Data-driven mean-variability optimization of PV portfolios with automatic differentiation

Matthias Zech, Lueder von Bremen
German Aerospace Center
Institute of Networked Energy Systems
Carl-von-Ossietzky-Str. 15, Oldenburg, Germany
matthias.zech@dlr.de

Abstract

Increasing PV capacities has a crucial role to reach carbon-neutral energy systems. To promote PV expansion, policy designs have been developed which rely on energy yield maximization to increase the total PV energy supply in energy systems. Focusing on yield maximization, however, ignores negative side-effects such as an increased variability due to similar-orientated PV systems at clustered regions. This can lead to costly ancillary services and thereby reduces the acceptance of renewable energy. This paper suggests to rethink PV portfolio designs by deriving mean-variability hedged PV portfolios with smartly orientated tilt and azimuth angles. Based on a data-driven method inspired from modern portfolio theory, we formulate the problem as a biobjective, non-convex optimization problem which is solved based on automatically differentiating the physical PV conversion model subject to individual tilt and azimuth angles. To illustrate the performance of the proposed method, a case study is designed to derive efficient frontiers in the mean-variability spectrum of Germany's PV portfolio based on representative grid points. The proposed method allows decision-makers to hedge between variability and yield in PV portfolio design decisions. This is the first study highlighting the problem of ignoring variability in PV portfolio expansion schemes and introduces a way to tackle this issue using modern methods inspired by Machine Learning.

1 Introduction

The large-scale deployment of renewable energy is the key pillar to achieve carbon-neutral energy systems and as part of this transition, PV has a crucial role to provide affordable, clean energy with expected increasing expansion rates in the coming years to meet climate targets (IEA, 2022). In the past, countries have decided on a multitude of different incentive schemes to increase the individual's willingness to invest in PV systems. Feed-in tariffs are the most prominent policy tool in Europe and show a positive track to increase PV shares over the last years (Sun and Nie, 2015; Lipp, 2007). However, remunerating each generated energy unit equally, as achieved by feed-in tariffs, only focuses on maximizing annual PV yield and therefore assumes that annual yield is the single important metric. This leads, however, to similar PV system orientations at regions with large solar irradiance potential and therefore to clusters of highly correlated PV systems. Meteorological events relevant for solar irradiance, such as clouds and fog, then more likely affect the same clustered region of PV systems which leads to large PV feed-in variability. These events often can not be reliably forecasted which results into a large mismatch between planned and available solar energy in these clustered regions. Energy systems need to balance this mismatch through costly ancillary services within the electrical grid. Whereas there is a large body of literature describing solar energy variability of existing PV farms (Hoff and Perez, 2012; Perez et al., 2012; Ranalli and Peerlings, 2021; Lave and Kleissl, 2013; Lave et al., 2013;

Widén, 2015; Elsinga, 2017), there is to the best of the authors knowledge no literature yet about how variability can be integrated into the investors decision to hedge between yield and variability.

We argue that the design of decentralized, smart PV portfolios have the potential to reduce the portfolio variability maintaining large levels of yield. We therefore formulate the PV portfolio decision as a non-convex PV portfolio optimization problem inspired from the well-known mean-variance portfolio optimization problem. Based on the advances in Machine Learning to design efficient batch-wise optimization algorithms for non-convex optimization problems, we formulate the optimization problem including the non-convex PV conversion process in a data-driven fashion and automatically differentiate based on the azimuth and tilt angles which are the most important controllable parameters of the PV conversion process (Saint-Drenan, 2015). In a case study covering Germany, we show that the proposed method can derive well-balanced PV portfolios in terms of risk and reward. This paper aims to create awareness for smarter PV portfolios in the mean-variability spectrum to create system-friendlier solar energy feed-ins by introducing an easy-to-use method based on automatic differentiation which can be relevant for academia and policy designers.

