



Prediction of Photovoltaic Power Generation from Cloud Imaging

Master Thesis

Amrollah Seifoddini

March 31, 2016

Advisors: Prof. Marc Pollefeys, Dr. Jan Poland
Department of Computer Science, ETH Zürich

Abstract

Managing fluctuation in photo-voltaic power plants which is frequent in cloudy days, is one of the big challenges that need to be solved in order to significantly increase its penetration into the power grid. One possible approach to predict short term variations is vision-based which includes a fish eye camera pointing into the sky, taking image sequences. Cloud states are predicted for near future using cloud segmentation and optical flow. In this context, we investigate the role of the cloudiness in shielding the direct sun irradiation and also reflecting the sunlight which increases the diffuse irradiation. Specifically, by analyzing our irradiation sensor measurements and image features we learn a soft sensor regressor for inferring irradiance of a given image. Direct component of irradiation is inferred using sun detection algorithm and cloud map. Finally the diffuse irradiation is estimated using several features derived from cloud state and date and time of image. Several regression algorithms are compared and Support Vector Ordinal Regression Machine delivers the best result with $34W/m^2$ error and $33W/m^2$ error standard deviation.

Acknowledgements

This work would not have been feasible without the clever ideas and the patient explanations of my supervisor Dr. Poland. Also, I would like to thank ABB for providing me with the infrastructure, data, and human support which allowed me to conduct my thesis in the best conditions.

Contents

Contents	iii
List of Figures	1
List of Tables	2
1 Introduction	3
1.1 Power prediction approaches for a PV plant	3
1.1.1 Ground whole-sky imagery	4
1.2 Cloud segmentation, cloud tracking	4
1.3 Image irradiance estimation for power prediction	5
1.4 Irradiance components	6
1.5 Accuracy metrics	7
2 Related work	9
2.1 Estimate irradiance from zone types in sky images	9
2.2 Using clear sky irradiance model and binary cloud mask . . .	10
2.2.1 Cloud detection	11
2.2.2 Image un-distortion	11
2.2.3 Shadow mapping	12
2.2.4 Irradiance retrieval	12
2.2.5 Irradiance forecast	13
2.3 Retrieval of direct and diffuse irradiance from sky images . .	13
2.3.1 Image features	13
2.3.2 irradiance estimation	14
3 Estimating Diffuse Horizontal Irradiance (DHI) from sky image	15
3.0.1 Experimental Setup	15
3.0.2 Acquired Data	17
3.0.3 Sun positions and sun states in image	17
3.0.4 Clear-sky irradiance model	19

CONTENTS

3.0.5	Estimating diffuse from pyranometers	23
3.0.6	Key image features affecting DHI	23
3.0.7	Learning the relation between image features and DHI	23
3.0.8	Translating irradiance to power	23
4	Results	29
5	Discusion	31
6	Conclusion	33
	Bibliography	35

List of Figures

1.1	Cloud segmentation sample	5
1.2	Irradiance components	7
2.1	Three different zones identified in images (sun, cloud, sky) . . .	10
2.2	Correlation between estimated DNI and recorded DNI.	10
2.3	Comparing estimated k_{DHI}^* and k_{DNI}^* to measured values. source:[28]	14
3.1	wide-angle camera system used at Carviglia site	16
3.2	Two pyranometers (one horizontal, one 45 degrees titled to north) located close to the camera location	16
3.3	Sun position angles. source:[1]	18
3.4	Different sun states, from left to right: sun flag=4, 3, 2, 1 respec- tively.	19
3.5	Variation of sun appearance in the images.	20
3.6	GHI estimation of McClear vs Ineichen vs observation	24
3.7	Correlation of GHI simulation of McClear and Ineichen with ob- servations on selected days	25
3.8	DNI and DHI correlation of McClear to Ineichen	26
3.9	Comparison of DHI of McClear and Ineichen to observed GHI .	26
3.10	Variation of irradiance components during day	27
3.11	Overall simple system design for power prediction	27

List of Tables

LIST OF TABLES

Chapter 1

Introduction

Solar energy is one of the key alternative energy sources. The recent developments in solar panels and different business models around them made the photo-voltaic(PV) power plants more economical. However, the variability in PV output power make the integration into main energy grid risky and slow[26]. These fluctuation comes from cloud states in sky, and it has two different effects, one decreasing the power, the other one increasing the power. Firstly, if the clouds cover the sun completely or partially, some area of plant is shaded and does not receive direct sunlight, causing a power drop. On the other hand, if the clouds are not blocking the the sun completely or at all, based on their type, height, position and time, they can re-reflect some part of the irradiation¹ which is reflected by ground, back into the power plant. In this case, the input irradiance² and consequently the output power increases. In the electricity grid, the stability of power is vital. Therefore, if we want to integrate a PV power source into the grid we need to compensate for any power drop by using other electricity sources, and also restrain any excessive power. That is why we need to predict these short-time power changes in advance to design better strategies for handling them and ultimately provide a guaranteed stable power in the grid. In this chapter, we first explain different approaches towards this prediction problem, then we describe overview of the setup used in this study, and finally we talk about accuracy metrics for the result.

1.1 Power prediction approaches for a PV plant

The large variety of cloud characteristics such as motion, height, opacity and spatial distribution makes the cloud-induced fluctuations difficult to

¹Irradiation is the sum of irradiance over a time period, expressed in Wh/m^2

²Irradiance is understood as instantaneous density of solar radiation incident on a given surface, typically expressed in W/m^2

predict. However, according to these comprehensive surveys[12, 18], solar irradiance forecasting techniques have been successfully developed. These include numerical weather models (NWP)s[21] using pattern recognition of meteorological data for irradiance prediction, satellite-based forecasts using cloud motion vectors to determine sun occlusion based on fixed velocity model and predict power [15, 19, 10], statistical methods based on machine learning applied on past several years trend[34] and time series analysis[24] which are mostly developed for intra-day and day-ahead forecasts. However, for very short-term power prediction applications, the interest horizon stretches up to 30 minutes ahead. And therefore, these methods fall behind the required spatial or temporal resolution required on cloud-induced irradiance variability[12].

1.1.1 Ground whole-sky imagery

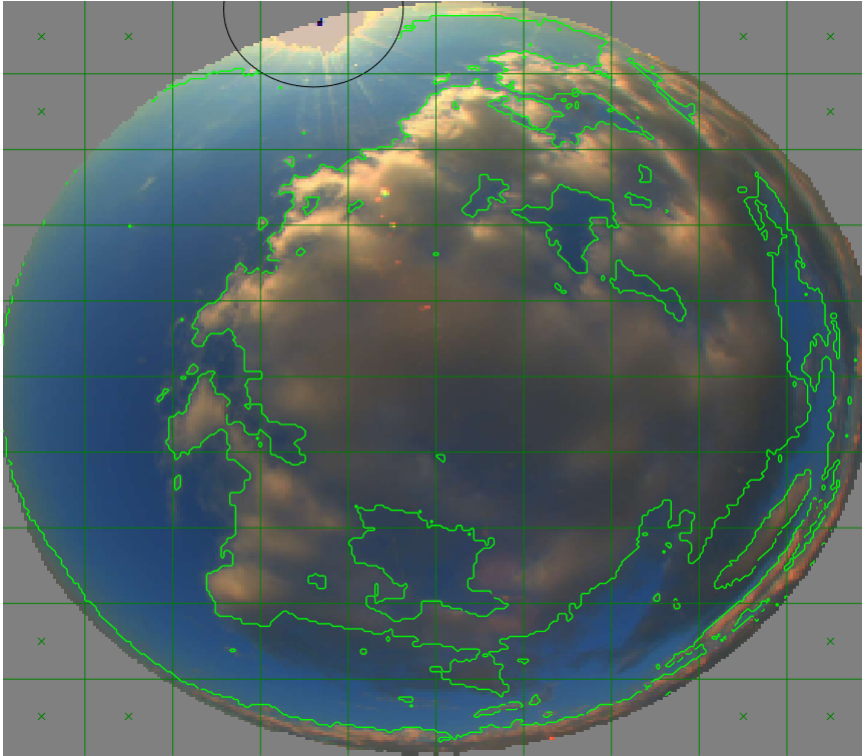
For achieving this high resolution forecasts, vision-based methods using total-sky-cameras are developed. The earliest works use cameras for monitoring cloud cover characteristics[22, 7] and aerosol properties[20, 6]. In recent years, using sky cameras for solar irradiance forecast has grown rapidly and several successful works have been developed by analysing motion, optical and distribution of clouds in the whole-sky images captured by a fisheye lense camera.(Chu et al., 2015;West et al., 2014; Quesada-Ruiz et al., 2014; Chu et al., 2013; Fu and Cheng, 2013; Yang et al., 2014; Bernecker et al., 2014; Chow et al., 2011; Marquez and Coimbra, 2013). Using only one camera for data input, some methods[33] predict sun occlusions perceived by the camera and therefore, their forecast is only valid for the point very close to the camera, since areas further away might be sunny or cloudy and camera's point of view is not covering that area. However, one can incorporate cloud base height to calculate projection of clouds shadows on the ground. This information is usually acquired by using a laser based cloud base sensor (ceilometer) and make the area irradiance forecasts more accurate[35]. It is theoretically possible to use two or more cameras in the site mounted with a distance about 50m and by applying stereo algorithm, find the cloud height; however, this method has not been investigated in practice yet. In this research we focus on ground whole-sky imagery methods aimed for very short-term irradiance forecast.

