QUALIFIED FORECAST OF ENSEMBLE POWER PRODUCTION BY SPATIALLY DISPERSED GRID-CONNECTED PV SYSTEMS

Elke Lorenz*, Johannes Hurka*, Giota Karampela*, Detlev Heinemann*, Hans Georg Beyer +, Martin Schneider°

* University of Oldenburg, Institute of Physics, Energy and Semiconductor Research Laboratory,

Energy Meteorology Group, D-26111 Oldenburg, Germany, elke.lorenz@uni-oldenburg.de

+ University of Applied Sciences (FH) Magdeburg-Stendal, Institute of Electrical Engineering,

Breitscheidstr.2 D-39114, Magdeburg, Germany

° Meteocontrol GmbH, Spicherer Straße 48, D-86157 Augsburg, Germany

ABSTRACT: The contribution of power production by Photovoltaic (PV) systems to the electricity supply is constantly increasing. An efficient use of the fluctuating solar power production will highly benefit from forecast information on the expected power production. This forecast information is necessary for the management of the electricity grids and for energy trading.

This paper presents an approach to predict regional PV power output based on irradiance forecasts provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). In the first part of the paper we introduce and evaluated different approaches to refine the irradiance forecasts. The second part of the paper addresses the power prediction for ensembles of PV systems. Here, in view of the data handling problems associated with the high number of individual systems contributing to the total PV generation within a region, the identification of representative subsets and representative system characterizations reflecting the power characteristics of the total ensemble is discussed.

Keywords: PV system, grid-connected, solar radiation, forecasting

1 INTRODUCTION

Power generation from solar energy systems is highly variable due to its dependence on meteorological conditions. An efficient use of this fluctuating energy source requires reliable forecast information for management and operation strategies. Today, wind power prediction systems have already shown their strong economic impact and improve the integration of wind energy into the electricity grid (see e.g. [1]). Accordingly, the prediction of solar yields will become more and more important for utilities, which have to integrate increasing amounts of solar power, especially for countries where legislation encourages the deployment of solar power plants. The Spanish feed-in law already includes incentives for correct predictions of solar yields for the next day. In Germany the installed photovoltaic (PV) power amounts to more than 3 GW. In Bavaria, where a large number of these PV systems is installed, the share of solar power injected on power grids can reach around 10 % during peak hours on sunny days.

In this paper we present further developments of an approach to forecast the hourly regional PV power production, presented in [2]. A focus will be on further elaboration of the forecast of ensemble power production for PV-Systems within sections of the untidy network.

In addition we address a problem arising when aiming on a broad application of the scheme. The analysis in [2] was based on ensembles of selected PV systems, where all system parameters necessary for simulation were given. Under real conditions, the knowledge on the power production of all PV systems that contribute to a control area of the supply grid is necessary. Here, the problem arises that in general no detailed information on all PV systems for a control area is available. Especially for smaller PV systems that largely contribute to the overall power production the necessary parameters for PV simulation (coordinates, orientation and tilt angle of the systems, characteristics of inverter and modules) will hardly be available. Statistical methods to select representative systems that adequately characterize the actually given ensemble will be presented. These investigations are based on a selected database of 460 operating PV systems in Germany.

The next section presents the empirical data sets – for both, irradiance and power output - used for model improvements and model tests, followed by an introduction of the statistical measures used in that validation. The subsequent section gives a short description of the basic irradiance forecasting scheme, its recent refinements and performance. Following, the scheme for the derivation of the power output of the PVsystems is introduced with special focus on the handling of incomplete information on the systems. An evaluation of the power prediction using this representative system model is given, including a detailed analysis of the quality of ensemble power prediction. The last section addresses the upscaling of the power prediction of a subset of systems in order to represent the power production of the complete ensemble.

2 GROUND MEASURED DATA

2.1 Irradiance data

For these investigations, measured irradiance data of more than 200 meteorological stations in Germany were available, partly operated by the German Weather service DWD and partly operated by Meteomedia GmbH. The distribution of these stations in Germany is shown in Figure 1. The period of evaluation is 1.1.2007 to 31.10.2007.

2.2 Measurement of power output

PV power forecasts were evaluated using hourly power output data for 460 PV systems with a total installed power of about 28 MW, distributed over Germany as shown in Figure 2. The evaluation was performed for the months April to October of the year 2007.



