Meteorol. Z. (Contrib. Atm. Sci.), PrePub DOI 10.1127/metz/2020/1024

ICEM 2019

A new sky imager based global irradiance forecasting model with analyses of cirrus situations

Anna Dittmann*, Nicolas Holland and Elke Lorenz

Fraunhofer Institute for Solar Energy Systems ISE

(Manuscript received December 20, 2019; in revised form April 20, 2020; accepted May 9, 2020)

Abstract

High resolution irradiance forecasts based on sky imagers are valuable for applications that require short term decisions based on ramps of solar irradiance. Here, we present our sky imager based forecasting algorithm, using images of a low cost surveillance camera. Model development and evaluation is done separately for the different steps in sky imager forecasting, starting with cloud detection, followed by estimation and extrapolation of cloud movement, and finally deriving irradiance forecasts from the predicted cloud images. We distinguish between clear and cloudy conditions and especially evaluate the effect of cirrus situations on the different forecasting steps. To create binary cloud masks, we adapted a pixel value based cloud algorithm using a set of manually classified pixels. In an independent validation dataset 90.3 % of the pixels are classified correctly. For the circumsolar region, where cloud decision is known to be especially challenging and crucial for the forecasting of the direct component of the solar radiation, we introduce a correction procedure using real-time irradiance measurements and object recognition methods. Applying this method we can significantly improve the cloud detection in the circumsolar region and increase the forecast skill of the cloud decision forecast. The development of the irradiance algorithm is a special focus of this paper. Real-time irradiance measurements and cloud decision information are used as input to our irradiance model. The algorithm is developed using cloud decision information derived from measurements instead of sky imager cloud decision forecasts to exclude the influence of errors in cloud decision and cloud motion methods for model development. Afterwards, the irradiance algorithm is applied to sky imager based cloud decision forecasts. Even though we start with binary cloud information, the distribution of the clear sky index from our forecasts is in very good agreement with the distribution of the measurements. In a validation dataset of 46 days, we receive a positive forecast skill for all forecast horizons larger than 1 min. We also apply our forecast chain to a dataset of two month from an independent measurement station resulting in a comparable forecasting performance.

Keywords: sky imager, short-term solar forecasting, irradiance model, cloud detection, circumsolar region

1 Introduction

Solar irradiance fluctuates with high amplitude on very short time scales due to passing clouds. Applications that are affected by these irradiance ramps require short term irradiance forecasts. Forecasts of solar irradiance based on Numerical Weather Predictions (NWP) or satellite data give valuable information for many applications (Lorenz et al., 2014). However, they cannot resolve the passing of single clouds and are thus not suited well for forecast horizons up to 15 min. Some applications require forecasts of single clouds with very high resolution on short time scales. In photovoltaic(PV)-Diesel systems, for example, fuel can be saved if the generator is only turned on when clouds are passing (Peters et al., 2018). Also a distribution grid can benefit from short term forecasts, since the fluctuating load in the grid increases due to an increase of large consumers like heat pumps. Short term forecasts can be used, e.g., to control heat pumps combined with PV-systems to be even beneficial for the grid operator (Dittmann et al., 2019). One way to create such high resolution forecasts is to use measurement-based statistical approaches of time series analyses like described in Chowdhury (1990) and Inman et al. (2013). The downside of these approaches is that they cannot forecast the timing of single clouds passing. This can be achieved by physical approaches.

Information on clouds and their motion in the surroundings of a given site can be obtained from ground-based cameras with fisheye lenses, so called sky imagers. The first steps of sky imager based forecasts are typically cloud detection and cloud motion estimation. To create a forecast, the cloud motion is extrapolated into the future and as a last step irradiance forecasts are derived from the predicted clouds. The maximum forecast horizon depends on the cloud situation, especially on cloud height and cloud velocity. Typical values are around 15 min (if one sky imager is used). The forecast resolution can be as small as seconds.

Here, we present our sky imager based forecasting algorithm for global horizontal irradiance (GHI), using images of a low cost surveillance camera. Our camera is installed in Freiburg, Germany, since 2017. We evaluate

*Corresponding author: Anna Dittmann, Fraunhofer Institute for Solar Energy Systems ISE, Heidenhofstrasse 2, 79110 Freiburg, Germany, e-mail: anna.dittmann@ise.fraunhofer.de

irradiance forecasts for the position of the camera for forecast horizons up to 15 minutes. Model development and evaluation is done separately for the different steps (described above), with a focus on the irradiance model.

The first who developed a sky imager based forecasting system were Chow et al. (2011). After that several studies addressed new ideas and developments in the area (e.g. Ghonima et al., 2012; Nguyen and Kleissl, 2014; West et al., 2014; Cazorla et al., 2015; Chauvin et al., 2016; Liandrat et al., 2017; Kuhn et al., 2018). New cloud detection methods have been suggested for example by Ghonima et al. (2012) and Cheng and Lin (2017). Here, the approach by Schmidt et al. (2016) is adapted. We optimize the parameters of the pixel value based cloud decision algorithm using a set of manually classified pixels. Additionally, we address cloud detection in the circumsolar region, which is known to be difficult, due to high saturation (Urquhart et al., 2013; Pfister et al., 2003). Yang et al. (2014) found that the cloud decision in the circumsolar region is one of most important factors affecting the forecast accuracy. In this study, we introduce a novel procedure to improve the circumsolar region of the cloud masks using real-time irradiance measurements and object recognition methods.

Different approaches to determine cloud motion from sky imager pictures have been investigated e.g. by Huang et al. (2013) and Chauvin et al. (2016). Furthermore, sky imager based forecasting can benefit from new methods published in the field of optical flow research which has been very active during the last years [1]. Here, we use the DeepFlow algorithm by Weinzaepfel et al. (2013) for calculation of cloud motion, which is suitable to detect also large displacements.

The last step in sky imager forecasting, the translation of the cloud forecasts to solar irradiance, is a focus of this study. A simple but often used method for irradiance retrieval is based on a histogram of the clear sky index from recent measurements of GHI (Yang et al., 2014). The clear sky index is defined as the ratio of measured GHI to the modelled clear sky irradiance ($k^* = GHI/GHI_{\text{clear}}$). Other methods calculate the irradiance components direct normal irradiance (DNI) and diffuse horizontal irradiance (DHI) separately from a sky image using image information (Kurtz and Kleissl, 2017; Schmidt, 2017). Nouri et al. (2019) developed a procedure to forecast DNI by determining the transmittance for each cloud from recent pyrheliometer and cloud height measurements and probability estimation from historical measurements.

Here, we address GHI forecasting for PV applications and use as a baseline the histogram method based on GHI measurements in combination with binary cloud decision information similar as proposed by Yang et al. (2014). We propose and evaluate several modifications to adapt and improve the irradiance retrieval for different weather situations. The algorithm is developed and optimized using cloud decision information derived from measurements at the forecast valid time instead of sky imager cloud decision forecasts in order to exclude the influence of errors in cloud decision and cloud motion methods for model development. Finally, overall evaluations for the entire forecasting chain from sky imager pictures to irradiance forecasts are performed for our site in Freiburg and, additionally, for a test-site in Blaustein.

Another focus of this study is the effect of cirrus clouds on cloud decision and irradiance forecasting. Schmidt et al. (2016) and Rodriguez-Benitez et al. (2018) analysed the forecast performance depending on the cloud type. They found that the cloud detection as well as the forecast performance for cirrus situations is difficult. Here, we discuss the effect of cirrus clouds on different forecasting steps and possible implications for future developments.

In the following sections we describe the development and validation of our sky imager based forecasting algorithm, as summarized in Fig. 1. In Section 2 we give an overview of the datasets we use. In Section 3 we introduce the error metrics used for the validation. In Section 4 we explain and evaluate our cloud detection algorithm and our procedure to improve cloud segmentation in the circumsolar region. In Section 5 we show our optical flow and cloud movement methods along with a validation. The main focus of this study is the irradiance forecasting. We explain the development and optimization of the algorithm in Section 6 and show the results of the final algorithm in Section 7. In Section 8 we discuss sky imager based forecasting for cirrus situations. The conclusion is given in Section 9.

2 Data

The camera used for the development of our algorithm is situated in Freiburg located in the Upper Rhine Valley next to the Black Forest in Southern Germany. This is an interesting location for sky cameras due to the formation of convective clouds over the mountains.

