Title
Algorithm for detecting clear sky images

Permalink
https://escholarship.org/uc/item/2xt425wn

Author
Pawar, Prathamesh Vijay

Publication Date
2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Algorithm for detecting clear sky images

A Thesis submitted in partial satisfaction of the requirements
for the degree Master of Science

in

Engineering Sciences (Mechanical Engineering)

by

Prathamesh Vijay Pawar

Committee in charge:

Professor Jan Kleissl, Chair
Professor Renkun Chen
Professor Lynn Russell

2016

The Thesis of Prathamesh Vijay Pawar is approved and acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

Chair

University of California, San Diego

2016

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to express my deep gratitude to my adviser, Professor Jan Kleissl, for his constant support throughout the research work. Also, for always replying to my queries via emails as quickly as possible. I appreciate his patient guidance, encouragement and useful critiques for this thesis. It is a great honor to be a part of his innovative team which has done phenomenal work in the area of solar energy forecasting.

I would also like to offer my special thanks to Keenan Murray, whose valuable advice always helped me come up with new ideas for the research. He has always been a person I would turn to for any programming help.

I wish to thank Ben Kurtz and Felipe Mejia for their valuable insights and for always having answers to all my questions. I would also like to thank Zack Pecenak for setting me up with the required software and user interface required for the project on Windows. And thanks to Nishank Sheth for always being there as a good friend throughout the journey.

Finally, I would like to thank my parents and sister for their encouragement and support throughout my study.

Chapters 1, 3, 4 and 5, in part, are currently being prepared for submission for publication of the material. Pawar, Prathamesh; Murray, Keenan; Kleissl, Jan. The thesis author was the primary investigator and author for this material.

ABSTRACT OF THE THESIS

Algorithm for detecting clear sky images

by

Prathamesh Vijay Pawar

Master of Science in Engineering Science (Mechanical Engineering)

University of California, San Diego, 2016

Professor Jan Kleissl, Chair

Many solar forecast algorithms based on ground based sky imagery apply the red-blue ratio (RBR) method to classify image pixels as clear or cloudy, by comparing the current image with the corresponding image from a clear sky library (CSL). The CSL needs to be updated regularly due to change in clear sky conditions over time caused by aerosols and imager dome properties. This clear sky library is typically created by visually scrutinizing daily sky videos and selecting appropriate clear sky periods. This practice takes a significant amount of time and manual intervention can result in human errors. To avoid this, an automated CSL algorithm (ACSL) was developed which filters each image for clear sky features such as maximum green pixel brightness, average RBR, and red channel difference. The relative root mean square error (RMSE) between the

image RBR of the manually created CSL and the one created using ACSL for November 2013 at UC San Diego was observed to be less than 5% over the range of solar zenith angles and it was found to be more representative of clear conditions than its manual counterpart.

1. INTRODUCTION

Ground based sky imagers have been developed by many groups [1, 2, 3, 4] for short term solar forecasts. Various solar forecast algorithms based on ground based sky imagery use a clear sky library(CSL) for cloud detection, which is generally created by visual inspection of sky images. After visually scrutinizing the images for clear sky periods, the periods chosen as clear sky are passed on to the CSL which stores the RBR for each image pixel as a function of sun pixel angle (SPA), image zenith angle (IZA) and solar zenith angle(SZA) [5, 6]. IZA is the angle between any pixel projected on the dome and the vertical line drawn from the center of the imager, whereas SPA represents the angular distance from the sun [7]. That is, the RBR difference between current image and CSL image, helps to classify the clouds based on the threshold for thin and thick clouds. But these RBR values change with time due to natural events or instrument properties such as change in aerosol concentration or size distribution in the sky, scratches or soiling on the imager dome, etc. [8]. Studies have found that natural and anthropogenic aerosols scatter and absorb solar radiation, thus altering the clear sky radiation properties [9, 10]. For accurate cloud detection the clear sky period reference should therefore be as recent as possible, which is why the CSL needs to be updated regularly. This process can consume significant man-hours and introduce human errors.

For example, daily sky image videos may be played back in 1 minute when 20 images per second are shown and the image capture resolution is 30 sec. This requires

1

great attention of the viewer, but often small clouds around the horizon or near the sun are missed. In particular, thin   clouds are often missed by the viewer, which results in contamination of the CSL with clouds and ultimately can lead to classification of thin clouds as clear sky while running the daily forecast. To avoid this, an automated clear sky library(ACSL)was developed (Section 3). ACSL images for Nov 2013 are compared against manual selection (Section 4). Conclusions are provided in Section 5.

Chapter 1, in part, is currently being prepared for submission for publication of the material. Pawar, Prathamesh; Murray, Keenan; Kleissl, Jan. The thesis author was the primary investigator and author for this material.

