module parameters. In this study, we show that these assumptions can lead to large deviations between the reported and estimated energy, as shown for the case of photovoltaic energy in Germany. We propose a novel gradient-based end-to-end framework that can learn local representative photovoltaic capacity factors from aggregated PV feed-ins. As part of the end-to-end framework, we compare physical and neural network model formulations to obtain a functional mapping from meteorological data to photovoltaic capacity factors. We show that all the methods developed have better performance than commonly used reference methods. Both physical and neural network models have much better performance than reference models whereas operational use cases may prefer the neural network due to higher accuracy while interpretable, physical models are more suited to academic settings.

* Corresponding author.
E-mail address: matthias.zech@dlr.de (M. Zech).

https://doi.org/10.1016/j.apenergy.2024.122923
Received 23 September 2023; Received in revised form 16 January 2024; Accepted 23 February 2024
Available online 28 February 2024

1. Introduction

The consumption of fossil fuels is responsible for the majority of anthropogenic greenhouse gas emissions [1] with the electricity sector accounting for more than a third of the energy-related carbon dioxide [2]. A major pillar in reducing global greenhouse gas emissions is the transition from a fossil-based energy system to a renewable energy system. A wide range of studies show that energy systems based on 100% renewable energy until 2050 are achievable with current technology [3]. Solar photovoltaic (PV) energy has a crucial role within this energy transition that is often still underestimated due to ignorance of public incentive schemes, non-monetary preferences, and rapid technology learning [4]. The expansion of PV technology has experienced remarkable global growth, with annual PV installations increasing from around 30 GW in 2012 to around 240 GW in 2022. This has resulted in a total installed capacity of approximately 1.2 TW worldwide [2]. This trend is expected to continue according to the IEA Sustainable Development Scenario, which projects the PV capacity to increase to 4240 GW in 2040 [5].

The large share of PV capacities calls for accurate modeling of PV energy generation. Physically, the solar resource-to-energy process can be described for single PV modules with known technical properties as described in fundamental solar energy textbooks [6] and implemented with publicly available software [7,8]. Larger PV simulations as in energy system models rely on detailed information of the PV modules, as tilt and azimuth angles, which are, however, not available in publicly accessible worldwide datasets [9,10]. Whereas there is a large body of literature dealing with the optimal determination of tilt angles for site installations [11], determining existing solar module parameters is much less investigated. Recent progress has been achieved to derive locations of PV systems based on Deep Learning in aerial images and satellite imagery, such as, for instance, for the United States [12], regions of Germany [13,14], China [15] and, more recently, even worldwide [10]. However, a publicly available dataset which includes the precise location, capacity, azimuth and tilt angles is still missing. This is crucial for PV generation modeling as individual PV system orientations are not determined based on yield-optimal configurations but are subject to individual regional conditions such as roof angles, leading to wide distributions of tilt and azimuth angles as observable in actual data [16].

A concept widely applied in energy system modeling to handle this missing data problem and to further reduce computational efforts is the usage of *representative* sites which aggregate multiple sites into single ones. PV energy is then modeled on the basis of simplified physical models with assumed module parameters or empirical power curves. This concept is used by energy system models such as the `pypsa-eur` model [17] or implementations of the `eTraGo` model [18]. Other datasets sample solar module parameters from distributions as in Scholz [19] or within the `Renewables.ninja` dataset [20] which is a widely used dataset [21–25] and is the default weather dataset in energy system models such as the `Dispa-SET` model [26] or the `DynElmod` model [27].

Methodologically, the `Renewables.ninja` dataset [20] exploit both reanalysis and satellite data to derive realistic PV energy feed-ins with randomly sampled azimuth and tilt angles to match reported Transmission System Operator (TSO) PV feed-ins [20]. The method from Pfenninger and Staffell [20] has significantly improved the state-of-the-art of large-scale photovoltaic simulation, but comes with some drawbacks: The assumption that azimuth and tilt angles follow a normal distribution is not observable in studies of actual PV systems [16] which mostly show right-skewed distributions for tilt angles and very sharp peak (leptokurtic) distributions for azimuth angles. Furthermore, it is not very intuitive to assume that tilt angles are symmetric around a mean, as local installations clearly follow local patterns as latitudinal position and the tilt of existing roofs which are more likely to be similar in close areas than to be sampled from a distribution. In addition, sampling from a distribution leads to different solutions for each sampling and to non-interpretable tilt and azimuth angles. Lastly, a subsequent bias correction is needed leading to a two-step approach in which it is unclear what the bias actually corrects.

This study proposes an alternative approach that, instead of sampling module parameters, learns module tilt and azimuth angles from reported TSO feed-ins. As solar module parameters are needed locally for each site and solar feed-ins come as aggregates from TSOs, we formulate an integrative model which is able to learn lower level capacity factors from spatially aggregated PV feed-ins. This translates into a hard mathematical problem not only because of the inherent hierarchical dependency but also as PV feed-ins depend non-linearly, due to the underlying nonlinear physical process, on representative site azimuth and tilt angles. For one time step, the problem is ill-posed as necessarily more representative sites than aggregated feed-ins exist. However, given large enough time series of solar energy feed-ins, we show that reliable estimates can be derived with a data-driven Machine Learning approach. We refer to this framework as an end-to-end framework, as it directly connects the gridded meteorological information with the aggregated PV feed-in time series. Different physical conversion models can be integrated into this framework, for which we compare two *physical* formulations, global and local site characteristics, and one *neural network* formulation. In a case study, we show that the proposed framework enables us to learn local capacity factors in Germany from the aggregated PV feed-ins from the four TSOs. More concretely, we investigate the following hypotheses:

Hypothesis 1. Estimated PV feed-ins through traditional method highly deviate from aggregated reported PV feed-ins in both accuracy and statistical relationship

Hypothesis 2. The proposed end-to-end model enables to efficiently learn capacity factors from spatially aggregated PV feed-ins

Hypothesis 3. The neural network is most performant, but the physical model is more interpretable and transferable

Hypothesis 1 is also the main motivation for this paper as large deviations between observed and estimated feed-ins justify the need for more better estimation methods. Hypothesis 2 aims to show that the end-to-end model framework is able to learn from aggregated feed-ins despite the problem is ill-posed. As part of the end-to-end model, the neural networks are expected to have higher performance than physical models as they are known to be very efficient at fitting nonlinear relationships. As neural networks are black-box models, it is expected that physical models have higher interpretability. This trade-off between interpretability and accuracy, and its implications for the usage within the energy context, is investigated in Hypothesis 3. None of these hypotheses have been investigated in the literature so far, highlighting the novelty in this work.

The paper contains a methodological and a practical contribution. We first introduce the methodological proposals of this paper in Section 2 which formulates the end-to-end model framework and the capacity factor models. The case study is then introduced (Section 3) and is based on the hypotheses evaluated (Section 4). After discussing the results (Section 5), this study is finally concluded and future research directions are formulated (Section 6).

2. Model

2.1. End-to-end model framework

To learn local capacity factors from aggregated information, the process of converting meteorological input to aggregated energy feed-ins for all aggregated entities is embedded in an integrative model

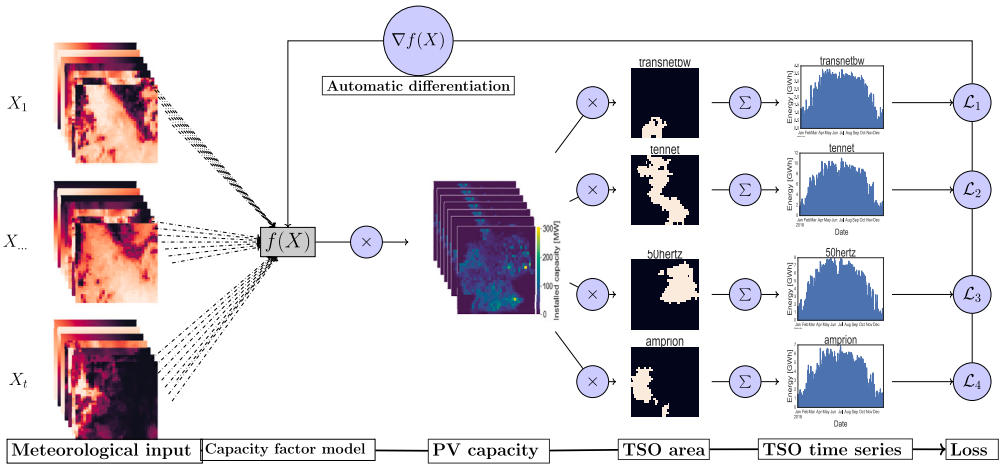


Fig. 1. Schematic illustration of the end-to-end model framework to map meteorological data to TSO time series. Blue circles represent mathematical operators. Note that the automatic differentiation is illustrated for the sake of clarity around the model chain whereas in its implementation it is propagated backwards through the mathematical operators.

which we refer to as *end-to-end framework*. This model allows to connect meteorological inputs $X_{t,s}$ for each model grid point $s \in S$ and time step $t \in T$ to the time series of the aggregated entities $s \ast\in S^*$. In the context of this paper, s and t represent the space–time grid of the meteorological gridded data and s^* represents the set of the TSOs. A forward pass through the model chain from inputs to outputs is described in the following to provide a better understanding its inner mechanics, which is also visualized in Fig. 1.