2 Mean-variability optimization of PV capacity factors

2.1 PV portfolio simulation

To simulate PV systems, a physical model is required which describes the energy conversion from incoming solar irradiance to actual usable PV feed-ins. This includes the calculation of the total irradiance reaching the tilted surface of the PV system and how much of it the PV system can transform into electricity. For this purpose, we use a simple PV capacity factor model inspired by the `atlite` (Hofmann et al., 2021) library and translate the capacity factor model into the `pytorch` (Paszke et al., 2019) library to allow automatic differentiation. This avoids to manually derive any gradients and allows to benefit from the rich set of optimization algorithms implemented in `pytorch`. Further information on the formulation of the PV model is provided in the appendix. The output of the PV conversion model represents the maximum potential energy at given time t and site s which we refer to as the capacity factors $c_{t,s}$. The variables of interest within this study are the orientation of the PV system which includes the orientation (azimuth angle γ^{pv}) and the slope of the PV system (tilt angle β^{pv}).

2.2 Mean-variability optimization

The major goal of PV system investors is to derive PV configurations which maximize the annual PV yield and thereby maximize profits. This can be expressed by the total annual energy yield whereas we refer to the mean annual yield in alignment with mean-variance optimization literature. Maximizing annual yield at given site unsurprisingly received large attention in academic literature and heuristics based on particle swarm optimization, genetic algorithms and simulated annealing have been proposed (Yadav and Chandel, 2013). Considering variability is usually not part of the investment decision which also reflects in no mention in the academic literature about PV investment optimization. We therefore propose to consider variability in the expansion decision in accordance with the academic consensus of its crucial relevance for energy systems (Hoff and Perez, 2010; Ranalli and Peerlings, 2021; Perez et al., 2016; Lohmann and Monahan, 2018). Solar energy variability is defined in accordance with Hoff and Perez (2010) as the spread of the feed-in differences

$$\Delta c_{t,s} = \sigma(c_{t,s}(\beta_s^{pv}, \gamma_s^{pv}) - c_{t-1,s}(\beta_s^{pv}, \gamma_s^{pv})) \tag{1}$$

Note that maximizing the PV yield is a spatially separable optimization problem whereas optimizing variability is not due to the interaction between correlated sites. Maximizing annual yield and minimizing variability are conflicting objective functions as simple reflections can show: Without any electricity generation, there is no variability whereas only maximizing yield leads to similar PV configurations which amplifies overall variability due to accumulation. To find a good compromise between these two goals, we define a biobjective optimization problem which formulates

$$\min_{\delta_s,\xi_s} z = \left\{ (1-\lambda) \underbrace{\sigma(\Delta c_{t,s}(\beta_s^{pv}, \gamma_s^{pv}))}_{\textbf{variability}} - \lambda \underbrace{\frac{1}{TS} \sum_{t,s} c_{t,s}(\beta_s^{pv}, \gamma_s^{pv})}_{\textbf{mean}} \right\} \qquad \forall t \in \{2, 3, ..., T\}, \tag{2}$$

with decision variables $\beta_s^{pv} \in \{0°, 90°\}, \gamma_s^{pv} \in \{0°, 360°\}$. The parameter λ specifies the risk preference of the decision-maker meaning that $\lambda = 1$ reduces the optimization problem to a yield-maximization problem and $\lambda = 0$ to a variability-minimization problem. The premise of formulating the problem as a biobjective optimization problem is that between the competing objective terms, there are promising PV orientation combinations in terms of risk and reward. A single value of λ is difficult to determine as no knowledge about risk preference is available and scaling factors of the objective terms need to be considered. Testing multiple values of λ is a strategy to obtain more information about the yield-variability spectrum and allows to derive the *efficient frontier* which expresses the maximum possible yield at given risk.

Note that a brute-force approach to derive a robust PV portfolio by simply simulating all possible PV portfolio configurations quickly exceeds the computational limits. For this purpose, consider the discretized case of simulating 900 different tilt angles $\{0°, 0.1°, ..., 90°\}$ and 360 different azimuth angles $\{0°, 1°, .., 360°\}$ for 10,000 sites. This would require around 3 billion simulations which, when considering around 0.1 seconds per simulation, would take around 10 years to simulate.

