Robert Ulbricht, Martin Hahmann, Hilko Donker, Wolfgang Lehner

Systematical Evaluation of Solar Energy Supply Forecasts

Erstveröffentlichung in / First published in:

Data Analytics for Renewable Energy Integration: Second ECML PKDD Workshop. Nancy, 19.09.2014. Springer, S. 108–121. ISBN 978-3-319-13290-7.

DOI: http://dx.doi.org/10.1007/978-3-319-13290-7_9

Final edited form was published in "Data Analytics for Renewable Energy Integration: Second ECML PKDD Workshop. Nancy 2014", S. 108–121, ISBN 978-3-319-13290-7
http://dx.doi.org/10.1007/978-3-319-13290-7_9

Systematical Evaluation of Solar Energy Supply Forecasts

Robert Ulbricht[2](✉), Martin Hahmann[1], Hilko Donker[2], and Wolfgang Lehner[1]

[1] Database Technology Group, Technische Universität Dresden, Dresden, Germany
[2] Robotron Datenbank-Software GmbH, Dresden, Germany
robert.ulbricht@robotron.de

Abstract. The capacity of renewable energy sources constantly increases world-wide and challenges the maintenance of the electric balance between power demand and supply. To allow for a better integration of solar energy supply into the power grids, a lot of research was dedicated to the development of precise forecasting approaches. However, there is still no straightforward and easy-to-use recommendation for a standardized forecasting strategy. In this paper, a classification of solar forecasting solutions proposed in the literature is provided for both weather- and energy forecast models. Subsequently, we describe our idea of a standardized forecasting process and the typical parameters possibly influencing the selection of a specific model. We discuss model combination as an optimization option and evaluate this approach comparing different statistical algorithms against flexible hybrid models in a case study.

Keywords: Solar energy · Energy forecast model · Classification · Ensemble

1 Introduction

The capacity of renewable energy sources (RES) constantly increases world-wide due to governmental funding policies and technological advancements. Unfortunately, most of the grid-connected RES installations are characterized by a decentralized allocation and a fluctuating output owed to the changing nature of the underlying powers. Coincidentally, today's available transformation and storage capabilities for electric energy are limited and cost-intensive, which is the primary reason for the increasing interference of renewable energy output with power network stability. Efficient and dedicated forecasting methods will help the grid operators to better manage the electric balance between power demand and supply in order to avoid unstable situations or even possible collapses in the near future. A lot of research has been conducted in the past years by different communities trying to cope with this challenge. Despite of the large amount of available related work and both scientific and practical optimization ideas, there is still no straightforward and easy-to-use recommendation for a standardized forecasting strategy. Comparing the results obtained while executing different

-

Final edited form was published in "Data Analytics for Renewable Energy Integration: Second ECML PKDD Workshop. Nancy 2014", S. 108–121, ISBN 978-3-319-13290-7
http://dx.doi.org/10.1007/978-3-319-13290-7_9

experimental approaches is difficult, as most of the presented cases are bound to a specific region including the corresponding real-world data-set. Further, there is no constant form of result evaluation across all publications, as different error metrics are applied to measure output quality.

In this paper, we address the problem of a systematical optimization for solar energy forecasting strategies conducting an analysis of state-of-the-art approaches. The paper is organized as follows: In Sect. 2 we review and classify models proposed in the literature to predict (1) weather influences and (2) the output of solar energy production units. In Sect. 3, the energy forecasting process is described and relevant parameter settings and exogenous influences for the model selection decision are discussed before that background. In Sect. 4 we evaluate the performance of an exemplary ensemble model which combines the forecast output of popular statistical prediction methods using a dynamic weighting factor. Finally, we conclude and outline additional research directions for our future work in Sect. 5.

2 Energy Supply Forecasting Approaches

The prediction of energy time series is a classical application of time series analysis methods. Thus, there is a long research history related to electricity load forecasting, where a range of sophisticated high-quality models has been developed and classified (i.e. compare the work of Alfares and Nazeeruddin [2]). In contrast, the need for energy supply forecasting is a much more recent topic, as the challenge of grid-connected RES penetrating the distribution systems has emerged just a couple of years ago. Nevertheless, both energy demand and supply forecasting approaches make use of similar techniques.

2.1 Weather Forecast Models

In order to make energy supply planning rational, forecasts of RES production have to be made based on the consideration of weather conditions as the most influencing factor for output determination for solar energy production is the quality of the solar irradiation forecast. Consequently, the use of precise weather forecast models is essential before reliable energy output models can be generated. Although this step is orthogonal to a grid operator's core activities (weather data usually is obtained from meteological services), a basic understanding of the underlying principles is helpful when choosing a specific energy output model.

