Title
Solar Resource Assessment with Sky Imagery and a Virtual Testbed for Sky Imager Solar Forecasting

Permalink
https://escholarship.org/uc/item/25v718t5

Author
Kurtz, Benjamin Bernard

Publication Date
2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Solar Resource Assessment with Sky Imagery and a Virtual Testbed for Sky Imager Solar Forecasting

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Engineering Sciences (Mechanical Engineering)

by

Benjamin Bernard Kurtz

Committee in charge:

      Professor Jan Kleissl, Chair
      Professor Carlos Coimbra
      Professor Amato Evan
      Professor Joel Norris
      Professor George Tynan

2017

The dissertation of Benjamin Bernard Kurtz is approved, and it is acceptable in quality and form for publication on micro-film and electronically:

_____

_____

_____

_____

_____
Chair

University of California, San Diego

2017

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

As in most things in life, the support and feedback of those around me has been critical to the success of my dissertation.

First and foremost, I am indebted to my advisor, Jan Kleissl, and the rest of my committee for their advice and guidance over the last six years. I have also come to appreciate the wisdom of Thijs Heus with regard to the appropriate scope of a doctoral thesis. Janet Shields knows more about sky imagers than perhaps anyone, and was kind enough to impart some of her knowledge to me.

Furthermore, I have been luck to have ideas, support, and assistance from many other students here. Thanks particularly to Bryan Urquhart, Chi Wai Chow, Elliot Dahlin, Andu Nguyen, Handa Yang, Oytun Babacan, Felipe Mejia, Guang Wang, and Keenan Murray for their help building, maintaining, and cleaning the sky imagers.

Finally, and perhaps most importantly, my family and my friends, especially Priya Nayak, Daniel Freilich, Michael Erickstad, and Elaine Denny, have helped me to stay happy and sane through my time in this hell we call San Diego. Of course, none of this would have been possible without my parents. Thank you!

Chapter 2, in full, is a reprint of the material as it appears in "Measuring diffuse, direct, and global irradiance using a sky imager" (B. Kurtz and J. Kleissl) in *Solar Energy*, 2016. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, is a reprint of the material as it is expected to appear (accepted pending minor revisions) in "A Virtual Sky Imager Testbed for Solar Energy Forecasting" (B. Kurtz, F. Mejia, and J. Kleissl) in *Solar Energy*, 2017. The dissertation author was the primary investigator and author of this paper.

VITA

2011	BS in Physics, Caltech
2011-2017	Graduate Research Assistant, University of California, San Diego
2012	MS in Engineering Sciences (Mechanical Engineering), University of California, San Diego
2017	PhD in Engineering Sciences (Mechanical Engineering), University of California, San Diego

PUBLICATIONS

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*, 2014.

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

G. Wang, **B. Kurtz**, and J. Kleissl, "Cloud base height from sky imager and cloud speed sensor," *Solar Energy*, 2016.

A. Nguyen, M. Velay, J. Schoene, V. Zheglov, **B. Kurtz**, K. Murray, B. Torre, and J. Kleissl, "High PV penetration impacts on five local distribution networks using high resolution solar resource assessment with sky imager and quasi-steady state distribution system simulations," *Solar Energy*, 2016.

F. Mejia, **B. Kurtz**, K. Murray, L. Hinkelman, M. Sengupta, Y. Xie, and J. Kleissl, "Coupling sky images with radiative transfer models: a new method to estimate cloud optical depth," *Atmospheric Measurement Techniques*, 2016.

B. Urquhart, **B. Kurtz**, and J. Kleissl, "Sky camera geometric calibration using solar observations," *Atmospheric Measurement Techniques*, 2016.

Z. Pecenak, F. Mejia, **B. Kurtz**, A. Evan, and J. Kleissl, "Simulating irradiance enhancement dependence on cloud optical depth and solar zenith angle," *Solar Energy*, 2016.

B. Kurtz and J. Kleissl, "Measuring Diffuse, Direct, and Global Irradiance Using a Sky Imager," *Solar Energy*, 2017.

B. Kurtz, F. Mejia, and J. Kleissl, "A Virtual Sky Imager Testbed for Solar Energy Forecasting," *Solar Energy*, accepted pending revisions.

ABSTRACT OF THE DISSERTATION

Solar Resource Assessment with Sky Imagery and a Virtual Testbed for Sky Imager Solar Forecasting

by

Benjamin Bernard Kurtz

Doctor of Philosophy in Engineering Sciences (Mechanical Engineering)

University of California, San Diego, 2017

Professor Jan Kleissl, Chair

In recent years, ground-based sky imagers have emerged as a promising tool for forecasting solar energy on short time scales (0 to 30 minutes ahead). Following the development of sky imager hardware and algorithms at UC San Diego, we present three new or improved algorithms for sky imager forecasting and forecast evaluation.

First, we present an algorithm for measuring irradiance with a sky imager. Sky imager forecasts are often used in conjunction with other instruments for measuring irradiance, so this has the potential to decrease instrumentation costs and logistical complexity. In particular, the forecast algorithm itself often relies on knowledge of the

current irradiance which can now be provided directly from the sky images. Irradiance measurements are accurate to within about 10%.

Second, we demonstrate a virtual sky imager testbed that can be used for validating and enhancing the forecast algorithm. The testbed uses high-quality (but slow) simulations to produce virtual clouds and sky images. Because virtual cloud locations are known, much more advanced validation procedures are possible with the virtual testbed than with measured data. In this way, we are able to determine that camera geometry and non-uniform evolution of the cloud field are the two largest sources of forecast error.

Finally, with the assistance of the virtual sky imager testbed, we develop improvements to the cloud advection model used for forecasting. The new advection schemes are 10–20% better at short time horizons.

Chapter 1

Introduction

In recent years, there has been a significant effort to shift toward the generation of energy via renewable sources such as wind and solar. A significant drawback of these compared to conventional energy sources is that they cannot easily be controlled by the operator; there is no way to "turn up" the sun when more energy is needed. Worse still, even predicting the natural variability of wind and solar is non-trivial.

As solar forecasters, our goal is to provide accurate estimates of power availability to allow the operators of the electrical grid sufficient advanced warning to ensure that there is always enough energy available on the grid to meet customer demands. Today, this means that sudden drops in output from renewable generators must be offset with additional power from fossil fuel plants or energy storage systems. In the future, utilities will likely also have some ability to request that customers defer low-priority loads during times of reduced availability of renewable energy; energy storage is expensive, while in many cases the cost of running the dishwasher a few hours later is negligible.

1.1   Background

Predicting availability of solar energy can be subdivided into two main tasks: predicting the available solar resource and determining how much of that resource can be converted into electricity by a given photovoltaic (PV) panel or solar thermal plant. Although modeling of the conversion process is important, it is understood to relatively high accuracy for most PV technologies [48]; we will focus here on resource prediction.

Solar resource is measured as irradiance—radiative power per unit area. Global irradiance includes all the light falling on a surface, while direct irradiance (sometimes called beam irradiance) does not consider the diffuse light that has been scattered while passing through the atmosphere. Because the angular distribution of light is non-uniform, and because the projected area of a beam changes as it falls on a surface at different angles, irradiance measurements also depend on the orientation of the measurement surface. Global Horizontal Irradiance (GHI) is measured on a horizontal plane, and is a standard measure when no more specific requirements are present. However, to predict the energy produced by a given PV system, which will likely be installed at a tilt, the Plane-of-Array (PoA) irradiance is a more useful measure. Furthermore, for systems that concentrate sunlight—either to produce higher temperatures in a solar thermal plant or for use in a concentrated-PV system which can have much higher efficiency than standard PV—only the direct beam can be concentrated, so the most relevant measure is Direct Normal Irradiance (DNI), the direct beam irradiance on a plane facing the sun.

For timescales less than a few days, the dominant factors influencing solar irradiance are clouds and the diurnal cycle of the sun; on longer timescales, the most accurate forecasts must also account for changes in atmospheric aerosols, although the influence of these is generally small in comparison to clouds. As changes in aerosols have a smaller influence and the diurnal motion of the sun is well understood and easily accounted for,

we will interest ourselves primarily with forecasting clouds.

In what follows, we will largely treat the forecasting of DNI and GHI as being similar, though in fact it differs slightly. Clouds scatter light, effectively converting irradiance from direct to diffuse (in addition to scattering some of it back into space). As such forecasting DNI is less forgiving because small errors in cloud cover tend to lead to much larger changes in irradiance. On the other hand, forecasting GHI depends on knowledge of cloud state over a wider area, and produces a wider variety of results (DNI can often be well approximated as a binary value).

Solar irradiance forecasting can be accomplished by a number of means, typically dependent on how far into the future one wants to forecast, hereafter referred to as the "forecast horizon." For the very shortest time horizons—from a few seconds out to 30 minutes—forecasts are typically based on local measurements or imagery. For example, one might try to guess future irradiance based on a history of recent irradiance measurements. A more sophisticated technique uses a sky imager to capture photographs of clouds as they move overhead, which allows forecasting over a larger area, and with more information about what clouds will be arriving in the future (as opposed to clouds that have already affected the sensor) [70]. Another related technique uses a camera mounted at a high point looking downward to capture images of cloud shadows moving past on the ground [26]. For longer forecast horizons, say, out to a few hours, knowledge of past conditions and very local clouds are not typically sufficient, so forecasts switch to being based on satellite imagery. Satellites can see clouds over a much larger area, however the images are recorded much less frequently, and have much lower spatial resolution (0.5 km) than local imagery, so satellite forecasts are not capable of predicting or measuring fast fluctuations in irradiance. Beyond 4-6 hours, the current state of the clouds is not sufficient to predict the future state of the atmosphere. To predict from a few hours up to (with decreasing accuracy) as much as two weeks ahead, forecasters

rely on numerical weather prediction (NWP) models, which attempt to calculate the fluid dynamics and thermodynamics of the atmosphere, and can therefore predict the formation and dissipation of clouds, which none of the previous models can do. Although modern computers are pushing the bounds of what is possible, NWP models require steadily increasing computational power to resolve details at small spatial and temporal scales, and so forecasts based on satellite or local imagery are typically preferred over NWP models for applications where high spatial or temporal resolution and low forecast horizon is needed.

For all methods on all timescales, it is common to do some additional statistical post-processing to a forecast in order to produce more accurate results. For example, if historical trends indicate that a NWP model tends to under-predict clouds at UC San Diego on average, forecast outputs can be adjusted accordingly and will typically then perform better in the future. Statistical comparisons with historical data can also be used to combine forecast models into an "ensemble" forecast, which may simply include multiple slight variations of a given model, or can be used to incorporate different types of models as well in order to perform well across all time scales. Of note, machine learning has recently become a popular [10, 67, 20] tool for forecasting, however by-and-in-large it can be considered in the category of statistical post-processing, as it typically combines current measurements and results of other physical forecasts to produce a more accurate forecast, albeit using very sophisticated techniques.

1.2 Overview of the Dissertation

In this dissertation, our focus will be on physical models for forecasting using local sky imagery. These forecasts consist, at a high level, of three major steps (illustrated in Figure 1.1): 1. detection and geolocation of clouds in the sky image, 2. detection of

| Example USI Image | Detect Clouds | Map Clouds to 3D Position | Forecast Future Cloud Positions | Calculate Irradiance at Surface |
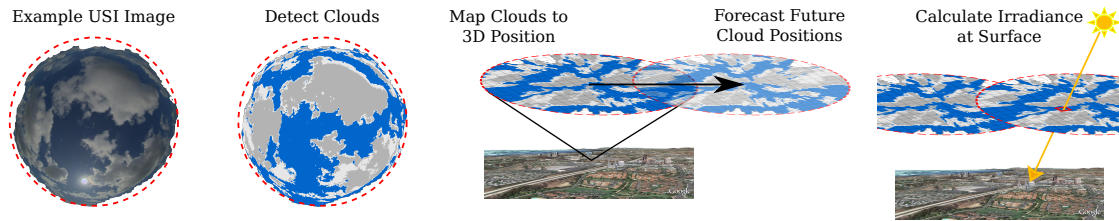
Figure 1.1: Steps in a typical sky imager forecast

Figure 1.2: Photographs of UC San Diego Sky Imager (USI) hardware

cloud motion and prediction of future cloud positions, and 3. prediction of irradiance based on those positions. Our team at UC San Diego has developed both sky imaging hardware and a corresponding forecast algorithm. The imaging hardware (shown in Figure 1.2) has been described and characterized in multiple publications [63, 64] and consists of a high-quality CCD camera with a fisheye lens housed in a weather-tight, climate-controlled enclosure. The hardware system also includes an embedded CPU which manages image capture and data transfer, and can be used to process images using the forecast algorithms when such real-time processing is desired. The forecast algorithms [9, 15, 70] will be the basis for our present discussion. More details on the functionality of the forecast algorithm are given in Section 3.2.3, as well as in our previously published work [70].

While we have primarily been interested in sky imagers as a technique for solar forecasting, other methods (particularly satellite imagery, as data rates and spatial

resolution have increased in recent years) have begun to encroach on the traditional domains where sky imagery is preferred. The remaining domains—very high spatial and temporal resolution for time horizons less than 15-20 minutes on domains only a few km in size—are of questionable interest to the electrical grid operators who are conventionally the users of a solar forecast. In particular, utilities often do not need detailed solar forecasts at these small spatial and temporal scales because power is usually balanced on much larger spatial scales on which spatial aggregation effects [27] dramatically reduce the effects of local variability, making forecasting on those larger spatial scales more relevant. Despite this, we consider studies of sky imagers to have some merit. First, there remain a few cases—for example, large PV generating stations—where spatial aggregation/smoothing may not be sufficient to alleviate need for a fine-grained forecast. In addition, we have used sky imagers (in a non-forecasting capacity) to study congestion due to PV in electrical distribution systems [39], for which satellite resolution is insufficient. Furthermore, while satellite resolution may increase to the point that ground-based imagers are no longer relevant, many of the techniques for forecasting cloud positions based on imagery will remain relevant. Finally, ground-based imagers will always have an advantage where resolution is concerned, and the ability to accurately detect clouds in these images (particularly in 3D [32]) could be very beneficial for research that attempts to understand the details of cloud structure and evolution, for example to inform development of advanced NWP models.

The chapters that follow describe work by the author to further the application of sky imagers to the field of solar resource assessment and forecasting. Chapter 2 documents a technique that allows DNI and GHI to be measured with a sky imager. While the accuracy of the sky imager does not match traditional instruments for measuring irradiance, it will be sufficient for many purposes, and in particular we hope to use it to reduce the dependence of our sky imager forecast algorithm on external data streams,

as needing to combine data from multiple sources greatly complicates operational deployment. Chapter 3 describes a setup for validating sky imager forecast algorithms at a high level of detail using simulated clouds and irradiance. This allowed us to determine with greater certainty than ever before that the lion's share of the errors in our forecast algorithm originate from the non-3D nature of our cloud map and from the frozen cloud advection (clouds stay a constant size and shape; they only move) that we originally used to predict future clouds. A colleague is addressing the 3D issue as part of his dissertation, and Chapter 4 investigates more advanced techniques for cloud advection and evolution. Finally, Chapter 5 provides some concluding remarks and thoughts on directions for future work.

Chapter 2

Measuring Diffuse, Direct, and Global Irradiance Using a Sky Imager

2.1   Introduction

Global horizontal irradiance (GHI) and direct normal irradiance (DNI) are two essential parameters in solar resource assessment. GHI is the standard measure of total available solar radiation, and is typically measured using a thermopile or photodiode pyranometer. DNI is the intensity of the direct solar beam and is the portion of the solar resource used by concentrating solar technologies; it is also important for calculating total irradiance on a tilted plane. DNI measurement is typically accomplished using a pyrheliometer on a solar tracker. Radiation coming from the rest of the sky, diffuse horizontal irradiance or DHI, can be measured by shading a pyranometer with a shade ball mounted on a solar tracker [56]. Measurements of GHI, DNI, and DHI made with high-quality, well-maintained instruments have 95% confidence intervals around $\pm$2% for DNI and from $\pm$3% to $\pm$10% for GHI, depending on the solar zenith angle [56]. Researchers at the National Renewable Energy Laboratory (NREL) find that compared

to standard-class reference measurements of GHI, thermopile pyranometers have 1-hour RMSE between 1.5 and 5% depending on the configuration (ventilation, thermal corrections, etc), while silicon-based pyranometers such as the LI-200 have 1-hour RMSE of 3.5% [19]. Accuracy of silicon-based sensors can be improved by correcting for their angular and spectral response. For precise angular measurements, scanning radiometers or spectroradiometers have narrow fields of view and can be positioned at will to sample radiance from various parts of the sky.

Alternative techniques intended to have lower upfront and maintenance costs are often proposed. Rotating Shadowband Irradiometers (RSIs) are devices that have a single pyranometer and a rotating shadowband to measure both GHI and DHI from which DNI can be computed. Typical 95% confidence intervals for the resulting DNI measurement are around $\pm 5\%$ [56]. However, the single axis of rotation and the typical use of a diffuser rather than an optical dome mean that RSIs are much less sensitive to maintenance schedules [56]. Another approach, taken by the SPN1 sunshine pyranometer, involves measurements using seven miniature thermopiles underneath a complex shading dome. The shading dome is constructed so that for any sun position, at least one sensor is fully shaded and at least one is fully exposed, so that in principle direct and diffuse components of the irradiance can be calculated. This eliminates moving parts, but has been found to result in a systematic positive bias in DNI measurements of $1.1 - 4.1\%$ with 1-minute RMSE of 8 to 14% even after advanced calibration as reported in Badosa et al. [2]. An NREL study found that the SPN1 had reasonable MBE (mean bias error) and RMSE (of 3% and 5%, respectively) on 1-hour GHI, but larger errors (MBE 7.2% and RMSE 9.5%) on 1-hour DNI [19].

There is a long history of using cameras to measure the brightness pattern of the sky. They have an advantage in angular resolution over hemispherical sensors, and an advantage in speed over scanning radiometers. As early as 1970, film cameras [11] and

later CCDs [55, 38] were used to measure the sunshapes—the normalized azimuthal average profile of broadband radiance about the sun. Techniques generally involve imaging the sun through one or more dark filters to bring it in range of the camera, and researchers were generally concerned only with relative brightness. However, Kaluza and Neumann detail the use of a CCD in concert with a flux gauge to produce absolute readings [24]. A modern commercial device called the Sun and Aureole Measurement system uses two cameras with different filter settings to image a larger region around the sun, and can be calibrated using a sun photometer to give absolute radiance as well [66]. On the wide angle front, many researchers over the years have used fisheye lenses to measure the radiance distribution of the sky [57, 50, 49, 71, 62], generally with absolute calibrations, but the direct beam is generally either blocked or saturated and therefore not measured, although Tohsing et al. estimated the direct beam at either its clear-sky value or 0 depending on whether the sun was obscured by clouds [62]. Several of these researchers were interested in luminance distributions (weighted for the human visual spectrum) for daylighting applications rather than solar energy production, but the procedures are much the same.