The end-to-end model begins with the conversion of spatially gridded meteorological information into PV capacity factors. For the sake of clarity, this is described here as a function f which converts the continuous meteorological input variables $X_{s,t}$ for each spatial point $s \in S$ and for each time step $t \in T$ into capacity factors. Different implementations for the capacity factor models f are described in the next two subsections. The capacity factors are thus formulated here as

$$c_{s,t} = f(X_{t,s}) \tag{1}$$

By definition, the capacity factors have values between 0 and 1 because negative energy cannot be generated and the available energy is upwardly limited by the installed capacity. To convert the capacity factors into energy, the capacity factors $c_{s,t}$ are multiplied element-wise by the available PV capacities

$$G_{s,t} = c_{s,t}^{max} \circ c_{s,t} \tag{2}$$

leading to generated PV energy time series for each site. To assign the spatial sites s to the upper-level entities s^*, the PV time series are modified through

$$G_{s,s^*,t} = G_{s,t} \circ M_{s,s^*} \qquad \forall s^* \tag{3}$$

with M_{s,s^*} representing a two-dimensional binary matrix ($s \times s^*$) which maps from lower spatial levels s to aggregated spatial levels s^*. The entries in M_{s,s^*} indicate the belonging of a site to the aggregated entity and zeros indicate no belonging, respectively. More precisely in the context of the case study: If one site belongs to a TSO, it is indicated by 1 in the matrix M and 0 otherwise. Note that the indices do not match in the element-wise multiplication. The notation means that the element-wise multiplication is broadcasted by the index s^* meaning that the tensor $G_{s,t}$ is multiplied element-wise $|s^*|$ times by the binary matrix M_{s,s^*}. This leads to the tensor $G_{s,s^*,t}$ which represents for each aggregated entity s^* at each site s its respective PV feed-in time series.

To make the estimations match with the aggregated time series, the tensors are summed up over all sites

$$G_{s^*,t} = \sum_s G_{s,s^*,t} \tag{4}$$

which results in the final estimated PV generation time series of the TSOs. Through the loss function, the divergence between reported $y_{s^*,t}$ and calculated hourly PV generation $G_{s^*,t}$ is calculated. The models are trained on the L1-Loss in this study which leads to the loss

$$\mathcal{L}(y_{s^*,t}, G_{s^*,t}) = \frac{1}{N}\frac{1}{S}^* \sum_n \sum_{s^*} |y_{s^*,t} - G_{s^*,t}| \tag{5}$$

The L1-loss is chosen for its robustness towards outliers and its high interpretability, as it directly reflects the average magnitude of errors. Based on the loss, the gradients are calculated and reported back to the capacity factor model. The capacity factor model then adapts its model weights by the gradients in the direction of improving the loss and restarts the forward pass of the end-to-end framework.

The gradients are not derived analytically but based on the `pytorch` library [28] which provides efficient implementations to automatically calculate gradients through efficient exploitation of the chain rule [29]. Despite it is more known as a method to train deep neural networks, automatic differentiation has shown high promise in this study to learn parameters with the hierarchical structure of the end-to-end framework.

2.2. Capacity factor model

To derive solar capacity factors, this section describes two different approaches to define the function f in Eq. (1) which maps the meteorological information to capacity factors. Both proposals are implemented with standard mathematical operators and linear algebra which makes them trainable through gradient-based approaches. Therefore, they can both be implemented in a modular fashion in the end-to-end framework from Section 2.1.

2.2.1. Physical model

The physical model aims to physically describe the PV energy conversion process and to optimize the most relevant parameters of the physical model. For this purpose, a simplified PV conversion model is developed,inspired by the physical formulation of the `atlite` library [30] and implemented in `pytorch` [28] to enable automatic differentiation.

To calculate the total irradiance on a tilted plane, we use a simple trigonometric model based on Sproul [31]. The total irradiance I_{glob}^T on a tilted plane consists of the direct I_{Dir}^T and diffuse irradiance I_{Dif}^T

$$I_{glob}^T = I_{Dir}^T + I_{Dif}^T \tag{6}$$

The direct irradiance on a tilted surface I_{Dir}^T can be derived in accordance with Sproul [31] following

$$I_{Dir}^T = I_{Dir}^N \cos\theta \tag{7}$$

where I^N_{Dir} represents the direct irradiance on a plane normal to the direct beam irradiance and θ the sun incidence angle of the direct irradiance. Following the simple vector analysis in Sproul [31], the last term can be calculated from

$$cos\theta = \sin \boldsymbol{\beta^{pv}} \cos \alpha^s \cos(\boldsymbol{\gamma^{pv}} - \gamma^s) + \cos \beta \sin \alpha^s \quad (8)$$

with β^{pv} PV tilt angle, α^s sun altitude, γ^{pv} azimuth angle of the PV system and γ^s azimuth angle of the sun.

The diffuse irradiance on the tilted plane is retrieved from Reindl et al. [32] through

$$I^T_{Dif} = \frac{1 + \cos \boldsymbol{\beta^{pv}}}{2} I^N_{Dif} + \rho I^N_{Glob} \frac{1 - \cos \boldsymbol{\beta^{pv}}}{2} \quad (9)$$

with I^{Glob}_n total influx on a horizontal plane, ρ the ground albedo which is derived in accordance with Hofmann et al. [30] through

$$\rho = \frac{I^N_{Glob} - I^N_{Net}}{I^N_{Glob}} \quad (10)$$

whereas I^N_{Net} represents the solar radiation reaching a horizontal plane minus the ground albedo. This parameter can be retrieved from reanalysis models, as from the ERA5 model [33] used in this study, and allows a more accurate description of the ground albedo than commonly used constant values as in comparable studies [19,20].

The tilted irradiance is next translated into PV energy using the PV module performance model from Huld et al. [34] which only depends on the module temperature and the in-plane irradiance through

$$P(I^T_{glob}, T_{mod}) = P_{STC} \frac{I^T_{glob}}{I_{STC}} \eta_{rel}(I', T') \quad (11)$$

with P_{STC} indicating the power at standard test conditions (STC) of $I_{STC} = 1000\,\mathrm{W/m^2}$ and $T_{mod,STC} = 25\,°\mathrm{C}$ respective temperature. The relative efficiency η_{rel} is calculated as proposed from Huld et al. [34]

$$\eta_{rel}(I', T') = 1 + k_1 \ln I' + k_2 [\ln I']^2 + T'(k_3 + k_4 \ln I' + k_5 [\ln I']^2) + k_6 T'^2 \quad (12)$$

where I' and T' are normalized parameters to STC values with $I' = I^T_{glob}/1000$ and $T' = T_a + 0.035G - 298.15$ [34]. The parameters $k_1, \ldots, k_6$ are empirical coefficients that are fitted in practice to modules [34] but in this study are based on the standard carbon silicon PV module from the `atlite` library [30]. As a last step in the physical model, the energy is reduced by the inverter efficiencies that are assumed to be 90% in accordance with Hofmann et al. [30].

In this study all the described parameters are fixed except for the tilt angle β^{pv} and the solar azimuth angle γ^{pv} printed in bold in Eq. (8). In a data-driven way, these are learned within the end-to-end model $f(cf_{s,t}|\beta^{pv}, \gamma^{pv}, X)$ using the feature matrix X. Note that β^{pv} and γ^{pv} can both be globally defined or with locally different values $\beta^{pv}_s, \gamma^{pv}_s$ which is both tested in this study. Using global values aligns with other studies applying one reference module whereas locally different values are the more realistic choice.

