Development of a Solar Forecasting System and Characterization of Irradiance Variability in Hawaii

Prepared for the

U.S. Department of Energy
Office of Electricity Delivery and Energy Reliability

Under Cooperative Agreement No. DE-EE0003507
Hawai‘i Energy Sustainability Program

Subtask 3.1: Photovoltaic Systems

Submitted by

Hawai‘i Natural Energy Institute
School of Ocean and Earth Science and Technology
University of Hawai‘i

December 2012

Acknowledgement: This material is based upon work supported by the United States Department of Energy under Cooperative Agreement Number DE-EE0003507.

Table of Contents

List of Figures

Introduction

In the state of Hawai‘i, photovoltaic systems generate a significant amount of power. Variability in this resource, primarily modulated by cloud formation and advection, is problematic in this region due to the small size of the island grids. This report presents: (1) statistical techniques, developed at HNEI, to characterize irradiance variability, and (2) the ongoing development of a forecasting system that predicts downwelling solar irradiance at minute-to-day time scales.

1. Characterizing Solar Irradiance

In this section, we present an overview of work submitted to *Renewable Energy* for publication:
Matthews, D.K., M. Dubarry, S. Buesquet, K. Stein, and R. Rocheleau: Characterizing Irradiance Time-Series in Hawai‘i for Application to Solar Power Generation and Forecasting.

This study is a statistical analysis of irradiance patterns and variability on time scales from seconds to months, using pyranometer observations collected on four of the main Hawaiian Islands. Study sites are compared using the following metrics: (i) frequency distributions of irradiance and clearness index at the seasonal and diurnal time-scale, (ii) estimates of the dominant timescale and variability using the semivariogram technique, and (iii) frequency distributions of quick changes in irradiance (ramps) on the scales of seconds to minutes.

Variability of solar irradiance contains both deterministic and stochastic signals. Deterministic seasonal and diurnal variations can be determined using simple astronomical relationships. However, atmospheric conditions, such as water vapor, turbidity, and most importantly clouds, randomly influence the amount of solar radiation that reaches the ground. Cloud formation in the Hawaiian Islands is complex, governed by the interaction of sea breezes and trade winds over dramatically varying topography through orographic lifting and blocking of the strong and constant pacific trade winds.

To characterize irradiance variability, eight years of pyranometer measurements from five test sites around the State of Hawai‘i are used. These test sites are located on four of the main islands in different climate conditions. Figure 1 shows the location of each PV test site and the topography of the island chain. For brevity, we show only results from a Hukseflux LP02-10 pyranometer, installed by HNEI as part of the Green Holmes Hall Initiative (GHHI) on the University of Hawai‘i at Manoa campus. This time series runs from July 1, 2011 to July 1, 2012.

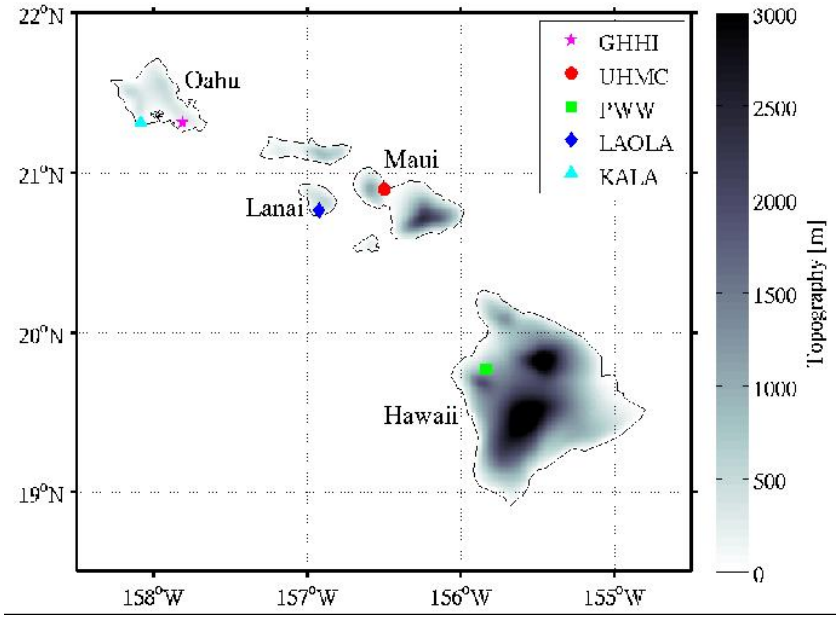

Figure 1: Topography of the Hawaiian Islands, with locations of the data sets used in the analysis.

It is possible to remove deterministic signals from the irradiance time series by dividing the pyranometer measurements by the top of the atmosphere (TOA) irradiance, leaving a time series of clearness index (CI). TOA irradiance data is provided by the National Renewable Energy Laboratory's (NREL) Measurement and Instrumentation Data Center (MIDC) solar position (SOLPOS) calculator. CI data is particularly useful in characterizing atmospheric attenuation since it no longer has a diurnal signal and is used along with the irradiance observations in the analysis.

1.1 Frequency Distributions of Irradiance & Clearness Index at Seasonal & Diurnal Time-scales

To characterize surface irradiance at seasonal and diurnal time scales, we bin data as a function of the time of day and month of the observation and calculate the mean irradiance of each bin. To characterize atmospheric attenuation at each site, we generate bivariate histograms of CI, binning data as a function of the observation time (either time of day or month) and CI level. With this information we can easily compare the time series from different areas of the islands.

