BENCHMARKING OF DIFFERENT APPROACHES TO FORECAST SOLAR IRRADIANCE

Elke Lorenz[1], Jan Remund[2], Stefan C. Müller[2], Wolfgang Traunmüller[3], Gerald Steinmaurer[4], David Pozo[5], José Antonio Ruiz-Arias[5], Vicente Lara Fanego[5], Lourdes Ramirez[6], Martin Gaston Romeo[6], Christian Kurz[7], Luis Martin Pomares[8], Carlos Geijo Guerrero[9]

[1]University of Oldenburg, Institute of Physics, Energy and Semiconductor Research Laboratory,
Energy Meteorology Unit, Carl von Ossietzky Strasse 9-11, 26129 Oldenburg, Germany, elke.lorenz@uni-oldenburg,de
[2]Meteotest, Fabrikstrasse 14, 3012 Bern, Switzerland
[3]Blue Sky Wetteranalysen, Steinhüblstr.1, 4800 Attang Puchheim, Austria
[4]ASiC – Austria Solar Innovation Center, Roseggerstraße 12,4600 Wels, Austria,
[5]University of Jaén, Department of Physics, Campus Lagunillas, 23071, Jaén, Spain
[6]CENER, Ciudad de la Innovación 7, 31621 Sarriguren (Navarre), Spain
[7]Meteocontrol GmbH, Spicherer Straße 48, 86157 Augsburg, Germany
[8]Ciemat, Energy Department, Av. Complutense 22, 28040, Madrid, Spain
[9]AEMet, Leonardo Prieto Castro, 28071, Madrid, Spain

ABSTRACT: Power generation from photovoltaic 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. Due to the strong increase of solar power generation the prediction of solar yields becomes more and more important. As a consequence, in the last years various research organisations and companies have developed different methods to forecast irradiance as a basis for respective power forecasts. For the end-users of these forecasts it is important that standardized methodology is used when presenting results on the accuracy of a prediction model in order to get a clear idea on the advantages of a specific approach.

In this paper we introduce a benchmarking procedure to asses the accuracy of irradiance forecasts and compare different approaches of forecasting. The evaluation shows a strong dependence of the forecast accuracy on the climatic conditions. For Central European stations the relative *rmse* ranges from 40 % to 60 %, for Spanish stations relative *rmse* values are in the range of 20 % to 35 %.

Keywords: Forecasting, Solar Radiation, Grid-connected, PV System

1 INTRODUCTION

Power generation from photovoltaic 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 fluctuation wind power into the electricity grid (see e.g. [1]). Due to the strong increase of solar power generation the prediction of solar yields also becomes more and more important.

As a consequence, in the last years various research organisations and companies have developed different methods to forecast irradiance as a basis for respective power forecasts. For the end-users of these forecasts it is important that standardized methodology is used when presenting results on the accuracy of a prediction model in order to get a clear idea on the advantages of a specific approach.

A common benchmarking procedure was set up in the framework of the IEA SHC Task 36 „Solar resource knowledge management" (http://sunbird.jrc.it/iea-shc-task36/) and the European project MESoR "Management and Exploitation of Solar Resource Knowledge" (http://www.mesor.net).

As a basis for the benchmarking we have defined a common ground measurement data set. Previous evaluations have shown that site and period may significantly influence the performance of a given forecasting system (see e.g. [2]). Therefore, it is important to use the same ground data for evaluation when comparing different prediction methods. The selected data set with hourly measured irradiance values covers four regions in Europe with different climatic conditions: Southern Germany, Switzerland including mountain stations, Austria, and Spain. The period of evaluation is July 2007 until June 2008.

To assess the accuracy of the predictions different statistical error measures and exploratory methods may be used. For users of the forecast the verification scheme should be kept as simple as possible and we propose to use the root mean square error (*rmse*), the mean absolute error (*mae*), and the bias as basic set of accuracy measures. Another useful quality check is whether a prediction system performs better than any trivial reference model, which is the result of simple considerations and not of modelling efforts. Persistence of the cloud situation was identified as a suitable reference model for irradiance forecasts.

Many users will be interested in the application of the forecasts for a certain region. Hence, the forecast quality is evaluated for regions with different climatological conditions separately.

The proposed procedure of benchmarking was applied to compare hourly irradiance forecasts up to 3 days ahead of seven participants of the IEA SHC task 36. Different methods to forecast irradiance are proposed by the task members. These include the use of mesoscale numerical weather prediction models, the application of statistical post-processing tools to forecasts of a numerical weather prediction (NWP) model, and also a synoptic approach combing different forecasting models.

The paper first presents the ground measurement data sets used for the validation. Then, we give an overview on the different forecasting approaches. Following, the concept of evaluation is introduced, and finally, we provide the results of the application of the proposed benchmarking procedure.

2 GROUND MEASUREMENT DATA

The selected data set with hourly measured irradiance values covers four regions in Europe with different climatic conditions: Southern Germany, Switzerland including mountain stations, Austria, and Southern Spain. Figure 1 gives an overview on the locations of the ground measurement stations. Detailed information on coordinates and station heights is provided in Table 1.

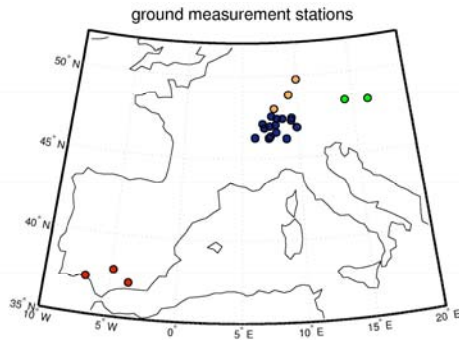The period of evaluation is July 2007 until June 2008.

Figure 1: Locations of ground measurement stations. Red: Spanish stations, orange: German stations, blue: Swiss stations, green: Austrian stations.