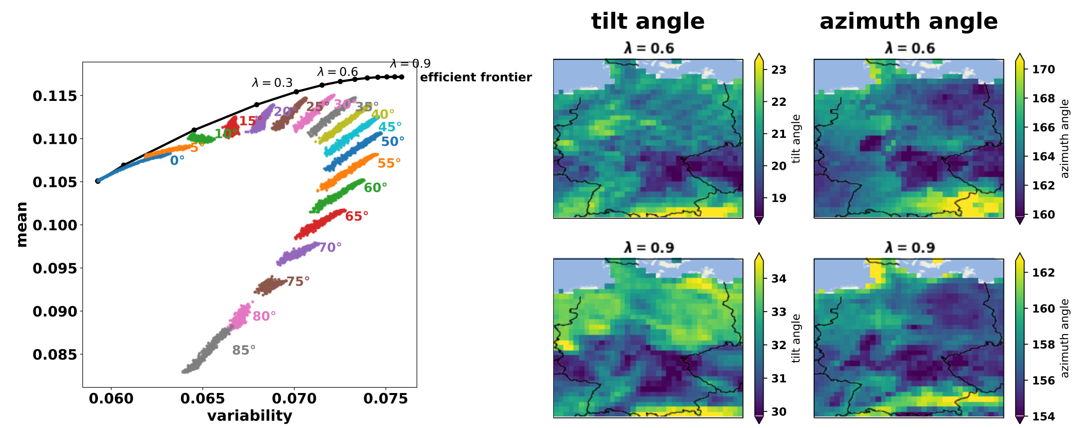
The proposed method contains the physical conversion process and does not rely on learning the actual PV conversion process by using, for instance, neural networks as in comparable studies Yadav and Chandel (2013). This has the advantage that the model remains physically valid and the optimization solution aligns with the solution obtained by running physical PV conversion simulations with the same parameters.

2.3 Relationship to modern portfolio theory

The proposed method is heavily inspired by the mean-variance optimization literature as introduced by Markowitz (1952) which is widely used to derive asset portfolios to maximize expected returns at given levels of risk. Both PV portfolios and stock portfolios estimate balanced portfolios based on historic performance, measured through historical weather years or stock performance respectively. Diversification can be achieved in asset portfolios through uncorrelated assets which would translate in the PV context to distant PV systems with different orientations. Mathematically, mean-variance optimization in the context of asset portfolio optimization is formulated as a convex optimization problem whereas the physical model of the PV system is non-convex making the problem much harder to solve. The proposed method is able to derive, however, sophisticated solutions as shown in the case study despite no global optimality can be guaranteed.

3 Case Study: Germany's PV portfolio

The mean-variability optimization method is evaluated based on a case study covering Germany's PV portfolio. As input data, solar irradiance data is retrieved from the satellite-based SARAH-2 data (Pfeifroth et al., 2019) and temperature data from the ERA5 reanalysis model (Hersbach et al., 2020). To match satellite with reanalysis data, the satellite data is temporally averaged to hourly values and spatially to the ERA5 grid ($0.25° \times 0.25°$) which represents approximately 30 km distance. When applied on Germany, this translates into 1147 sites as representative grid points s. The training period covers 5 years (2012-2016) whereas one batch covers one entire year with 8760 hours. The results are evaluated on the left-out year 2017. The optimization problem is solved using the Adam optimizer (Kingma



(a) Efficient frontier and random simulations (100,000 for each tilt angle) to illustrate the mean-variability spectrum

(b) Tilt and azimuth angle distribution as output from the mean-variability optimization for different values of λ

Figure 1: Performance and result of the proposed method for the test year 2017

and Ba, 2015) with a cyclical learning rate derived from the range test as proposed by Smith (2015).

Performance of randomly sampled portfolios: To compare the optimization results to alternative PV portfolios, random PV portfolios are generated around fixed tilt angles. Randomly sampling around near-optimal angles is a common method to reproduce historical feed-in time series due to missing site information (Pfenninger and Staffell, 2016). Therefore, for each fixed tilt angle $(0°, 5°, ..., 90°)$, 10,000 PV configurations are sampled from a normal distribution with varying standard deviations for tilt and azimuth angles between 0 and 20° which approximately aligns with distributions of existing PV systems in Germany (Killinger et al., 2018). The randomly simulated portfolios are illustrated in Figure 1a) which leads to a boomerang-alike shape of the sampled PV portfolios. More specifically, low tilt angles $(< 45°)$ show large mean yield with moderate variability whereas large tilt angles $(> 50°)$ show a steep decrease of mean annual yield while remaining high variability levels. Flat PV portfolios are therefore more preferable in the mean-variability spectrum compared to steeper tilt angles.