Numerical Weather Prediction. Complex global *numerical weather prediction* (NWP) models is a modern and common method to predict a number of variables describing the physics and dynamic of the atmosphere, which is then used to derive the relevant weather influences at a specific point of interest. These are e.g. the *European Center for Medium-Range Weather-Forecasts Model*[1] (ECMWF), the *Global Forecast System* (GFS) from National Centers for

[1] http://www.ecmwf.int

Final edited form was published in "Data Analytics for Renewable Energy Integration: Second ECML PKDD Workshop. Nancy 2014", S. 108–121, ISBN 978-3-319-13290-7
http://dx.doi.org/10.1007/978-3-319-13290-7_9

Environmental Prediction[2] or the *North American Mesoscale Model*[3] (NAM). As they have a coarse spatial and temporal resolution, several post-processing and correction techniques are applied in order to obtain down-scaled models of finer granularity (e.g. Model Output Statistics). A quality benchmark was conducted by Lorenz et al. [15], where European ground measurement data is used to compare the performance of each NWP including different scientific and commercial post-processing approaches (Fig. 1).

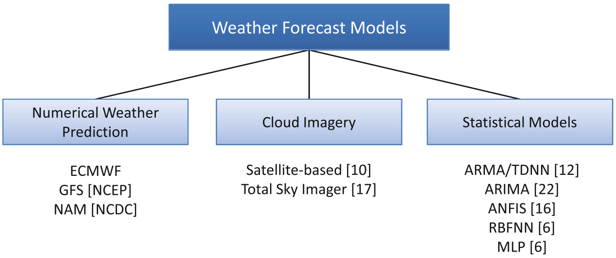

Fig. 1. Classification of weather forecasting models

Cloud Imagery. The influence of local cloudiness is considered to be the most critical factor for the estimation of solar irradiation, especially on days with partial cloudiness where abrupt changes may occur. The use of satellite data can provide high quality short-term forecasts, as geostationary satellites like METEOSAT provide half-hourly spectrum images with a resolution from 1 to 3 square kilometers. Clouds are detected by processing these images into cloud-index images. To predict the future position of a cloud over ground, two consecutive cloud-index images are interpolated using motion vectors [10]. A similar method is the use of *Total Sky Imagers*, which enables real-time detection of clouds in hemispherical sky images recorded by ground-based cameras using sophisticated analytical algorithms [17].

Statistical Models. Furthermore, there are several studies treating the forecasting of solar radiation based on historical observation data using common time series regression models like ARIMA, *Artificial Neural Networks* (ANN) or *Fuzzy-Logic* models (FL). An analysis published by Reikard shows that after comparing various regression models, ARIMA in logs with time-varying coefficients performs best, due to its ability to capture the diurnal cycle more effectively than other methods [22]. Ji and Chee [12] propose a combination of ARMA and a *Time Delay Neural Network* (TDNN). Dorvlo et al. discuss the usage of two ANN-models: *Radial Basis Functions* (RBF) and *Multilayer*

[2] http://www.ncep.noaa.gov
[3] http://www.ncdc.noaa.gov

Final edited form was published in "Data Analytics for Renewable Energy Integration: Second ECML PKDD Workshop. Nancy 2014", S. 108–121, ISBN 978-3-319-13290-7
http://dx.doi.org/10.1007/978-3-319-13290-7_9

Perceptron (MLP) [6]. Martin et al. [16] compare the performance of autoregressive models (AR) against ANN and Adaptative-network-based fuzzy inference system (ANFIS). As such statistical models usually are considered being domain-neutral, their characteristics are discussed more in detail in the subsequent section.

2.2 Energy Forecast Models

Any output from the weather models described above must then be converted into electric energy output. According to the underlying methodology, the existing solutions can be classified into the categories of *physical*, *statistical* and *hybrid* methods as presented in Fig. 2.

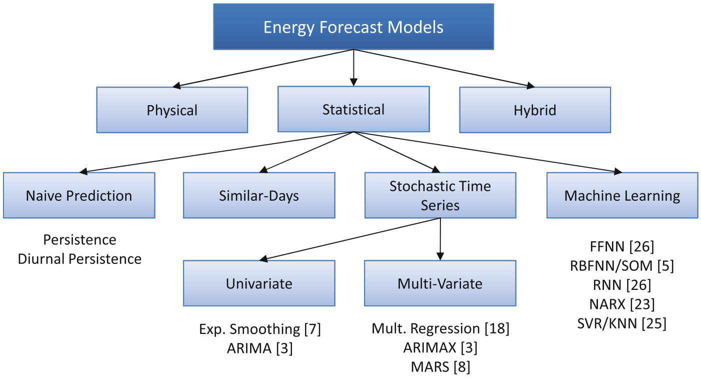

Fig. 2. Classification of energy forecasting models

Physical Models. All forecasting approaches mainly relying on a renewable power plant's technical description concerning its ability to convert the introduced meteorological resources into electrical power are summarized by the term *physical model*. Taking into account external influences derived from NWP, atmospheric conditions and local topography, once they are fitted they are accurate and do not require historical output curves. Especially the latter makes them suitable for estimating the future output of planned or recently installed RES units. Applications of physical models are more frequently found for wind power prediction, but are also used for solar energy forecasts. For example, if we consider the electrical energy P_E extracted from the NWP for global radiation G_{nwp} by a PV panel, the equation for a simplyfied model is as follows:

$$P_E = \alpha G_{nwp} A \tag{1}$$

where α is the conversion efficiency of the solar panel and A is its surface size. Improvements of this method are demonstrated by Iga and Ishihara [11] including the outside air temperature, or Alamsyah et al., using the panel temperature

Final edited form was published in "Data Analytics for Renewable Energy Integration: Second ECML PKDD Workshop. Nancy 2014", S. 108–121, ISBN 978-3-319-13290-7
http://dx.doi.org/10.1007/978-3-319-13290-7_9

[1] as additional parameters. The major disadvantage of physical models is that they are highly sensitive to the NWP prediction error. Furthermore, they have to be designed specifically for a particular energy system and location. As a consequence, the usage of such models requires detailed technical knowledge about characteristics and parameters of all underlying components, thus making them more relevant for energy plant owners or producers than for grid operators.