2.2.2. Neural network formulation

As an alternative to the physical capacity factor model, a multi-layer perceptron (MLP) neural network formulation is proposed which ignores any physical knowledge. Neural networks are very efficient at fitting non-linear relationships which makes them a promising tool to learn the PV energy conversion model, which is, for example, already shown for single PV module measurements [35]. More precisely, the neural network learns the energy conversion process including the PV panel efficiencies from scratch. The input features of the neural network are selected based on the relevant parameters of the physical model to make the results more comparable. According to *Occam's razor*, the simplest possible model should always be preferred over more complex high-parametric models [36]. The final neural network architecture is thus determined by starting with a simple 2-layer dense neural network and its complexity is gradually increased until the training and validation error did not improve anymore. This finally leads to a simple 3-layer deep neural network as shown in Fig. 2(b) with layers of 7, 8 and 5 neurons. As in addition to the weights, a bias term is added to each neuron as shown in Fig. 2(a), the total number of learnable parameters consists of the weights ($7 * 8 + 8 * 5 + 5 = 101$) and the bias ($|l_1| + |l_2|) = 13$ leading in total to 114 learnable parameters. Note that this network is small compared to neural networks applied to image recognition tasks which usually contain more than a million parameters [37]. The layers are connected with Rectifier Linear Unit (ReLU) activation functions which allows the neural network to learn non-linear relationships as stated by the universal approximation theorem [38] which is necessary as illustrated in the description of the physical process in Section 2.2.1. The final outputs of the neural network are forced to be between 0 and 1 using the sigmoid function ($sig(z) = \frac{1}{1+e^{-z}}$) with z representing the output of the last hidden layer of the neural network. Using the sigmoid function is a trick to restrict the codomain of the capacity factor model function to the feasible set of capacity factors in accordance with the end-to-end framework formulation making physical and neural network outputs directly comparable. Furthermore, the sigmoid function and its optimization behavior are well studied in neural network implementations due to its standard usage as final outputs for binary classifications. Before feeding the data into the neural network, each parameter is standardized through $x_{t,s} = \frac{x_{t,s} - \bar{x}_{t,s}}{\sigma(x_{t,s})}$ in which the mean $\bar{x}_{t,s}$ and the standard deviation $\sigma(x_{t,s})$ are retrieved from training data to prevent data leakage. Standardization improves the numerical stability of the optimizer.

In contrast to the physical model, the neural network does not have any location-specific information, but aims at finding a general mapping of the meteorological inputs to capacity factors. It is therefore comparable to a physical model with global azimuth and tilt angles.

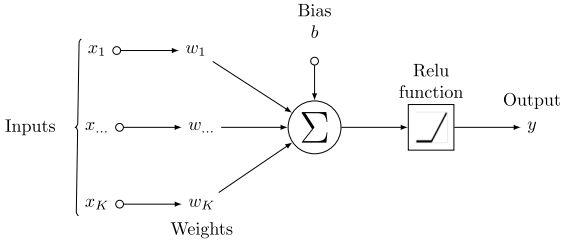
3. Case study

3.1. Models

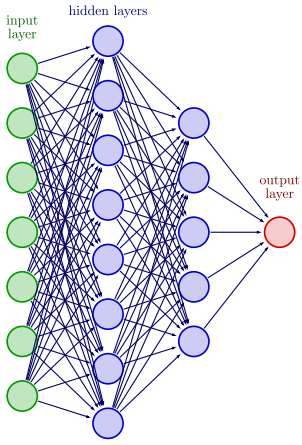
To illustrate the performance of the proposed models, the end-to-end model is tested for three different formulations of capacity factor models. The *physical global model* (**PhyGlob**) assumes that all representative sites have the same trainable azimuth and tilt angle within Germany. The *physical local model* (**PhyLoc**) allows different trainable tilt and azimuth angles for each site which is a more realistic setting given the large heterogeneity of PV module orientations. The *neural network model* (**MLP**) only assumes one global model without local information. Therefore, the MLP and PhyGlob models do not contain local site information, while the PhyLoc model allows for different azimuth and tilt angles.

To verify the models three reference models are formulated based on the same physical model from Section 2.2.1, but with assumed fixed azimuth and tilt angles as implemented in the `atlite` library [30]. Two reference models are investigated following the implementations which we refer to as *conservative model* (**RefCon**) and the *optimal layout model* (**RefOpt**). The RefCon model assumes fixed tilt angles of 30° for all representative sites. The RefOpt model refers to a layout of tilt and azimuth angles that maximize the energy yield. In case of Germany with latitudes ranging from roughly 44° to 55°, all latitudes between 44° to 50° obtain a tilt angle equal to the latitude and for all values above 50° a tilt angle of 40° is assumed in accordance with the `atlite` library [30]. A third reference model uses the RefCon model and applies a multiplicative bias correction scheme based on TSO feed-ins which we name **RefCon+**. Applying bias correction is a widely used technique to avoid systematic over- or underestimation for the meteorological data [39] and also for the converted PV energy feed-ins [20].

The RefCon model is inspired by similar assumptions of tilt (30°) and azimuth angles (0°) in other studies using PV simulation for Germany and Europe. The `pypsa-eur` model uses a fixed tilt angle of 35° for entire Europe [17]. Furthermore, the examples within the `atlite` and the `feedinlib` library also propose to use 30° tilt angles. The `Renewables.ninja` dataset uses a normal distribution of both tilt ($\mathcal{N}(35°, 15°)$) and azimuth angles ($\mathcal{N}(180°, 40°)$) [20].



(a) Structure of neurons used within hidden layers



(b) Used deep neural network architecture with 7 input variables, 2 hidden layers and sigmoid output

Fig. 2. Structure of used neural network model in this study.

3.2. Data

The main data source for the meteorological data originates from the global atmospheric reanalysis model ERA5 from the European Centre for Medium-Range Weather Forecasts with a spatial resolution of $0.25° \times 0.25°$ representing 31 km [33]. The retrieved parameters are the global horizontal irradiation (I^N_{Glob}), the surface net solar irradiation (I^N_{Net}) necessary to retrieve the ground albedo ($\rho = I^N_{Glob} - I^N_{Net}$), the direct solar irradiation (I^N_{Dir}) to derive the diffuse irradiation ($I^N_{Dif} = I^N_{Glob} - I^N_{Dir}$) and the temperature at 2 meters (T_a). This leads to four meteorological parameters retrieved from the ERA5 data ($I^N_{Dif}, I^N_{Dir}, \rho, T_a$). To obtain the other necessary input data, the other parameters specific to the sun (solar elevation angle, solar azimuth angle and top of the atmosphere insulation) are calculated based on the solar model from Reindl et al. [32] through the `atlite` library.

Besides reanalysis data, satellite-derived irradiance is often used within energy system modeling. They are known to provide accurate solar irradiance estimates due to the high spatio-temporal resolution and the ability to directly measure the solar radiation reaching the Earth's surface [40]. Therefore, as a second meteorological data source, the SARAH-2 [41] data from the Meteosat-based CM-SAF satellites is used to compare the generalizability of the models to different datasets. To ensure that the spatiotemporal resolutions of both ERA5 (1 h, 0.25°) and SARAH-2 (30 min, 0.05°) are identical, the SARAH-2 data is aggregated to hourly values and averaged to the ERA5 spatial grid. The variables retrieved from SARAH-2 are the direct I^N_{Dir} and diffuse irradiation I^N_{Dif} whereas for all other parameters ERA5 data are used. The meteorological data used in this study are shown for one hour in Fig. 3. A comparison of the direct and diffuse irradiance from SARAH-2 and ERA5 reveals the greater level of detail in SARAH-2.

The installed PV capacities are retrieved from the Open Power System Data (opsd) database [42] on postal code levels and are remapped to the ERA5 grid resolution. The capacities are considered as yearly energy capacities and are based on all available PV systems until the mid of the respective year. The PV generation time series from the four German TSOs 50Hertz, Amprion, TenneT and TransnetBW are also retrieved from the opsd database [42]. The download and the postprocessing of the meteorological data are based on scripts within the `atlite` library [30].

3.3. Model training and verification

Even if the physical and neural network model formulations are fundamentally different from a technical point of view, mathematically they are both non-convex optimization problems which are computationally hard to solve. This is the reason why a data-driven approach for both models is chosen. Whereas for the neural network this is a common approach, we show that also the physical models can be trained to reasonable values using automatic differentiation. To prevent from overfitting, the models are trained using a 5-fold cross-validation approach. This means that each year between 2012 and 2017 is left out for validation once and the model is trained on the other years. This is repeated until all years are once in the validation set (5 times) which represents one training epoch. The models are trained on both ERA5 and SARAH-2 datasets and the validation score is averaged over both datasets. This not only makes the models more robust but allows one to use the same models for both datasets. The training procedure is stopped when the optimizer is unable to improve the validation accuracy during the last 3 epochs, which is more commonly known as *Early stopping* [36]. The L1-Loss is chosen to be optimized for each single TSO as a multi-task loss with the Adam [43] optimization algorithm and a batch size of one entire year (8760 h). Different learning rates $\{7.5e^{-4}, 1e^{-3}, \dots, 2.5e^{-1}\}$ are used to test the sensitivity to the learning rate. The linear layers of the neural network are initialized based on the standard uniform distribution ($\mathcal{U}_{[-a,a]}$ with $a = \frac{1}{\sqrt{n}}$ and n neurons). The physical models are initialized based on the conservative reference model of 30° tilt angle and 180° azimuth angle perturbed by Gaussian noise with tilt angles to sample from $\mathcal{N}(30°, 5°)$ and azimuth angles sampled from $\mathcal{N}(180°, 1°)$. The final verification metrics are evaluated on the left-out year (2018) not used within training and verification.