More recently, as whole-sky cameras have become common tools for solar energy forecasting, a few groups have extracted image features (such as textures, colors, and cloud coverage [54] or cloud locations [14]) to build empirical models for GHI and DNI without trying to measure the exact spatial distribution of radiance.

We propose a system that allows the use of a camera system described by Urquhart et al. [63] (the UCSD Sky Imager, or USI) to measure both DHI and DNI based on the photon fluxes incident on the image sensor. This physicality differentiates our method from those which rely on an image classification layer [54, 14], while our ability to make a non-binary determination of DNI is an improvement over radiance-based methods [62]. After introducing our data sources (section 2.2), we proceed with a discussion (section

2.3) of the methods used to measure and calibrate irradiance, with particular emphasis on the technique we have developed for measuring DNI using CCD smear. We then compare these results against LI-200 GHI and SPN1 DNI measurements (section 2.4) and show good qualitative and quantitative agreement. Finally, we conclude (section 2.5) with some ideas for improvement and additional applications.

2.2   Data

USI images for this study were captured every 30 seconds during daylight hours beginning on August 18, 2014 and ending on August 17, 2015. Data points corresponding to sun positions less than 20 pixels (2.5°) above the obstructions on the horizon were omitted in order to ensure that both the reference sensor and the USI were unaffected by said obstructions.

The reference GHI sensor, used for both calibrations and validation, is a LI-COR LI-200 silicon pyranometer located approximately 10 m from the camera. Reference data points are 3-second averages of 20 Hz samples centered at the time of the image. The data logger was intermittently overloaded with data acquisition instructions, and occasionally dropped up to a few minutes worth of data. Additionally, there were a few periods of longer downtime. In total, 15% of 3-second average data points at times corresponding to camera images were missing from the data set.

In April 2015, a Delta-T SPN1 sunshine pyranometer was added on a rooftop 1.25 km from the camera in order to be able to compare DNI and DHI as well. The SPN1 uses seven miniature thermopiles and a complex shading mask to measure both GHI and DHI, from which it calculates DNI. Data is sampled at 1 Hz. However, these validation data presented issues due to the spatial displacement. Purely random differences in high resolution data due to spatial displacement could be reduced by temporal averaging,

but in this coastal climate (1 km from the ocean) systematic differences in cloud cover exist with a bias towards more cloudy conditions near the ocean. While these clouds are typically thin and do not significantly affect GHI, DNI differences can be large. Since occasionally different cloud conditions persist for several hours these data are therefore suitable for quantitative comparisons only during time periods that are cloud-free. While differences in aerosol loading between the two sites could also change the partition from GHI to DNI, the strong sea breeze flow typically causes air masses of oceanic origin to exist at both sites.

2.3   Methods

We will decompose GHI into three separate components to be extracted from each set of images.

$$\text{GHI} = \text{DHI}_{\text{raw}} + \text{DNI} \cdot \sin(\alpha) - \text{DHI}_{\text{stray}}, \tag{2.1}$$

where α is the solar elevation above the horizon. The raw measurement of DHI will be derived from most of the pixels in the 16-bit high dynamic range (HDR) image captured by the camera described in Urquhart et al. [63]. Three images with adjacent exposure times varying by a factor of 4 are combined by averaging pixel readings exposed in the linear range of the detector. Unfortunately, without a filter changer, the camera does not have enough dynamic range to capture the intensities in the solar region directly—the sun is at least 15 times as bright as the brightest object the camera can record at its shortest exposure time. Therefore, pixels near the sun are always saturated when the sun is unobscured. DNI will be measured instead by measuring the intensity of CCD smear that results from the presence of very bright light sources such as the sun. As smear occurs during readout, it is expected to be independent of exposure time. This can lead

to complications in the HDR compositing process which expects pixel values to scale with exposure time. For this reason, a single exposure at the camera's shortest exposure time (75 μs) is used for smear measurement. The short exposure time enhances the brightness of smear relative to the rest of the image. Finally, DHI is corrected downward to account for stray light that is scattered off the camera optics when they are illuminated by direct sunlight. Note that the stray light correction does not have to be added back to the DNI measurement since the DNI method utilizes a scaling factor that is also expected to correct for stray light losses to the DNI signal.

Although insufficient to measure the direct solar beam, the dynamic range of the camera is generally sufficient to capture the remainder of the sky scene with a fixed set of exposures under all sky conditions. This simplifies the conversion from pixel values to radiance. Capture time for the sequence of four images (three for HDR plus a minimal-time exposure) is approximately 830 ms, of which roughly 200 ms is delay between frames.

All three components of Equation 2.1 require calibration in order to provide values in meaningful physical units. While optical setups could be calibrated in the lab, small changes in focus and alignment can lead to nontrivial changes in calibration factors; our objective is therefore to present an algorithm that can be self-calibrated using field measurements.

2.3.1   Diffuse Irradiance from HDR Image

Methods for absolute radiance calibration of the camera are fairly standard and are presented by other researchers [57, 62]. Essentially, a uniform, calibrated, Lambertian reflectance plaque is illuminated by a calibrated light source, and the distance from the source to the plaque is varied along with the camera exposure time, thereby recording the mapping from pixel values v and exposure times t into radiance L. DHI is then

a straightforward sum of the product of radiance, cosine of the angle of incidence θ, and solid angle $\delta\Omega$ over all pixels above the horizon and outside the direct beam region (typically between 2 and 5 degree radius from the sun, depending on the desired characteristics).

$$\text{DHI}_{\text{raw}} = \sum_{\text{sky hemisphere}} L(v,t) \cos\theta \, \delta\Omega \tag{2.2}$$

However, under many circumstances, a precise radiance calibration for a given instrument will not be available. Even a lab calibration alone may not be enough, as the relative position of the lens and the sensor may change slightly during camera deployment. When an accurate field calibration is unavailable, the pixel's radiance L is assumed to be linearly proportional to the pixel value v and exposure time t. Previous work has found pixel brightnesses to be linear within 10% for this camera [63]. The proportionality constant β typically has a small spatial variability due to nonuniform pixel sensitivity and optical effects. Pixel sensitivity is generally small and in any case varies from camera to camera. On the other hand, rolloff (the decreased transmission of the optics with increasing view-angle, sometimes known as vignetting) is primarily a function of the optics and the distance from the optical axis and is expected to be comparable between cameras of the same model and optical configuration. Radiance is therefore approximated as

$$L(v,t) = \beta_0 \beta'(r) vt, \tag{2.3}$$

where β_0 is now a constant across the image (assuming homogeneous pixel sensitivity) which can be estimated using reference GHI data. $\beta'(r)$ can be measured by capturing images under uniform illumination (e.g. several rotations of an integrating hemisphere). In the event that measurements for $\beta'(r)$ are unavailable, it is expected that reasonable results could be obtained taking $\beta'(r) = 1$ independent of r (i.e. assuming rolloff is negligible). Thus, a range of accuracy levels can be achieved, depending on the calibration

equipment available: fully calibrated radiance should give the best accuracy, but may be hard to achieve in most scenarios; GHI-calibrated, radially variable response is easier to achieve; and GHI-calibrated constant response requires no extra calibration of the camera, but results may be lower in accuracy. Here, we have taken the second approach, i.e. approximating $L(v,t)$ as indicated in equation 2.3 with $\beta'(r)$ dependent on r. We note that many low-cost cameras compress pixel values for storage with the result that values are no longer linearly related to brightness; users of such cameras may require extra steps to recover radiance from pixel values.

To determine β_0, DHI is calculated during times when DNI and therefore stray light contributions in equation 2.1 are expected to be nearly zero. β_0 is then determined using a linear fit against data from a nearby GHI sensor (see Fig. 2.11 later for an example). In our experience, one day with at least 40% low-DNI conditions is sufficient for the determination of β_0. Low-DNI conditions can be conveniently defined in this context as 1. average pixel brightness in the solar region ($< 2.5°$ from the sun) is less than 2.5 times the average pixel brightness outside the solar region, 2. raw DNI (h from equation 2.6 below) is nearly zero, specifically h is less than 1% of the 98th percentile, and 3. reference GHI is between 75 and 250 W/m^2. The first condition ensures that the difference between DHI and GHI is less than about 0.25%, which the second confirms by checking that measured DNI is low. The third condition avoids low solar elevations as well as bright and dark conditions in which nonlinearities were observed.

A standard pyranometer measurement includes the influence of any obstructions to its field of view. To replicate those conditions, we have opted not to mask trees and distant structures along the horizon before summing diffuse irradiance. However, an advantage of camera-based radiance measurements is that such obstructions can be removed if desired. Masked regions could then be filled with the average value from the rest of the sky, or with values interpolated from nearby portions of the sky using an
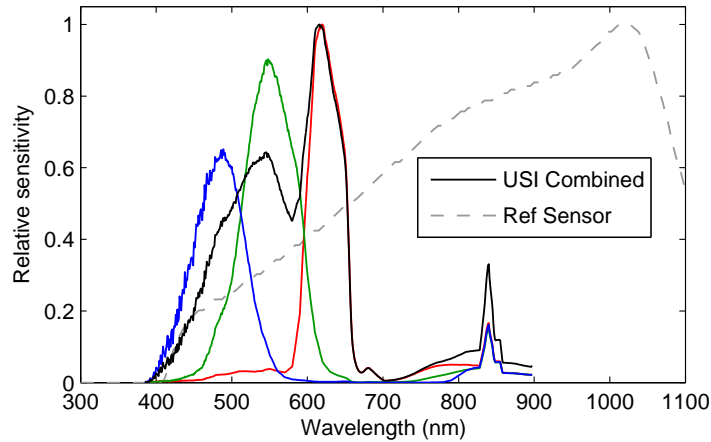
Figure 2.1: Red, green and blue lines illustrate the relative spectral sensitivity of the associated color channel in the USI image sensor [1, 12]. For comparison, the effective combined (RGB weights 1, 0.65, 0.5) sensitivity for DHI is shown in black, along with the sensitivity of the LI-200 reference instrument (dashed gray) [28].

appropriate model for anisotropic diffuse radiance.

The spectrum of DHI is markedly different under clear and overcast skies. It is therefore advisable to use a weighted combination of the different color channels in order to obtain a desired spectral response. Thermopile pyranometers are typically designed to have uniform spectral response (measurement of 1 W/m^2 is independent of the spectral content of the radiation). However, for some resource assessment applications, it may be desirable to select weights to match a particular instrument or a specific model of photovoltaic device. For the camera used in this study, weighting factors of 1 for red, 0.65 for green, and 0.5 for blue were found to provide the best match to the reference sensor (based on a linear least-squares fit of the spectral response in the 400-700 nm range where the camera has significant sensitivity), as illustrated in Figure 2.1.

2.3.2 Stray Light Correction

Previous work found stray light to contribute as much as $10-20\%$ of the observed radiance in some areas of the image [63]. Therefore a correction term was introduced in
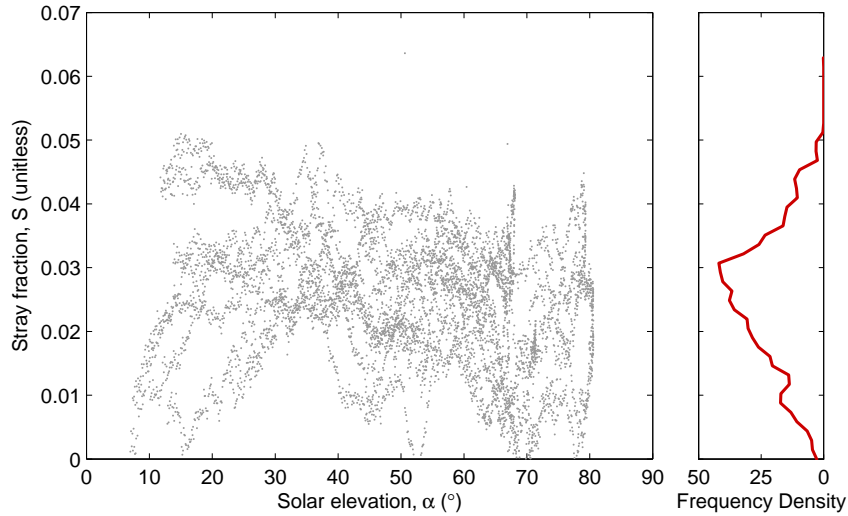
Figure 2.2: Difference between DHI_{raw} and DHI_{SPN1} as a fraction of DNI for several clear days between April and July 2015. The red line at the right indicates the frequency of a given stray fraction across all solar elevations. Strong dips (e.g. at $\alpha = 70$) arise from times where no single SPN1 sub-sensor is completely unshaded.

Eq. 2.1. Stray light is assumed to be due only to reflections and scattering of the direct beam in the optics. In the general case, we expect that the amount of scattering might depend on solar elevation, α, due to changes in angle of incidence and intersected area. We shall refer to the ratio of stray DHI to DNI as the *stray fraction*, S:

$$S \equiv \frac{DHI_{stray}}{DNI} = S(\alpha) \tag{2.4}$$

Although there is potential for a more complex relationship between stray fraction and solar elevation, Figure 2.2 indicates that the camera used in this study does not exhibit any appreciable relationship at all, i.e. $S(\alpha) = $ const. We suspect that this is not true for all cameras.

One potential issue with modeling stray light as a function of DNI is that localized areas of strong scattering in the optics—for example, scratches, smudges, or dirt—tend to cause increased stray light while simultaneously resulting in a lower DNI measurement

due to the increased attenuation. As this effect is similar to that of passing thin clouds, it is nontrivial to correct for, and we do not attempt to do so in this work.

2.3.3 Direct Beam Estimation

Smear and blooming in CCD sensors

As DNI of 5 W/m^2 is sufficient to saturate pixels in the solar region, it is impossible to measure the direct beam radiance directly using pixel values. Instead we turn to the smear effect of CCD sensors. In an interline-transfer CCD such as is used in the USI, charges are collected in a capacitor-like potential well at the site of each pixel during integration, and then shifted off of the chip into an A/D converter. Charges are shifted into and then down a so-called "interline transfer" column (illustrated in Figure 2.3), which is behind a mask to prevent additional light from changing pixel values during readout. However, in the presence of very bright light, some photons or excess electrons can still overflow into the interline transfer column during readout, causing smear. Note that as smear occurs during readout, the intensity of the smear effect will depend on frame readout time, rather than exposure time. For a review of CCD operation or more details on smear, readers are referred to [69, 41].

Smear is often accompanied by another effect known as blooming, which describes pixels that overflow into their neighbors during the exposure. These effects are illustrated in Figure 2.4. Blooming is difficult to separate from the saturation in the solar region, whereas smear extends through the entire column, making it easier to measure.

Smear intensity

In order to separate the smear from the rest of the image, we calculate the average value of each column, and then look for columns that have elevated values in
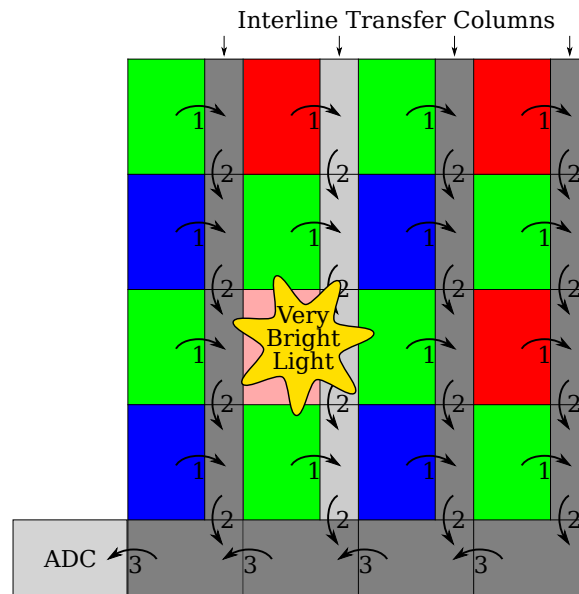
Figure 2.3: Illustration of CCD smear at the sensor level. Colored rectangles represent light-sensitive pixels of various colors, while gray rectangles indicate regions of the sensor behind a light-blocking mask. The CCD readout process first shifts charges into the interline transfer column (1), then shifts them down to a horizontal shift register one row at a time (2), and finally off the chip into the A/D converter (3), repeating 3 and 2 as necessary until all pixels have been read. Smear is introduced when photons or electrons from a brightly illuminated pixel leak into adjacent interline transfer pixels. The extra charges join those already collected for each pixel in the column as the original charges move past the brightly illuminated pixel.
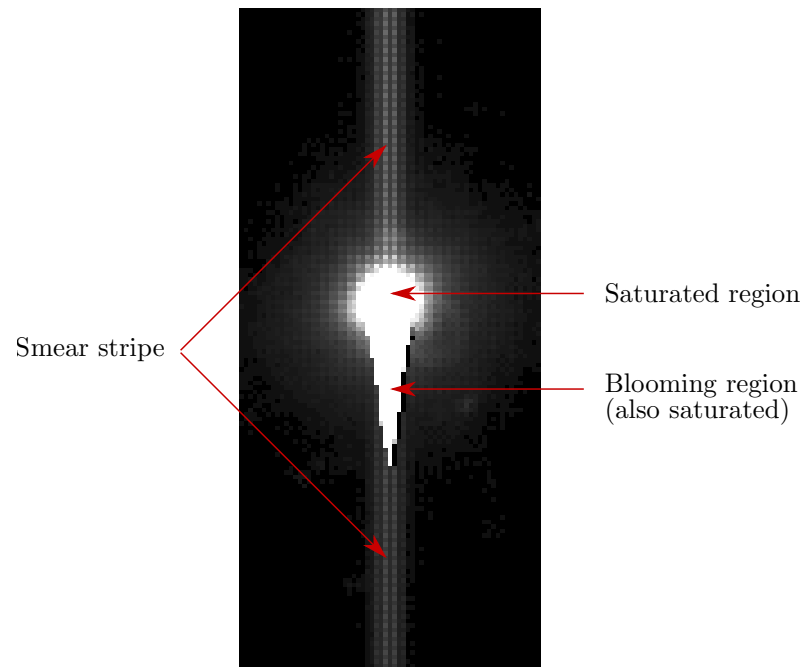
Figure 2.4: Solar region (67×147 pixels) of a 75 μs exposure exhibiting blooming and smear. Blooming is the overflow of CCD cells into neighboring cells, resulting in a pointy "tail" on the sun. Smear results in excess signal penetrating the interline transfer columns, which brightens the entire column as charges are shifted off the chip to the A/D converter.

comparison with those nearby. Removing 5 to 10% of the rows immediately around the sun reduces the influence of the saturated portion of the image on the column-average values. Additionally, some rows need to be removed due to side-effects of hardware cropping (discussed in Section 2.6.1) of the image.