We use the surveillance camera FE9381 from Vivotek for this study. Images have been stored every 10 s since February 2017. Next to the camera, a sun tracker from EKO (STR-22G) is positioned, measuring DNI with a pyrheliometer (MS-56) and DHI with a pyranometer shaded by a shadow ball (MS-802F). Irradiance is measured instantaneously with a temporal resolution of 20 s. For this study GHI is derived from the irradiance components and the sun zenith angle Θ_z using $GHI = DHI + DNI \cdot \cos(\Theta_z)$. In the following, the GHI derived like this is referred to as measured GHI. For the detection of cirrus situations (Section 8) we also use cloud height measurements from a ceilometer operated by the German weather service DWD 1.5 km from the camera position. The cloud height measurements are 10 min averages.

In addition, we use data from a camera of the same type located in Blaustein, about 160 km from Freiburg,

[1] http://sintel.is.tue.mpg.de/results

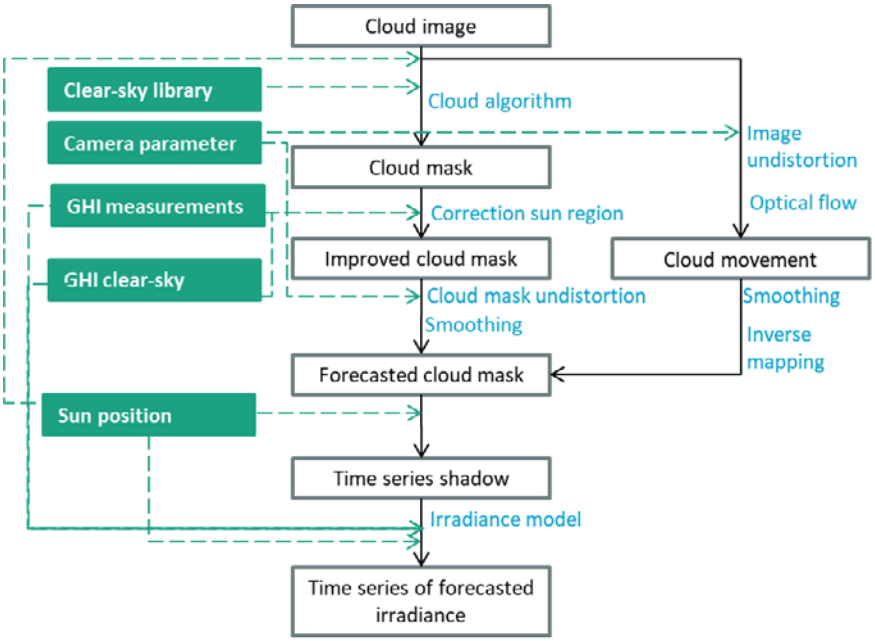


Figure 1: Steps of our sky imager based forecast.

for validation. GHI is measured with a ventilated pyranometer from EKO (MS80) next to the camera with a temporal resolution of 1 s.

For both sites all instruments are cleaned weekly. The irradiance data is filtered using a quality control procedure based on the Baseline Surface Radiation Network (BSRN) recommendations (Long and Dutton, 2002). Besides local measurements, clear sky irradiance values are used in this study to derive the clear sky index. For calculating clear sky irradiance, we use the clear sky model introduced by Dumortier (1995) with the turbidity described by Dumortier (1998) and Bourges (1992).

An overview of the datasets used in this study is given in Table 1. Model development and detailed evaluations were done using data from the Freiburg site. For the cloud detection algorithm two datasets were created, one for optimization (CloudDecision1) and one for validation (CloudDecision2 – in brackets). They consist of 28 (16) images selected carefully to give a good representation of cloud types and zenith angles. Cloud types of the images were assigned manually. Additionally, clear sky libraries were created from clear sky images in temporal proximity for each data set. In these images, 1068 (569) randomly selected pixels were classified manually into the categories cloudy or clear. From the first set of images an additional dataset containing 289 pixels within a radius of 20° around the sun was created (CloudDecisionSun) to validate the cloud detection in the circumsolar region.

For analysing the shadow and irradiance forecasts, a dataset from summer and autumn 2018 is used. The dataset contains 46 days and covers the periods 13 June 2018–5 July 2018 and 23 August 2018–14 September 2018. Only sun zenith angles up to 75° are analysed. Forecasts were started every 5 min with a maximum forecast horizon of 15 min and a resolution of 20 s. We choose these values to get a sufficient number of forecasts for validation and keep the computation time acceptable. This dataset (ForecastFreiburg) includes the sky images, calculated and predicted cloud masks, time series of shadow forecasts for the site of the camera and irradiance measurements and forecasts.

Additionally, we created a cloud shadow reference time series based on this dataset. With a cloud shadow reference time series, we mean a binary forecast created from measured GHI which is supposed to have the change from cloud to clear sky at exactly the right time. We used this dataset to validate our shadow forecast and as a basis to develop the irradiance algorithm in Section 6. For simplification we call a cloud shadow at the measurement site "cloudy", and no cloud shadow at the measurement site "clear" in the following. The cloud shadow reference time series is derived from measured GHI values by calculating clear sky index values and applying a threshold of $k^* = 0,9$ to distinguish between cloudy ($k^* < 0.9$) and clear ($k^* \geq 0.9$) as illustrated in Fig. 2. This threshold was chosen from visual analyses of time series and histograms of the clear sky index (see also Fig. 7). It is a balance between detecting thin clouds as clear (threshold too low) and detecting periods with high turbidity as cloudy (threshold too high). Evidently, applying a simple threshold is a simplification considering e.g. very thin clouds. Approaches for detecting clear sky periods from the literature like described in Tina et al. (2012) or Reno and Hansen (2016) are not suitable for our purpose here. They are designed to find stable clear sky periods using criteria like moving averages and variability information. They are not designed to detect short periods without cloud shadows

Table 1: Overview of the used datasets.

Dataset	Time range	Overview	Count	Purpose
CloudDecision1	May 2017	28 images	1068 classified pixels	Optimization of the cloud decision algorithm
CloudDecisionSun	″	″	289 classified pixels	Validation of the cloud decision in circumsolar region
CloudDecision2	June 2017	16 images	569 classified pixels	Validation of the cloud decision algorithm
ForecastFreiburg	June–Sept. 2018	46 days	5908 forecasts	Development of the forecasting algorithm
ForecastBlaustein	May–June 2019	61 days	9410 forecasts	Validation of the forecasting algorithm

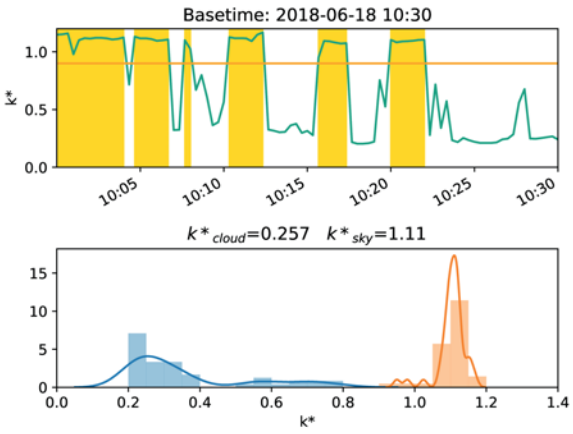


Figure 2: Top: time series of k^* from GHI measurements; a threshold of $k^* = 0.9$ (orange line) is applied to distinguish between clear (yellow) and cloud shadow (white). Bottom: normalised histograms and KDE for cloudy ($k^* < 0.9$, blue) and clear sky ($k^* \geq 0.9$, orange) for a forecast starting at 1030 UTC from the above time series; Each distribution and corresponding histogram are normalised separately.

