Title
Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed

Permalink
https://escholarship.org/uc/item/5813b9j0

Journal
Solar Energy, 85(11)

ISSN
0038092X

Authors
Chow, Chi Wai
Urquhart, Bryan
Lave, Matthew
et al.

Publication Date
2011-11-01

DOI
10.1016/j.solener.2011.08.025

Peer reviewed
Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed

Chi Wai Chow¹, Bryan Urquhart¹, Matthew Lave¹, Anthony Dominguez¹, Jan Kleissl¹, Janet Shields², Byron Washom³
¹Department of Mechanical and Aerospace Engineering
²Marine Physical Laboratory, Scripps Institution of Oceanography
³Director, Strategic Energy Initiatives
University of California, San Diego

Phone: +1 858 534 8087
e-mail: jkleissl@ucsd.edu

Abstract

A method for intra-hour, sub-kilometer cloud forecasting and irradiance nowcasting using a ground-based sky imager at the University of California, San Diego is presented. Sky images taken every 30 seconds were processed to determine sky cover using a clear sky library and sunshine parameter. From a two-dimensional cloud map generated from coordinate-transformed sky cover, cloud shadows at the surface were estimated. Limited validation on four partly cloudy days showed that (binary) cloud conditions were correctly nowcast 70% of the time for a network of six pyranometer ground stations spread out over an area of 2 km². Cloud motion vectors were generated by cross-correlating two consecutive sky images. Cloud locations up to five minutes ahead were forecasted by advection of the two-dimensional cloud map. Cloud forecast error increased with increasing forecast horizon due to high cloud cover variability over the coastal site.
1. Introduction

Integration of large amounts of photovoltaics (PV) into the electricity grid poses technical challenges due to the variable nature of the solar resource. The ability to forecast solar irradiation will allow grid operators to better accommodate the variable electricity generation in their scheduling, dispatching, and regulation of power. Currently, physically based forecasting is primarily conducted using numerical weather prediction (NWP) and satellite cloud observations. NWP provides information up to several days ahead, however there are significant biases and random errors in the irradiance estimates (Remund et al. 2008, Lorenz et al. 2009, Perez et al. 2010, Mathiesen and Kleissl 2011). The spatial resolution of NWP is coarse at about 100 km², but there is active research on high-resolution rapid refresh models with grid cell areas of less than 10km² (Benjamin et al. 2010, Lara-Fanego et al. 2011). Either way, most clouds will remain unresolved in NWP. Frozen cloud advection based on GOES satellites images can provide accurate forecasts up to six hours ahead (Perez et al. 2010, Schroedter-Homscheidt 2009) at a resolution of 1 km².

To achieve high temporal and spatial resolution for intra-hour forecasts, NWP and Satellite forecasts are currently inadequate. Ground observations using a sky imager present an opportunity to fill this forecasting gap and deliver a sub-kilometer view of cloud shadows over a central PV power plant or an urban distribution feeder.

The advanced smart 42 MWp microgrid of the University of California, San Diego (UCSD) is one of the world’s most densely monitored environments with over 18,000 measurement points per km², including a hemispherical sky imager (Fig. 1). The UCSD microgrid provides a globally unique testbed for a customer side of the meter smartgrid with renewable generation, thermal, and electricity energy storage, demand side management, and demand response. In this paper, we demonstrate how intra-hour cloud shadows are forecasted and irradiance is nowcasted for the UCSD microgrid. In Section 2, we describe our experimental setup, present methods to obtain sky cover, cloud motion, irradiance, and to forecast cloud locations and irradiance. Existing literature on sky imaging systems and image processing is also reviewed in Section 2. In Section 3, we present the results for cloud shadow
mapping onto ground stations as nowcasting and cloud forecasting using sky imagery. Conclusions are presented in Section 4.

3 Nomenclature

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$alt$</td>
<td>altitude [m]</td>
</tr>
<tr>
<td>$am$</td>
<td>relative optical airmass</td>
</tr>
<tr>
<td>$dt$</td>
<td>forecast horizon [sec]</td>
</tr>
<tr>
<td>$e_m$</td>
<td>matching error</td>
</tr>
<tr>
<td>$e_{mp}$</td>
<td>matching error assuming cloud persistence</td>
</tr>
<tr>
<td>$e_{cap}$</td>
<td>cloud-advection-versus-persistence error</td>
</tr>
<tr>
<td>$f_{h1}$</td>
<td>$\exp(-alt/8000)$</td>
</tr>
<tr>
<td>$f_{h2}$</td>
<td>$\exp(-alt/1250)$</td>
</tr>
<tr>
<td>$GHI$</td>
<td>global horizontal irradiance [W m$^{-2}$]</td>
</tr>
<tr>
<td>$GHI_{csk}$</td>
<td>clear sky global horizontal irradiance [W m$^{-2}$]</td>
</tr>
<tr>
<td>$\Delta H$</td>
<td>cloud base height above ground level at TSI location [m]</td>
</tr>
<tr>
<td>$\Delta H(i,j)$</td>
<td>cloud base height above ground level at location i, j [m]</td>
</tr>
<tr>
<td>$l_o$</td>
<td>solar constant, 1366.1 [W m$^{-2}$] modulated by the inverse square of the earth-sun distance</td>
</tr>
<tr>
<td>$i, j$</td>
<td>indexed location on ground grid corresponding to a particular latitude, longitude and altitude</td>
</tr>
<tr>
<td>$kt$</td>
<td>Clear sky index, $GHI/GHI_{csk}$</td>
</tr>
<tr>
<td>$\ell_{lat}$</td>
<td>length of a unit latitude at the TSI latitude [m]</td>
</tr>
<tr>
<td>$\ell_{lon}$</td>
<td>length of a unit longitude at the TSI longitude [m]</td>
</tr>
<tr>
<td>$\ell_{lat}(0)$</td>
<td>latitude of TSI [°]</td>
</tr>
<tr>
<td>$\ell_{lon}(0)$</td>
<td>longitude of TSI [°]</td>
</tr>
<tr>
<td>$\ell_{i,j}$</td>
<td>longitude of location $i,j$ [°]</td>
</tr>
<tr>
<td>$P_{false}$</td>
<td>Number of false pixels</td>
</tr>
<tr>
<td>$P_{total}$</td>
<td>Number of total pixels</td>
</tr>
<tr>
<td>$\eta_{lat}$</td>
<td>resolution of the ground latitude grid [°]</td>
</tr>
<tr>
<td>$\eta_{lon}$</td>
<td>resolution of the ground longitude grid [°]</td>
</tr>
<tr>
<td>$T_L$</td>
<td>Linke Turbidity coefficient</td>
</tr>
<tr>
<td>$\ell_{lat}$</td>
<td>length of a unit latitude at the TSI latitude [m]</td>
</tr>
<tr>
<td>$\ell_{lon}$</td>
<td>length of a unit longitude at the TSI longitude [m]</td>
</tr>
<tr>
<td>$\ell_{lat}(i,j)$</td>
<td>column pixel coordinate in cloud map of solar vector intersection at height $\Delta H$</td>
</tr>
<tr>
<td>$\ell_{lon}(i,j)$</td>
<td>row pixel coordinate in cloud map of solar vector intersection at height $\Delta H$</td>
</tr>
<tr>
<td>$\theta$</td>
<td>pixel zenith angle [°]</td>
</tr>
<tr>
<td>$\phi$</td>
<td>solar azimuth angle [°]</td>
</tr>
<tr>
<td>$\phi_i$</td>
<td>pixel azimuth angle [°]</td>
</tr>
</tbody>
</table>

