Accounting for Circumsolar and Horizon Cloud Determination Errors in Sky Image Inferral of Sky Cover.

C. N. Long, Pacific Northwest National Laboratory

1) Introduction

In observing the cloudless sky, one can often notice that the area near the sun is whiter and brighter than the rest of the hemisphere. Additionally, even a slight haze will make a large angular area of the horizon whiter and brighter when the sun is low on the horizon. The human eye has an amazing ability to handle a range of light intensity spanning orders of magnitude. But one of the persistent problems in using sky images to infer fractional sky cover is the intensity range limitations of the camera detector. It is desirable to have bright enough images to be able to detect thin clouds, yet this often means the part of the image near the sun and near the horizon for low sun is "white" in the images, not because that is the color perceived by the human eye, but because the CCD elements are recording near their maximum signal. Thus, these pixels are often interpreted as "cloudy" in the sky imager retrievals when a human observer would label them as "cloudless". The Total Sky Imager (TSI) software allows user configurable settings to keep separate additional accounting of clear/thin/opaque determinations for these two "problem areas“ (see Fig. 1). Additional analysis such as that described in Pfeister et al. (2003, JAM), which was originally developed for the first prototype Hemispheric Sky Imager (Long and DeLuisi, 1998, AMS Conf.), can be performed to further help determine if these "problem" areas should indeed be counted as cloudy or not. This analysis technique is refined and applied to TSI data.

2) Methodology Overview

One of the characteristics of the “sun circle” and “horizon area” cloud misidentification problem is that the erroneously retrieved fractional cloud amount in these areas varies slowly with time. On the other hand, clouds moving across the sky also move through these areas, causing an increase in variability of the fractional cloud amount in these areas. Thus, low variability through time can be used as an identifier of when the error might be occurring.

Another identifying characteristic is that there is generally a larger fractional cloud amount in the problem areas, yet a smaller fractional cloud amount in the remaining parts of the image outside these areas, again with little variability through time in this remaining area.

Finally, it is essentially forward scattering of the direct sun that increases the sky whiteness and brightness causing the problem. Thus the “sun circle” can exhibit circumsolar brightness at all solar elevations, and is a more persistent problem compared to the “horizon area” where the problem occurs most often for low sun elevations. Thus, the typical characteristics noted above can be used to test these problem areas and determine whether they should be considered cloudy or not.

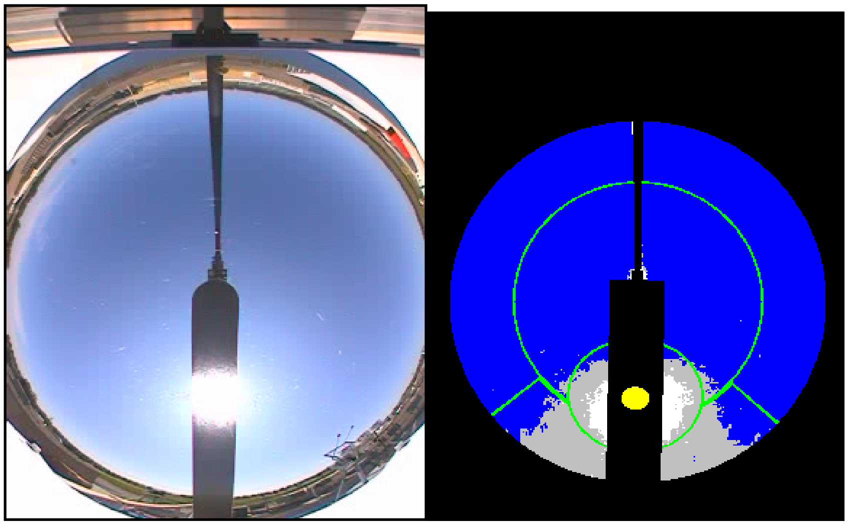

Figure 1: A sample HSI image (left) and corresponding cloud decision image (right) taken at 1300 Local time, 20040904 at PNNL. The yellow dot on the sun blocking mask in the cloud decision image is the location of the sun in the image. The area around the sun outlined in green is the “sun circle” area that is problematic for cloud detection as described above, and separate additional pixel counts are kept for this area. Similarly, the area outlined in green centered on the sun blocking mask and extending in an arc to either side is the problematic “horizon area”. As shown in this example, the two areas contain pixels erroneously determined as “cloud” while the sky image shows clear-sky. The large circle centered on zenith is the “zenith circle” discussed in Section 5.