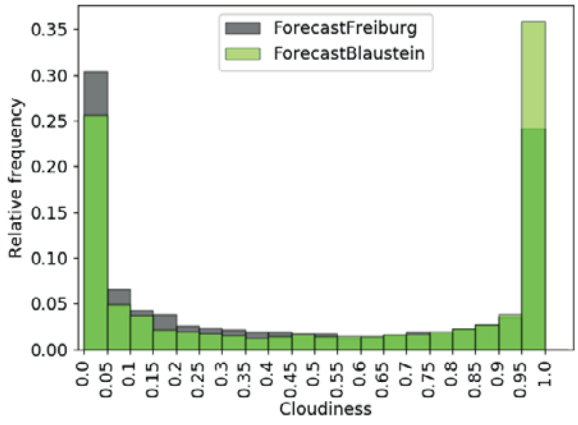


Figure 3: Histogram of the calculated cloudiness derived from the improved cloud masks as described in Section 4. Cloudiness is defined here as the fraction of pixels classified as cloudy: 0 corresponds to cloudless sky, 1 corresponds to a completely cloudy sky. Datasets ForecastFreiburg and ForecastBlaustein.

on a partly cloudy sky with fast fluctuations between clear and cloudy, which is essential for our method. We choose to use a method based on GHI, not on the direct and diffuse components, since these are not available for many measurement sites. With our method, we get a good indication on the presence of cloud shadows at the camera position and the timing of the change between cloudy and clear for many conditions.

As a completely independent validation dataset we use measurements and forecasts from the station in Blaustein (ForecastBlaustein). The dataset contains forecasts and measurements for May and June 2019. These forecasts have a resolution of 10 s.

Figure 3 shows a comparison of the cloud situations of the two datasets ForecastFreiburg and ForecastBlaustein based on the evaluation of the cloud masks derived with the algorithm described in Section 4. The dataset ForecastFreiburg (ForecastBlaustein) contains 30.4 % (25.7 %) clear, 45.4 % (38.5 %) partly cloudy and 24.2 % (35.8 %) overcast situations. Here, clear is defined as less than 5 % cloudy pixels, overcast as more than 95 % cloudy pixels and partly cloudy as everything in-between. The dataset ForecastBlaustein contains more overcast situations and less clear and partly cloudy situations than the dataset ForecastFreiburg.

Table 2: Contingency table.

	Model	
Observations	Clear	Cloudy
Clear	true negatives	false alarms
Cloudy	misses	hits

3 Error metrics

To validate the binary classification of the cloud forecasts, a contingency table is computed as shown in Table 2. This table is used to calculate the accuracy, defined as the number of correct classified samples divided by the total number of samples N.

$$\text{Accuracy} = \frac{\text{hits} + \text{true negatives}}{N} \tag{3.1}$$

For the validation of the irradiance forecasts p_i against the measurements m_i the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE) is calculated.

$$RMSE = \sqrt{\frac{\sum_{i=0}^{N}(p_i - m_i)^2}{N}} \tag{3.2}$$

$$MAE = \frac{\sum_{i=0}^{N}|p_i - m_i|}{N} \tag{3.3}$$

The skill of a forecast evaluates its performance in comparison to a trivial reference forecasts for a given score (here accuracy or RMSE) as defined e.g. in Sengupta et al. (2015). It is calculated from the score of the modelled forecast (S_{fc}), the score of a reference forecast (S_{ref}) and the score of a perfect forecast (S_{perf}).

$$\mathrm{Skill} = \frac{S_{\mathrm{fc}} - S_{\mathrm{ref}}}{S_{\mathrm{ref}} - S_{\mathrm{perf}}} \tag{3.4}$$

S_{perf} is one for the accuracy and zero for the RMSE. The skill has positive values if the score of the modelled forecast is better than the score of the reference forecast. It has a maximum value of 1.

As reference we use a persistence forecast derived from the last measurement. For the cloud shadow forecasts, we use persistence of the cloud shadow reference time series derived from measurements. For the irradiance forecasts, we combine persistence of the clear sky index with the clear sky irradiance to account for the diurnal cycle as defined e.g. in Sengupta et al. (2015), where t is the forecast valid time and t_0 is the start time of the forecast.

$$GHI_{\mathrm{pers}}(t) = GHI_{\mathrm{clear}}(t) \cdot k^*(t_0) \tag{3.5}$$

4 Cloud detection

For the calculation of the cloud masks the algorithm developed by Schmidt et al. (2016) is adapted. The algorithm is based on the evaluation of the red-to-blue-ratio of the cloud image pixels in relation to a reference clear sky image. Clear sky libraries were created manually for the different data sets described in Table 1. From the clear sky library, we automatically select the image with the smallest sun angle difference between the cloudy and the clear sky image. The algorithm contains three parameters that were adapted to our camera using the manually classified pixels from the dataset CloudDecision1. The parameters were optimized simultaneously by maximising the accuracy (Eq. 3.1) with an automated procedure using systematic parameter variation.

The contingency tables of the training and the validation dataset are summarized in Table 3 and Table 4, the calculated accuracy can be found in Table 5. In the independent validation dataset CloudDecision2, 90.3 % of the pixels are classified correctly. Using the cloud type classification of the images we find that many errors occur for cirrus situations. The pixels of cirrus clouds are hard to classify into cloudy and clear sky, since they have no sharp boundaries like other cloud types. Excluding four cirrus images, 96.5 % of the pixels could be classified correctly. Comparing the contingency tables (Table 3, Table 4) we see that this improvement is mainly caused by a reduction of false alarms, i.e. pixels classified wrongly as cloudy.

Due to the high saturation at the sun position and intense forward scattering in the circumsolar region, the cloud segmentation near the sun is known to be difficult (Yang et al., 2014; Urquhart et al., 2013). For validating the performance of our algorithm in the circumsolar region we created the dataset CloudDecisionSun from the images used for the CloudDecision1 dataset, containing only pixels in the circumsolar region. Excluding again cirrus situations (seven out of 28 images) we find that the accuracy in the circumsolar region of 94,1 % is lower than in the evaluation of the whole image (96.4 %).

Since the circumsolar region is of special importance for the forecasts, a novel procedure was developed for our forecasting system to correct the circumsolar region of the cloud masks. The procedure is based on real-time GHI measurements and pattern recognition in the cloud masks. There are two cases of false cloud classification in the circumsolar region: 1. The optical thickness of a cloud in front of the sun is not high enough to conceal the sun completely. It is classified as clear sky in a circle around the sun (Fig. 4, top). This occurs for nearly all cloud types, not only for cirrus clouds. 2. In clear sky

Table 3: Contingency tables for the cloud decision model for different datasets and of manually classified pixels using all available images.

Manual classification	Model	
CloudDecision1	Clear	Cloudy
Clear	532	31
Cloudy	21	479
CloudDecision2	Clear	Cloudy
Clear	245	43
Cloudy	12	269
CloudDecisionSun	Clear	Cloudy
Clear	164	67
Cloudy	7	152
CloudDecisionSun improved	Clear	Cloudy
Clear	166	65
Cloudy	4	155

Table 4: Contingency tables for the cloud decision model for different datasets of manually classified pixels excluding cirrus images.

Manual classification	Model	
CloudDecision1	Clear	Cloudy
Clear	286	7
Cloudy	21	475
CloudDecision2	Clear	Cloudy
Clear	151	4
Cloudy	11	257
CloudDecisionSun	Clear	Cloudy
Clear	120	10
Cloudy	7	152
CloudDecisionSun improved	Clear	Cloudy
Clear	120	10
Cloudy	4	155

Table 5: Accuracy of the cloud decision for different datasets of manually classified pixels using all images and excluding cirrus images.

Dataset	CloudDecision1	CloudDecision2	CloudDecisionSun	CloudDecisionSun improved
All images	95.1	90.3	81.0	82.3
Excluding cirrus images	96.4	96.5	94.1	95.2

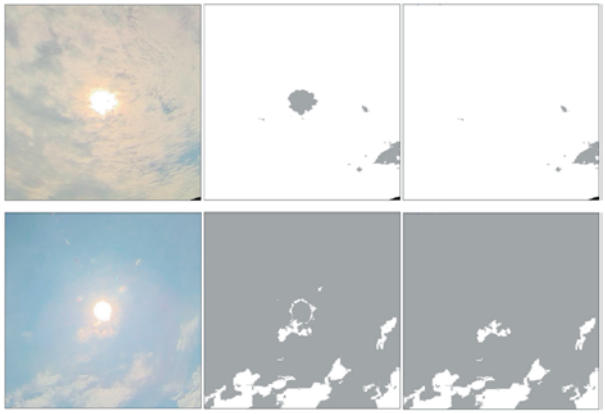


Figure 4: Improvement of the cloud mask in the circumsolar region; left: original image; middle: original cloud mask (cloud=white); right: improved cloud mask; the upper case shows a not detected cloud in the circumsolar region. The lower case shows a false detected ring around the sun.

situations, often a ring shaped cloud is detected around the sun as well as sparkles of clouds due to dirt on the camera dome (Fig. 4, bottom). To identify the first case, the cloud shadow reference time series derived from GHI measurements at the camera position is compared to the cloud classification for the sun position. If the cloud shadow reference indicates a cloud shadow, a circle detection (Yuen et al., 1990) is applied on the cloud mask to detect the cloud hole in the circumsolar region. This circle is then filled as cloud. The second case is identified if the cloud shadow reference shows clear sky. Using circle and object detection, clouds that form a ring around the sun and very small clouds in the circumsolar region are recognised and then removed. Examples of original and corrected cloud masks for the two cases are shown in Fig. 4.