1.2 Cloud segmentation, cloud tracking

For predicting the future state of clouds in the sky we need to first know where are the clouds. This is done by applying a dynamic-threshold segmentation algorithm on red to blue channel ratio of RGB images. Several studies has shown that the red to blue ratio is a good criterion for cloud segmentation, but the threshold for this ratio should be set in a way that

discriminate clouds locally but be smooth globally. This method handles cloud color variation on different cloud types very well. One sample result of such segmentation is figure 1.1

Figure 1.1: Cloud segmentation sample



After detecting the clouds in a sequence of images, one can apply optical flow algorithm on some points of interest in the first image and extract the cloud movement as motion vectors of optical flow. This cloud tracking pipeline gives us the clouds position in the sky for very short time in future. The result of experiment on several future time horizon has shown that accuracy decreases considerably after 30 minutes, specially in high speed cloud motions.

1.3 Image irradiance estimation for power prediction

The final step in power prediction is to associate a potential power estimate to any time in several minutes ahead knowing the cloud positions and their characteristic in that time. The generated power of a PV plant depends on several factors including received irradiance, operational temperature and panels specification. However, the only factor which changes rapidly and has a huge impact on the output power is irradiance. Therefore, in our

power prediction framework, we use a power prediction adaptation method with recorded data of previous minutes to derive the power output from future irradiance estimations. The scaling factor for the result is calculated by dividing the recorded power output and irradiance sensor measurements. Thus, this power adaptation mechanism, separates and defines our main problem as estimating the received irradiance from a clouds attributes in a specific time and location.

1.4 Irradiance components

The total solar radiation -GHI³- which hits the surface of solar panels consists of three basic components, direct -DNI⁴-, diffuse -DHI⁵- and reflected. The direct part comes from the sunlight beams directly raying from sun direction to the solar module. While passing through atmosphere, some amount of sunlight scatters in every direction by dense particles. The portion of this scattered light which hits the module forms the diffuse irradiation for solar panels. Reflected irradiance represents sunlight that is reflected off the clouds or ground around the array of panels. The source of this reflected radiation can be DNI or DHI. Rate of the reflection depends on clouds coverage, size of the ground that is visible from the module and their albedo coefficient⁶. The albedo coefficient for ground is typically 0.2, though it can be higher during snowy periods in cold climates. The albedo coefficient for clouds depends on their type, density, temperature and etc. These components are shown in figure 1.2

Total irradiation is related to other three components with this formula:

$$GHI = DNI \times \cos(Z) + DHI + reflected$$

where Z is the solar zenith angle-the angle between the direction of the sun and the line directly overhead. Since distinguishing between the reflected and diffuse irradiation is practically hard and also there is not any ground truth value for training, we decided to combine both of them as the diffuse component. Thus, the formula changes to:

$$GHI = DNI \times \cos(Z) + DHI \tag{1.1}$$

where DHI is sum of all non direct irradiations.

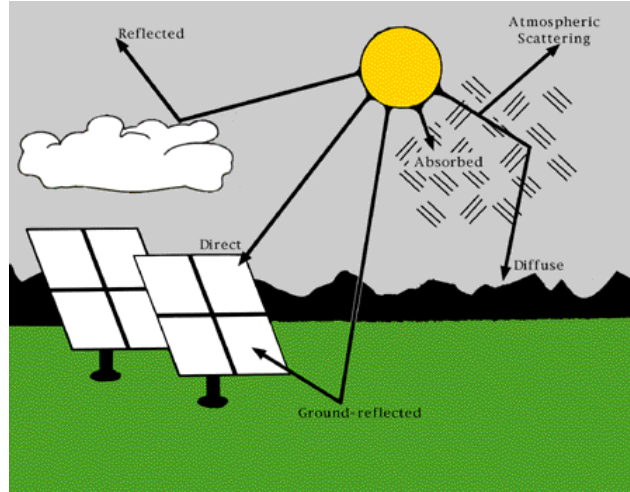
³Global Horizontal Irradiation

⁴Direct Normal Irradiance

⁵Diffuse Horizontal Irradiance

⁶The portion of the incident irradiance that is reflected

Figure 1.2: Irradiance components



1.5 Accuracy metrics

For measuring accuracy of the result we can use popular error measures such as RMSE-Root mean square error. However, since there are not any other publicly available study on this specific problem to compare our RMSE with, we can define our own error measures which quantify our solution quality better. For example, apart from RMSE, we calculate a relative error as well that penalize errors for irradiances higher than a specific value more. And errors for irradiances lower than that value, will be penalized less. This cutting value is set as 100 and the scale is logarithmic, because when the irradiance is less than 100, the power output of plant is very low and practically not useful in grid. On the other hand, we want more precise result for very high irradiances. This means errors for small irradiances are less important than errors in big big ones. We also use MBE (mean bias error) and R^2 for correlation coefficient.

Chapter 2

Related work

In this chapter we survey the studies which focus on problem of irradiance estimation using sky images. The usage of ground-based cameras for studying effect of clouds on irradiation has a long history, as early as 1977 when Borkowski et al.[4] developed the first whole-sky camera system for investigating effects of clouds on middle ultraviolet global radiation. In this study, the degree of solar obstruction and cloud coverage were determined visually from the images. Later in 1998, Jeff Sabburg and Joe wong[25] developed and evaluated the first automated, ground-based, sun-centered sky camera system for cloud assessment. However, since the purpose of study was the clouds effect on UVB¹ radiation they only considered a small area around the sun for cloud and sun obstruction detection which is of paramount importance for this rays. They use a threshold-based approach on gray scale pixel intensities for cloud detection. They also use solar radiation measurements in a image processing algorithm to reduce reflections from the sun on the camera system being mistaken for cloud in the images.