Site	Latitude	Longi-tude	Altitude
Fürstenzell	48.55° N	13.35° E	476 m
Stuttgart	48.83° N	9.20° E	318 m
Würzburg	49.77° N	9.97° E	275 m
Basel-Binningen	47.54° N	7.58° E	316 m
Payerne	46.81° N	6.94° E	490 m
La Chaux-de-Fonds	47.09° N	6.80° E	1018 m
Bern-Liebefeld	46.93° N	7.42° E	565 m
Buchs-Suhr	47.38° N	8.08° E	387 m
Napf	47.00° N	7.94° E	1406 m
Zürich SMA	47.38° N	8.57° E	556 m
Säntis	47.25° N	9.34° E	2490 m
St. Gallen	47.43° N	9.40° E	779 m
Genève-Cointrin	46.24° N	6.12° E	420 m
Sion	46.22° N	7.34° E	482 m
Montana	46.31° N	7.49° E	1508 m
Jungfraujoch	46.55° N	7.99° E	3580 m
Locarno-Magadino	46.16° N	8.88° E	197 m
Weissfluhjoch	46.83° N	9.81° E	2690 m
Davos	46.81° N	9.84° E	1590 m
Linz	48.30° N	14.30° E	266 m
Vienna	48.20° N	16.37° E	171 m
Huelva	37.28° N	6.54° W	19 m
Córdoba	37.50° N	4.51° W	91 m
Granada	37.08° N	3.37° W	687 m

Table 1: Locations and heights of ground measurement stations.

3 FORECAST DATA

We have compared hourly irradiance forecasts up to 3 days ahead provided by seven participants of the IEA SHC task 36. The different forecasting approaches are all based on a global numerical weather prediction model, as e.g. the global model of the European Centre for Medium-Range Weather Forecasts (ECMWF) or the Global Forecast System (GFS) model of the National Centers for Environmental Prediction (NCEP). These global models have a coarse temporal and spatial resolution and do not allow for a detailed mapping of small-scale features. Different methods to derive optimized hourly and site specific irradiance forecasts are proposed by the IEA SHC Task 36 participants. These include the use of mesoscale NWP models, the application of statistical post-processing tools or a combination of both, and also a synoptic approach combining different forecasting models.

Table 2 gives an overview on the forecasting approaches with the used NWP models. In the following we shortly introduce the main features of each approach.

The forecasting scheme proposed by the University of Oldenburg is based on irradiance forecasts up to 3 days ahead provided by the ECMWF ([3]). The post-processing procedure developed by [4] implies temporal interpolation of the forecast data to hourly values utilizing a clear sky model in order to account for the diurnal course of the solar irradiance. In addition, a bias correction in dependence on the cloud situation is performed, based on historic ground measured irradiance values for Germany and Switzerland, and on satellite derived irradiance values ([5]) for the other countries.

The contribution of Blue Sky (Austria) includes two forecasting schemes. The first "traditional" method is a result of operational synoptic cloud cover forecasts, using different global and local NWP models in combination with the expert-knowledge of meteorologists. This approach is applicable only for the Austrian stations. The more general second approach is a result of statistical post-processing applied to the GFS-NWP model using different methods of data mining like ridge degression, automatic quadratic models or neuronal networks, based on meteorological inputs (see [6]).

The Meteocontrol forecasts are provided by Meteomedia GmbH. They are based on forecast data from ECMWF, combined with post-processing based on ground measurements (Model Output Statistics, MOS). MOS are statistical relations between model forecast variables and observed weather variables used for the correction of the model forecast values. They take the form of multi-linear regression equations derived by screening potential model forecast variables as predictors. MOS forecasts show significantly reduced systematic errors compared to the output of the numerical weather prediction model. To derive the irradiation forecasts historic ground measurements of irradiation are used (German stations). At stations, where no irradiation measurements are available, irradiance values derived from Meteosat satellite data ([5]), provided by University of Oldenburg, are taken instead (Swiss stations).

Cener's solar global irradiance prediction scheme is mainly based on the regional weather forecasting system Skiron. This wheather forecast system was developed at the Helenic National Meteorological Service ([7] and

[8]), and is used at CENER with initial and GFS boundary conditions, which provide a 1° x 1° horizontal resolution. The execution of Skiron implemented in CENER generates weather forecasts in a grid with a resolution of 0.1° x 0.1° and hourly frequency. The covering domain includes Europe and North African regions. To increase the accuracy of the global irradiance predictions, CENER applies a statistical post process based on learning machines [9]. So, it needs historical data for the training and therefore the final results will depend of the quality and quantity of these data.

Team & abbreviation	Approach	NWP model with spatial and temporal resolution
University of Oldenburg, Germany, ECMWF-OL	Statistical post-processing in combination with a clear sky model	ECMWF - 0.25°x 0.25° - 3 hours
Blue Sky, Austria, a) SYNOP b) BLUE	a)"human" cloud cover forecast (by meteorologists) b) BLUE FORECAST: statistic forecast tool	for b) GFS - 1° x 1° and 0,5°x 0.5° - 3 hours and 6 hours
Meteo-control, Germany, MM-MOS	MOS (model Output Statistics) by Meteomedia GmbH	ECMWF - 0.25°x 0.25° - 3 hours
Ciemat, Spain, HIRLAM-CI	Bias correction	AEMet-HIRLAM - 0.2°x 0.2° - 1 hour
CENER, Spain, CENER	Post-processing based on learning machine models	Skiron/GFS - 0.1°x 0.1° - 1 hour
Meteotest, Switzerland, WRF-MT	Direct model output of global irradiance, averaging of 10x10 pixels	WRF/GFS - 5km x 5km - 1 hour
University of Jaen, Spain, WRF-UJAEN	Direct model output of global irradiance	WRF/GFS - 3km x 3km - 1 hour

Table 2: Overview on forecasting approaches.

The model of Ciemat is based on predictions of the HIRLAM operational model from the Spanish weather service (AEMet). The temporal resolution of the model is 1 hour for the next 72 hours and the spatial resolution is 20 km x 20 km. The solar irradiance predictions of the model are corrected subtracting the bias obtained from a dataset independent of the benchmarking period.

The basis of the forecasting scheme proposed by Meteotest is the numerical Weather Research and Forecasting model (WRF) (see [10]). Input data is derived from GFS 1° x 1° data and three nestings are performed with a grid size of 5 km x 5 km of the innermost domain. For the post-processing Meteotest averaged 10 x 10 model pixels around the point of interest. This corresponds to an area of 50 x 50 km.