Rows are therefore removed from about 300 pixels above to 200 pixels below the position of the sun. When there are saturated pixels in the image, the sun is taken to be at the center of the saturated region; otherwise, a modeled sun position is used, based on the geometrical calibration procedure described by Urquhart et al. [64]. If neither a calibration nor saturated pixels are available, sun position can be approximated based on the brightest pixels in the image.

Note that when taking column averages, it is important to separate the color channels of the image, as they are apparently affected differently. More specifically, half of the columns of the image contain red and green pixels, while the remainder contain green and blue as shown in Figure 2.3. Red filters have the highest transmissivity and blue the lowest, so alternating columns experience different amounts of smear. For reasons that are not well understood, the red pixels also exhibit higher smear than the green pixels in the same column, though the green and the blue in the other column are nearly the same as might be expected given that they share the same interline transfer column (see Section 2.6.2). We have opted to work primarily with the red pixels, as they exhibit the highest smear signal on the lowest background, since Rayleigh scattering is weakest for longer wavelengths (the sky is blue). Our feeling is that the reduction in error realized by a more careful consideration of spectral effects is likely to be small in comparison with the improvement from resolving outstanding difficulties with the DNI calibration procedures.

Next, the location and magnitude of the stripe are determined. The coarse sampling (only every second column contains red pixels, see Figure 2.3) and nonzero
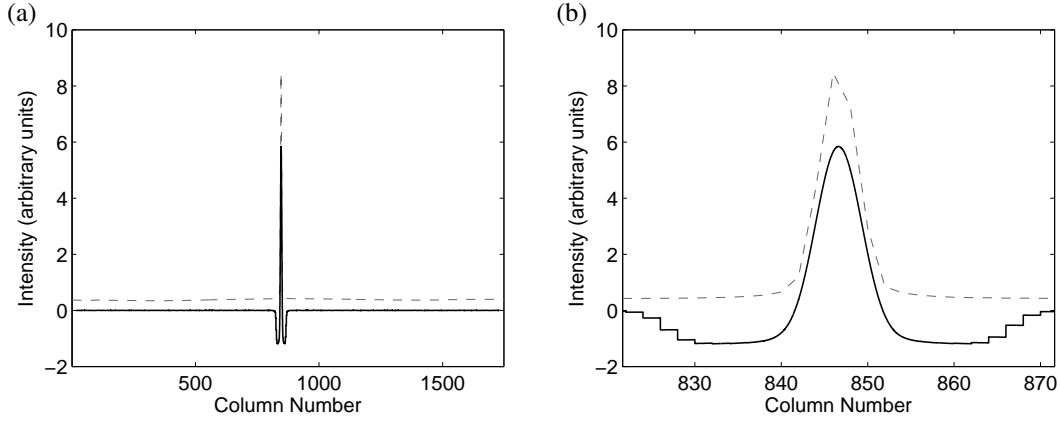
Figure 2.5: Column means for (a) all columns and (b) columns near the smear stripe, of the red channel before (dashed gray line, $\overline{v_x}$) and after pattern matching (solid black line, $h^\star$) with a 1.07 pixel Gaussian template as defined in Eqns. 2.5 and 2.6.

background (natural variation of red irradiance across the image) mean that using the maximum point of the column-wise average signal $\overline{v_x}$ is incorrect. Instead, taking the cross-correlation of $\overline{v_x}$ with a Gaussian template (kernel) yields a robust pattern-matching metric (Figure 2.5).

$$h^\star(x_0) = \sum_{x_0-19<x\leq x_0+19} \overline{v_x} \cdot \left(T(x-x_0) - \langle T \rangle \right) \tag{2.5}$$

Here, $T(x) = \exp(-x^2/2\sigma^2)$ is the Gaussian template, which has mean $\langle T \rangle$ over the range of x. A template width of $\sigma = 1.07$ pixels with a window size of 39 pixels (20 red pixels) was empirically chosen.

Figure 2.5 shows column-wise means of the red channel before and after pattern matching. The coordinates of the maximum point in this curve give the smear intensity (h) and the location of the smear stripe.

$$h = \max_{x_0}(h^\star) \tag{2.6}$$

Note that $h^\star(x_0)$ can be evaluated for non-integer x_0.

Calibration

The computed smear intensity, h, should be proportional to DNI, however as sensor manufacturers typically design to avoid smear rather than provide consistent smear, sensitivity and linearity vary considerably over the image. We expect the magnitude of the smear effect to depend on the location of the pixel on the sensor. This may be due, for example, to spatial differences in pixel sensitivity, and differences in the alignment of the metal mask covering the sensor. We also expect that there may be some variation with solar elevation, which affects various angles of incidence and the path length through the optics.

The raw smear intensities h therefore need to be calibrated to yield DNI. We begin by applying the same $\beta'(r)$ roll-off correction discussed for DHI, but an additional calibration factor (C) is required.

$$\text{DNI} = C(x, \alpha)h' = C(x, \alpha)\beta'(r)h \qquad (2.7)$$

$C(x, \alpha)$ are determined by one of two methods, and then the roll-off-corrected smear intensity h' is multiplied by the calibration factor to yield DNI. The simpler of the two methods is to assign a fixed calibration to each column ($C(x)$) based on a look-up table. The more complicated method also takes into account the elevation angle for each pixel ($C(x, \alpha)$), and uses an artificial neural network (ANN) for obtaining the relationship. The advantages of the simple method are its ease of implementation and predictable behavior for data outside the training period. The ANN generally produces better DNI values when interpolating within the range of solar elevations covered by the training data, but may not be reliable outside.

$C(x, \alpha)$ is trained using data from several clear days. Predominantly clear days were selected by manually screening a GHI time series, followed by removal of short

cloudy periods by visual inspection of images. As clouds will cause significant distur-bance in the calibration factors, a cursory inspection of calibrated results helps confirm that clear training data was used.

As a reliable source of DNI is not generally available to calibrate against, linearity of smear response is assumed and calibration is performed against the residual GHI after DHI (including stray light) is accounted for:

$$\text{DNI}_{\text{nom}} = \frac{\text{GHI} - \text{DHI}_{\text{raw}}}{\sin(\alpha) - S}, \tag{2.8}$$

where the GHI measurement here is taken from the reference sensor, while DHI comes from the camera. If available, data from a nearby well-maintained DNI sensor would be preferred and could alleviate the requirement of an empirical model for stray light, but such a sensor with sufficient data quality was not available for the present study.

Calibrations were found to change significantly when camera focus or other optics were adjusted, so the year of data was divided into three periods for calibration, as documented in Section 2.7. Example values of C as well as fits using both methods for one day are illustrated in Figure 2.6a. For comparison against the results in Section 2.4, the figure lists errors calculated using the standard metrics defined in Section 2.8.

The lookup-table based method estimates C from x alone. Training data are placed into 10-pixel bins in x, and the median value of each bin is used for the lookup table. Cubic spline interpolation in x is used between bin centers, and nearest neighbor extrapolation is applied to x below or above those present in the training data.

The ANN is implemented using MATLAB's neural network toolbox (`fitnet`) with a single 5-node hidden layer and is trained using Bayesian regularization (`trainbr`). Standard procedure for training neural networks is to divide the calibration data into a *training* set and a *validation* set. The error is minimized over the training set, while the
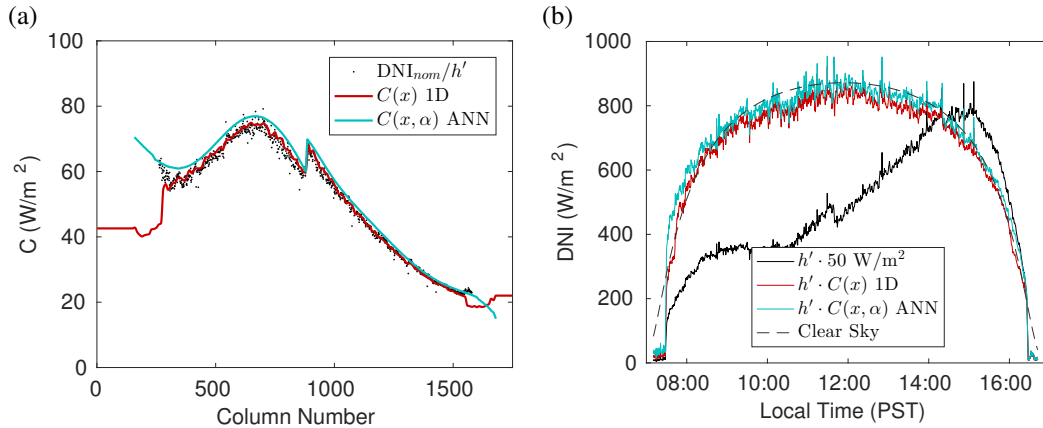
Figure 2.6: (a) Calibration factors (Eq. 2.7) and (b) DNI on a clear day (January 3, 2015) before and after calibration using a 1D lookup table (as a function of image column) and a neural network (which considers image column and solar zenith angle). On this day, squared errors are below average and bias errors are larger than average for both the ANN calibration (RMSE 4.5%, MBE 3.9 W/m^2) and the 1D lookup table calibration (RMSE 2.3%, MBE -7.0 W/m^2).

validation set is used as an independent measure of the error, which avoids overfitting by allowing training to stop before the minimum error is reached on the training set if validation set errors are increasing. In order to preserve the independence of the validation set, data from a given day are placed in one of the three sets (training, validation, testing). Specifically, in this case, the first two and last two clear days (generally the days with the highest and lowest solar elevations at solar noon) in each calibration period were placed in the validation set, while the remaining days were divided evenly between the training and testing sets. The ANN is trained 10 times with random starting weights, and the network with lowest errors on training and validation days combined is selected as the final calibration model. Since the USI employs separate A/D converters for the right and left sides of the sensor, there is a discontinuity in $C(x)$ near the center of the image. As a result, separate ANNs are trained for the left and right halves of the image.

The 1D lookup table method is less sensitive to the selection of clear days than the ANN-based algorithm. Nevertheless, best results are obtained if the clear days selected span the full range of seasons; of particular interest are long summer days which

potentially provide data for a wider range of x values than shorter winter days. For comparison purposes, we chose to train the lookup table method using the same clear days that were used for the training and validation sets for the ANN.

Sample calibrated results are shown in Figure 2.6b, along with DNI from the clear sky model of Ineichen and Perez (PV Lib implementation) [47, 44]. While the raw data do not correspond to a typical DNI signal, the calibrated data follow the clear-sky DNI with an RMSE of about 5%. About one sixth of the RMSE is due to short-term random fluctuations from localized variations in the optics.

2.3.4   Validation

GHI data during the study year are compared at all times when the sun is more than 20 pixels (approx. 2.5°) above the horizon and data are available from both the camera and the reference sensor, leaving a total of 424,033 images for comparison. Of these, 47,122 were used for calibration and validation, leaving 376,911 images for testing. RMSE and MBE against the reference sensor are computed using the standard definitions, given in Section 2.8 for completeness.

As previously mentioned in Section 2.3.3 and enumerated in Section 2.7, calibration was performed in three separate periods due to optical configuration changes. However, errors were similar across all three periods and are reported without reference to a particular period.

2.4   Results

Figure 2.7 shows example results from a partly cloudy day. The camera-derived GHI generally follows the reference sensor, increasing as the sun rises above the horizon in the morning and then dropping steeply as clouds move in the afternoon. Of particular
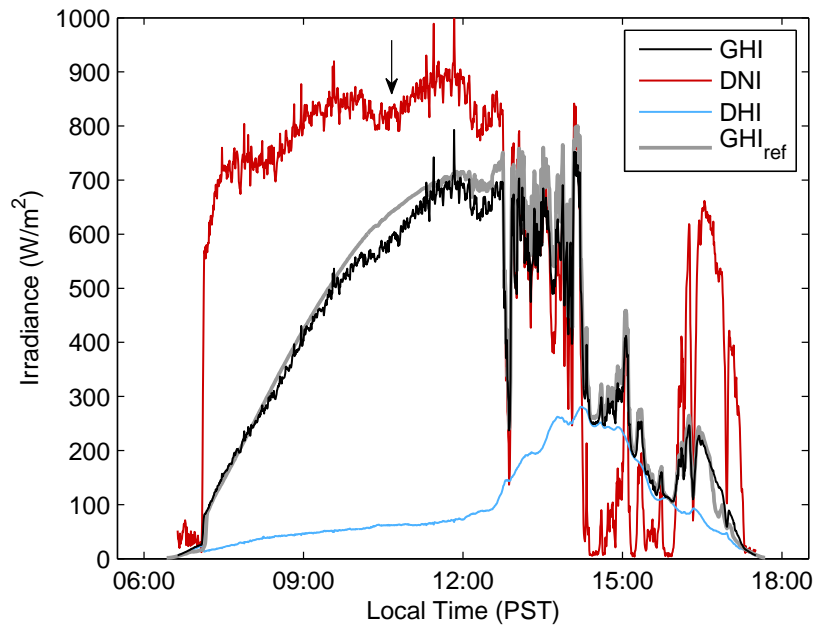
Figure 2.7: Results for a typical day, 2015-02-11, with the ANN calibration. GHI on this particular day exhibits slightly-below-average RMSE (8.1%) but has worse-than-average negative bias (−5.6%). Inconsistencies in the DNI calibration give rise to the dip (indicated by an arrow) in the DNI curve at 10:30 PST and contribute significantly to the negative bias.

note is the accuracy with which the camera follows the steep drops that result from shading of the direct beam. However, this example also illustrates several of the most prevalent sources of error. Although reference DNI measurements are unavailable for this day, conditions are essentially clear until 11:00 PST. This would normally suggest a relatively smooth DNI curve, which is not observed. In actuality, there appears to be significant random noise and a drop in DNI appears around 10:30 PST, both of which are carried over to GHI as well. Inspection of the raw smear intensity and calibration signal reveals that while the dip in DNI around 10:30 PST originates in the smear intensity, on other nearby days (including the two nearest clear days used for calibration), this dip is smaller and is properly corrected. For unknown reasons, some days (including 2015-02-11 and the second clear day before it) have a larger dip in smear, which in this case leads to poor performance even after calibration. Analogous errors on other days can lead to positive bias in DNI as well. Meanwhile, the increased noise in the camera's DNI measurement arises primarily from localized optical defects. When target calibration factors are displayed as a function of image position, some of these optical defects can be visually associated with dirt and scratches on the dome. To verify that noise is not caused by the sensor, for one day a burst of 10 images was taken within 0.7 sec rather than a single image every 30 sec. For each set of 10 images, the path through the optical system is essentially identical. Over the course of two clear hours, fractional variation of smear intensity within a burst of images was only 0.15% to 0.6% compared with 12% for images that were 30 seconds apart. Finally, DHI measurements are too small during times of high diffuse; on this day, this can be seen most clearly shortly before 15:00 PST when DNI ≈ 0, but is also demonstrated in Figure 2.11 later.

Daily error statistics for instantaneous GHI are presented in Figure 2.8. Errors tend to be lower on and near days used for calibration. Yearly statistics broken down by training and testing are shown in Figure 2.9a with RMSE for instantaneous GHI around

10%, and MBE $< 2\%$. RMSE for 1-hour average GHI is around 7%. The data generally follow the expected trends that data used for training or cross-validation should have lower errors than testing data, that the neural network should perform better than the 1-D lookup method, and that 1-hour averaging helps reduce errors due to random processes. The relatively small MBE is not surprising as the calibration of both DHI and DNI is based on the reference sensor.

A distribution of errors is shown in Figure 2.9b. Large overpredictions are significantly less common for the ANN calibration than for the lookup table calibration. Both calibrations have similar probabilities for large underpredictions. Figure 2.10 breaks out errors for low-DNI periods and clear days, which exhibit fewer large errors than the overall data set. However, the error distributions for hourly-averaged data no longer show this distinction; this suggests that many of the large errors may be due to the larger spatial and temporal variability in GHI under partly cloudy conditions, rather than because measurements are inherently better under clear or low-DNI conditions.

The fact that relative hourly errors for low DNI are comparable to other conditions suggests a prominent source of errors in conditions dominated by diffuse. By definition, low-DNI periods have GHI $\approx$ DHI, and therefore can be used to evaluate the quality of the DHI measurement. Figure 2.11 presents these errors for three different days as a function of GHI. June 27 was the day used to calibrate DHI for this period, and shows both positive and negative errors with no overarching trend. In contrast, data from June 15 and 16 have a strong non-linear trend toward more negative errors with larger GHI, which is also apparent on several other nearby days (not shown). Mornings and afternoons on these days also differ slightly, though not as much as from the conditions on June 27. Some of the error in DHI appears to be related to the difference between spectral channels; for example, on June 27, the errors tend to be positive when the blue channel is brighter than the red, and negative otherwise. However, attempts to correct
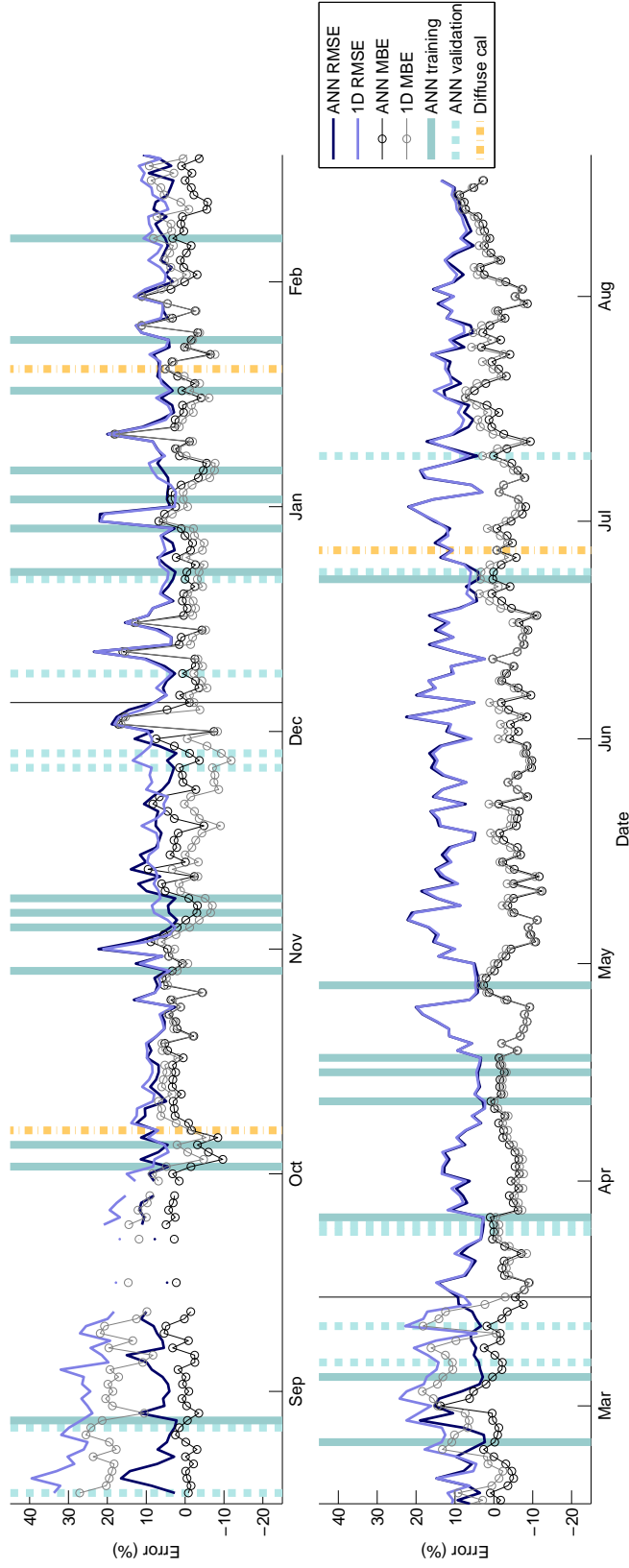
Figure 2.8: Daily GHI error normalized by daily mean GHI. Colored stripes in background indicate days used for training and calibration. Thin black vertical lines indicate separation between the different calibration periods (see Table A-1). Because the 1D lookup calibration does not distinguish between training and validation, days used for ANN validation are lumped into the training set for the 1D calibration.