4 2. Methods

5 2.1 Experimental Setup

A rooftop mounted TSI 440A total sky imager (TSI, Yankee Environmental Systems, setup at 32.8852°N, 117.2400°W, 124 m MSL) has taken images of the sky over UCSD since August 2009 (Fig. 2). The instrument consists of a spherical mirror and a downward pointing camera. Images are taken every 30 seconds when the sun is above an elevation angle of three degrees. The camera provides images that are 640 by 480 pixels and the mirror occupies 420 by 420 pixels.
The image output settings by the TSI data acquisition software are default and cannot be changed. Images are 24-bit compressed jpgs. The compression process induces a small loss of information in the image. Additionally the system has an automatic gain adjustment to provide a larger dynamic range for the 8 bits on each of the red, green and blue channels. This adjustment process causes the intensity histogram of each channel to redistribute and thus the relative spectral composition of the images to fluctuate slightly with incoming signal strength. To determine sky cover obstructed by the camera arm and the shadowband, image masks are generated, edges of the masks are identified, and pixel values of the edge region are used to interpolate the region within.

### 2.2 Cloud Decision Algorithm

Sky imagers provide a visual measurement of a whole sky dome with high temporal and spatial resolution. Spatial resolution is determined by the camera's ability to resolve a single picture element (pixel), the solid angle subtended by this pixel, and the distance of the cloud from the imager which is a function of its cloud base height (CBH) and zenith angle (ZA). Temporal resolution is dictated primarily by the operational ability of the image processing algorithm. Atmospheric properties such as aerosol
optical depth, cloud optical depth, cloud cover, cloud type, and CBH have been retrieved using sky
imagers developed by research groups (Cazorla et al. 2008, 2009, Heinle et al. 2010, Kassianov et al.
( CCD) camera, a fisheye lens, an environmental housing, and possibly a solar occultor depending on
the choice of CCD sensor and application.

The TSI data acquisition software provides a cloud decision output image, however, we developed
our own algorithm and all images processing are done on raw image output. Our cloud detection
 technique is based on the concept of the ratio of the red channel to the blue channel developed at the
molecular (Rayleigh) scattering of shorter wavelengths causes the clear sky to be blue and the signal
measured in the blue channel to be larger. Since clouds scatter the visible wavelengths more evenly,
the red signal is similar to the blue signal. The red-blue-ratio (RBR) indicates whether the dominant
source of the scattered light incident on a pixel is the clear sky or a cloud.

The RBR in the whole sky hemisphere in clear conditions is not uniform due to aerosol and airmass
effects. Due to the size distribution of aerosols relative to the wavelength of incident visible light,
scattering by aerosols shows a weaker wavelength dependence than scattering by molecules
(McCartney 1976), which results in the scattered light appearing white. This is apparent near the
horizon where relative optical airmass increases and more total aerosols occupy the optical path
through the atmosphere. In the circumsolar region the forward scattering effects of aerosols are also
apparent as a white region around the sun. Because of these effects, a single RBR threshold for the
entire image is not appropriate. Rather a clear sky library (CSL based on Shields et al., 2009) as a
function of zenith and sun-pixel-angle was calculated from images on a cloud-free day (section 3.2). A
clear sky background image is then generated for each sky image based on the current solar ZA (Fig.
3b). A pixel is classified as cloudy if its RBR (Fig. 3c) is larger than the RBR in the clear sky
background image by more than a threshold value. The RBR of clouds is usually larger than the blue
sky. Outside the circumsolar region (small sun-pixel-angles), the method using the threshold from CSL is able to detect clouds accurately. However, within the circumsolar region, thick opaque clouds appear darker, have a lower RBR than the clear sky, and cannot be detected using the CSL. A sunshine parameter (SP, defined by Pfister et al. 2003) is used in addition to the CSL to improve cloud decision in the circumsolar region. SP is computed as the average RBR of the pixel area on the shadowband around the sun position (indicated in Fig. 3a). SP is a dynamic single-valued RBR threshold that changes from image to image, and is typically small when the sun is obscured. Pixels with RBR > SP are classified as cloudy even if the CSL indicates otherwise. By adding the clouds detected using the CSL (Fig. 3e) to the clouds detected using SP (Fig. 3d), the overall cloud decision image is obtained (Fig. 3f).