Statistical Models. *Naive Prediction.* The most straightforward approach to determine a time series' future value denoted as P'_{t+1} would be a naive guess, assuming that next periods' expected energy output will be equal to the observations of the current period P_t. This method is called *naive* or *persistent prediction*. The distinctive cycle of solar energy is expressed by choosing a period of 24 h for *diurnal persistence*, so forecasts are obtained by

$$P'_t = P_{t-k} \tag{2}$$

with k being the number of values per day, i.e. $k = 96$ having a time series granularity of 15 min. Although very limited due to its inability to adopt to any influences and therefore providing results of low preciseness, it is easy to implement and commonly used as a reference model to evaluate the performance of concurrent, more sophisticated forecasting approaches. Using complex forecasting tools is worthwhile only if they are able to clearly outperform such trivial models.

Similar-Days Model. Based on the concept of diurnal persistence, improved forecasts can be computed by selecting similar historical days using suitable time series similarity measures like e.g. Euclidean distance. These models are very popular for load forecasts (e.g. compare [19]), where weather-awareness plays a minor part compared to the influence of consumption-cycle patterns derived from historical data. As for solar energy forecasts, such models are used whenever there is no NWP available at all or the prediction error included naturally in the NWP is estimated as too high to provide reliable energy output forecasts.

Stochastic Time Series. Depending on the number of influencing parameters, two groups of models can be distinguished: *Uni-* and *Multivariate* models. Univariate models are calculated based on the time series' history only. Well known representatives of that group are *Auto-Regressive (Integrated) Moving Average* models (ARMA/ARIMA), which can be described best as a stochastic process combining an auto-regressive component (AR) with a moving average component (MA). Dunea et al. [7] propose the consideration of *Exponential Smoothing* as an effective alternative for one-period ahead forecasts. In contrast, multivariate models allow for the integration of exogenous parameters. *Multiple Regression* methods like ARIMAX (ARIMA with exogenous influences) are a popular choice whenever there is a linear correlation structure expected in two time series [18]. In the case of solar energy prediction, this is given by the dominating dependency of energy output on the global radiation values from the NWP. Historical

Final edited form was published in "Data Analytics for Renewable Energy Integration: Second ECML PKDD Workshop. Nancy 2014", S. 108–121, ISBN 978-3-319-13290-7
http://dx.doi.org/10.1007/978-3-319-13290-7_9

observation data is used to derive the regression coefficients. Bacher et al. demonstrate the performance of an ARIMA model using a clear-sky-normalization for short-term forecasts [3]. As an extension to linear modeling *Multivariate Adaptive Regression Splines* (MARS), a methodology developed by Friedman [8], is used in the energy domain to generate more flexible, nonlinear models.

Machine Learning. The use of machine learning methods is a common approach to forecast a time series' future values, as they are seen as alternative to conventional linear forecasting methods. Reviewed literature shows that ANN have been successfully applied for forecasts of fluctuating energy supply. ANN learn to recognize patterns in data using training data sets. For example, the use of neural networks is proposed by Yona et al. [26] due to their examination of the *Feed-Forward* (FFNN), the *Radial Basis Function* (RBFNN) and the *Recurrent Neural Network* (RNN) for solar power forecasting based on NWP input and historical observation data. A similar approach is described by Chen et al. [5], where a RBFNN is combined with a weather type classification model obtained by a *Self Organizing Map* (SOM). Wolf et al. compare *k-Nearest Neighbors* (KNN) and *Support Vector Regression* (SVR) finding that the latter outperforms KNN on non-aggregated data [25]. In contrast, Tao et al. compute hourly energy forecasts using an adaptive NARX network combined with a clear-sky radiation model, which allows for forecasts without including NWP data and still outperforms non-adapting regression-based methods [23].

Hybrid Models. Any combination of two or more of the above described methods is known as a *hybrid model*. The use of such hybrid approaches has become more popular as it offers the possibility to take advantage of the strongest points of different stand-alone forecasting techniques. The basic idea of combining models is to use each methods' unique features to capture different patterns in the data. Theoretical and empirical findings from other domains suggest that combining linear and non-linear models can be an efficient way to improve the forecast accuracy (e.g. [27]), so hybrid models seem to be a promising approach that can potentially outperform non-hybrid models individually. A successful application of this idea is provided e.g. by the work of Ogliari et al. [21].

