Real Time Cloud Tracking via Sky Imagery for Power Output Prediction of Solar Plants

by

Burak Zeydan

Submitted to the Department of Computer Science
in partial fulfillment of the requirements for the degree of

Master of Science in Computer Science

at the

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

March 2014

Author . . . .
Department of Computer Science
March 14, 2014

Certified by. . . . .
Prof. Sabine Süsstrunk
sabine.susstrunk@epfl.ch
Thesis Supervisor

Certified by. . . . .
Dr. Mehmet Mercangoez
mehmet.mercangoez@ch.abb.com
Thesis Supervisor

Real Time Cloud Tracking via Sky Imagery for Power Output Prediction of Solar Plants

by

Burak Zeydan

Abstract

In this thesis the research and development procedure of a novel photovoltaic (PV) power output estimation system is described. Since the power output of a PV panel is closely related to the incident global horizontal irradiance (GHI) on it, GHI level on the PV plant area should be predicted. The method utilized for the prediction consists mainly of three steps: estimation of a baseline power input to the system, the estimation of possible losses on this power input, and the integration of the estimated drops to the baseline. Development of the system is done in a modular fashion, where each module implements a different step of the method. The first module generates an estimate of the clear-sky GHI level for the future. This gives the maximum level of the power input that will be incident on the panels. The second module generates estimates on the occlusions caused by clouds, using the sky-imager developed during this study. If an occlusion on panels is predicted, during the integration step, the clear-sky GHI level is decremented by a suitable amount. For tracking the clouds, a frame-based persistence method is utilized. For any frame, first clouds in the frame are detected using the pixel-wise ratio of red channel value to blue channel value (r/b ratio), and a binary image that just captures detected clouds is obtained as a result of this image. Since this image is much more detailed than needed, in terms of the edges, a median filter is applied on it to smoothen the edges. Then a clustering algorithm is applied to obtain the contours of clouds. For each contour, a rectangular bounding box is fitted, and the center of this box is taken as the representative centroid of the cloud object. Representing the previous frames, are the tracks containing the Kalman Filters and many other statistics including centroid, area, and bounding box of the trails of cloud objects observed up to the current frame. The locations of these tracks are predicted, to be compared with the ones observed in the current frame, using their Kalman Filter estimates. Then, using these predicted locations, the current detections are matched to tracks through the Hungarian Algorithm. After this matching, the tracks are updated with the observation that is assigned to them. During the update, the Kalman Filter, the contour points, bounding box, centroid, and all the other information kept in a track are updated. Then finally, for all the tracks, their state several minutes later is predicted using their Kalman Filter. At this stage of the tracking algorithm, a 3D model of the clouds and the sun for all the minutes in the prediction horizon is present.

Using this model, an occlusion can be foreseen hence a power-input-drop prediction for all the minutes in the prediction horizon can be estimated. Afterwards the integrator module, integrates drops and baseline, and comes up with a numeric value for the future GHI. In this thesis the details of the development of this system, along with the associated background information, will be covered.

Thesis Supervisor: Prof. Sabine Süsstrunk
Title: sabine.susstrunk@epfl.ch

Thesis Supervisor: Dr. Mehmet Mercangoez
Title: mehmet.mercangoez@ch.abb.com

Acknowledgments

To be added

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

Energy generation using photovoltaic (PV) cells is becoming increasingly popular on both stand-alone PV systems and grid-connected PV plants, and indeed this popularity is expected to further increase in the future [83]. It is an obvious fact that the power output of a PV plant is related to the solar irradiation level, which varies with varying atmospheric conditions [18, 60, 83]. As the large-scale grid connected PV plants are increasing in number, necessity of the studies regarding problems that can be related to such a variable power production scheme become obvious [83].

Random variations in the power output of a grid-connected plant, causes great complications in the management of regional power systems [83]. With randomly fluctuating power inputs to the grid, security and stability issues can arise in the regional power systems [62, 83]. Because of these issues, some regulations for stand-alone plants concerning the power production to the grid are expected to come into force. The most probable ones being some restrictions on the ramp rate of the produced energy prior to its integration into a grid [57]. To be able to meet such requirements, a plant or a grid operator would require predictions on the future level of solar irradiation [17, 31, 51, 57, 60, 83].

The solar irradiation forecasting methods can be classified, in terms of the models they utilize, into three main classes: empirical models, physical models, and statistical models [60, 83]. Solar irradiance is a measure of the electromagnetic energy reaching from the

sun to a surface per unit time per unit area [23]. This energy decays as it travels due to the inverse square law [23]. Also, when traveling through a medium, due to absorption, scattering, and reflection of the molecules, the amount of decay can increase [23, 83]. Difference between the amount of irradiance reaching, the top level of atmosphere, and the ground is owing to this fact. Empirical models try to estimate this decay rate fitting an empirical correlation on the recorded data [72], that contains the global location, altitude, sunshine duration in hours, and, if possible, site related measurements [8]. Physical models are founded on complex non-linear mathematical equations describing the physical and dynamical state of the atmosphere [83]. Since these are complex models, it takes significant computational effort to solve them, and usually solutions are approximated using numerical methods, because of which they are also called numerical weather prediction models (NWP) [83]. Statistical models make use of the trends in the data. Many factors including: location, time of the year, time of the day, temperature, humidity, aerosol density of the atmosphere, and the cloud conditions determine the solar irradiation [18, 46, 60]. Essentially, statistical models try to discover the relation between: different parameter values measured within the same time period, or the same parameter value measured in consecutive time periods. The discovered relations are usually very hard to observe by a user, hence sophisticated methods such as Artificial Neural Networks (ANNs) and Markov chains are utilized [60]. Statistical properties of a data can be analyzed only in the presence of a big amount of data points; statistical models can only be used if there are sufficiently many data points [60].

The amount of drop on the incident solar irradiance is most seriously affected by cloud events [54]. Jayadevan et al.(2012) states that clouds cause more than 20 % of the drop on power production estimated for clear-sky conditions, and that the majority of energy loss is caused by cloud events that cause more than 70 % drop on the power production and lasts 1 to 6 hours [37]. This makes it obvious that, the predicting model should take into account the state of cloudiness. For instance NWP models are capable of forecasting cloudiness conditions several days ahead; however, cloud arrival time predictions are precise only up to several hours [37]; exact position and extend of individual clouds and cloud fields cannot be predicted using NWP models [67]. For using the information related to clouds, there are models that make use of satellite images [27], models that make use of a sensor grid on

the ground [51], and models that make use of sky cameras [63]. Using NWP and satellite cloud observations, several days ahead predictions with a spatial resolution around 100km^2 to 3km^2 are available [46, 52, 74]. But predictions on such coarse ranges are not very useful since they cannot predict spontaneous fluctuations, and hence they introduce significant errors [52, 74].

Depending on the purpose of use, predictions with different horizons are required. For instance, to determine the price on the energy, predictions up to 72h ahead are useful [57]. For operational efficiency purposes, daily and more recent predictions are needed. If the predictions on the future level of irradiance are provided, and the plant operator detects a drop on the predicted levels, there are several possible actions. Operator may start a fossil-fuel based plant, or charge a battery to later discharge, to compensate for the predicted drop. But since a fossil-fuel based plant consumes power before producing it, and a battery would need time to charge to the level required, the operator requires a notice of drop at least ten minutes before [37]. For this reason, an intra-hour prediction method is required.

Except for the cloud conditions, all of the factors affecting solar irradiation are locally stable with smooth transitions; changes on these can easily be foreseen using a simple time-series approach [57, 60]. What cannot be predicted using a simple approach, and in fact the focus of this project, is a local cloud event that can affect the direct sunlight incident on the PV panels, and hence that can cause serious fluctuations on the power production curve [26, 37, 51, 54, 65]. In a recent study by Jayadevan et al.,2012, it is mentioned that the majority of the cloud events cause a drop of 30-40 % and last less than a minute, stressing the significance of local, spontaneous cloud events in power intermittency [37]. For predicting such local events, there are several techniques utilized. The usage of a sensor grid on the ground that can track the movement of decay trend on the geographical landscape [51], usage of ground cameras to observe the sky and track clouds [34, 49, 50, 77–79], and usage of satellite images for tracking clouds [25, 76] are among the most popular ones. Satellite imagery is not suitable for our purpose, since the images obtained do not have a good temporal and spatial resolution, and they cannot catch spontaneous cloud events with the accuracy we require [12, 17, 28, 86]. Usage of a sensor grid is a novel approach but can get relatively costly, if there is no existing infrastructure, which is usually the case

around solar power plants. Usage of ground cameras, in single [12, 54, 86] or multi-camera setups [17, 28, 36], is increasingly popular since this is a less costly alternative, requiring only several cameras, and the temporal and spatial resolution of the system can be tuned up to the requirements of a particular site.

To summarize, the irradiance incident on a point on earth is directly related to the extraterrestrial irradiation incident on the corresponding point on the top of the atmosphere. The extraterrestrial irradiation is decreased during the propagation of the rays through the atmosphere. There are some models trained to predict that decrease during a clear-sky day. They take into consideration some averaged atmospheric factors and come up with some predictions on the level of irradiance. These models give us a baseline on the irradiance level for that time period; the actual observed irradiance can only be lower than this baseline due to a further drop caused by a cloud event. Our strategy for this project is, first getting the baseline clear sky irradince model to predict the maximum level of irradiance arriving on our panels, then getting a possible drop estimation on that level due to a cloud event. Obviously according to this strategy, we need two main functional blocks to implement our system.

- Baseline solar irradiation estimator block

- Cloud tracking and forecasting block

In the remainder of this chapter, information on the basic concepts that contribute to the understanding of the whole picture is provided.

1.2 Key Concepts

There are some key concepts that are important in establishing a solid understanding of the subject. In this section, brief explanation on these is going to be provided.

1.2.1 Solar Radiation

As already mentioned in section 1.1 on page 13, power production of a PV plant is closely related to the amount of solar power input incident on it. This power is measured by the

irradiance (W/m^2) [23]. The irradiance incident on a point at the top of the atmosphere, is refered to as the extraterrestrial irradiance, G_o [82]. Since this is energy originating from the Sun, depending on the position relative to the Sun, the amount of incident energy changes. Since the position of a point relative to the Sun changes with date and time, we can expect the dependence of incident irradiation to be a function of these parameters. In equation 1.1, where the extraterrestrial radiation is calculated, G_s denotes the solar constant which has the value of 1367 W/m^2, D denotes the day number in the year as a number in [1, 365], and $\cos\theta_z$ represents the cosine of the solar zenith angle, which takes into account the location and the time of the day information [82]. Extraterrestrial component represents the maximum amount of solar energy reaching the top of the atmosphere. This value decays as the sun rays travel through the atmosphere due to absorption, reflection, and scattering effects [23]. As can be seen in equation 1.2, the total amount of radiation on a surface G_t (W/m^2) is the sum of, the direct or beam radiation G_b (W/m^2), the diffuse radiation G_d (W/m^2), and the reflect radiation G_r (W/m^2) on the same surface [18].

$$G_o = G_s \left[1 + 0.033 \cos\left(360 \frac{D}{365} \right) \right] \cos\theta_z \qquad (1.1)$$

$$G_t = G_b + G_d + G_r \qquad (1.2)$$

The beam component is the amount of solar irradiance directly incident on the surface; it is related to the extraterrestrial radiation incident on the point corresponding to the same location on the boundary of the atmosphere, and the atmospheric conditions. This constitutes the major amount of energy, and it may decay due to interactions with the molecules or objects in the atmosphere. According to Duffie and Beckman,1985, the decay is caused either by scattering or by absorption [23]. What is referred to as scattering is the Rayleigh scattering due to air molecules, water vapor and droplets, and dust in the atmosphere [23]. The amount of scattering is dependent on the number of molecules and particle sizes [23]. The absorption is due to O_3, H_2O, and C_2O present in the atmosphere, and the band of absorption, in terms of the band of wavelength absorbed, is dependent on the type of molecule that interacts [23]. The amount of atmospheric molecules the solar light beams must traverse, is referred to as the air mass [81]. The amount of air mass depends on the solar position

of the observation point, and directly affects the reduction on the direct component [81].
For instance, when a cloud casts a shadow on a surface, because of the energy absorbed,
and reflected by this cloud the beam radiation reduces. Not all the decay in beam radiation
is lost. Diffuse and reflect components represent the part of irradiation that is not directly
coming from the sun, but are redirected due to reflections and scattering mentioned.