All models, except the reference models, were trained on GPU architectures (Tesla V100, 16 GB) which took a few hours. Training the model and performing inference are, however, also possible on CPU-only architectures as the utilized `pytorch` [28] library supports both GPU and CPU. It is important to note that this leads to longer training times, as the GPU proved to be approximately twice as fast compared to training on a CPU-only architecture (20 cores).

4. Results

4.1. Training

Fig. 4 depicts the validation error (MAE) during the training procedure of the two physical models (PhyLoc, PhyGlob) and the neural network model (MLP). The validation error subject to training iterations (Fig. 4(b)) shows the learning behavior that provides information

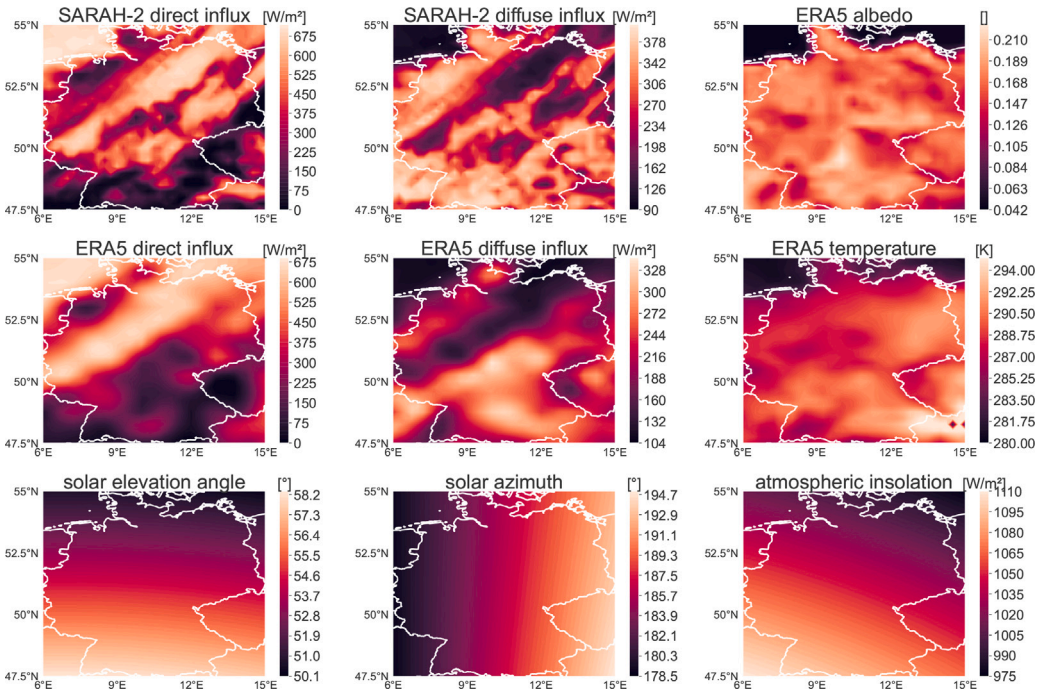
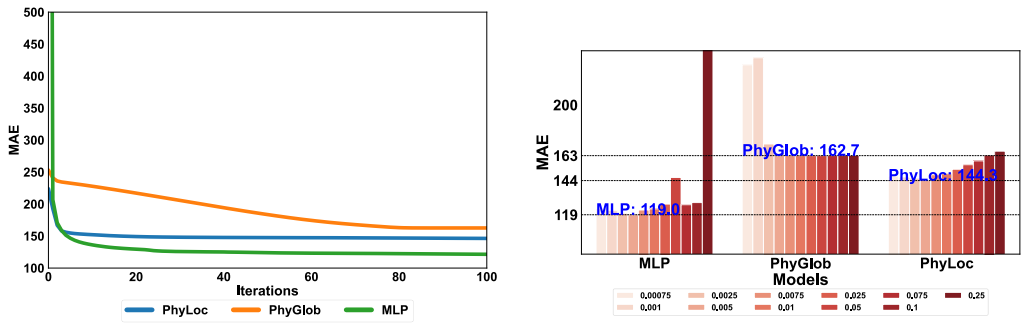


Fig. 3. Exemplary data used to derive capacity factors. (2018-05-03 12:00).



(a) Validation loss by iteration (best learning rate) (b) Validation loss by learning rate

Fig. 4. Learning performance of the different model implementations.

about the bias of the model and the risk of overfitting. The models are relatively resistant to overfitting, except for the case of 0.25 for the MLP, visible in no sudden increases of validation error. Regarding the learning rate (Fig. 4(a)), a wide range of learning rates lead to similar results. Comparing the best performance between the models, the neural network achieves a lower validation error compared to both physical models. Between the physical models, the local physical model performs better than the global physical model (12%).

4.2. Performance

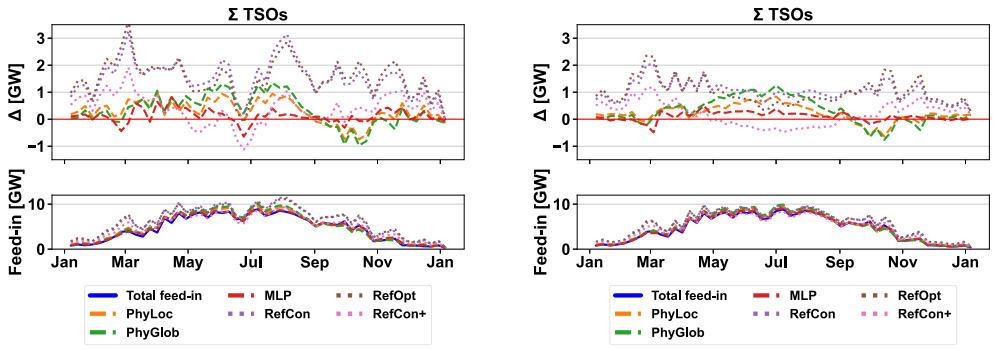
The performance of the different proposed methods and the reference models is shown in Table 1. The MLP is strictly better than the other methods. This translates into bias reductions of up to 90% (RefOpt, SARAH-2) and MAE reductions of up to 74% (RefOpt, SARAH-2). Compared to its closest competitor (PhyLoc), the MLP still shows better metrics with more than 50% lower bias for the SARAH-2 dataset and 60% for the ERA5 dataset. The MLP is also more performant for both datasets in all metrics compared to the other methods. Very high correlation coefficients are observable for all methods and datasets whereas the neural network provides close to perfect positive correlation (0.997).

Table 1
Model characteristics and performance based on left-out year (2018). Favorable properties are printed in bold.

	PhyLoc	PhyGlob	MLP	RefCon	RefOpt	RefCon+
MAE ERA5	0.64	0.72	**0.42**	1.77	1.86	0.83
MAE SARAH-2	0.48	0.60	**0.31**	1.17	1.19	0.55
Correlation ERA5	0.988	0.986	**0.993**	0.975	0.968	0.975
Correlation SARAH-2	0.993	0.991	**0.997**	0.991	0.986	0.991
Bias ERA5	0.31	0.35	**0.13**	1.56	1.57	0.40
Bias SARAH-2	0.21	0.31	**0.10**	1.03	1.11	0.20
# Fitted parameters	$2\|s\|$	2	114	–	–	2

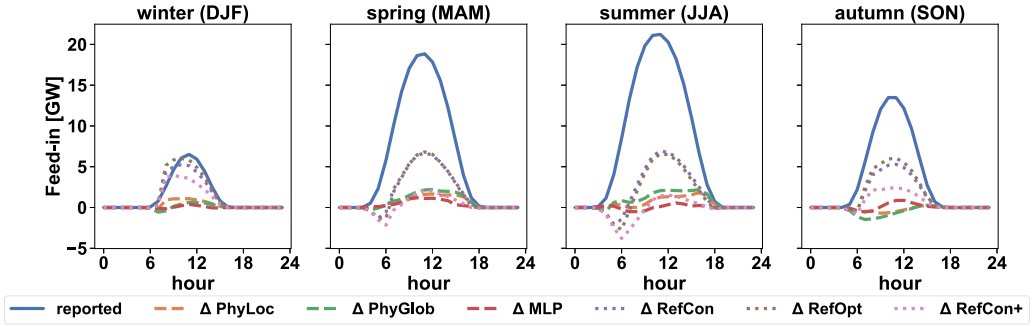
Regarding the choice of the meteorological datasets, the SARAH-2 shows better performance in all metrics compared to the ERA5 dataset. Using SARAH-2 instead of ERA5 improves the bias and MAE by up to 34% (RefCon model).