2. THEORY

Solar energy is one of the major sources of renewable energy available to mankind in abundance. Earth receives around 1366 Watts of direct solar radiation per square meter at the top of its atmosphere. Clearly, solar energy has a huge potential in meeting all the energy demands of human beings, but there are certain limitations. Clouds can be a major cause of variability and ramp events in energy production. So it is very essential, for solar plant operators to have forecasts to adjust their energy production process so that continuous energy is supplied to end users. Various forecasting techniques are in use, like numerical weather prediction (NWP), satellite imagery, etc. But these techniques aren't good enough for short term forecasts and also their resolution is limited [7]. For short term forecasts having better resolutions, various ground based sky imagers have been developed. The sky imager used at UCSD, applies RBR method to classify sky image pixels as thin, thick or clear. The sky imager consists of a fish eye lens facing sky having a field of view of 180°. The sky information is stored in each pixel of the image calibrated according to SZA, IZA and SPA.

2.1 Clear sky RBR

Due to molecular scattering of shorter wavelengths (Rayleigh scattering) blue light is scattered more than other wavelengths which gives the clear sky its blue color. On the other hand, clouds scatter visible radiation uniformly (Mie scattering) and thus cloudy areas in the sky appear white. So, the pixels covered with clear sky will measure

more blue signal than red, and pixels covered with clouds will measure both signals equally or with slight difference depending on the thickness of the cloud. Thus, RBR of clear sky is much lower than that of cloudy sky. But clear sky RBR is not uniform due to presence of aerosols which have weaker wavelength dependency for scattering light as compared to molecules. Thus, they tend to increase the RBR of clear sky. Circumsolar region also tends to have higher RBR because of the forward scattering of direct sunlight by aerosols.

Figure 2.1.1 shows the average RBR values for clear sky and overcast period from the month of November, 2013.
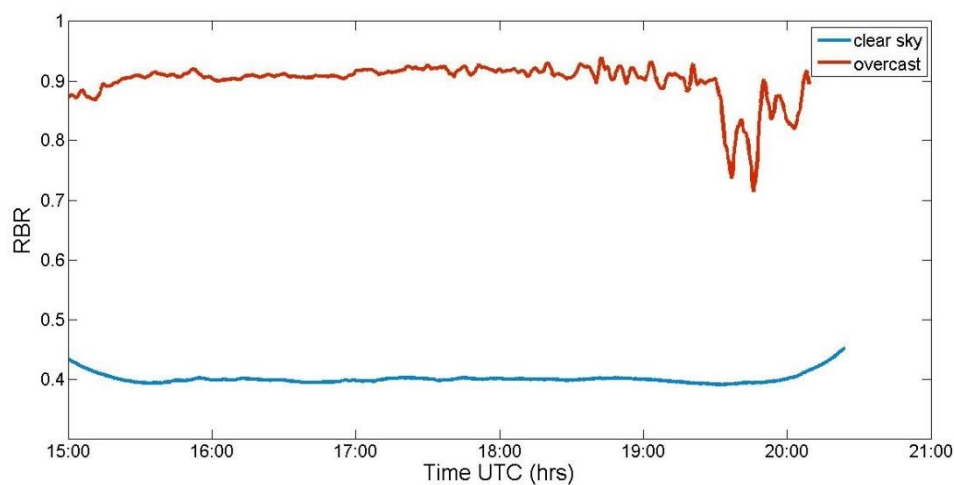

Figure 2.1.1: RBR comparison of clear sky and overcast periods from Nov, 2013

Figure 2.1.2 shows the average red-green-blue (RGB) pixel intensity values for 1st Nov, 2013 for clear sky period.
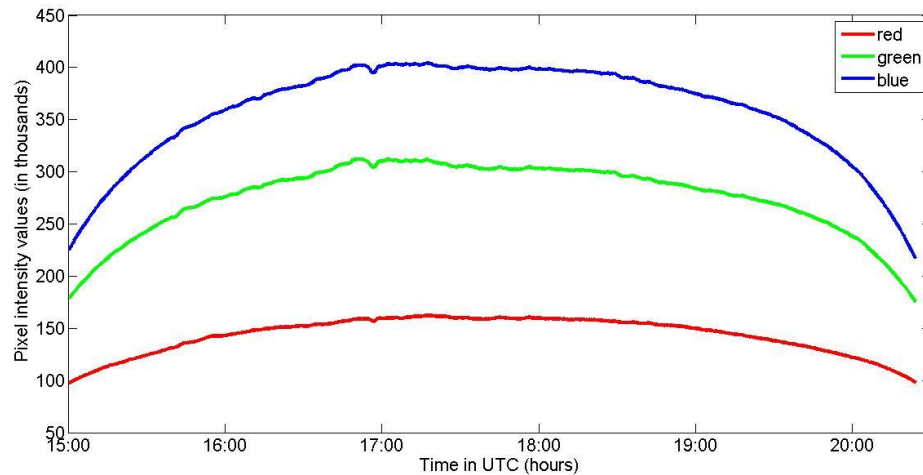
Figure 2.1.2: RGB pixel intensity values for clear sky period for 1st Nov, 2013. As explained, it can be seen that more blue light is captured for clear sky