Figure 2a presents the diurnal and seasonal irradiance patterns found by binning data as a function of observation time (using 15-minute and 1-month bins) and computing the mean of each bin. GHHI irradiance levels are relatively high throughout the year compared to other sites around the state, averaging ~800 W/m^2 in the hours around midday. The pyranometer at this site has a 20 degree tilt,

facing south to maximize power output through the year. This also reduces the amplitude of the seasonal cycle of irradiance. At a 20° tilt, TOA irradiance peaks at the Spring and Fall Equinoxes and has its minimum in the summer. GHHI observations follow this pattern, except during the Hawaiian winter, when cloud cover increases slightly. This is clearly shown in Figure 2b, which contains the bivariate histogram of CI binned according to the month of observation and CI level. GHHI was found to be remarkably clear, with a high percentage of CI values in the 0.7 to 0.9 range throughout the year. GHHI was also found to be clear throughout the day. Figure 2c, contains a bivariate histogram of CI with data binned according to time of day (using 15-minute bins), showing the highest percentage of CI values above 0.7 from 8:00 to 16:00. Finally, we summarize the average daily irradiance from GHHI data. Figure 2d shows average diurnal irradiance, CI, and TOA irradiance, along with the diurnal irradiance variability (given in standard deviation). Mean irradiance is over 800 W/m^2 at midday, with a mean CI value above 0.6.

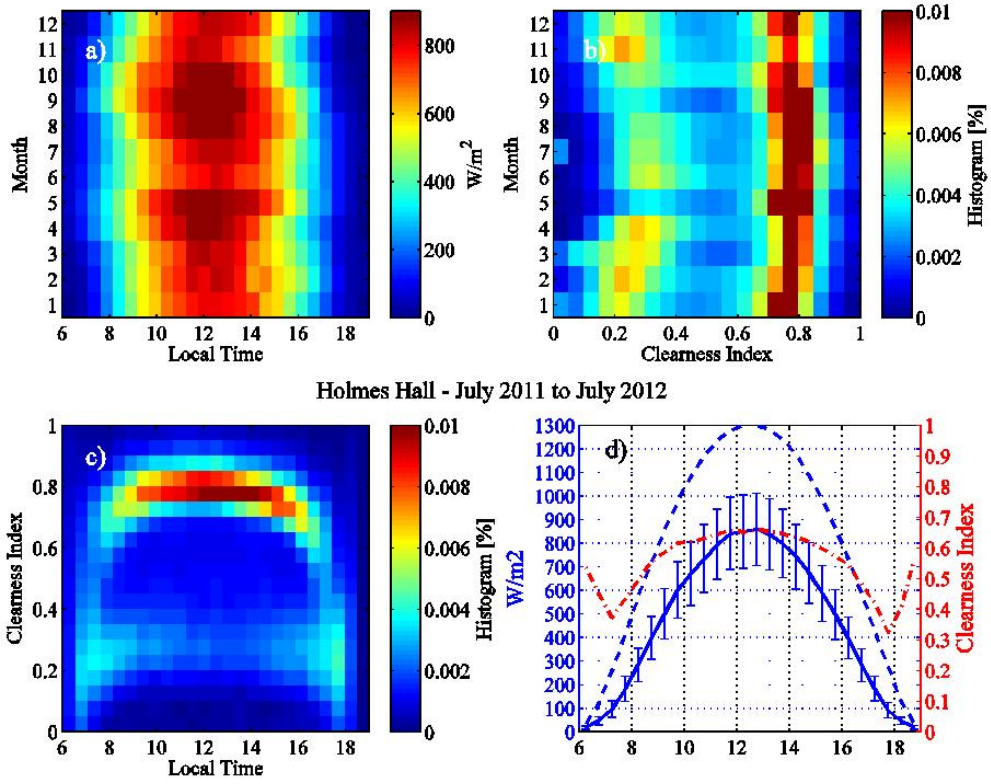

Figure 2: Statistics from irradiance and clearness index observations taken at GHHI (west of Honolulu, on the island of O'ahu). Shown in a) is the mean irradiance of data binned according to observation time (hour and month). In b) and c) we show bivariate histograms of clearness index data binned according to month and local time. In d) we show the mean observed irradiance binned every 30 minutes (solid line, with the STD given by error bars), against the mean TOA irradiance (dashed line); also shown is the mean CI (dot-dashed line).

Figure 2a presents the diurnal and seasonal irradiance patterns found by binning data as a function of observation time (using 15-minute and 1-month bins) and computing the mean of each bin. GHHI irradiance levels are relatively high throughout the year compared to other sites around the state, averaging ~800 W/m^2 in the hours around midday. The pyranometer at this site has a 20 degree tilt, facing south to maximize power output through the year. This also reduces the amplitude of the seasonal cycle of irradiance. At a 20° tilt, TOA irradiance peaks at the Spring and Fall Equinoxes and has its minimum in the summer. GHHI observations follow this pattern, except during the Hawaiian winter, when cloud cover increases slightly. This is clearly shown in Figure 2b, which contains the bivariate histogram of CI binned according to the month of observation and CI level. GHHI was found to be remarkably clear, with a high percentage of CI values in the 0.7 to 0.9 range throughout the year. GHHI was also found to be clear throughout the day. Figure 2c, contains a bivariate histogram of CI with data binned according to time of day (using 15-minute bins), showing the highest percentage of CI values above 0.7 from 8:00 to 16:00. Finally, we summarize the average daily irradiance from GHHI data. Figure 2d shows average diurnal irradiance, CI, and TOA irradiance, along with the diurnal irradiance variability (given in standard deviation). Mean irradiance is over 800 W/m^2 at midday, with a mean CI value above 0.6.

