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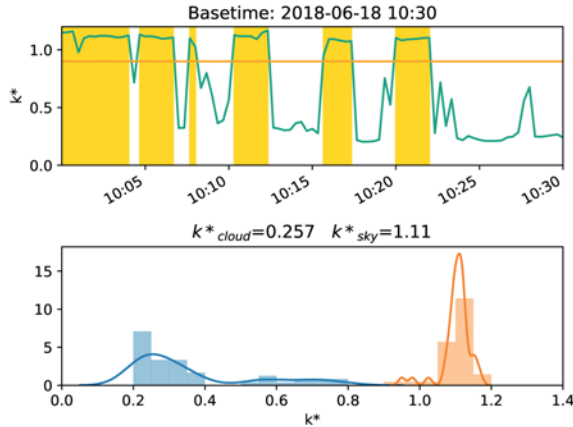
Figure 2: Top: time series of k^* from GHI measurements; a threshold of $k^* = 0.9$ (orange line) is applied to distinguish between clear (yellow) and cloud shadow (white). Bottom: normalised histograms and KDE for cloudy ($k^* < 0.9$, blue) and clear sky ($k^* \geq 0.9$, orange) for a forecast starting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

Table 2: Contingency table.

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

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.
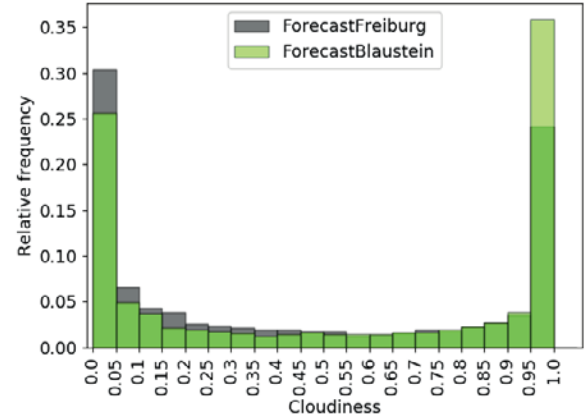
3 Error metrics

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.

$$\text{Accuracy} = \frac{\text{hits} + \text{true negatives}}{N} \qquad (3.1)$$

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.

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

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

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 SEN-GUPTA 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}).

$$\text{Skill} = \frac{S_{fc} - S_{ref}}{S_{ref} - S_{perf}} \quad (3.4)$$

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.

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

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

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

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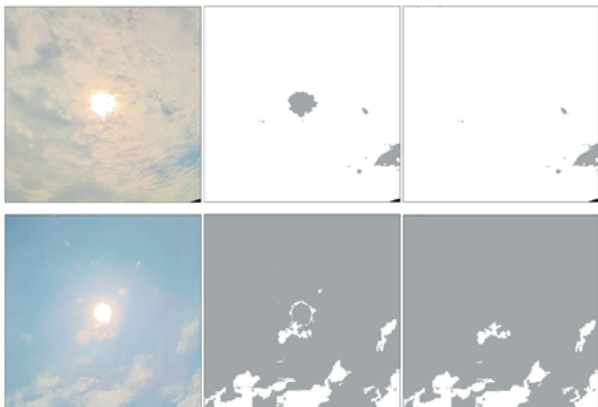
Figure 4: Improvement of the cloud mask in the circumsolar region; left: original image; middle: original cloud mask (cloud=white); right: improved cloud mask; the upper case shows a not detected cloud in the circumsolar region. The lower case shows a false detected ring around the sun.

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 Cloud-DecisionSun 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:

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

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 %.