This algorithm was developed using a subset of the ForecastFreiburg dataset and evaluated on the CloudDecisionSun dataset. From the 21 cloud masks (without cirrus situations), nine show a ring around the sun. Six of these cloud masks are corrected. A wrong cloud hole is found in six cloud masks. Three of them could be corrected. The validation of the corrected cloud masks with the classified pixels in the circumsolar region results in 95.2 % of correctly classified pixels. The contingency table (Table 4) shows that due to the correction procedure the number of missed cloudy pixels is reduced by three. The enhanced cloud masks have a big effect on the cloud decision forecasts (see Section 5).

5 Cloud and shadow forecasts

In order to create irradiance forecasts from cloud masks those masks have to be extrapolated into the future. Two consecutive images are used to compute an optical flow, which can then be used to compute cloud mask forecasts. The field of optical flow research is very active with new state-of-the-art methods being published every year[1]. We choose to use the DeepFlow algorithm by Weinzaepfel et al. (2013) also used in Dittmann et al. (2018) which was state-of-the-art in 2013 and is suitable for large displacements, which is necessary for fast moving clouds. The DeepFlow algorithm is applied to the undistorted images. For undistortion the calibration parameters of the camera lens are needed which were computed from images of a chessboard. For the implementation we use OpenCV (Bradski, 2000) for both the DeepFlow as well as the fisheye camera module.

Cloud mask pixels ($cm_t(x, y)$) are moved according to the optical flow using inverse mapping:

$$cm_{t+\triangle t}(x, y) = cm_t(-u(x, y) \cdot \lambda_{\triangle t} + x, -v(x, y) \cdot \lambda_{\triangle t} + y) \tag{5.1}$$

For each forecast step, the optical flow components u and v are multiplied with the factor $\lambda_{\triangle t}$ that is calculated as the forecast horizon $\triangle t$ divided by the time difference between the images used for flow calculation (here 10 s).

We found that applying smoothing on both the flow as well as the cloud mask yields better performance. Time series of shadow forecasts for the camera location are derived from the resulting cloud mask forecasts by evaluating the pixel corresponding to the position of the sun, because it determines whether the sun will be seen by the camera. The pixel position of the sun in the cloud mask is calculated using the sun angles and the calibration parameters of the camera lens.

We evaluate the shadow forecasts against the cloud shadow reference time series derived from measurements, using the dataset ForecastFreiburg. Fig. 5 shows the skill of the shadow forecasts using the original cloud masks and using the cloud masks with enhanced sun region. The correction considerably improves the forecasts, especially for small forecast horizons where there is no skill before. Since the circumsolar region is closest to the pixel that shades the camera, the probability that the pixels in the circumsolar region move over the sun pixel is higher than for pixels further away from the sun and small forecast horizons are affected more strongly by the correction. Now, the cloud decision forecasts have a positive skill from a forecast horizon of 1 min of up to 12 %.

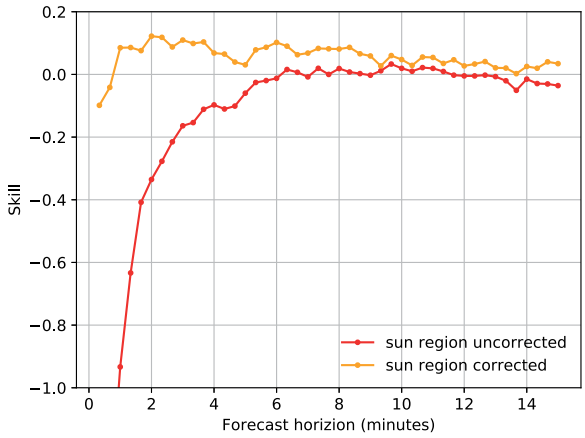


Figure 5: Influence of the enhanced cloud maps in the circumsolar region on the cloud shadow forecasts; the skill of the accuracy is plotted against the forecast horizon.

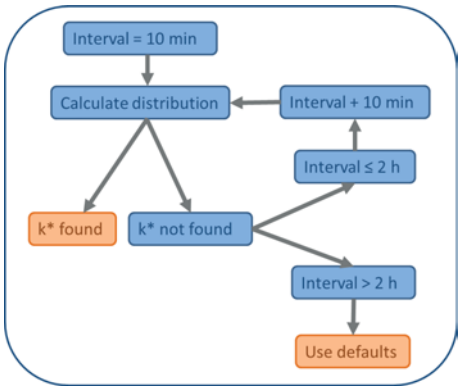


Figure 6: Schema for retrieving k^* peaks using a weather dependent length of the time interval of measurements.

6 Irradiance algorithm

The final step in irradiance forecasting is the translation of a binary shadow forecast to an irradiance forecast, i.e. assigning irradiance levels to both states cloudy and clear in order to describe the influence of turbidity, cloud optical depth and other factors.

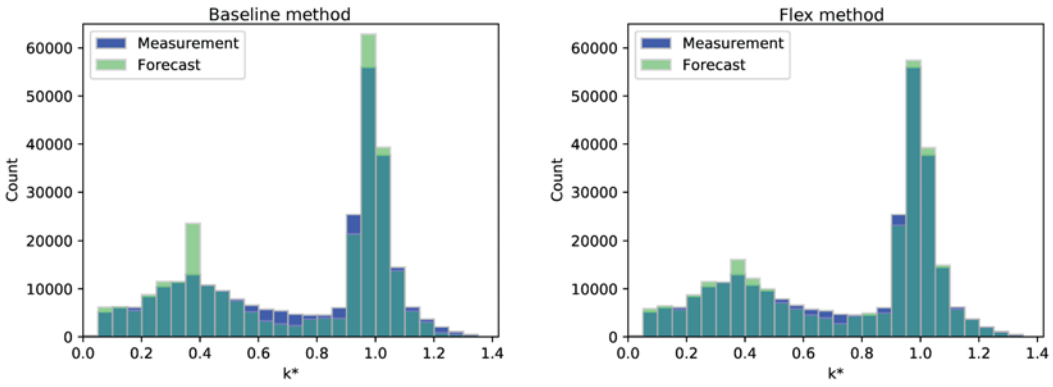
We tested various options using real-time irradiance measurements to estimate future irradiance levels for cloudy and clear. The most promising concepts are presented here. In order to investigate the performance of the irradiance algorithms independent of errors in sky imager based cloud decision and motion estimation, we use the following set-up to develop and optimize the algorithms: Instead of sky imager based cloud shadow forecasts we use the cloud shadow reference derived from GHI measurement at forecast valid time. We call this dataset perfect shadow forecasts, since the changes between cloudy and clear are supposed to be at exactly the right time, knowing that this is a simplification as described in Section 2. The perfect shadow forecasts are combined with recent irradiance measurements available at forecast calculation time in order to forecast the irradiance for the two levels. This study is performed using the irradiance measurements of the dataset ForecastFreiburg.

We develop the irradiance retrieval inspired by the histogram method introduced by Yang et al. (2014). This original method is based on the estimation of the levels of k^* for cloudy and clear from the peaks in a k^* histogram of recent measurements. The clear sky index values are combined with a clear sky model to derive the irradiance levels. If no peak is found, a default value is used. Here, we do not use a histogram, but a kernel density estimation (KDE) using the python package SciPy (Virtanen et al., 2019) to find the values of the k^* peaks as illustrated in Fig. 2. In this example, k^*_{cloud} is 0.225 for the histogram method and 0.257 for the KDE method. k^*_{sky} is 1.125 for the histogram method and 1.110 for the KDE method. One advantage of the KDE method is that the resulting k^* peaks are not binned.