This analysis was also performed on the other data sets shown in Figure 1 (results can be found in the publication, Matthews et al. 2014). The statistical results agreed with the known climatological cloud conditions for each island. Test sites on the island of O‘ahu showed the highest mean irradiance levels and clear skies throughout the day. The test site on Maui showed high irradiance levels and clear skies in the morning. However, beginning in the early afternoon (14:00), clouds form, which reduces irradiance values. Test sites on the islands of Lanai and Hawai‘i showed high irradiance levels and clear skies only in the early morning, after which, afternoon clouds severely limited the amount of irradiance observed. All sites demonstrated seasonal shifts in irradiance and CI values that tracked seasonal changes in trade wind strength and variability. At each site, however, these shifts varied.

Examination of results also suggested a bimodal distribution of CI values, with a majority of values indicating very clear or very cloudy skies. Below, in Figure 3, histograms of CI values from each of the five data sets help further examine these results. We find all data sets show a bimodal distribution, with two distinct peaks, indicating either cloudy (CI < 0.3) or clear skies (CI > 0.7).

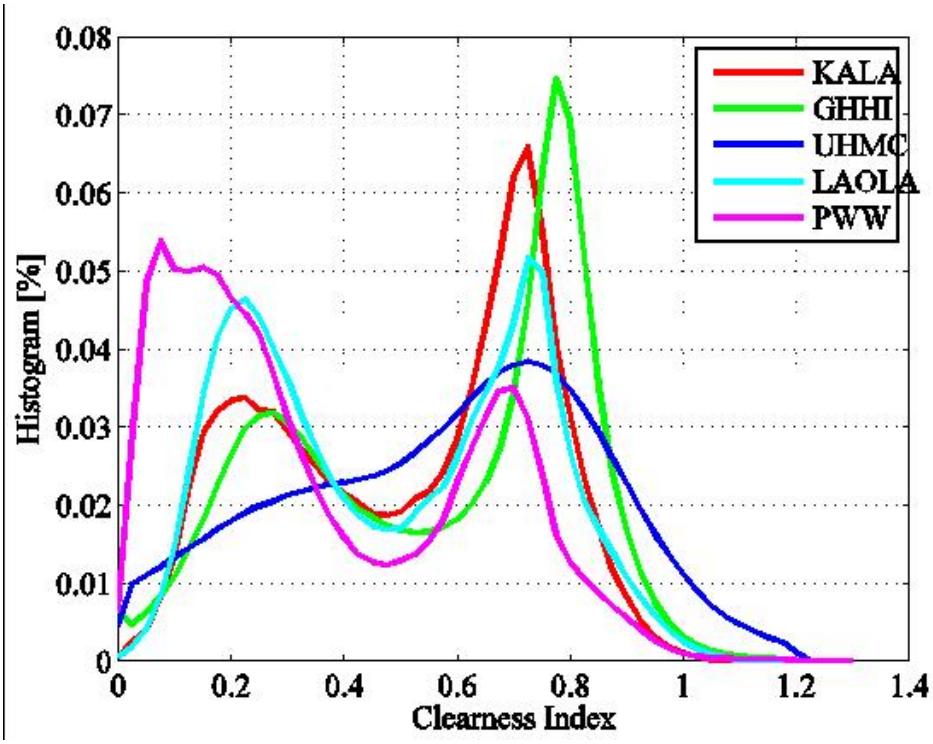

Figure 3: Frequency distribution of CI from observations from the five study sites. Where KALA and GHHI are located on the island of O'ahu, UHMC is on Maui, LAOLA is on Lanai, and PWW is on the island of Hawai'i.

1.2 Dominant Time Scale and Variability

CI time series lack the diurnal and seasonal signals of the irradiance measurements and are assumed to be stochastic, stationary time series. Fluctuations in CI are driven by random atmospheric processes that occur at varying time scales. This section briefly describes the semivariogram method (Journel and Huijbregts, 1978), a statistical technique that describes the temporal or spatial dependence of a data set that has been used to study dynamical processes (Matthews et al., 2011). We then apply this method to the CI time series to find the dominant time scale and variability imposed by the atmosphere on the irradiance observations at each test location. Estimating the dominant time scales and variability is essential for solar forecasting, as they are required by both optimal interpolation and data assimilation schemes, and give an idea of the predictability of irradiance at each site.

The semivariogram describes the covariance of data as a function of the data distance (time or space) by summing residual squared values according to a time lag. The empirical semivariogram is defined as the expected squared difference of the data values separated by the lag distance bins. For a full description of the semivariogram method, please see Banerjee et al., 2004. We fit a statistical model to

the empirical semivariogram to provide quantitative metrics that define the temporal dependence of the observations, specifically the dominant time scale and variability. This also provides a mathematical representation of the variance used in data assimilation methods.

Figure 4, below, shows the semivariogram calculated from the GHHI CI time series, using time lag bins of ten minutes. The empirical semivariogram is shown by the black line (circles indicate each bin) and the variability (in standard deviation) in each lag bin is indicated by the error bars. The figure also shows the dominant time scale (vertical red dashed line) and geophysical variability (difference between the two horizontal red dashed line), as determined by the exponential model fit (red solid line). For the GHHI observations the dominant time scale is 166.9 minutes and the CI geophysical variability is 0.031.