(a)

(b)
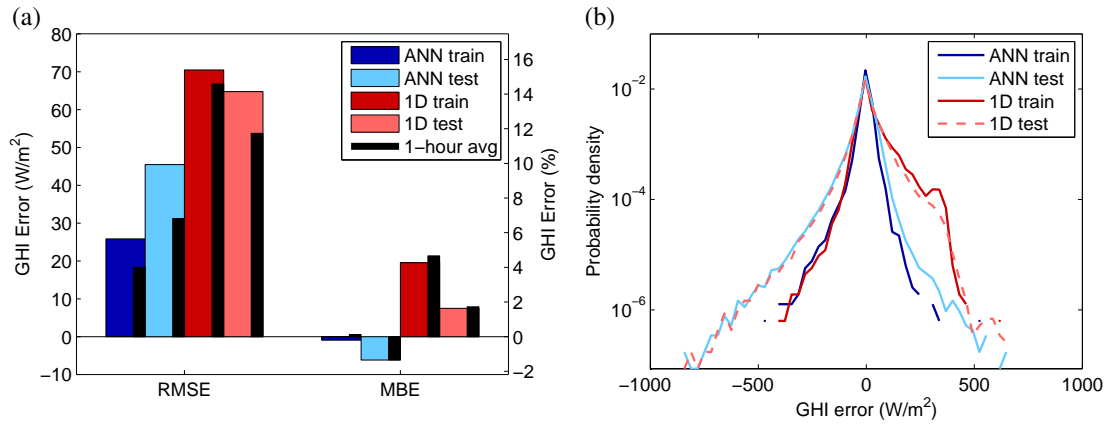
Figure 2.9: Error statistics for the test data set over the course of approximately one year. In (a), colored bars indicate error rates for instantaneous data, while black bars are for 1-hour averages that are often used in performance simulations. The right axis is scaled to the left axis using the mean of the testing data (cf. Equation 2.9); therefore the error percentages are not entirely accurate for the training data. Probability densities of instantaneous GHI errors are shown in (b) with the same color scheme as in (a).

(a)

(b)
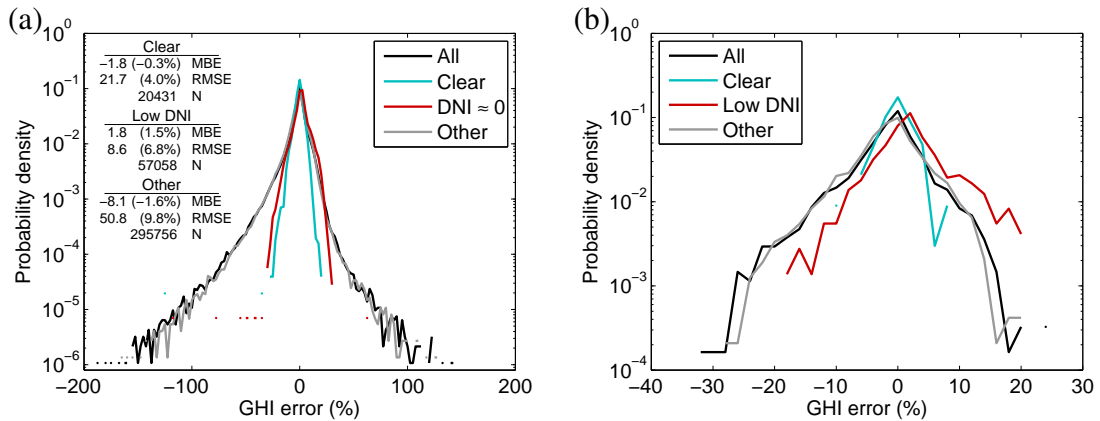
Figure 2.10: Error distributions for clear and low-DNI conditions, for (a) instantaneous data, and (b) 1-hour averaged data. The black "All" curve in (a) represents essentially the same data as the light-blue "ANN test" line in Figure 2.9. Low DNI periods are selected using the first two criteria from DHI calibration (no limitations on GHI, see Section 2.3.1), while clear conditions were determined manually. Error metrics in W/m^2 and image counts are given in (a).
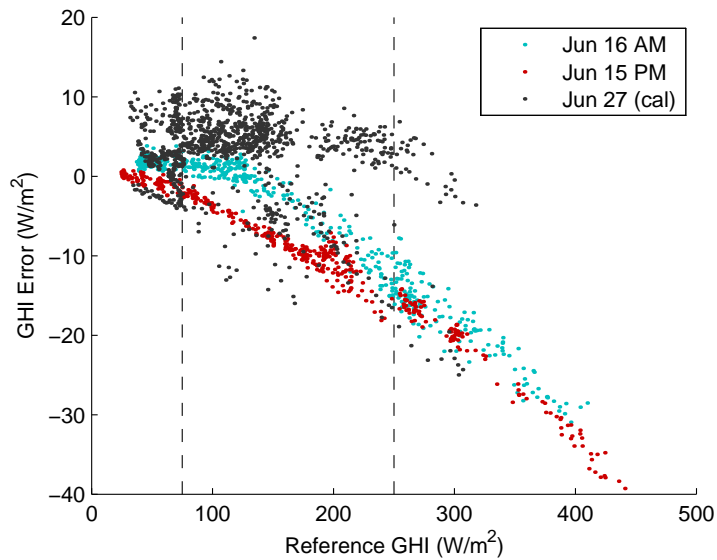
Figure 2.11: Error as a function of GHI under low-DNI conditions, where GHI ≈ DHI. Data shown are for the ANN calibration. Dashed vertical lines indicate the limits used for DHI calibration on June 27 and two other days (Figure 2.8).

this by altering the balance between the channels required unphysical combinations with negative weights on at least one channel, and were still unable to remove trends entirely. Some additional spectral mismatch is expected due to the near-zero sensitivity of the USI in the region from 700 to 1100 nm where the reference sensor has significant response (see Figure 2.1). Additionally, previous work has suggested nonlinearity on the order of 5% in camera sensitivity, which would lead to errors in DHI measurement without better calibration. Despite these issues, the vast majority (95%) of these low-DNI periods have errors either $< 15\%$ or < 15 W/m^2.

To validate the DNI performance of the camera, Figure 2.12 compares the DNI measured by the camera and the SPN1 sensor located 1 km inland. Due to the spatial separation between the sites and the proximity to the ocean, the sites are not expected to experience matching DNI in the presence of clouds. To focus on data points where spatially homogeneous conditions prevailed, a reduced set of data from clear periods between April 15 and September 22, 2015 is shown. Because images on these days
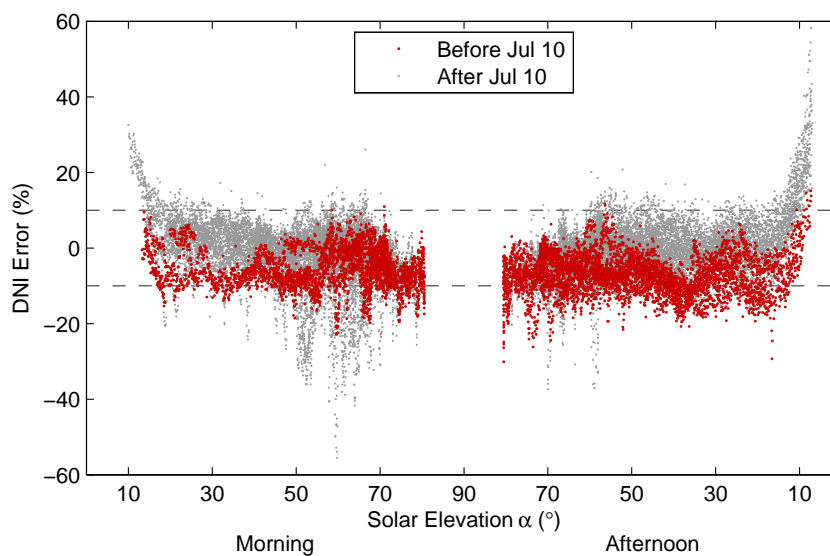
Figure 2.12: DNI validation of the USI against SPN1 measurements under clear skies. July 10 was the last full clear day used for calibrating DNI measurements, so measurements after that exhibit larger errors, particularly when the sun is low in the sky. DNI errors: RMSE of 66.0 W/m^2 (7.7%), MBE of -26.0 W/m^2 (-3.0%). DHI errors: RMSE of 45.4 W/m^2 (53%), MBE of 41.7 W/m^2 (49%). GHI errors: RMSE of 51.8 W/m^2 (7.1%), MBE of 0.2 W/m^2 (0.02%).

during these periods were visually observed to be completely clear, both sites should be cloud-free, so that DNI should be comparable. While the measurements generally agree to within 10%, the USI produces lower measurements on average, except at very low solar elevations where a strong positive bias is present. Additionally, there are several dozen instances of very localized larger negative bias, which were found to correspond to the sun passing behind scratches or dirt on the camera dome. While some of the overall negative bias is likely due to calibration errors in the camera, other researchers [2] have also observed a positive bias between 4.8 and 14% in DNI measurements from SPN1 instruments due to the min/max procedure used to separate diffuse and direct irradiance. Although it would be possible to calibrate DNI against the SPN1, we are hesitant to do so given its reputation for overestimation of DNI.

2.5   Conclusions

This paper details the use of an all-sky camera for measuring DNI and DHI from which GHI can be computed. While measuring DHI is relatively trivial by integrating over radiance across the sky hemisphere, the method for measuring DNI is novel and has not been presented before in the literature to our knowledge. Our results confirm that the method is generally effective, albeit with lower accuracy than can be achieved with well-calibrated conventional sensors. However, compared to comparable technologies, accuracy is actually improved. For example, Schmidt et al. [54] report GHI RMSE of nearly 25% (10-second instantaneous samples) for their image-based method, while Gauchet et al. [14] managed RMSE of 17% for five-minute average data using cloud position information. Both relative errors are normalized by average daytime irradiance. However, as errors are generally lower under clear conditions, the comparison of data from different sites does not constitute conclusive evidence as to which method is superior. Different spectral sensitivity of the reference instrumentation further complicates comparisons. These other algorithms generally suffer from limited information about DNI due to saturation in the solar region, and as a result are unable to follow some of the large fluctuations that originate primarily in the direct beam; their performance improves considerably with time averaging (e.g. the algorithm by Schmidt et al. achieves RMSE of 11.7% on hourly-averaged data). Meanwhile, 1-hour satellite-derived GHI data experiences RMSE from 12% to 25% and more, depending on the location [45].

Furthermore, in many cases, the ability to measure radiation using a camera may be useful where only a camera is present. For example, it can simplify data acquisition for a sky-imager based solar forecast [70], since it alleviates the need for a separate GHI data source to assign cloud optical depth. Separating DNI and DHI is also of use when forecasting output from tilted systems, and for determining cloud thickness near the

sun where pixels are generally too saturated to contain useful information. Additionally, the ability of a camera to capture the directional dependence of radiance may prove a valuable tool for developing non-isotropic diffuse radiation [7] and transposition [16] models.

The primary source of error in our methodology is expected to stem from the poorly understood relationship between CCD smear intensity and DNI. Although we have derived means of empirically calibrating away these irregularities as a function of pixel position and solar elevation, their exact causes are unknown and it is therefore unlikely that our empirical method captures all underlying relationships and parameters. Furthermore, some observational evidence suggests that smear is nonlinear in intensity (distinct from sensitivity varying spatially), or that smear may not account for all the irradiance typically considered DNI. Indeed, DNI is typically defined to include everything within $2.5°$ of the sun, but the smear stripe only results from very bright pixels, likely only within the solar disk itself (closer to $0.5°$). If the ratio of brightness within the solar disc to brightness within the standard DNI cone remains constant, then this discrepancy would be accounted for by calibration; however, if the ratio changes as atmospheric conditions change, it would mean that smear stripe alone does not contain enough information to measure DNI. Due to the difficulty of determining the proper ratio in which to combine the two signals, we have elected not to consider the remainder of the solar region or any nonlinearity of smear at this time.

Additional sources of error include variations in the transmissivity of the camera dome due to dust, dirt, and scratches on the outside of the camera dome, but these are expected to primarily result in random errors. Dust and dirt affect all radiation measurement systems, and present special challenges because they change on a daily basis. We have anecdotally observed the camera-measured DNI increasing by several percent on a day when the optics were cleaned. Temperature also affects the camera

sensor, but it has a relatively small influence on the value of "normal" pixels, though there is no guarantee this applies to smear.

Although the calibration procedures developed here are general, i.e. they could be applied to any compatible camera, the overall technique places several restrictive requirements on the camera system. In order to capture sufficient dynamic range to measure GHI and DNI, multiple exposures with different integration times are required. In a system with limited dynamic range, allowing the exposure time to respond dynamically to the observed radiation intensity, and using smear from that single exposure rather than a minimal integration should still work; data quality would, however, be expected to be lower. Furthermore, use of dynamic exposure times complicates the conversion between camera measurements and radiance/irradiance. Finally, only certain sensor types actually exhibit smear at all. Indeed, CMOS sensors (which are increasingly common) use in-pixel read-out electronics and therefore are not subject to smear; in fact, CMOS sensors are preferred for sky imaging by some groups because they are immune to smear.

Potential improvements to the method presented in this paper revolve primarily around the DNI calibration procedure. In particular, it would be valuable to have a field calibration procedure that can eliminate the need to extrapolate DNI calibration factors to different solar zenith angles, and to determine with more certainty which properties of the various calibrations are potentially affected by redeployment or configuration changes of the instrument. Additionally, further automation of calibration procedures would be beneficial, particularly with respect to detection of clear-sky periods. Nevertheless, we expect the correlation between smear and DNI to be robust and applicable to all interline-transfer CCDs, and we expect that in the near future we will make use of this data in our solar forecast algorithm, if not in more demanding resource-assessment applications.

2.6   Supplemental Camera Details

2.6.1   Cropping

As mentioned above, hardware cropping is performed in the camera by skipping the A/D conversion step for whole rows. This means that as pixels are shifted off the image sensor, rows immediately adjacent to the sun will have less time to be "contaminated" by smear, resulting in a lower smear signal in these rows.

Empirically, it is observed that cropping the top of the image by N pixels affects the N rows immediately below the sun, while cropping the bottom of the image by M rows affects M rows somewhat above the sun in the following image. The distance above the sun is unfortunately somewhat dependent on time between image captures; in a test data set, it was approximately 130 pixels.

Our current recommendation is to remove the entire region affected by cropping (from 300 rows above to 200 rows below the sun for the USI), or perhaps, better yet, to disable hardware cropping for the exposure used to measure DNI and an extra frame before it, which can be discarded.

Astute readers will note that, since cropping the top of the image affects rows below the sun that this implies that the horizontal shift register is actually located at the top of the image sensor, which is contrary to the way the sensor is drawn in most literature!

2.6.2   Color Channel Differences of Smear

Since smear results from overflow of signal into interline transfer columns, we would expect the increase in pixel value due to smear to be uniform throughout each column of the image since each image column shares an interline transfer column. Specifically, while smear is expected to be higher in the red/green columns than in

the green/blue columns due to higher average filter transmissivity, we do not expect to observe any difference between the two different colors within a column. Although more light will cause smear at the location of the red pixel on the sensor, the signal which is altered could have originated in either a red or a green pixel.

Observed smear signals, in contrast, are highest in red pixels, followed by green pixels in the red column. Blue pixels and green pixels in the blue column exhibit lowest smear, and are both about the same, as is expected. We are unsure as to the cause of this, but would welcome proposals, as they might shed further light on the underlying physics, which might in turn help us to improve our calibration procedures and accuracy.

2.7 Calibration Periods

Period 1 begins at the start of the study period. The starts of periods 2 and 3 each coincide with attempts to improve image focus. While at least some of the changes in calibration are due to the focus change directly, it is anticipated that accidental changes in alignment between the dome, lens, and camera body could require re-calibration any time the device is opened.

The first two and last two days in each "DNI Calibration" set were placed in the validation set for the ANN, while the other listed days were used as training data.

Table 2.1: Calibration Periods

	Period 1	Period 2	Period 3
Start [UTC]	2014-08-18 08:00:00	2014-12-05 19:54:00	2015-03-16 19:10:00
End [UTC]	2014-12-05 19:45:00	2015-03-16 18:40:00	2015-08-18 08:00:00
DNI Calibration	Aug 18, 27, 28; Oct 02, 05, 29; Nov 04, 06, 08, 26, 28	Dec 09, 22, 23, 29; Jan 02, 06, 17, 24; Feb 07, 24; Mar 05, 07, 12	Mar 25, 26, 27; Apr 12, 16, 18, 28; Jun 23, 24; Jul 10
DHI Calibration	Oct 7	Jan 20	Jun 27

2.8 Error Metrics

Error rates are reported in RMSE (root-mean-square error) and MBE (mean bias error), following their standard definitions:

$$\text{RMSE} = \sqrt{\frac{\sum (\text{GHI} - \text{GHI}_{\text{ref}})^2}{N}}, \qquad \text{MBE} = \frac{\sum \text{GHI} - \text{GHI}_{\text{ref}}}{N}, \qquad (2.9)$$

where GHI is the sky imager, and GHI_{ref} is the reference pyranometer measurement, or analogously for DNI or DHI. When RMSE (or MBE) is reported in percent, it has been normalized by the average value of the reference sensor during the same time period:

$$\frac{\text{RMSE}}{\frac{\sum \text{GHI}_{\text{ref}}}{N}}. \qquad (2.10)$$

In addition, in some places we report errors for hourly-averaged data. In order to avoid biasing the results for or against hours near sunrise/sunset (which are much more likely than others to have partial data due to horizon exclusion rules), hours in which 45 or more (of 120 possible when sampling every 30 seconds) data points were missing or excluded from were omitted from the hourly statistics.