Figure 1: Distribution of the meteorological stations.



Figure 2: Distribution of PV systems. Red dots represent the complete data base, blue squares show 11 PV systems with complete system information. The green box approximates the area of the transmission system operator EnBW.

The forecast quality is investigated not only for single PV systems but also for ensembles of PV systems, as the ensemble production of all PV systems contributing to a control area is of interest for the utility companies. The accuracy assessment was performed for the complete German ensemble and for an area in Southern Germany approximating the control area of the German transmission system operator EnBW. In this area a large number of PV systems is installed.

3 MEASURES OF MODEL ACCURACY

We use the root mean square error

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_{forecast,i} - x_{measured,i})^2}$$
(1)

as a main score for the assessment of prediction accuracy. Here, N is the number of data pairs evaluated. The variable x is replaced by I_{glob} for the evaluation of the prediction of the global irradiance, for the evaluation of the power forecast x is set to P_{AC}/P_{nom} . The normalisation of the alternating current (AC) power output P_{AC} to the nominal power P_{nom} of the PV system allows for a better comparison of different PV systems.

In addition, we use the mean value of the errors:

$$BIAS = \frac{1}{N} \sum_{i=1}^{N} (x_{forecast,i} - x_{measured,i})$$
(2)

to describe systematic deviations of the forecast.

The accuracy measures are calculated for hourly values. Only hours with daylight ($I_{glob,i} > 0$) are considered for the calculation of *BIAS* and *RMSE*, night values with no irradiance are excluded from the evaluation.

Relative values of the error measures (*rRMSE*, *rBIAS*) are obtained by normalization to the mean ground measured irradiance or PV power production of the considered period.

4 IRRADIANCE FORECAST

4.1 Basic forecast information

The scheme to derive predictions of PV power output is based on irradiance and temperature forecasts up to 3 days ahead provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). The global model run by the ECMWF provides these forecasts with a temporal resolution of 3 hours and a spatial resolution of $0.25^{\circ} \times 0.25^{\circ}$. [3] provides a description of physical process of the implementation of the ECMWF model during the evaluation period.

4.2 Refinement of the forecasts

We have investigated different approaches to derive an optimized hourly and site-specific irradiance forecast, that are shortly described in the following. A more detailed description is given in [4].

In a first step, a spatial averaging procedure is applied. An analysis of the forecast accuracy in dependence on the area of averaging revealed that best results are achieved with a region of approx. 100 km x 100 km.

For the temporal interpolation two different approaches are considered.

- A very simple approach to derive hourly resolved forecast is the linear interpolation of the three hourly mean values that are provided by the ECMWF (version V1).
- To better account for the diurnal course of the irradiance we combine the forecast data with a clear sky model ([5]). The temporal interpolation is performed for the clear sky index *kt**, characterizing the transmission of irradiance through the atmosphere and defined as ratio of the global

irradiance I_{glob} to the modelled irradiance for clear sky conditions $I_{clearsky}$ (version V2).

Figure 3 shows the *RMSE* and the *BIAS* over the predicted clear sky index kt^* for the two different versions of temporal interpolation. For situations with no or few clouds predicted ($kt^* \approx 1$) the combination with a clear sky model (V2, red line) significantly improves forecast quality compared to simple linear interpolation (V1, black line). The forecast quality is high for these situations with an *rRMSE* of less than 15 %.

For cloudy situations a systematic overestimation of the irradiance is found with clear sky index values between $kt^*=0.3$ and $kt^*=0.8$.

In order to avoid these systematic deviations, we introduce a situation specific bias correction. The *BIAS* is modelled as a function of the clear sky index kt^* and the solar zenith angle θ_z . The final version of the forecast (version V3) is obtained by subtracting the modelled *bias*(kt^*, θ_z) from the predicted values obtained with version V2.

To allow for independent testing of this adaptation to ground data, the data were divided into training and test data. The first 15 days of each month form the training set used to derive the fit-function, the remaining data were used for the evaluation.