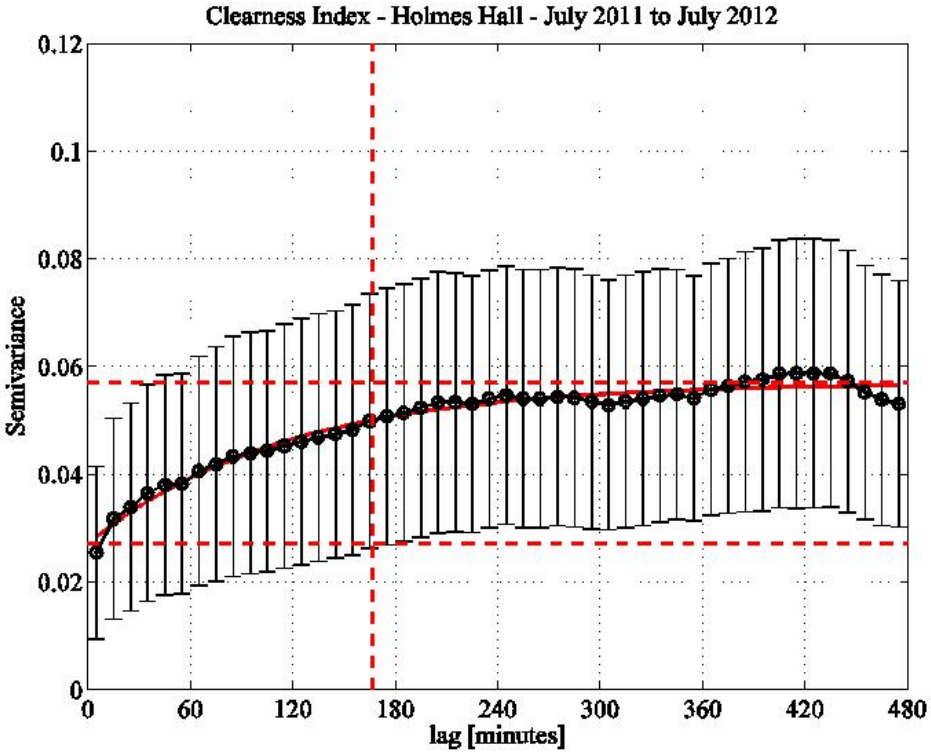

Figure 4: Semivariogram calculated from clearness index data from the GHHI test site at the campus of the University of Hawai'i. The empirical semivariogram is indicated by the black line and error bars show the variability of the data in each 10-minute bin. The red line indicates the exponential model fit and the nugget, sill, and range are represented by the red dashed lines.

From the evaluation of the semivariograms from the different study sites around Hawai'i, we find that for sites with primarily clear skies, time scales are ~2 hours—significantly shorter than sites dominated

with afternoon clouds. At these sites, the dominant time scales are likely driven by the average time for a cloud to pass through the line of sight from the sensor to the Sun. As for the sites dominated by cloudy skies, time scales are ~4-6 hours. These longer time scales are likely driven by the persistence of afternoon clouds, suggesting that once the site is cloudy it will remain so. Also, clear skies are inherently less variable than cloudy skies and at the sites with primarily clear skies the semivariogram indicates less geophysical variability.

1.3 Characterizing Ramp Events

In this section, we characterize the magnitude, duration, and frequency of occurrence of changes in irradiance that occur on very short time scales (ramps). Ramps are captured in the irradiance time series using a data compression technique, the swinging door algorithm, which defines the dominant points in a time series. Ramp magnitudes and durations then are calculated from the differences in irradiance levels and time between the dominant points.

The swinging door compression algorithm is a data-trending technique that identifies dominant points in a time series (we point readers to Bristol [1990] for a more detailed description of the algorithm). The method obtains the optimal straight-line approximation to a segment of a time series, given the initial point in that line. The line is optimal in the sense that it is the longest straight line possible for the data, given a set error, and that it is computed with the minimum number of computations possible.

This method is illustrated below in Figure 5. Starting at an initial point, the method sets two pivot points: one above and one below the initial point. Lines set between the initial point and each pivot point can be thought of as closed doors, set at a distance that defines the allowable error in the straight-line estimation. The doors are then swung open at the pivot points, along the time series, with the top door swinging up and the bottom swinging down. As each new point in the time series is considered, the doors are opened further to include that point within the area bounded by the doors. When a data point is reached, beyond which the doors would open past parallel, a midline segment is generated, starting at the initial point and extending to the time of the last point considered. These two points are considered the dominant points in the time series. The last data point considered is then used as the initial point in a new data segment, continuing the process to the end of the time series.