2.1 Estimate irradiance from zone types in sky images

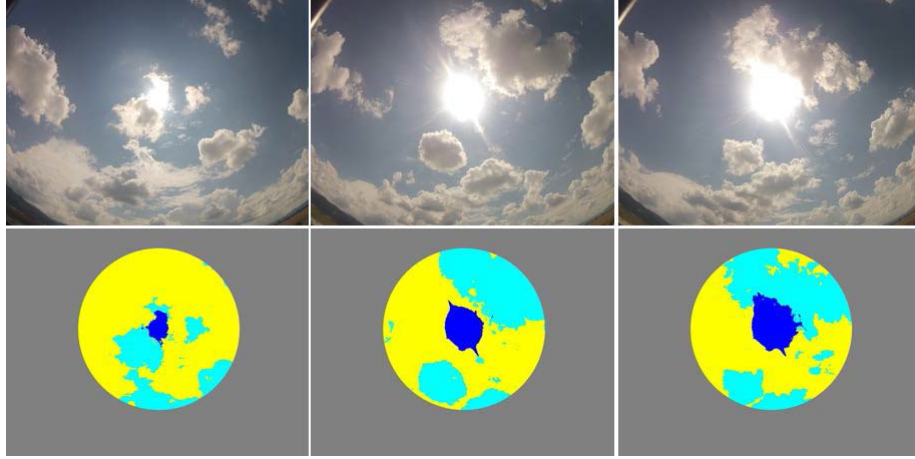
One of the recent researches done in this area is held as a collaboration between Universitatea Transilvania din Braşov in Romania and Cyprus University of Technology[31][32]. This work uses sets of two consecutive images taken by wide-view angle GoPro Hero2 camera(one with normal exposure and the other one under-exposed) and extracts their RGB², HSV³ components. Then by learning several intensity ranges, they segment four zone type in each image: sun, blue sky, thin clouds and thick clouds. One sample of segmentation is shown in figure2.1.

¹Ultraviolet B

²(Red, Green, Blue)

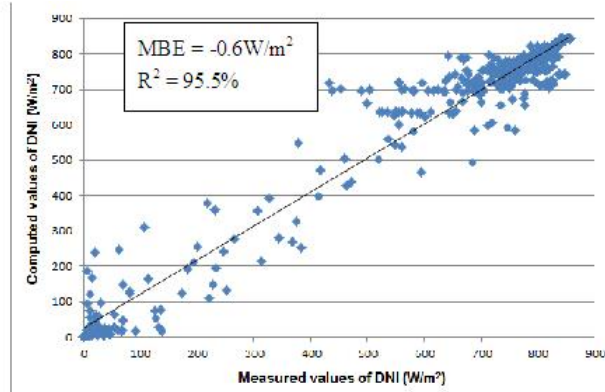
³(Hue, Saturation, Value)

Figure 2.1: Three different zones identified in images (sun, cloud, sky)



The irradiance (direct, diffuse, global) is recorded using the equipment Kipp & Zonen, Solys2 at the same time of image capturing. Finally, a regressor used to estimate direct irradiation (DNI) based on a feature vector consisting the number of pixels of different zone types in the images. The correlation in the result is shown in figure2.2.

Figure 2.2: Correlation between estimated DNI and recorded DNI.



2.2 Using clear sky irradiance model and binary cloud mask

The work done by T. Schmidt et al.[29] at University of Oldenburg in Germany is a very recent and relevant work on irradiance forecast using sky imager pictures. The experimental setup consists of a wide-view camera, one ceilometer (cloud base height sensor) located next to the camera, and a grid of 99 pyranometer distributed uniformly over 10km by 12km in the

area close to camera. The aim is to forecast irradiance of every pyranometer up to 30 minutes. The training data is recorded from the pyranometers and the camera for two months every 10 seconds during daytime. In order to determine clouds projection on the ground, they apply a series of image processing algorithms.

2.2.1 Cloud detection

Firstly, they use Red-to-Blue Ratio (RBR) threshold for cloud detection which was first developed by Scripps Institution of Oceanography [13, 30] and is been used in many sky-imager-based forecast applications such as [8]. The RBR values close to 1 are usually cloud, and values very less than 1 are blue sky, since the blue channel which is in denominator dominates the red channel. However, since the RBR is not homogeneously distributed over the whole field of view, using a fixed global threshold for cloud detection brings a lot of misclassification for the areas close to the sun and also dark thick clouds or very transparent clouds. Therefore, they correct the RBR values based on clear-sky RBR values for each pixel. A Clear Sky Library (CSL) is created from images of one clear day. Then, the closest distance of current position and sun positions of CSL images is used to choose the reference RBR image map. This RBR map is used in correction formula 2.1 to decrease RBR threshold in circumsolar area to counter effect of sun saturation there. The correction also decreases RBR threshold for dark areas and increases it for bright pixels of image in order to detect thick and thin clouds.

$$R_{mod,i,j} = R_{orig,i,j} - R_{CSL,i,j} \times (a \times S - b \times (I_{i,j} - 200)) \quad (2.1)$$

Where $0 < S < 1$ is the average pixel intensity in circumsolar area. For more detailed discussion on results, they also apply a image-based cloud type classification using several visual cloud characteristics , and classify them into 7 different cloud types.

2.2.2 Image un-distortion

Since the raw image is from a fisheye lens, they apply a transformation to project it into geometric coordinates for convenience in other calculations. For that, intrinsic parameters of camera are determined using Scaramuzza Matlab toolbox [27] which solves a fifth-degree polynomial function of point-mapping between fisheye image and plain image. The extrinsic parameters are calculated as the best rotation which matches position of sun reprojection (derived mathematically) and sun position in the image. They calculate sun zenith and azimuth by using solar geometry2 algorithm[3].

2.2.3 Shadow mapping

In this step, shadow of cloud pixels are projected on the ground. For this, besides incidence and azimuth angle of every cloud pixel (which is derived using camera calibration function), cloud base height is needed. The cloud base height is estimated using a ceilometer for every point in time. However, to smooth the data, median of last 30 measurements is used. Even though the ceilometer supports multi-layer clouds as well, in this work they only use the lower-level cloud height. The distance of every cloud pixel to the camera is derived using $d_{i,j} = h \times \tan(\theta_{i,j})$. Given the distance $d_{i,j}$, incidence angle $\theta_{i,j}$, pixel's azimuth angle $\varphi_{i,j}$ and current sun position angles, horizontal distance of the cloud's shadow on the ground from the camera is calculated using Eq. 2.2.

$$\begin{aligned} dx_{i,j} &= h \times \tan(\theta_{i,j}) \times \sin(\varphi_{i,j}) + h \times \tan(\theta_{sun}) \times \sin(\varphi_{sun}) \\ dy_{i,j} &= h \times \tan(\theta_{i,j}) \times \cos(\varphi_{i,j}) + h \times \tan(\theta_{sun}) \times \cos(\varphi_{sun}) \end{aligned} \quad (2.2)$$

The shadow pixel points are mapped to a grid of 20km to 20km with resolution of 20m, and coordinates are interpolated if the shadow map resolution is lower than grid resolution, otherwise the central pixel of the dense shadow area is used for that grid point. Finally, a Gaussian filter is applied to smooth the cloud edges for more realistic result.

2.2.4 Irradiance retrieval

Upon determining the shadowed and sunny grid points on the ground area of experiment, they use the histogram of clear-sky index (k^*) to estimate the GHI irradiance. The clear sky index is defined as ratio of measured global horizontal irradiance GHI_{meas} and a clear sky reference value GHI_{clear} (Eq. 2.3).

$$k^* = \frac{GHI_{meas}}{GHI_{clear}} \quad (2.3)$$

Clear sky irradiance is obtained from the mode of Dumortier [9] and turbidity values of Bourges [5] which is validated according to Ineichen's work [11]. For adapting to smooth changes of irradiance caused by factors other than clouds, this histogram is generated with measurements of last 30 minutes. As it is shown in figure ?? this histogram usually has two peaks which correspond to sunny and shadow states on the specified point of ground. Now, for every point on the ground, based on its state (shadow, no-shadow), the corresponding k^* is used from the peaks of the histogram to estimate GHI following Eq. 2.4. In case two distinct peaks could not be detected in histogram due to homogeneous irradiance condition, default values of 0.4 and 1, have been assigned for shadow and no-shadow states.

$$GHI = k_{hist}^* \times GHI_{clear} \quad (2.4)$$

2.2.5 Irradiance forecast

For forecasting the cloud map, they use optical flow algorithm (Lucas-Kanade) on cloud edges and corners (Shi-Tomasi method) in the past 2 minutes to extract clouds' motion vector. Then, by applying that motion vector to current cloud state, cloud map at different time horizons is estimated, and ray tracing from sun position at that times through cloud maps gives the shadow map on the ground. After determining shadow or no-shadow states for points on the ground, forecast GHI is estimated using Eq. 2.4 . They compare results of using GHI histogram from the nearest pyranometer station versus using only the pyranometer close to the camera as a representation for the whole area. These comparison is separately done for different cloud types, and the results shows for cumulus clouds using one pyranometer close to camera is enough to forecast irradiance for up to 2km radius. However, the forecast skill is highly varies depending on cloud types and overall does not do better than persistent model which uses median of past several minutes' irradiance.