Between the reference models, the RefCon model leads to better metrics for all datasets than the RefOpt model showing that this is the more competitive benchmark model. Applying a bias correction scheme further improves the RefCon model showing competitive bias results for the SARAH-2 dataset (0.55) which is even better than the PhyGlob (0.6)



(a) ERA5 (b) SARAH-2

Fig. 5. Residuals (upper plot) and model estimates (below) of weekly mean solar energy feed-in for different meteorological data sources (ERA5, SARAH-2) in 2018.



(a) ERA5

(b) SARAH-2

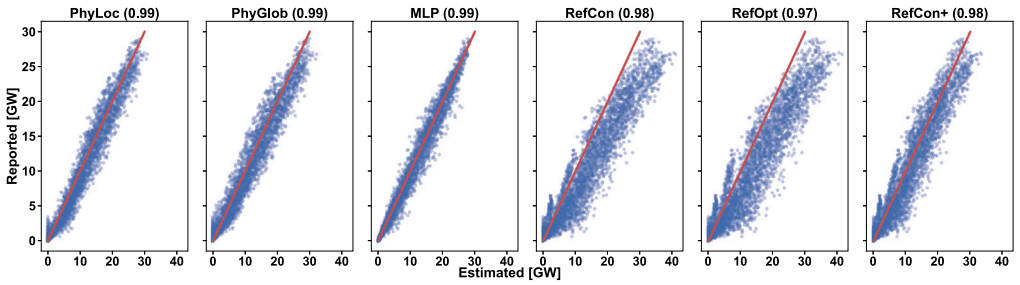
Fig. 6. Observation (solid line) and model estimation residuals (dashed lines) averaged over seasons in 2018.

model. In terms of MAE, the RefCon+ shows worse MAE values than the PhyLoc (0.83 vs. 0.64) for ERA5 and SARAH-2 (0.55 vs. 0.48).

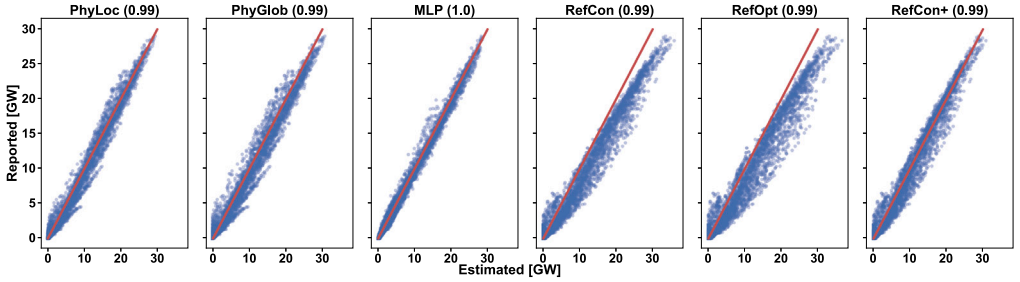
Note that the application of a multiplicative correction scheme can be interpreted as compressing ($0 < a < 1$) or decompressing a point cloud ($a > 1$). Therefore, applying a bias correction scheme as in the RefCon+ model does not improve the correlation coefficient and therefore also does not improve the statistical relationship between feed-ins and estimates. This is shown in Table 1 as Refcon and RefCon+ model have the same correlation coefficient. A more rigorous proof of this statement is formulated in the proof in Appendix to which the interested readers are referred.

To understand the aggregated metrics better, we decompose them into weekly, daily, and hourly components. Fig. 5 shows the weekly mean of the PV feed-ins and the residuals between model estimations and reported feed-ins. Uncorrected reference models lead to large overestimation during the winter season peaking in March and November regardless of the meteorological data used. Whereas bias correction can improve the bias on average, the yearly cycle shows that this comes at the cost of a negative bias during summer as shown by the blue dotted line in Fig. 5 with still large a large positive bias in winter. The fitted physical models PhyGlob and PhyLoc are able to reduce the winter bias compared to the reference models but to the cost of overestimation during the summer season. The neural network only has minor errors, without apparent seasonal bias. Large differences between the weather datasets can be observed not only in magnitude but also in the structural behavior of the time series. For example, in August, all methods show a large positive bias using ERA5 data while this is not observable using the SARAH-2 data.

The diurnal bias for the different seasons is depicted in Fig. 6. The uncorrected reference models show large positive bias during the winter period which is for the ERA5 dataset as large as the actual feed-in. Applying a bias correction reduces this large winter overestimation to the cost of larger underestimations in summer. Furthermore, all reference models suffer from a negative bias during the early hours of the summer days, which is worsened by bias correction. The fitted physical models show much smaller bias during the evening hours in



(a) ERA5



(b) SARAH-2

Fig. 7. Scatterplot between observed and estimated feed-ins (2018).

summer a positive bias can be observed which is more dominant in the SARAH-2 dataset. Again, the neural networks do not show any bias.

To evaluate the actual fit of the different models, Fig. 7 compares the hourly estimates against the reported values and its respective correlation. The neural network accurately captures the statistical relationship as shown by the thin lines for both weather datasets. The SARAH-2 dataset has a smaller spread shown by narrower lines for all methods. The reference models show a distinct positive bias indicated by the large number of values below the red lines. Bias correction can aid to improve this, yet at the cost of a larger number of underestimations. The physical models also show a very distinct positive bias leading to spikes in the case of the SARAH-2 dataset which aligns with the summer evening hour bias observed in the diurnal cycle in Fig. 6. The reference models show for both datasets high deviations as visible through the fan-like pattern in the graph origin for both datasets. The fitted models are able to largely reduce these deviations.

4.3. Interpretability

In contrast to the blackbox neural network model, the physical models derive actual tilt and azimuth angles and therefore provide interpretable physical entities. The spatially different azimuth and tilt angles of the PhyLoc model including its distributions are shown in Fig. 8. The distribution of tilt angles is located between 0° and 60° with a mean of 10.4° showing a right skewed distribution. The azimuth angles (Fig. 8b) have a predominant south-east orientation in the range from 135° to 181° with a mean of around 168°. A large peak around 180°, meaning perfect south orientation, is observable. The PhyGlob model has a single azimuth angle of 181° and a single tilt angle of 1° which represents a flat PV system oriented southward. The PhyLoc model, therefore, leads to a more realistic setting covering a wide range of plausible PV panel configurations.

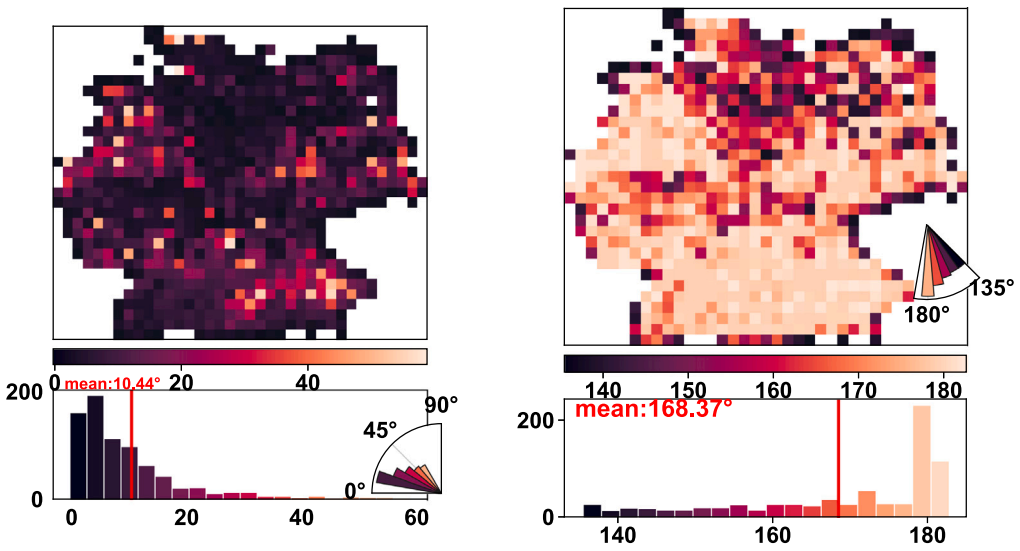
In addition to having interpretable weights, both physical and neural network models have the advantage of providing capacity factors. Fig. 9 shows the global horizontal irradiation (GHI) and the respective estimated model capacity factors. The difference between global models (PhyGlob, MLP) and the model with local information (PhyLoc) becomes evident when looking at the spatial smoothness of the capacity factors. The PhyLoc model leads to a pixelated image, whereas the other models learn a general applicable capacity factor model. It therefore leads to a more reasonable solution as different azimuth and tilt angles necessarily also lead to different capacity factors, which is not true for the global models.