Figure 3: Processing chain of a sky image on October 4, 2009 15:45:30 PST (a) to obtain the cloud decision image. The sunshine parameter is 0.85 and is evaluated around the sun position indicated by the blue cross. The dotted black lines show the borders of the circumsolar region defined as solar azimuth ±35° and the solid black line shows ZA at 65°. (b) Clear sky RBR (colorbar) background image plus the threshold. (c) RBR (colorbar) image. (d) Pixels in (c) with RBR > SP or (e) RBR > CSL are assumed to be cloudy. (f) Shows the final cloud decision image. Green areas are clouds and blue areas are clear skies.
Once a cloud decision image has been generated, the cloud fraction is computed as the number of cloudy pixels divided by the total number of sky pixels with the solid angle correction described by Long et al. (2006). The cloud decision image is transformed to sky coordinates to geolocate clouds for cloud shadow mapping and forecasting using a geometric transformation that assumes constant CBH throughout the image

\[
\begin{bmatrix} x \\ y \end{bmatrix} = \Delta H \tan \theta_p \begin{bmatrix} \sin \phi_p \\ \cos \phi_p \end{bmatrix}.
\]

The CBHs were obtained from a ceilometer at Miramar Naval Air Station (KNKX, 32.8759°N, 117.1366°W, 141 m MSL, 9.5 km east of the TSI) since the imagery was taken before a ceilometer was installed at the site. The sky coordinate image (hereinafter cloud map) is restricted to ZAs less than 65° since the coordinate transformation near the horizon is not valid (Beaubien and Freedman 2001).

2.3 Irradiance Estimation

Prior techniques for estimating global horizontal irradiance (GHI) over an extended area consisted of satellite-derived cloud cover coupled with an empirical clear sky model (Cano et al. 1986, Schmetz 1989, Zelenka et al. 1999, Perez et al. 2002, Rigollier et al. 2004, Pereira et al. 2000, Martins et al. 2007). On the other hand, ground based instruments are typically point sensors and do not provide spatial irradiance information.

The primary factor modulating GHI from its clear sky value is the presence of a cloud between the ground location and the sun as this directly attenuates the solar beam irradiance. Variations in diffuse irradiance caused by changing cloud distribution and optical depth are smaller and generally negatively correlated with beam irradiance (a possible exception is cloud enhancement effects which are beyond the scope of this paper, Luoma et al. 2011). To estimate the sky conditions between a given ground point and the sun, the intersection of the vector to the sun from that point on the ground to the cloud map is computed. The ground around the TSI is gridded (hereinafter ground map) into 10 x 10 m cells
extending 5 km in each of the four cardinal directions. The pixel coordinate of the intersection of the solar vector with the cloud map for the \((i,j)\) element of the ground map is:

\[
x_p(i,j) = \frac{\tan \theta \sin \phi}{\ell_{\text{lon}} r_{\text{lon}}} \Delta H(i,j) + \frac{\text{lon}(i,j) - \text{lon}(0)}{r_{\text{lon}}},
\]

\[
y_p(i,j) = -\frac{\tan \theta \cos \phi}{\ell_{\text{lat}} r_{\text{lat}}} \Delta H(i,j) + \frac{\text{lat}(i,j) - \text{lat}(0)}{r_{\text{lat}}}.
\]

Actual spatial coverage of the GHI estimates within the 100 km² region considered varies with the sun’s position, cloud height and topography, e.g. when the clouds are low, the horizontal component of the distance to the cloud is smaller and thus the coverage is smaller.

To parameterize GHI, the Ineichen clear sky model (Ineichen 2002) as modified by Perez et. al. (2002) is used:

\[
\text{GHI}_{csk} = a_1 I_o \cos \theta \exp \left( -a_2 m \left( f_{h1} + f_{h2}(T_L - 1) \right) \right) \exp(0.01 a m^{1.8}),
\]

\[
a_1 = 5.09 \cdot 10^{-5} \text{alt} + 0.868,
\]

\[
a_2 = 3.92 \cdot 10^{-5} \text{alt} + 0.0387.
\]

This model requires the Linke turbidity factor (Linke 1922) as input and has a reported mean bias error of -6 W m⁻², and a root mean square error of 19 W m⁻² (Ineichen 2006). When the sun is occluded the GHI is assumed to equal 40% of the clear sky value. This value is chosen because the clouds that occurred on the days selected for validation were optically thick and reduced the irradiance to approximately this level. While the computed irradiance is not used in a quantitative evaluation, it is useful for a qualitative evaluation of the method as presented in section 3.4.

2.4 Cloud Motion

Cloud velocity and direction of motion is determined through the cross-correlation method (CCM) applied to two consecutive sky images (Hamill & Nehrkor n 1993). Before applying the CCM, images
are projected into sky coordinates to remove geometric distortion. The projected sky image is partitioned into subsets of pixels of equal size such that each subset is about 1% of the sky image area. The CCM finds the position that best matches each given subset of pixels in the previous sky image within the current image (Fig. 4). The CCM yields the wind vector (direction and speed) with the largest cross-correlation coefficient (CCC) that specifies the quality of the match. Since the CCM is computationally expensive, the search area is limited to a search distance that is five times the size of the original subset. This search distance effectively sets the largest detectable cloud speed. In general, the CCM obtains a larger correlation coefficient on an area with a high contrast of pixel values, such as a cloud with a sharp boundary in a clear sky. The CCM is performed on the red channel image which has a higher contrast between clear sky and cloud than the blue and green channels (the cloud decision image would have even higher contrast because it is a binary image but no texture information). Future work will use the RBR image because this has higher contrast than the red image.