The method developed by Yang et al. (2014) uses a fixed length of the time interval of measurements. We define this as our "baseline method" where we use a time interval of 30 min. However, different interval lengths may be suitable for different weather situations. On the one hand, the interval has to be long enough to collect enough data for shaded as well as not shaded measurements. On the other hand, for most situations it should be favourable to use a time interval as close to the present as possible, because the irradiance values associated with cloudy/clear can vary on small time scales depending on e.g. cloud type, cloud coverage and turbidity. In situations with fast fluctuation between cloud and clear sky, the time interval can be allowed to be shorter than in situations with larger or slower moving clouds.

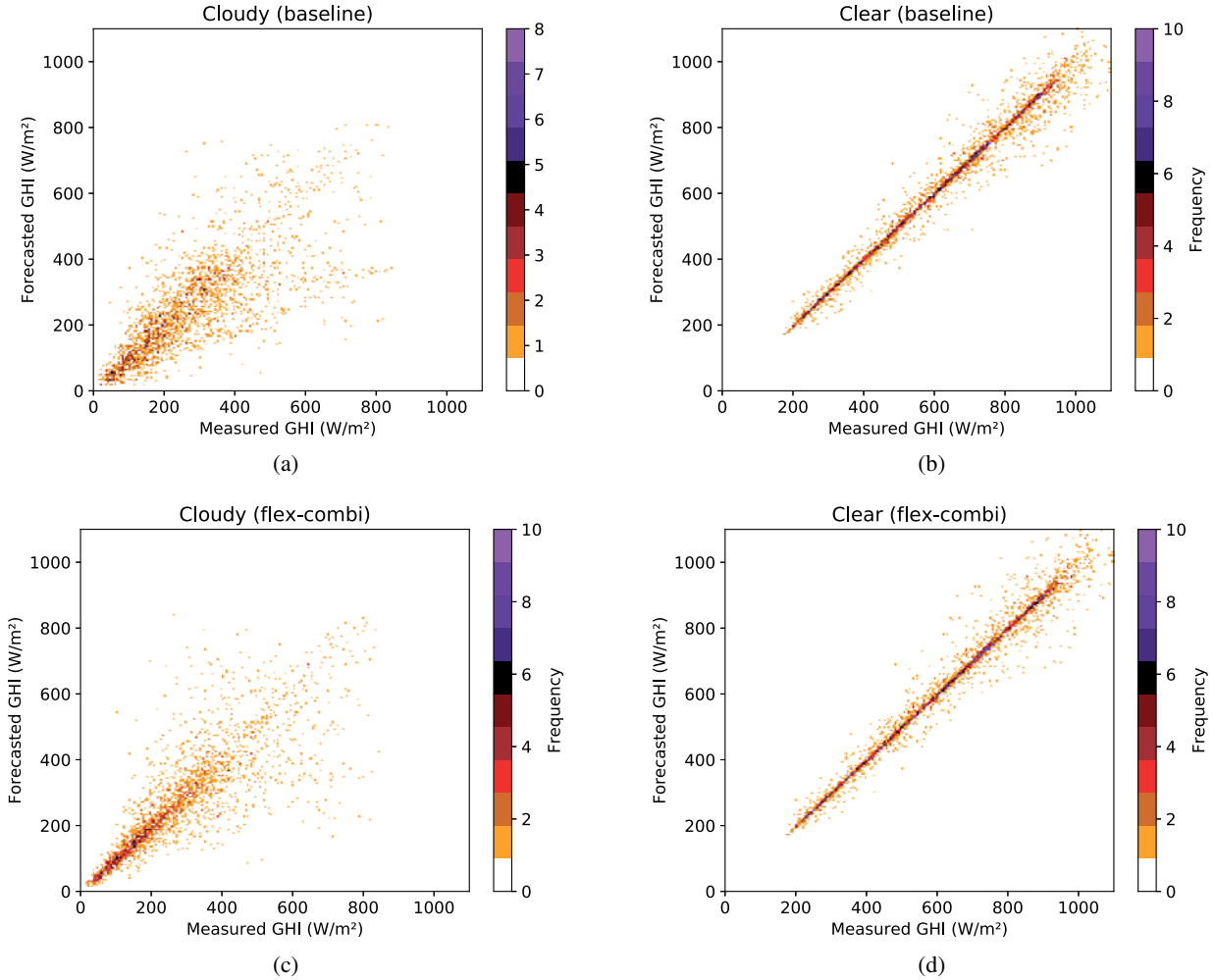
Therefore, we introduce a weather situation dependent time interval of measurements (Fig. 6). The length of the time interval is increased from 10 min to 2 h in steps of 10 min, as long as the k^* peak is not found. We tested two minimal intervals (10 min and 20 min) and two maximal intervals (2 h and 6 h) and found the combination of 10 min and 2 h to perform best. The forecast performance is found not to be very sensitive to these parameters. This method is applied for the k^* peak for cloudy and clear separately. If no k^* peak is found, 0.4/1.0 is used as default values for cloudy/clear. This irradiance retrieval method is called "flex method".

Fig. 7a and Fig. 7b show histograms of k^* for measurements and irradiance forecasts based on the perfect shadow forecasts for the baseline and the flex method. For the flex method the distribution of the forecasted k^* is more realistic compared to the measurements. Noteworthy are the peaks at the position of the default values for the baseline method. Due to the flexible length of the time interval the number of used default values is reduced for the flex method. For situations where the cloud decision derived from the last measurement (at forecast start time) and the forecasted cloud decision



(a) (b)

Figure 7: k^* Histogram for measurements and forecasts based on the perfect shadow forecast for the dataset ForecastFreiburg; (a) baseline method; (b) flex method.



(a) (b) (c) (d)

Figure 8: Forecasted against measured irradiance based on the perfect shadow forecast with a forecast horizon of 5 min; the correlation coefficient R is given in brackets. (a) baseline method, forecasted cloudy (R = 0.78); (b) baseline method, forecasted clear (R = 0.98); (c) flex-combi method, forecasted cloudy (R = 0.84); (d) flex-combi method, forecasted clear (R = 0.99).

are the same (i.e. both clear or both cloudy), another possibility is to use only the last measurement instead of a time-range of measurements. For these situations we can use persistence as defined in Equation 3.5. We found it valuable to analyse and optimize the irradiance retrieval method separately for clear and cloudy conditions. The difference between clear and cloudy is visualised in Fig. 8a and 8b for a forecast horizon of 5 min for the baseline method. The scattering is much larger for cloudy conditions than for clear sky conditions re-

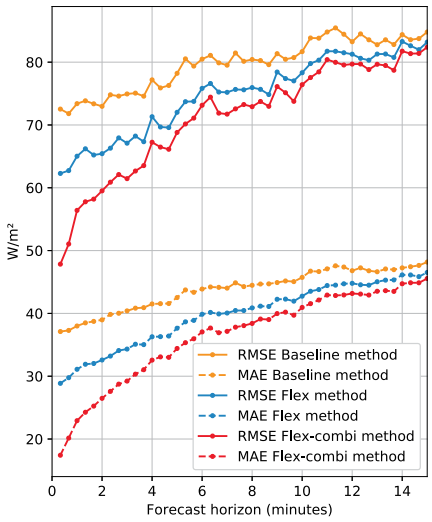


Figure 9: RMSE and MAE for three different irradiance retrieval methods based on the perfect shadow forecast.

sulting in a lower correlation coefficient of 0.78 for cloudy conditions compared to 0.98 for clear sky conditions. The k^* value for cloudy conditions is characterized by a higher variability and faster fluctuations. One reason for this is that the range of possible values of k^* is wider for cloudy than for clear (Fig 7). Furthermore, the clear sky index for cloudy conditions may show fast fluctuations due to variations in the cloud optional thickness. For clear sky, k^* is fluctuating much less, mainly influenced by comparatively slow turbidity variations.

This potentially fast fluctuations of irradiance levels for cloud shadows are a reason that for these situations persistence represents the irradiance at forecast valid time better than a longer time range of recent measurements, if persistence is cloudy as well. For clear conditions both methods perform similar. To forecast irradiance for all situations we use a combined method, called "flex-combi". We use

1. persistence, if cloud decision is cloudy for forecast calculation time and forecast valid time.