Furthermore, the physical models have an obvious but useful advantage. As the weights are standalone physical entities, they can be used within existing physical model implementations.

5. Discussion

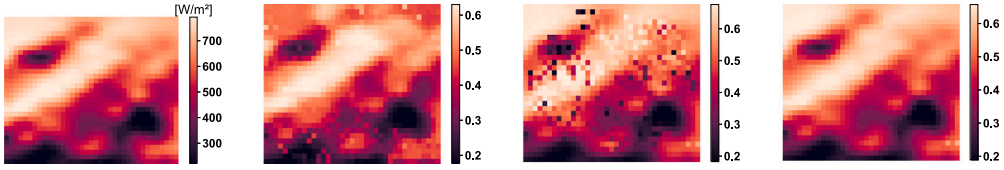
This section summarizes and discusses the results in alignment with the formulated hypotheses and the existing literature.

Existing reference models fail to accurately model PV energy feed-ins. Both conservative layout (RefCon) and yield-optimal layout (RefOpt) lack model performance. Without any correction, both reference models highly overestimate the actual PV feed-in leading to a bias of 100% in the winter months. Applying a multiplicative bias correction as in Pfenninger and Staffell [20] can reduce the yearly bias, but still shows a large positive bias during the winter months and a strong negative bias during the summer months. In the energy context this is undesirable, as during times of high load (winter) PV feed-in are overestimated and during times of low load (summer) PV feed-ins are heavily underestimated. This could lead to underestimations during critical winter periods of backup and storage capacities. Furthermore, underestimated PV feed-ins during summer underrate the risk of PV curtailments. In addition, we show and prove that a multiplicative bias correction cannot improve the correlation. Therefore, systematic misalignments of PV modules cannot be corrected by bias correction. Furthermore, applying a bias correction directly on the feed-ins hamper the link to the source of the bias making it impossible to find the reason for the bias. By consequence, a bias could correct a multitude of different aspects of the model chain as for instance a meteorological bias, insufficient physical model descriptions, erroneous capacity assumptions or the PV module efficiency. As tilt and azimuth angles are besides the installed capacity the most sensitive parameters for PV output [44], it is likely that the bias corrects incorrectly assumed tilt and azimuth angles while an end-to-end model, as proposed in this study, is able to learn these directly. This is also shown in the results of this study as a large proportion of the errors cannot be reduced with bias correction while the proposed fitted models can.



(a) Tilt angles (b) Azimuth angles

Fig. 8. Local physical model weights and its distributions (red line represents mean). The polar coordinate plot indicates distribution within conventions (logarithmic y-axis for illustration).



(a) GHI (b) MLP model (c) PhyLoc (d) PhyGlob

Fig. 9. Global horizontal irradiance and model capacity factors (ERA5, 2018-05-03 12:00) for the spatial extent of Germany.

The proposed end-to-end framework is modular and efficient. All models within the framework show superior performance over all metrics compared to the uncorrected reference models. In the case of the bias corrected reference model, the PhyLoc and the neural network still show strictly better performance. The neural network outperforms the other models evident by smaller errors compared to the RefCon model (MAE: 74%) and the RefCon+ model (MAE: 44%). Regarding the learning behavior, it is observable that all models show clean learning behavior without overfitting. The physical models start from a much better start than the neural network which can be explained by their physical backbone. Embedding the learning within physical equations provide a smaller learning space which leads to good initial learning values but also. However, if the physical model is too simple, it can also lead to restrictive capabilities of the model. All models seem relatively robust, as even with a different set of learning rates solutions of similar accuracy are achievable. Between the capacity factor models, the best performing learning rate is similar. Therefore, we show that the proposed end-to-end framework is able to learn effectively, making it an attractive methodology to connect the inherent hierarchical dependencies within a learnable framework.

The modularity of the end-to-end framework provides rich further research avenues. We show that two very different capacity factor models based on neural networks and physical models are able to learn within the framework. Due to its modular formulation, the end-to-end framework is applicable to other research areas with hierarchically organized data as the only requirement of the model chain is that the mathematical operators need to follow tensor operations implemented in common deep learning frameworks. For instance, the derivation of wind capacity factors given an appropriate wind capacity factor model for f can be easily formulated in the same way as the PV capacity factor model. Furthermore, other parameters of the model can be adapted, such as the energy capacities $c_{s,t}^{max}$ and the mapping matrix M_{s,s^*}. The capacity matrix could include weights to estimate the data uncertainty of installed power capacities. A modification of M_{s,s^*} would allow to incorporate other hierarchical information organizations: In this study, one location can only belong to one TSO, however, spatial points could only belong partly ($M \in \mathbb{R}^+$) or to multiple aggregated time series ($\sum_{s*} M_{s*,s} > 1; M \in \{0,1\}$). The latter would apply under the consideration of multiple hierarchies, as for instance, the simultaneous consideration of postal code feed-ins and TSO feed-ins. This setting resembles the problem of forecast reconciliation which is an active research area within the forecasting literature as described in Hyndman et al. [45] that found successful applications in solar forecasting to model temporal [46] and geographical [47] hierarchical dependencies. Embedding these hierarchies in the end-to-end framework from this study and using the end-to-end framework within forecasting contexts are promising research avenues.

Neural networks are most performant, whereas physical models are slightly less performant but highly interpretable. The neural network highly outperforms the reference models but also the best physical (PhyLoc) model. It has a much lower MAE (35%), better correlation (0.997 vs. 0.993) and lower bias (53%) than the PhyLoc model and does not show specific seasonal or diurnal errors. This is remarkable as the model only contains one model for all sites which shows that one neural network without site information is able to correct wrong assumptions of installed capacities, weather data or the physical conversion process based on the provided features. However, the PhyLoc model still shows large success compared to the best reference model

RefCon+ with MAE improvements of 23% (ERA5) and 13% (Sarah-2). A possible avenue to increase the PhyLoc's performance may be to include a term correcting wrongly assumed installed capacities or module net efficiencies which are treated as constants in the fitted models. This is particularly promising regarding the immense performance gain of the RefCon+ over the RefCon model (MAE SARAH-2: 54%, MAE ERA5: 55%). To assess how realistic the derived tilt and azimuth values of the PhyLoc model are, we compare the calculated values with the actual ones. Killinger et al. [16] conducted a study analyzing tilt and azimuth angles from a commercial database in multiple countries including Germany. In case of Germany, the tilt angles resemble a right-skewed distribution ranging from 0° to 50° whereas the calculated tilt angle distribution in this study is located in the same range with a smaller spread. With respect to orientation, actual PV azimuth angles face all directions from north to south [16] with a predominant number of systems orientated towards the south. Compared to the fitted azimuth angles in this study, the heavy south orientation is also observable, as shown in Fig. 8, while the range of azimuth angles is smaller in this study.

The preferred model depends on the use-case The higher interpretability of the physical models also has a fundamental advantage over the neural network as the modeler has entire control over the model. Single model factors can be adapted as for instance a factor describing installed capacity uncertainty as a specific bias correction factor. A blackbox neural network model provides no link between parameters and errors. The only possibility to improve the neural network performance is to make them more complex and to use more data which does not help in understanding the underlying reason. In contrast, reducing model errors in physical models necessarily leads to a better understanding of the entire physical process, making them more reliable in academic settings. In particular, for studies with long-term focus as intended by Pfenninger and Staffell [20], physical models provide possibilities to adapt model coefficients such as increasing module efficiencies or different installed capacities. Using a blackbox model makes it difficult to understand whether the investigated variability is raised by climate variability or just by an overfitted neural network model. Therefore, a strong academic interest should be to derive interpretable and modifiable models as with the physical model approach (PhyLoc). Large-scale studies deriving renewable sites based on remote sensing data [14] can help to derive new site information to be incorporated in physical model approaches. Yet, the neural network provides more accurate results which are interesting for use cases more interested in accuracy rather than interpretability as, for instance, in optimal control settings.