2.2 Current method and its limitations

CSL RBR for current image, which is being processed for cloud decision, is obtained from the CSL through lookup tables, searching for similar sun position in the sky image. Then, CSL RBR is subtracted from current sky image RBR to obtain a difference matrix. So, clear sky areas will have same difference values, and clouds can be classified as thin or thick based on certain threshold difference values. But, aerosols cause a change in clear sky RBR which can result in faulty cloud decision if the CSL being used is not very recent. Aerosol optical depth (AOD) was found to affect the clear sky RBR directly [8]. So, the CSL needs to be updated to reflect recent changes in clear sky properties. At present, this is done by manually scrutinizing daily sky videos and selecting clear sky periods. ACSL helps to avoid manual intervention and errors related to that, by screening through the clear sky images and applying clear checks to store the clear sky period automatically in the CSL.
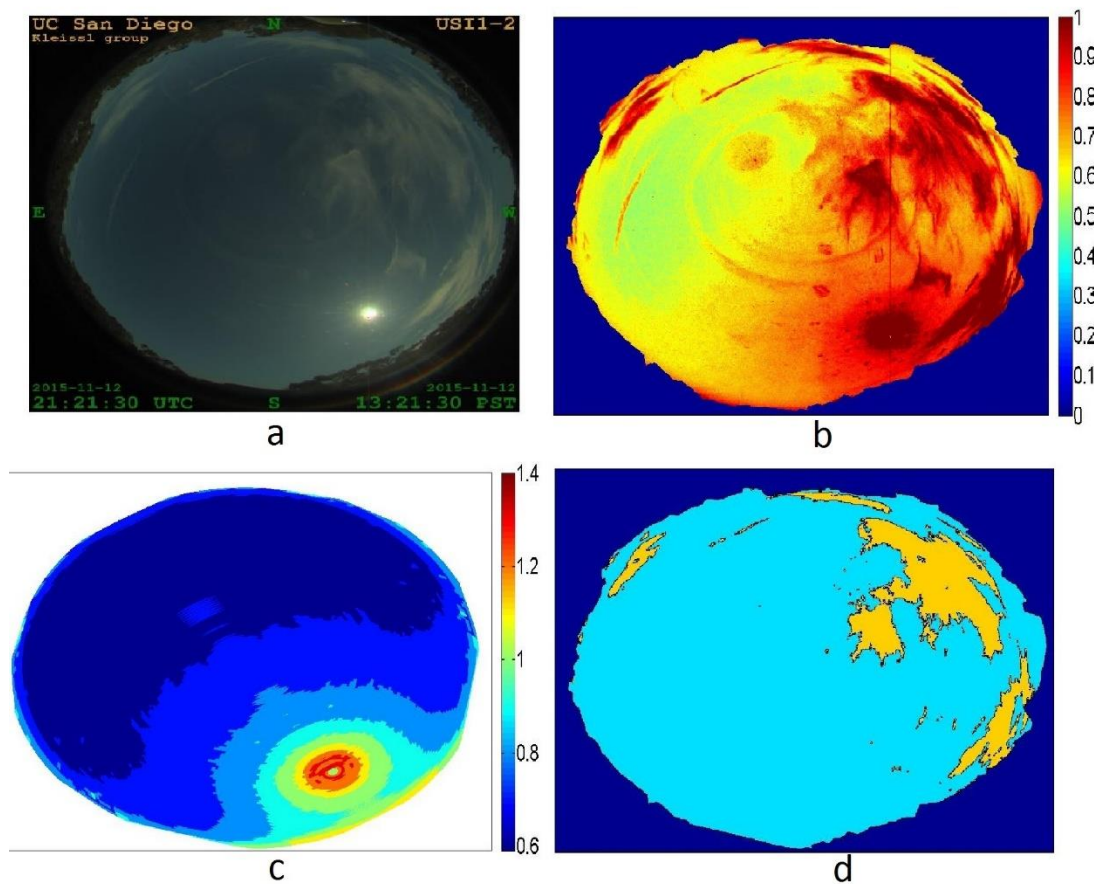
Figure 2.2.1: Forecast run for Nov 12 2015 showing a) Raw image of the sky, b) RBR of the image, c) CSL RBR with similar sun position, d) Cloud decision

2.3 ACSL

The main logic behind creating ACSL is that cloud advection causes successive images with clouds to show differences in their RGB pixel values. For example, if a pixel denoted as (p, q) in the sky image has clouds, then in some time due to cloud advection the pixel intensity value of (p, q) will change. ACSL uses this principle to sift out clear sky images from the total stack of images given to it.

3. METHODOLOGY

Sky images are subjected to several tests to check for clouds; if all conditions are satisfied a measurement period is classified as clear sky period and stored in the CSL.

3.1 Single Image Tests

The circumsolar region (SPA<15°) may confuse the algorithm because pixels in this area can be saturated due to forward scattering by aerosols and clouds [11]. Before removing the circumsolar region, ACSL tests for clouds in the circumsolar region by checking the maximum green pixel value in the image. If the maximum green value is less than a threshold (<15000 where the full range is 65519), the sun is assumed to be obscured by a cloud and the image is classified as cloudy. The green pixel value is used rather than RBR because the RBR luminous distribution is noisier and the brightness difference between clear sky and clouds is less for RBR as compared to individual RGB channels. Also, due to the use of a color filter array(CFA) in digital cameras green pixel intensity is twice that of red or blue pixels [12].