2.9 Acknowledgements

Chapter 2, in full, is a reprint of the material as it appears in "Measuring diffuse, direct, and global irradiance using a sky imager" (B. Kurtz and J. Kleissl) in *Solar Energy*,

2016. The dissertation author was the primary investigator and author of this paper.

Chapter 3

A Virtual Sky Imager Testbed for Solar Energy Forecasting

3.1   Introduction

In recent years, whole-sky imagers have become popular for forecasting solar energy availability on short time horizons [70, 14, 53, 6, 43]. However, validation of these forecasts can be tricky; reference data is often limited to at most a few irradiance sensors, and even in the case where many sensors are present over a large area, detailed validation data on the cloud field itself is uniformly unavailable. Under these circumstances, validation can determine the forecast accuracy, but apportionment of the forecast error to different components of the algorithm is difficult due to the lack of data about the actual state of the atmosphere and the resulting radiation field. Therefore prioritization of forecast development work is usually not well-informed and is unable to follow cost-benefit principles.

We propose to address some of these limitations by producing a virtual sky imager testbed, in which the configuration of the clouds and resulting irradiance is

known. The purpose of this paper is to describe the setup of the virtual testbed and briefly illustrate its potential through a case study. The virtual testbed is used to design and test improvements to whole-sky imager forecast methodology developed at UC San Diego, but it is straightforward to adapt it to any other algorithm.

Simulating clouds is one of the grand challenges of atmospheric physics as it includes scales from micrometers (cloud condensation nuclei) to kilometers (cloud size), multiple phases (vapor, liquid, ice), and even chemistry (hydrophobicity of aerosol species). In terms of short-term (order of 10 minutes) cloud dynamics that are most relevant to sky imager solar forecasting, the multi-scale and multi-phase fluid dynamics need to be represented. In particular atmospheric turbulence plays a critical role in cloud formation (e.g. thermals) and cloud dynamics. Not only do clouds "live" in the turbulent atmospheric boundary layer flow field, but they also generate their own turbulence due to longwave radiative cooling at the cloud top and latent heat release. Large Eddy Simulation (LES) is a uniquely suited tool to simulate these boundary layer and cloud dynamics. In LES the large turbulent eddies that are responsible for most of the momentum, heat, and moisture transport are explicitly resolved and simulated faithfully based on the Navier Stokes equations. The small scales (less than about 10 meters) cannot be resolved due to computational cost and are parameterized through subfilter scale models [33]. LES also simulates all modes of heat transfer, water vapor transport and phase change, as well as cloud microphysics. LES is a mature field in engineering and atmospheric science and the resolution, subfilter scale models, and microphysics models have been continually improved over the past decades [35, 61].

Virtual cloud fields will be produced using LES. Surface-level irradiance fields and simulated whole-sky images will be derived from a 3-dimensional radiative transfer model (3D RTM). These tools (LES and 3D RTM) are significantly more physically grounded and accurate than current sky imager forecast algorithms, so there is con-

siderable scope for improving sky imager forecasts based on the virtual testbed. It is worth noting that the virtual testbed need not reproduce a given observed cloud field for this to be useful, so long as the virtual clouds behave similarly to real clouds. Why not just use the LES and 3D RTM for forecasting in the first place? First, while recent GPU-accelerated LES codes [51] approach the speeds necessary to produce operational forecasts, the computational requirements for LES and 3D RTM tools are currently too large to be feasible for short-time-horizon forecasting. Furthermore, even in those cases where LES has been run operationally on a wide variety of measured data [18, 37], the cloud fields are statistically accurate on timescales from tens of minutes to hours. To produce meaningful forecasts of individual clouds, LES would require input of a detailed state of the atmosphere including detailed humidity and velocity fields which, as noted, are generally unavailable. Even here, the virtual testbed is useful, as it allows improved testing of 3D cloud detection algorithms for whole-sky imagers, which could eventually be used as input to an LES-based forecast.

In section 3.2, we present the virtual testbed and whole-sky imager forecast. Section 3.3 compares the results of the sky imager forecast to those of the virtual testbed, paying special attention to the newfound ability to determine errors of difficult-to-measure quantities such as wind speed aloft and 3D cloud structure. Differing geometrical perspectives and cloud field dynamics constitute the largest sources of error in the current forecast, with geometry playing a larger role at short forecast horizons, and cloud evolution dominating the error for further-ahead forecasts. Discussion and conclusions are provided in Section 3.4.

3.2 Virtual Testbed Components

3.2.1 Large Eddy Simulation

LES are carried out using the UCLA LES [59, 58, 60], which has been thoroughly validated and tested for a number of cases including continental cumulus [5], raining cumulus [61], and stratocumulus clouds [58]. The UCLA LES uses the Smagorinsky sub-gridscale model, and parameterizes cloud microphysics following Stevens and Seifert [61]. Interactive radiation is implemented via a Monte Carlo version [46] of the delta-four-stream model [29]. Cloud droplet radius for both radiation and microphysics is modeled by assuming a fixed cloud droplet mixing ratio.

A single 14.5 hour simulation was carried out using example input data modeled for continental cumulus clouds, following the base case in [21], which is itself based on a detailed LES study of measurements taken at the Southern Great Plains (SGP) site of the Atmospheric Radiation Measurement (ARM) program [5]. Following prior simulations [21], precipitation was disabled in the microphysics model, leaving cloud liquid water diagnosed as the total water mixing ratio in excess of the saturation mixing ratio, and with the fixed cloud droplet mixing ratio of 70×10^6/kg. Initial profiles of atmospheric temperature and humidity, as well as input surface fluxes are shown in Figure 3.1. Small volumetric forcings are applied as in [5] in order to represent observed large-scale advection in the periodic simulation domain. This day represents typical formation of a convective boundary layer due to surface heating, with cumulus clouds forming at the top of the (initially clear) boundary layer. As the day progresses, the cloud base rises from 1000 m to around 1500 m, with maximum cloud thickness of around 1250 m. Both the boundary layer and the clouds continue to deepen until late afternoon when solar radiation has decreased significantly. Typical horizontal cloud size is 400 m. Hemispherical cloud cover peaks just above 65% around solar noon; Figure 3.6 later
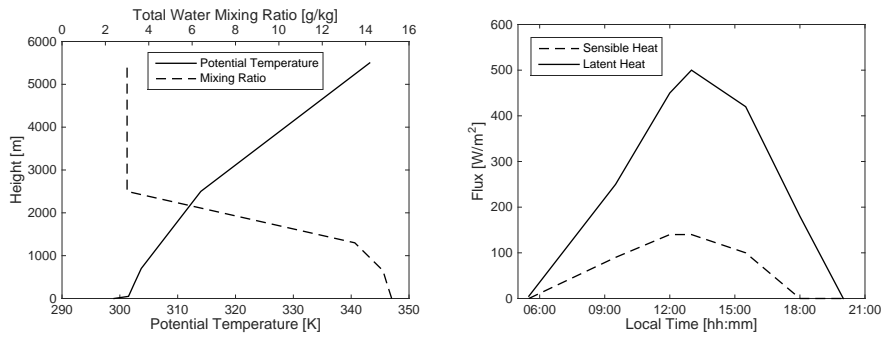
Figure 3.1: Profiles of temperature and humidity at simulation start, along with surface convective heat fluxes during the simulation.

shows hemispherical cloud cover over the course of the day.

LES grid cells are 50 meters across in both horizontal dimensions and 40 meters high, spanning a 6.4 km domain that is 5.1 km deep. Periodic boundary conditions are used in the horizontal dimensions. A 10-cell thick sponge layer is used at the top of the domain to prevent wave reflection, while the lower surface uses a no-slip boundary with roughness length of 0.035 m, representative of long grass.

LES requires on the order of an hour of simulation time to properly "spin-up" the turbulent flow and cloud field. After spin-up, the 3D state of the atmosphere (velocity, temperature, pressure, humidity, and liquid water content) is saved every 60 seconds of simulation time for input into the 3D RTM and reference against the sky imager forecast results.

3.2.2   3D Radiative Transfer Model

The Spherical Harmonic Discrete Ordinate Method (SHDOM) [13] is used to solve the 3D Radiative Transfer Equation. SHDOM is the most computationally intensive portion of the virtual testbed, requiring over half of the approximately 5000 CPU-core-hours used for the run presented here. SHDOM inputs are derived from the liquid water content output by UCLA LES, combined with the aerosol loading shown in Figure 3.2,
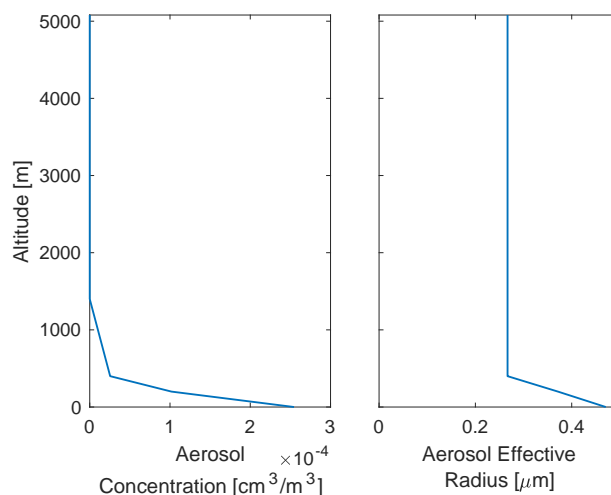
Figure 3.2: Aerosol loading and effective radius used to produce blue sky in SHDOM.

which is based on the `nauru19990707` data file included with SHDOM adjusted to match the observed annual-average aerosol concentration, and effective radius at the ARM SGP AERONET site in 2013. This rapid decrease in aerosol concentration with height matches the exponential decay proposed in [17]. SHDOM also uses atmospheric temperature when computing scattering properties; input vertical temperature profiles were derived from LES outputs. In order to simplify interpretation of the results, SHDOM is run with a constant sun position (solar zenith angle of $45°$) for the entire simulation time period; this avoids changing clear sky irradiance and geometric perspectives.

At each time step, SHDOM produces a map of surface global horizontal irradiance (GHI) across the simulation domain. In addition, it produces one or more simulated sky images (essentially a map of radiance versus direction at a single location) that can be fed into the sky imager forecast routines. SHDOM results at three different wavelengths (450 nm, 550 nm, and 670 nm) are combined to produce full-color images, and are averaged to approximate broadband GHI. As in the LES, periodic boundary conditions are used.

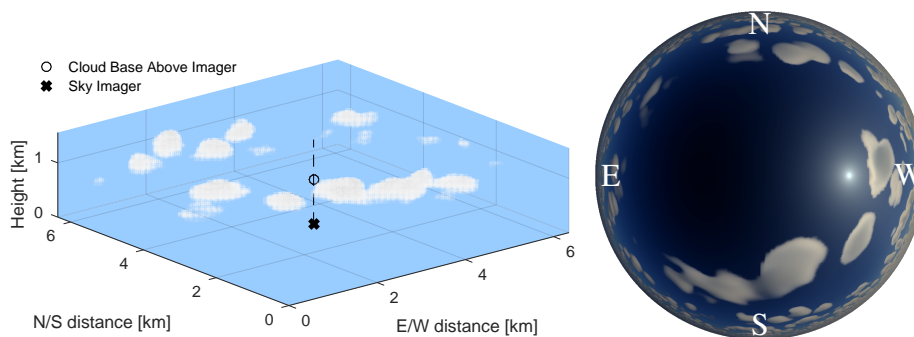Figure 3.3 shows an example of clouds from the LES and the corresponding

Figure 3.3: Example LES clouds and virtual sky image at 10:43 local time.

Table 3.1: Naming shorthands for modified versions of the forecast algorithm. The standard forecast is si-pix-kthist.

Cloud Map		Cloud Motion		Radiation	
Sky Image	(si)	Pixel Motion	(pix)	k_t Histogram	(kthist)
LES Converging Ray	(conv)	LES Layer Mean	(llm)	Per-class Mean	(ktmean)
LES Zenith Parallel Ray	(zen)			No Quantization	(noquant)
LES Sun Parallel Ray	(sun)			k_t Advection	(ktadv)

virtual sky image from SHDOM.

3.2.3   Sky Imager Forecast

The sky imager forecast [70] investigated here models clouds as occurring in a single plane at the height of the cloud base. Current cloud positions are detected based on the color of the input image, and future positions are forecast using the "frozen cloud advection" assumption, which assumes that the entire cloud field moves in a uniform direction without changing shape. Inputs to the sky imager forecast are a sky image, cloud base height (usually derived from lidar (Light Detection and Ranging) data), and recent measured GHI—used to estimate average cloud optical thickness, which is difficult to determine from the image. Figure 3.4 illustrates data flow through the sky imager forecast algorithm, along with inputs from the virtual testbed. In addition, several variations of the algorithm are discussed as part of the virtual testbed; naming conventions for these variations are given in Table 3.1.
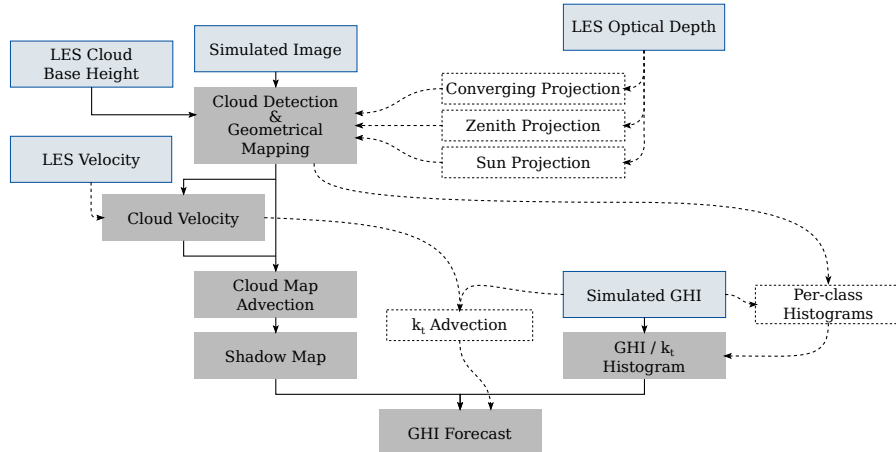
Figure 3.4: Data flow through sky imager forecast algorithms with inputs from virtual testbed. Solid arrows indicate the standard flow of data through the algorithm, while dashed lines show where "correct" data from the virtual testbed can be used in place of a step in the forecast algorithm. Outputs of LES or SHDOM are shown with a thin solid outline, while derived results have a dashed outline; steps in the basic sky imager forecast have no outline.

Cloud Detection and Geometrical Mapping

In the virtual sky imager testbed, cloud base height is determined based on the first grid cell to have significant liquid water content. As lidar point measurements of cloud base height are generally accurate, the "correct" LES-derived cloud height is used directly for forecasting. In practice, errors would be introduced in the process of interpolating point measurements of cloud height into an accurate height for an entire layer, particularly in the presence of topography or heterogeneous land surface and over larger areas. In the interest of brevity, we do not address these errors here.

Cloud detection operates on the virtual sky images in the same manner as real sky images, and classifies each pixel of the input image as clear sky, thin cloud, or thick cloud, by applying thresholds to the difference between the red-blue ratio (RBR) of the image being analyzed and RBR of a clear sky. Pixels with $\text{RBR} - \text{RBR}_{\text{clear}} \geq 0.4591$ are considered thick cloud, while those with $0.4591 > \text{RBR} - \text{RBR}_{\text{clear}}$ and $\text{RBR} - \text{RBR}_{\text{clear}} \cdot \text{HCF} \geq 0.3044$ are considered thin cloud. These thresholds generally
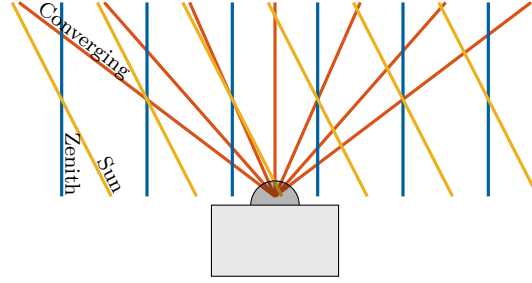
Figure 3.5: Ray geometry for cloud projections. All three methods are used for computing reference cloud maps from LES. Converging-ray projection is also implicitly used when producing cloud maps from sky images, and "sun projection" is used to map from the cloud plane to shadows on the ground.

vary with camera and location, and these values were manually selected specifically for use with the virtual testbed based on five images. HCF is the haze correction factor, and helps distinguish thin clouds from background haze. It is iteratively determined for each frame so that the average RBR of portions of the image detected as clear matches the RBR of the haze-corrected clear sky.

Reference cloud maps are derived from LES optical depth, with optical depths greater than 1.5 considered thick clouds and any smaller non-zero optical depth considered thin cloud. Optical depth is the integral of extinction coefficient μ along the rays of the projection, normalized by ray orientation.

$$\text{Optical Depth} = \int \mu \frac{dz}{ds} ds \tag{3.1}$$

$$\mu = \frac{3}{2} \frac{\text{LWC}}{\rho_l r_e}, \tag{3.2}$$

where LWC is the liquid water concentration in kg/m^3, ρ_l is the density of water, and r_e is the effective droplet radius, here fixed at 8 μm.

As there is no obviously "correct" way to compress a 3D cloud into a plane, reference optical depth maps are computed using three different geometries (illustrated in Figure 3.5): zenith projection, sun projection, and converging-rays projection. Zenith and

sun projected cloud maps compute the cloud optical depth along parallel rays, while the converging-ray projection computes cloud optical depth along rays emanating from the location of the camera. Because it uses the same projection function as the camera, the converging-ray projection is representative of the best results we can expect to achieve with a pixel-by-pixel cloud detection on a sky image, while the sun projection is most relevant to the actual irradiance received at ground level. The zenith projection is similar to the view from a satellite positioned directly overhead.

Cloud Velocity and Cloud Map Advection

The sky imager forecast computes cloud speeds based on pixel motion between adjacent frames. Motion vectors are determined for small regions of the image, and then clustered and averaged to produce a single wind vector that will be used to advect the entire cloud field. Assuming that clouds travel on the background flow, reference wind vectors can be obtained directly from the LES as the vector average wind at the cloud base height.