Performance of the proposed method: The black line in Figure 1a) shows the efficient frontier which illustrates the best trade-off between risk and reward derived by running the proposed method with a range of different risk preferences $(\lambda \in \{0, 0.05, ..., 1.\})$. The efficient frontier always sits above the randomly simulated points which shows that the automatic differentiation is constantly able to derive better configurations of PV tilt and azimuth angle orientations. This is particularly evident at mean capacity factors above 0.11 at which the distance between the best sampled portfolio and the efficient frontier increases. The course of the efficient frontier shows that the mean-variability spectrum is valuable to be exploited even for small spatial extents such as Germany. For example, the comparison of $\lambda = 0.6$ with $\lambda = 0.9$ only looses 1.5% of the mean yield while reducing variability around 6.5%.

Interpretation of the tilt and azimuth angle distributions: Figure 1b) shows the tilt and azimuth angle distributions for $\lambda = 0.6$ and $\lambda = 0.9$. The *azimuth angles* for both cases show similar spatial distributions, yet the lower variability portfolio $(\lambda = 0.6)$ shows a more pixelated situation with a slightly larger range of azimuth angles (160°-170° vs. 154°-162°) and therefore a larger spread of orientations. There is a shift towards east facing systems noticeable which may indicate a better exploitation of morning hours than evening hours which is physically explainable due to the rising thermals from solar heating over the course of the day which leads to often more cloudy afternoons than mornings. The *tilt angle* distribution between both mean-variability hedged portfolios show different, distinct patterns. The low-variability portfolio $(\lambda = 0.6)$ shows structures similar to waves particularly in the northern part of Germany where no large changes of topography are. A plausible explanation for this pattern is that in case of western wind conditions, which are dominant in Western Europe, the determined distribution of tilt angles shown in the upper left image of Figure 1b) is then orthogonal to this wind direction. Cloud conditions then affect a region with a larger spread of tilt angles which smoothes the variability. The larger-variability portfolio ($\lambda = 0.9$) illustrates no such waves and therefore less smoothing and larger variability.

4 Concluding discussion

This paper proposes a novel method to estimate robust PV portfolios with smartly selected tilt and azimuth angles by exploiting the mean-variability spectrum of PV feed-ins. We show that making the connection between yield and variability can be valuable even for smaller regions such as Germany as only small yield losses are noticeable while reducing the variability largely.

A limitation of this study is that no global optimality can be guaranteed as the problem is non-convex. Further research avenues may be to benchmark the solution against alternative algorithms in the literature to derive tilt angles to maximize annual yield as described in Section 2.2. Furthermore, more meteorological understanding is needed about the relationship between tilt and azimuth angles and the underlying meteorological phenomena. We have provided plausible physical explanations, yet more scientific significance is needed. A further promising research avenue is the consideration of different capacities at the sites as an additional degree of freedom.

References

IEA. International Energy Agency (IEA) World Energy Outlook 2022. Technical report, 2022. URL `www.iea.org/tc/`.

Peng Sun and Pu yan Nie. A comparative study of feed-in tariff and renewable portfolio standard policy in renewable energy industry. *Renewable Energy*, 74:255–262, 2015. ISSN 09601481. doi: 10.1016/j.renene.2014.08.027. URL `http://dx.doi.org/10.1016/j.renene.2014.08.027`.

Judith Lipp. Lessons for effective renewable electricity policy from Denmark, Germany and the United Kingdom. *Energy Policy*, 35(11):5481–5495, 2007. ISSN 03014215. doi: 10.1016/j.enpol.2007.05.015.

Thomas E. Hoff and Richard Perez. Modeling PV fleet output variability, 8 2012. ISSN 0038092X.

Richard Perez, Sergey Kivalov, Jim Schlemmer, Karl Hemker, and Thomas E. Hoff. Short-term irradiance variability: Preliminary estimation of station pair correlation as a function of distance, 8 2012. ISSN 0038092X.

Joseph Ranalli and Esther E.M. Peerlings. Cloud advection model of solar irradiance smoothing by spatial aggregation. *Journal of Renewable and Sustainable Energy*, 13(3), 2021. ISSN 19417012. doi: 10.1063/5.0050428.

Matthew Lave and Jan Kleissl. Cloud speed impact on solar variability scaling - Application to the wavelet variability model. *Solar Energy*, 91:11–21, 2013. ISSN 0038092X. doi: 10.1016/j.solener.2013.01.023. URL `http://dx.doi.org/10.1016/j.solener.2013.01.023`.