Figure 3: Forecast errors (rRMSE and rBIAS) for hourly global irradiance I_{glob} over clear sky index kt^* for the different forecasting approaches. (V1: linear interpolation of I_{glob} , V2: linear interpolation of kt^* , V3: linear interpolation of kt^* and bias correction). Solid lines represent the *RMSE* and dashed lines show the *BIAS*.

The systematic overestimation of the irradiance for cloudy situations with clear sky index values between $kt^{*}=0.3$ and $kt^{*}=0.8$ is adjusted with the situations specific bias correction also on the test set, as shown in Figure 3 (V3, light grey line). Still, the *RMSE* is high for these situations that are often related to variable cloud cover.

4.3 Overall evaluation of prediction accuracy

An overall evaluation of the forecast accuracy for the different approaches in dependence on the forecast horizon is given in Figure 6, where the results on the test data set are displayed in dependence on the forecast day. Results for the first forecast day integrate forecast horizons up to 24 hours, the second forecast day integrates forecast horizons from 25 hours to 48 hours, and the third day includes forecast horizons from 49

hours to 72 hours. The upper image shows the results for single stations, in the lower image the mean irradiance of all stations was evaluated.



Figure 4: Relative forecast errors for hourly global irradiance I_{glob} in dependence on the forecast horizon for the different forecasting approaches. (V1: linear interpolation of I_{glob} V2: linear interpolation of kt^* , V3: linear interpolation of kt* and bias correction). Coloured bars represent the *rRMSE*, the respective white bars show the *rBIAS*. Upper image: single station. Lower image: ensemble of all stations in Germany.

Best results are achieved with approach V3. While for single sites the differences between the three approaches are only small, for regional forecasts a significant improvement is achieved by correction of systematic deviations. All following results and the PV power forecasts are derived using approach V3.

For single sites the *rRMSE* is 36.9 % for the first forecast day, and increases to 46.3 % for the third forecast day. Due to spatial averaging effects the forecast accuracy for the average irradiance of an ensemble of distributed stations is much higher than for a single system. The *rRMSE* for the ensemble of all stations amounts to 13.4 % for the first forecast day, for the third forecast day the *rRMSE* increases to 22.5 %.

To illustrate the forecast quality for single stations and ensembles, respective time series of predicted and measured irradiance are given in Figure 5. Confidence intervals indicate the maximum expected uncertainty for a given situation in dependence on the cloud situations and solar elevation. The method to derive these situationspecific confidence intervals is described and evaluated in [4].



Figure 5: Forecast of global irradiance I_{glob} with confidence intervals compared to measured irradiance for six days in May 2007 for one site (upper image) and for the ensemble of all stations shown in Figure 1 (lower image).

5 PREDICITION OF POWER OUTPUT

5.1. PV system model

From the site-specific, hourly forecasts of the global horizontal irradiance, the effective irradiance on the module plane of the PV-modules is derived using the anisotropic-all-sky model formulated by [6]. Based on these values a PV simulation model ([7]) provides the power forecasts. This model is based on a characterization of the PV-efficiency as function of irradiance and module temperature:

$$\eta_{MPP}(I_{POA}, T_m) = [a_1 + a_2 I_{POA} + a_3 \ln(I_{POA})][1 - \alpha(T_m - 25^{\circ}C)]$$
(3) with:

$$T_m = T_a + \gamma I_{PO}$$

 I_{POA} : irradiance on module plane in W/m² T_a : ambient temperature in °C T_m : amodule temperature in °C $a_1, a_2, a_3, \alpha, \gamma$: module specific parameters.

With the efficiency and nominal power of the generator (standard test conditions, STC), the DC-output of the modules of the generator can be calcualted. Using a standard characterization of the ohmic DC-losses, the inverter efficiency according to [8], AC-losses and reasonable assumptions on miscellaneous other losses [9] the forecasted AC power output of a generator is determined. The applicability of this modeling scheme has been analyzed in [9].

5.2. Representative system parameters

In [2] we have presented an evaluation of power output prediction using the described PV simulation model for a small ensemble of systems with complete system description available. The system parameters and installations details could be derived from the respective data sheets and documentations.

In general this information is not available for the majority of the systems. In addition, the derivation and management of the system specific parameters is cumbersome when aiming at the correct representation of all systems within a network region. Thus, it would be desirable to identify a unique set of parameters for the system components (PV-modules, inverters) characterising a representative system performance (given by its efficiency versus irradiance and temperature), aiming at reasonable reflection of the ensemble performance.