If the circumsolar region is clear of clouds, it is removed from the image, and cloudy conditions are tested by checking average RBR of the remaining image pixels. If RBR is greater than 0.6 [13], the image is defined as cloudy and is not included in the CSL.

3.2 Multiple Image Test

Since clouds affect the red channel the most [7] a Red Channel Difference (RCD) check is applied. ACSL calculates RCD between consecutive images as,

$$RCD_i(p,q) = RC_i(p,q) - RC_{i-1}(p,q); \ i = [2: M], (p, q) = pixel \ position,$$

where, $p = 1: P, q = 1: Q$ are the pixel indices, i is the current time step, $M = \frac{N}{I} + 1$, $N = 5$ min and $I = 0.5$ min in our case. Thus, the images are divided into sets of 5 minutes with 30s interval between each image.

The RCD matrix contains P x Q = 1683 pixels x 1683 pixels = 2.8 x 10^6 pixels and M such matrices for one set.

The horizon is represented by the area with large IZAs (>80) [5, 7]. The horizon requires special treatment, because if clouds near the horizon move tangential to the image circle then due to the use of fish eye lens and its field of view, for a certain amount of time, there will be a negligible difference in RCD of two consecutive images because our sky image area is circular and this motion will be misinterpreted as stationary because these clouds aren't moving into or out of the image circle but are moving along the circumference. To avoid this situation, ACSL checks the horizon area separately by selecting a ring of pixels along the horizon having IZA>80°. In this area the change in the fraction of pixels between two consecutive images having a difference above a particular RCD threshold (here 50) is counted. If this fraction is greater than the decided threshold (0.04% out of 360,000 pixels in the horizon ring), then that set is discarded. This test also helps to capture the fact that, if given a cloud with certain velocity, it will

affect more pixels in a particular time span 't' if it is near the center of the image as compared to when it is near the horizon. A mathematical explanation of the above scenario is given below.
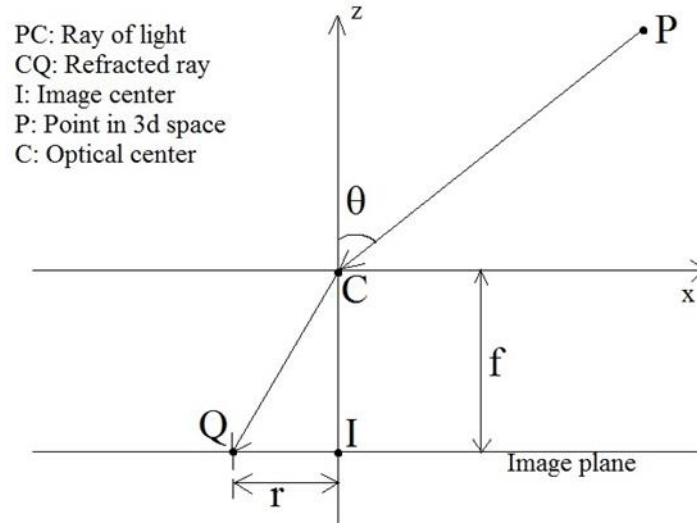

Figure 3.2.1 Schematic representation of an image for a fish eye lens

Figure 3.2.1 shows the lens projection diagram for an object point P in space. For an equisolid angle projection, which is used for UCSD sky imager (USI), the radial distance of the image point from the image center *I* is given by,

$$r = 2fsin\left(\frac{\theta}{2}\right)$$

where, θ is the IZA in radians and f is the focal length. Suppose P represents a point cloud at a base height *h* from mean sea level. And *x* is the horizontal distance measured from the optical center.

Using trigonometry, it can be easily shown that,

$$\theta = \arctan\left(\frac{x}{h}\right)$$

Assuming the cloud speed constant at *v*,

$$v = \frac{dx}{dt}$$

and the velocity of image point corresponding to the point cloud P can then be written as,

$$\frac{\partial r}{\partial t} = \frac{\partial r}{\partial \theta} \frac{\partial \theta}{\partial x} \frac{dx}{dt}$$

as f is constant. Simplification of the above equation gives the following result,

$$\frac{\partial r}{\partial t} = K. cos^2\theta. \cos\left(\frac{\theta}{2}\right)$$
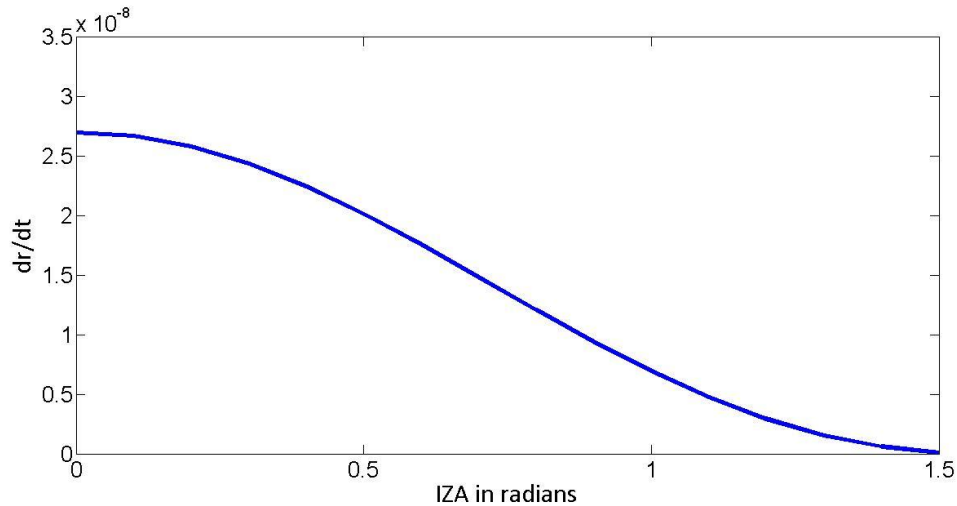
where, $K = v. f / h^2$