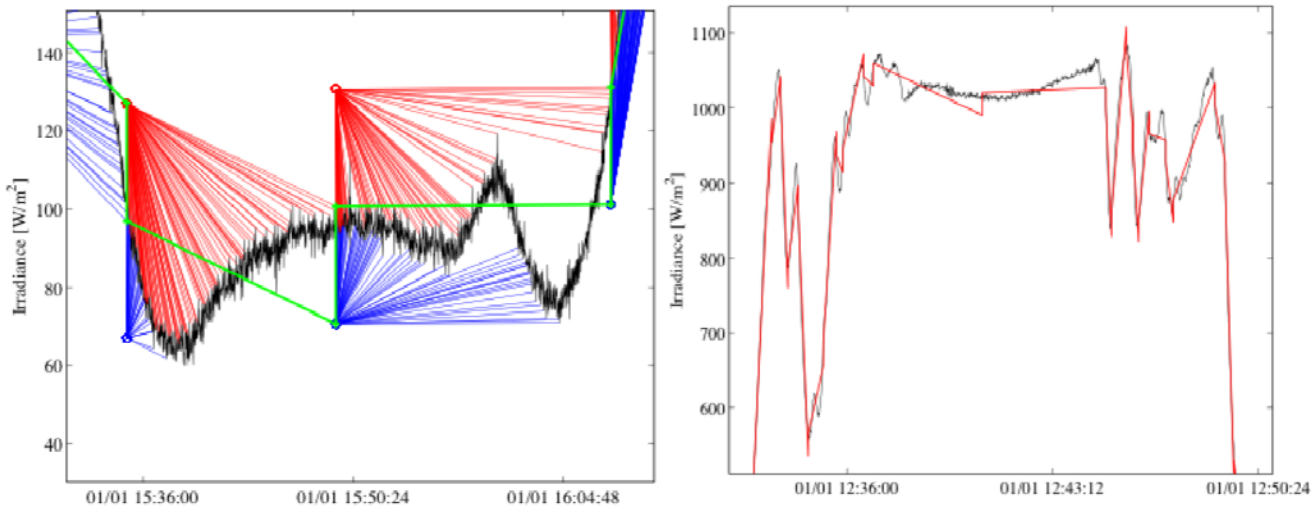

Figure 5: Illustration of the swinging door algorithm applied to 1-second pyranometer data (a), along with a small 15-minute section of the resulting dominant points (red line) and raw data (black line) (b).

After experimenting with a range of values, the following results for GHHI observations were obtained using a door of 75 W/m^2. The swinging door method reduced the 1-second data by 99.57%. We focused on the larger, more significant, changes in irradiance by filtering all ramps with magnitudes less than ±50 W/m^2. The average ramp magnitude was 257.9 W/m^2, with a standard deviation of 140.8 W/m^2. The mean duration between dominant points was 1.93 min, with a STD of 5.3 min. Figure 6 shows the bivariate distribution of ramp magnitudes and durations using bins of 50 W/m^2 and 12 sec. The majority of the ramp events were less than two minutes, accounting for over 81%. Also, we found that ramps of decreasing irradiance were on average 40 sec faster than positive ramps.

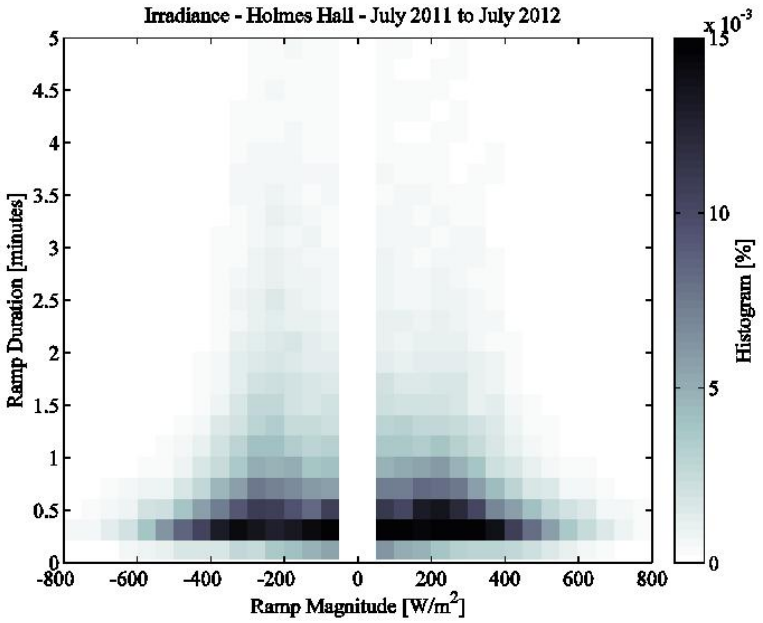

Figure 6: Bivariate distribution of ramp magnitudes and durations from 1-second pyranometer observations taken on the campus of the University of Hawai‘i.

The time of day that these ramp events occur is important when considering PV power variability and grid management. In Figure 7, we show bivariate histograms of absolute ramp magnitude and ramp duration binned at the hour of ramp occurrence. These histograms show that ramps are primarily in the 12 to 36 sec range, with the largest ramp events occurring in the hours around midday, and the highest percentage of ramps in the 200-300 W/m^2 range. The constant ramps seen through the day indicate the persistent advection of small clouds.