2. the flex method for all other situations.

Compared to the baseline method, the scattering is reduced for this method especially for forecasted cloudy conditions (Fig. 8c). The correlation coefficient for cloudy conditions increases from 0.78 for the baseline method to 0.84. Fig. 9 shows the RMSE and the MAE for the three compared methods for our study setup with the perfect shadow forecasts. The flex method reduces the RMSE as well as the MAE especially for low forecast horizons compared to the baseline method. The flex-combi method further decreases the errors for all forecast horizons. The forecast skill of the optimized method lies around 50 % for the perfect shadow forecast (Fig. 10a). It represents the maximal possible skill for our irradiance model using a binary cloud decision.

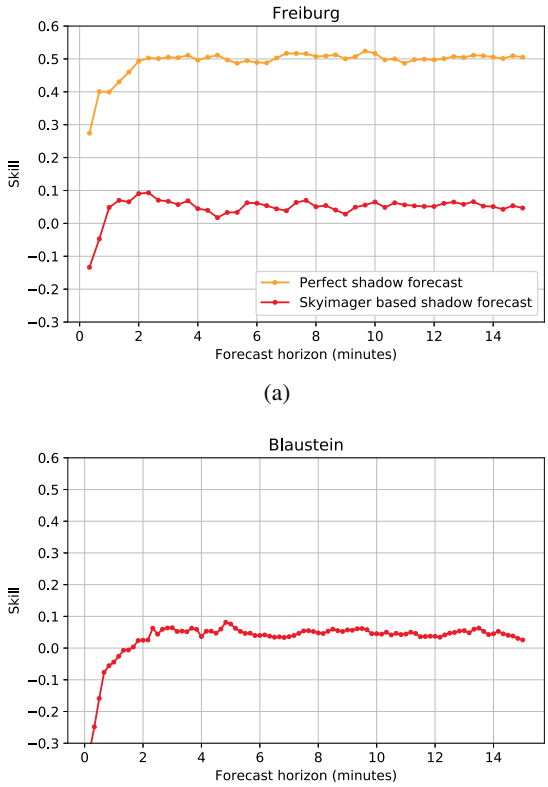


Figure 10: (a) Skill of the irradiance forecasts for the datasets Forecast Freiburg based on the perfect shadow forecast and the sky imager based shadow forecast; (b) Skill of the irradiance forecasts for the dataset ForecastBlaustein.

Based on these results we choose the model flex-combi and use it as our irradiance model for further evaluations.

7 Results of the sky imager based forecasts

Finally, we evaluate the complete forecasting chain combining the optimized irradiance model flex-combi with our sky imager based shadow forecasts for Freiburg and Blaustein. Fig. 11 shows the k^* distribution of the optimized sky imager based irradiance forecasts from the dataset ForecastFreiburg. It reveals a good agreement between the distributions of forecasts and measurements. This is a good performance of the irradiance model, considering that we started with a binary cloud decision forecast.

For the sky imager based shadow forecast, as expected, the skill is much lower than for the study case described above due to uncertainties in the forecasted cloud positions (Fig. 10a). Nevertheless, we get a positive skill for forecast horizons larger or equal to 1 min. The skill is in the same range as for the cloud decision forecast (Fig. 5).

To validate our forecast chain with an independent dataset we use data from the camera installed in

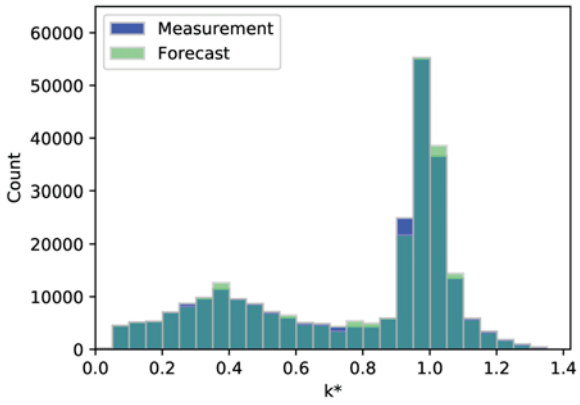


Figure 11: k^* Histogram for measurements and forecasts based on sky imager based shadow forecasts from the dataset ForcastFreiburg for all forecast horizons.

Blaustein. Fig. 10b shows that the skill is positive from a forecast horizon of 100 s on and in a similar range as for the station in Freiburg. The forecasts of different stations and time periods are not directly comparable since there are differences in cloud occurrence and weather situations. As shown in Fig. 3 this dataset contains a higher fraction of overcast situations than the Freiburg dataset. Also the horizon lines of the cameras differ. Nevertheless, we show that we can transfer our forecasting model to a different location without reducing the forecast performance.

8 Forecasting during cirrus situations

Since cirrus situations are identified as a challenge for cloud classification (Section 4), in agreement with other studies (Schmidt et al., 2016; Rodriguez-Benitez et al., 2018), the performance of the irradiance forecast is also investigated specifically for these situations. As simple filter for cirrus images, the cloud height measured by a ceilometer is used. In contrast to the images with spatial and instantaneous information, the ceilometer data gives temporal averaged point measurements of the cloud height. To select situations where a large part of the sky is covered by cirrus clouds, a moving minimum of three consecutive measurements of the lowest cloud layer is used. Thus, situations with lower clouds in part of the images or with mainly clear sky are mostly excluded. An image is categorized as cirrus, if this value is larger than 6000 m. The result was checked manually and was found to be an effective method to filter cirrus situations. 17 % of the images of the dataset Forecast-Freiburg are categorised as cirrus.

In Section 4 we found that the performance of the cloud detection is worse for cirrus situations than for other cloud situations. Fig 12a confirms this also for the cloud forecasts with a reduced skill for cirrus situations compared to all situations for all forecast horizons. Since cirrus clouds have no sharp boundaries and no homogeneous optical thickness, it is difficult to classify pixels with a binary system. The corresponding GHI measurements have no clear defined ramps since there is a smooth transition between cloud and clear sky (Fig. 13). Thus cirrus situations have a reduced accuracy for the persistence forecast as well as the sky imager based cloud decision forecast (Fig. 12c). In these situations, a binary classification is hard even manually and harder to train an algorithm for this task. Also our measurement derived cloud shadow reference is of limited reliability in these cases.

Looking at the irradiance forecast, the outcome is not so clear. For low forecast horizons the skill strongly fluctuates and is on average higher for cirrus situations than for all situations, for forecast horizons from 9 min on it is reverse (Fig. 12b). Cirrus clouds have a very low optical thickness and reduce GHI much less than other cloud types (Fig. 13). Therefore, the RMSE and the MAE are significantly lower for these situations for the persistence as well as for the sky imager based forecasts (12d, 12e). A low RMSE for persistence forecasts during cirrus situations compared to other cloud types was also found by Schmidt (2017). Since cirrus situations can be recognised well using cloud height information or more advanced algorithms, it is suggested to develop adaptations in the forecast procedure for these situations.

9 Conclusions

In this study, we present our sky imager based forecasting system. We describe all relevant steps of the forecast procedure and put a focus on cloud detection in the circumsolar region and on the step from the binary cloud forecast to the irradiance forecast.

Cloud detection is known to be especially challenging in the circumsolar region. We introduce a procedure to reduce the number of wrongly classified pixels in the circumsolar region using real time irradiance measurements and object recognition methods. We can significantly improve cloud detection in the circumsolar region and increase the forecast skill of the cloud decision forecast.