Figure 3.2.2 Graph of dr/dt vs IZA for v=6m/s, f=4.5mm and h=1000m

Thus, it can be seen that the image point velocity, which will be proportional to the pixel velocity (pixels/s) in the image area, is proportional to a function of the cosine of the zenith angle θ, for a given cloud height h. Thus, for a given cloud at same height *h,* and moving with velocity *v,* the number of pixels affected will be less if the cloud was near the horizon (higher IZA) as compared to if it was near the center of the image (lower IZAs).

After the horizon test, for the rest of the image, if P ($RCD_i > 50$ counts) > 0.06% then ACSL declares this 5-minute time period as cloudy. The value 50 (corresponding to 0.08% of the full range of 65519 counts) and the fraction 0.06 were chosen empirically. In an image without optical defects the fraction of pixels that differs by 50 counts should be zero. In real images, small areas can have difference above 50 due to optical effects of the imager dome such as sun glint. Note that the sky area represents only 1.94×10^6 pixels out of 2.8×10^6 total pixels, but the remaining non-sky pixels were included in the count for simplicity. Since these black pixels will not show a RCD above the threshold, the fractions in this paper have be to be adjusted for use with other camera systems.

If all clear sky checks are positive, the images from that time period are passed on to creating the CSL. To understand how RCD is affected by varying sky conditions refer to Figure 3.2.3. Figure 3.2.3.a was selected as clear sky whereas the other two images were discarded because they failed the clear sky checks.
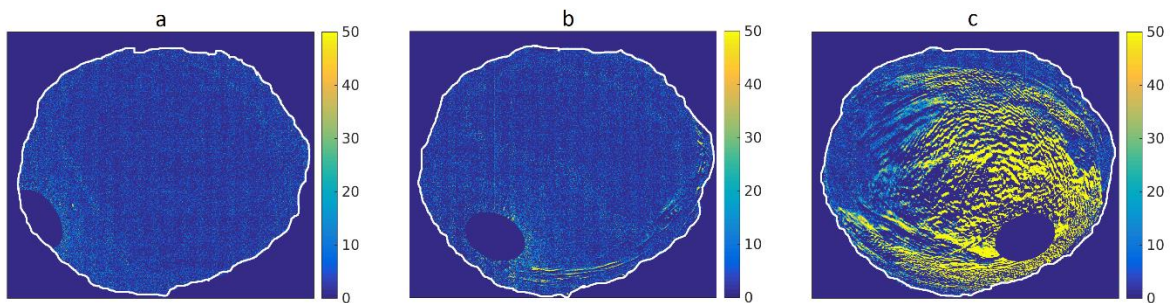

Figure 3.2.3: RCD between images from Nov 2, 2013 at a) 15:05:00 and 15:05:30 UTC. This image was classified as clear sky, b) 17:10:00 and 17:10:30 UTC. This image was not selected as clear sky as the fraction of pixels with RCD > 50 is 0.1% (>0.06%) due to thin clouds, c) at 21:30:00 and 21:30:30 UTC. This image was not selected as clear sky, as the fraction is 10.7% due to widespread thick cloud cover.

The idea behind doing this was that, if a given set (5 minute) represents clear sky, then the RGB pixel values won't change much between consecutive images. If clouds enter the sky image area, then that will result in high difference in RGB pixel values (i.e. above threshold) and that will be captured by the algorithm. 5 minute sets were chosen to provide sufficient time for cloud advection, whereas at the same time to avoid losing a considerable clear sky period because of clouds in only one or two of the images.

Chapter 3 in part, is currently being prepared for submission for publication of the material. Pawar, Prathamesh; Murray, Keenan; Kleissl, Jan. The thesis author was the primary investigator and author for this material.