3) Algorithm Outline

Define:
Ssc = Fractional sky cover in Sun circle area
Hsc = Fractional sky cover in horizon area
Rsc = Remainder sky cover, all area outside sun circle and horizon areas
Sadj = Sun circle adjustment factor = (1 – Rcf)
Sfact = maximum limit for Sadj
Sdev = standard deviation of Ssc over 21 data centered on data of interest
Hdev = standard deviation of Hsc over 21 data centered on data of interest
Rdev = standard deviation of Rsc over 21 data centered on data of interest
AdvLim = limit of allowable Hdev and Sdev

RdvLim = limit of allowable Rdev
SscLim = minimum limit for Ssc
HscLim = minimum limit for Hsc
RscLim = limit of allowable Rsc
SCpix = total number of cloudy pixels in Sun circle

First Guess for Sun Circle:
If Sadj > Sfact, then Sadj = Sfact
Subtract (Sadj * SCpix) from total sky cover

Sun Circle:
IF
(Sdev < AdvLim)
AND (Ssc > SscLim)
AND (Rsc < RscLim)
AND (Rdev < RdvLim)
THEN do not count Sun circle cloud pixels

Horizon Area:
IF
(Hdev < AdvLim)
AND (Hsc > HscLim)
AND (Rsc < RscLim)
AND (Rdev < RdvLim)
THEN do not count horizon area cloud pixels

Smoothing:
Calculate a running 11 data average of the amount of correction applied centered on the data of interest, and apply average correction to the data of interest.

4) Example case

To illustrate the method, Figure 2 shows a series of sky images and corresponding cloud decision images for 20040904 at PNNL, as in Figure 1. The morning was clear (see Fig. 2), with cloudiness moving in at about 1320 and lasting through about 1520 when skies cleared again. More cloudiness then moved slowly in again at around 1700, slowly moving off through about 1840. As the cloud decision images of Figs. 1 and 2 show, this day exhibited significant haze producing the erroneous identification problem. Figure 3 shows the total sky cover for this daylight period, including the original retrieval, the “first guess” Sun circle adjustment, and the final adjusted retrieval. The “first guess” is intended to account for the probability of some error near the sun due to persistent forward scattering for times when the other tests outlined in Section 3 do not subtract the sun circle and horizon areas cloud pixels.