The University of Jaén forecasting approach is also based on the WRF mesoscale model, but using a more recent version of the model. The WRF model (version 3, [11]) was run using initial and boundary conditions from the NOAA National Climatic Data Center (NCDC; http://nomads.ncdc.noaa.gov) global forecast system. The initial mother domain was selected to have a horizontal resolution of 27 km x 27 km and two successive nestings (increasing the horizontal resolution successively by a factor of three) were applied to lead up to a final inner domain with a resolution of 3 km x 3 km which cover the study region (Andalusia, Southern Spain). The number of vertical levels simulated was twenty-seven. The choice of the different parameterizations was based on a calibrating experiment for MM5, a former version of the WRF model, carried out in the study region by [12]. Hence, the model was operated with the Thompson microphysic scheme ([13]), the unified Noah land-surface model, the YSU planetary boundary layer scheme ([14]), the Kain-Fritsch cumulus scheme ([15]), the RRTM scheme for long-wave radiation ([16]) and the scheme of Dudhia for shortwave radiation ([10]). In the last nesting, the cumulus and planetary boundary layer schemes were turned off. The integration time step was set to 30 seconds and the time step between successive calls to the radiation scheme was set to 60 seconds.

Summarizing the description of the forecasting methods, we can group the algorithms to four categories:

- The approaches ECMWF-OL, BLUE, and MM-MOS combine a global NWP model with a post-processing method involving historic ground or satellite derived irradiance values.
- The models proposed by CENER and Ciemat apply a post-processing procedure based on historic ground measurement data to predictions of a mesoscale NWP model.
- Meteotest and the University of Jaen provide forecasts of the mesoscale model WRF without any integration of ground measured values.
- The SYNOP approach by Bluesky involves human interpretation of numerical weather forecasts.

4   CONCEPT OF EVALUATION

To assess the performance of a forecast algorithm a lot of different aspects have to be taken into account. In this paper, we focus on few measures of accuracy that are considered to be most relevant for the intended application of power prediction.

Following common practice for the evaluation of wind power predictions (e.g. [17]), we use the root mean square error

$$rmse = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(I_{forecast,i} - I_{measured,i})^2} \qquad (1)$$

as a main score to compare predicted irradiance $I_{forecast,i}$ to measured irradiance $I_{measured,i}$. Here, N is the number of evaluated data pairs.

Using the *rmse* large deviations between predicted and measured values are weighted stronger than small deviations. For most users this is a suitable measure, because large forecast errors often result in disproportionately high financial losses. However, some users prefer a linear weighting of all deviations, and the

$$mae = \frac{1}{N}\sum_{i=1}^{N}\left|I_{forecast,i} - I_{measured,i}\right| \qquad (2)$$

is given as an additional measure.

Furthermore, we calculate the mean value of the errors:

$$bias = \frac{1}{N}\sum_{i=1}^{N}(I_{forecast,i} - I_{measured,i}) \qquad (3)$$

to describe systematic deviations of the forecast.

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

Relative values of the error measures are obtained by normalization to the mean ground measured irradiance of the considered period.

Another useful quality check is, whether a prediction system performs better than any trivial reference model, which is the result of simple considerations and not of modelling efforts. It is worthwhile to implement and run a complex forecasting tool if it is able to clearly outperform trivial models. The probably most common reference model for short term forecasts is persistence, which means that the actual or recently measured values are considered as forecast values. Persistence of the average cloud situation of the previous day in combination with a clear sky model was identified as a suitable reference model for irradiance forecasts. The clear sky model describes the daily course of the irradiance for clear sky situations.

Forecast accuracy is evaluated for the four countries separately. Often, users of the forecasts, e.g. a utility company, are interested in the application of the forecasts and hence also accuracy information for a certain region. Also from a scientific point of view the performance of the forecasting algorithms for regions with different climatological conditions is of interest. Furthermore, there is a practical reason for separating the evaluation for the four countries: Not all forecasting approaches were processed for all countries. Hence, for each country a different subset of the forecasting methods was available.

In addition to the overall results with respect to the statistical error measures, we provide selected results of a more detailed evaluation. Time series of forecasted and measured irradiance illustrate the performance of the forecast for different meteorological conditions. A station by station evaluation with respect to the *rmse* gives further insight to the influence of local climate on the performance of the prediction methods. The seasonal dependence of forecast accuracy is shown exemplarily for Germany, and an evaluation of the frequency distributions is provided for Switzerland.

5   RESULTS

The following sections provide results of the benchmarking of the different forecast algorithms for Germany, Switzerland, Austria, and Spain.

5.1 Germany

For Germany, irradiance forecasts of the University of Oldenburg, Meteocontrol, Meteotest, CENER, and Blue Sky were available for the evaluation.

An overview on the statistical error measures for the complete German data set consisting of three stations is given in Table 3, Table 4, and Table 5 for the first, second, and third forecast day, respectively. The relative *rmse* in dependence on the forecast horizon is also visualized in Figure 2.

approach	*rmse* in W/m^2	*mae* in W/m^2	*bias* in W/m^2
ECWMF-OL	92 (40.3 %)	59 (26.2 %)	-7 (-2.9 %)
BLUE	94 (41.5 %)	62 (27.2 %)	-3 (-1.4 %)
MM-MOS	99 (43.5 %)	67 (29.6 %)	-2 (-1.0 %)
CENER	113 (49.9 %)	72 (31.5 %)	13 (5.9 %)
WRF-MT	118 (51.8 %)	74 (32.6 %)	-1 (-0.3 %)
persistence	144 (63.5 %)	92 (40.5 %)	-4 (-1.8 %)

Table 3: *Rmse, mae*, and *bias* for the five forecasting approaches and persistence, first forecast day, complete German data set.