3 Energy Forecasting Process

As shown in the previous section, there are plenty of possibilities to compute forecasts for fluctuating energy production units. But choosing the optimal forecasting model for a given use case is an important decision to make and requires expert knowledge. Figure 3 describes the forecasting process steps: First, the raw data has to be *preprocessed* according to the specific requirements of the wanted forecast. Second is the *selection* of a suitable algorithm to best describe the observations. Next, the parameters for the chosen model have to be *estimated* before the *forecasting* task is executed. After *evaluating* the obtained results, this decision might be reconsidered in case of too high and therefore unsatisfying prediction errors. From the description of the forecasting techniques mentioned in

Final edited form was published in "Data Analytics for Renewable Energy Integration: Second ECML PKDD Workshop. Nancy 2014", S. 108–121, ISBN 978-3-319-13290-7
http://dx.doi.org/10.1007/978-3-319-13290-7_9

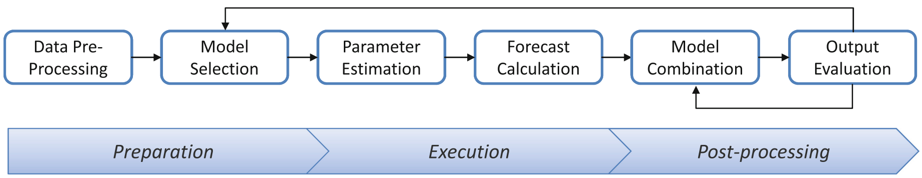

Fig. 3. Typical forecasting process steps

the introduction, we derive that the choice of the appropriate forecasting models depends on the amount and quality of external information, the applied forecast horizon, the data aggregation level and the availability of historical observation data. Furthermore, we consider *model combination* as an optimization option, a strategy also known as *ensemble prediction*. Ensembles can be created manually based on user preferences and experiences or by using machine driven optimization approaches. However, another fact to consider when choosing among forecasting models is their efficiency: There is an economic preference for inexpensive and easy-to-use methods if they promise satisfying results.

Context Information. The availability of weather forecasts is an essential condition for both physical and multiple-regression forecasting models, most importantly the quality of solar irradiation values. Predicted outside air temperature can be used to estimate a panel's future surface temperature, as energy output is reduced significantly on hot cells. In a similar manner, wind speed can indicate cooling effects. Further, technical information like the panels inclination angle and production year (due to the fact that their conversation efficiency decreases over age) are interesting. As for environmental influences, cleaning cycles are relevant because polluted panels will produce significantly less energy, which is a considerable influence in dry and dusty areas. Also, in some regions, detected snow coverage might prevent any energy output at all.

Forecast Horizon. Studies show that the forecast horizon for which a generated model has to be applied is an important parameter while choosing an appropriate prediction approach. In the energy domain, forecast horizons are determined depending on the requirements of the underlying business process. Usually, we can distinguish the following categories: now-casts (up to 4 h ahead), short-term (up to 7 days ahead) and long-term forecasts (more than 7 days ahead). Focusing on the grid operators activities related to renewable energy integration, we find that intra-day and day-ahead horizons represent the most relevant time scales for operations [22], while long-term predictions are of special interest for resource and investment plannings.

Spatial and Temporal Aggregation. Forecast quality of statistical or physical models will vary strongly depending on the level of spatial aggregation of the underlying energy time series. Since it is well known that the overall impact of single peak values decreases with increasing size of the chosen aggregates, forecasts computed on single or disaggregated time series usually contain the risk of higher average prediction errors. In contrast, following an agglomerative or bottom-up

Final edited form was published in "Data Analytics for Renewable Energy Integration: Second ECML PKDD Workshop. Nancy 2014", S. 108–121, ISBN 978-3-319-13290-7
http://dx.doi.org/10.1007/978-3-319-13290-7_9

approach by creating different aggregation levels might lead to better results on higher levels (e.g. average cloudiness in a region can be predicted more accurately than cloudiness at a particular site [14]), but complicates the integration of available context information, especially in the case of influences characterized by strong locality. The use of clustering techniques to create hierarchical aggregations of RES time series is a matter of a separate study [24] in progress. Temporal aggregation can be considered if the granularity of source time series needs to be reduced in order to obtain forecasts of lower resolution.

History Length. Stochastic approaches create a forecast model over the historical supply curves. The size of available history influences the accuracy of the forecasting result, as a longer history length might be suitable for learning repeatable patterns, while a shorter history length is more beneficial for strongly fluctuating time series. The latter requires a continuous adaption of the forecast models and, possibly, also of the history length. However, determining the best model parameters involves multiple iterations over the time series history which is an expensive process especially on large data sets. Reducing the history length can therefore speed up model creation significantly. Previous research in this area [9] proposes an I/O-conscious skip list data structure for very large time series in order to determine the best history length and number of data points for linear regression models.

4 Model Selection - A Case Study

In this section we analyze the impact of the previously described energy model selection parameters on the forecast output. After briefly introducing the forecasting methods to be assessed, we provide a description of our experimental setting including the used data set, the applied methodologies and the output evaluation criteria before we discuss the obtained results.