For developing and optimising the irradiance algorithm we introduce a measurement derived perfect shadow forecast in order to exclude errors and biases in the cloud decision and cloud motion methods. The irradiance algorithm determines k^* values for cloudy and clear sky conditions using recent measurements of GHI. Depending on the initial situation it uses either a Gauss based distribution of a weather situation dependent time interval of recent measurements or only the last measured GHI value. Even though we start with a binary cloud decision forecast, the distribution of the clear sky index from our forecasts is in very good agreement with the distribution of the measurements. In a validation dataset of Freiburg of 46 days, we receive a positive

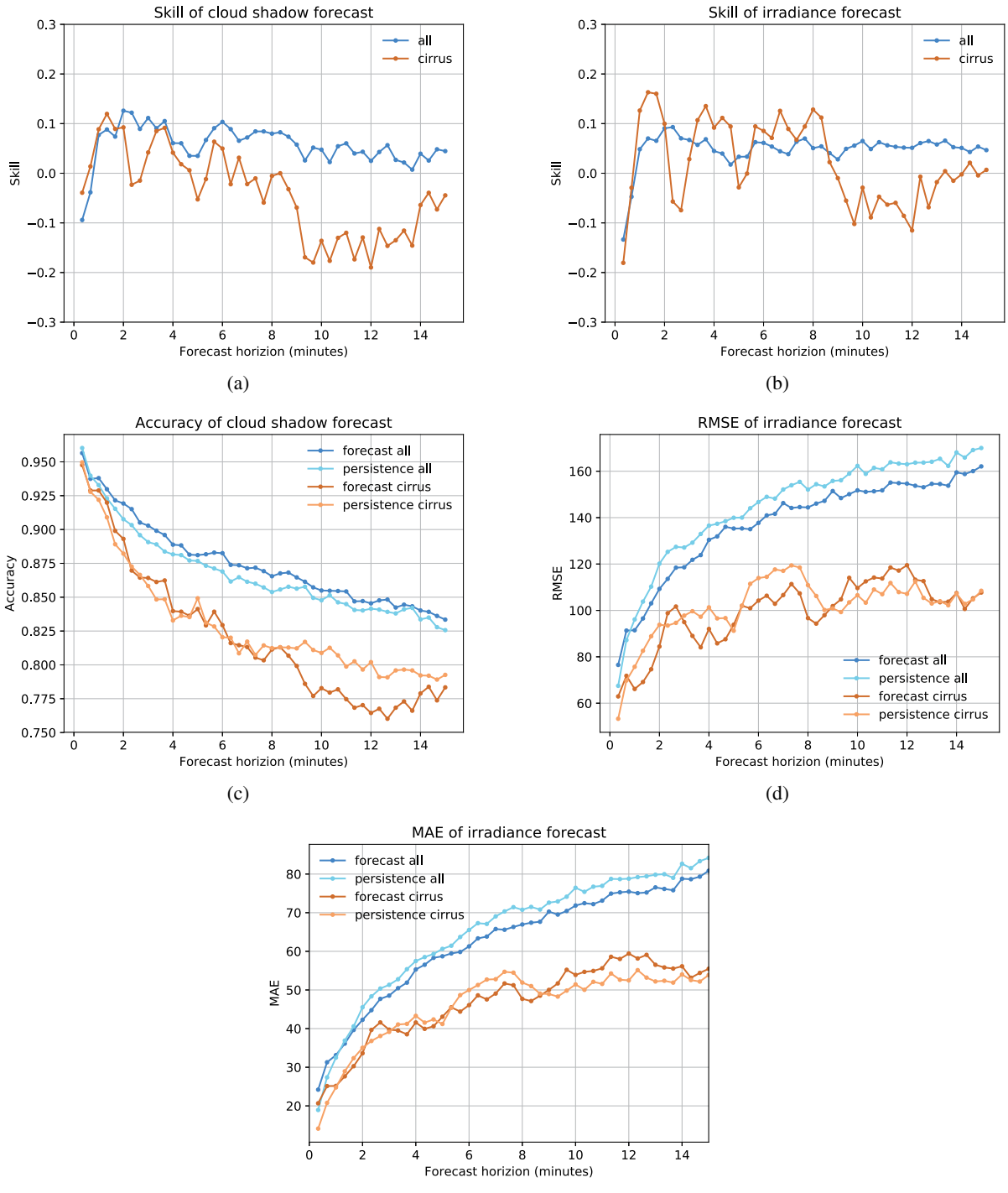


Figure 12: Validation of all data from the dataset ForecastFreiburg compared to cirrus situations; (a) Skill of the cloud shadow forecast; (b) Skill of the irradiance forecast; (c) Accuracy of the sky imager based cloud shadow forecast and the persistence; (d) RMSE of the sky imager based irradiance forecast and the persistence; (e) MAE of the sky imager based irradiance forecast and the persistence.

forecast skill for all forecast horizons larger than 1 min. Evaluations using the independent site in Blaustein show a similarly good performance.

We examine the performance of our forecasting model for different cloud situations and find that cloud detection is a challenge during cirrus situations. Since cirrus clouds have no sharp boundaries and no homogeneous optical thickness, it is difficult to classify pixels with a binary classification scheme into cloudy and clear sky. This shows the limitations of a binary cloud decision system. Though we find cloud detection is on average worse for cirrus situations, the impact on the quality of the irradiance forecasting is limited due to the low optical thickness of cirrus clouds.

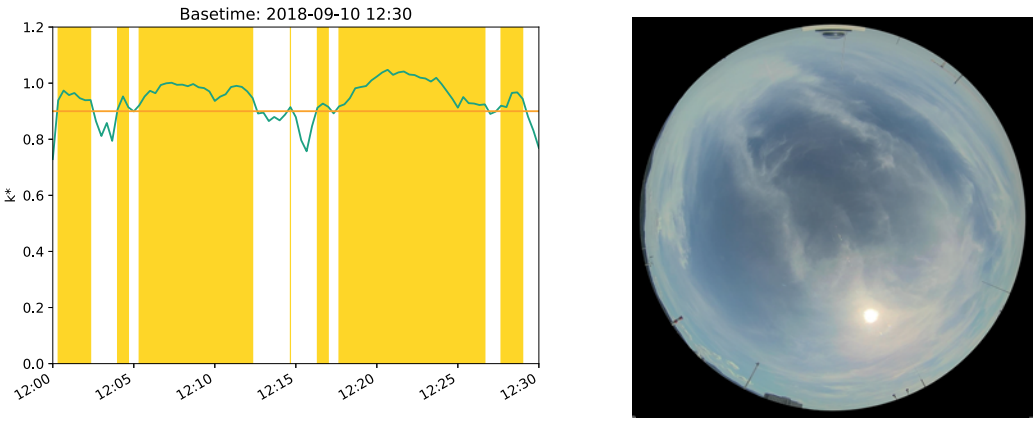


Figure 13: Left: Time series of k^* calculated from GHI measurements for a cirrus situation; Right: corresponding camera image with time stamp 2018-09-10 12:30.

In a next step we will investigate our forecasts using a network of eight measurement stations that we installed in a radius of 10 km around the camera in Freiburg. We plan to use these distributed observations combined with sky imager based shadow projections to develop an irradiance retrieval method based on machine learning and image features. Furthermore, we are currently using the forecasting chain introduced here to employ our forecasts for the management of heat pumps combined with a rooftop PV system. Especially, we investigate additional error metrics relevant for this application and optimise our forecasts with respect to the most relevant metrics.

Acknowledgments

Part of this study was funded by the "Ministerium für Umwelt, Klima und Energiewirtschaft" of Baden-Württemberg, as part of the project "PV Vorhersage für die netzdienliche Steuerung von Wärmepumpen".

References

Bourges, B. (Ed.), 1992: Climatic data handbook for Europe, volume 13537. – Kluwer Acad. Publ, Dordrecht.

Bradski, G., 2000: The opencv library. – Dr. Dobb's Journal of Software Tools.

Cazorla, A., C. Husillos, M. Antón, L. Alados-Arboledas, 2015: Multi-exposure adaptive threshold technique for cloud detection with sky imagers. – Solar Energy **114**, 268–277, DOI: 10.1016/j.solener.2015.02.006.

Chauvin, R., J. Nou, S. Thil, S. Grieu, 2016: Cloud motion estimation using a sky imager. – In: SOLARPACES 2015: International Conference on Concentrating Solar Power and Chemical Energy Systems, AIP Conference Proceedings, 150003. DOI: 10.1063/1.4949235.

Cheng, H., C. Lin, 2017: Cloud detection in all-sky images via multi-scale neighborhood features and multiple supervised learning techniques. – Atmos. Mea. Techn. **10**, 199–208, DOI: 10.5194/amt-10-199-2017.

Chow, C.W., B. Urquhart, M. Lave, A. Dominguez, J. Kleissl, J. Shields, B. Washom, 2011: Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed. – Solar Energy **85**, 2881–2893, DOI: 10.1016/j.solener.2011.08.025.