Figure 4: Illustration of the cross-correlation method using two images on October 4, 2009 at 16:18:30 (a) and 16:19:00 PST (b). Each region of pixels from (a) is correlated to (b) within a search distance. The location of the highest correlation is found and a motion vector is defined for each box. The cloud velocity determined from this pair of images was 5.2 m s\(^{-1}\). While the method is illustrated here on a full color sky image, the method is actually applied to the coordinate transformed (Eq. 1) red channel image.

Assuming spatial homogeneity of cloud velocity, the vector field obtained through the CCM is further processed using several quality controls to yield an average cloud velocity across the image.
The circumsolar region is often offscale bright (when a pixel registers the highest possible 8-bit value) and shows little variation over time in clear conditions. This results in erroneously small motion vectors and hence vectors in this region are removed. Using the information in the cloud decision image the CCM is performed only on areas that contain clouds. Only motion vectors with CCC higher than 0.8 are considered. Finally, vectors that are more than one standard deviation from the mean in either the \( x \) or \( y \) component of the velocity are eliminated. The remaining vectors are vector-averaged to yield the advection velocity of the cloud field. We note that this method needs further development for the case of several cloud layers with different velocity vectors.

### 2.5 Cloud Forecasting

To forecast cloud cover, the cloud map at time \( t_o \) is advected at the speed and direction of the global vector determined from cross-correlating the images at time \( t_o \) and \( t_o - dt \) (\( dt = 30 \) seconds). To determine accuracy, the actual cloud map at time \( t_o + dt \) (Fig. 5b) is overlaid onto the advected cloud map (Fig. 5c) to determine the pixel-by-pixel forecast error (Fig. 5d). Since the circumsolar region has a large potential for erroneous cloud decision, the sky image is divided into the circumsolar region (within 35° of the solar azimuth angle) and the outer region (Fig. 3a). The matching error between the two cloud maps is

\[
e_m = \frac{P_{\text{false}}}{P_{\text{total}}} \times 100\%,
\]

which is the ratio between the number of falsely forecasted pixels and the number of total pixels in a region. The cloud-advection-versus-persistence (cap) error:

\[
e_{\text{cap}} = \frac{e_m}{e_{m,p}} \times 100\%
\]

describes the forecast error obtained by cloud advection (Fig. 5d) divided by the error obtained if the image at \( t_o \) was assumed to persist until \( t_o + dt \) (no advection). An \( e_{\text{cap}} < 1 \) implies that the cloud advection improves the forecast compared to persistence.
Figure 5: The forecast cloud map (b) is produced by advecting the cloud map at time $t_0$ (a) in the direction of the motion vector (here XX degrees). To determine the forecast error (d), the future cloud map at time $t_0 + 30$ seconds (c) is compared to the forecast (b). Blue and red colors in (d) show forecast errors (blue: pixel forecast cloudy and but actually clear; red: pixel forecast clear but actually cloudy) and white indicates accurate forecasts.

3. Results

3.1 Selection of Days

To assess the accuracy of sky imager forecasts, we selected relevant scenarios with typical sky conditions and high image quality. In our binary (cloud / no cloud) system, entirely clear or overcast days will always result in perfect forecast and these days are eliminated. Furthermore days with multiple cloud layers moving in different directions, rapid cloud deformation, formation, or evaporation were removed. Lastly the heating system on the imager failed for part of the year leading to excessive dirt aggregation on the mirror because of dew that made cloud detection impossible. Single layer broken clouds without rapid deformation moving across the sky are thought to be the best scenario to test the potential of the method. Sky condition occurrence frequency measured from a ceilometer at KNX in 2010 (Table 1) provides context as to the fraction of time this method is useful in the present form. The following days were analyzed: September 14, 2009; October 4, 2009; March 4, 2010; and March 10, 2010.

<table>
<thead>
<tr>
<th>Sky conditions</th>
<th>% Occurrence</th>
<th>Layers during Partly Cloudy</th>
<th>% Occurrence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear sky</td>
<td>18.2</td>
<td>Multiple cloud layers</td>
<td>34.8</td>
</tr>
<tr>
<td>Overcast</td>
<td>22.5</td>
<td>Single cloud layer</td>
<td>65.2</td>
</tr>
<tr>
<td>Partly cloudy</td>
<td>59.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.2 Cloud Decision

Since cloud decision images can only be validated visually, we describe here qualitative experiences of applying our method. In Fig. 6a, sky imager ZAs larger than 75° correspond to the horizon, and small sun-pixel-angles (approximately < 35°) correspond to the circumsolar region, respectively. It can be seen in Fig. 6a that clear pixels away from the sun and horizon have rather uniform RBR. Decreasing the distance from sun and horizon increases RBR because the forward scattering of sunlight turns the horizon and the circumsolar region whiter and brighter.