A respective analyses is reported in [10]. Using a case study involving 11 well described systems, it could be shown, that a representative system may be defined by a set of parameters reflecting the average values of the PV efficiency (see Figure 6) and the average efficiencies at 3 part-load levels of the inverters.



Figure 6: Normalized MPP efficiency curves (25°C) for a set of modules (blue lines) and the representative efficiency curve (red lines) reflecting the average efficiency values at non STC-irradiances.

In the same way the average of all temperature coefficients could be applied.

This representative model for the maximum power point efficiency is used to forecast the power output for the complete ensemble in combination with information on the installed power of the single systems and the respective irradiance forecasts for the given sites.

For the final version of the power forecast, the given values of installed power are adjusted with the average power output of the systems during the complete period of evaluation:

$$P_{pred,i} = P_{pred,0,i} \frac{mean(P_{meas,i})}{mean(P_{pred,0,i})}.$$
(4)

The information on the annual power production of the single PV systems is available at the utility companies, while measurements with hourly resolution usually are not performed for the majority of small PV systems.

5.3. Results of power prediction

An evaluation of the described procedure to predict power output based on the irradiance forecast with situation specific bias correction (V3) is given in Fig. 7. The Figure shows the *rRMSE* and *rBIAS* of the original power forecast $P_{pred,0}$ (dark green bars), the power forecast with adjustment of the installed power P_{pred} (light green bars), and the irradiance forecast (yellow bars) in comparison. The evaluation of the irradiance and the power forecast is performed using the two different data sets introduced in section 2. The respective error measures are given for single stations, the average of the ensemble approximately covering the area of the utility company EnBW in Southern Germany (see Figure 2), and the average of the German ensemble.



Figure 7: Relative errors of original and scaled power output prediction, ($P_{pred,0}$ and P_{pred}) and global irradiance forecast I_{glob} for single systems, an ensemble covering the area of EnBW, and the German ensemble (see Figure 2). Coloured bars represent the *rRMSE*, the respective white bars show the *rBIAS*.

Figure 7 shows that the adjustment of the installed power to measured mean values, not only reduces the *rBIAS* but also the *rRMSE* of the power forecast. All following results are given for the power forecast with adjustment of the installed power P_{pred} .

For single stations the *rRMSE* of the power forecast is dominated by the *rRMSE* of the irradiance forecast. The slightly larger error of the power forecast is mainly due to the conversion of the irradiance on the tilted plane and due to imperfect PV system description. The absolute error of the final version of the power forecast amounts to $RMSE_P = 112 \text{ W/kW}_{peak}$.

For the ensemble covering the area of EnBW, with a size of approximately 220 km x 220 km, the *rRMSE* of the irradiance forecast is reduced with the error reduction factor $f=RMSE_{ensemble}/RMSE_{single}=0.56$ and the *rRMSE* of the power forecast is reduced by f=0.61. This corresponds to an absolute $RMSE_P$ of 63 W/kW_{peak}.

For the German ensemble the quality of the irradiance forecast is further increasing in comparison to the smaller ensemble (f=0.35), while $RMSE_P = 52W/kW_{peak}$ of the power forecast for the German ensemble is only slightly smaller than for the Southern German Ensemble (f=0.51).

This difference in the error reduction factor for irradiance and power forecast is analyzed in detail in the next section.

6 ANALYSIS OF ERROR REDUCTION FOR ENSEMBLES OF SYSTEMS

The reduction of prediction errors, when considering an ensemble of systems instead of a single system, is determined by the characteristics of the ensemble (number of systems, area, distribution of systems and installed power) and by the cross correlation of forecast errors, both aspects are addressed in the following.

6.1. Cross correlation of forecast errors

A statistical model to estimate the error reduction factor for a given ensemble in dependence of the cross correlation of forecast errors has been introduced in [11]. With decreasing cross correlation of the forecast errors of the systems of the ensemble, the accuracy of the prediction of the ensemble power output is increasing.