approach	*rmse* in W/m^2	*mae* in W/m^2	*bias* in W/m^2
ECWMF-OL	95 (41.6 %)	63 (27.5 %)	-7 (-3.1 %)
BLUE	100 (43.9 %)	66 (29.0 %)	-3 (-0.9 %)
MM-MOS	103 (45.1 %)	70 (30.5 %)	-2 (-1.4 %)
CENER	121 (52.9 %)	77 (33.6 %)	9 (3.7 %)
WRF-MT	125 (54.9 %)	79 (34.5 %)	-10 (-4.5 %)
persistence	161 (70.2 %)	104 (45.6 %)	-6 (-2.7 %)

Table 4: *Rmse, mae*, and *bias* for the five forecasting approaches and persistence, second forecast day, complete German data set.

approach	*rmse* in W/m^2	*mae* in W/m^2	*bias* in W/m^2
ECWMF-OL	102 (44.9 %)	67 (29.5 %)	-8 (3.3 %)
BLUE	104 (45.6 %)	69 (30.4 %)	-2 (-0.8 %)
MM-MOS	107 (46.8 %)	73 (32 %)	-2 (-1.1 %)
WRF-MT	138 (60.6 %)	86 (37.7 %)	-13 (-5.7 %)
persistence	167 (73.3 %)	109 (47.9 %)	-7 (-2.9 %)

Table 5: *Rmse, mae*, and *bias* for the four available forecasting approaches and persistence, third forecast day, complete German data set.

The proposed approaches show a very low *bias* for the evaluated forecast horizons. Only the forecasts by CENER show a systematic overestimation of the irradiance of more than 5% for the first forecast day.
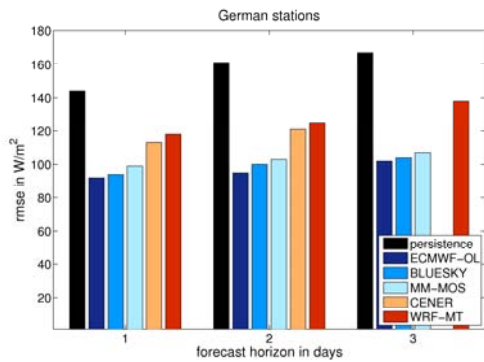
Figure 2: Relative *rmse* for the five forecasting approaches and persistence in dependence on the forecast horizon.

With respect to the *rmse* and *mae* the forecasting approaches combining a global model with post-processing (ECMWF-OL, MM-MOS, BLUE) show better results than mesocale models.

The two mesocale models show a similar performance, with slightly lower *rmse* values for the CENER approach. Here, it has to be considered that the WRF forecasts are direct model output of global irradiance and only an averaging procedure is applied, while the post-processing proposed by CENER involves historic ground data.

Comparing the approaches based on a global model, best accuracy is found for the ECMWF forecasts combined with the post-processing proposed by the University of Oldenburg, closely followed by the statistic forecasts by Blue Sky, and the MM-MOS.

All approaches show a significant improvement in comparison to persistence for all forecast horizons. For persistence a comparatively strong decrease of forecast accuracy occurs for a forecast horizon of two days, while the NWP based approaches only show a small increase of the forecast errors from one day ahead to two days ahead.
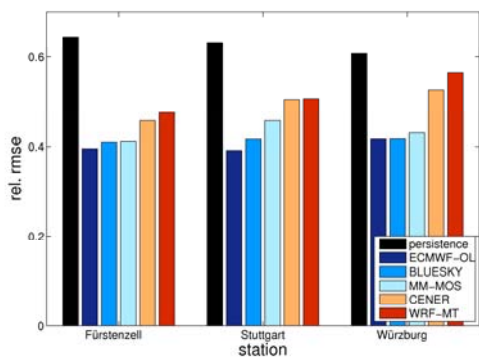

Figure 3: *Rmse* for the five forecasting approaches and persistence per station for the first forecast day.

Figure 3 shows the relative *rmse* for the first forecast day for the three ground measurement stations. The *rmse* for the global model based forecasts is very similar for the three stations. For the mesoscale model forecasts the *rmse* values vary from 47% for Fürstenzell to 55% for Würzburg.

To illustrate the performance of the different forecasting methods, predicted irradiances are compared to measured irradiance for an example time series in Figure 4. The figure shows a good agreement between forecast and measurement for clear sky days for all forecasting algorithms. For days with variable clouds significant deviations between the different forecasts and the measurements may occur.
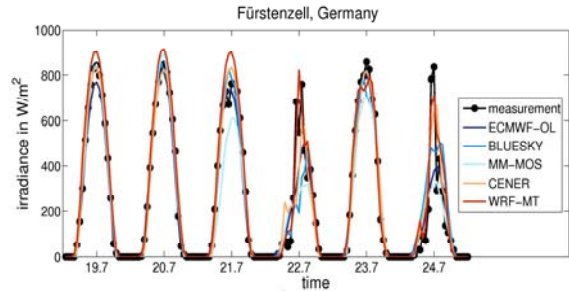

Figure 4: Measured irradiance in comparison to predicted irradiances for six days in July 2007 for the station Fürstenzell.

In order to evaluate the seasonal dependence of forecasts errors, the absolute *rmse* and *bias* are given over the month in Figure 5. During winter with low solar elevations and low clear sky irradiances absolute errors are small and relative errors are large. The improvement in comparison to persistence is low during December, for all other month the NWP based forecasts perform significantly better than persistence.
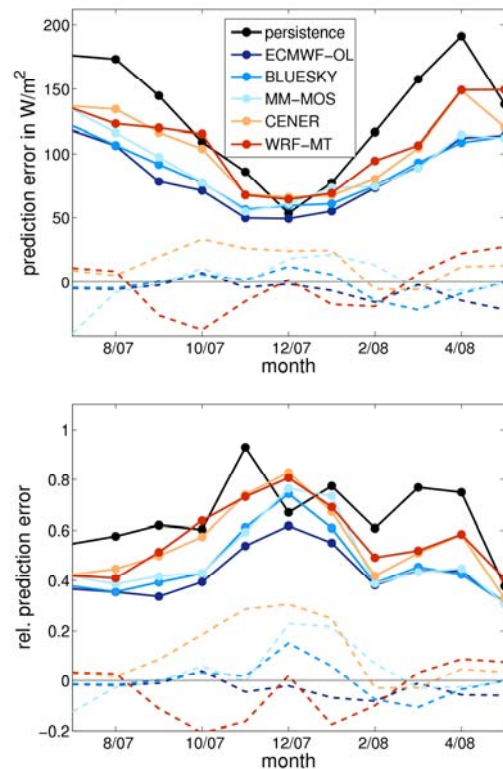

Figure 5: Absolute (top) and relative (bottom) forecast errors; *rmse* (full lines with circles) and *bias* (dashed lines) for the five forecasting approaches and persistence in dependence on the month for the first forecast day using the complete German data set.