Figure 6: (a) Average RBR (colorbar) as a function of zenith angle and sun-pixel-angle based on TSI imagery on a clear day (September 24, 2009). This image is used to generate the clear sky library in Fig. 3b. (b) Standard deviation (colorbar) of clear sky RBR and (c) Number of occurrence (colorbar) for each zenith angle and sun-pixel-angle pair.
Figure 6b presents the standard deviation of the clear sky RBR to quantify the RBR variation during the (clear) day in Fig. 6a. Larger variation of RBR leads to less accurate cloud decisions since it makes it more likely that a clear sky RBR would be larger than the threshold and classified as cloudy. For sun-pixel-angles greater than 50° and sky imager ZA less than 60° (i.e. in the region with the highest occurrence, Fig. 6c) the variation in the clear sky RBR is small and our method performs well. However, the standard deviation of the RBR in the circumsolar region is found to be largest when the sun is near the horizon (low sun-pixel-angle and high zenith angle). Anecdotal observations confirm that clear conditions in the circumsolar region at low sun elevation are often misclassified as cloudy while the horizon and circumsolar region appear white. In addition, at larger solar ZAs less sunlight scattered by clouds reaches the TSI due to shading by other clouds, increased optical path length through the atmosphere, and near 90° scattering angles. These effects cause clouds to be dark in color at the cloud-base and results in a smaller RBR than the CSL leading to incorrect clear classification.

Optically thin clouds are bluish because the background clear sky is visible through the cloud and are notoriously difficult to detect automatically (Shields et al. 1998) as their RBR is only slightly larger than the clear sky. If the RBR threshold is lowered to detect thin clouds it will increase false cloudy (cloud detected but no cloud exists) detections, especially in the circumsolar region. For our application of solar forecasting and sky imagers located in proximity to solar power plants, the CSL threshold is optimized to provide a more accurate cloud decision in the circumsolar region at the expense of thin cloud detection.

Optically thick clouds have smaller RBR than white clouds and bright regions. Optically thick clouds can be detected accurately in regions with small CSL RBR (Fig. 6a). However, near the horizon or the circumsolar region, thick clouds are often misclassified depending on the sun obscuration. If the sun is unobscured (not shown), SP is large and dark clouds near the sun are misclassified as clear. When the sun is obscured, dark clouds are correctly classified due to the SP algorithm (Fig. 7).
To summarize the errors: cloud decision when the sun is near the horizon is not reliable; false clouds appear on the horizon and in the circumsolar region; false clear conditions occur for dark clouds on the horizon and in the circumsolar region; and optically thin clouds are not detected. Cloud decision behind the shadow band and camera arm is expected to be unreliable since its RBR is interpolated from surrounding areas. Differing aerosol optical depth between the day when the CSL was generated and the day under consideration can degrade the accuracy of the cloud decision.

3.3 Cloud Motion

Figure 8 illustrates the computed motion vector field at several stages in the quality control process. A reduction in the CCC usually implies smaller accuracy of motion vectors. Reduced CCC can be the result of cloud formation, deformation, and evaporation. Also the appearance of a cloud may change over time due to different camera white balancing, lighting on the cloud, and different viewing geometries.

After removing vectors in the circumsolar and clear sky region, and vectors with CCCs less than 0.8, most remaining vectors are uniform (Fig. 8b). By computing the average and standard deviation of the vector field, the remaining outlier vectors are eliminated (Fig. 8c).
Figure 8: Illustration of procedures for obtaining a motion vector field on a sky image in sky coordinates on March 4, 10:21:30 PST. (a) Original vector field resulting from the cross-correlation method, (b) after removing vectors near the sun and vectors with a correlation of less than 0.8, and (c) after removing outliers. The average cloud velocity was 4.0 m s\(^{-1}\).

Incorrect motion vectors are found primarily in overcast and nearly clear sky conditions (Fig. 9). In overcast skies, the sky contrast is reduced resulting in low CCC. In nearly clear sky conditions, clouds are often thin and evaporate. If the sky contains clouds with sharp cloud boundaries, CCM generally performs well. Due to the general persistence of cloud velocities (usually within 2 m s\(^{-1}\) over several minutes), erroneous decision vectors can be identified since they are associated with large cloud velocity fluctuations (Fig. 9a). If the cloud velocity changes by more than two standard deviations of velocity over the past 5 minutes, the current cloud velocity is replaced by the average velocity over the past 5 minutes.
3.4 Nowcast of Binary Cloudy Conditions

Using the binary irradiance technique described in section 2.3, GHI time series were produced for the ground stations listed in Table 2 Location and status information for weather stations used to validate TSI cloud shadow maps.

### Table 2 Location and status information for weather stations used to validate TSI cloud shadow maps

<table>
<thead>
<tr>
<th>Station Identifier</th>
<th>RIMC</th>
<th>BMSB</th>
<th>EBU2</th>
<th>TIOG</th>
<th>HUBB</th>
<th>MOCC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Location</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>latitude [deg]</td>
<td>32.8852</td>
<td>32.8758</td>
<td>32.8813</td>
<td>32.8790</td>
<td>32.8673</td>
<td>32.8782</td>
</tr>
<tr>
<td>longitude [deg]</td>
<td>-117.2401</td>
<td>-117.2362</td>
<td>-117.2330</td>
<td>-117.2434</td>
<td>-117.2535</td>
<td>-117.2231</td>
</tr>
<tr>
<td>altitude [m]</td>
<td>124</td>
<td>134</td>
<td>104</td>
<td>158</td>
<td>12</td>
<td>104</td>
</tr>
</tbody>
</table>

Figure 9: Time series plot of eastward and northward cloud velocities (a) and cloud fraction (b) on March 10, 2010. The velocities are persistent under partial cloud cover (0800-1030 PST and 1530-1730 PST), but fluctuate largely in overcast and almost clear sky conditions.
distance to TSI [m] | 15 | 1,107 | 789 | 760 | 2,362 | 1,770
--- | --- | --- | --- | --- | --- | ---
Sept 14, 2009 | | | | | |  
Oct 4, 2009 | | | | | |  
Mar 4, 2010 | N/A | N/A | | | |  
Mar 10, 2010 | | | | N/A | N/A |  
| Operational Status |  
| Operational | N/A | Unavailable | Cloud too low, out of range

Figure 10, showing MOCC on October 4, indicates that GHI variability can be estimated using a TSI.