4.   RESULTS AND DISCUSSIONS

The ACSL was run for Nov 2013 and the RBR of images found to be clear were stored. For the same time period clear sky images were selected manually and stored as RBR images as well. The total number of clear sky images selected was 3224 for manually created CSL(MCSL) and 4200 for ACSL.
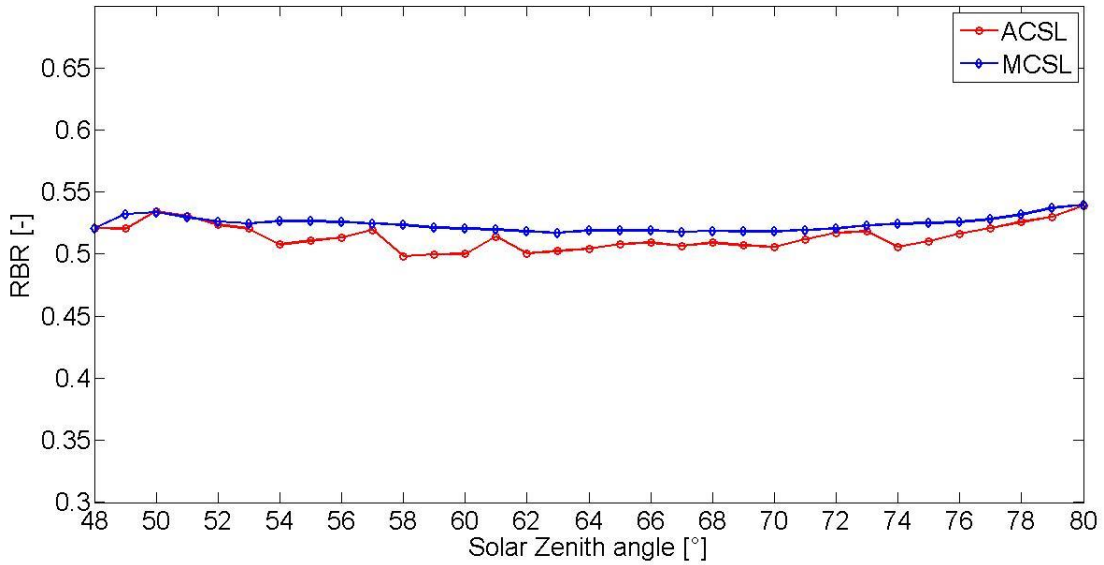

Figure 4.1: Comparison of mean RBR versus SZA for ACSL and MCSL, Nov 2013

Mean RBR at different SZAs are compared in Figure 4.1 for both methods. The maximum mean absolute percentage error (MAPE) of ACSL RBR with respect to MCSL RBR for full range of SZAs is 2% and ACSL always reports a lower RBR than MCSL. MAPE for each SZA is calculated as,

$$MAPE = \frac{100}{P*Q} \sum_{p=1}^{P} \sum_{q=1}^{Q} \left| \frac{RBR_{MCSL_{p,q}} - RBR_{ACSL_{p,q}}}{RBR_{MCSL_{p,q}}} \right|$$

13

Figure 4.2 shows the difference between ACSL and MCSL for RBR averaged over SZAs during midday and near sunrise and sunset.

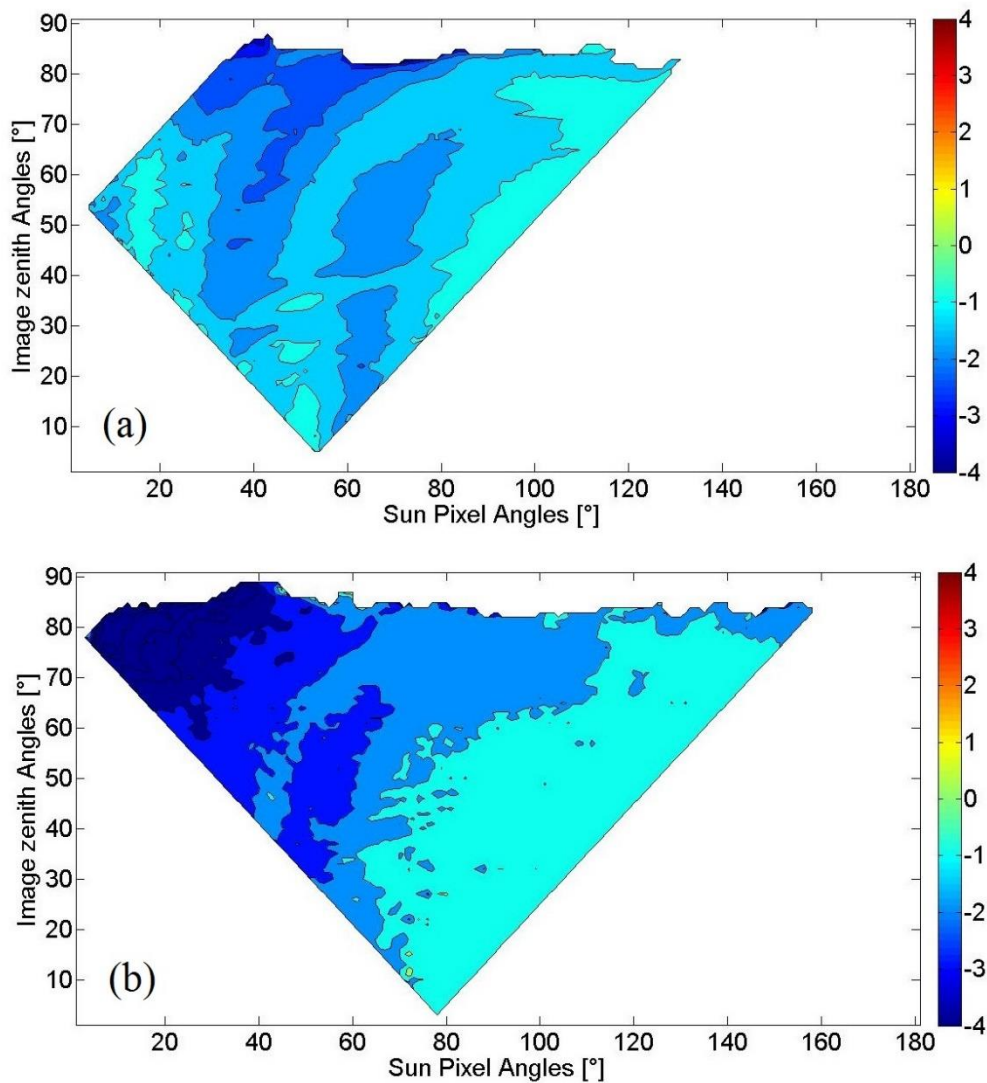*Figure 4.2: RBR difference in % of the mean RBR between ACSL and MCSL for a) SZAs 48° to 57°, and b) SZAs 74-80° for November 2013*