The correlation coefficient of the forecast errors of two systems depends on the distance between the systems. This is illustrated in Figure 8, where the correlation coefficient of the power prediction errors between two systems is displayed over the distance between the systems. The dependence of the correlation of the forecast errors on the distance of two systems can be modelled with an exponential function as proposed in [11]. The light green line represents the model curve for power forecast errors. For comparison the model curve for irradiance forecast errors is also shown (red line).

Figure 8 shows that the correlation of errors for the power prediction is larger than for the irradiance forecast. This partly explains that power prediction errors are not reduced as much as irradiance forecast errors for an ensemble covering the same area.



Figure 8: Correlation coefficient of forecast errors of two systems over the distance between the systems. The blue dots represent measured values for the power output P, the light green line gives the model curve for P, the red line shows the model curve for global irradiance I_{glob} .

6.2. Regional distribution of sites

The influence of the regional distribution of sites on the quality of ensemble power prediction was investigated using subsets of 50 systems. Two different cases were considered. In the first case the regional distribution of stations in the subsets should reflect the regional distribution of the complete ensemble, as a second case we investigated subsets with an almost uniform distribution all over Germany. For both cases 50 subsets with 50 stations were selected and the mean RMSE for these subsets was calculated. For the first case a random selection of the systems was performed. For the second case, in order to achieve an approximation to a uniform distribution, the complete area of evaluation was divided to five regions of equal size and ten systems were selected randomly for each region.

For the subsets representing the distribution of the complete ensemble with a large share of systems in the South of Germany (see Figure 2) the average RMSE amounts to $60 W/kW_{peak}$. For the subsets with uniform distribution a smaller average *RMSE* of 50 W/kW_{peak} is found.

These results demonstrate that the regional distribution of stations has a significant influence on the quality of the power prediction for an ensemble of stations. The different distributions of irradiance measurement stations and PV systems according to Figure 1 and Figure 2 contribute to the difference in the error reduction factors for irradiance and power forecast.

A comparison of the regional distribution of the PV systems in the evaluation data base to the real distribution of PV systems in Germany based on data published in [Photon, 2007] revealed, that in the evaluation data set the installed power in Northern Germany is strongly underrepresented. Hence, for an ensemble reflecting the real distribution of PV systems in Germany, a further improvement compared to the results for the present evaluation set can be expected.

7 UPSCALING

In this section the upscaling of the power output prediction of a subset of systems in order to represent the power production of the complete ensemble is addressed. The representative subset should reflect the distribution of systems and their orientation (azimuth orientation and tilt angle of generator) and show a similar response to the irradiance conditions as the full ensemble.

In a first step, we investigated the *RMSE* of the ensemble power prediction for a subset in dependence on the number of systems. Figure 9 shows the average *RMSE* (red line with circles) for 50 randomly selected subsets together with the corresponding standard deviation (red bars) over the number of systems that contribute to a subset. A very fast decrease of forecast errors can be observed for up to ten systems. The *RMSE* is further decreasing with the number of systems. For about 150 systems the accuracy for the subset approximates the accuracy of the full ensemble (black dashed line).

In a second step, the forecast of the power output of the subset is compared to the measured power of the complete ensemble (upscaling). The power output values are normalized to the installed power of the ensemble (see section 3), which allows for a direct comparison of the power prediction for the subsets to the measured power of the complete ensembles. The respective mean *RMSE* values and their standard deviations are displayed in blue colour in Figure 9. The average *RMSE* values are very similar to the *RMSE* values when comparing the forecast for the subsets to the measured power output for the respective subset (red line with circles). With a subset of a least 150 systems the power output of the complete ensemble of 500 systems can be predicted with a good approximation to the accuracy that is obtained for the forecasts based on the complete ensemble.



Figure 9: Average *RMSE* for 50 randomly selected subsets together with the corresponding standard deviation over the number of stations. Red: comparison of predicted and measured power of the subset. Blue: comparison of predicted power of the subset to measured power of the complete ensemble. Black squares correspond to uniform regional distribution of the systems of the subset.

In addition to the average *RMSE* values for randomly selected subsets, the average *RMSE* values for subsets with uniform distribution of the systems are given in Figure 9. While the power prediction for the subset with uniform regional distribution (black-red squares) shows higher accuracy than for the randomly selected subsets (see also section 7.2), the *RMSE* is increasing, when evaluating the power output prediction for the subset against the measured power of the complete ensemble (black-blue squares). This shows the importance of a correct representation of the regional distribution for the upscaling approach.