1.2.2   Cloud Characteristics

Chapter 2

Related Work

2.1 Introduction

In the previous chapter, the motivation of this study along with key concepts that are required by the reader for a better understanding of the subject is provided. In this chapter, a detailed literature survey on the methods that are investigated during the development of the overall system is presented.

2.2 Irradiance Prediction Methods

Irradiance is a measure of the amount of solar power input to a PV plant. To predict the power output, one can start by predicting the total amount of irradiance incident on the plant. In this part, the models are going to be investigated in two main groups:

- Models that do not take as input sky cloud conditions.

- Models that take as input sky cloud conditions from recent observations, and make use of this cloud information during prediction.

In this section, a brief explanation of these methods with examples from the literature is going to be provided.

2.2.1 Models that do not take as input sky cloud conditions

These methods do not make use of any means of sky imaging. They make use of the statistical or the physical phenomena to predict the solar irradiance. The most important examples of these are: numerical weather prediction (NWP) methods, and statistical methods.

Numerical Weather Prediction (NWP) Methods

These methods model the physical phenomena behind the irradiance level reaching the ground. As explained in 1.2.1 on page 16 total irradiance has several components: beam, reflect, and diffuse, and these components are a result of the interaction of extraterrestrial irradiance with the atmosphere before reaching the ground level. If one assumes that, he has the knowledge of the state of the atmosphere in terms of, for instance, temperature, humidity, pressure, and molecular density, using the associated models in physics, he would be able to estimate the amount of scattering and absorption a traversing solar beam would experience before reaching the ground. Furthermore, the values for these variables change in space and time according to the laws of dynamics and physics, so if we know their state and the rate of change of state at any time, we can predict their state for the next time. For instance, if one has the knowledge of the humidity values and the rate with which these change in space and time, he would be able to predict the future state of those values using the appropriate equation in thermodynamics [2, 83]. Applying the same procedure for several times, one can even obtain the state further into the future. These were the ideas that led to the development of the initial weather prediction models by Cleveland Abe, Vilhelm Bjerknes, and Lewis Fry Richardson [53]. Since there are many parameters to be predicted, in such a model the number of equations to be solved is big, and it is not possible to solve such a system with analytical methods [53]. Because of this reason, solutions to the system of equations can only be approximated by numerical methods, which is why these methods are called numerical weather prediction methods [83]. GFS model (2003) [1], ECMWF (2010) [3], WRF (2010) [4], MM5 (1995) [24], and ARPS (1995) [84] are some of the widely known NWPs.

NWPs predict weather related data at global (GFS, ECMWF) and regional (WRF) di-

mensions [67]. Since the computation required is intense, there are some limitations on the temporal and spatial resolution of these methods. Predictions are not done for all the possible locations, but for a grid of locations. The distance between grid points determines the spatial resolution of the model. Moreover the frequency with which the predictions are generated determines the temporal resolution of the model. For one model, with different accuracies there may be several resolutions. For instance GFS has two different temporal resolution modes, one running up-to 192 hours ahead with a finer spatial resolution, while the other running up-to 16 days ahead with a lower spatial resolution [2]. To increase temporal and spatial resolution, mesoscale models such as MM5 [24] and ARPS [84] downscale the data from NWPs like GFS [1], using the standard outputs for grids as initial and boundary conditions, to a finer physical resolution through physical equations.

NWPs can carry out the prediction of irradiance values in two ways. They can either use radiative transfer scheme [19] for long term predictions, or a simple model for short term predictions [86]. The former group is computationally expensive, which makes them hard to use for short term predictions [19]. The simple models, are physical models that can give an estimate of the total irradiance value incident on a surface, knowing weather related variables such as density, humidity, temperature, and pressure at certain altitudes of the atmosphere [70, 82, 87]. These methods use the following steps for the prediction:

- First predict the future values of weather related variables using NWPs.

- Then use those future values in physical models that determine the irradiance level depending on these [87].

An important point to note here is the slight ambiguity in the literature as to what is a NWP and how it is used in predicting the irradiance level. Originally NWPs are used to predict the state of atmospheric variables only. After these variables are predicted, the irradiance value corresponding to those variables is calculated using a separate model. So actually the model that predicts the irradiance is some other model than the NWP model itself, that predicts variables required by the irradiance model. Nonetheless, these two-stage methods are sometimes also referred to as NWPs in the literature. But in fact, they make use of NWPs and are not the NWPs themselves. For instance, to predict the ground level

of solar irradiance, the ARPS surface solar radiation model [85], by Xue et al., utilizes a physical model, Atwater and Ball's model (1981) [6], that make use of water absorption transmittance by Lacis and Hansen [45], and transmittance for Rayleigh scattering and permanent gas absorption. Here, the key point is, both the water absorption transmittance and the transmittance for Rayleigh scattering and permanent gas absorption make use of variables air mass, surface pressure, and precipitable path length, that are predicted by a standard NWP, in this case the ARPS system by Xue et al. [85], but overall prediction of irradiance is completed using the physical model by Atwater and Ball [6]. Model by Yang and Koike [87] that estimates the surface solar irradiation from upper-air humidity is another example to such models. First, a clear-sky solar radiation is calculated using the model by Yang et al. [88]. Then, using the relative humidity profiles for three atmospheric sublayers from any NWP system such as ARPS [84] and MM5 [24], a sky clearness indicator is generated. Afterwards through a radiation-damping process by Rasmussen et al., the clear-sky value is adjusted. Models that make use of sunshine duration [88][More citation needed!] or cloud fraction [6][More citation needed!] are some other examples of such models.

Regressive Models

Models of this kind are obtained by fitting a function, that takes as input the measured weather and location related variables and matches these to irradiance measurements. Usually the function fitted represents the physical phenomena behind the actual reality, but the model parameters tuned on the data. To be able to generalize the success of these models for more than one site, data from a variety of sites is used during the model training procedure. Since as input to these models, weather data is used, predictions by NWPs on the weather related data can be plugged into the model to generate predictions on irradiance values. In fact the models explained in 2.2.1 on page 20, the Atwater and Ball's model (1981) [6] and the model that estimates the surface solar irradiation from upper-air humidity by Yang and Koike [87], are also examples to these kind of models.

   The simplest models take the attenuation of irradiance due to atmosphere as a matter of the distance traveled by the solar ray. The matter is approached from a geometrical perspective. The attenuation of the extraterrestrial irradiance is considered as a function

of the solar zenith angle, since the higher the zenith angle the higher the air-mass (AM), these models mostly scale the extraterrestrial irradiance with a factor of the cosine of the solar zenith angle [5]. Their approach is geometric; they don't consider cloudy conditions, and from this respect they can be referred to as clear-sky models. Also, because these models are calibrated for certain sites, one should be careful before applying them on a different site [5, 7]. Daneshyar-Paltridge-Proctor(DPP) model (1978) [66], Kasten-Czeplak (KC) model (1980) [40], Haurwitz model (1945) [29, 30], Berger-Duffie model (1979) [7], Adnot-Bourges-Campana-Cicquel (ABCG) model (1979) [7], and Robledo-Soler model(RS) (2000) [75] are models of this kind. In eq. 2.1 the BD model is given [5, 7]. In the equation, I_0 stands for the extraterrestrial irradiance which is calculated using eq. 1.1 on page 17, and the θ_z is the solar zenith angle, which has a geometric formulation depending on the day of the year and the geographical location of the site. As can be seen in the eq. 2.1, it has a very simple formulation that only considers the length of the path the solar ray travels, and the constant factor 0.70 is probably found through statistical or empirical methods for the site, for which the model is fitted, and need to be adapted for the specific site before used.

$$GHI = I_0 \times 0.70 \times \cos(\theta_z) \tag{2.1}$$

There are also some simple models, though they are more advanced than the group of models introduced in the previous paragraph, that, in addition to the zenith angle, make use of basic atmosphere related parameters; such as, air pressure, temperature, relative humidity, aerosol content, and Rayleigh scattering [5, 7]. The Kasten model [5] [cite from the original!] that, as given in eq. 2.2, takes as input Linke Turbidity (TL), air mass (AM), and elevation (h) is an example to such models. Kasten model is later improved by the addition of some correction terms by Ineichen and Perez as can be seen in eq. 2.4 on the next page [35].

$$GHI = 0.84 \times I_0 \times \cos(\theta_z) \times \exp(-0.027 \times AM \times (f_{h1} + f_{h2}(TL - 1))); \tag{2.2}$$

$$\text{with } f_{h1} = \exp{-h/8000} \text{ and } f_{h2} = \exp{-h/1250}. \qquad (2.3)$$

$$\text{GHI} = c_{g1} \times I_0 \times \cos(\theta_z) \times \exp(-c_{g2} \times \text{AM} \times (f_{h1} + f_{h2}(\text{TL} - 1))) \times \exp(0.01 \times \text{AM}^{1.8}); \quad (2.4)$$

$$\text{where } c_{g1} = 5.09e^{-5} \times h + 0.868 \text{ and } c_{g2} = 3.92e^{-5} \times h + 0.0387. \qquad (2.5)$$

Apart from these, there are complex models that takes atmospheric parameters such as ozone, aerosols, and perceptible water [5]. These models are denoted as some of the most accurate models by Ren, but since the number of parameters needed is high for many of them, since all the required parameters are not readily available for all the sites, though, at the expense of reduced accuracy, some parameters can be approximated by constants, it is not possible to use them everywhere [5]. The MAC model [20, 21] that accounts for ozone layer absorption, Rayleigh scattering by molecules, the extinction by aerosols, and the absorption by water vapor, taking air mass, humidity, temperature, and an aerosol transmissivity factor, is one example to models of this kind [5]. Also Bird model [10] that takes as input the turbidity, perceptible water vapor, surface pressure, and ground aldebo, and using those considers transmittance of atmosphere for Rayleigh scattering, aerosol attenuation, water vapor absorption, ozone absorption, and uniformly mixed gas absorption for the prediction of irradiance [5, 10]. Interested reader can find more examples of this kind of models in the study by Ren in [5].