The morning period is dominated by false clouds due to MOCC’s projected sky position being located in the shadowband area of the circumsolar region. Once the projected sky position moved into the outer region the TSI nowcasts the sky conditions between the sun and MOCC accurately, correctly predicting the sky condition 68.3% in clear conditions and 80.4% in cloudy conditions, where clear is defined as a clear sky index \( (kt) \) greater than 0.7.

![Figure 10](image)

**Figure 10**: Time series (a) of MOCC station for October 4, 2009 showing measured GHI (blue) versus the TSI nowcast. The TSI shadow map is separated into circumsolar (red) and outer-region (green). Nowcasts of the TSI after 1100 PST (b),(c) and (d) match the trends in the measured data well.

Nowcast accuracy is expected to decrease with distance from the TSI due to the lower resolution at large sky imager ZAs (in an extreme case on March 4, low clouds over the HUBB station 2.4 km to the southwest were out of sight of the TSI). However, the TSI was able to capture the trends in GHI at the
HUBB station (Figure 11) on the other days. When HUBB was in the outer region, the TSI correctly predicted clear skies 60.0% of the time, and cloudy skies 90.6% of the time.

*Figure 11:* Time series (a) of HUBB station for March 10, 2010 showing measured GHI (blue) versus TSI nowcast, with the latter separated into circumsolar (red) and outer-region (green). Zooming in (b), (c) and (d) shows the ability of the TSI to produce nowcasts for a distant site.

Since binary descriptive statistics are most appropriate to validate the binary TSI cloud estimate the clear-cloudy co-occurrence was tabulated (Table 3) for images taken when the cloud fraction was in the range $0.2 \leq cf \leq 0.8$. Cloud fraction is defined as percent hemispheric cloud coverage weighted by solid angle (as opposed to by pixel). This restriction on cloud fraction is used to ensure that the results are not biased by nearly cloud free or overcast conditions which would result in a near perfect sky condition nowcast. For the four days chosen and for all available stations (Table 2 Location and status information for w), the TSI correctly estimated the condition of the sky 69.7% of the time in the outer region. This ratio was significantly different in measured clear (59.4%) and cloudy conditions (81.5%). It was more often predicted cloudy when it was clear than predicted clear when it was cloudy. A cause of this asymmetry could be that the technique described in section 2.3 assumes that clouds have a negligible vertical extent at a uniform height; when the side of a cloud is measured it increases the total cloud coverage measured by TSI over the actual amount.

**Table 3** Distribution of measured versus TSI nowcast clear (CLR) and cloudy (CLD) conditions for skies
with cloud fraction $0.2 \leq cf \leq 0.8$ (fraction of hemispheric solid angle). In line 1, the first term denotes measured clear sky index, and the second is the TSI estimate, thus CLD/CLR means the ground station measured cloudy conditions ($kt \leq 0.7$) but the TSI estimated clear. CLR$_m$ and CLD$_m$ are the number of clear and cloudy measurements, respectively, in the outer region. The last column shows the overall statistics (not conditioned on cloudy or clear). Data is for all stations.

<table>
<thead>
<tr>
<th></th>
<th>CLR/CLR [%]</th>
<th>CLR/CLD [%]</th>
<th>CLR$_m$</th>
<th>CLD/CLR [%]</th>
<th>CLD/CLD [%]</th>
<th>CLD$_m$</th>
<th>% correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sept 14, 2009</td>
<td>47.1</td>
<td>52.9</td>
<td>594</td>
<td>28.3</td>
<td>71.7</td>
<td>368</td>
<td>56.6</td>
</tr>
<tr>
<td>Oct 4, 2009</td>
<td>76.2</td>
<td>23.8</td>
<td>441</td>
<td>10.3</td>
<td>89.7</td>
<td>419</td>
<td>82.8</td>
</tr>
<tr>
<td>Mar 4, 2010</td>
<td>63.1</td>
<td>36.9</td>
<td>157</td>
<td>33.3</td>
<td>66.7</td>
<td>144</td>
<td>64.8</td>
</tr>
<tr>
<td>Mar 10, 2010</td>
<td>52.7</td>
<td>47.2</td>
<td>110</td>
<td>7.0</td>
<td>93.0</td>
<td>201</td>
<td>78.8</td>
</tr>
<tr>
<td>Total</td>
<td>59.4</td>
<td>40.6</td>
<td>1302</td>
<td>18.5</td>
<td>81.5</td>
<td>1132</td>
<td>69.7</td>
</tr>
</tbody>
</table>

The error metrics of the sky imager GHI estimates (Fig. 10, 11) indicate a negative bias (not shown). The negative bias in the circumsolar region is primarily due to false clouds decreasing the estimated GHI. In the outer region the negative bias is mostly caused by the cloudy bias noted above and cloud enhancement effects whereby the solar disk is not obscured but forward scattering of radiation by clouds located near the solar disk increases the diffuse radiation component and thus the overall GHI. Since the TSI estimate assumes a clear sky index of unity when the sun is not obscured it underestimates the clear sky GHI as well.