Chowdhury, B.H., 1990: Short-term prediction of solar irradiance using time-series analysis. – Energy Sources **12**, 199–219, DOI: 10.1080/00908319008960199.

Dittmann, A., E. Lorenz, N. Holland, C. Braun, W. Heydenreich, 2018: High resolution irradiance measurement network for validation and optimization of sky imager based forecasts. – http://presentations.copernicus.org/EMS2018-291_presentation.pdf.

Dittmann, A., E. Lorenz, P. Kober, O.V. Mier, H. Ruf, P. Schad, M. Schmidt, 2019: Optimierung der PV-Speisung von Wärmepumpen durch Kurzfristprognosen mit Wolkenkameras. – 34. Symposium photovoltaische Solarenergie, Staffelstein.

Dumortier, D., 1995: Modelling global and diffuse horizontal irradiances under cloudless skies with different turbidities. – Final Report Daylight II, JOU-CT92-0144, Vol 2.

Dumortier, D., 1998: The satellight model of turbidity variations in europe: Ecole nationale des travaux publics de l'etat, vaulx-en-velin, france. – Report for the sixth satellight meeting in Freiburg,Germany.

Ghonima, M.S., B. Urquhart, C.W. Chow, J.E. Shields, A. Cazorla, J. Kleissl, 2012: A method for cloud detection and opacity classification based on ground based sky imagery. – Atmos. Mea. Techn. **5**, 2881–2892, DOI: 10.5194/amt-5-2881-2012.

Huang, J., M. Korolkiewicz, M. Agrawal, J. Boland, 2013: Forecasting solar radiation on an hourly time scale using a coupled autoregressive and dynamical system (cards) model. – Solar Energy **87**, 136–149, DOI: 10.1016/j.solener.2012.10.012.

Inman, R.H., H.T. Pedro, C.F. Coimbra, 2013: Solar forecasting methods for renewable energy integration. – Prog. Energy Combustion Sci. **39**, 535–576, DOI: 10.1016/j.pecs.2013.06.002.

Kuhn, P., M. Wirtz, S. Wilbert, J.L. Bosch, G. Wang, L. Ramirez, D. Heinemann, R. Pitz-Paal, 2018: Field validation and benchmarking of a cloud shadow speed sensor. – Solar Energy **173**, 229–245, DOI: 10.1016/j.solener.2018.07.053.

Kurtz, B., J. Kleissl, 2017: Measuring diffuse, direct, and global irradiance using a sky imager. – Solar Energy **141**, 311–322, DOI: 10.1016/j.solener.2016.11.032.

Liandrat, O., S. Cros, A. Braun, L. Saint-Antonin, J. Decriox, N. Schmutz, 2017: Cloud cover forecast from a ground-based all sky infrared thermal camera. – In: Comerón, A., E.I. Kassianov, K. Schäfer (Eds.): Remote Sensing of Clouds and the Atmosphere XXII, Volume 10424, 19–31. – SPIE, DOI: 10.1117/12.2278636.

Long, C.N., E.G. Dutton, 2002: Baseline surface radiation network: Global network recommended QC tests V2.0. – BSRN techn. Report.

Lorenz, E., J. Kühnert, D. Heinemann, 2014: Overview of irradiance and photovoltaic power prediction. – In: Troccoli, A., L. Dubus, S.E. Haupt (Eds.): Weather Matters for Energy. – Springer New York, New York, NY, 429–454, DOI: 10.1007/978-1-4614-9221-4_21.

Nguyen, D., J. Kleissl, 2014: Stereographic methods for cloud base height determination using two sky imagers. – Solar Energy **107**, 495–509, DOI: 10.1016/j.solener.2014.05.005.

Nouri, B., S. Wilbert, P. Kuhn, N. Hanrieder, M. Schroedter-Homscheidt, A. Kazantzidis, L. Zarzalejo, P. Blanc, S. Kumar, N. Goswami, R. Shankar, R. Affolter, R. Pitz-Paal, 2019: Real-time uncertainty specification of all sky imager derived irradiance nowcasts. – Remote Sens. **11**, 1059, DOI: 10.3390/rs11091059.

Peters, D., T. Kilper, M. Calais, T. Jamal, von K. Maydell, 2018: Solar short-term forecasts for predictive control of battery storage capacities in remote pv diesel networks. – In: Sayigh, A. (Ed.): Transition Towards 100 % Renewable Energy, Volume 155. – Innovative Renewable Energy Ser. V, Springer, Cham, 325–333, DOI: 10.1007/978-3-319-69844-1_29.

Pfister, G., R.L. McKenzie, J.B. Liley, A. Thomas, B.W. Forgan, C.N. Long, 2003: Cloud coverage based on all-sky imaging and its impact on surface solar irradiance. – J. Appl. Meteor.**42**, 1421–1434, DOI: 10.1175/1520-0450(2003)042<1421:CCBOAI>2.0.CO;2.

Reno, M.J., C.W. Hansen, 2016: Identification of periods of clear sky irradiance in time series of ghi measurements. – Renewable Energy **90**, 520–531, DOI: 10.1016/j.renene.2015.12.031.

Rodriguez-Benitez, F., M.A. Pamos-Urena, C. Arbizu-Barrena, M.A. Cuesta, J. Tovar-Pescador, A.D. Pozo-Vazquez, 2018: Comparing sky-camera vs satellite solar radiation nowcasts. – EMS Annual Meeting, Budapest, 3–7 September 2018.

Schmidt, T., 2017: High resolution solar irradiance forecasts based on skyimages. – Phd Dissertation, Oldenburg University.

Schmidt, T., J. Kalisch, E. Lorenz, D. Heinemann, 2016: Evaluating the spatio-temporal performance of sky-imager-based solar irradiance analysis and forecasts. – Atmos. Chem. Phys. **16**, 3399–3412, DOI: 10.5194/acp-16-3399-2016.

Sengupta, M., S. Kurtz, A. Dobos, S. Wilbert, E. Lorenz, D. Renné, D. Myers, S. Wilcox, P. Blanc, R. Perez, 2015: Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications. – National Renewable Energy Laboratory (NREL).

Tina, G.M., S. de Fiore, C. Ventura, 2012: Analysis of forecast errors for irradiance on the horizontal plane. – Energy Conv. Man. **64**, 533–540, DOI: 10.1016/j.enconman.2012.05.031.

Urquhart, B., M. Ghonima, D. Nguyen, B. Kurtz, C.W. Chow, J. Kleissl, 2013: Chapter 9 – sky-imaging systems for short-term forecasting. – In: Kleissl, J. (Ed.): Solar Energy Forecasting and Resource Assessment. – Academic Press, Oxford, UK, 195–232, DOI: 10.1016/B978-0-12-397177-7.00009-7.

Virtanen, P., R. Gommers, Oliphant T.E., M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, van der Walt S.J., Brett, M. Wilson, J., K. Jarrod M., N. Mayorov, A.R.J. Nelson, E. Jones, R. Kern, E. Larson, C.J. Carey, Polat, Feng, Yu, Moore Eric W., J. Vander Plas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E.A. Quintero, C.R. Harris, A.M. Archibald, A.H. Ribeiro, F. Pedregosa, P. van Mulbregt SciPy 1.0 Contributors, 2019: Scipy 1.0 – fundamental algorithms for scientific computing in Python. – https://www.nature.com/articles/s41592-019-0686-2.pdf.

Weinzaepfel, P., J. Revaud, Z. Harchaoui, C. Schmid, 2013: Deepflow: Large displacement optical flow with deep matching. – ICCV – IEEE International Conference on Computer Vision 1385–1392.

West, S.R., D. Rowe, S. Sayeef, A. Berry, 2014: Short-term irradiance forecasting using skycams: Motivation and development. – Solar Energy **110**, 188–207, DOI: 10.1016/j.solener.2014.08.038.

Yang, H., B. Kurtz, D. Nguyen, B. Urquhart, C.W. Chow, M. Ghonima, J. Kleissl, 2014: Solar irradiance forecasting using a ground-based sky imager developed at UC San Diego. – Solar Energy **103**, 502–524, DOI: 10.1016/j.solener.2014.02.044.

Yuen, H.K., J. Princen, J. Illingworth, J. Kittler, 1990: Comparative study of hough transform methods for circle finding. – Image Vision Comput. **759**, 71–77, DOI: 10.1016/0262-8856(90)90059-E.