4.1 Predictor Description

Several forecasting algorithms have been chosen for our evaluation: (1) The Similar-Days model using Euclidean distance and (2) the univariate *Autoregressive Fractionally Integrated Moving Average* (ARFIMA) model that both are weather-unaware. Regression-based models are represented by (3) Mirabel[4], a scientific model based on principal component analysis and multivariate regression and (4) Epredict[5], a commercial library using the non-linear MARS algorithm. Additionally, (5) a domain-neutral multiple linear regression model from the OpenForecast[6] library is included. Hence, all classes of statistical models are covered except machine learning. The benchmark will be conducted against a naive model using diurnal persistence.

[4] http://www.mirabel-project.eu/
[5] http://www.robotron.eu/
[6] http://www.stevengould.org/software/openforecast/

Final edited form was published in "Data Analytics for Renewable Energy Integration: Second ECML PKDD Workshop. Nancy 2014", S. 108–121, ISBN 978-3-319-13290-7
http://dx.doi.org/10.1007/978-3-319-13290-7_9

Table 1. Sample data properties

Time series	Aggregation level (extension)	Peak P_{max}	Installations (capacity)
DIA	None	42.75 kW	1 (351.3 kW)
DSA	Distribution system (23 km^2)	257.45 kW	7 (3,440 kW)
TSA	Transmission system (109,000 km^2)	1,616.04 MW	107,216 (7,008 MW)

4.2 Methodology

The Data. To cope with the recently introduced model selection criteria of spatial aggregation, we include three observed solar energy output curves into our scenario: (1) A single, disaggregated PV-installation located in central Germany denoted as DIA, (2) an aggregate built of all measured PV-installations available in the same local distribution system denoted as DSA and (3) an aggregate build of all PV-installations attached to the superior transmission system denoted as TSA. DIA and DSA were provided by a cooperating distribution system operator[7], while TSA was obtained from a public website[8]. All time series have a resolution of 15 min and cover all of the year 2012. Corresponding weather data including measurements of solar irradiation, air temperature, and wind speed with a resolution of 1 h is available from a weather station run by a meteorological service[9], located within the distribution networks' range. Using weather observations instead of weather forecasts eliminates the naturally included NWP prediction error thus allowing for a unbiased evaluation of the energy model performance itself (Table 1).

Operational Framework. We use the first 11 months of historical data from our source time series for model training. Forecasts are computed for the remaining month, thus providing a test data set of 2976 predicted values according to our time series' resolution. To cover both intra-day and day-ahead terms with our scenario, we define varying forecast horizons of 2, 12 and 24 h ahead. After computing a forecast the training period is adopted by adding the forecast horizon length, thus extending the available history length accordingly with each completed iteration. A suchlike *moving origin* approach simulates the integration of newly arriving observations in the forecasting process, which can then be compared with the latest forecast output and used to adjust the forecast model. Therefore, the number of forecasting models required to cover the whole month is 372 for a horizon of 2 h, 62 (12 h) and 31 (24 h) respectively. Finally, the test data is split into a calibration, and an evaluation period.

Combination Approach. Various forms of combining forecasts have been developed like subjective human judgment or objective approaches, the latter

[7] http://www.en-apolda.de
[8] http://www.50hertz.com
[9] http://wetterstationen.meteomedia.de

Final edited form was published in "Data Analytics for Renewable Energy Integration: Second ECML PKDD Workshop. Nancy 2014", S. 108–121, ISBN 978-3-319-13290-7
http://dx.doi.org/10.1007/978-3-319-13290-7_9

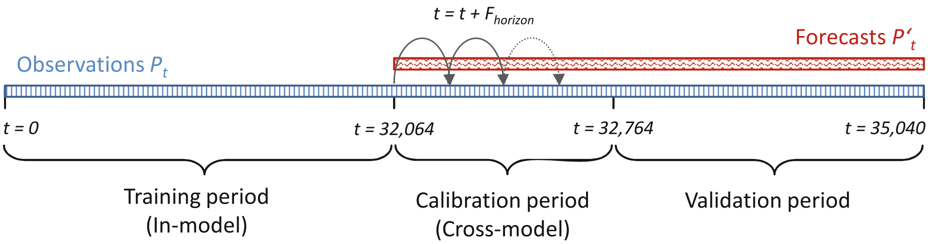

Fig. 4. Operational benchmark framework using moving origin

extending from the simple arithmetic mean to more sophisticated methods such as neural networks. A classification and identification of the most common methods can be found in [4]. In our study, we apply an objective method to combine the forecasts through a linear combination of n non-biased individual forecasts using an unrestricted weight factor λ_n for each forecast. The final energy forecast P'_t is then computed by

$$P'_t = \sum\nolimits_1^n \lambda_n P'_{nt} \tag{3}$$

where P'_{nt} is the forecasted value from model n for a timestamp t. In order to derive the optimal weight factors we use Nelder-Mead function minimization [20], which aims at reducing the error in the forecast output during the calibration period. After experimenting with different sets of input parameters, best results were obtained using the RMSE as target function and a calibration period of 700 values (approx. 1 week) as depicted in Fig. 4. Finally, several ensembles were computed using the n-best models in terms of RMSE denoted as *Ensble-nB* or all available individual models denoted as *Ensble-All*.