Statistical Models

This class of models are based on relating the past observations to irradiance values in the future to obtain predictions [57]. Because of this assumed time causality in irradiance, these models can also be referred to as time series models. Simplest models use the irradiance measurements of previous time-steps directly as inputs to explain the future predictions. More complex models use statistical summaries of different atmospheric measurements and seek for relation between those and the future irradiance values accordingly. Also Diagne et al. puts models that relate the observed atmospheric data for time t to the observed irradiance data at time t, which is the same as what mentioned in regressive models. Since the statistical methods, such as Neural Networks, that is used for this purpose are not mentioned

in the previous section, they are going to be covered here as well.

The Persistence models, are the most basic examples of statistical models in this category. The idea behind is that the changes in irradiance values are smooth, so the most reliable prediction of the irradiance value in the next time step is the irradiance value in the current time step [22, 57] as can be seen in eq. 2.6.

$$GHI_{t+1} \approx GHI_t \tag{2.6}$$

An advanced version of this approximation, as presented by Perez et al. and Marquez and Coimbra, is assuming that the factors that affect the observed irradiance value at a time t would persist for time $t + 1$ and hence their effects will be the same [55, 68]. For instance, as can be see in eq. 2.7, where I_{0_t} stands for the extraterrestrial irradiance at time t, and τ_{ex} is the turbidity factor that attenuates the expected extraterrestrial irradiance to the observed total solar irradiance; according to the stated assumption, the turbidity factor stays the same for the next time step and accounting for the changes in the extraterrestrial irradiance due to geometry changes, one can generate a prediction using the calculated turbidity for time $t+1$. As noted by Diagne et al., this is a naive method that is vulnerable to changes in cloud conditions, since these can change fast and significantly alter the turbidity factor between time steps. As a result of this, persistence methods are used for a baseline comparison to newly introduced methods [22] as done by Perez et al. in [68].

$$GHI_t = \tau_{ex}I_{0_t} \text{ and } GHI_{t+1} \approx \tau_{ex}I_{0_{t+1}} \tag{2.7}$$

Auto Regressive Moving Average (ARMA) models can be used if the time series to be predicted has is stationary. These are two part models with an auto regressive part and a moving average part, and both parts has an associated order, p and q respectively, that determines the model complexity [22]. The model is refered to as ARMA(p,q) due to the order of its respective parts. However according to the Augmented Dickey-Fuller (ADF) test, the solar irradiance data is non-stationary, and contains trends [22]. Because of this, Diagne et al. mentions methods for detrending for obtaining stationary series, but these are beyond the scope of this text. Ji and Chee in [38] applied an ARMA model to the stationary

solar data time series, after the dtrending phase, and they concluded that the best model in terms of order is ARMA(1,1) [22]. Also there are ARIMA (Auto-Regressive Integrated Moving Average) techniques that can be applied on non-stationary time series, and used to predict irradiance by Reikard in [73] where these are compared to other models, and noted to catch the sharp transitions in irradiance at the 24-h horizon more accurately then other models.

As mentioned in the first paragraph, some of the regression models can be also considered as statistical models, since to learn the model function machine learning and statistics related methods are used. These can be investigated under three groups in terms of the model inputs [22]:

- Models mapping weather related variables at time t to irradiance measurements at time t, these are the models of the form in eq. 2.8. For obtaining these models Multi Layer Perceptrons (MLPs) and Radial Basis Functions (RBFs) are used [22].

- Models mapping a series of irradiance measurements at previous time steps, between time $t - n$ and t, to irradiance measurements in the future time steps, time $t + l$, these are the models of the form in eq. 2.9 on the next page. For obtaining these models, time series models that can be solved using Neural Networks, wavelet-networks, and fuzzy-wavelet-networks can be used [22]. These models are also refered to as Nonlinear Auto Regressive (NAR) models by Marquez and Coimbra in [55][also the other one comparison of clear-sky models for evaluating]. In [13–15, 39, 42, 59, 80] examples of these models can be found.

- Models mapping weather related variables and irradiance measurements at previous time steps, time $t - n$, to irradiance measurements in the future time steps, time $t + l$, these are the models of the form in eq. 2.10 on the facing page. For obtaining these models, Auto Regressive techniques, Markov Chains and Time Delay Neural Netwroks can be used [22]. These models are also refered to as Nonlinear Auto Regressive with Exogenous inputs (NARX) models by Marquez and Coimbra in [55, 56].

$$GHI_t = f(T_t, S_t, P_t, Lat, Lon) \tag{2.8}$$

$$GHI_{t+l} = f(GHI_t, GHI_{t-1}, GHI_{t-2}, ..., GHI_{t-n}) \qquad (2.9)$$

$$GHI_{t+l} = f(GHI_t, GHI_{t-1}, GHI_{t-2}, ..., GHI_{t-n}, T_t, S_t, P_t, Lat, Lon) \qquad (2.10)$$

The variables that are seen in eq. 2.8 on the preceding page and eq. 2.10 can vary from one model to the other, the important idea is that they are variables related to weather, such as temperature (T_t), sunshine duration (S_t), and air pressure (P_t), and variables that are related to the location of the site under consideration, latitude (Lat) and longitude (Lon).

Neural Networks are known for their ability to learn non-linear functions in the data, provided that there is a large number of examples [60], they are widely employed in irradiance forecast as well. Unlike the physical models, Neural Networks do not require any knowledge about the process through which the data is generated, they just explore the patterns, that are hard to recognize, in the data [60]. Some of the proposed models use only the raw weather and irradiance related data, while others may also use statistical features extracted from a time-frame of the data. Method presented by Sfetsos and Coonick, that uses the raw hourly averaged irradiance data to make one-step predictions for the next hour [80] is a good example for the former case, while the method presented by Wang et al. [83] that predicts solar irradiance using statistical features extracted from the irradiance and ambient temperature data is an example of the latter.

Martín et al. presents a Neural Network based time series prediction technique in [57]. They predict half daily irradiance for the next 3 days, and due to non-stationary nature of the irradiance data, they use the clearness index (K_T) and lost component (L_C), where the clearness index is defined as the ratio between the ground measured global and extraterrestrial solar irradiance, and the lost component is the difference between the extraterrestrial and the ground measured global solar irradiance. Time series methods for both variables including the Auto Regressive, Neural Network, Addaptive Network Based Fuzzy Inference models are evaluated against the baseline, persistence model; the Neural network based model seen to provide the best improvement [57].

Similarly Wang et al. build a statistical feature vector as input to a neural network with four layers, to predict hourly solar irradiance values for the next day in a time series . Feature vector consists of the day number to account for date, the average surface irradiance and the

average ambient temperature of the previous day to account for general weather conditions, the maximum of the discrerete third order derivative of the difference between calculated extraterrestrial and measured ground global solar irradiance time-series data (lost component) of the previous day to account for variations and fluctuations in weather conditions, and the mean-square difference between the calculated extraterrestrial and measured ground global solar irradiance time-series data of the previous day to account for the changes in weather conditions [83]. In [83], this feature based approach is compared by Wang et al. against the approach that uses the hourly irradiance measurements belonging to the previous day, and it is seen that the results of feature based model are accurate for the sunny days. Also the cloudy day performance of feature based model is better than the other. The RMSE is found to be 42.29 W/m2 for clear days and 84.65 W/m2 for cloudy days for the feature based method, while it is 43.07 W/m2 for clear and 254.66W/m2 for cloudy days for the other method.

Mellit and Pavan developed an MLP that takes mean daily irradiance and the mean daily temperature of the previous day to predict 24 h ahead, hourly solar irradiance [58]. Sfetsos and Coonick investigated different meteorological variables such as wind speed, temperature, and pressure as for forecasting hourly solar radiation using ANNs along with other techniques [22, 80]. Their comparison yield the result that, in terms of training time and prediction errors, ANN trained using The Levenberg-Marquardt (LM) algorithm performs superior to other methods.

There are also alternative methods other than neural networks. Mellit et al. developed an adaptive α-model that predict the hourly global, direct, and diffuse solar irradiance using sunshine duration, relative humidity, air temperature, and air pressure of the previous time step, in a time-series fashion. As can be seen in eq. ?? on page ??, model is similar to ones that are mentioned in the third category, eq. 2.10 on the preceding page. In case of unavailability of some of the parameters, any combination of the parameter set is tested for the accuracy of the estimation they provide. Mellit et al. compared this adaptive α-model with a neural network based model with the same inputs, for a site at Jeddah. They reported a 97% correlation between the predicted and the measured data for the adaptive α-model against 98% correlation for the neural network based model. But since the bias error for addaptive

α-model is less than 0.8, which is good, and the implementation and the formulation of the proposed addaptive α-model is much simpler, in terms of the model complexity, than a neural network, it is a presented as a good alternative for neural networks [60].

$$GHI_{t+1} = f(T_t, S_t, P_t, H_t) \tag{2.11}$$

Many more different methods can be presented in this category. Since predictions of this kind are required for weather and cloud condition predictions, literature contains many alternative methods. As stated by Ineichen and Perez, in general models used are chosen not only by their accuracy in predicting correctly, but also the availability of the required data and the implementation simplicity. Because of this, only the models that are possible to be used in this study in terms of the available data and implementation simplicty arre covered here.

2.2.2 Models that take as input sky cloud conditions from recent observations, and make use of this cloud information during prediction

As alreay mentioned, models presented in 2.2.1 on page 20 are not successfull in predicting the short term cloud events, and local changes in weather related data. Lonij et al. states that, WRF, an example to NWPs, can predict a big cirrus, a major cloud event, 50-h ahead, while it fails to predict a cumulus, a local cloud event, with nothing better than a 1-h temporal accuracy [51]. NWPs generally lack the spatial resolution as according to Chow et al., the finest spatial resolution of such methods are as low as $1km^2$, and that these models are not able to generate sub-hour predictions accurately [17]. Since as already presented in the previous section, many of the methods are based directly on the NWPs, using their predictions on weather data as input, their temporal accuracies are also upper bounded by those of NWPs. Also in [37] Jayadevan et al. present the effects of cloud events statistically. Comparing the recorded GHI values to computed clear-sky GHI values, they report that clouds cause more than 20% drop on power output calculated for clear-sky conditions. Also they state that the majority of events cause a power drop 20-30% and less then a minute, while the majority of energy loss is due to events that cause a drop in power of more than 70% and last between 1

and 6 hours. Given the temporal accuracy of the stated methods , it is sufficient to introduce the visual observation of cloud conditions to prediction models [17].

When the literature of methods using sky observations are analyzed, it is seen that they provide predictions using a simple idea. This idea is outlined by Chow et al. in [17]. It is stated that the primary factor attenuating the GHI is the presence of cloud between the ground and the sun, since this directly attenuates the solar beam irradiance. Due to this reason, studies that use cloud observations for predictions tend to predict the cloud conditions that will affect the site under consideration, and decrease the irradiance prediction if the site is predicted to be occluded. In the following, studies that work in this direction are presented.