Matthew Lave, Jan Kleissl, and Joshua S. Stein. A wavelet-based variability model (WVM) for solar PV power plants. *IEEE Transactions on Sustainable Energy*, 4(2):501–509, 2013. ISSN 19493029. doi: 10.1109/TSTE.2012.2205716.

Joakim Widén. A model of spatially integrated solar irradiance variability based on logarithmic station-pair correlations. *Solar Energy*, 122:1409–1424, 2015. ISSN 0038092X. doi: 10.1016/j.solener.2015.10.043.

Boudewijn Elsinga. *Chasing the Clouds: Irradiance Variability and Forecasting for Photovoltaics*. PhD thesis, 2017.

Yves-Marie Saint-Drenan. *A Probabilistic Approach to the Estimation of Regional Photovoltaic Power Generation using Meteorological Data*. PhD thesis, 2015.

Fabian Hofmann, Johannes Hampp, Fabian Neumann, Tom Brown, and Jonas Hörsch. atlite: A Lightweight Python Package for Calculating Renewable Power Potentials and Time Series. *Journal of Open Source Software*, 6(62):3294, 2021. ISSN 2475-9066. doi: 10.21105/joss.03294.

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Köpf, Edward Yang, Zach DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. PyTorch: An imperative style, high-performance deep learning library. *Advances in Neural Information Processing Systems*, 32, 12 2019. ISSN 10495258. URL `http://arxiv.org/abs/1912.01703`.

Amit Kumar Yadav and S.S. Chandel. Tilt angle optimization to maximize incident solar radiation: A review. *Renewable and Sustainable Energy Reviews*, 23:503–513, 7 2013. ISSN 13640321. doi: 10.1016/j.rser.2013.02.027.

Thomas E. Hoff and Richard Perez. Quantifying PV power Output Variability. *Solar Energy*, 84(10):1782–1793, 2010. ISSN 0038092X. doi: 10.1016/j.solener.2010.07.003. URL `http://dx.doi.org/10.1016/j.solener.2010.07.003`.

Richard Perez, Mathieu David, Thomas E. Hoff, Mohammad Jamaly, Sergey Kivalov, Jan Kleissl, Philippe Lauret, and Marc Perez. Spatial and Temporal Variability of Solar Energy. *Foundations and Trends® in Renewable Energy*, 1(1):1–44, 2016. ISSN 2328-8892. doi: 10.1561/2700000006.

Gerald M. Lohmann and Adam H. Monahan. Effects of temporal averaging on short-term irradiance variability under mixed sky conditions. *Atmospheric Measurement Techniques*, 11(5):3131–3144, 2018. ISSN 18678548. doi: 10.5194/amt-11-3131-2018.

Harry Markowitz. Portfolio Selection. *The Journal of Finance*, 7(1):77–91, 1952. ISSN 15406261. doi: 10.1111/j.1540-6261.1952.tb01525.x.

U. Pfeifroth, S. Kothe, J. Trentmann, R. Hollmann, P. Fuchs, J. Kaise, and M. Werscheck. Surface Radiation Data Set - Heliosat (SARAH) - Edition 2.1, 2019.

Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, Adrian Simmons, Cornel Soci, Saleh Abdalla, Xavier Abellan, Gianpaolo Balsamo, Peter Bechtold, Gionata Biavati, Jean Bidlot, Massimo Bonavita, Giovanna De Chiara, Per Dahlgren, Dick Dee, Michail Diamantakis, Rossana Dragani, Johannes Flemming, Richard Forbes, Manuel Fuentes, Alan Geer, Leo Haimberger, Sean Healy, Robin J. Hogan, Elías Hólm, Marta Janisková, Sarah Keeley, Patrick Laloyaux, Philippe Lopez, Cristina Lupu, Gabor Radnoti, Patricia de Rosnay, Iryna Rozum, Freja Vamborg, Sebastien Villaume, and Jean Noël Thépaut. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, 2020. ISSN 1477870X. doi: 10.1002/qj.3803.

Diederik P. Kingma and Jimmy Lei Ba. Adam: A method for stochastic optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 12 2015. URL `http://arxiv.org/abs/1412.6980`.

Leslie N. Smith. Cyclical Learning Rates for Training Neural Networks. 6 2015. URL `http://arxiv.org/abs/1506.01186`.