With respect to the inter-comparison of the five NWP based approaches a strong dependence on the month is observed. For July and August 2007 the forecasting approaches perform quite similar. For October 2007 and for April and May 2008 the *rmse* of the mesoscale model forecasts is considerably larger than for the global model based approaches.

Summarizing it can be stated, that for Germany global model based forecasts show higher accuracy than the forecasts based on a mesoscale model. The detailed analysis has shown that the difference between mesoscale models and global models with post-processing strongly depends on the month of evaluation. Furthermore, the accuracy of the mesoscale model forecasts shows a stronger dependence on the location of evaluation. All five NWP based approaches perform significantly better than persistence.

5.2 Switzerland

For Switzerland global radiation forecasts of the University of Oldenburg, Meteocontrol, Meteotest and Blue Sky were available for benchmarking.

Table 6 to 8 and Figure 6 give an overview on the evaluation results for the three forecasting days. The relative *rmse* for the first forecast day varies between 39 % and 45 % with best performance of the unbiased Oldenburg approach and the slightly biased BLUESKY approach. The *rmse* values for the MM-MOS and WRF-MT forecasts are very similar for all forecast horizons which is different for Switzerland in comparison to Germany.

approach	*rmse* in W/m^2	*mae* in W/m^2	*bias* in W/m^2
ECWMF-OL	107 (39.6 %)	70 (25.8 %)	-1 (-0.3 %)
BLUE	109 (40.5 %)	73 (27.0 %)	-9 (-3.3 %)
MM-MOS	122 (45.0 %)	85 (31.5 %)	-18 (-6.6 %)
WRF-MT	119 (44.2 %)	76 (28.0 %)	4 (1.3 %)
persistence	158 (58.4 %)	104 (38.7 %)	-17 (-6.3 %)

Table 6*: Rmse, mae*, and *bias* for the four forecasting approaches and persistence, first forecast day, complete Swiss data set.

approach	*rmse* in W/m^2	*mae* in W/m^2	*bias* in W/m^2
ECWMF-OL	113 (41.8 %)	74 (27.4 %)	-1 (-0.4 %)
BLUESKY	113 (41.9 %)	76 (28.0 %)	-3 (-1.2 %)
MM-MOS	125 (46.3 %)	87 (32.4 %)	-18 (-6.7 %)
WRF-MT	125 (46.2 %)	79 (29.1 %)	2 (0.9 %)
persistence	173 (64.0 %)	116 (42.9 %)	-18 (-6.6 %)

Table 7: *Rmse, mae*, and *bias* for the four forecasting approaches and persistence, second forecast day, complete Swiss data set.

approach	*rmse* in W/m^2	*mae* in W/m^2	*bias* in W/m^2
ECWMF-OL	115 (42.7 %)	76 (28.1 %)	-2 (-0.6 %)
BLUESKY	120 (44.3 %)	80 (29.8 %)	-4 (-1.4 %)
MM-MOS	130 (48.1 %)	91 (33.7 %)	-19 (-7.2 %)
WRF-MT	136 (50.5 %)	86 (31.8 %)	-2 (-0.9 %)
persistence	180 (66.8 %)	121 (44.9 %)	-19 (-7.0 %)

Table 8: *Rmse, mae*, and *bias* for the four forecasting approaches and persistence, third forecast day, complete Swiss data set.

With respect to the *mae*, the WRF forecasts show a slightly better performance than the MM-MOS. Persistence shows a considerably higher relative *rmse* and *mae* than all NWP based forecasts. Persistence and the Meteocontrol approach show a negative *bias* of about 7%, the *bias* for all other approaches is small.

Figure 7 shows the relative rmse for selected stations. Due to large variations with respect to climate and topography, considerable differences in the results for the different stations are found. Best performance by all approaches is achieved for the sunny stations, like for example Sion. For the high alpine station Jungfraujoch the MM-MOS has higher rmse values than persistence, while for all other stations all approaches perform better than persistence.
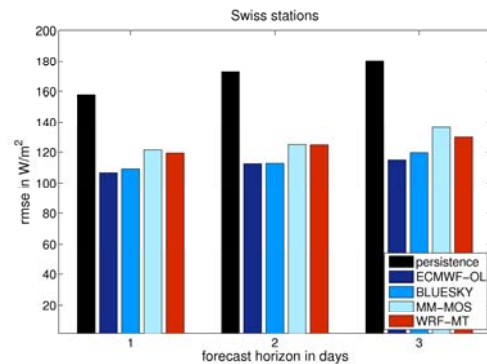

Figure 6: *Rmse* for the four forecasting approaches and persistence in dependence on the forecast horizon.
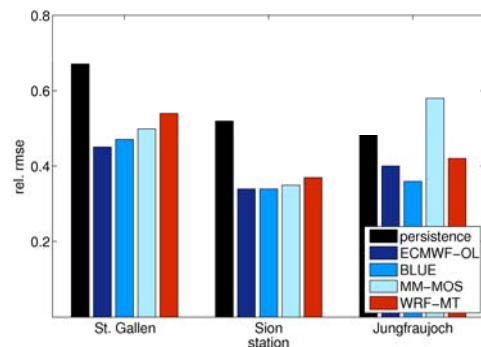

Figure 7: Relative *rmse* for the four forecasting approaches and persistence per station for the first forecast day.