Chow et al. propose a method in [17] using a ground-based sky imager for cloud observation. For GHI predictions, Ineichen clear sky model that is later improved by Perez [35] is used. It requires the Linke Turbidity factor [48] as input. To account for cloud effects, they use the projection of cloud decision image onto sky, assuming a constant CBH value that is fed from a ceilometer close to their site. For each 10mx10m cell in the 10kmx10km grid centered at the camera, sun occlusion is calculated using a geometrical formula and the sky projected cloud image, and if an occlusion is detected, the clear sky prediction by Ineichen is attenuated by 60%. Selection of the attenuation factor is explained to be due to optically thick clouds observed during development period. Also depending on the CBH, it is stated that the area covered by the camera changes, according to the triangle inequality. The cloud motion detection and prediction methods are validated comparing the GHI values predicted and measured. The nowcast performance, the ability to perceive and understand the current scene, of the method is found to be correct 70% of the time. The improvement of irradiance prediction due to cloud forecast is evaluated comparing: the prediction obtained using cloud forecast, with predictions obtained using the persistence of cloud conditions in the forecast horizon. It is found that, the irradiance forecast error reduced by 50-60% when the cloud forecast is used instead of cloud persistence.

Hashimoto and Nagakura present an early phase of an ongoing study in [28], where the irradiance is not predicted directly but it is claimed that the prediction can be generated with the methods explained. The cloud positions can be measured in 3D with a range direction

error of less than 5%. It is claimed that this information is sufficient for determining the direction of movement and the speed of clouds. Accordingly the way of providing impact time predictions for clouds is explained . Also the irradiance prediction procedure corresponding to the impact of clouds is explained. During the detection phase, clouds are classified as thin or thick depending on the light-dark variation of their pixels. Then according to a statistical study presented, it is reported that, thick clouds cause a drop of 70-80%, and thin clouds cause a drop of 10-50%. It is also mentioned that in 100 cases the drop estimation method proved to be 85% accurate. So putting these two parts together, it is claimed that predictions on the local irradiance level could be provided. However, as already mentioned, the actual method is not experimented in the study.

Nova et al. present an alternative method in [63]. Attenuation over the extraterrestrial irradiance, due to atmosphere, is calculated using a persistence based method that takes the ratio between the measured and predicted solar irradiation of the previous time step as the turbidity factor. A neural network with two layers is trained to model the attenuation process. Using MLP that takes previous 20 attenuations as inputs and gives predicted next 10 attenuations as outputs. Also using the cloud predictions, the time when the sun is going to be occluded is predicted and the cloud cover, as a ratio of cloud pixels to total number of pixels on the cloud decision image, is calculated, but it is not clearly explained how these are used for irradiance prediction.

In [54], Marquez and Coimbra present a method that does not consider area shadow due to clouds, for sky cover estimation, but projects clouds to a flat sky surface. A ground based sky imager is used for sky observation. Using the cloud tracking algorithm based on cross correlation method, a representative cloud motion vector is obtained. Using this vector, grids of 20x20 pixel cells, oriented towards the inverse cloud motion direction starting from the sun, are formed, and the cloud cover for each of these cells are computed as the ratio between the cloud pixels and the total number of pixels in one cell. Then for each different horizon a different grid cell's cloud cover value is used for irradiance prediction. The idea behind is, if it takes 1 min for a cloud to travel 10 pixels on the image, then in 3 minutes it would travel 30 pixels, the cell that is 30 pixels away from the sun is the second cell, so the current cloud cover of this cell gives us a prediction for the 3 minute later cloud

cover around the sun, provided that the assumptions about the point sensor and the device location are precise. Using the predicted cloud cover denoted as $X_i(t)$, associated irradiance estimation is obtained using the equation presented in eq. 2.12 where the cloud cover is inversely related to irradiance as expected. Comparing the irradiance forecast RMS errors for different horizons, to corresponding naive persistence model forecast RMS errors, it is seen that for a 5 minute horizon, the best improvement over persistence model is obtained., which is between 20-40%.

$$X_i(t)' = DNI_{max}(1 - X_i(t)) \tag{2.12}$$

An alternative method that makes use of distributed PV installations from 80 locations for cloud observation is presented in [51] by Lonij et al.. Method collects all the measurements from 80 locations to a central unit, assuming any loss over the clear-sky value is due to clouds. Formulates a clearness index (K) as a function of location and time as can be seen in eq. 2.13 on the next page. Then using the cloud velocity (v_x, v_y), the future clearness index for a location (x, y) can be predicted using eq. 2.14 on the facing page. For the areas that are between measurement points, a four point interpolation method is applied, where the median of four closest measurements is considered as the measurement for the location under consideration. To obtain cloud speed, three methods are investigated:

- Wind velocity estimation from NWP as an estimator of the cloud velocity,

- Cloud velocity inference from the network of PV systems,

- A constant velocity throughout the day that is numerically optimized to minimize RMS error.

The cloud velocity inference from the network is done finding the best solution (v_x, v_y) to the equation eq. 2.15 on the next page, for all the locations in the system (x_i, y_i). Using all three cloud velocity estimators, the proposed method is compared against the persistence method on time harizons ranging from 15 minutes to 75 minutes. It is reported that proposed method, that uses the numerically optimized cloud speed, outperforms the persistence method on partially cloudy days by reducing the RMS prediction error by 20%, that are very hard cases for forecasts of this sort. Also a cloud tracking method, that uses ground based sky images

from a sun tracking camera, is mentioned. It is stated that using this method, 10 minute ahead predictions of cloud impacts can be provided, and these can further improve the irradiance predictions. Also in [37], an early phase of this ongoing study where no precise method for irradiance prediction is given, a method to predict impact time for clouds is presented, along with suggestions for possible use in irradiance prediction, by Jayadevan et al..

$$K(x, y, t) = \frac{GHI(x, y, t)}{GHI_{clear}(x, y, t)} \tag{2.13}$$

$$K(x, y, t + dt) = K(x - v_x dt, y - v_y dt, t) \tag{2.14}$$

$$K_i(t) = K(x_i - v_x dt, y_i - v_y dt, t - dt) \tag{2.15}$$

2.3 Cloud Observation

As already explained, NWPs can predict the amount of cloud cover and major cloud events using cloud models that are dependent on weather related data. However, with a spatial resolution of 100km-1km, and a temporal resolution of 3-10 days, it is not possible to generate precise predictions on the cloud conditions. Because of this reason, most prediction methods, use the information that is obtained by a cloud observation method. For NWPs, the source of observation is satellite data. It is reported that using the satellite data, NWPs can generate accurate predictions up to half an hour [17, 25]. It is also a known fact that, NWPs and satellite forecasts do not perform as good as persistence based methods at sub-kilometer and intra-hour forecasts [17]. As a result of this, to provide a sub-kilometer and intra-hour resolution data on clouds, ground based sky imaging systems are introduced as local observation methods [17]. In this section, the methods that are used for cloud observation and prediction are presented.

2.3.1 Cloud Detection and Classification

To track an object, one has to detect and identify it first. In this part, methods used for cloud detection and classification are discussed.

Cloud Detection

Cloud detection is the basic step of a system that predicts irradiance values using cloud observations [86]. Several methods developed for cloud detection, and Yang et al. classify these methods into two classes: methods that do not employ any a-priori knowledge, and methods that use a-priori knowledge [86]. Clear sky, because of the molecular scattering, scatters the blue component of the light more than its red component [50, 69]. Due to this fact the sky is observed as blue (as can be seen in the top left image of figure 2-1 on page 52), except for the horizon, where because of the increased air-mass, all the components are scattered far more, and a the observed color is close to white [69]. When the red and blue channel components of the image pixels corresponding to the clear sky are observed, it is seen that the red channel value is low and the blue channel values are higher since red light is not scattered as much as the blue light. As a result, as the difference between clear sky areas in the bottom right and left images of figure 2-1 on page 52 can be seen, the clear sky ares are observed much brighter in the blue channel image than the red channel image. Clouds scatter both the red and the blue components more than the clear sky, as a result of this the both the red and the blue channel components of a pixel corresponding to a cloudy area have considerably high values then a pixel corresponding to the clear sky area. As a result, the cloud areas look brighter in both red and blue channel images (bottom right and left images of figure 2-1 on page 52). Also because the clear sky areas do not scatter the red component as much as the cloud areas do, the cloud areas look brighter on the red channel image (figure ?? on page ??, bottom left) and the contrast between cloud and sky pixels are greater. When we look at the red-to-blue ratio (RBR R/B)intensity image (figure 2-1 on page 52, top left), as a result of the stated facts, the ratio values are greater for the cloudy areas than they are for the clear sky areas. Using this observations, methods that use a fixed threshold on the whole RBR image, to decide as cloud or clear sky for each pixel, are designed. Koehler et al. mention one such fixed threshold method for thin and opaque cloud decision. For pixels with RBR above the opaque cloud threshold the decision is opaque cloud, while the pixels with RBR values between opaque and thin cloud threshold are classified as thin cloud, and the other pixels with R/B values below the thin

cloud threshold are classified as clear sky. Through empirical observations, decisions for opaque clouds are said to be fairly good, while the thin cloud decisions are said to perform not as well [43]. Long et al. apply a similar approach and classify the pixels with RBR greater than 0.6 as cloudy and other pixels as clear sky. To determine the threshold at 0.6, an empirical hand-tuning procedure is employed on 100 images, and with visual inspection and comparison to the original image, the threshold is set to 0.6. Heinle et al. evaluates the idea by Long et al. in [32], and mention problems in detecting thick clouds and classifying circumsolar pixels at the same time, and come up with an alternative fixed threshold on R-B. In [32], it is mentioned that the performance of the introduced criteria, with the fixed threshold at $R - B = 30$ is better than the RBR threshold. In Figure. 2-2 on page 53 top right, an application of this criterion on a sample sky image is shown. As can be seen, the cloud areas are darker than the sky areas. This is because, the cloud areas scatter both the red and blue components high, while the clear sky areas scatter blue component much more than the red one. As a result the difference between red and blue channel values for clear sky are expected to be considerably higher than the cloud areas. Hence, areas occupied by clear sky are expected much brighter than areas occupied by clouds, as can also be seen in Figure. 2-2 on page 53 on the top right.

When we compare the two criteria (RBR and R-B), we see that the threshold works in the opposite directions for them. For RBR, pixels with RBR higher than a threshold are decided as cloudy, while for R-B pixels with R-B lower than a threshold are decided as cloudy. Another important point to note is, as can be seen on the original image (Figure. 2-3 on page 54 left), towards the bottom right area, in the middle of clear sky there is a cloud. This can be observed as well on the RBR image (Figure. 2-3 on page 54 middle) while it is very hard to notice on the R-B image (Figure. 2-3 on page 54 right). Also when the bottom part of the original image (Figure. 2-3 on page 54 right) is observed, it is seen that there is a little cloud occurrence along the bottom. This cloud is detected bigger than actually it is on the RBR criterion (Figure. 2-3 on page 54 middle) while the R-B criterion accurately captures that (Figure. 2-3 on page 54 right). Considering these, it can be inferred that the superiority of these methods to each other is not really clear and depending on the site, camera, and the other parameters of the corresponding setup, the result may change.