Output Evaluation. To evaluate the quality of the predicted values, different statistical accuracy metrics can be used for illustrating either the systematic or random errors included in the results. The *root mean square error* (RMSE) is the recommended measure and main evaluation criterion for intra-day forecasts, as is addresses the likelihood of extreme values better [13]. As the RMSE returns absolute values, normalization is applied in order to allow for model performance comparison on time series having different aggregation scales. The *normalized root mean square error* (nRMSE) is achieved by

$$nRMSE = \frac{100}{P_{max}} * \sqrt{\frac{\sum_{t=1}^{n}(P_t - P'_t)^2}{n}} \tag{4}$$

with P_{max} being the maximum observed power output in the validation data set. Beside the nRMSE, the systematic error can be expressed by computing the average error over the whole evaluation period. The *normalized mean bias error* (nMBE) is found by

$$nMBE = \frac{100}{P_{max}} * \frac{\sum_{t=1}^{n}(P_t - P'_t)}{n} \tag{5}$$

Final edited form was published in "Data Analytics for Renewable Energy Integration: Second ECML PKDD Workshop. Nancy 2014", S. 108–121, ISBN 978-3-319-13290-7
http://dx.doi.org/10.1007/978-3-319-13290-7_9

Table 2. Quality of forecast results using nRMSE evaluation metric

Predictor	DIA2	DIA12	DIA24	DSA2	DSA12	DSA24	TSA2	TSA12	TSA24
Naive	30.39	30.39	30.39	29.59	29.59	29.59	11.04	11.04	11.04
SimDays	21.28	21.59	26.77	19.77	20.32	24.23	6.08	7.17	8.54
ARFIMA	18.26	24.47	23.00	19.44	28.09	26.01	3.51	16.40	17.37
OpenFC	16.19	16.19	16.19	14.14	14.14	**14.14**	6.67	6.67	**6.67**
MARS	**13.99**	14.78	16.10	**12.81**	**13.33**	14.24	4.94	5.96	6.96
Mirabel	15.96	15.96	15.96	14.58	15.57	17.85	3.73	6.55	7.86
Ensble-2B	14.33	**14.61**	**15.71**	15.23	16.68	18.22	**2.49**	5.56	7.57
Ensble-3B	14.14	15.47	15.80	14.82	16.75	18.61	2.54	5.50	7.27
Ensble-4B	14.29	15.66	18.86	14.45	16.73	20.25	2.61	5.15	7.13
Ensble-All	14.25	16.18	18.12	14.56	17.03	18.12	2.62	**5.12**	7.05

and can be used to detect a systematic bias in the forecast, as according to Eq. 5 negative values represent over-estimations and vice versa. Note that non-daylight hours (values with timestamps before 8 am and after 4 pm) and all resting zero observation values are excluded from error calculation. The latter also implies that the effects of snow coverage or measurement failures are removed completely from the results.

4.3 Experimental Results

Our results listed in Table 2 show that in terms of RMSE, almost all forecasting models clearly outperform the naive benchmark with two exceptions being ARFIMA on TSA12 and TSA24. It is also visible that the uni-variate stochastic models SimDays and ARFIMA perform rarely better than those able to integrate external influences, especially on time series with lower aggregation levels. Regarding the impact of the chosen forecast horizon, we observe that the fitness of most of the models is decreasing with longer horizons (compare Fig. 5). In contrast, OpenFC seems to be completely unaffected by that parameter and provides constant values, thus leading to good results for day-ahead forecasts. We suspect that OpenFC can even outperform the sophisticated energy predictors MARS and Mirabel on short- and mid-term forecasts, which have not been covered by the presented scenario.

An analysis of the combined models shows that both improvements and degradations of individual results were obtained. The best results were provided using the 2 best models denoted as Ensble-2B: The RMSE of the best individual model could be reduced by 28.95 % for the TSA2 forecast (compare Fig. 6) and slight improvements were obtained on DIA12 and DIA24 with 1.11 % and 1.60 % respectively. Ensble-3B and Ensble-4B only outperformed the individual models once with the TSA12 forecast. According to expectations all evaluated models show the lowest preciseness on the individual PV-installation denoted as DIA,

Final edited form was published in "Data Analytics for Renewable Energy Integration: Second ECML PKDD Workshop. Nancy 2014", S. 108–121, ISBN 978-3-319-13290-7
http://dx.doi.org/10.1007/978-3-319-13290-7_9