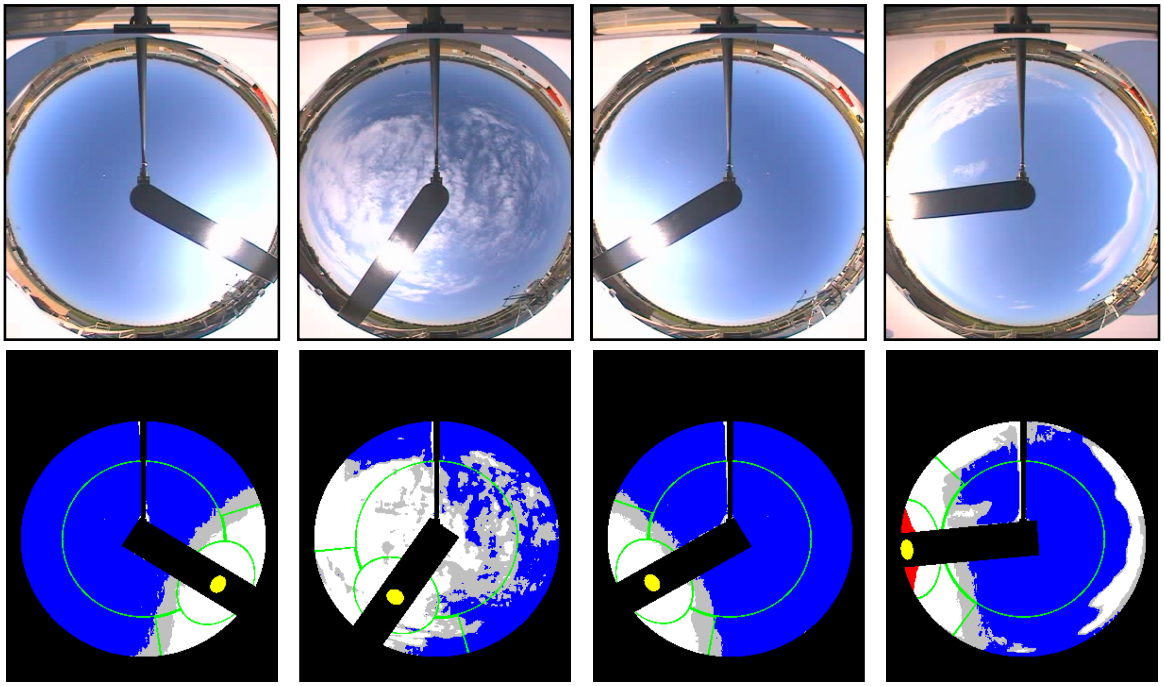

Figure 2: Series of sky images from 20040904 at PNNL. Local times of images are (upper left to right) 1000, 1430, 1600, and 1800. Lower images are the corresponding cloud decision images.

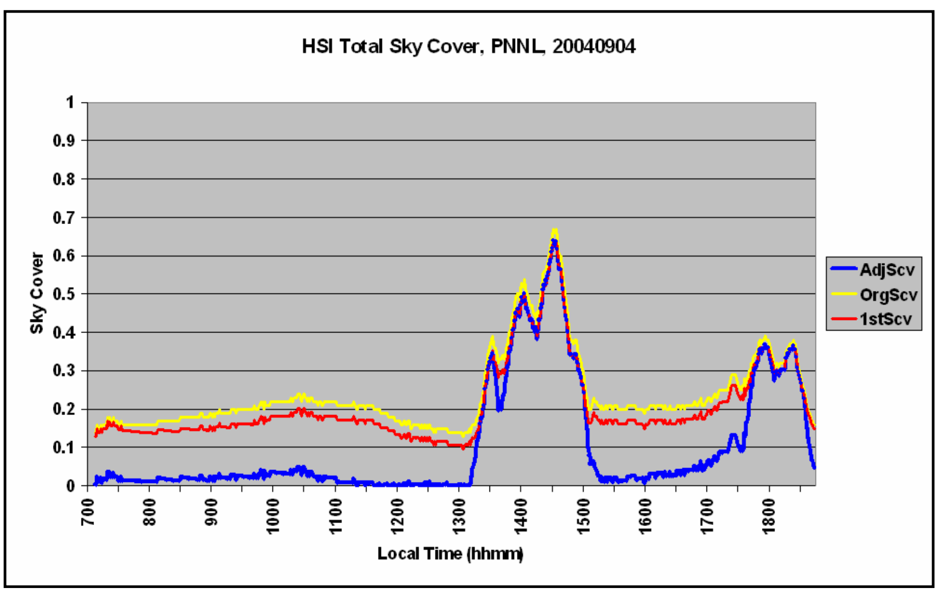

Figure 3: Total sky cover retrieval for 20040904 at PNNL corresponding to images in Figs. 1 and 2. Yellow line is the original retrieval, red is the retrieval including the “first guess” adjustment of the sun circle area, blue line is the final result including all adjustments and smoothing as outlined in Section 3.