RBR of ACSL is almost always smaller than RBR of MCSL, consistent with what was observed from Figure 4.1. This can be attributed to hazy conditions or small clouds neglected during manual selection of clear sky images, because clear sky has

lower value of RBR than cloudy sky [14]. That means that the ACSL does not detect cloudy conditions as clear sky and prevents "polluting" the clear sky library with cloudy images which in turn would have rendered the whole cloud detection process faulty. Figure 4.3 shows root mean square error (RMSE) as a function of SZA.

Following formula was used for calculating the RMSE:

$$RBR_{diff}(p,q) = RBR_{ACSL}(p,q) - RBR_{MCSL}(p,q);$$

$$p=1: P, q=1: Q;$$

$$RMSE = \sqrt{\frac{\sum_{p=1}^{P} \sum_{q=1}^{Q} \left(RBR_{diff}(p,q)\right)^2}{P * Q}};$$

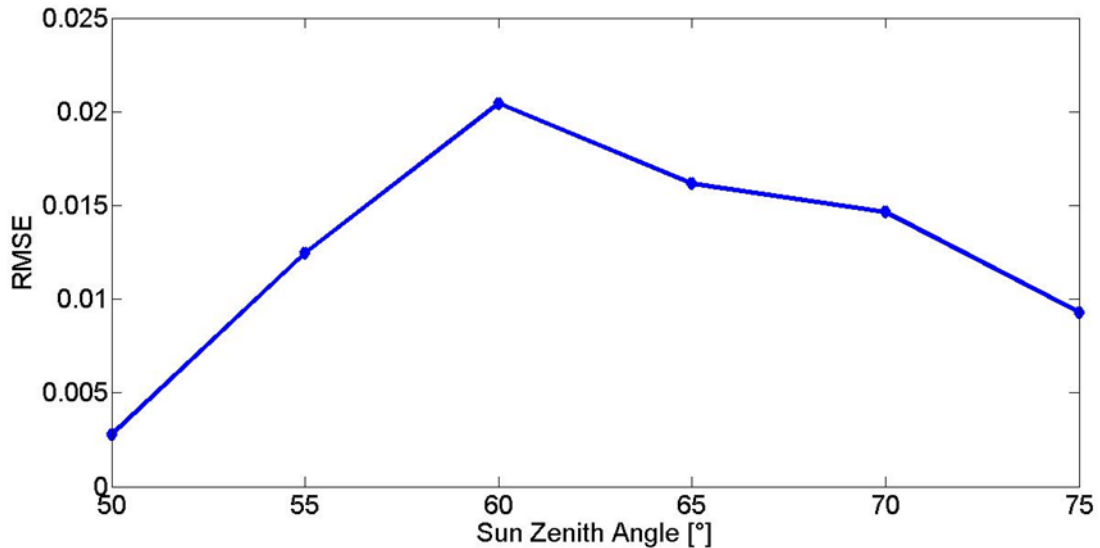where, *P* and *Q* are the total number of pixels along the axes of image area.

Figure 4.3: RMSE between RBR of ACSL and MCSL. Each marker represents the center of a bin spanning a SZA range of 5°, and the RBR values are averaged over these SZAs before taking RMSE.

The relative root mean square error was calculated by dividing the RMSE by the mean RBR. The relative RMSE range for RBR values of ACSL and MCSL was found

to be in the range of 0.8 to 5%. RMSE values are in the range of 0.004 to 0.021 which is small considering the scale of RBR which is 0 to 1.

From all these results it is evident that ACSL captures clear sky images successfully, and it is very robust in doing that. It is meticulous and cautious in selecting clear sky images as compared to MCSL, and also it has the capability to capture discrete clear sky periods which is very difficult to achieve by visual inspection. This was one of the reasons why ACSL actually captured more clear sky images than MCSL. It clearly possesses the ability to update the CSL with recent clear sky periods without any manual intervention and avoiding any scope for human errors.