In addition to the statistical measures characterizing the agreement of single values, the distribution functions of predicted and measured time series are evaluated for the example of Switzerland. The frequency distributions of forecasted and ground measured data were compared with a Kolmogorov-Smirnof test. The *OVER* parameter ([18]) quantifies the coincidence of the two distributions. This parameter is less important than *rmse* and *mae* for forecast evaluation with respect to solar energy applications, however it provides additional insight to the performance of the different approaches. As shown in Figure 8, the WRF-MT approach shows a good agreement between the frequency distribution of forecasted and measured values. The other three approaches, here represented by the statistical model of Blue Sky approach, do not reproduce the highest irradiance values. Similar results are obtained for Germany and Austria.
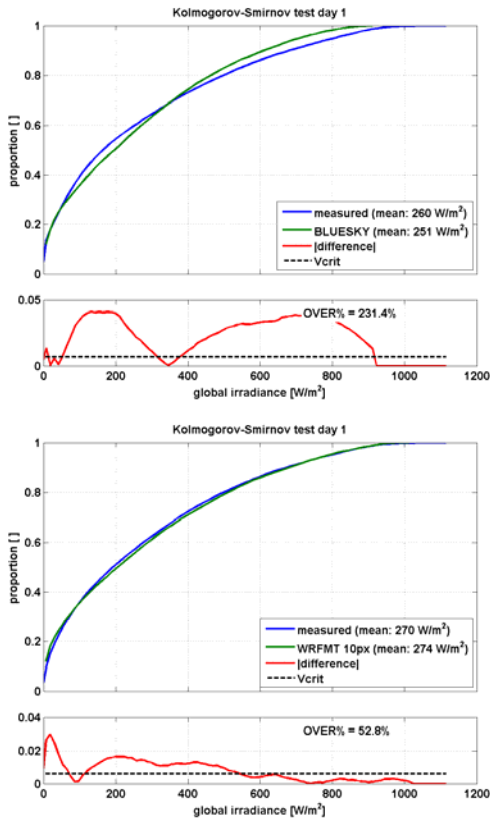

Figure 8: Kolmogorov-Smirnof test for the forecasting approaches of Blue Sky (top) and Meteotest (bottom) for the first forecast day.

Overall, it can be concluded that for Switzerland the approaches of University of Oldenburg and Blue Sky achieve higher accuracy than those of Meteocontrol and Meteotest. The Meteotest approach produces the most realistic distribution of irradiance values. Due to variations in topography and climate all forecasting approaches show a strong dependency of the forecast accuracy on the station.

5.3. Austria

For Austria, in addition to the forecasts of the University of Oldenburg, Meteotest, CENER and the statistical method of Blue Sky, also the traditional synoptic method of the meteorologists of Blue Sky was evaluated. This method is spatially restricted and only available for Austria.

approach	*rmse* in W/m^2	*mae* in W/m^2	*bias* in W/m^2
ECWMF-OL	101 (45.6 %)	65 (29.2 %)	16 (7 %)
BLUE	99 (44.6 %)	61 (27.6 %)	1 (0.5 %)
CENER	129 (58.1 %)	87 (39.0 %)	30 (13.6 %)
WRF-MT	123 (55.4 %)	77 (34.7 %)	47 (21.0 %)
SYNOP	112 (50.4 %)	70 (31.5 %)	9 (0.4 %)
persistence	142 (64.3 %)	91 (41.2 %)	-14 (6.4 %)

Table 9: *Rmse, mae*, and *bias* for the five forecasting approaches and persistence, first forecast day, complete Austrian data set.

approach	*rmse* in W/m^2	*mae* in W/m^2	*bias* in W/m^2
ECWMF-OL	107 (47.4 %)	69 (30.6 %)	14 (6.1 %)
BLUE	101 (45.2 %)	64 (28.4 %)	1 (0.5 %)
CENER	134 (59.5 %)	90 (39.8 %)	21 (9.4 %)
WRF-MT	132 (58.5 %)	84 (37.1 %)	30 (13.1 %)
SYNOP	111 (49.3 %)	70 (30.9 %)	7 (2.9 %)
persistence	158 (70.2 %)	103 (45.5 %)	-14 (-6.6 %)

Table 10: *Rmse, mae*, and *bias* for the five forecasting approaches and persistence, second forecast day, complete Austrian data set.

approach	*rmse* in W/m^2	*mae* in W/m^2	*bias* in W/m^2
ECWMF-OL	112 (50.5 %)	72 (32.3 %)	17 (7.5 %)
BLUE	107 (48.1 %)	69 (30.9 %)	5 (2.1%)
WRF-MT	140 (62.9 %)	88 (39.7 %)	30 (13.5 %)
persistence	160 (71.7 %)	103 (46.5 %)	-15 (-6.5 %)

Table 11: *Rmse, mae*, and *bias* for the three available forecasting approaches and persistence, third forecast day, complete Austrian data set.

The overall results (see Tables 9 to 11, and Figure 9) are similar to Germany and Switzerland. The methods using statistical post-processing based on a global model have smaller errors than the other models. Best results are found for the statistical approach of Blue Sky. Also the synoptic method of the meteorologists of Blue Sky achieves good results. The two mesoscale models show higher errors, in particular there is a considerable

overestimation of the irradiance. However, still a considerable improvement is achieved in comparison to persistence.
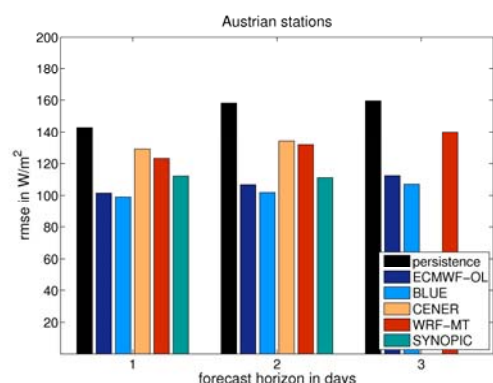

Figure 9: *Rmse* for the five forecasting approaches and persistence in dependence on the forecast horizon.