Shadow Mapping and GHI Forecast

The final step of the forecast is to place cloud shadows and estimate $GHI(x,y,t)$. The correct way to estimate surface GHI is to run a 3D RTM on a 3D field of extinction coefficients, which accounts for attenuation of the direct beam and 3D photon transport for diffuse radiation. Sky imager forecasts require simplifications both because 3D fields are not available and due to the computational complexity of 3DRTM. At present (kthist in Table 3.1), effects on direct and diffuse irradiance are lumped by assigning a clear-sky index k_t (fraction of clear-sky GHI that will be received) to each cloud class:

$$GHI(x,y,t) = GHI_{csk}(t) \times k_t(\text{cloud class}(x,y,t)) \qquad (3.3)$$

with cloud classes projected from the cloud plane to the ground using "sun projection" geometry from Figure 3.5. The k_t for each cloud class is selected by finding three peaks (modes) in the histogram of measured GHI data from the past 2 hours. If fewer than three peaks are found, defaults of 0.42, 0.70, or 1.06 (for thick, thin, and clear respectively) are used. "Correct" k_t for each class is determined by averaging the SHDOM GHI of pixels located in the shadows of each class.

In addition to reference GHI computed in SHDOM, we also compare several other radiation schemes, designed to illuminate the errors that arise in the existing forecast model. 1. Following the current sky imager forecast method, but using the "correct" k_t for each class (ktmean). 2. Converting directly from optical depth (Eq. 3.1, any projection) to k_t at each point via an exponential model fit at each time step, without quantizing into cloud classes (noquant). 3. k_t advection, i.e. $k_t(x, y, t) = k_t(x - ut, y - vt, 0)$, for clouds moving with velocity (u, v), without reference to detected clouds (ktadv). 4. Persistence, i.e. $k_t(x, y, t) = k_t(x, y, 0)$. Method (1) removes errors in the k_t assignment, while (2) removes errors due to quantization. Methods (3) and (4) are initially perfect, and are included primarily to illustrate model performance as the cloud field changes. We note that methods (2) and (3) require more detailed information about the cloud field than is generally available outside the virtual testbed.

Error Calculations

Comparison of each of these intermediate forecast quantities to the reference values can obviously be done directly, but it is also beneficial to compare the relative effects of errors at each step. For example, it is not clear how a cloud-speed error of 1 m/s relates to an error in cloud-cover of 10%. For this purpose, we compare the final forecast errors that result from substituting various reference values into subsequent forecast steps. For example, we might calculate forecast cloud positions and shadows using the "correct"

sun projection reference cloud map rather than the cloud map derived from the sky image (corresponding to sun-pix-kthist in Table 3.1). This and other varying paths through the forecast algorithm are drawn in Figure 3.4. Naming conventions for variations are summarized in Table 3.1.

Note that domain-average GHI is nearly constant over short periods of time, so errors are computed for all points, rather than for the domain average. Errors thus obtained are representative of validating sky imager forecasts against point measurements at weather stations. Forecasts for power plants exhibit reduced random error magnitudes due to spatial averaging. Forecasting and error reporting commence 15 minutes before the formation of the first clouds and extend through the end of the simulation.

When comparing the error E of different methods to a baseline case, it is also useful to define forecast skill,

$$\text{forecast skill} = 1 - \frac{E}{E_{\text{ref}}}, \tag{3.4}$$

which is small positive number (up to 1 for a perfect forecast) if a method performs better than the baseline, and a negative number if the method under consideration is worse.

3.3   Results and Discussion

3.3.1   Errors in Intermediate Quantities

Time series of cloud cover, cloud velocity, and k_t results are illustrated in Figure 3.6 and demonstrate the forecast's ability to match overall atmospheric conditions. During the simulation run, the sky imager forecast had errors (RMS) of 2.0 m/s and 1.7 degrees for the detected cloud velocities compared to LES wind at the cloud base height. Considering multiple cloud classes, 83% of pixels were correctly classified, with 7% that
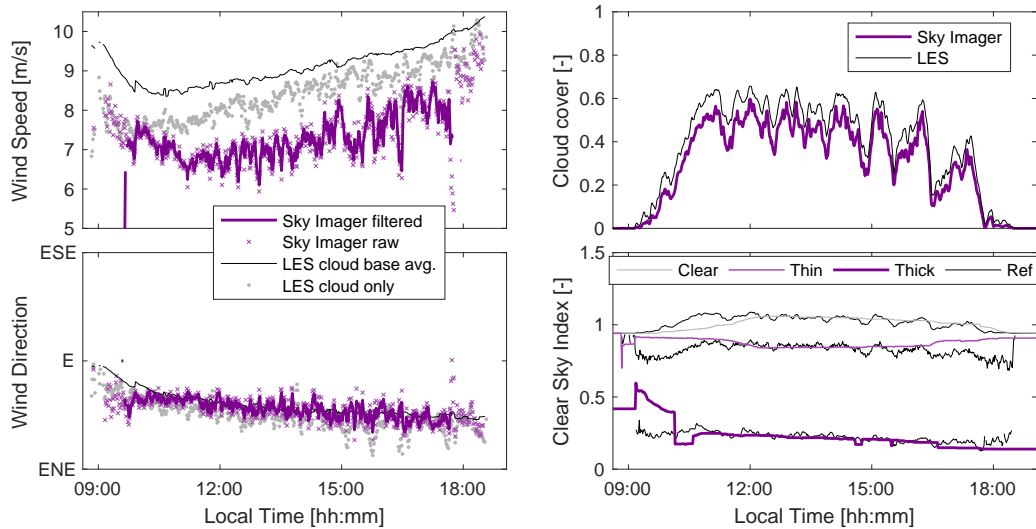
Figure 3.6: Time series comparison of cloud motion, hemispherical cloud cover, and per-class k_t against reference values derived from LES. Cloud motion filtering smooths data and removes points where cloud cover < 0.05. Clear-sky k_t exceeds 1 during much of the simulation due to cloud enhancement.

were classified as a cloud of the wrong class, and the remaining 10% classified as clear when they should have been cloud or vice versa. Detected k_t values from the existing histogram-based method were also relatively reliable, with errors (RMS) of 0.033, 0.078, and 0.079 for clear, thin, and thick categories.

Based purely on these error numbers, only the cloud speed error appears large enough to be of concern; the following sections consider the relative importance of these different errors to the GHI forecasts. Errors at short time horizons will mainly be influenced by the cloud mapping and radiation models, while longer forecasts rely significantly on the ability to predict the evolution of the cloud field.

3.3.2 Projection

Figure 3.7 illustrates the difference between the different cloud projection schemes. The standard sky imager forecast errors and persistence forecast errors follow the trend
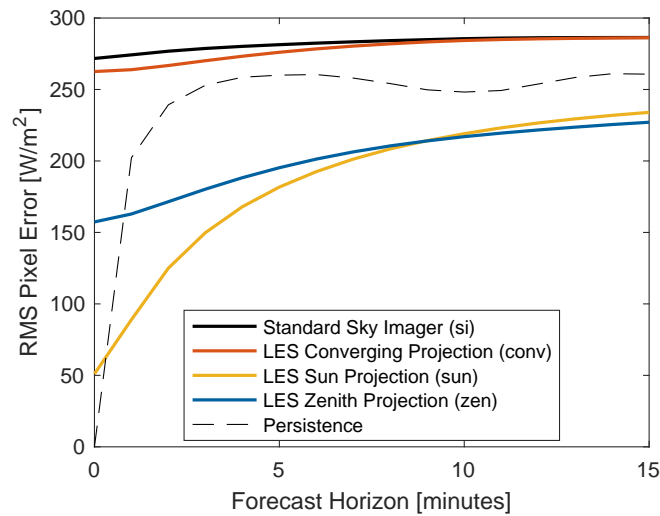
Figure 3.7: GHI forecast errors for several methods (*X*-pix-kthist) of mapping 3D clouds onto a horizontal plane at the cloud base height. All methods increase in error as the forecast horizon grows, however the methods with converging rays (Sky Imager and LES Converging) are unable to beat a persistence forecast. For reference, the typical range of GHI at any given time is around 670 W/m².

observed in previous work involving real-world data [70]. The converging-ray reference cloud map produces slightly better short term forecasts, but does no better at longer time horizons. Most notable, however, is the significant improvement that comes from using one of the parallel-ray projections, particularly at short time horizons. The sun projection method works best for short forecasts because it best matches the actual path light takes through the atmosphere, while zenith projection seems to work better at longer time horizons. We suspect this is because cumulus clouds form convectively, and as a result are more dynamic in the vertical dimension (which is hidden in the zenith projection) than the horizontal dimensions. Converging-ray projection was generally known (e.g. [25]) to cause some degree of perspective error, but the authors had not previously realized just how much of the error (over 2/3 at the shortest time horizons) was a result of this. The remaining error at zero time horizon ("nowcast") is due to cloud detection (thresholding of optical depth) and the complex 3D diffuse irradiance field that

is not captured by the k_t assignment; this error is further investigated in the following section. The inadequacy of the frozen cloud advection hypothesis and to a lesser extent, cloud speed errors (Figure 3.6), result in all the methods having larger errors at long time horizons.

3.3.3 Radiation

To investigate the remaining nowcast errors, we consider the radiation component of the forecast algorithm. The current algorithm makes two significant approximations. First, it treats GHI as depending only on the value in the 2D cloud map at a single point. This is accurate for the direct beam, but not at all representative of how diffuse irradiance propagates. Secondly, as a result of this single-point approximation and our quantized cloud map, the cloud shadows are also quantized. To assess the performance implications of these assumptions, the results of relaxing each of these assumptions are demonstrated in Figure 3.8. The sun projection is used for this comparison as it is most physically representative, and performs best (Figure 3.7) at short time horizons.

Nowcast errors are independent of cloud motion and therefore reveal the radiation model errors. Choosing the optimal (mean observed at zero horizon) k_t for each class (red line) results in modest (around 12%) improvements in the radiation model. However, even eliminating the quantization (blue line) leaves over 40% of the nowcast error. The remainder requires properly dealing with diffuse irradiance and 3D cloud structure.

At longer time horizons, the difference between the various methods decays as the advected static cloud map becomes less representative of the real cloud field. The k_t advection scheme uses the initial measured $k_t(x, y)$, and is thus perfect initially, but by 5 minutes is hardly any better than the standard algorithm. Interestingly, the mean k_t method actually performs better at long time horizons, presumably because localized fluctuations about the mean values tend to change more quickly with time and smoothing
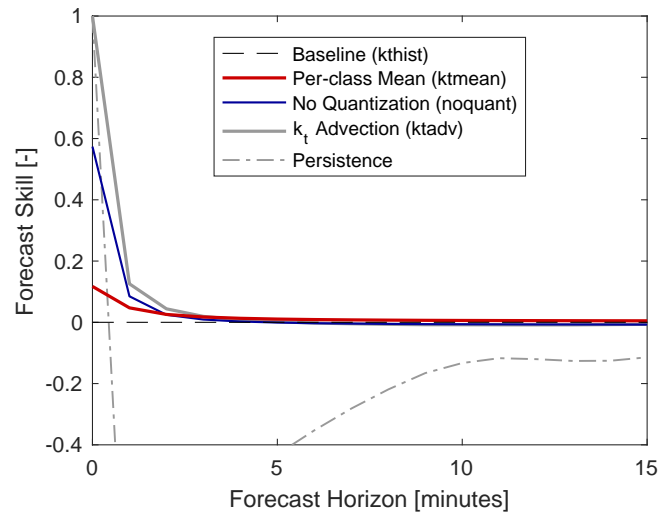
Figure 3.8: Forecast skill for various methods of modeling radiation. The baseline (reference) method is sun-pix-kthist, and other illustrated methods are sun-pix-*X*. The mean k_t observed for each class at the time of forecast gives best results for a quantized, single-point radiation model. Results are also shown for a single-point model without quantization and the full (3D with diffuse) radiation model run at the time of forecast. For comparison, a persistence forecast (constant k_t at each point) is also shown.

forecast fields therefore tends to reduce errors.

It should also be noted that this cloud scene contains only medium-thickness fair-weather cumulus clouds which probably tends to improve the performance of the baseline radiation model compared to conditions with a mix of thin and thick clouds. In particular, the algorithm would likely have more difficulty selecting the correct peaks from a more complicated k_t histogram.

3.3.4 Cloud Evolution

To address errors at longer forecast horizons, additional comparisons were run using the nominal average wind vector from LES. As illustrated in Figure 3.9, using the nominal wind vector from LES results in less than 4% improvement in forecast accuracy. For the sun and zenith projections, these improvements are relatively small (median 1.1% and 1.6% respectively across forecast horizons) in comparison to the overall increase
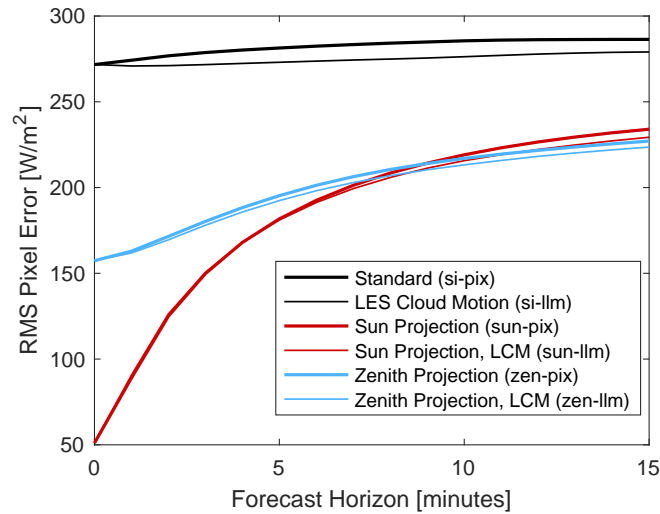
Figure 3.9: Forecast errors for frozen cloud advection compared with reference motion vectors from LES. Algorithm variations shown are X-Y-kthist. Similar behavior is observed for other forecast variants not shown.

in error with forecast horizon, suggesting that the current sky imager forecast's motion vector algorithm works well (at least, for this simple, one-layer cloud case), and that we have essentially saturated the capabilities of the frozen cloud advection model; further improvements would require a more dynamic model for cloud development. After a forecast horizon of 5 minutes, a forecast that assumes constant k_t thoughout the domain (not shown) outperforms all other forecast variants. Thus, 5 minutes can be considered to be the decorrelation time scale of this cloud field and an upper bound for the validity of the frozen cloud assumption; the decorrelation time scale is expected to vary with atmospheric conditions.

Some additional attention is required to the motion estimation algorithm as applied to the sky image or converging cloud map. In Figure 3.6 previously, a significant deviation was observed between the detected cloud speed and the LES reference speed—the pixel motion estimation consistently under-predicts speed. While the contribution to overall error is still always less than 4% (median 2.9%) in this case, approximately half of the forecast-horizon-dependent error is attributable to this velocity under-prediction.

This under-prediction appears to be related to the vertical geometry of the cloud, as Figure 3.10 shows that the detected speeds in the sun and zenith projections match the
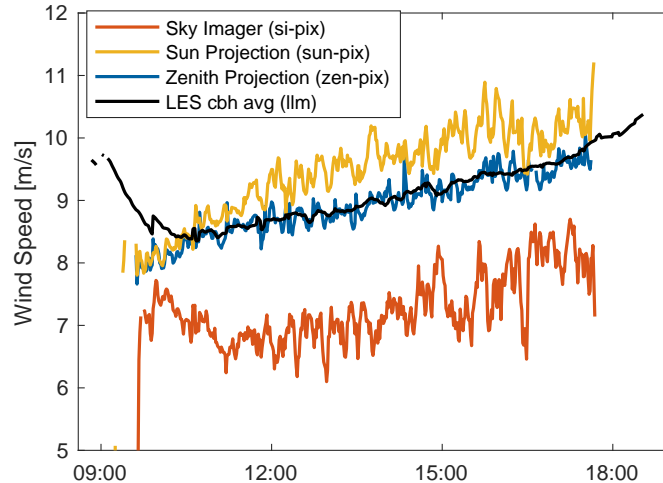
Figure 3.10: Cloud speed estimates based on pixel motion for the different projections in comparison to the LES reference speed. The black (llm) and red (si-pix) lines are also shown in Fig. 3.6.

LES results much more closely. Furthermore, experiments with non-physical clouds occupying only a single grid layer showed no issues with motion estimation, suggesting that cloud depth or wind shear is involved. At present, a complete explanation for this under-prediction of velocity is lacking; it will be investigated in more detail in future work. As noted above, the more accurate projections also yield more accurate motion estimates without additional work, so this investigation is primarily of interest until it becomes possible to generate 3D cloud maps from sky imagery.

3.4   Discussion and Conclusions

The virtual sky imager testbed is a valuable and versatile tool, allowing us to validate the quality of outputs from many steps of the sky imager forecast algorithm, and to assess the source of remaining errors. Here, the testbed demonstrated that for

a simple cloud scene with scattered cumulus clouds, nowcast errors already negated most of the utility of sky imager forecasting. Nowcast errors primarily originated in the converging-rays projection of 3D clouds into a 2D plane, while cloud detection contributed relatively minor errors.

Sky imager forecast errors further increase from the nowcast errors, never managing to outperform a persistence forecast. The virtual sky imager testbed allowed cloud motion estimation errors to be examined separately and these errors were found to be small except for converging-ray projections, and of minor consequence there. Further, the virtual sky imager testbed demonstrated that even with projection errors in the nowcast corrected, the frozen-cloud-advection assumption for forecasting future cloud positions increasingly deteriorates forecast accuracy at longer time horizons.

However, the virtual testbed suffers from a number of limitations as well. LES is mostly limited to boundary layer clouds over flat and homogeneous or at least idealized (periodic) ground surfaces. The current LES setup is therefore limited in its ability to produce high clouds, including cumulonimbus and cirrus, as well as multiple cloud layers and topographic clouds. In principle, use of a larger domain, non-idealized measured inputs, and advances in numerical codes can enable simulations of these other cloud types (e.g. as in [52, 18]), but with considerable computational and human resource investments. Varying types of clouds and topography would likely influence the measured errors quantitatively, but qualitative conclusions would likely be similar to those for cumulus clouds. For example, clouds with smaller vertical extent such as stratocumuli would likely reduce projection errors, but sun or zenith projection would still be expected to outperform converging-ray projection. Therefore, while not necessarily sufficient to validate forecasts under the variety of conditions seen in the real world, for development of generic forecast algorithms it is preferable to utilize simpler-to-implement, well-studied cases. Multiple cloud layers, on the other hand, considerably complicate cloud

detection (shadows of upper layers on lower layers), cloud mapping (single-cloud-plane model is no longer accurate), and motion estimation (distinguish multiple layers moving independently), and are therefore more likely to reveal qualitatively different results. In a future iteration of the virtual testbed, multiple cloud layers might be approximated by running multiple separate LES simulations and stacking the results, though this is obviously not physically realistic. Finally, the process of producing virtual sky images currently omits both stray light and sensor noise. Noise, and in particular stray light tend to cause issues with cloud detection, so cloud detection in the virtual testbed is likely more accurate than for real images. Models for noise and stray light could be added in a future version of the virtual sky imager testbed.