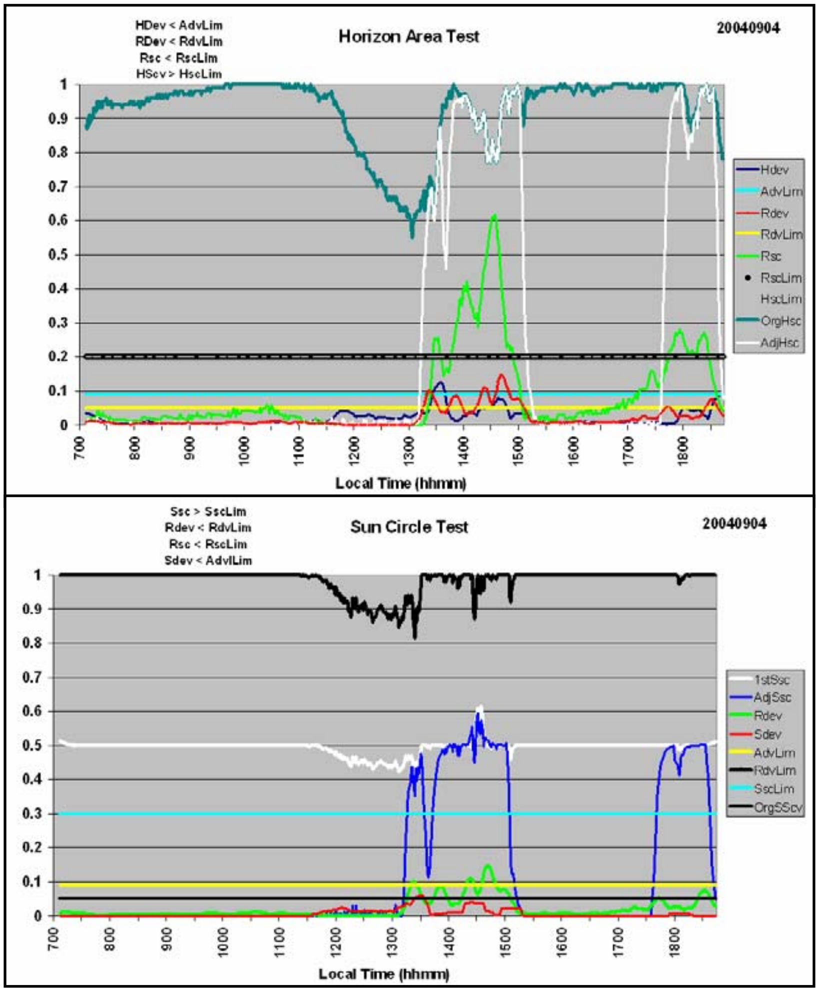

Figure 4: Horizon area (top) and Sun circle (bottom) testing results for 20040904 at PNNL corresponding to Fig. 3 adjustments. See Section 4 for explanation.

Figure 4 illustrates the testing involved in the method as outlined in Section 3. The top plot is for the horizon area testing. Originally, the horizon area cloud amount is retrieved as nearly 100% for most of the morning and afternoon, with some decrease centered on about 1300 (teal line). All of the horizon area cloudiness is discounted except for the periods from about 1315 through 1510, and 1740 through 1840. At different times during these latter cloudy periods, the plot shows that one or more of the Rsc (green line) > RscLim (black horizontal), Rdev (red) > RdvLim (yellow), or Hdev (dk. Blue) > AdvLim (lt. blue) tests were failed, thus the horizon area cloud pixels were not discounted. The result is the white line, the horizon area cloudiness that was kept as part of the total sky cover amount.

Similarly, the bottom plot in Figure 4 shows the results for the Sun circle testing. Originally, the Sun circle cloudiness (black line) was 100% almost all day. The “first guess” adjustment (white) in general decreases the Sun circle area amount by about half. For the same periods as noted above for the horizon area, again the Rsc (top plot) and Rdev (green line) tests were failed, causing the Sun Circle cloudiness to revert to the “first guess” amount. Thus the amount of Sun circle cloudiness included in the total sky cover is shown by the blue line, where all of the Sun circle cloudiness is discounted except for the two previously identified cloudy periods, where about half has been discounted. Note in Figure 3 that the actual amount of sky area for this Sun circle, which does not include the sun blocking strip mask, is minor. Yet as can be seen in the example sky and cloud decision images there is often some overestimation of cloud amount in the Sun circle area, thus the reasoning behind the “first guess” adjustment.