2.3 Retrieval of direct and diffuse irradiance from sky images

In another recent work, T. Schmidt et al.[28] aims to estimate components of irradiance (direct, diffuse) instead of just GHI from the sky images using machine learning on image features. They hope this kind of irradiance detail helps in estimating GHI in cloudy and partial sunny states more accurately. The Experimental setup includes a fish-eye camera with sample rate of 10 sec and a pyranometers package next to it that records direct, diffuse and global irradiance every second.

2.3.1 Image features

As image features they calculate several local and global features including:

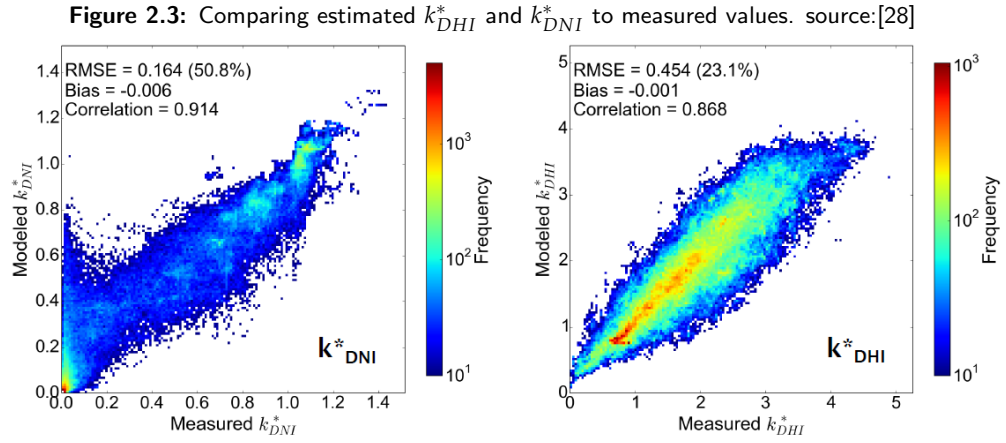
- Texture properties of the Grey Level Co-occurrence Matrix (GLCM)
- Color statistics (RGB space)
- Inter-color relations (e.g. Red-Blue-Ratio)
- Statistics of saturated pixels in circumsolar area in RGB and HSV color space
- Derived features like cloud coverage
- Solar elevation angle

For prediction, two k-nearest-neighbor (kNN) models are trained that estimate the clear sky index of diffuse horizontal (k_{DHI}^*) and direct normal

(k_{DNI}^*) components which are defined as ratio of each component to their clear-sky values obtained from Ineichen's algorithm[11]. Since the initial feature list contained 37 features, for reducing computation time and avoiding over-fitting they apply a feature selection using decision tree feature ranking algorithm to choose the optimal feature set among them.

2.3.2 irradiance estimation

The result of KNN predication for DHI and DNI clear-sky indexes compared to measured values shows a correlation around .085 in Figure 2.3. In forecast applications, GHI can be derived from predicted irradiance components using 1.1. However, predictability of some of the used image features such as color statistics in this method is not robust enough. This leads to some errors in irradiation forecast.



Chapter 3

Estimating Diffuse Horizontal Irradiance (DHI) from sky image

In this chapter, a new approach that we developed for estimating DHI from sky images is explained. First, our experimental setup and data is presented, then we talk about clear-sky model used here, and why estimating DHI is very important in predicting GHI. Furthermore, DHI estimation using the irradiance sensors and also sky-images is discussed. Afterwards, machine learning regression methods for obtaining DHI from sky-image are studies. Finally, the general strategy for power prediction in a photovoltaic power plant using the forecast irradiance components is proposed.

3.0.1 Experimental Setup

This study is conducted on one of the photovoltaic(solar) power plants operated by ABB company. This PV plant which is located in Caviglia region in Italy is chosen as a pilot site for the "forecasting power prediction using sky-imagery" project. Therefore, it is equipped with the following instruments for recording irradiation and sky images:

- A customized wide-angle high resolution (4MP) camera system with a fisheye lens covering 185 degrees of field of view. The camera is in a packaging attached to a pole on rooftop of the building next to the site. Figure 3.1 shows the camera and its position next to the PV plates.
- Two GHI pyranometer (irradiance sensor); located close to the camera on rooftop. One of the sensors is horizontally facing sky, and the other one facing north with around 40 degrees angle to the horizontal plane. It's worth mentioning that the PV plates are titled to south with a fixed angle around 30 degrees to get more sun exposure. Th sensors are depicted in Figure 3.2.
- A thermometer for recording the temperature at the site.

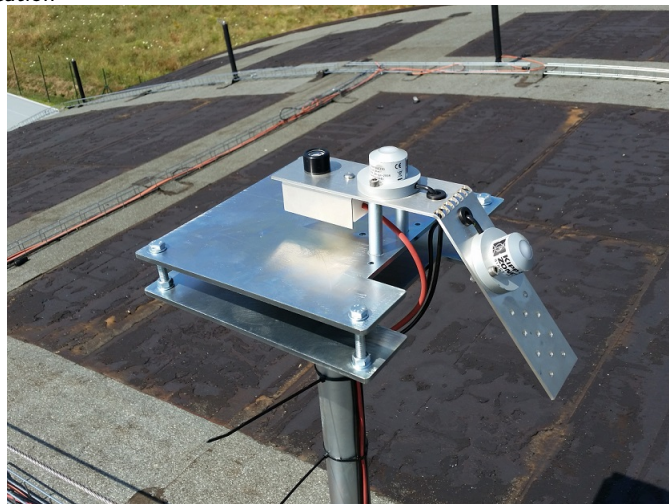
3. ESTIMATING DIFFUSE HORIZONTAL IRRADIANCE (DHI) FROM SKY IMAGE

- A PC which is connected to the camera, pyranometers and thermometer via their software interfaces in order to configure sample rates and store taken images and irradiance measurements. The data of power generated by the PV plant is also sampled and stored for every day. All the data recorded during each day is been automatically transfered to the company samba sever at midnight using a control software running on the PC.

Figure 3.1: wide-angle camera system used at Carviglia site



Figure 3.2: Two pyranometers (one horizontal, one 45 degrees tilted to north) located close to the camera location



3.0.2 Acquired Data

The camera system captures several images from the whole-sky every 8 seconds with different narrow exposure ranges. These images which are labeled according to capture time, are combined to create an HDR (High Dynamic Range) image for every sample time. The original narrow exposure images are generally deleted except the images at each hour time (i.e. around 7:00, 8:00, 9:00 etc.). Since capturing images at night is not useful for power prediction applications, camera is instructed to only take pictures during daytime (i.e. sunrise to sunset) which is obtained for that specific location for each day using mathematical models. Nevertheless, according to the captures images, this is not enough and still there are some black images taken at the minutes before sunrise and after sunset. Therefore, while processing the images on the application side, we exclude those images using a threshold on average pixel intensity of each image. This threshold is determined empirically. The images are further filtered against a sky mask which is been designed to exclude nearby mountains and buildings in the field of view and also limit the field of view to 170 degrees since transforming the points which are further in horizon is not accurate enough and the sun light is not negligible when the sun is in those points. It is worth mentioning that using HDR images in the image processing step is one of the key advantages of this study to related works specifically [28].