Stefan Pfenninger and Iain Staffell. Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data. *Energy*, 114:1251–1265, 2016. ISSN 03605442. doi: 10.1016/j.energy.2016.08.060.

Sven Killinger, David Lingfors, Yves Marie Saint-Drenan, Panagiotis Moraitis, Wilfried van Sark, Jamie Taylor, Nicholas A. Engerer, and Jamie M. Bright. On the search for representative characteristics of PV systems: Data collection and analysis of PV system azimuth, tilt, capacity, yield and shading. *Solar Energy*, 173:1087–1106, 2018. ISSN 0038092X. doi: 10.1016/j.solener.2018.08.051.

Alistair B. Sproul. Derivation of the solar geometric relationships using vector analysis. *Renewable Energy*, 32(7):1187–1205, 2007. ISSN 09601481. doi: 10.1016/j.renene.2006.05.001.

D. T. Reindl, W. A. Beckman, and J. A. Duffie. Diffuse fraction correlations. *Solar Energy*, 45(1):1–7, 1990. ISSN 0038092X. doi: 10.1016/0038-092X(90)90060-P.

Thomas Huld, Ralph Gottschalg, Hans Georg Beyer, and Marko Topič. Mapping the performance of PV modules, effects of module type and data averaging. *Solar Energy*, 84(2): 324–338, 2 2010. ISSN 0038092X. doi: 10.1016/j.solener.2009.12.002.

A PV capacity factor model

To calculate the total irradiance on a tilted plane, we use a simple trigonometric model based on Sproul (2007). The total irradiance I^T_{glob} on a tilted plane consists of the direct I^T_{Dir} and diffuse irradiance I^T_{Dif}

$$I^T_{glob} = I^T_{Dir} + I^T_{Dif} \tag{3}$$

The direct irradiance on a tilted surface I^T_{Dir} can be derived in accordance with Sproul (2007) following

$$I^T_{Dir} = I^N_{Dir} \cos\theta \tag{4}$$

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 (2007), the last term can be calculated from

$$\cos\theta = \sin\beta^{pv}\cos\alpha^s\cos(\gamma^{pv} - \gamma^s) + \cos\beta\sin\alpha^s \tag{5}$$

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. (1990) through

$$I^T_{Dif} = \frac{1+\cos\beta^{pv}}{2} I^N_{Dif} + \rho I^N_{Glob} \frac{1-\cos\beta^{pv}}{2} \tag{6}$$

with I^{Glob}_n total influx on a horizontal plane, ρ the ground albedo which is derived in accordance with Hofmann et al. (2021) through

$$\rho = \frac{I^N_{Glob} - I^N_{Net}}{I^N_{Glob}} \tag{7}$$

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 (Hersbach et al., 2020) used in this study, and allows a more accurate description of the ground albedo than commonly used constant values as in comparable studies Pfenninger and Staffell (2016).

The tilted irradiance is next translated into PV energy using the PV module performance model from Huld et al. (2010) which only depends on the module temperature and the in-plane irradiance through

$$P(I^T_{glob}, T_{mod}) = P_{STC} \frac{I^T_{glob}}{I_{STC}} \eta_{rel}(I', T') \tag{8}$$

with P_{STC} indicating the power at standard test conditions (STC) of $I_{STC} = 1000W/m^2$ and $T_{mod,STC} = 25$ respective temperature. The relative efficiency η_{rel} is calculated as proposed by Huld et al. (2010)

$$\eta_{rel}(I', T') = 1 + k_1 \ln I' + k_2[\ln I']^2 + T'(k_3 + k_4 \ln I' + k_5[\ln I']^2) + k_6 T'^2 \tag{9}$$

where I' and T' are normalized parameters to STC values with $I' = I^T_{glob}/1000$ and $T' = T_a + 0.035G - 298.15$ (Huld et al., 2010). The parameters $k_1, ..., k_6$ are empirical coefficients which are fitted in in practice to modules (Huld et al., 2010) but in this study are based on the standard carbon silicon PV module from the `atlite` library Hofmann et al. (2021). The last step in the physical model, the energy is reduced by the inverter efficiencies which are assumed to be 90% in accordance with Hofmann et al. (2021).

Within this study all the described parameters are fixed except the tilt angle β^{pv} and the azimuth angle γ^{pv} in Equation 5.