The meteorological dataset matters. A side-product of this study is the superiority of SARAH-2 data over ERA5 data for aggregated PV energy feed-in simulations. All models, including reference models, show large performance gains through the use of the SARAH-2 data over the ERA5 dataset. This aligns with the observations on irradiance (GHI) conducted by Urraca et al. [40] who state that despite recent improvements of reanalysis data (ERA5, COSMO-REA6), satellite data still have a higher accuracy than reanalysis data [40]. Urraca et al. [40] observe a positive bias in the ERA5 GHI values which also aligns with the observed positive bias of PV generation in this study. Using SARAH-2 instead of ERA5 data can reduce the MAE of the GHI by around 36% ($1 - \frac{12.44}{19.45}$) [40] which is similar to the improvements of the PV feed-ins in this study. The MAE of the reference models can also be significantly reduced using SARAH-2 data (RefCon: 34%, RefOpt: 36%, RefCon +: 34%), while fitted models within the end-to-end framework benefit less from different meteorological datasets (PhyLoc: 25%, PhyGlob: 17%, MLP: 26%). This shows the models robustness regarding different weather data which certainly comes to the cost of learning one general relationship suited to both datasets. For operational settings, training the models only on satellite data is promising to obtain even higher performances. The trained models in this study are forced to suit two different datasets with one (ERA5) having arguably lower accuracy probably reducing the overall algorithm performance. Training on both datasets however was chosen to obtain one general model for both datasets which is certainly an artificial competitive disadvantage over the bias-corrected reference model.

Reliable large-scale feed-ins are difficult to obtain.

The formulation of the representative capacity factor model within the end-to-end framework assumes that learning based on reported aggregated feed-ins can improve the solar capacity factor model. Therefore, a critical remark is how reliable the reported TSO PV feed-ins are. This question is also raised in Pfenninger and Staffell [20] who reports that TSO feed-ins are often not measured but calculated through representative sites. In addition, the reported feed-ins represent actual outputs which include non-physical effects such as market inefficiencies, grid restrictions, technical defects or curtailments. By contrast, energy system models aim to exclude these effects as these are model outputs and not inputs. Even if curtailment is less an issue for the investigated data in Germany with only 0.3% of the potential PV output in 2018 [48], this issue needs to be addressed when the model is applied to other regions such as California [48].

6. Conclusion

This study shows that estimated PV feed-ins based on reference methods deviate from TSO reported feed-ins for the case study of Germany. Both accuracy and statistical behavior do not match with reported feed-ins, raising the question of how to more reliably derive local PV capacity factors from aggregated PV feed-ins. To address this problem, we propose an end-to-end model framework that can recreate PV energy feed-ins more realistically than existing reference models. This study proposes different solar capacity factor models based on blackbox neural network models and based on physical models with learnable tilt and azimuth angles (global and local configuration). The efficiency of this framework is shown by the fact that all model implementations have higher performance than the compared reference models. Regarding the model implementations, the neural network model is more performant than the physical models, making neural networks attractive for studies with a greater focus on accuracy than interpretability. The physical models have high interpretability and better performance than the reference models, making them more appropriate for academia.

This paper opens multiple avenues for future research. As the end-to-end model is broadly formulated, it can be used in a wide range of research contexts. This includes the transfer to other renewable energy sources, such as wind energy, or the consideration of further hierarchical dependencies, such as lower-level distribution grids or postal codes. Another interesting research topic might be to study the sensitivity of the model accuracy to the number of sites, as a minimum number of sites is probably needed to obtain reliable estimates. Furthermore, a future study could investigate how reliable the azimuth and tilt angles are when the model is applied to other systems, including urban systems, with known characteristics as a tool to solve missing data problems.

CRediT authorship contribution statement

Matthias Zech: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Lueder von Bremen:** Writing – review & editing, Supervision, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The presented work has been carried out within the national research project "SOLREV" (FKZ 03EE1010E) funded by the Federal Ministry for Economic Affairs and Climate Action (BMWK) on the basis of a decision by the German Bundestag. M. Zech has been supported by the German Federal Environmental Foundation within the PhD scholarship (grant no. 20020/667-33/2). The authors thank Hendrik-Pieter Tetens for helpful feedback during the review phase. Furthermore, the authors would like to thank the project members of the SOLREV project and the energy meteorology group for interesting discussions which have improved this paper.

Appendix. Proof: invariance of scale for correlation (multiplicative bias correction)

We want to show that the multiplicative bias correction does not change the Pearson correlation coefficient. Multiplicative bias corrections can be interpreted mathematically as a linear transformation aY with $a > 0$ of the original time series Y. This translates into the hypothesis

$$corr(X, Y) = corr(X, aY) \tag{A.1}$$

The Pearson correlation coefficient is defined as

$$corr(X, Y) = \frac{cov(X, Y)}{\sigma(X)\sigma(Y)} \tag{A.2}$$

The linear transformation of the covariance can be simplified to

$$\begin{aligned} cov(X, aY) &= n^{-1}(X - \bar{X})(aY - a\bar{Y}) \\ &= an^{-1}(X - \bar{X})(Y - \bar{Y}) \\ &= acov(X, Y) \end{aligned} \tag{A.3}$$

and the standard deviation to

$$\begin{aligned} \sigma(aX) &= \sqrt{Var(ax)} \\ &= \sqrt{a^2 Var(X)} \\ &= \sqrt{a^2}\sqrt{Var(X)} \\ &= a\sigma(X) \end{aligned} \tag{A.4}$$

Finally, we can show that

$$corr(X, aY) = \frac{acov(X, Y)}{a\sigma(X)\sigma(Y)} = \frac{cov(X, Y)}{\sigma(X)\sigma(Y)} = corr(X, Y) \tag{A.5}$$

which is, quod erat demonstrandum, the Hypothesis Eq. (A.1) and finishes the proof.

References

[1] Edenhofer O, Madruga RP, Sokona Y, Seyboth K, Matschoss P, Kadner S, et al. Renewable energy sources and climate change mitigation: Special report of the intergovernmental panel on climate change. In: Renewable energy sources and climate change mitigation: special report of the intergovernmental panel on climate change. Cambridge University Press; 2011, p. 1–1075. http://dx.doi.org/10.1017/CBO9781139151153.

[2] IEA. International Energy Agency (IEA) World Energy Outlook 2022. Technical Report, 2022, https://www.iea.org/reports/world-energy-outlook-2022/executive-summary, URL https://www.iea.org/countries/egypt.

[3] Brown TW, Bischof-Niemz T, Blok K, Breyer C, Lund H, Mathiesen BV. Response to 'burden of proof: A comprehensive review of the feasibility of 100% renewable-electricity systems'. Renew Sustain Energy Rev 2018;92(September 2017):834–47. http://dx.doi.org/10.1016/j.rser.2018.04.113.

[4] Creutzig F, Agoston P, Goldschmidt JC, Luderer G, Nemet G, Pietzcker RC. The underestimated potential of solar energy to mitigate climate change. Nature Energy 2017;2(9):17140. http://dx.doi.org/10.1038/nenergy.2017.140.

[5] International Energy Agency I. World Energy Outlook 2018. Technical Report, 2018, URL www.iea.org/weo.

[6] Kalogirou SA. Solar Energy Engineering: Processes and Systems. Second Edition. 2014, p. 1–819. http://dx.doi.org/10.1016/C2011-0-07038-2.

[7] F. Holmgren W, W. Hansen C, A. Mikofski M. pvlib python: A python package for modeling solar energy systems. J Open Source Softw 2018;3(29):884. http://dx.doi.org/10.21105/joss.00884.

[8] Chouder A, Silvestre S, Taghezouit B, Karatepe E. Monitoring, modelling and simulation of PV systems using labview. Sol Energy 2013;91:337–49. http://dx.doi.org/10.1016/j.solener.2012.09.016.

[9] Dunnett S, Sorichetta A, Taylor G, Eigenbrod F. Harmonised global datasets of wind and solar farm locations and power. Sci Data 2020;7(1). http://dx.doi.org/10.1038/s41597-020-0469-8.

[10] Kruitwagen L, Story KT, Friedrich J, Byers L, Skillman S, Hepburn C. A global inventory of photovoltaic solar energy generating units. Nature 2021;598(7882):604–10. http://dx.doi.org/10.1038/s41586-021-03957-7.

[11] Yadav AK, Chandel S. Tilt angle optimization to maximize incident solar radiation: A review. Renew Sustain Energy Rev 2013;23:503–13. http://dx.doi.org/10.1016/j.rser.2013.02.027.

[12] Yu J, Wang Z, Majumdar A, Rajagopal R. DeepSolar: A machine learning framework to efficiently construct a solar deployment database in the United States. Joule 2018;2(12):2605–17. http://dx.doi.org/10.1016/j.joule.2018.11.021.