Despite these limitations, the virtual testbed is expected to be a valuable tool for validating and improving sky imager forecast algorithms. The authors would be happy to share the virtual sky images and ancillary data with other researchers.

3.5   Acknowledgements

Chapter 4

Cloud Advection

4.1 Motivation

Inspired by our work with the virtual sky imager testbed in Chapter 3, we know that the projection of 3D clouds onto a 2D hemisphere (the imaging surface of a fisheye lens) and the evolution of clouds over time are the two largest sources of error in the forecast. Other researchers have successfully demonstrated binary 3D cloud detection using multiple sky imagers [40], and another student in our lab is working to develop full 3D reconstruction techniques [32], so I will focus here on cloud evolution. Rather than struggle with the geometrical issues of real sky images, we will take advantage of the virtual testbed to use clouds in the "sun projection" cloud maps that yield best results for the virtual testbed (see Sections 3.2.3, 3.3.2). Although it is currently impossible to generate such cloud maps from real images, the aforementioned work in 3D cloud reconstruction should enable it in time. In addition, the techniques discussed in this section should be generally applicable to satellite imagery as well.

4.2   State of the Art

Most sky imager forecasts use either the frozen cloud advection assumption (a uniform, constant velocity field) [70, 43, 31] or a simple warping model (constant, non-uniform velocity field) [23]. Warping models typically act by taking the original grid of points for which cloud values are known and adjusting their locations so that the known cloud values lie at new (irregularly spaced) coordinates. An alternative approach often used in NWP models is to maintain a regular grid and adjust the values in the grid based on local gradients and wind. We will refer to this class of methods as "numerical advection" as they implement the standard advection-diffusion equation numerically (often with diffusion set to zero). We are not aware of any sky imager forecasts that also evolve the velocity field over time, nor any that use numerical advection or allow formation of clouds. At least one group has used numerical advection and non-constant velocity fields for forecasting with satellite images, by using the Weather Research and Forecasting (WRF) NWP model to perform their cloud advection [34], however they also derive their motion vectors from WRF rather than from the satellite image.

For motion detection, many authors track points, either on a grid [70] or by locating "feature" points with strong gradients [53]; others make use of various "optical flow" methods that generate a dense motion vector field across the entire image. For sky images, a variational method for optical flow (such as the Horn-Schunck method [22]) is typically used. These methods use a regularization parameter to encourage spatial uniformity of the resulting velocity fields, which is important for cloud scenes where much of the scene (i.e. the clear sky) may not appear to be moving at any given moment. More recently, a number of authors have found that variational methods that use a robust error function (one that reduces the error penalty for very large discontinuities in velocity) can help to work around natural edges in a scene [30, 4, 8].

Of note, these optical flow algorithms are typically derived from the computer vision field, where they are expected to give a relationship between two known frames and only those two known frames; typically if the relationships to a third frame are needed, they are simply calculated based on that additional frame. In contrast, for solar forecasting, we would ideally like to apply a given motion field to predict future cloud maps that are unavailable at the original time.

4.3   Proposed Methods

4.3.1   Warping

Flow Estimation

For a robust variational optical flow method, we turn to Liu [30], who provides both a methodology and a publicly available implementation in C++. His code uses successive over-relaxation and iteration over smoothed downscaled images to minimize an error term

$$E(u,v) = \int \psi\left(|I_0(\mathbf{p}-\mathbf{w}) - I_1(\mathbf{p})|^2\right) + \alpha\phi\left(|\nabla u|^2 + |\nabla v|^2\right) d\mathbf{p} \qquad (4.1)$$

Where I_0 and I_1 are two images, $\psi(x)$ and $\phi(x)$ are the robust function $\sqrt{x^2 + \varepsilon}$ and α is a regularization parameter.

Our group has previously dabbled in optical flow for cloud advection, with promising results [8]. However, in more recent work with the virtual testbed (discussed in Chapter 3), we found that the optical flow method was consistently outperformed by frozen cloud advection. This turns out to be due in part to the necessity of choosing an optimal value of the regularization parameter α. As can be seen in equation 4.1 above, for a given degree of smoothness in the flow field, the corresponding value of α actually

scales with the image brightness $I(\mathbf{p})$; that is, if we compute the optical flow for the contrast-adjusted images $I' = \gamma I$ (for some constant γ) and wish to obtain the same flow field as for the original image, we should use a regularization parameter $\alpha' = \gamma\alpha$. On the other hand, we can see that adding a constant to each image I would not change the required value of α, since the two frames are subtracted from each other.

Motivated by this, we tried several different approaches for automatically selecting α, including basing it on the mean value of I and on the standard deviation. At present, we have settled on scaling by the average magnitude of the image gradient

$$\alpha = 1.5\sqrt{\overline{|\nabla I_1|^2}}, \tag{4.2}$$

however we feel there is further room here to optimize. In particular, selection of the best choice of α is complicated because the error in forecast cloud states depends not only on the regularization parameter used to determine the velocity field, but also on the advection scheme used to apply the that velocity field to produce a future cloud state. The typical procedure of minimizing model error therefore involves optimizing over a large domain. Unlike with some other parts of the virtual testbed, it is not immediately clear that we can rely on the LES wind velocities for reference values, because the LES wind applies to the 3D clouds while the sky imager is working with 2D clouds.

Advection

Typically [8], cloud fields at points in the future are forecast by a "warping" operation. We shall refer to **forward** warping as warping that uses velocities at a given given grid position to project point positions forward in time

$$I_2(\mathbf{p} + \mathbf{w}) = I_1(\mathbf{p}), \tag{4.3}$$

while a **backward** warping model looking to find the value at a point **p** looks backward
along the velocity vector to predict a value from the previous time

$$I_2(\mathbf{p}) = I_1(\mathbf{p} - \mathbf{w}). \tag{4.4}$$

Forward warping feels intuitively more correct, because it aligns well with our mental
model of how the system works (backward warping appears to move cloud values using
velocity vectors from different points in the cloud map), but the mathematical structure
of backward warping bears more similarity to Equation 4.1. This bears some similarity
to the difference between explicit and implicit methods for numerical integration.

Additionally, we have considered separating the motion into a mean flow and
a perturbation, $\mathbf{w} = \overline{\mathbf{w}} + \mathbf{w}'$. The perturbation would then be applied first in order to
produce a distorted cloud map, followed by a simple linear advection of the clouds.
This distinction is unimportant when used with forward warping as it is mathematically
identical there (at least for single time steps), but tends to lead to improved results for
the backward model. Intuitively, this is because it results in the cloud type and motion
vectors being derived from closer points in the cloud map; often when backward warping
is done for the full velocity vectors, non-uniform velocities derived for the cloud are
often applied to empty spaces between clouds where they do little good. We have found it
productive to think of this as a **hybrid** warping scheme, because the distinction between
backward and forward warping is irrelevant for uniform motion, so this is somewhat like
using the backward scheme for distortion and the forward scheme for mean motion.

Both warping schemes will result in a warped coordinate grid and therefore
require interpolation to convert back to the regular grid used for storing cloud maps and
computing forecast errors. This interpolation is more efficient for the backward grid
because in that case we wish to sample from a regular grid at warped grid points; in

contrast, for the forward warping scheme, the known values lie on the warped grid, which is more difficult to interpolate from (e.g. MATLAB proposes to perform a triangulation on the warped grid, which can then be used to look up nearest neighbors of regular grid points more efficiently).

In our experience, the hybrid warping scheme produces results with the lowest errors, however it is not immediately apparent that it should outperform forward warping, and it remains a possibility that this feature is due as much to an incorrect motion field as to any superiority of the hybrid warping scheme.

Furthermore, there are two additional considerations for producing forecasts more than one time step in the future. Forecasts can either be produced directly

$$I_N(\mathbf{p}) = I_1(\mathbf{p} - (N-1)\mathbf{w}) \tag{4.5}$$

or iteratively

$$I_N(\mathbf{p}) = I_{N-1}(\mathbf{p} - \mathbf{w}). \tag{4.6}$$

If we are iterating, we must also consider whether the velocity field is to be distorted in each step, or only the cloud field.

Results for these different procedures will be discussed in more detail in Section 4.4 below.

4.4   Results and Discussion

After consideration of several of the options discussed above (Sec. 4.3.1), we settled on 1.5 as the best value for the proportionality constant in Equation 4.2, so that α is 1.5 times the root-mean-square magnitude of the image gradient. This was coupled with a hybrid warping scheme applied iteratively for subsequent time steps. Physically,
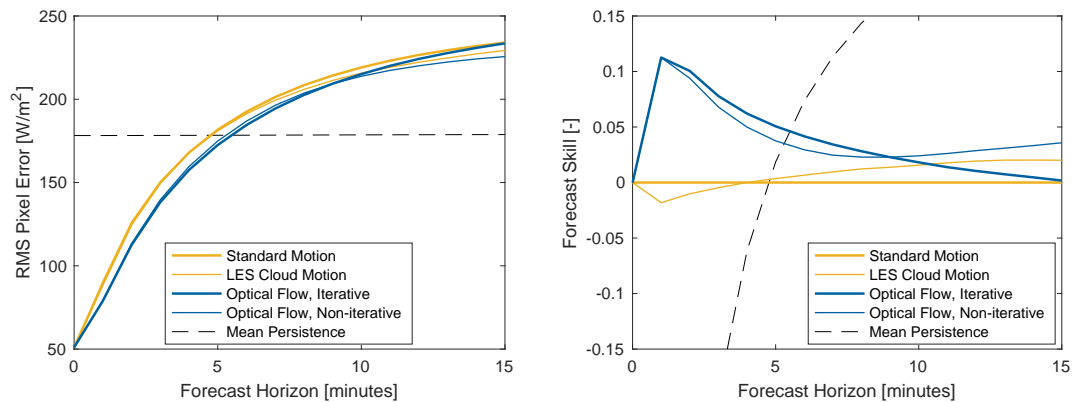
Figure 4.1: Forecast errors showing results of improved motion algorithm. Forecast skill is computed relative to the baseline sun-projection forecast, and results using the LES-derived cloud speed and for a mean-persistence forecast are also shown.

one would expect an iterative approach that also adjusts the velocity field to perform best. In practice, we found that results with adjusted velocity fields to be disappointing.

Results are shown in Figure 4.1. We found that while the iterative approach produces better short term forecasts, the resulting forecasts are actually worse by the 15-minute-ahead mark. In contrast, a non-iterative hybrid warping scheme did slightly worse in the first several minutes of the forecast, but performed better at the 15-minute mark. Although better long-term forecast performance is obviously desirable, we have elected to prefer the iterative approach at present because it is more intuitively correct, and because both forecasts have higher error than a mean persistence forecast after 6 minutes, at which point any operation that pulls the cloud field toward its mean value tends to reduce error.

Figure 4.2 shows forecast skill as a function of time through the day. Forecast skill is positive for the new method for most of the day except during times of strong cloud formation and dissipation at the beginning and end of the day respectively. Additionally periodic sharp drops in skill during the course of the day correspond to changes in cloud height in the model that were not properly accounted for during the first run. The model has been updated to address this issue, but the changes are not reflected in the plots above.
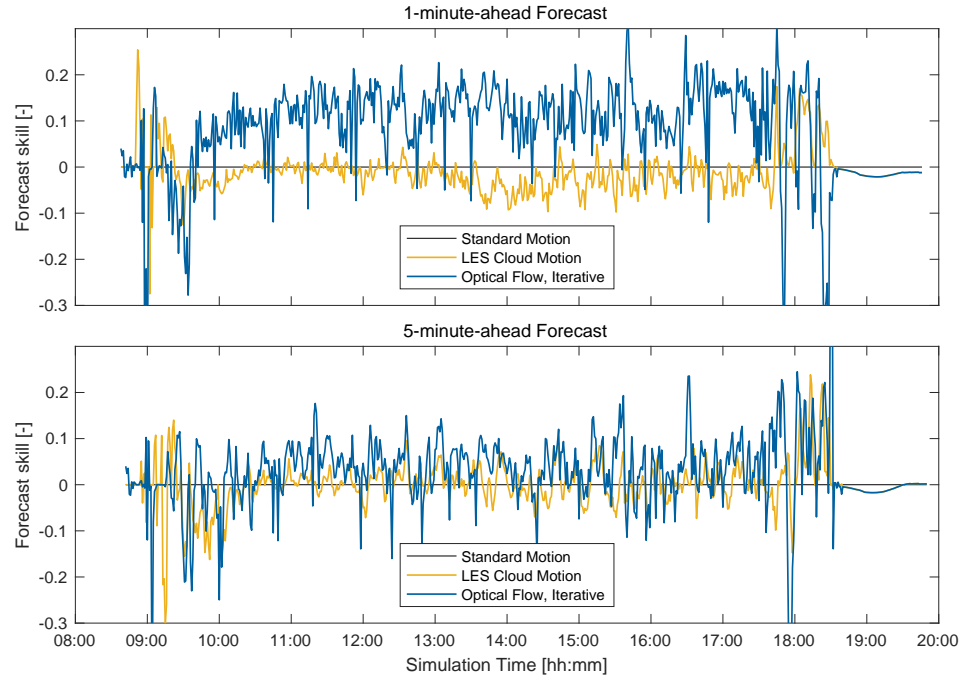
Figure 4.2: Forecast skill through the day for improved motion algorithm. Skill is computed relative to the sun-projection forecast using uniform motion computed via cross correlation.

4.4.1   Directions for Future Work

In the immediate future, a more in-depth investigation of parameters in the optical flow setup could produce further improved results. Specifically, it seems likely that a more exhaustive search of the regularization parameter (α) space would likely lead to a more robust rule for selecting that parameter. This search would necessarily also include a variety of different warping schemes as it's possible that poor determination of the flow field causes the otherwise-best warping scheme to perform poorly. We strongly suspect that a truly good warping scheme should also distort the motion field at each time step, and that our failure to observe promising results is due to either an implementation error or poor velocity fields.

In addition, we would very much like to see numerical advection tested for cloud forecasting. In particular, the framework of numerical advection nicely accommodates a

source term, which might particularly aid forecasting during times of strong formation and dissipation of clouds.

Finally, some effort will likely be required to extend this work to multiple cloud layers which may be moving in different directions. Other authors have suggested using detected cloud blocks to split the cloud into potential layers which can be assigned consistent motion [42], or initializing optical flow search based on cross-correlation motion tracking at strong gradients [68].

Chapter 5

Concluding Remarks

In the preceding chapters, several novel algorithms for sky imager forecasts are discussed.

Chapter 2 proposed algorithms for measuring DNI and GHI using the UC San Diego sky imager. Although other authors have proposed similar techniques, both our correction for scattered light and our technique for measuring DNI using CCD smear are unique; other authors typically neglect to account for stray light and either do not attempt to measure DNI or approximate it as a binary function.

In Chapter 3, we document our virtual sky imager testbed, which we then use to validate the individual steps of our sky imager forecast algorithm. The testbed shows that while both cloud detection and motion estimation work well by themselves, that geometrical limitations of a single sky imager limit the accuracy of initial cloud positions and the assumption of constant, uniform velocity limits our ability to predict cloud positions no matter how accurately the mean velocity may be. While these were previously known to be significant drawbacks of the forecast algorithm, we can now say with certainty that they account for the vast majority of the errors and should therefore be our primary targets for future algorithm development.

In keeping with that, Chapter 4 investigates two more advanced advection schemes inspired by optical flow techniques from computer vision. These techniques show moderate improvements in forecast error, but still leave considerable room for improvement.

5.1   Directions for Future Work

As touched on in Chapter 3, LES is a powerful tool, and with modern GPUs, it has become feasible to simulate a domain of the size used for sky imager forecasting in real time. Perhaps the holy grail of sky imager forecasting would be full 3-dimensional cloud detection followed by forecasting cloud positions with LES and radiation with SHDOM. GPU accelerated LES codes already exist [65]; SHDOM or another 3D radiative transfer model would need to be correspondingly accelerated. However, the most challenging aspect of such an endeavor would be to develop an algorithm for ingest sky imager data into the LES model; LES requires humidity and momentum initializations at every grid point, not just those with condensed water vapor that can be tracked with a sky imager. Perhaps simply repeatedly nudging the LES model towards sky imager observations would be sufficient to produce good results, but perhaps not.

No matter how accurate a forecast becomes, it will never be error-free. As important as low errors is the ability to tell forecast users the expected uncertainty of the forecast. For example, most solar forecasts are more accurate under clear skies than partly cloudy skies. However, there will be other important factors as well. If sky imager forecasts are ever to be useful operationally, they will need to come with uncertainty bounds, perhaps even adjusting these bounds in real-time. A few authors have proposed sky imager forecast methodologies with support for uncertainty forecasts [3]. Other promising approaches include quantile regression [36] or propagation of uncertainties with linkages determined using the virtual sky imager testbed.

Bibliography

[1] Allied Vision Technologies GmbH. *Allied Vision Prosilica GE Technical Manual*, March 2015. Accessed: 2016-04-11. <https://www.alliedvision.com/fileadmin/content/documents/products/cameras/Prosilica_GE/techman/Prosilica_GE_TechMan.pdf>.

[2] J. Badosa, J. Wood, P. Blanc, C. N. Long, L. Vuilleumier, D. Demengel, and M. Haeffelin. Solar irradiances measured using SPN1 radiometers: uncertainties and clues for development. *Atmospheric Measurement Techniques*, 7(12):4267–4283, 2014. doi:10.5194/amt-7-4267-2014.

[3] David Bernecker, Christian Riess, Elli Angelopoulou, and Joachim Hornegger. Continuous short-term irradiance forecasts using sky images. *Solar Energy*, 110:303–315, 2014. doi:10.1016/j.solener.2014.09.005.

[4] Michael J. Black and P. Anandan. The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. *Computer Vision and Image Understanding*, 63(1):75–104, 1996. doi:10.1006/cviu.1996.0006.

[5] A. R. Brown, R. T. Cederwall, A. Chlond, P. G. Duynkerke, J.-C. Golaz, M. Khairoutdinov, D. C. Lewellen, A. P. Lock, M. K. MacVean, C.-H. Moeng, R. A. J. Neggers, A. P. Siebesma, and B. Stevens. Large-eddy simulation of the diurnal cycle of shallow cumulus convection over land. *Quarterly Journal of the Royal Meteorological Society*, 128(582):1075–1093, 2002. doi:10.1256/003590002320373210.