8 CONCLUSIONS

We have presented and compared several approaches to derive hourly site-specific irradiance forecasts from three-hourly ECMWF forecasts as a basis to predict PV power output. We have shown that a considerable improvement of the quality for regional irradiance forecasts can be achieved by correction of systematic deviations of the original forecasts.

Furthermore, we have proposed and evaluated a first approach to predict the PV power output of a large ensemble of systems applying a representative system model. The evaluation of the PV power prediction scheme resulted in an *RMSE* of 0.11 kW/kW_{peak} for single systems. For the ensemble power prediction for an area of 220 km x 220 km an *RMSE* of 0.06 kW/kW_{peak} was found, and for a larger ensemble covering the area of Germany the *RMSE* of 0.05 kW/kW_{peak}. This accuracy is in the same range as the accuracy of operational wind power predictions systems that are already applied to improve the integration of wind energy into the electricity grid.

ACKNOWLEDGEMENTS

We thank the European Centre for Medium-Range Weather Forecasts (ECMWF) for providing forecast data. The research on solar irradiance forecasting presented in this paper was financed by the virtual Institute of Energy Meteorology (vIEM) of the German Helmholtz Association.

REFERENCES

- [1] Lange, M., 2006: A mature market? The history of short-term prediction services. POW'WOW Best-Practises Workshop, Delft, 2006, <u>http://powwow.risoe.dk/publ/ems-</u> research to business powwow Delft 2006.pdf.
- [2] Lorenz E., Heinemann D., Wickramarathne H., Beyer H.G., Bofinger S. (2007): Forecast of ensemble power production by grid-connected PV systems. Proc. of the 22th European PV Conference, 3.9-7.9 2007, Milano, Italy.
- [3] ECMWF European Centre for Medium-Range Weather Forecasts (2007):*IFS Documentation Cy 31r1. IV. Physical Processes.* <u>http://www.ecmwf.int/research/-</u> <u>ifsdocs/CY31r1/index.html</u>.
- [4] Lorenz E., Hurka J., Heinemann D., Beyer H.G. (2008): Irradiance forecasting for the power prediction of grid connected Photovoltaic systems. Special Issue "GEOSS: Benefits of Earth Observations for Renewable Energies", in review, 2008.
- [5] D. Dumortier (1995): Modelling global and diffuse horizontal irradiances under cloudless skies with different turbidities. Daylight II, JOU2-CT92-0144, Final Report Vol. 2, 1995. Technical Report
- [6] Klucher, T.M. (1979): Evaluation of models to predict insolation on tilted surfaces. Solar Energy 23, 111-114.
- [7] Schmidt, H., Sauer, D.U., 1996. Wechselrichter-Wirkungsgrade. Sonnenenergie 4, 43-47 (in German).
- [8] Beyer, H.G., Betcke, J., Drews, A., Heinemann, D., Lorenz, E., Heilscher, G., Bofinger, S.: 2004, *Identification of a General Model for the MPP Performance of PV-Modules for the Application in a Procedure for the Performance Check of Grid Connected Systems*. 19th European Photovoltaic Solar Energy Conference and Exhibition, Paris, France, 07.06.-11.6.2004, pp3073-3076
- [9] Drews A., H.G. Beyer, U. Rindelhardt (2007): Quality of performance assessment of PV plants based on irradiance maps. (in press), doi:10.1016/j.solener.2008.04.009.
- [10] Karampela P., Evaluation of Ensemble PV-Power Production Forecast by applying Representative System Parameters. Master Thesis, Postgraduate Programme Renewable Energy, Oldenburg University.
- [11] Bofinger S., Beyer H.G., Lorenz E., Wikramarathne H. (2007): Validierung und Optimierung eines Solarstromprognosesystems unter Berücksichtigung von Ensemble-Effekten, 22. Symposium Photovoltaische Solarenergie, 6.3.-9.3.2007, Bad Staffelstein, Germany.

 [12] Installierte Photovoltaikleistung in Deutschland zu Jahresende 2006.
 http://www.photon.de/photon/photon-aktion_installleistung.htm.