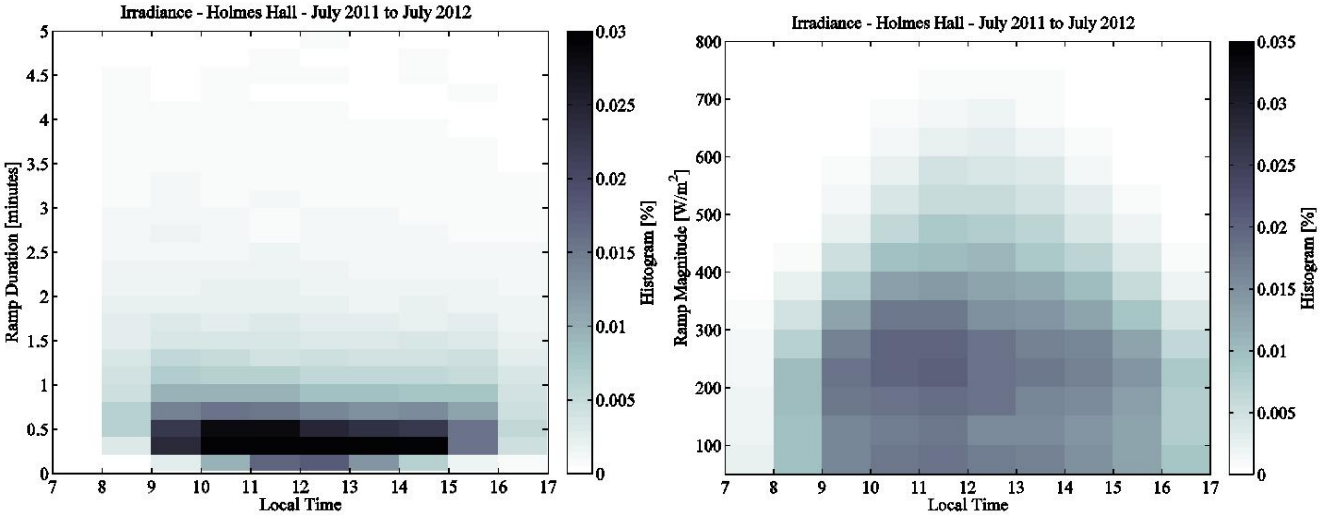

Figure 7: Bivariate histograms of ramp magnitude, duration, and time. Ramps are calculated using the swinging door method applied to data taken from east of Honolulu.

2. Solar Forecasting System

The solar forecasting system being developed is made up of three separate components: (1) a numerical weather prediction (NWP) model; (2) satellite data; and (3) ground based data. Each component generates forecasts that are most accurate at different forecasting horizons.

2.1 NWP Model

For regional, intraweek forecasts (hours to a few days ahead), the system employs the Weather Research and Forecasting (WRF) model. WRF is a widely used mesoscale NWP model developed at the National Center for Atmospheric Research (NCAR). We employ three separate WRF grids to generate statewide forecasts of irradiance, but focus on the island of O'ahu, where a majority of the population of the state resides. Figure 8 shows the spatial coverage of each domain. The parent and coarsest grid covers the entire Hawaiian Island chain at a 4-km resolution. Inside that grid, a 1-km grid covers the islands of Maui, O'ahu and Hawai'i. Finally, the highest resolution grid covers the island of O'ahu at 800 m. A long spin-up of the model dynamics is currently being run. After the spin-up is completed the model will run operationally, producing a two-day forecast of irradiance every morning. After the model is operational, we will begin testing the system, focusing on refining WRF to most accurately produce solar forecasts for the region. In future work, we will focus on integrating data assimilation techniques into the forecasting system, taking advantage of the satellite and sky imager observations used in the other solar forecasting system components.

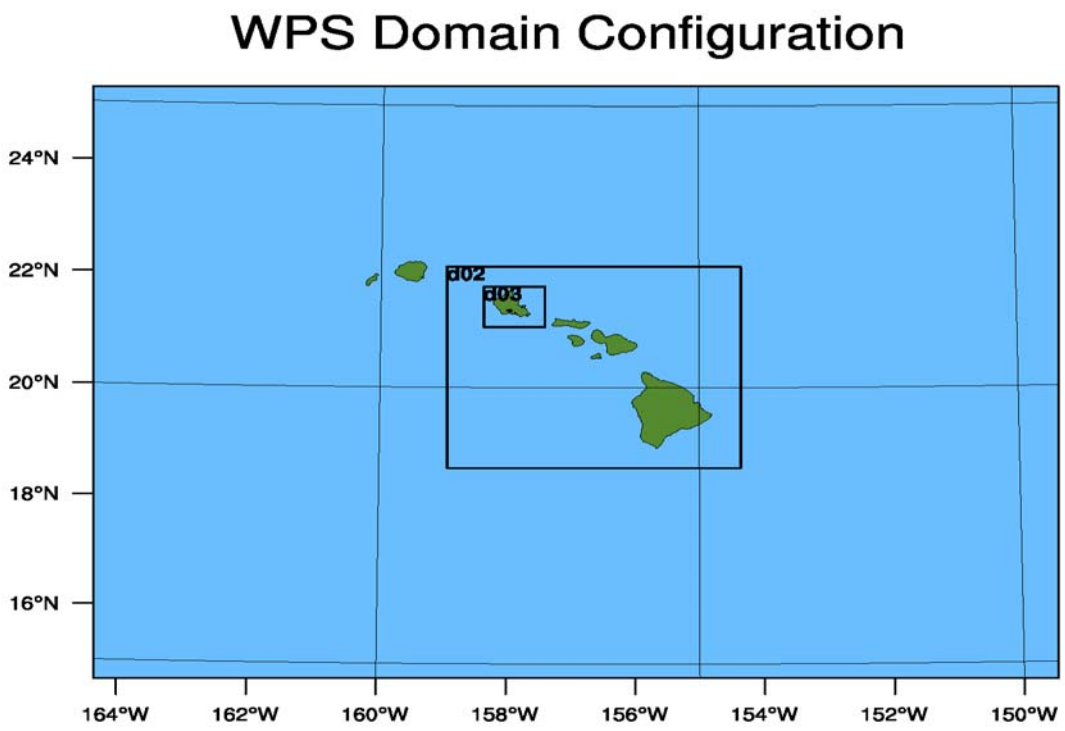

Figure 8: The spatial coverage of each grid of the NWP component. The largest grid is at a 4-km resolution, while domain 2 (d02) is at 1 km and domain 3 (d03) at 800 m.