The pyranometers measure GHI values with the sample rate of 6 seconds. The temperature is also recorded with the same sample rate. However, the generate power is measured and logged every 3 seconds. These different sample rate make it necessary to interpolate the available data to find the estimated data for a time which there is no data available. Therefore, we can assign total irradiance, temperature and generated power to any given image using its capture time. This data acquisition setup has been running since 7th July 2015 until present. However, there are some short periods of time (usually lasting several days up to two weeks) which one of the sensors (pyranometers or the camera) had problems or the power plant was not working to produce power data. Considering the relatively small sample rate (less than 8 seconds), the amount of recorded data is big enough to make those off-days negligible in data processing steps. The data used in this study spans from 15th July to 10 February, meaning that many summer, autumn and winter days are available in the dataset to make it a good representation for the whole year.

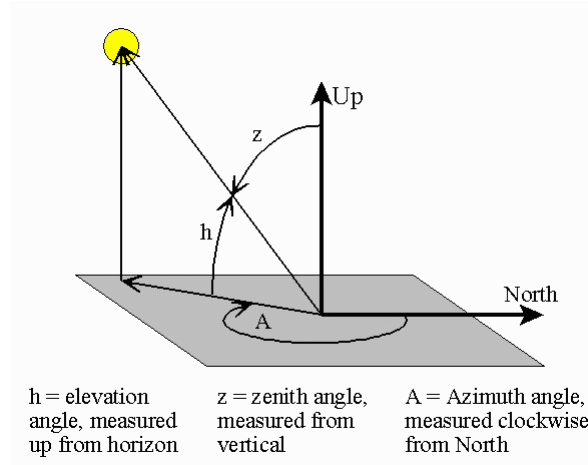
3.0.3 Sun positions and sun states in image

Knowing position of sun is very important both in cloud segmentation and in irradiance estimation. First of all, since our images are from a wide-angle fisheye lens, they need to be transformed into geometric coordinates

3. ESTIMATING DIFFUSE HORIZONTAL IRRADIANCE (DHI) FROM SKY IMAGE

by an un-distortion algorithm to make them ready for further image processing steps including sun position, cloud segmentation and etc. This can be done by multiplying the raw image coordinates to camera transformation matrix which consists of intrinsic and extrinsic parameters of the camera system. As described in section 2.2.2, intrinsic parameters are calculated using image of a chessboard[27] and extrinsic parameters are estimated using Kabsch algorithm[14] based on position of the sun appeared in the image versus the expected position of sun in image. The theoretical sun positions are calculated for the location of our PV plant site at every image time-stamp using NREL algorithm [23] in Matlab represented in unit sphere polar coordinates. As shown in Figure 3.3 this position is described as two angles (zenith and azimuth) which are converted to Cartesian coordinates using sphere to Cartesian conversion and later on are scaled to the image size to correspond with an image pixel. That pixel will be assumed as center of the sun.

Figure 3.3: Sun position angles. source:[1]



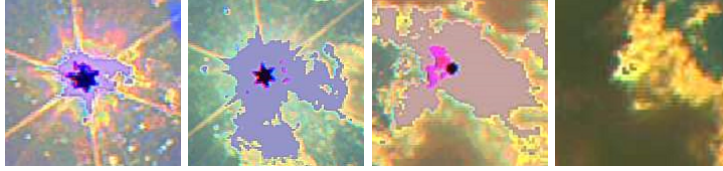
In cloud segmentation, which is not the focus of this study, sun position is used to treat pixels close to sun according to different threshold than other pixels. Furthermore, a sun state detection inspect the area around sun position to classify sun state in the image into following 4 categories:

- `sun_flag=4`: indicating the sun is visible in the image and it appears as a star shape emitting 6 symmetric strong rays.
- `sun_flag=3`: indicating the sun is visible in the image and it appears as a star shape but with 5 or less symmetric strong rays.
- `sun_flag=2`: indicating the sun is visible in the image but it does not appear as a star shape. Instead, it appears as a small black dot with no strong rays.

-
- sun_flag=1: indicating the sun is not visible in the image and it is either covered by clouds or the sun position is out of field of view in the image.
 - sun_flag=-1: indicating there is an unexpected situation around sun position, for example star shape sun is detected far away from expected sun position which could be because of strange cloud formations.

One sample for each one of these categories is depicted in Figure 3.4. We are able distinguish between this states thanks to HDR images, otherwise this fine classification would not be possible.

Figure 3.4: Different sun states, from left to right: sun flag=4, 3, 2, 1 respectively.



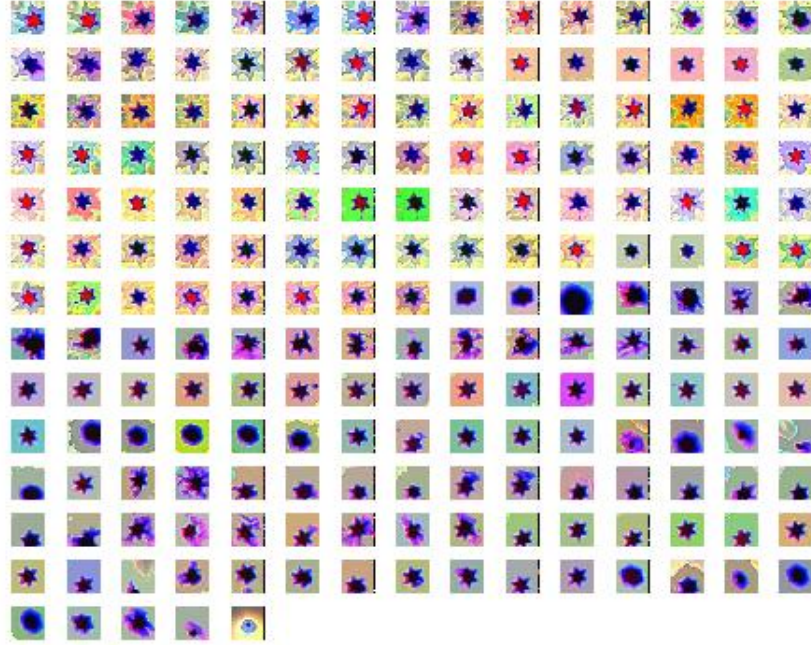
The variation of sun color in the images is so much that using Support Vector Machine for detecting sun is works very poorly. This issue is visible in Figure 3.5 which shows some clear sun samples from the images. Therefore, another approach using gray-scale images and geometrical symmetry detection is employed.

Using these sun states along with the sun position in cloud segmentation algorithm, will lead to better segmentation results close to sun that is particularly a difficult area for cloud segmentation due to highly saturated pixels with different sky or cloud colors. In irradiance estimation which is the main focus of this study, sun position is used to create two specific feature vectors in a bounding circle around the sun. These features are explained section 3.0.6 in detail.

3.0.4 Clear-sky irradiance model

Before dealing with the problem of irradiance estimation from cloudy images we should first know how much the irradiance would be at any time in clear sky conditions when there is no cloud in the sky at all. Fortunately, there are a few number of models which provide irradiance components at each given time for many locations on the Earth. In this study we investigate two of the these methods, Ineichen [9] and McClear [17]. Both of these methods use location coordinates (latitude, longitude) and query time (consisting of year,day,month,day,minute,second) to calculate sun angles internally and return the irradiation based on optic formulas which determine how much sun should reach the ground in as direct sunlight and

Figure 3.5: Variation of sun appearance in the images.



how much should be scattered in hemisphere and forms the diffuse irradiance. The amount of scattered sunlight is varying throughout the year and also depends on the location, ground albedo¹ and aerosol parameters including pressure, ozone column content, water vapour column content, optical depth and Angstrom coefficient.