[13] Mayer K, Rausch B, Arlt M-L, Gust G, Wang Z, Neumann D, et al. 3D-PV-locator: Large-scale detection of rooftop-mounted photovoltaic systems in 3D. Appl Energy 2022;310:118469. http://dx.doi.org/10.1016/j.apenergy.2021.118469.

[14] Zech M, Ranalli J. Predicting PV areas in aerial images with deep learning. In: Conference record of the IEEE photovoltaic specialists conference, vol. 2020-June, Institute of Electrical and Electronics Engineers Inc.; 2020, p. 0767–74. http://dx.doi.org/10.1109/PVSC45281.2020.9300636.

[15] Hou X, Wang B, Hu W, Yin L, Wu H. SolarNet: A deep learning framework to map solar power plants in China from satellite imagery. In: ICLR 2020 workshop on tackling climate change with machine learning. 2020, URL http://arxiv.org/abs/1912.03685.

[16] Killinger S, Lingfors D, Saint-Drenan YM, Moraitis P, van Sark W, Taylor J, et al. On the search for representative characteristics of PV systems: Data collection and analysis of PV system azimuth, tilt, capacity, yield and shading. Sol Energy 2018;173:1087–106. http://dx.doi.org/10.1016/j.solener.2018.08.051.

[17] Hörsch J, Hofmann F, Schlachtberger D, Brown T. PyPSA-Eur: An open optimisation model of the European transmission system. Energy Strategy Rev 2018;22:207–15. http://dx.doi.org/10.1016/j.esr.2018.08.012, URL https://github.com/PyPSA/GridKit/.

[18] Müller UP, Schachler B, Scharf M, Bunke WD, Günther S, Bartels J, et al. Integrated techno-economic power system planning of transmission and distribution grids. Energies 2019;12(11). http://dx.doi.org/10.3390/en12112091.

[19] Scholz Y. Renewable Energy Based Electricity Supply at Low Costs - Development of the REMix Model and Application for Europe. Technical Report, 2012, p. 1–199, URL http://elib.uni-stuttgart.de/opus/volltexte/2012/7635/.

[20] Pfenninger S, Staffell I. Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data. Energy 2016;114:1251–65. http://dx.doi.org/10.1016/j.energy.2016.08.060.

[21] Brown T, Schlachtberger D, Kies A, Schramm S, Greiner M. Synergies of sector coupling and transmission reinforcement in a cost-optimised, highly renewable European energy system. Energy 2018;160:720–39. http://dx.doi.org/10.1016/j.energy.2018.06.222, http://arxiv.org/abs/1801.05290.

[22] Staffell I, Pfenninger S. The increasing impact of weather on electricity supply and demand. Energy 2018;145:65–78. http://dx.doi.org/10.1016/j.energy.2017.12.051.

[23] Collins S, Deane P, Ó Gallachóir B, Pfenninger S, Staffell I. Impacts of inter-annual wind and solar variations on the European power system. Joule 2018;2(10):2076–90. http://dx.doi.org/10.1016/j.joule.2018.06.020.

[24] Heuberger CF, Rubin ES, Staffell I, Shah N, Mac Dowell N. Power capacity expansion planning considering endogenous technology cost learning. Appl Energy 2017;204:831–45. http://dx.doi.org/10.1016/j.apenergy.2017.07.075.

[25] Grams CM, Beerli R, Pfenninger S, Staffell I, Wernli H. Balancing europe's wind-power output through spatial deployment informed by weather regimes. Nature Clim Change 2017;7(8):557–62. http://dx.doi.org/10.1038/NCLIMATE3338.

[26] Quoilin S, Gonzalez IH, Zucker A. The Dispa-SET 2.1 open-source model. Technical Report, 2017, http://dx.doi.org/10.2760/25400, URL https://ec.europa.eu/jrc.

[27] Gerbaulet C, Lorenz C. Dynelmod: A dynamic investment and dispatch model for the future European electricity market. In: Energy weekly news. 2017.

[28] Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. Pytorch: An imperative style, high-performance deep learning library. Adv Neural Inf Process Syst 2019;32.

[29] Paszke A, Gross S, Chintala S, Chanan G, Yang E, Facebook ZD, et al. Automatic differentiation in PyTorch. Technical Report, 2017.

[30] Hofmann F, Hampp J, Neumann F, Brown T, Hörsch J. Atlite: A lightweight python package for calculating renewable power potentials and time series. J Open Source Softw 2021;6(62):3294. http://dx.doi.org/10.21105/joss.03294.

[31] Sproul AB. Derivation of the solar geometric relationships using vector analysis. Renew Energy 2007;32(7):1187–205. http://dx.doi.org/10.1016/j.renene.2006.05.001.

[32] Reindl DT, Beckman WA, Duffie JA. Diffuse fraction correlations. Sol Energy 1990;45(1):1–7. http://dx.doi.org/10.1016/0038-092X(90)90060-P.

[33] Muñoz-Sabater J, Dutra E, Agustí-Panareda A, Albergel C, Arduini G, Balsamo G, Boussetta S, Choulga M, Harrigan S, Hersbach H, Martens B, Miralles DG, Piles M, Rodríguez-Fernández NJ, Zsoter E, Buontempo C, Thépaut JN. Era5-land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst Sci Data 2021;13:4349–83. http://dx.doi.org/10.5194/essd-13-4349-2021.

[34] Huld T, Gottschalg R, Beyer HG, Topič M. Mapping the performance of PV modules, effects of module type and data averaging. Sol Energy 2010;84(2):324–38. http://dx.doi.org/10.1016/j.solener.2009.12.002.

[35] Karatepe E, Boztepe M, Colak M. Neural network based solar cell model. Energy Convers Manage 2006;47(9–10):1159–78. http://dx.doi.org/10.1016/j.enconman.2005.07.007.

[36] Goodfellow I, Bengio Y, Courville A. Deep Learning, vol. 521, (7553):MIT Press; 2017, p. 785. http://dx.doi.org/10.1016/B978-0-12-391420-0.09987-X.

[37] Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data 2021;8(1). http://dx.doi.org/10.1186/s40537-021-00444-8.

[38] Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. Neural Netw 1989;2(5):359–66. http://dx.doi.org/10.1016/0893-6080(89)90020-8.

[39] Frank CW, Wahl S, Keller JD, Pospichal B, Hense A, Crewell S. Bias correction of a novel European reanalysis data set for solar energy applications. Sol Energy 2018;164:12–24. http://dx.doi.org/10.1016/j.solener.2018.02.012.

[40] Urraca R, Huld T, Gracia-Amillo A, Martinez-de Pison FJ, Kaspar F, Sanz-Garcia A. Evaluation of global horizontal irradiance estimates from ERA5 and COSMO-REA6 reanalyses using ground and satellite-based data. Sol Energy 2018;164(March):339–54. http://dx.doi.org/10.1016/j.solener.2018.02.059.

[41] Pfeifroth U, Kothe S, Trentmann J, Hollmann R, Fuchs P, Kaise J, et al. Surface radiation data set - heliosat (SARAH) - edition 2.1. 2019, http://dx.doi.org/10.5676/EUM{_}SAF{_}CM/SARAH/V002{_}01.

[42] Wiese F, Schlecht I, Bunke W-D, Gerbaulet C, Hirth L, Jahn M, et al. Open power system data – frictionless data for electricity system modelling. Appl Energy 2019;236:401–9. http://dx.doi.org/10.1016/j.apenergy.2018.11.097.

[43] Kingma DP, Ba JL. Adam: A method for stochastic optimization. In: 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings. 2015.

[44] Saint-Drenan Y-M. a probabilistic approach to the estimation of regional photovoltaic power generation using meteorological data (Ph.D. thesis), 2015.

[45] Hyndman RJ, Lee AJ, Wang E. Fast computation of reconciled forecasts for hierarchical and grouped time series. Comput Statist Data Anal 2016;97:16–32. http://dx.doi.org/10.1016/j.csda.2015.11.007.

[46] Yang D, Quan H, Disfani VR, Rodríguez-Gallegos CD. Reconciling solar forecasts: Temporal hierarchy. Sol Energy 2017;158:332–46. http://dx.doi.org/10.1016/j.solener.2017.09.055.

[47] Yang D, Quan H, Disfani VR, Liu L. Reconciling solar forecasts: Geographical hierarchy. Sol Energy 2017;146:276–86. http://dx.doi.org/10.1016/j.solener.2017.02.010.

[48] O'Shaughnessy E, Cruce JR, Xu K. Too much of a good thing? global trends in the curtailment of solar pv. Solar Energy 2020;208:1068–77. http://dx.doi.org/10.1016/j.solener.2020.08.075.