Long et al. report problems due to usage of a fixed threshold. It is mentioned that, with fixed threshold detection of thin clouds and detection of clouds occurring near the circumsolar region are encountered. Considering the RBR image seen in Figure. 2-3 on page 54 middle, we see that there is a change in RBR value in clear sky areas, that, normally, is not expected. This suggests use of a variable threshold across the image. In [17] Chow et al. explain a method that uses variable threshold depending on the pixel and the sun position. Because of the non-uniformity of RBR in the whole sky in clear conditions due to aerosol and airmass effects, the proposed method combines two different methods for cloud decision. RBR is used with a dynamic threshold that is obtained from the Clear Sky Library (CSL). The dynamic threshold is calculated as a function of zenith and sun-pixel-angles from images on a clear-sky day. For each pixel on the image, a clear sky image based on the current solar zenith angle is generated, and the CSL threshold is defined as the sum of RBR in the clear sky background image and an additional threshold value. A pixel is classified as cloudy if its RBR is larger than the corresponding CSL threshold. Afterwards, it is seen that the thick and dark clouds close to sun are not detected accurately, and a Sunshine Parameter (SP) is introduced for correction. SP is calculated as the average RBR around the sun, and the pixels with RBR > SP are classified as cloudy, even if the CSL states otherwise. SP is calculated for each image dynamically, and if the sun is occluded, this value decreases. (the reason why it works is also given). Using CSL and SP together, the cloud detection is performed. Then the decision image is projected onto sky assuming a constant CBH value that is obtained from a ceilometer close to field, to obtain a cloud cover value. Also Marquez and Coimbra [54] explain an alternative variable threshold cloud detection algorithm to better detect thin clouds and clouds near the sun. In this algorithm an adaptive threshold is used for different parts of the image according to the study by Li et al. [47]. Normalized r/b ratio is calculated using (b-r)/(b+r). Through Minimum Cross Entropy (MCE) algorithm, based on cross entropy minimization between normalized RBR and segmented image threshold, an adaptive threshold is determined [54]. Yang et al. [86] propose a method for that improves detection of cirrus clouds. The method consists of three steps. First through red-to-blue channel operation a single channel image is obtained from the original RGB image. Different red-to-blue channel operations including, R/B, $\log(R)/\log(B)$, $R-B$, and $\frac{R-B}{R+B}$ are considered.

To select the method that would help the most in cirrus cloud detection, 181 cirrus cloud images are considered and the normalized difference ($\frac{R-B}{R+B}$) that keeps the biggest gray-scale information is selected. Second, nonuniform illumination background image is estimated using morphology opening, and this background is subtracted from the image obtained in the previous step. Third, an adaptive threshold is calculated and applied on the output of the second step, resulting in a binary cloud decision image. The proposed algorithm is compared against fixed threshold and adaptive threshold methods, and reported to outperform them with a 85% mean accuracy where the other methods score 46% and 60% respectively [86].

Alternatively there are methods that employ neural networks for cloud decision task. Nova et al. uses a Multi Layer Perceptron (MLP) based method in [63]. Input to the MLP is RGB value of a pixel, and the output is the cloud or sky decision. To train the MLP, pixels from images that are labeled for the cloud and sky regions by an operator providing a binary classification on the pixel level, are used. The MLP is trained on each pixel, using the R, G, B values as input and the binary cloud-sky decision as output. The error for the training phase is determined with respect to the labels provided by the operator. The performance of this algorithm is not mentioned in [63]. Also in [16] a more extensive study of the use of MLPs for cloud decision is used. It is referred to as extensive because, along with R, G, and B 15 different statistical and individual features of a pixel are examined, and the an optimal set of inputs is obtained using a genetic algorithm on the set of possible inputs. In [16], the mentioned learning algorithm is employed to classify any pixel into three classes: thin cloud, thick cloud, and clear sky. Existence of a relation between 18 different features that can be extracted from a pixel with its 8 neighbors is assumed. For the learning, an MLP is employed. The proposed MLP, takes each feature as an input and gives one of the three answers mentioned as a decision. For training, from 50 images, patches are taken, and labeled as belonging to one of the classes thin cloud, thick cloud, and clear sky. It is mentioned that, the patches are selected mainly from the areas, such as the cloud edges, and areas close to the image boundaries, that are hard to classify for any algorithm. For each pixel, a feature vector containing features extracted from the pixel and its neighboring area, forms an input sample. The input vector contains, R, G, B, gray-scale, R/B, R/G, G/R, G/B, B/R, and B/G values for the pixel, and mean and variance R, G, B, and gray-scale

values for the pixel and its 8 neighbors (3 pixel by 3 pixel rectangle centered at the pixel under consideration), and the label provided by an operator. As a result of this sampling, it is reported that, about 1000 samples are obtained. A neural network is trained using all the components in the samples, and a performance with 82% correct classification is obtained. Afterwards, using a genetic algorithm, the number of inputs to the neural network is minimized such that the performance of the neural network improves. It is said that, this is done to avoid unnecessary information. As a result, an optimum MLP is found that takes 3 inputs, the average of the pixel and its neighbors in red and blue channel, and the variance of of the pixel and its neighbors in red channel, and has a performance of 85% correct classification for opaque clouds, while it performs even better for clear sky pixels with 88% and slightly worse for thin cloud pixels with 61%.

Hashimoto and Nagakura proposes a cloud decision algorithm that relies on the pixel brightness values for cloud decision [28]. This algorithm follows a completely different method than any algorithm that is explained in here. During the decision phase, clouds are classified into two categories: thin and thick. The detection algorithm relies on an empirical observation, that for thick clouds, the surface of the cloud looking towards the sun is significantly brighter than the other side, while this is not the case for thin clouds. Depending on this, the cloud image is checked for each pixel along the radial line from the sun. The brightest pixel is found and the gray level decay along the line is calculated. If a pixel that is brighter than the previous pixel is detected, this is counted as a new starting point. If lots of detections are made, the cloud is classified as thick, otherwise thin. This method is tested on 50 random samples against the traditional meteorological methods, and with a difference rate of less than 10% it is said to be reliable.

Cloud Classification

There is a difference between the tasks of cloud detection and cloud classification. Cloud detection refers to deciding whether a pixel refers to clear sky or to a cloud. On the other hand, cloud classification refers to the task of determining the meteorological type of the corresponding cloud seen on the image. As can be expected, the classification task is much more complicated than the simple detection task, and in fact most of the methods proposed

for cloud classification require cloud decision done on the considered scenario [12, 32, 86]. To stress the difficulty of this task, Calbó and Sabburg mentions that, even the definition of a cloud is a problematic issue depending on the equipment used for observation and the site under consideration [12]. It is also mentioned that, even two trained human observers may disagree on the amount and/or type of clouds present, signifying the difficulty of the problem under consideration. Several methods that rely on feature extraction, and classification depending on the extracted features are proposed. Here, the ones that are seen as easy to understand and implement are discussed.

Heinle et al. [32] propose a method to classify the cloud condition classes of the International Cloud Classification System (ICCS) published in WMO (1987) [64]. Due to possible confusions between several types, these types are combined to avoid systematical misclassification. Used cloud condition types are [32]:

- Clear sky,

- Cumulus,

- Cirrus and Cirrostratus,

- Cirrocumulus and Altocumulus,

- Stratocumulus,

- Stratus and Altostratus,

- Cumulonimbus and Nimbostratus.

. For the multi-class classification, k-nearest-neighbor (kNN) is used on the feature vectors extracted from labeled images of different classes. For a sample to be classified, kNN goes through the training samples, and compares the new item to those to obtain a distance between the new item and all the items in the training set. As a distance metric, Manhattan Distance (eq. 2.16) is used.

$$d(x, x^i) := \sum_{j=1}^{dim} \|x_j - x_j^i\| \qquad (2.16)$$

Then using the labels of k training samples that have the smallest distance to the new item as votes, the algorithm decides on the class that has the vote of the majority. Each training image is processed for feature extraction. Prior to feature extraction, the binary cloud decision image is prepared using the R-B method mentioned in the previous part. After that, spectral features and textual features are extracted from the image. A feature used is selected through the Fisher distance obtained between classes assuming that this feature is used for classification. Features that provide a good separation between classes (high Fisher distance) are used. Spectral features are the mean of R and B, standard deviation of B, skewness on B (a third order statistic), and the differences between mean values of R, G, and B (R-G, R-B, and G-B). As can be seen, the spectral features describe the average color and tonal variation of an image, and used for discriminating thick dark clouds such as cumulonimbus, from brighter clouds, such as high cumuli-form clouds and transparent cirrus clouds. Since spectral features are not sufficient for a discrimination between types, textual features are also used. For the texture, energy showing the homogeneity of gray-level differences, entropy as a measure of randomness of gray-level differences, contrast as a measure of local variations of gray-level differences, homogeneity as a measure of similarity of adjacent gray-levels, and cloud cover as a measure of average cloudiness, to discriminate between cloud types that cover the whole sky like strati-form clouds from others, are used. The features are normalized to ensure that all features are equally weighted. The algorithm is evaluated for all the classes separately to understand which types are confused. As a result it is reported that the algorithm classifies correctly with a mean 74% performance. It is also reported that, cirrus and cumulus clouds in case of less than 30% cloudiness, and cumulus and high cumulus clouds due to their similarity in color and smooth transitions are frequently confused by the algorithm. It is also stated that due to the occurrence of more than one type of cloud in an image, some confusions can happen, and a grid based sub image evaluation method is proposed to avoid that problem [32].

Calbó and Sabburg propose a method to classify an alternative set of cloud classes that do not confirm to ICCS [64], considering their uses in radiative transfer studies [12]. The considered classes are [12]:

- Scattered. Sun may or may not be obscured.

- Broken to almost overcast. Sun is not obscured.

- Overcast sky with a quiet uniform layer of clouds, thick or thin. Sun is obscured, but may or may not be visible.

- Thin clouds not covering the sky. Sun obscured but visible.

- Thin clouds not covering the sky. Sun is not obscured.

- Broken to almost overcast. Sun is usually obscured.

. To differentiate between these classes, statistical features, features based on the Fourier transform of the image, and features requiring distinction of cloudy and sky pixels are employed. To obtain statistical features, RBR image and the gray-scale image as a mean value of R, G, and B values for each pixel are used. Statistical features on the possible pixel levels of intensity and RBR are used for the texture analysis on the images. As the statistical features, mean value of pixel levels as an indicator of the overall intensity and RBR of the image, standard deviation of pixel levels as a measure of contrast; smoothness, between 0 (for constant) and 1 (for highly variable), as a measure of the variability; third moment of pixel levels as a measure of histogram skewness; uniformity, as a measure of existence of different intensities and RBR values in the image; entropy of pixel levels as a measure randomness in the level of values in the image, are used. Features based on the Fourier transform of the image are, the correlation with clear that assess the similarity between the spectral power function of an image and the spectral power function of a clear sky image; and the spectral intensity that quantifies the distribution of spectral power along a range of wave-numbers. For the correlation with clear, possible values are between 0 and 1, and higher values indicate the more uniform aspect of the sky. For the features based on the thresholded image, fixed threshold on the RBR values is used. From the obtained cloud decision image, fractional cloud cover, as a ratio between cloudy pixels and total number of pixels; and the cloud brokenness, as the ratio between number of pixels on the cloud edges and total number of cloudy pixels, is used. Cloud brokenness is expected to attain its minimum for overcast days, and its maximum for days with patchy clouds. It s not mentioned, what is used as a learning method, but as the training data feature vectors extracted from labeled

sky images are used. The proposed classification method is evaluated, and a 62% correct classification is reported for the proposed eight classes, while the performance gets better to 76% when some classes are merged, and the same procedure is applied on five classes. It is also reported that, some of the features are found to be redundant, as they are highly correlated with others and do not provide additional information during classification [12].