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

[7] Rémi Chauvin, Julien Nou, Stéphane Thil, and Stéphane Grieu. Modelling the clear-sky intensity distribution using a sky imager. *Solar Energy*, 119(0):1–17, 2015. doi:10.1016/j.solener.2015.06.026.

[8] Chi Wai Chow, Serge Belongie, and Jan Kleissl. Cloud motion and stability estimation for intra-hour solar forecasting. *Solar Energy*, 115:645–655, May 2015. doi:10.1016/j.solener.2015.03.030.

[9] Chi Wai Chow, Bryan Urquhart, Matthew Lave, Anthony Dominguez, Jan Kleissl, Janet Shields, and Byron Washom. Intra-hour forecasting with a total sky imager at the UC san diego solar energy testbed. *Solar Energy*, 85(11):2881–2893, November 2011. doi:10.1016/j.solener.2011.08.025.

[10] Yinghao Chu, Mengying Li, and Carlos F.M. Coimbra. Sun-tracking imaging system for intra-hour dni forecasts. *Renewable Energy*, 96:792–799, 2016. doi: 10.1016/j.renene.2016.05.041.

[11] Adarsh Deepak and Richard R. Adams. Photography and photographic-photometry of the solar aureole. *Appl. Opt.*, 22(11):1646–1654, June 1983. doi:10.1364/AO.22. 001646.

[12] Eastman Kodak Company. *KODAK WRATTEN 2 Optical Neutral Density No. 96 Filter / 3.0*. Accessed: 2016-04-11. <http://motion.kodak.com/KodakGCG/ uploadedfiles/motion/Kodak/motion/Products/Lab_And_Post_Production/Kodak_ Filters/W96-3-0ND.pdf>.

[13] K. Franklin Evans. The spherical harmonics discrete ordinate method for three-dimensional atmospheric radiative transfer. *Journal of the Atmospheric Sciences*, 55(3):429–446, 1998. doi:10.1175/1520-0469(1998)055<0429:TSHDOM>2.0.CO; 2.

[14] Charlotte Gauchet, Philippe Blanc, Bella Espinar, Bruno Charbonnier, and Dominique Demengel. Surface solar irradiance estimation with low-cost fish-eye camera. In *Workshop on "Remote Sensing Measurements for Renewable Energy"*, Risoe, Denmark, May 2012. <https://hal-mines-paristech.archives-ouvertes.fr/ hal-00741620>.

[15] 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*, 5(11):2881–2892, 2012. doi:10.5194/amt-5-2881-2012.

[16] Christian A. Gueymard. Direct and indirect uncertainties in the prediction of tilted irradiance for solar engineering applications. *Solar Energy*, 83(3):432–444, 2009. doi:10.1016/j.solener.2008.11.004.

[17] Christian A. Gueymard and Didier Thevenard. Monthly average clear-sky broadband irradiance database for worldwide solar heat gain and building cooling load calculations. *Solar Energy*, 83(11):1998–2018, 2009. doi:10.1016/j.solener.2009. 07.011.

[18] William I. Gustafson, Andrew M. Vogelmann, Xiaoping Cheng, Satoshi Endo, Bhargavi Krishna, Zhijin Li, Tami Toto, and Heng Xiao. Large-eddy simulation (LES) ARM symbiotic simulation and observation, 2016. doi:10.5439/1256454.

[19] A. Habte, S. Wilcox, and T. Stoffel. Evaluation of radiometers deployed at the national renewable energy laboratory's solar radiation research laboratory. Technical Report NREL/TP-5D00-60896, NREL, 15013 Denver West Parkway, Golden, CO 80401, February 2014. <http://www.nrel.gov/docs/fy14osti/60896.pdf>.

[20] Sue Ellen Haupt, Branko Kosovic, Tara L. Jensen, Jared Lee, Pedro Jimenez Munoz, Jeffrey K. Lazo, James R. Cowie, Tyler McCandless, Julia M. Pearson, Gerry M. Wiener, Stefano Alessandrini, Luca Delle Monache, Dantong Yu, Zhenzhou Peng, Dong Huang, John Heiser, Shinjae Yoo, Paul Kalb, Steven Miller, Matthew Rogers, and Laura Hinkleman. The Sun4Cast® solar power forecasting system: The result of the public-private-academic partnership to advance solar power forecasting. Technical Report NCAR/TN-526+STR, NCAR, 2016. doi:10.5065/D6N58JR2.

[21] Laura M. Hinkelman, Bjorn Stevens, and K. Franklin Evans. A large-eddy simulation study of anisotropy in fair-weather cumulus cloud fields. *Journal of the Atmospheric Sciences*, 62(7):2155–2171, July 2005. doi:10.1175/JAS3463.1.

[22] Berthold K.P. Horn and Brian G. Schunck. Determining optical flow. *Artificial Intelligence*, 17(1):185–203, 1981. doi:10.1016/0004-3702(81)90024-2.

[23] H. Huang, J. Xu, Z. Peng, S. Yoo, D. Yu, D. Huang, and H. Qin. Cloud motion estimation for short term solar irradiation prediction. In *2013 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 696–701, October 2013. doi:10.1109/SmartGridComm.2013.6688040.

[24] Jens Kaluza and Andreas Neumann. Measurement of solar radiation with CCD-cameras: Influence of the spectral characteristic. In *Solar Engineering 1998*, Proceedings of the International Solar Energy Conference, pages 425–428, Albuquerque, New Mexico, June 14-17, 1998. American Society of Mechanical Engineers.

[25] Evgueni Kassianov, Charles N. Long, and Mikhail Ovtchinnikov. Cloud sky cover versus cloud fraction: Whole-sky simulations and observations. *Journal of Applied Meteorology*, 44(1):86–98, 2005. doi:10.1175/JAM-2184.1.

[26] P. Kuhn, S. Wilbert, C. Prahl, D. Schüler, T. Haase, T. Hirsch, M. Wittmann, L. Ramirez, L. Zarzalejo, A. Meyer, L. Vuilleumier, P. Blanc, and R. Pitz-Paal. Shadow camera system for the generation of solar irradiance maps. *Solar Energy*, 157:157–170, 2017. doi:10.1016/j.solener.2017.05.074.

[27] Matthew Lave, Jan Kleissl, and Ery Arias-Castro. High-frequency irradiance fluctuations and geographic smoothing. *Solar Energy*, 86(8):2190–2199, 2012. Progress in Solar Energy 3. doi:10.1016/j.solener.2011.06.031.

[28] LI-COR, Inc. *LI-200SA Pyranometer Sensor*. Accessed: 2016-04-11. <https://www.licor.com/env/pdf/light/200.pdf>.

[29] Kuo-Nan Liou, Qiang Fu, and Thomas P. Ackerman. A simple formulation of the delta-four-stream approximation for radiative transfer parameterizations. *Journal of the Atmospheric Sciences*, 45(13):1940–1948, 1988. doi:10.1175/1520-0469(1988)045<1940:ASFOTD>2.0.CO;2.

[30] Ce Liu. *Beyond Pixels: Exploring New Representations and Applications for Motion Analysis*. PhD thesis, Massachusetts Institute of Technology, May 2009. <http://people.csail.mit.edu/celiu/OpticalFlow/>.

[31] Ricardo Marquez and Carlos F.M. Coimbra. Intra-hour DNI forecasting based on cloud tracking image analysis. *Solar Energy*, 91:327–336, 2013. doi:10.1016/j.solener.2012.09.018.

[32] Felipe Mejia, Ben Kurtz, and Jan Kleissl. Cloud tomography applied to sky images: Part I: A virtual testbed. *In preparation for submission to Solar Energy*, 2017.

[33] Charles Meneveau and Joseph Katz. Scale-invariance and turbulence models for large-eddy simulation. *Annual Review of Fluid Mechanics*, 32(1):1–32, 2000. doi:10.1146/annurev.fluid.32.1.1.

[34] S. D. Miller, M. A. Rogers, J. M. Haynes, and M. Sengupta. A satellite-initialized model-advected scheme for short-term solar energy forecasting. *In preparation for submission to Solar Forecasting*.

[35] Chin-Hoh Moeng. A large-eddy-simulation model for the study of planetary boundary-layer turbulence. *Journal of the Atmospheric Sciences*, 41(13):2052–2062, 1984. doi:10.1175/1520-0469(1984)041<2052:ALESMF>2.0.CO;2.

[36] Jan Kloppenborg Møller, Henrik Aalborg Nielsen, and Henrik Madsen. Time-adaptive quantile regression. *Computational Statistics & Data Analysis*, 52(3):1292–1303, 2008. doi:10.1016/j.csda.2007.06.027.

[37] R. A. J. Neggers, A. P. Siebesma, and T. Heus. Continuous single-column model evaluation at a permanent meteorological supersite. *Bulletin of the American Meteorological Society*, 93(9):1389–1400, 2012. doi:10.1175/BAMS-D-11-00162.1.

[38] Andreas Neumann, Andreas Witzke, Scott A. Jones, and Gregor Schmitt. Representative terrestrial solar brightness profiles. *Journal of Solar Energy Engineering*, 124(2):198–204, 2002. doi:10.1115/1.1464880.

[39] Andu Nguyen, Maxime Velay, Jens Schoene, Vadim Zheglov, Ben Kurtz, Keenan Murray, Bill Torre, and Jan Kleissl. High PV penetration impacts on five local distribution networks using high resolution solar resource assessment with sky imager

and quasi-steady state distribution system simulations. *Solar Energy*, 132:221–235, July 2016. doi:10.1016/j.solener.2016.03.019.

[40] Daniel Oberländer, Christoph Prahl, Stefan Wilbert, Sebastian Müller, Badrudin Stanicki, and Natalie Hanrieder. Cloud shadow maps from whole sky imagers and voxel carving. In *ICEM*, Boulder, June 2015. <http://elib.dlr.de/97913/>.

[41] ON Semiconductor. *Interline CCD Smear*, 2 edition, September 2014. Accessed: 2016-04-11. <http://www.onsemi.com/pub_link/Collateral/AND9184-D.PDF>.

[42] Zhenzhou Peng, Dantong Yu, Dong Huang, John Heiser, and Paul Kalb. A hybrid approach to estimate the complex motions of clouds in sky images. *Solar Energy*, 138:10–25, 2016. doi:10.1016/j.solener.2016.09.002.

[43] Zhenzhou Peng, Dantong Yu, Dong Huang, John Heiser, Shinjae Yoo, and Paul Kalb. 3D cloud detection and tracking system for solar forecast using multiple sky imagers. *Solar Energy*, 118:496–519, 2015. doi:10.1016/j.solener.2015.05.037.

[44] R. Perez, P. Ineichen, E. Maxwell, R. Seals, and Zelenka A. Dynamic global-to-direct irradiance conversion models. *ASHRAE Transactions - Research Series*, 98(1):354–369, 1992.

[45] Richard Perez, Sergey Kivalov, James Schlemmer, Karl Hemker Jr., David Renné, and Thomas E. Hoff. Validation of short and medium term operational solar radiation forecasts in the US. *Solar Energy*, 84(12):2161–2172, 2010. doi:10.1016/j.solener.2010.08.014.

[46] Robert Pincus and Bjorn Stevens. Monte carlo spectral integration: a consistent approximation for radiative transfer in large eddy simulations. *Journal of Advances in Modeling Earth Systems*, 1(2), 2009. 1. doi:10.3894/JAMES.2009.1.1.

[47] PV LIB Toolbox. Version 1.2. <https://pvpmc.sandia.gov/resources-for-members/pv_lib-toolbox/>.

[48] J.F. Randall and J. Jacot. Is AM1.5 applicable in practice? modelling eight photovoltaic materials with respect to light intensity and two spectra. *Renewable Energy*, 28(12):1851–1864, 2003. doi:10.1016/S0960-1481(03)00068-5.

[49] R. Román, M. Antón, A. Cazorla, A. de Miguel, F. J. Olmo, J. Bilbao, and L. Alados-Arboledas. Calibration of an all-sky camera for obtaining sky radiance at three wavelengths. *Atmospheric Measurement Techniques*, 5(8):2013–2024, 2012. doi:10.5194/amt-5-2013-2012.

[50] Elton G. Rossini and Arno Krenzinger. Maps of sky relative radiance and luminance distributions acquired with a monochromatic CCD camera. *Solar Energy*, 81(11):1323–1332, 2007. doi:10.1016/j.solener.2007.06.013.

[51] Jerôme Schalkwijk, Eric J. Griffith, Frits H. Post, and Harm J. J. Jonker. High-performance simulations of turbulent clouds on a desktop pc: Exploiting the gpu. *Bulletin of the American Meteorological Society*, 93(3):307–314, 2012. doi:10.1175/BAMS-D-11-00059.1.

[52] Jerôme Schalkwijk, Harmen J. J. Jonker, A. Pier Siebesma, and Fred C. Bosveld. A year-long large-eddy simulation of the weather over cabauw: An overview. *Monthly Weather Review*, 143(3):828–844, 2015. doi:10.1175/MWR-D-14-00293.1.

[53] T. Schmidt, J. Kalisch, E. Lorenz, and D. Heinemann. Evaluating the spatio-temporal performance of sky-imager-based solar irradiance analysis and forecasts. *Atmospheric Chemistry and Physics*, 16(5):3399–3412, 2016. doi:10.5194/acp-16-3399-2016.

[54] Thomas Schmidt, John Kalisch, Elke Lorenz, and Detlev Heinemann. Retrieval of direct and diffuse irradiance with the use of hemispheric sky images. In *International Conference on Energy & Meteorology 2015*, Boulder, CO, June 2015. <http://icem2015.org/resources/oral-presentations/>.

[55] M. Schubnell. Sunshape and its influence on the flux distribution in imaging solar concentrators. *Journal of Solar Energy Engineering*, 114(4):260–266, 1992. doi:10.1115/1.2930015.

[56] M. Sengupta, A. Habte, S. Kurtz, A. Dobos, S. Wilbert, E. Lorenz, T. Stoffel, D. Renné, C. Gueymard, D. Myers, S. Wilcox, P. Blanc, and R. Perez. Best practices handbook for the collection and use of solar resource data for solar energy applications. Technical Report NREL/TP-5D00-63112, NREL, 15013 Denver West Parkway, Golden, CO 80401, February 2015. <http://www.nrel.gov/docs/fy15osti/63112.pdf>.

[57] Janet E. Shields, Monette E. Karr, Richard W. Johnson, and Art R. Burden. Day/night whole sky imagers for 24-h cloud and sky assessment: history and overview. *Appl. Opt.*, 52(8):1605–1616, March 2013. doi:10.1364/AO.52.001605.

[58] Bjorn Stevens, Chin-Hoh Moeng, Andrew S. Ackerman, Christopher S. Bretherton, Andreas Chlond, Stephan de Roode, James Edwards, Jean-Christophe Golaz, Hongli Jiang, Marat Khairoutdinov, Michael P. Kirkpatrick, David C. Lewellen, Adrian Lock, Frank Müller, David E. Stevens, Eoin Whelan, and Ping Zhu. Evaluation of large-eddy simulations via observations of nocturnal marine stratocumulus. *Monthly Weather Review*, 133(6):1443–1462, 2005. doi:10.1175/MWR2930.1.

[59] Bjorn Stevens, Chin-Hoh Moeng, and Peter P. Sullivan. Large-eddy simulations of radiatively driven convection: Sensitivities to the representation of small scales. *Journal of the Atmospheric Sciences*, 56(23):3963–3984, 1999. doi:10.1175/1520-0469(1999)056<3963:LESORD>2.0.CO;2.

[60] Björn Stevens, Thijs Heus, Bart van Stratum, and Axel Seifert. UCLA LES. https://github.com/uclales/uclales, 2015.

[61] Björn Stevens and Axel Seifert. Understanding macrophysical outcomes of microphysical choices in simulations of shallow cumulus convection. *Journal of the Meteorological Society of Japan. Ser. II*, 86A:143–162, 2008. doi:10.2151/jmsj.86A.143.

[62] Korntip Tohsing, Michael Schrempf, Stefan Riechelmann, Holger Schilke, and Gunther Seckmeyer. Measuring high-resolution sky luminance distributions with a CCD camera. *Appl. Opt.*, 52(8):1564–1573, March 2013. doi:10.1364/AO.52.001564.

[63] 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*, 8(2):875–890, 2015. doi:10.5194/amt-8-875-2015.

[64] Bryan Urquhart, Ben Kurtz, and Jan Kleissl. Sky camera geometric calibration using solar observations. *Atmospheric Measurement Techniques Discussions*, 2016:1–43, 2016. doi:10.5194/amt-2015-277.

[65] C. C. van Heerwaarden, B. J. H. van Stratum, T. Heus, J. A. Gibbs, E. Fedorovich, and J.-P. Mellado. Microhh 1.0: a computational fluid dynamics code for direct numerical simulation and large-eddy simulation of atmospheric boundary layer flows. *Geoscientific Model Development Discussions*, 2017:1–33, 2017. doi:10.5194/gmd-2017-41.

[66] Stefan Wilbert, Robert Pitz-Paal, and Joachim Jaus. Comparison of measurement techniques for the determination of circumsolar irradiance. *AIP Conference Proceedings*, 1556(1):162–167, 2013. doi:10.1063/1.4822222.

[67] Björn Wolff, Jan Kühnert, Elke Lorenz, Oliver Kramer, and Detlev Heinemann. Comparing support vector regression for PV power forecasting to a physical modeling approach using measurement, numerical weather prediction, and cloud motion data. *Solar Energy*, 135:197–208, 2016. doi:10.1016/j.solener.2016.05.051.

[68] Tianfan Xue, Michael Rubinstein, Ce Liu, and William T. Freeman. A computational approach for obstruction-free photography. *ACM Trans. Graph.*, 34(4):79:1–79:11, July 2015. doi:10.1145/2766940.

[69] Tetsuo Yamada. CCD image sensors. In Junichi Nakamura, editor, *Image Sensors and Signal Processing for Digital Still Cameras*, chapter 4. CRC Press, 2005.

[70] Handa Yang, Ben Kurtz, Dung Nguyen, Bryan Urquhart, Chi Wai Chow, Mohamed Ghonima, and Jan Kleissl. Solar irradiance forecasting using a ground-based sky imager developed at UC San Diego. *Solar Energy*, 103:502–524, May 2014. doi:10.1016/j.solener.2014.02.044.

[71] Hideki Yatsuzuka and Yoshiaki Uetani. Measurement of the all sky spectral radiance distribution using a digital camera with a fisheye lens. *J. Environ. Eng., AIJ*, 78(690):623–629, 2013. Japanese. doi:10.3130/aije.78.623.