2.2 Satellite Data

For regional, intraday forecasts (minutes to hours ahead) that cover the entire Hawaiian Island chain, the system uses a variety of satellite data to monitor current cloud conditions, along with cloud identification, image pattern tracking, linear advection algorithms, and an irradiance model to generate solar forecasts.

In this section, we briefly describe the steps and algorithms used to generate solar forecasts from satellite data. The Geostationary Operational Environmental Satellite (GOES-WEST) is the main source of data for this component, because the satellite constantly monitors the same portion of the Earth and gives continuous updates of current cloud conditions. Real-time GOES, 1-km, visible imagery is provided by the Marshall Space Flight Center Earth Science Office. This data is made available for downloading around 15 minutes after the image is collected by the satellite. After the data is downloaded, it is converted into radiance units and the Hawai‘i region is extracted from the full disk image. The data is then converted and saved into the Network Common Data Form (NetCDF) format.

NetCDF allows for easy access and sharing of array-oriented scientific data and is widely used in the meteorological and oceanographic modeling communities.

Cloud velocities are then determined by applying the maximum cross-correlation (MCC) technique to sequential GOES images, nominally 15 minutes apart. The MCC technique, illustrated in Figure 9, is a fully automated, robust pattern tracking algorithm that calculates the displacement of small sub-regions from one image to another. The procedure cross correlates a template sub-window in an initial image with all possible sub-windows of the same size that fall within the search window of a second image. The location of the sub-window in the second image that produces the highest cross correlation with the sub-window in the first image indicates the most likely displacement of that feature. A velocity vector is then calculated by dividing the displacement vector by the time separation between the two images. Because the method is fully automated and conditions may exist in which there are little to no features to track, filtering the raw MCC vector field is essential. We employ both a correlation cutoff, to ensure only the very similar patterns are matched, and a next-neighbor filter, which allows only a spatially smooth velocity field.

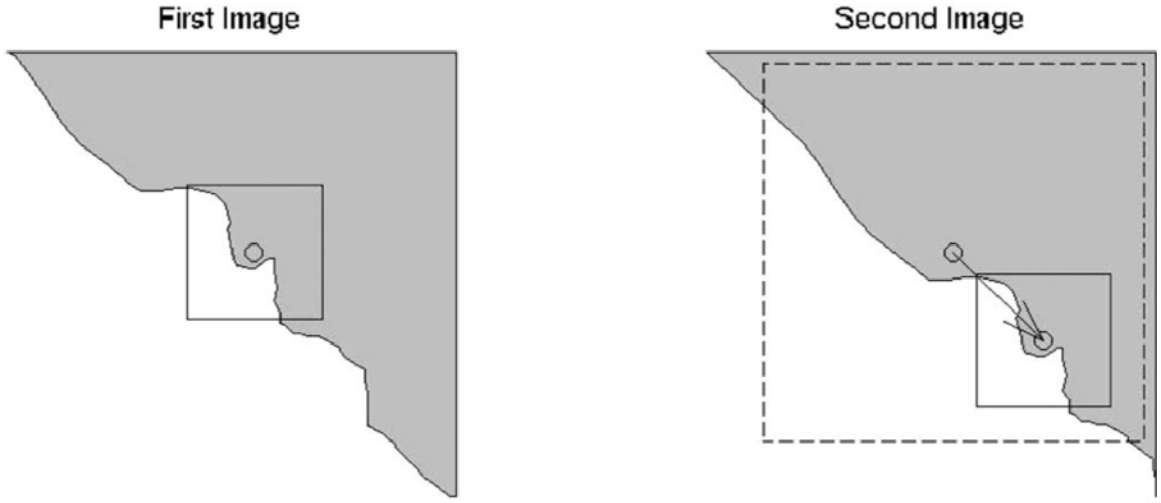

Figure 9: Illustration of the Maximum Cross Correlation method.

The most recent GOES image is then advected using the cloud velocities from the MCC method and an advection algorithm. The advection method used, developed by Stam 2002, is based on the Navier - Stokes equations. The method was chosen for its stability and speed, and because it can be advanced to arbitrary time steps. This allows for forecast of cloud locations at any time past the most recent GOES

image. The current system generates cloud forecasts from 15 minutes to six hours, at a resolution of 15 minutes.

To derive surface irradiance from the GOES imagery, we employ the HELIOSAT method (Hammer et al., 2002). This method deals with atmospheric and cloud extinction separately. In the first step, the clear sky irradiance is calculated using the Linke turbidity factor to describe atmospheric extinction. We estimate this factor for Hawai'i using more than 12 years of pyranometer observations from across the state. In a slight variation from the normal application of the HELIOSAT method, we vary the turbidity factor both diurnally and spatially. A manuscript on the analysis turbidity in Hawai'i is currently being prepared for publication. The second step uses both empirical and physical relationships to determine how much light passes through a cloud based on the brightness of that cloud, where the reflected radiance of a cloud, measured in the visible channel, is approximately proportional to the reduction of surface irradiance by that cloud.