Hashimoto and Nagakura propose an alternative method to classify the clouds according to their meteorological types using Hough transform on the contours as a feature extraction method [28]. According to their experimental results, if the Hough transformation results in circles with small radius, the cloud can be classified as a convective cloud, whereas if the circles have big radii, they are classified as stratus. Also adding a parameter to the Hough transformation, they can detect some spheres in 3D, and if the detected spheres has a large radii, the cloud is classified as a fleecy cloud or an assembled cloud. If the detected spheres have small radii, the cloud is classified as cumulonimbus cloud. Also the cloud classification is mentioned to be a challenging task because of the infinite possibility of cloud forms, and transparency. They also mention that the proposed algorithm is not lightweight, and takes long computations [28].

2.3.2   Cloud Tracking

To be able to predict the future states of clouds, knowledge about their movements is required. To obtain this knowledge, clouds are tracked in time. The first application of cloud tracking started with the use of satellite images for observation. A group of methods, using cloud contours and texture, are developed for cloud tracking using satellite imagery. Later, due to sparse temporal and spatial resolution of satellite imaging based methods, ground based cameras are introduced. During the literature study, the most important observation for us is the fact that, methods used for ground based imagery are mostly inspired from the methods used for satellite imagery. The resolution changes due to the amount of spatial information stored in one image element. As an example to satellite images, the GOES data is a 256 x 256 grid, where each grid element contains a 15 x 15 pixel block with 11km resolution [25, 61], while a typical ground based sky imager captures 2000 x 2000

pixels with 4 to 10 meters per pixel, depending on the cloud base height. Cross-correlation based methods are important examples for methods that are adapted from satellite imagery to ground based imagery. Basics of the method stays the same, but the spatial resolution increases due to the explained fact about the amount of data stored in a pixel [17, 33, 61]. In this part, starting with satellite based methods, the cloud tracking methods present in the literature are going to be investigated.

2.3.3 Methods using satellite Imagery

Nehrkorn present a good overview of satellite based cloud tracking methods. Since, for this project there is no intention of using satellite images, a more detailed explanation will not be provided, but since, as already mentioned, the cloud tracking methods for ground based imagers are mostly inspired by their satellite based counterparts, in this part a brief overview is provided.

Contour based methods

Early methods for cloud forecasting, based on satellite images, use the bounding contours of the clouds for tracking [61]. Algorithms model clouds using their contours, and the changing factor across different algorithms is the way they represent contours [61]. As an example to methods that are used for contour generation, bounding areas that are colder than -50°C is given by Nehrkorn. A sample illustration can be seen in Figure. 2-4 on page 55, where the red lines around clouds represent the contours used. Nehrkorn present "persistence" and "movable persistence" as the most simple techniques. "Persistence method" simply assumes that the cloud conditions for a time-step would remain the same as they are in the time-step previous to that time-step. This method is used for verification purposes [61]. "Movable persistence" assumes that the shape and the size of the contour does not change, and extrapolating the centroid location from the latest position using a linear least-squares regression and moving the contour accordingly, this method generates a forecast for the future [61]. The only tricky part of this method is presented as the centroid allocation for a contour. Nehrkorn state that, using the center of mass of the centroid, rather than

the average of contour points, would improve the forecast results as in this way, long and thin features, that easily change are not heavily weighted as they would be with averaging. For this method, how a contour is defined can be seen in Figure. 2-5 on page 56, and how the forecast is generated through linear interpolation is presented in Figure. 2-6 on page 57. Bianco and Huang [9] present the whole contour technique. According to this method, contour can be sampled at regular points, and the resulting x, y coordinates are periodic. Because of the periodicity, when a Fourier transformation is applied, with the number of waves to be used is an input from the user, a spectral representation of points can be obtained. With the increasing number of waves, the representation gets better. Also centroid positions, as an average of contour points, and aspect ratios (length/height) are also calculated and all the calculated parameters, including each Fourier components, are linearly interpolated into the future. To generate a forecast, contours are regenerated using an inverse transformation on the interpolated data, the obtained contour is scaled using the interpolated aspect ratio data, and the forecast placement of the contour is determined using the interpolated centroid data [9]. Nehrkorn state that, using two consecutive images, only linear interpolation can be used for forecasts, but the forecast accuracy can be improved using least-squares regression, if more than three observations are used. It is also stated that, using more number of components for Fourier transform improves the performance [61].

Bianco and Huang [9] also describe the "segmentation method" [61] as the "angle displacement" method [9]. This method defines a contour using it centroid position and the length of the lines, at equal angular displacements, where the angular separation is determined by the user, connecting the centroid to the contour, an illustration of the contour definition is provided in Figure. 2-7 on page 58. Linear least squares fit is used to forecast future line lengths and centroid positions. After the forecast is generated, contour is reconstructed connecting the end points of lines [9, 61].

Kavvas and Chen [41] present a method that is really similar to the "Segmentation Technique" presented by Bianco and Huang. This methods also uses the centroid position and line lengths radiating from the centroid to describe the contour (Figure. 2-7 on page 58), but it uses a fixed angular displacement of 22.5°. The extrapolation is done by adaptive exponential smoothing smoothing, which is a weighted least-squares fitting [61]. The

smoothing coefficient is used to determine the sensitivity of the algorithm in fluctuations in the observation, where a small smoothing coefficient weights the older observations heavily [61].

Cross-correlation based methods

Hamill and Nehrkorn [25] present the lagged cross-correlation method. Unlike the previous methods, this method does not extract the cloud contours, or does not require a precise detection of cloudy areas. The contour based methods cannot handle cases where a cloud represented as a single contour in a previous image splits into several pieces or a several contours combine to form a single one. Cross-correlation based methods are reported to perform better for these cases with providing accurate forecasts at a horizon of 0.5 to 2.5 hours [25]. The algorithm works in the following way, the image corresponding to the recent time-step is divided into cells as can be seen in Figure. 2-8b on page 59. Then the correlation coefficient for each of these cells is found against the attention window in the previous image. In Figure. 2-8b on page 59, the cell with red surrounding is presented as an example. The correlation between this cell and the area inside the blue circle on the previous time-step image (Figure. 2-8a on page 59) is computed and the location corresponding to maximum correlation is found, from which the motion vector is calculated. For this example case, the maximum correlation is found with the yellow cell in Figure. 2-8a on page 59, and the motion vector is determined as can be seen with the red arrow connecting the yellow box to red box, in Figure. 2-9b on page 60. This motion vector generation is applied on all the cells, and as a result a vector field is obtained for the whole image (Figure. 2-9a on page 60). After the motion vectors are obtained, to eliminate the outliers, a k-means algorithm is run, and all the motion vectors that are different from the cluster mean, are either eliminated, or replaced by the cluster mean. Then the prediction image is obtained using the motion vectors are to move the cells to their associated positions in the next time-step. them In Figure. 2-9b on page 60, application of this procedure is illustrated. As explained in the previous step, the red box is found to have the red motion vector radiating from its center, using this vector, red box is moved to the position designated by the green box in the prediction image (Figure. 2-9b on page 60). Applying this procedure to the whole image,

generates the prediction image. The methods that use cross correlation on any scale, uses a similar procedure to the one explained here, just on images of different spatial resolution. Also Hammer et al. improve this approach by adapting the motion vectors using a model of cloud motion and a maximum a-posteriori estimation criterion [26].

2.3.4 Methods using ground based imagery

Lonij et al. states in [51] that the satellite aided NWP based cloud observation methods are outperformed by persistence based methods under 1 hour. The main reason contributing to this fact is, as also stated by Jayadevan et al. in [37], that the majority of cloud events cause a power drop 20-30% and less then a minutes. This indicates that any method with a temporal resolution bigger than several minutes would not be able to catch and predict those majority of events. Such short term predictions are is critical in integration of a PV plant with a smart grid, as they allow the grid operator to foresee and compensate for the predicted power fluctuations manually using an alternative, for instance a fossil fuel based plant, power source, or via the smart grid software control automatically [33]. These short-term predictions are also crucial for the operation of individual plants, as they allow the plant operator to compensate for the fluctuations using a set of actions including, charging and discharging of batteries, starting an alternative power plant, and disconnecting some panels early to manage the ramp rate.

As will be seen in the following paragraphs, these methods use ground based sky cameras for observing the local state of clouds. Using mostly using methods that are inspired by the ones used for satellite observations at a smaller scale and at a higher resolution, they generate predictions for the future states of local clouds. Resolutions and area coverage of these methods are not constant like the satellite based methods, as the span of the observed cloud area changes highly with the changes in cloud base height. Though the resolution changes, it can be said that it is roughly on the order of 10-20 meters, as opposed to reported several hundred meters reported by satellite based methods. The cloud observation can be used to obtain parameters, such as, cloud cover, cloudiness, and cloud-base-height, that are used as input by irradiance forecast methods, or they can be used to obtain the information

as to when a certain part of the plant is going to be occluded by a cloud. Irregardless of their use, methods that are applicable, in terms of implementation simplicity and data availability, for this work are going to be presented in this part.

Huang et al. [33] present a hybrid method combining cross-correlation, phase-correlation, and local-feature based cloud motion estimation. Cross correlation and phase correlation are applied on the current (I_t) and the previous (I_{t-1}) frames, and using the resulting motion vectors determined, the current frame is projected to the future to obtain a prediction of the next frame (I_{t+1}). The correlation is applied on image sub-blocks, to determine a motion vector for each block. Afterwards, blocks are moved according to their motion vector to the associated area in the next image. The explained procedure is identical to the correlation based method presented for satellite images in the previous section, the only distinction is that this method is applied on a smaller scale where the considered blocks correspond to smaller geographical areas, on the order of tens of meters, than for the satellite images, on the order of hundreds of meters. Lets assume that, after the correlation method is applied, the block with size $m \times m$ that starts at pixel position (i, j) on time t attains the motion vector (u, v). This means that at time $t - 1$ the associated block had its starting point on the pixel position $(i - u, j - v)$, as can be seen in eq. 2.17.

$$I^t_{(i,j)(i+m,j+m)} \approx I^{t-1}_{(i-u,j-v)(i-u+m,j-v+m)} \tag{2.17}$$

With the assumption that the cloud motion and shape does not change during the observation period, an easy prediction for the position of the associated block would be that it will have its starting point at pixel position $(i+u, j+v)$ in the next time step, as can be seen in eq. 2.18.

$$I^{t+1}_{(i+u,j+v)(i+u+m,j+v+m)} \approx I^t_{(i,j)(i+m,j+m)} \tag{2.18}$$

Also through using local features, another set of motion vectors, each one corresponding to the neighboring area of a single feature, are obtained. It is mentioned that, combining these motion vectors with the motion vectors coming from the correlation based methods, a 25% improvement on the regular correlation based algorithms is obtained. In Figure. 2-10 on page 61, the motion vectors obtained using the implementation of this method can be seen.

As can be seen in Figure. , the vectors are found for all the blocks of the corresponding image.