Ineichen method

Ineichen model statistically and physically relates some irradiance measurements to the aforementioned parameters as Linke Turbidity profiles which are available for different locations. We used the default Linke Turbidity values which come as a separate file in PV-Lib toolbox [16] in Matlab and is representative for most locations in Europe. However, one might need to use other appropriate Linke Turbidity profiles for other locations to get more accurate results.

¹The fraction of solar energy (shortwave radiation) which is reflected from the Earth back into space. It is a measure of the reflectivity of the earth's surface. Ice, especially with snow on top of it, has a high albedo.

McClea r method

On contrary to this approach, McClea r uses a fully physical model that exploits recent aerosol properties, total column content in water vapour and ozone produced by the MACC project (Monitoring Atmosphere Composition and Climate). The MACC project, funded by the European Commission, uses data of many Numerical Weather Prediction (NWP) centers distributed around the world to provide a global aerosol property forecasts together with physically consistent total column content in water vapour and ozone. In other words, since McClea r uses synthetic data of NWP's, it does not depend on any local atmospheric observations for irradiance prediction. For the sake of speed, McClea r irradiance estimates are pre-computed for the the location of these measurement centers and are interpolated for all other point on th Earth using a look-up table approach. Of course the closer we are to one of these measurement centers, the more accurate McClea r estimate will be. The McClea r irradiance estimates are available worldwide for every minutes from 2004 to present with 2 days lag under this web service [2]. This means that if we want to get the irradiance for today, we need to interpolate data of several days or weeks before to get an estimate for the current time. Nevertheless, one can use the original data of past 2 days for current time as well, since irradiance does not change considerably in two days.

Comparison of simulation results

To evaluate performance of these two clear-sky models we choose several days which have a significant clear part during the day (just because complete clear days are very rare). The chosen days should be from different months of year in order to represent performance of models for the whole year better. After observing the irradiance logs and images for verification, the following days were selected for comparison: 2015/07/19, 2015/08/03, 2015/09/21, 2015/10/24, 2015/11/24 and 2016/02/05. As an example, simulated GHI of three days are plotted in Figure 3.6 next to the observation.

As it can be seen, both methods can predict the shape of irradiance curve very accurately around the year, but both have biases in the result such that the simulated values are always smaller than observed ones. During the summer days, McClea r and Ineichen results are biased almost equally, and as we get closer to winter days, the bias of McClea r gets smaller and bias of Ineichen get slightly bigger. Figure 3.7 shows the correlation of simulated GHI of both models plotted with respect to the actual measurements of all of the examined days.

This plot again shows that the bias of McClea r and Ineichen for large values of GHI is the same, and for small GHI values McClea r result is closer to the observed irradiance. however, it also suggests that since Ineichen methods

has lower variation in terms of bias, this bias can be compensated with a scaling factor more easily than the bias of McClear which shows larger variation throughout the year. Note that, the outliers in this figure are representing cloudy times. Furthermore, this bias of Ineichen method is strongly related to the Turbidity factors that are neglected in our case by using the default values. It would be not surprising to see smaller bias if one uses Turbidity factors which are verified for the location of PV plant. The correlation of DNI and DHI values of both methods are illustrated in Figure 3.8. One can see that DNI is predicted much higher in Ineichen method and DHI is also simulated slightly higher than McClear values. This behavior is intensified during autumn for DNI, but it is not varying a lot for DHI. Since we do not have observation values of DHI and DNI, we cannot compare correlation of methods' results to the observed values in this figure. However, we can hypothesize than during a complete sunny day, if at a very short time (i.e. seconds) a cloud covers the sun completely, the direct irradiance (DHI) is almost zero and all the irradiation only comes from DHI source which is the scattered light in the sky. Therefore, we can use our GHI irradiance observations as DHI and compare it to the DHI simulated values for that particular time.

In Figure 3.9 this has been shown for a moment around 13:00 which sun is occluded by a small thick cloud and therefore, GHI is dropped rapidly to a value close to 150 which is very close to simulated DHI values from McClear and Ineichen methods. Also, there is no other cloud in the sky to influence the DHI component. Furthermore, the simulated DHI values are very close to each other at any time and resemble the shape of GHI relatively well during whole day. Thus, we can conclude that these models can predict DHI with good accuracy, and since DNI is a function of GHI and DHI according to Eq 1.1, DNI values calculated from McClear and Ineichen models are also accurate enough and for our application. This hypothesize has been verified by looking at many other data points where GHI observation gets very close to DHI simulation values and there is a cloud obstructing the direct sun light.

We know that DNI should be always larger than DHI during clear-sky condition. Looking at Figure 3.10, pattern of changes in the simulated values of DNI and DHI during a day confirms this condition too.

Choosing the clear-sky model

As we discuss in this chapter, McClear and Ineichen models both predict the clear-sky irradiance components relatively accurately, however the bias for Ineichen method is more robust and manageable than McClear bias. Furthermore, obtaining the McClear values requires downloading the irradiance files from their web service since there is not offline library for calculating

them. All this considered together, we decided to use Ineichen as the clear-sky irradiance model for this study. The scaling factor for compensating the bias in Ineichen is set to 1.08 which is empirically calculated based on several clear day observations.

3.0.5 Estimating diffuse from pyranometers

Using clear-sun-flag and clear-sky DNI for calculating diffuse from tilted plate

ii. clear-sky DNI is a good approximation for actual DNI. Showing some clear or cloudy days to prove this point.

We only consider images with complete visible sun or no sun at all. ii. Comparing tilted diffuse with diffuse from the main irradiance plate, (then correcting tilted diffuse??). Showing its robustness. Show the result of relation of DNI DHI to irr1 irr2 with plots.my recent work.

3.0.6 Key image features affecting DHI

iii. Investigate different parameters which affect diffuse

Discuss cloud coverage geometrical polar feature around sun and also the whole image

Discuss saturation detection algorithm and its results

iv. Show their correlation to diffuse, discuss cases based on images and corresponding irradiance components

3.0.7 Learning the relation between image features and DHI

Regression, or svorim method for estimation Compare to other method to regress, or penalty [no result]

3.0.8 Translating irradiance to power

Estimating shadow ratio on the plant site

refer to Figure 3.11. Projecting plant coordinates into the sky using assumed cloud height Adaptation to smooth changes of power by using the histogram of irradiance over past 30 minutes