Table 3. Quality of forecast results using nMBE evaluation metric

Predictor	DIA2	DIA12	DIA24	DSA2	DSA12	DSA24	TSA2	TSA12	TSA24
Naive	4.45	4.45	4.45	4.33	4.33	4.33	0.69	0.69	0.69
SimDays	0.21	−2.55	−9.02	−0.37	−2.03	−6.01	0.07	−0.11	−0.91
ARFIMA	1.24	−0.97	7.93	**0.03**	−1.96	12.24	−0.85	−11.59	−11.44
OpenFC	4.02	4.02	4.02	0.30	**0.30**	**0.30**	−2.57	−2.57	−2.57
MARS	1.87	2.34	3.97	−0.36	−0.98	−0.47	−1.18	−2.32	−2.74
Mirabel	3.80	3.80	3.80	−3.48	−3.79	−4.17	**−0.06**	−0.60	**−0.35**
Ensble-2B	**−0.06**	2.39	**3.51**	1.08	1.72	2.02	−0.27	−1.39	−1.98
Ensble-3B	0.07	2.72	3.69	0.24	1.68	2.41	−0.27	−0.93	−1.52
Ensble-4B	0.20	2.23	4.27	0.24	1.61	5.37	−0.36	−0.45	−1.10
Ensble-All	−1.09	**0.49**	3.69	0.33	1.00	5.09	−0.38	**−0.05**	−0.65

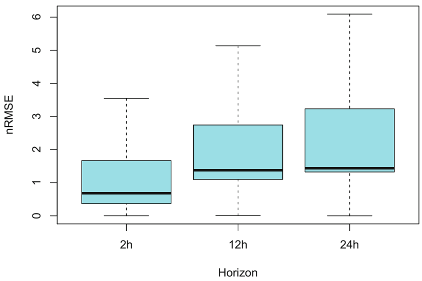

Fig. 5. Distribution of nRMSE values for varying horizons on TSA time series using MARS predictor

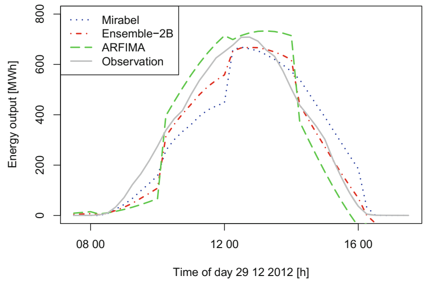

Fig. 6. Performance of Ensemble-2B and the underlying individual forecasts on TSA time series and 2h-ahead horizon

since single outliers can hardly be compensated as in the case of aggregated data like DSA or TSA. It is also noted that all models perform best on the transmission system level TSA. Although the correlation of weather information observed at only one specific location is not considered being a representative influence on that supra-regional level, the effect of weather-awareness seems to be completely neutralized by the impact of high aggregation. However, these results are not reflected using the nMBE evaluation criteria listed in Table 3, where in all cases except DIA24 at least one model could provide nearly unbiased forecasts having a nMBE value close to zero.

5 Conclusions

In this work we have shown that the forecasting of solar energy output is a two-step approach, typically requiring a weather- and an energy-forecast model. As for the energy-forecast, possible choices can be selected amongst physical, statistical, and hybrid models. The selection of an appropriate model depends

Final edited form was published in "Data Analytics for Renewable Energy Integration: Second ECML PKDD Workshop. Nancy 2014", S. 108–121, ISBN 978-3-319-13290-7
http://dx.doi.org/10.1007/978-3-319-13290-7_9

on the characteristics of the underlying data and the relevant evaluation criteria. Combining models offers additional optimization options whenever there is no model to be found that individually outperforms in all given situations, as demonstrated against the parameters of forecast horizon and spatial aggregation. However, deriving generalizable recommendations regarding the selection of appropriate models or model combinations based only on the evaluated use case remains challenging. For our future work, we think that conducting a more complex and global benchmark covering more state-of-the-art forecasting approaches and additional scenarios will provide useful information on how to systematically select an optimal energy model and might unlock the potential towards establishing industry standards regarding the application of forecasting strategies and output evaluation criteria.

Acknowledgment. The work presented in this paper was funded by the European Regional Development Fund (EFRE) and the Free State of Saxony under the grant agreement number 100081313 and co-financed by Robotron Datenbank-Software GmbH. We thank Claudio Hartmann and Andreas Essbaumer for supporting our work.

References

1. Alamsyah, T.M.I., Sopian, K., Shahrir, A.: Predicting average energy conversion of photovoltaic system in Malaysia using a simplified method. Renew. Energy **29**(3), 403–411 (2004)
2. Alfares, H.K., Nazeeruddin, M.: Electric load forecasting: literature survey and classification of methods. Int. J. Syst. Sci. **33**(1), 23–34 (2002)
3. Bacher, P., Madsen, H., Nielsen, H.A.: Online short-term solar power forecasting. Sol. Energy **83**(10), 1772–1783 (2009)
4. Mancuso, A.C.B., Werner, L.: Review of combining forecasts approaches. Indep. J. Manag. Prod. **4**(1), 248–277 (2013)
5. Chen, C., Duan, S., Cai, T., Liu, B.: Online 24-h solar power forecasting based on weather type classification using artificial neural network. Sol. Energy **85**(11), 2856–2870 (2011)
6. Dorvlo, A.S.S., Jervase, J.A., Al-Lawati, A.: Solar radiation estimation using artificial neural networks. Appl. Energy **71**(4), 307–319 (2002)
7. Dunea, A.N., Dunea, D.N., Moise, V.I., Olariu, M.F.: Forecasting methods used for performance's simulation and optimization of photovoltaic grids. In: IEEE Porto Power Tech Proceedings, Porto, Portugal (2001)
8. Friedman, J.H.: Multivariate adaptive regression splines. Ann. Stat. **19**(1), 1–141 (1991)
9. Ge, T., Zdonik, S., Tingjian, G., Stanley, B.Z.: A skip-list approach for efficiently processing forecasting queries. In: Proceedings of the VLDB Endowment, Auckland, New Zealand, pp. 984–995 (2008)
10. Heinemann, D., Lorenz, E., Girodo, M.: Forecasting of solar radiation. In: Wald, L., Suri, M., Dunlop, E.D. (eds.) Solar Energy Resource Management for Elitricity Generation from Local Level to Global Scale, pp. 83–94. Nova Science Publishers, New York (2006)