Similar to the previous work, Marquez and Coimbra use the same idea to obtain the cloud velocity field, but employ an alternative version of cloud tracking. For the cloud velocity field determination, cross correlation based method is used from "MPIV" software package developed by Mori and Chang (2003). For the tracking, instead of predicting the future states of cloud cover, the pixels that are predicted to cover the sun in the prediction horizon, are taken and a cloud cover is calculated for these pixels. k-means algorithm is used for velocity field representative identification. According to this representative vector, which part of the image to consider is determined [54].

In the work presented by Chow et al. [17] a single ground based sky imager (TSI) is used for cloud observation, and for cloud motion detection, cross correlation method is used. Cross correlation is applied on the red component of the contiguous images that are projected to sky coordinates to correct the perspective distortion. The red channel is used, because the red channel image is claimed to have a higher contrast between sky and cloud than other channels. Also, instead of cloud decision image, the red channel image is used to avoid possible errors due to incorrect cloud detections. Cross correlation method takes as input two consecutive images, divides the previous image into rectangular grid of cells, and for each cell it finds the maximum correlated position in the next image. The vector from the cell position in the previous image to maximum correlated position on the next image gives the associated cloud movement. This way, a vector field representing different cloud movements in different parts of the image is obtained. It is assumed that there is a spatial homogeneity of cloud velocity across the image, and the vector field obtained is processed using clustering methods to eliminate the outliers. This way, a representative cloud movement vector can be obtained. To perform cloud forecast, the representative cloud movement vector is used to advect the cloud decision image in that direction. The obtained forecast results are overlaid with the image corresponding to the observation at the forecast horizon, and a pixel-wise error metric, as a ratio of falsely forecast pixels to total number of pixels, is applied. The forecast horizons from 30 seconds up to 5 minutes are evaluated, and it is seen that the cloud matching error scales almost linearly from 6% to

23-30%. The errors are associated with the perspective, cloud detection, and nonuniform cloud speed issues.

Jayadevan et al., in [37], present the use of cross correlation method on an alternative setup using a sun tracking camera. Cross correlation method on pixel blocks is applied on the sky image to obtain cloud motion vectors. Through perspective correction, obtained motion vectors are projected to real world velocities assuming a constant cloud base height at 3000m. After the real world velocities are obtained, cloud arrival times for each motion vector is calculated using a linear interpolation method, assuming that the velocity would remain the same. Using the obtained arrival times, which cloud would arrive the soonest is determined and for that cloud the time for impact is estimated. This way, successful ten minutes ahead forecasts with an accuracy of one minute reported, provided that the circumstances for clouds are ideal (single-layer cloud with no deformation, splitting, disappearance or appearance) [37].

Hashimoto and Nagakura [28] present an alternative method using a stereo camera setup, with 1.2km apart cameras, for cloud observation. Using the stereo camera system with image registration and feature triangulation, it is said that the cloud positions can be measured in 3D with a range direction error of less than 5%. It is claimed that this information is sufficient for determining the direction of movement and the speed of clouds. It is further claimed that, using linear extrapolation, impact time predictions can be obtained for cloud detections, however no results are presented about any application of this. Also the correlation between ground level wind direction and cloud movement direction, and the correlation of the wind speed at high altitudes and cloud movement speed is studied. It is reported that the ground level wind direction and the cloud movement direction are approximately the same, and hence the ground wind direction can be used to obtain the cloud movement direction, possibly to aid the tracking methods. Furthermore, it is noted that there is a difference between the wind speed at cloud altitude, and the cloud movement speed. According to the presented study, the bigger the cloud in dimension is, the slower it is with respect to wind. The difference between the wind speed and the cloud speed is assumed to be due to the shape change in cloud during its movement. The possibility of a method to predict the wind speed at high altitudes using the cloud speed, and the scale of

shape change of the cloud, using the correlations, is also mentioned [28].

2.3.5 Methods using sensor networks on the ground

Though these are few and recent, there are methods that use an array of sensors that can communicate with a center but geographically distributed, to observe and predict the effects of clouds.

Cazorla et al. present one such method in [16]. The proposed method makes use of distributed PV installations from 80 locations for cloud observation. Collects all the measurements from these locations to a central unit, assuming any loss over the clear-sky value is due to clouds, formulates a clearness index ($K(x, y, t) = \frac{GHI_{x,y,t}}{GHI_{clear,x,y,t}}$) as a function of location and time. Then using the cloud velocity (v_x, v_y), the future clearness index, $K(x, y, t + dt)$ for a location (x, y) can be predicted using eq. 2.19.

$$K(x, y, t + dt) = K(x - v_x dt, y - v_y dt, t) \tag{2.19}$$

For the areas that are between measurement points, a four point interpolation method is applied, where the median of four closest measurements is considered as the measurement for the location under consideration. To obtain cloud speed, three methods are investigated: wind velocity from NWP, cloud velocity inference from the network of PV systems, a constant velocity thought the day that is numerically optimized to minimize RMS error. The cloud velocity inference from the network is done finding the best solution (v_x, v_y) to the equation eq. 2.20 for all the locations in the system (x_i, y_i).

$$K_i(t) = K(x_i - v_x dt, y_i - v_y dt, t - dt) \tag{2.20}$$

Using all three cloud velocity estimators, the proposed method is compared against the persistence method on time horizons ranging from 15 minutes to 75 minutes. It is reported that proposed method, that uses the numerically optimized cloud speed, outperforms the persistence method on partially cloudy days by reducing the RMS prediction error by 20%, that are very hard cases for forecasts of this sort. Also a cloud tracking method, that uses

ground based sky images from a sun tracking camera, is mentioned. It is stated that using this method, 10 minute ahead predictions of cloud impacts can be provided, and these can further improve the irradiance predictions.

Kuszamaul et al. introduce another method using 24 irradiance sensors communicating with 19 radio stations every second to characterize surface irradiance and track cloud shadows on the ground [44]. The cloud tracking is accomplished in an alternative way, through inference from irradiance measurements. To observe the cloud shadows, they generate animations of the irradiance field as they measured. As a result of this, they report that, for some cloud cases the cloud shadows and their movements along the area can be observed, while for some cases the shadow behaviour is "chaotic"; shadows appear and disappear randomly, or move in many directions. This chaotic behavior of shadows is associated to complex changes in weather conditions that affect the irradiance measurements, and to possible data losses and repetitions in the RF network of the sensors.

Bosch and Kleissl introduce an alternative representation to the use of irradiance sensors for cloud tracking in [11]. 8 irradiance sensors are put on the perimeter of a semi-circle with a radius of 6m (Figure. 2-11 on page 62), and the data acquisition is done at 20Hz. Two methods are developed for cloud tracking. The first method, assumes that when two sensors observe the same cloud with a time lag, the irradiance they measure would be similar. Using this assumption, the correlation between sensor pair measurements for different time-lags are calculated. For the maximum correlation, using the distance between the associated pair and the time lag between maximally correlated measurements, a cloud speed is estimated. All the sensors are paired with the sensor on the center, which gives 7 pairs, and using the alignment of the maximally correlated pair cloud motion vector is generated. The second method uses, just the central sensor, and sensors with labels N and E (Figure. 2-11 on page 62). Assuming that the cloud has a linear edge, and using geometric relations with the sensor placements, it generates a cloud speed estimation [11].
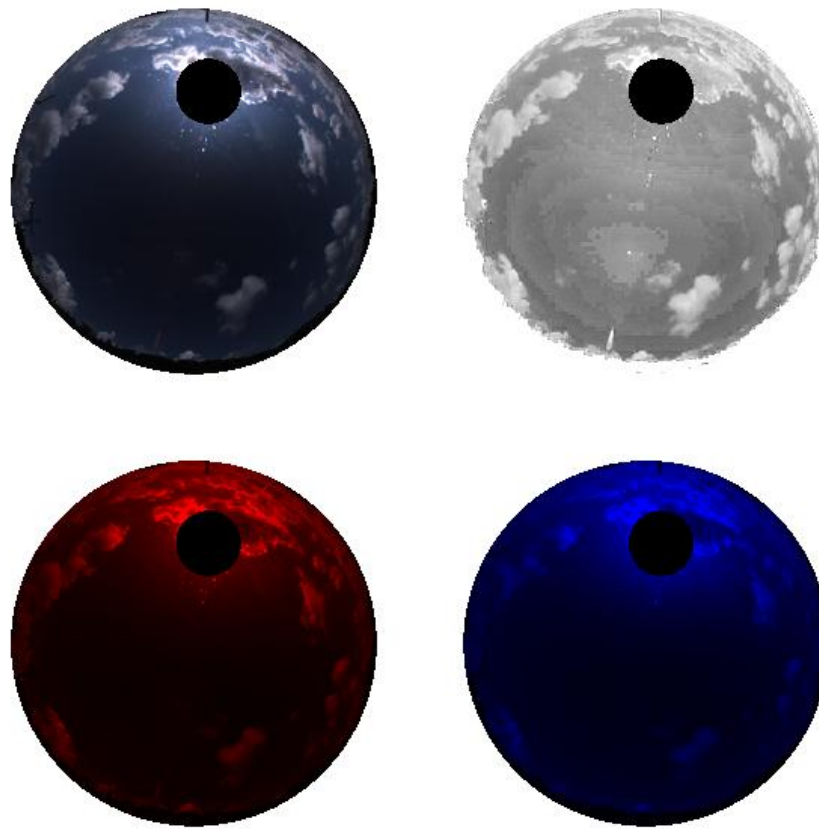
Figure 2-1: The different light scattering and the observed results are seen in this. The top left is the masked RGB image coming from the camera. As can be seen, the sun is covered with a black patch. The top right image corresponds to the pixel-wise R/B ratio of the RGB image. Due to scattering of clouds, higher values for both the red and blue channels are measured on the image, and these areas are observed as white on the RGB image. Sky areas scatter the blue component of light more than they scatter the red component, as a result on the measured image, the blue channel has a higher value than the red channel and the corresponding areas are observed as blue. Cloud pixels has a higher red channel value than sky pixels, as a result in the red channel image (bottom left) cloud pixels are observed much brighter than sky pixels. When we consider the blue channel image (bottom right) we see that the contrast between cloud and sky areas are not as much as the red channel. Correspondingly, the R/B value for a cloud pixel is higher than it is for a clear sky pixel. As can be seen, the cloud areas are much brighter than the clear sky areas on this image. Also the tone change in the clear-sky area under the sun patch suggests that the R/B value changes for pixels with different positions with respect to the image center and the sun. The bottom left is the R component of the image alone. The higher contrast between the clear sky and cloud areas is an important point to note for this image. Also, the clouds look brighter than the sky areas, which is because of the lower scattering of sky, with respect to cloud, on the red component. The bottom right image corresponds to the blue component of the RGB image. As can be seen on this, the contrast between cloud and clear sky areas is considerably lower than the red channel image. This is because, both the cloud and the sky scatters the blue component high. The last point to note is that the clear sky areas are brighter in this image than the red component image. This is because the blue component is scattered much more than the red component by the molecular scattering of the clear sky areas.
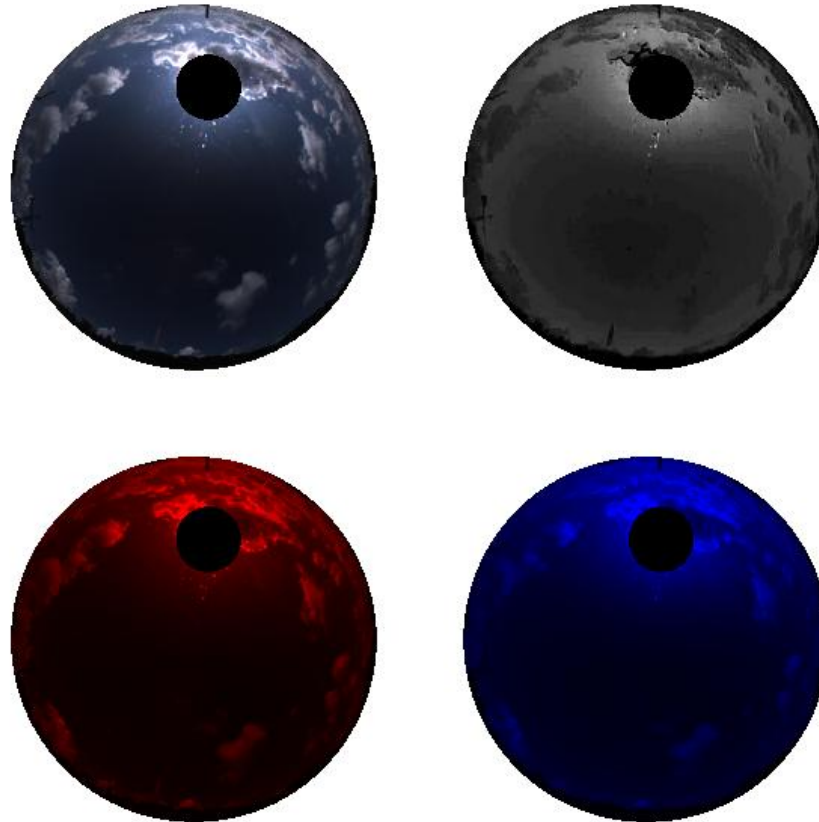
Figure 2-2: The different light scattering and the observed results are seen in this. The top left is the masked RGB image coming from the camera. As can be seen, the sun is covered with a black patch. The top right image corresponds to the pixel-wise R - B difference of the RGB image. As can be seen, the cloud-clear sky contrast in this image is better than the original image, and the cloud areas are much darker than the clear sky areas. The darker a pixel on this image, means the smaller the difference between the red and blue values for that pixel. As explained, the cloud areas scatter both the red and blue components high, and the clear sky areas scatter the blue and do not scatter the red component that much [69]. As a result of this, the difference for cloud areas is smaller than the difference for clear sky areas. Also the tone change in the clear-sky area under the sun patch suggests that the R - B value changes for pixels with different positions with respect to the image center and the sun. The bottom left is the R component of the image alone. The higher contrast between the clear sky and cloud areas is an important point to note for this image. Also, the clouds look brighter than the sky areas, which is because of the lower scattering of sky, with respect to cloud, on the red component. The bottom right image corresponds to the blue component of the RGB image. As can be seen on this, the contrast between cloud and clear sky areas is considerably lower than the red channel image. This is because, both the cloud and the sky scatters the blue component. The last point to note is that the clear sky areas are brighter in this image than the red component image. This is because the blue component is scattered much more than the red component by the molecular scattering of the clear sky areas.
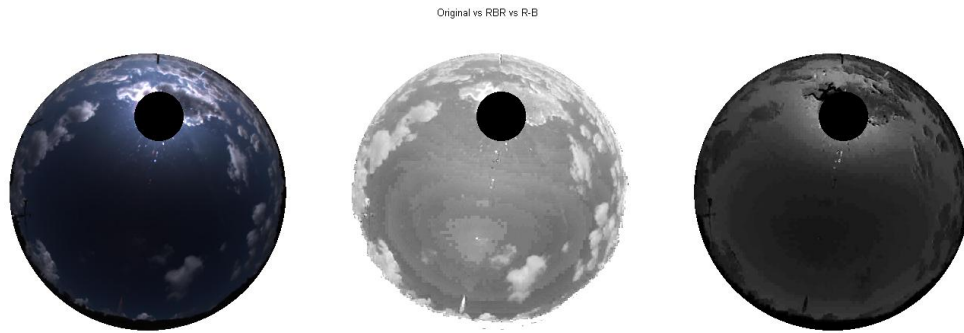
Original vs RBR vs R-B