3. ESTIMATING DIFFUSE HORIZONTAL IRRADIANCE (DHI) FROM SKY IMAGE

Figure 3.6: GHI estimation of McClear vs Ineichen vs observation

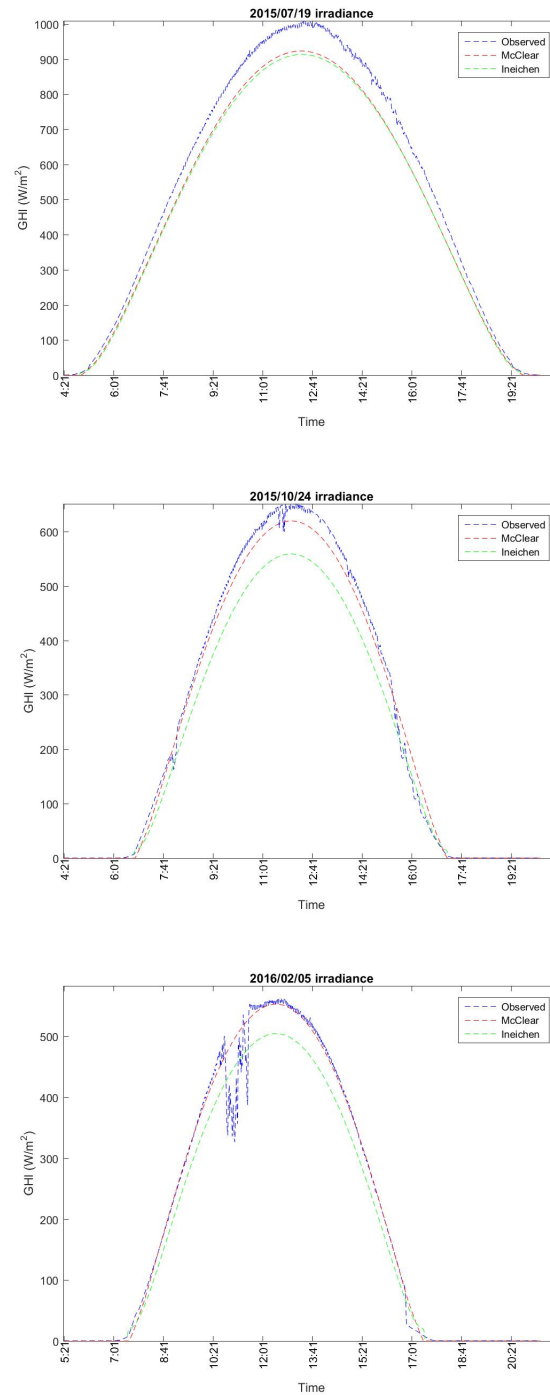
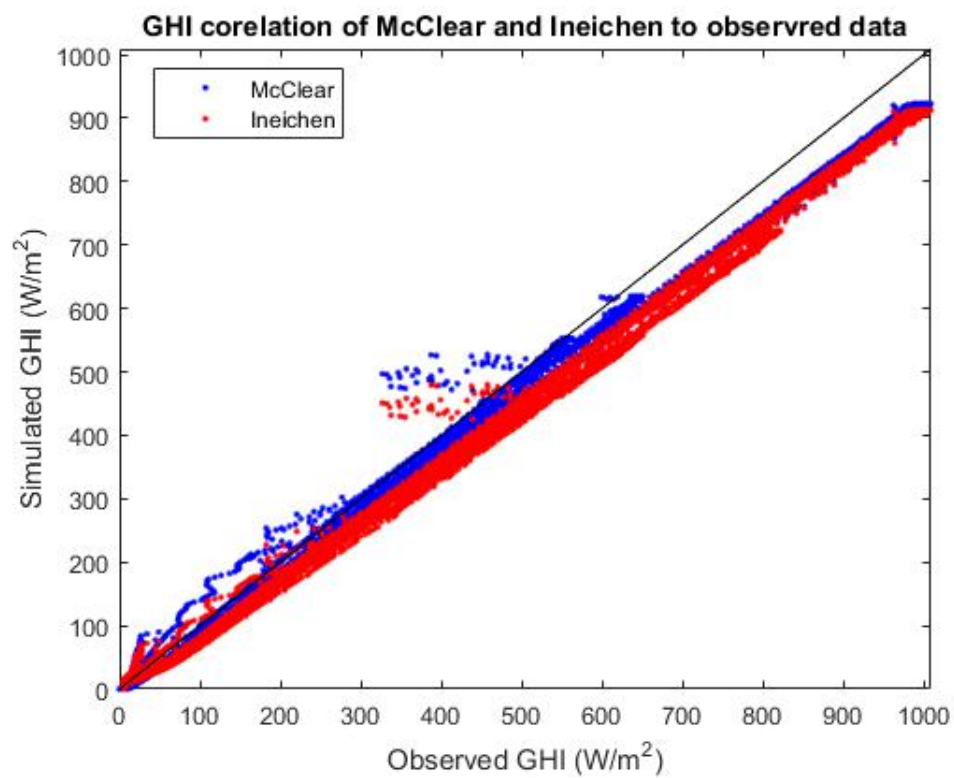


Figure 3.7: Correlation of GHI simulation of McClear and Ineichen with observations on selected days



3. ESTIMATING DIFFUSE HORIZONTAL IRRADIANCE (DHI) FROM SKY IMAGE

Figure 3.8: DNI and DHI correlation of McClear to Ineichen

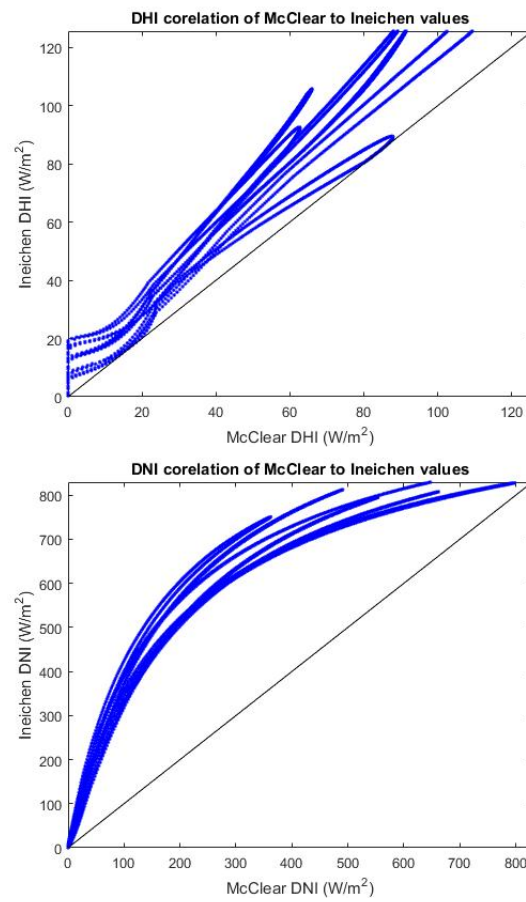


Figure 3.9: Comparison of DHI of McClear and Ineichen to observed GHI

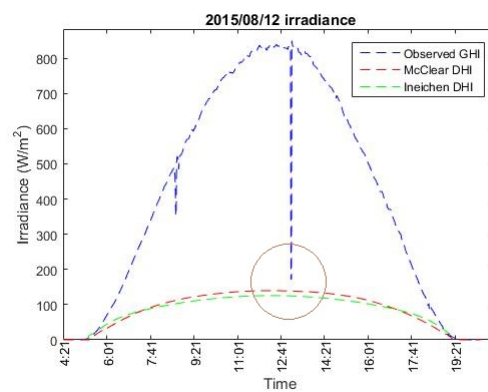


Figure 3.10: Variation of irradiance components during day

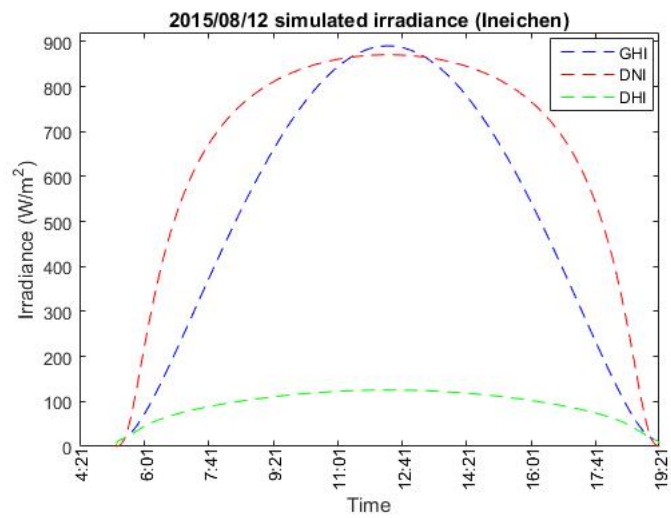
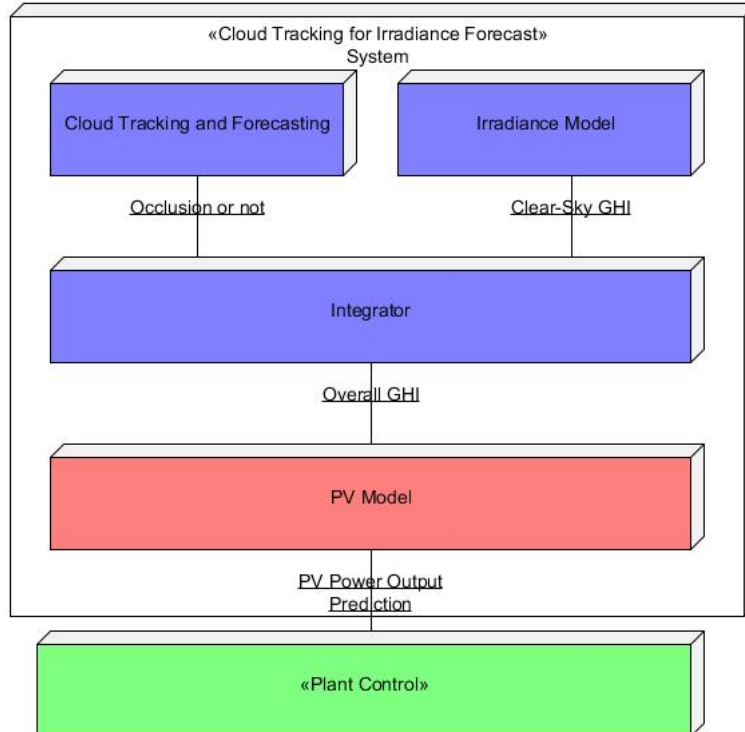