Chapter 4 in part, is currently being prepared for submission for publication of the material. Pawar, Prathamesh; Murray, Keenan; Kleissl, Jan. The thesis author was the primary investigator and author for this material.

5. CONCLUSIONS

ACSL has been observed to perform well for the selected month. RMSE of the RBR in the CSL between MCSL and ACSL was found to be in the range of 0.004 to 0.025. ACSL selected more clear sky images than MCSL, which means that it obtains a more statistically converged CSL over a shorter time span. The mean ACSL RBR was found to be lower than that of MCSL, which confirms that the algorithm is accurate and in fact performs better than the manual process, considering that RBR is lower for clearer sky. ACSL will help in storing clear sky periods which reflect the latest sky and imager conditions, like aerosol effects or imager dome scratches. These changes will not affect the working of ACSL because it works on relative differences, i.e. it compares two subsequent images which would experience similar aerosol or imager properties; thus ACSL can be applied to any sky imager in general. ACSL will be tested further on different sky imagers and thresholds can be adjusted for better performance.

Chapter 5 in part, is currently being prepared for submission for publication of the material. Pawar, Prathamesh; Murray, Keenan; Kleissl, Jan. The thesis author was the primary investigator and author for this material.

REFERENCES

[1] J. E. Shields, M. E. Karr, R. W. Johnson and A. R. Burden, "Day/night whole sky imagers for 24-h cloud and sky assessment: history and overview," *Applied optics,* pp. 1605-1616, 2013.

[2] R. Marquez and C. F. Coimbra, "Intra-hour DNI forecasting based on cloud tracking image analysis," *Solar Energy,* vol. 91, pp. 327-336, 2013.

[3] B. Urquhart, C. Chow, D. Nguyen, J. Kleissl, M. Sengupta, J. Blatchford and D. Jeon, "Towards intra-hour solar forecasting using two sky imagers at a large solar power plant," *Proceedings of the American Solar Energy Society, Denver, CO, USA,* 2012.

[4] B. Urquhart, M. Ghonima, D. Nguyen, B. Kurtz, C. Chow and J. Kleissl, "Sky imaging systems for short-term solar forecasting," in *Solar Energy Forecasting and Resource Assessment*, 2013.

[5] M. S. Ghonima, B. Urquhart, C. W. Chow, J. E. Shields, A. Cazorla and J. Kleissl, "A method for cloud detection and opacity classification based on ground based sky imagery," *Atmospheric Measurement Techniques,* pp. 2881-2892, 2012.

[6] J. Shields, M. Karr, A. Burden, R. Johnson, V. Mikuls and J. Streeter, "Research toward Multi-Site Characterization of Sky Obscuration by Clouds, Final Report for Grant N00244-07-1-009, Marine Physical Laboratory, Scripps Institution of Oceanography,University of California San Diego, Technical Note 274.," 2009.

[7] H. Yang, B. Kurtz, D. Nguyen, B. Urquhart, C. W. Chow, M. Ghonima and J. Kleissl, "Solar irradiance forecasting using a ground-based sky imager developed at UC San Diego," *Solar Energy,* pp. 502-524, 2014.

[8] M. S. Ghonima, "Electronic Theses and Dissertations UNIVERSITY OF CALIFORNIA , SAN DIEGO Aerosol effects on Red Blue Ratio of Clear Sky Images , and Impact on Solar Forecasting A Thesis submitted in partial satisfaction of the requirements for the degree Master of Scienc," 2011.

[9] P. Davison, D. Roberts, R. Arnold and R. Colvile, "Estimating the direct radiative forcing due to haze from the 1997 forest fires in Indonesia," *Journal of Geophysical Research, [Atmospheres],* vol. D10207, p. 109, 2004.

[10] N. Feng and S. A. Christopher, "Clear sky direct radiative effects of aerosols over Southeast Asia based on satellite observations and radiative transfer calculations," *Remote Sensing of Environment,* vol. 152, pp. 333-344, 2014.

[11] B. Urquhart, B. Kurtz, E. Dahlin, M. Ghonima, J. E. Shields and J. Kleissl, "Development of a sky imaging system for short-term solar power forecasting," *Atmospheric Measurement Techniques,* pp. 875-890, 2015.

[12] J. Yang, Q. Min, W. Lu, W. Yao, Y. Ma, J. Du, T. Lu and G. Liu, "An automated cloud detection method based on the green channel of total-sky visible images," *Atmospheric Measurement Techniques,* pp. 4671-4679, 2015.

[13] C. N. Long, J. M. Sabburg, J. Calbo' and D. Pages, "Retrieving Cloud Characteristics from Ground-Based Daytime Color All-Sky Images," *JOURNAL OF ATMOSPHERIC AND OCEANIC TECHNOLOGY,* pp. 633-652, 2006.

[14] C. W. Chow, B. Urquhart, M. Lave, A. Dominguez, J. Kleissl, J. Shields and B. Washomc, "Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed," *Solar Energy,* pp. 2881-2893, 2011.