5) Analysis results

Figure 5 shows relative frequency histograms of various instruments and time periods as noted in the Figure caption. In each case, the original retrievals (red) show a bias away from the “clear” bin (on the left) toward higher values. This result is at odds with expectations, where it is common that the frequency distribution includes about 1/3 clear-sky, 1/3 overcast, and the remaining 1/3 distributed in between. This expected distribution is indeed the case when the adjustments detailed here are applied to the retrievals (blue) for each case. In the top two plots, the third distribution (striped) is from the available 100-degree field-of-view “zenith circle” retrievals. This zenith area is far less susceptible to the misidentification problems we are addressing, since the entire horizon area is not included, and the sun circle problem is generally less for higher sun elevations. As is seen, there is much better agreement with the adjusted distributions than with the original. Similarly, the third distribution in the bottom plot, produced by the SW Flux Analysis VAP, agree better with the adjusted values than the original. All these results suggest that the adjustment methodology significantly improves the sky imager retrievals as intended.

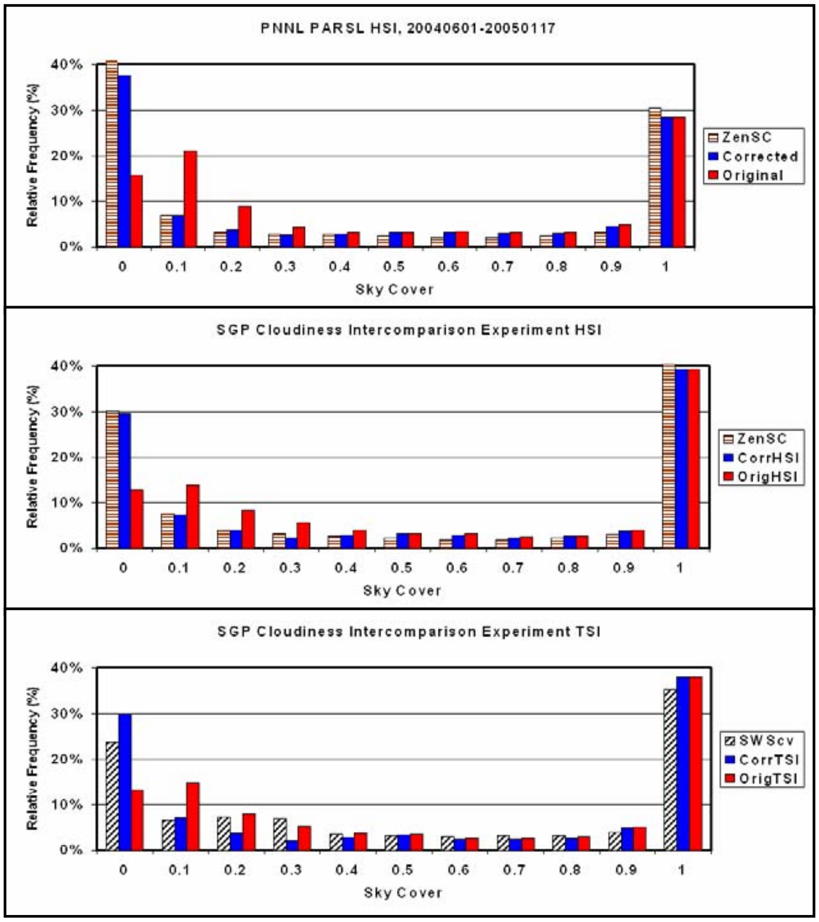

Figure 5: Sky cover frequency histograms for over 7 months of data at PNNL (top), and for the Hemispheric Sky Imager (middle) and Total Sky Imager (bottom) deployed during the 3 months of the ARM Cloudiness Intercomparison Experiment at the SGP. See Section 5 for details.

REFERENCES:

Long, C. N. and J. J. DeLuisi, (1998): Development of an Automated Hemispheric Sky Imager for Cloud Fraction Retrievals, Proc. 10th Symp. On Meteorological Observations and Instrumentation, Jan. 11-16, 1998, Phoenix, Arizona, Amer. Meteor. Soc., 171-174.

Pfister, G., R. L. McKenzie, J. B. Liley, A. Thomas, B. W. Forgan and C. N. Long (2003): Cloud Coverage Based on All-Sky Imaging and Its Impact on Surface Solar Irradiances, JAM, 42, 1421-1434.