3.5 Minutes Ahead Forecast

The performance of a minutes ahead forecast is discussed in this section for four days following the metrics established in section 2.5 and Fig. 5. Figure 12 show the time series of cloud speed, cloud fraction, and errors for a 30 second forecast on September 14. Since both clear and overcast skies always have perfect forecasts due to the use of binary cloud maps, cloud speeds and forecast errors are not assigned for these times.
On September 14, 2009 a transition from overcast to clear sky occurred between 1100 PST and 1200 PST with clouds at an altitude of 610 m moving from north to south at less than 2 m s\(^{-1}\). In the afternoon, thin clouds moved at high speed eastward and were evaporating and deforming rapidly, as evident by the variable cloud fraction. As a result, the matching error in the afternoon is much larger than between 1100-1200 PST. Despite the poor matching error, September 14 has similar cap error as the other days. The matching error is larger than on other days, but the large cloud speed causes large persistence errors which reduces the cap error. The invariance of the cap error over time suggests that the normalization by the persistence image provides a universal performance metric. While matching errors become small when the cloud cover fraction is small or large (the edge-to-area ratio is small in the latter case), the cap error is not sensitive to cloud cover fraction. The cap error is however sensitive to cloud speed.
In the circumsolar region, the large errors are a result of the poor cloud decision. While the SP dramatically increases the accuracy of the cloud decision in the circumsolar region, its high variability occasionally causes cloud decision errors. Additionally, a false cloud detection results when two different clouds bordering the shadowband on both sides are interpolated across. Both issues cause a step change in the cloud decision in a large area between consecutive images which leads to erroneous forecasts.

Mean and standard deviation of outer region error metrics for a 30 second forecast are summarized in Table 4. The mean matching errors are 3 to 7% and the cap errors are 40 to 45%. The cap error of less than unity indicates that translating the cloud map improves the forecast compared with persistence. The mean matching error scales with cloud speed. Otherwise, the comparison of different days does not yield strong correlations between different variables, presumably due to intra-day variability.

<table>
<thead>
<tr>
<th>Date</th>
<th>Mean $e_m$ (%)</th>
<th>Mean $e_{cap}$ (%)</th>
<th>Std $e_m$ (%)</th>
<th>Std $e_{cap}$ (%)</th>
<th>Avg cloud fraction [-]</th>
<th>Avg cloud speed [m s$^{-1}$]</th>
<th>Avg cloud height [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sept 14, 2009</td>
<td>6.26</td>
<td>42.8</td>
<td>3.24</td>
<td>14.2</td>
<td>0.135</td>
<td>7.79</td>
<td>840</td>
</tr>
<tr>
<td>Oct 4, 2009</td>
<td>4.09</td>
<td>40.0</td>
<td>1.48</td>
<td>7.45</td>
<td>0.473</td>
<td>3.75</td>
<td>1090</td>
</tr>
<tr>
<td>Mar 4, 2010</td>
<td>3.16</td>
<td>44.7</td>
<td>1.33</td>
<td>10.5</td>
<td>0.276</td>
<td>2.95</td>
<td>610</td>
</tr>
<tr>
<td>Mar 10, 2010</td>
<td>3.85</td>
<td>41.1</td>
<td>2.40</td>
<td>20.2</td>
<td>0.427</td>
<td>4.49</td>
<td>1220</td>
</tr>
</tbody>
</table>

Table 5 shows the mean total matching error and total cap error for a 30 seconds to 5 minutes ahead forecast. Mean forecast matching errors increase less than linearly from around 6% for 30 second forecasts to 23-30% for 5 minute forecasts. Mean cap errors are less than 80% out to 5 minutes, except on September 14, 2009 when large cloud speeds were observed. Note that while the outer region matching error was considered for 30 second forecasts, here the total (circumsolar and outer region) errors are considered as clouds are frequently advected between the circumsolar region and the outer region (depending upon cloud speed and direction).
Table 5 Mean total matching error and total cap error for a 30 second to 5 minute ahead forecast. The errors are averaged only during conditions that were not clear or overcast and when the entire scene had not been advected out of the field-of-view. Since errors during overcast and clear conditions are zero, the errors in the table are biased high.

<table>
<thead>
<tr>
<th>Time until advection out of scene [min]</th>
<th>30 sec</th>
<th>1 min</th>
<th>2 min</th>
<th>3 min</th>
<th>4 min</th>
<th>5 min</th>
<th>30 sec</th>
<th>1 min</th>
<th>2 min</th>
<th>3 min</th>
<th>4 min</th>
<th>5 min</th>
<th>30 sec</th>
<th>1 min</th>
<th>2 min</th>
<th>3 min</th>
<th>4 min</th>
<th>5 min</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sept 14, 2009</td>
<td>8.67</td>
<td>13.5</td>
<td>20.1</td>
<td>23.8</td>
<td>27.2</td>
<td>28.0</td>
<td>45.0</td>
<td>51.5</td>
<td>63.9</td>
<td>70.0</td>
<td>76.5</td>
<td>123</td>
<td>45.0</td>
<td>51.5</td>
<td>63.9</td>
<td>70.0</td>
<td>76.5</td>
<td>123</td>
</tr>
<tr>
<td>Oct 4, 2009</td>
<td>6.01</td>
<td>9.84</td>
<td>15.9</td>
<td>20.5</td>
<td>24.2</td>
<td>27.1</td>
<td>47.2</td>
<td>49.8</td>
<td>55.6</td>
<td>61.5</td>
<td>66.6</td>
<td>70.3</td>
<td>47.2</td>
<td>49.8</td>
<td>55.6</td>
<td>61.5</td>
<td>66.6</td>
<td>70.3</td>
</tr>
<tr>
<td>Mar 4, 2010</td>
<td>5.36</td>
<td>8.68</td>
<td>13.6</td>
<td>17.3</td>
<td>20.4</td>
<td>23.4</td>
<td>54.6</td>
<td>55.3</td>
<td>59.3</td>
<td>63.4</td>
<td>67.7</td>
<td>71.8</td>
<td>54.6</td>
<td>55.3</td>
<td>59.3</td>
<td>63.4</td>
<td>67.7</td>
<td>71.8</td>
</tr>
<tr>
<td>Mar 10, 2010</td>
<td>6.41</td>
<td>10.7</td>
<td>17.8</td>
<td>23.1</td>
<td>27.2</td>
<td>30.0</td>
<td>48.8</td>
<td>53.9</td>
<td>62.3</td>
<td>68.8</td>
<td>75.1</td>
<td>78.0</td>
<td>48.8</td>
<td>53.9</td>
<td>62.3</td>
<td>68.8</td>
<td>75.1</td>
<td>78.0</td>
</tr>
</tbody>
</table>