Generally, the performance of all models for the Austrian stations, especially for Linz (see Figure 10), is slightly worse than in the other countries, due to orograhic effects and local weather phenomena, for example fog.
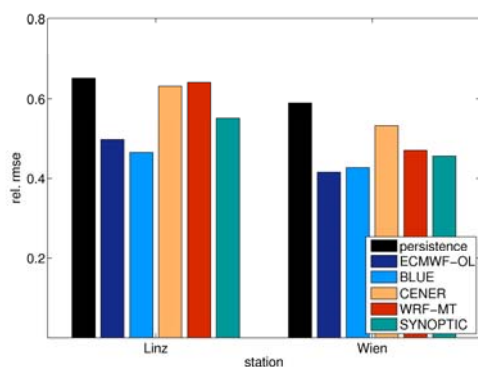

Figure10: Relative *rmse* of the five forecasting approaches and persistence per station for the first forecast day.

An advantage of the forecasts by meteorologists can be seen during a fog episode case study of 3 days in Linz. The numerical models and the models with statistical post-processing predicted sunny days with clear sky conditions. The meteorologists had the opportunity to adjust their forecasts during the fog episode, as illustrated in Figure 11.

Summarizing it can be stated, that model performance in Austria is similar to Germany and Switzerland with advantages of the global models with statistical post-processing compared to the mesoscale models. In addition, a synoptic approach was analyzed, with slightly higher *rmse* values than the global model based forecasts in general, but increased accuracy for special meteorological situations.
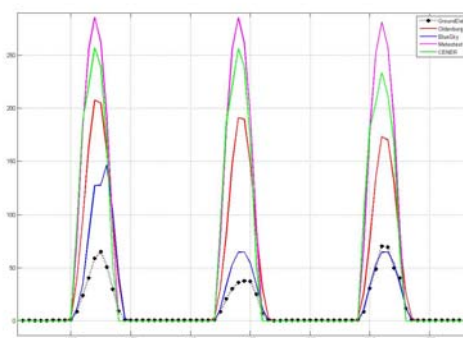

Figure 11: Hourly solar irradiance forecasts during a fog episode in Linz (13th-15th January 2008); black dots: ground measurements: red: ECMWF-OL, blue: SYNOP, pink: WRF-MT, green: CENER.

5.4. Southern Spain

In Spain, besides the approaches of CENER and University of Oldenburg, two additional approaches based on mesoscale models were integrated into the evaluation.

The overall results are given in Table 12 to Table 14, and in Figure 11. In general, forecast errors in Spain are smaller than in Central Europe. The relative *rmse* values in Spain range from of 20 % for the best forecasts to 35 % for persistence, in comparison to about 40 % for the best forecasts and 60 % for persistence in Central Europe. For sunny days, a better agreement of forecast and measurement is achieved, as illustrated in Figure 4. Hence, in Southern Spain with sunny climate, absolute error values are smaller than in Europe. Concerning relative error values, there is an additional effect by normalisation to a higher average irradiance in Spain.

approach	*rmse* in W/m^2	*mae* in W/m^2	*bias* in W/m^2
ECWMF-OL	81 (20.8 %)	48 (12.2 %)	1 (0.1 %)
CENER	100 (25.4 %)	64 (16.3 %)	-3.33 (-0.8 %)
HIRLAM-CI	124.06 (31.7 %)	79.87 (20.4 %)	58.87 (15.0 %)
WRF-UJAEN	89.64 (22.9 %)	57.64 (14.7 %)	28.26 (7.2 %)
persistence	125.88 (32.1 %)	64.91 (16.6 %)	2.44 (0.6 %)

Table 12: *Rmse, mae,* and *bias* for the four forecasting approaches and persistence, first forecast day, complete Spanish data set.

With respect to the inter-comparison of the different algorithms, again the ECMWF-OL algorithm performs best, closely followed by the WRF forecasts provided by the University of Jaen. The HIRLAM-CI forecasts are provided with the largest *rmse* and *mae* values and a systematic overestimation of the irradiance of about 15 %. Like for Central Europe, all algorithms show smaller errors than persistence.

approach	rmse in W/m^2	mae in W/m^2	bias in W/m^2
ECWMF-OL	83 (21.3 %)	49 (12.6 %)	0 (0.0 %)
CENER	106 (27.1 %)	67 (17.2 %)	-4 (-1.0 %)
HIRLAM-CI	144 (36.8 %)	94 (23.9 %)	61 (15.7 %)
WRF-UJAEN	96 (24.4 %)	61 (15.5 %)	29 (7.4 %)
persistence	140 (35.8 %)	75 (19.1 %)	5 (1.2 %)

Table 13: *Rmse, mae*, and *bias* for the four forecasting approaches and persistence, second forecast day, complete Spanish data set.

approach	rmse in W/m^2	mae in W/m^2	bias in W/m^2
ECWMF-OL	88 (22.4 %)	52 (13.2 %)	1 (0.3 %)
HIRLAM-CI	160 (40.9 %)	104 (26.7 %)	65 (16.7 %)
WRF-UJAEN	105 (26.9 %)	72 (18.3 %)	30 (7.6 %)
persistence	145 (37.0 %)	80 (20.4 %)	7 (1.7 %)

Table 14: *Rmse, mae*, and *bias* for the three available forecasting approaches and persistence, third forecast day, complete Spanish data set.
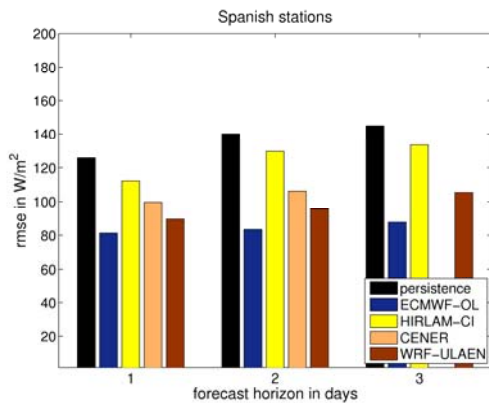

Figure 12: *Rmse* for the four forecasting approaches and persistence in dependence on the forecast horizon.

Figure 13 shows the evaluation for the single stations. Due to local climatic conditions again a considerable variation of the *rmse* values from station to station is found.