Final edited form was published in "Data Analytics for Renewable Energy Integration: Second ECML PKDD Workshop. Nancy 2014", S. 108–121, ISBN 978-3-319-13290-7
http://dx.doi.org/10.1007/978-3-319-13290-7_9

11. Iga, A., Ishihara, Y.: Characteristics and embodiment of the practical use method of monthly temperature coefficient of the photovoltaic generation system. IEEJ Trans. Power Energy **126**, 767–775 (2006)
12. Ji, W., Chee, K.C.: Prediction of hourly solar radiation using a novel hybrid model of ARMA and TDNN. Sol. Energy **85**(5), 808–817 (2011)
13. Kostylev, V., Pavlovski, A.: Solar power forecasting performance - towards industry standards. In: 1st International Workshop on the Integration of Solar Power into Power Systems, Aarhus, Denmark (2011)
14. Lorenz, E., Hurka, J., Heinemann, D., Beyer, H.G.: Irradiance forecasting for the power prediction of grid-connected photovoltaic systems. Sel. Top. Appl. Earth Observ. Remote Sens. **2**(1), 2–10 (2009)
15. Lorenz, E., Remund, J., Müller, S.C., Traunmüller, W., Steinmaurer, G., Pozo, D., Ruiz-Arias, J.A., Fanego, V.L., Ramirez, L., Romeo, M.G., et al.: Benchmarking of different approaches to forecast solar irradiance. In: 24th European Photovoltaic Solar Energy Conference, Hamburg, Germany, pp. 1–10 (2009)
16. Martin, L., Zarzalejo, L.F., Polo, J., Navarro, A., Marchante, R., Cony, M.: Prediction of global solar irradiance based on time series analysis: application to solar thermal power plants energy production planning. Sol. Energy **84**(10), 1772–1781 (2010)
17. Martins, I., Ferreira, P.M., Ruano, A.E.: Estimation and prediction of cloudiness from ground-based all-sky hemispherical digital images. In: Proceedings of the 2nd Ambient Computing Colloquium in Engineering and Education, Faro, Portugal (2011)
18. Meyer, E.L., van Dyk, E.E.: Development of energy model based on total daily irradiation and maximum ambient temperature. Renew. Energy **21**(1), 37–47 (2000)
19. Mu, Q., Wu, Y., Pan, X., Huang, L., Li, X.: Short-term load forecasting using improved similar days method. In: Asia-Pacific Power and Energy Engineering Conference (APPEEC), ChengDu, China, pp. 1–4 (2010)
20. Nelder, J., Mead, R.: A simplex method for function minimization. Comput. J. **7**(4), 308–313 (1965)
21. Ogliari, E., Grimaccia, F., Leva, S., Mussetta, M.: Hybrid predictive models for accurate forecasting in PV systems. Energies **6**(4), 1918–1929 (2013)
22. Reikard, G.: Predicting solar radiation at high resolutions: a comparison of time series forecasts. Sol. Energy **83**(3), 342–349 (2009)
23. Tao, C., Shanxu, D., Changsong, C.: Forecasting power output for grid-connected photovoltaic power system without using solar radiation measurement. In: 2nd IEEE International Symposium on Power Electronics for Distributed Generation Systems (PEDG), HeFei, China, pp. 773–777 (2010)
24. Ulbricht, R., Fischer, U., Lehner, W., Donker, H.: Optimized renewable energy forecasting in local distribution networks. In: Proceedings of the EDBT/ICDT Joint Conference Workshops, EDBT '13, Genua, Italy, pp. 262–266 (2013)
25. Wolff, B., Lorenz, E., Kramer, O.: Statistical learning for short-term photovoltaic power predictions. In: Proceedings of the ECML/PKDD Joint Conference Workshops, Prague, Czech Republik (2013)
26. Yona, A., Senjyu, T., Saber, A.Y., Funabashi, T., Sekine, H., Kim, C.-H.: Application of neural network to 24-hour-ahead generating power forecasting for PV system. In: IEEE Conversion and Delivery of Electrical Energy in the 21st Century, Pittsburgh, USA, pp. 1–6 (2008)
27. Zhang, G.: Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing **50**, 159–175 (2003)