The errors can decrease in time (Fig. 13) if for example, a cloud that was forecast incorrectly is advected out of the scene while another cloud that was forecast correctly remains. The time until the initially measured cloud field is advected out of the scene determines maximum possible total sky imager forecast horizons for the area surrounding the imager. The footprint of the cloud shadows on the ground varies with sun angle and cloud height and the forecast horizon is a function of cloud height and cloud speed. For the low cloud heights and cloud speeds observed, the maximum forecast horizon was about 5 to 25 minutes. The following sources of forecast error are observed:

(i) In general, cloud deformation contributes most to forecast error. In our low latitude coastal areas with mostly mesoscale weather patterns, changing cloud patterns are more common than in more synoptically driven areas. Large solar irradiance can lead to short convection and cloud evaporation time scales. The rapid condensation and evaporation and associated deformation of clouds introduce complexity to deterministic cloud forecasting.

(ii) As discussed in section 3.4, perspective and occlusion effects on clouds are another source of error resulting from the two-dimensionality of the sky image. The sides of approaching clouds are interpreted as cloud bottoms. Consequently, as the clouds approach the sky imager, the perspective changes until only the cloud bottom is visible leading to a perceived decrease in cloud sizes and change in cloud shape. The perspective error may also create the illusion that there is one cloud while there are two nearby but separate cloud. Generally, the perspective error is inversely proportional to CBH; it will be largest for tall thunderstorm clouds and small for thin stratus clouds. A network of sky imagers with different perspectives could reduce this forecast error.
(iii) Heterogeneity of cloud velocity. Usually, the cloud velocity computed from cross-correlation is representative of the velocity of all clouds in the image. In rare cases, however, cumulus clouds to the east of the sky imager were moving at a different velocity than stratiform clouds to the west. Since our forecast assumes a homogeneous cloud velocity (roughly the size weighted average velocity of the two cloud layers), large forecast errors result.

Figure 13: Total matching error for different forecast horizons on October 4, 2009. (a) 1 minute, (b) 2 minutes, (c) 3 minutes, (d) 4 minutes, (e) 5 minutes. Missing data occur because of missing images, clear or overcast skies, and when the entire scene had been advected out of the field-of-view.

4. Conclusions and Future Work

A technique for intra-hour, sub-kilometer cloud shadow nowcasting and forecasting using a ground-based sky imager was presented for selected days at the UC San Diego solar energy testbed. The use of sky imagery to assess the solar resource for solar energy applications shows much potential for augmenting the spatial and temporal resolution provided by satellite and numerical forecasting methods. Only short deterministic forecast horizons are feasible using a single TSI at our site due to low clouds
and large cloud variability at the fine spatial scale studied. Capturing these features deterministically is nearly impossible with satellites or numerical weather prediction.

The method’s accuracy is also limited by the shortcomings of the sky imager used. Part of the data within each image is lost due to obscuration by the camera arm and shadowband. Low dynamic range causes images to saturate easily near the sun, even with automatic gain control (AGC) of the CCD camera. AGC causes its own problems because the output signal strength fluctuates between images, leading to shifting intensity histograms and making spectral composition based cloud decisions less reliable. Inaccuracies in the forecast sky conditions are also due to cloud deformation, evaporation, and condensation, as well as uncertainty in cloud base height. The effect of these hardware limitations is evident by the increased errors in the circumsolar region (Fig. 12).

Despite these shortcomings cloud shadows in the outer regions are correctly nowcast 70% of the time. Cloud-advection-versus-persistence (cap) errors were found to be a universal metric to compare forecast performance (Table 4, 5); for 30 second forecast, a 50-60% reduction in forecast error compared to persistence was found. The error would be further reduced if it were spatially averaged over a large solar array rather than computed on a pixel-by-pixel basis. Matching forecast errors increase with cloud speed and forecast horizon.

Future work will include several elements to improve forecast accuracy. Replacing the TSI by a sky imager that does not use a shadowband (Cazorla et al. 2008, 2009, Heinle et al. 2010) will increase the usable fraction of the image. A CCD camera with large well-depth will be deployed to minimize saturation in direct sunlight. A fisheye lens with an equisolid angle projection will improve the image resolution at low elevation angles to increase forecast horizon and accuracy at long forecast horizons. On-site cloud height will be retrieved using a ceilometer collocated with the sky imager to improve the geometry-based irradiance nowcasting. The CSL will be generated using data from more than one day and will incorporate a third dimension by using SZA, which will help account for airmass and other diurnal effects. CSL thresholding based on daily aerosol optical thickness will improve the cloud decision since the spectral composition of clear sky is affected by scattering due to aerosols. Additional
Sky imagers will be installed at the UC San Diego energy testbed to increase the coverage area and forecast horizon. This will allow the vertical extent of clouds to be determined with a 3D-reconstruction of clouds using stereography. Sky imagers are also currently being installed at inland sites to study cloud conditions likely to be experience by solar power plants. The methods developed here for sky imagery will be combined with satellite- and numerically-based forecasts to yield a comprehensive and more accurate forecast product with different forecast horizons, with the goal of meeting the needs of utility, grid, and power plant operators.

Acknowledgements

We acknowledge (i) funding from DOE High Solar PV Penetration grant 10DE-EE002055, (ii) assistance with installation and maintenance of the sky imaging facility by Tex Xelowski, David Purtell, Mildred Mcmilion from RIMAC Sports Facilities, (iii) support from John Ratcliffe and Robert Serocki.

References