Figure 3.11: Overall simple system design for power prediction



Chapter 4

Results

Chapter 5

Discussion

Chapter 6

Conclusion

Dummy text.

Bibliography

- [1] Shperical model for position of sun. <https://pvpmc.sandia.gov/modeling-steps/1-weather-design-inputs/sun-position/>. Accessed: 2016-03-10.
- [2] Shperical model for position of sun. <http://www.soda-pro.com/web-services/radiation/cams-mcclear>. Accessed: 2016-03-20.
- [3] P. Blanc and L. Wald. A Library for Computing the Relative Position of the Sun and the Earth. Technical report, 2011.
- [4] J. Borkowski, Chai A.-T., Mo T., and Green A. E. O. Cloud effects on middle ultraviolet global radiation. 25(4):287–301, 1977.
- [5] B. D Bourges. Yearly variations of the Linke turbidity factor. page 61–64, 1992.
- [6] A. Cazorla. *Development of a Sky Imager for Cloud Classification and Aerosol Characterization*. PhD thesis, Universidad de Granada, Granada, Spain, 2010. PhD thesis.
- [7] A. Cazorla, F. J. Olmo, and L. Alados-Arboledas. Development of a sky imager for cloud cover assessment. 25:29–39, 2008.
- [8] C. W. Chow, B. Urquhart, M. Lave, A. Dominguez, J. Kleissl, J. Shields, and B. Washom. Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed. 85:2881–2893, 2011.
- [9] M. Fontoynt, D. Dumortier, D. Heinemann, A. Hammer, J. Olseth, A. Skarveit, P. Ineichen, C. Reise, J. Page, L. Roche, H. G. Beyer, and L. and Wald. Satellight: aWWWserver which provides high quality daylight and solar radiation data for Western and Central Europe. page 434–437, 1998.

- [10] A. Hammer, D. Heinemann, E. Lorenz, and B. Lückehe. Shortterm forecasting of solar radiation: a statistical approach using satellite data. 67:139–150, 1999.
- [11] P. Ineichen. Long Term Satellite Hourly, Daily and Monthly Global, Beam and Diffuse Irradiance Validation. 2013.
- [12] R. H. Inman, H. T. C. Pedro, and C. F. M. Coimbra. Solar forecasting methods for renewable energy integration, *Prog. Energ.* 39:535–576, 2013.
- [13] R. Johnson, W. Hering, and J. Shields. Automated Visibility and Cloud Cover Measurements with a Solid State Imaging System. 1989.
- [14] Wolfgang Kabsch. A solution for the best rotation to relate two sets of vectors. 32:922–930, 1976.
- [15] J. Kühnert, E. Lorenz, and D. Heinemann. Satellite-based irradiance and power forecasting for the German energy market. page 504, 2013.
- [16] Sandia National Laboratories. PVLIB Toolbox v1.2 for Matlab. 2012.
- [17] M. Lefèvre, A. Oumbe, P. Blanc, B. Espinar, B. Gschwind, Z. Qu, L. Wald, M. Schroedter-Homscheidt, C. Hoyer-Klick, A. Arola, A. Benedetti, J. W. Kaiser, and J.-J. Morcrette. McClear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions. 6:2403–2418, 2013.
- [18] E. Lorenz and D. Heinemann. Prediction of solar irradiance and photovoltaic power. 1:239–292, 2012.
- [19] E. Lorenz, D. Heinemann, and A Hammer. Short-term forecasting of solar radiation based on satellite data. page 841–848, 2004.
- [20] F. J. Olmo, A. Cazorla, L. Alados-Arboledas, M. A. Lopez-Alvarez, J. Hernandez-Andres, and J. Romero. Retrieval of the optical depth using an all-sky CCD camera. 47:182–189, 2008.
- [21] R. Perez, E. Lorenz, S. Pelland, M. Beauharnois, G. Van Knowe, K. Hemler Jr., D. Heinemann, J. Remund, S. C. Müller, W. Traunmüller, G. Steinmauer, D. Pozo, J. A. Ruiz-Arias, V. Lara-Fanego, L. Ramirez-Santigosa, M. Gaston-Romero, and L. M. Pomares. Comparison of numerical weather prediction solar irradiance forecasts in the US, Canada and Europe. 94:305–326, 2013.

-
- [22] G. Pfister, R. L. McKenzie, J. B. Liley, A. Thomas, B. W. Forgan, and C. N. Long. Cloud coverage based on all-sky imaging and its impact on surface solar irradiance. 42:1421–1434, 2003.
 - [23] I. Reda and A. Andreas. Solar position algorithm for solar radiation application. Technical report, 2003.
 - [24] G. Reikard. Predicting solar radiation at high resolutions: a comparison of time series forecasts. 83:342–349, 2009.
 - [25] J. Sabburg and J. Wong. Evaluation of a Ground-Based Sky Camera System for Use in Surface Irradiance Measurement. 16:752–759, 1998.
 - [26] S. Sayeef, S. Heslop, D. Cornforth, T. Moore, S. Percy, J. Ward, A. Berry, , and D. Rowe. Solar Intermittency: Australia’s Clean Energy Challenge: Characterising the Effect of High Penetration Solar Intermittency on Australian Electricity Networks. 2012.
 - [27] D. Scaramuzza. OCamCalib: Omnidirectional Camera Calibration Toolbox for Matlab. 2014.
 - [28] T. Schmidt, J. Kalisch, and E. Lorenz. Retrieving direct and diffuse radiation with the use of sky imager pictures. 17:12–17, 2015.
 - [29] T. Schmidt, J. Kalisch, E. Lorenz, and Heinemann D. Evaluating the spatio-temporal performance of sky-imager-based solar irradiance analysis and forecasts. 16:3399–3412, 2016.
 - [30] J. E. Shields, M. E. Karr, T. P. Tooman, D. H. Sowle, and S. T. Moore. The whole sky imager – a year of progress. page 23–27, 1998.
 - [31] R. Tapakis, A.G. Charalambides, M.D. Moldovan, and B.G. Burduhos. Cloudy sky irradiance model using sky images. 2015.
 - [32] R. Tapakis, A.G. Charalambides, M.D. Moldovan, and B.G. Burduhos. Effect of clouds on solar irradiance. Technical report, Universitatea Transilvania din Braşov, 2015.
 - [33] S. R. West, D. Rowe, S. Sayeef, and A. Berry. Short-term irradiance forecasting using skycams: motivation and development. 110:188–207, 2014.
 - [34] B. Wolff, E. Lorenz, and O. Kramer. Statistical learning for short-term photovoltaic power predictions. 2013.
 - [35] D. Yang, Z. Dong, T. Reindl, P. Jirutitijaroen, and W. M. Walsh. Solar irradiance forecasting using spatiotemporal empirical kriging and vector autoregressive models with parameter shrinkage. 103:550–562, 2014.