Summarizing the evaluation for Spain it shall be emphasized that forecasts errors are smaller than for Central Europe. Lowest errors are achieved by the ECMWF based forecast of University of Oldenburg and the WRF forecasts of the University of Jaen.

6   SUMMARY AND OUTLOOK

A procedure of benchmarking irradiance forecasts has been presented and applied to seven different forecasting algorithms.
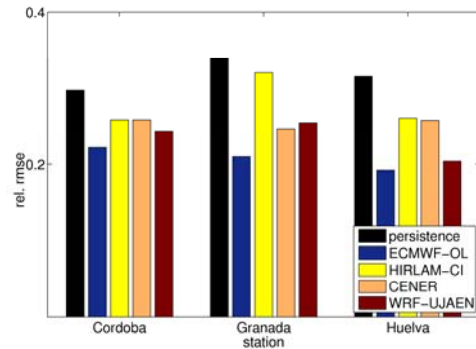

Figure13: Relative *rmse* of the four forecasting approaches and persistence per station for the first forecast day.

A strong dependence of the forecast accuracy on the climatic conditions is found. For Central European stations the relative *rmse* ranges from 40% to 60%, for Spanish stations relative *rmse* values are in the range of 20% to 35%. At the current stage of research, irradiance forecasts based on global model numerical weather prediction models in combination with post-processing show best results. All proposed methods perform significantly better than persistence.

There is ongoing development of the methods to predict irradiance by the IEA SHC Task 36 members. Accordingly, evaluation and comparison of the forecasts will be continued and extended to recent ground measurement data.

REFERENCES

[1] M. Lange: A mature market? The history of short-term prediction services. POW'WOW Best-Practises Workshop, Delft, http://powwow.risoe.dk/publ/ems-research_to_business_powwow_Delft_2006.pdf, 2006.
[2] J. Remund, R. Perez, and E. Lorenz: Comparison of Solar Radiation Forecasts for the USA. Proc. of the 23rd European PV Conference, 1.9-4.9 2008, Valencia, Spain.
[3] J.-J. Mocrette, H.W. Barker , M.J. Iancono, G. Mozdzynski, R. Pincus, D. Salmond, S. Serrar: A new radiation package. McRad. ECMWF_Newsletter 112, Summer 2007
[4] E. Lorenz, J. Hurka, D. Heinemann, and H.G. Beyer: Irradiance Forecasting for the Power Prediction of Grid-

connected Photovoltaic Systems IEEE Journal of Selected Topics in Applied Earth Observations and remote sensing, Vol. 2, No. 1, March 2009.

[5] A. Hammer, D. Heinemann, C. Hoyer, R. Kuhlemann, E. Lorenz, R.W. Mueller, H.G. Beyer: Solar energy assessment using remote sensing technologies. Remote Sensing of Environment 86, 423-432, 2003.

[6] T. Natschläger, W. Traunmüller, K. Reingruber, H. Exner: Lokal optimierte Wetterprognosen zur Regelung stark umweltbeeinflusster Systeme; SCCH, Blue Sky. Tagungsband Industrielles Symposium Mechatronik Automatisierung, pages 281-284, Clusterland Oberösterreich GmbH / Mechatronik-Cluster. 2008.

[7] Kallos, G. The Regional weather forecasting system SKIRON. Proceedings, Symposium on Regional Weather Prediction on Parallel Computer Environments, 15-17 October 1997, Athens, Greece

[8] Papadopoulos A, Katsafados P, and Kallos G. Regional weather forecasting for marine application. Global Atmosphere Ocean Systems 2001; 8: 219-237

[9] M. Gastón, E. Lorenz, S. Lozano, D. Heienmann, M. Blanco, and L. Ramírez: Comparison of global irradiance forecasting approaches, Proceedings of the 15th SolarPACES Symposium, 15.9.-18.9.2009, BErlin Germany.

[10]Dudhia, J. (1989), Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model, J. Atmos. Sci., 46, pp. 3077-3107.

[11]Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. G. Duda, X. Y. Huang, W. Wang, and J. G. Powers (2008), A Description of the Advanced Research WRF Version 3, Mesoscale and Microscale Meteorology Division, Natl. Cent. Atmos. Res., NCAR/TN-475+STR, Boulder, Colorado.

[12]Ruiz-Arias, J. A., D. Pozo-Vázquez, N. Sánchez-Sánchez, J. P. Montávez, A. Hayas-Barrú, and J. Tovar-Pescador (2008), Evaluation of two MM5-PBL parameterizations for solar radiation and temperature estimation in the South-Eastern area of the Iberian Peninsula, Il Nuovo Cimento C, 31, pp 825-842.

[13]Thompson, G., R. M. Rasmussen, and K. Manning (2004), Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis, Mon. Wea. Rev., 132, pp. 519–542.

[14]Hong, S. Y., Y. Noh, and J. Dudhia (2006), A new vertical diffusion package with an explicit treatment of entrainment processes, Mon. Wea. Rev., 134, pp. 2318–2341.

[15]Kain, J. S., and J. M. Fritsch (1993), Convective parameterization for mesoscale models: The Kain-Fritcsh scheme in The representation of cumulus convection in numerical models, edited by K. A. Emanuel and D.J. Raymond, Amer. Meteor. Soc.

[16]Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough (1997), Radiative transfer for inhomogeneous atmosphere: RRTM, a validated correlated-k model for the long-wave, J. Geophys. Res., 102(D14), pp. 16663–16682.

[17] T.S. Nielsen, H. Madsen, H. Aa. Nielsen, P. Pinson, G. Kariniotakis, N. Siebert, I. Marti, M. Lange, U. Focken, L. von Bremen, P. Louka, G. Kallos, G. Galanis (2006): Advanced Statistical Modelling and Uncertainty Assessment for Wind Power Forecasting. Proc. of the European Wind Energy Conference 2006, 27.2-2.3 2006, Athens, Greece.

[18] B. Espinar, L. Ramírez, A. Drews, H.G. Beyer, L.F. Zarzalejo, J. Polo, L. Martín, (2008): Analysis of different error parameters applied to solar radiation data from satellite and German radiometric stations, Solar Energy, Volume 83, Issue 1, January 2009