This satellite component of the forecasting system is run without human intervention every 30 minutes on a linux server at HNEI. The full process, from downloading the image to producing a six-hour surface irradiance forecast, takes around 30 minutes. Currently, we are testing post processing techniques to increase the accuracy of the forecast. These include: using cloud top height data to account for cloud shadows, filtering and smoothing techniques, and linear adjustments to the forecasted irradiance to take account for biases in the method.

2.3 Ground Based Data

For local, intrahour forecasts (seconds to one hour), the system will use a total sky imager, which provides images of the entire sky every 20 seconds, and a ceilometer, which provides cloud height data. This information is ingested into cloud identification, geometric transformation, image pattern tracking, and advection algorithms to generate solar forecasts within 10 to 15 km of the instrumentation, depending on cloud height. The sky imager is installed on the roof of the Hawai'i Institute for Geophysics (HIG) building at the University of Hawai'i at Manoa. The ceilometer will be installed in May 2014. Along with these instruments, two pyranometers will be placed on the roof of HIG. These instruments will provide data needed to perform an extensive validation and calibration of the solar forecasting system. In Figure 10, we show an example sky image and the ratio of the red and blue channels that are used to establish if a pixel contains a cloud or clear sky. The far-right panel in Figure 10 is the cloud map generated by this method. This information will be propagated through an advection algorithm similar to the GOES imagery to generate forecasts of cloud locations.

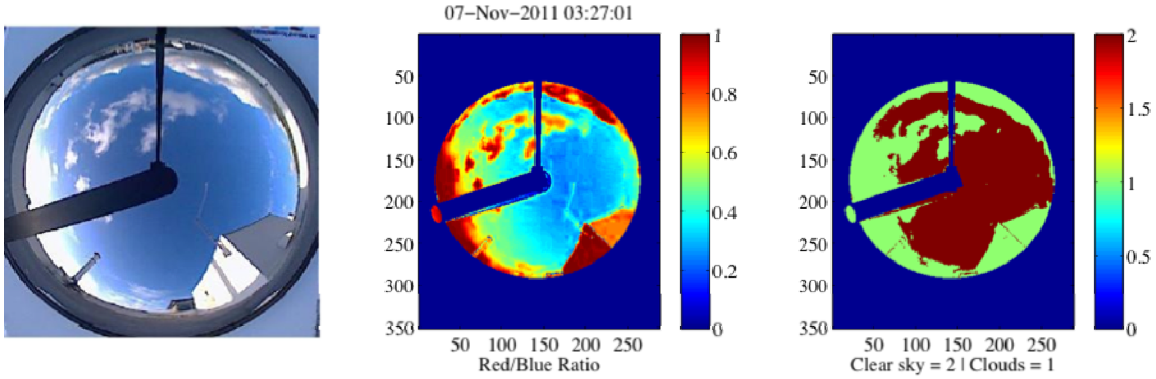

Figure 10: Processing steps of cloud detection on sky imagery. The raw sky image from November 7, 2011 shows relatively clear skies (left). The RBR is used to identify cloud containing pixels through the change in atmospheric scattering characteristics (center). Finally, the cloud map is generated from a simple RBR threshold (0.5), with clear sky shown as 2 and clouds as 1 (right).

Summary

The statistical analysis of pyranometer observations was done in preparation for development of a solar forecasting system. Dominant time scales and variability, seasonal/diurnal irradiance patterns, and the size and duration of ramp events are necessary information for building, calibrating, and validating such a system. These results can also be used to determine the uncertainty for energy production in Hawai‘i, which could be used for grid management. This work is completed and has been submitted for publication.

A solar forecasting system is being developed at HNEI using three separate components: a numerical weather prediction model, satellite data, and ground based data. The satellite component of this system has been operational since the beginning of 2014. The NWP and ground-based components will be operational in the spring of 2014. In future work, each component will be evaluated for forecasting accuracy using pyranometer observations from across the state. Also, the components will be merged into one forecasting system using the WRF data assimilation system.

References

Bristol, E., “Swinging Door Trending: Adaptive Trend Recording” 1990, ISA National Conference Proceedings.

Chow, C., Urquart, B., Lave, M., Dominguez, A., Kleissl, J., Shields, J., Washhom, B., ”Intra-hour

Forecasting with a Total Sky Im ager at the UC San Diego Solar Energy Testbed" Solar Energy, 2011, 85, 2881-2893.

Jedlovec, G., Haines, S., LaFont aine, F., "Spatial and Tem poral Varying Thresholds for Cloud Detection in GOES Imagery" IEEE Trans. on Geo. and Remote Sensing, 2008, 6, 1705-1717.

Journel, A.G., Huijbregts, C., 1978. Mining Geostatistics. Acedemic Press, Harcourt Brace Jovanovich Pubilshers, New York, New York 600 pp.

Kitanidis, P.K., 1997. Introduction to Geostatis tics, Applications in Hydrogeology. Ca mbridge University Press, New York, New York 249 pp.

Matthews, D., Powell, B., Milliff, R., "Dominant Spatial Variability Scales from Observations around the Hawaiian Islands", Deep Sea Res., 2011, Part I, 58(10), 979–987.

Zanifé, O. Z., Vincent, P., Am arouche, L., Dumont, D., Thibaut, P ., Labroue, S., "Comparison of the Ku-band Range Noise Level and the Relative Sea-state Bias of the Jason-1, TOPEX, and Poseidon-1 Radar Altimeters", Mar. Geod., 2003, 26, 201–238.