Figure 2-3: On the left is the original image on which two criteria are applied. On the middle is the RBR criteria applied. On the right is the R-B criteria applied. When we compare these two criteria, we see that the threshold works in the opposite directions. For RBR, pixels with RBR higher than a threshold are decided as cloudy, while for R-B pixels with R-B lower than a threshold are decided as cloudy. Another important point to note is, as can be seen on the original image, towards the bottom right area, in the middle of clear sky there is a cloud. This can be observed as well on the RBR image while it is very hard to notice on the R-B image.Also when the bottom part of the original image (right) is observed, it is seen that there is a little cloud occurrence along the bottom. This cloud is detected bigger than actually it is on the RBR criterion (middle) while the R-B criterion accurately captures that (right).

Figure 2-4: Here, an illustration is given for a satellite image. The contours are defined as can be seen (red lines around clouds) on the image.
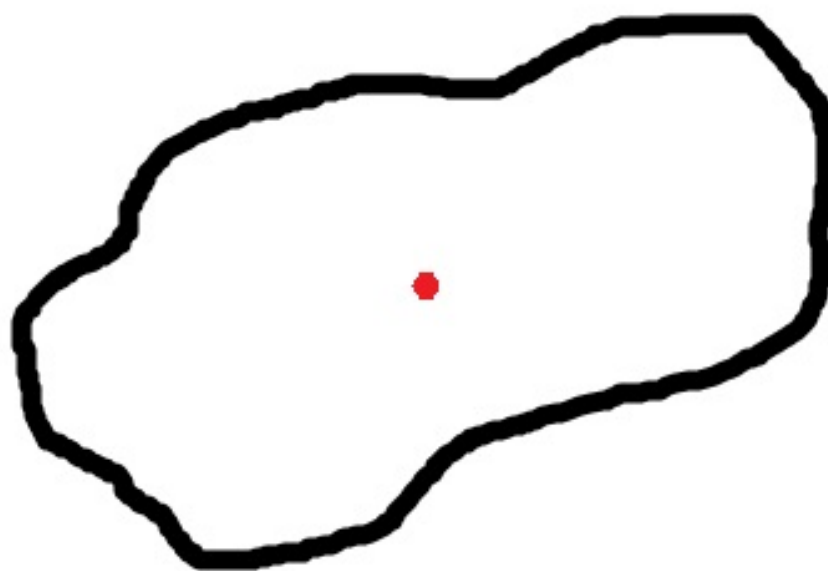
Figure 2-5: An illustration of how the contour and the centroid of the contour is defined.
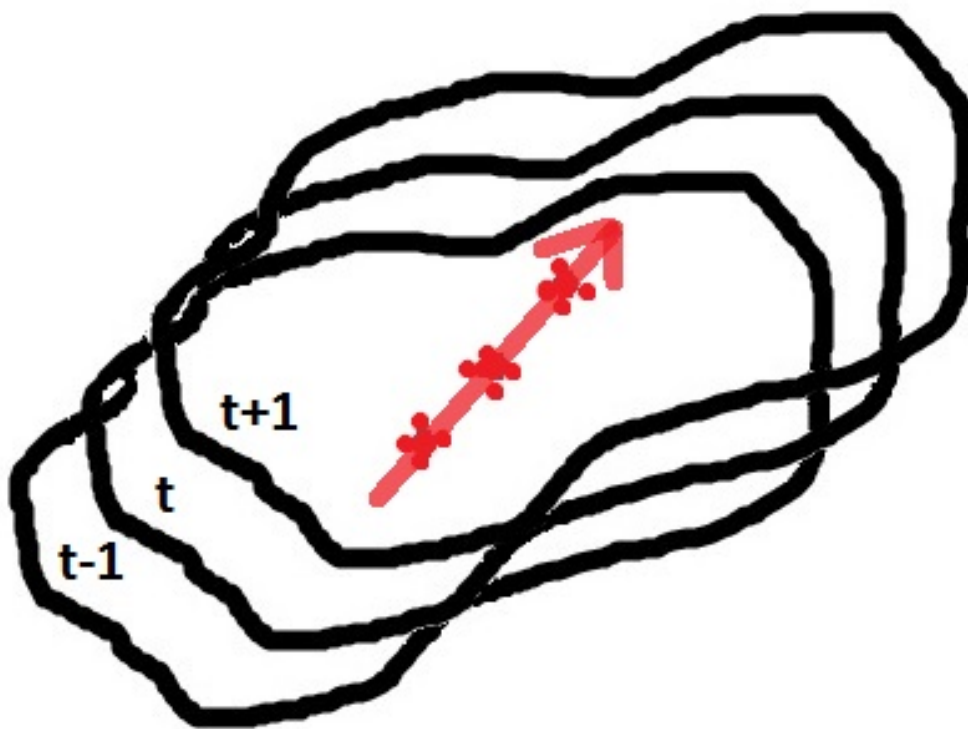Central red dot is the centroid.

Figure 2-6: An illustration presenting how linear interpolation on the contour centroid is applied for cloud forecast. As can be seen, using a linear leas-square regression previous centroid positions, as motion vector is determined, and all the contour points along with the centroid, are moved accordingly to generate the forecast.
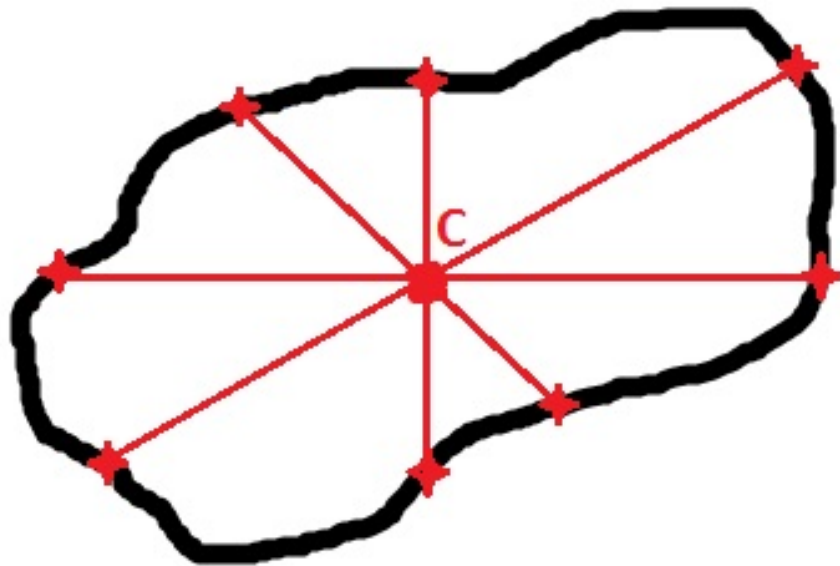
Figure 2-7: An illustration presenting how the contour is defined using the centroid point and the length of lines, with equal angular separation, connecting the centroid to contour. Centroid is the central red dot, while the lengths of representing lines are defined as the distance between centroid and the intersection point of the corresponding line and the contour.
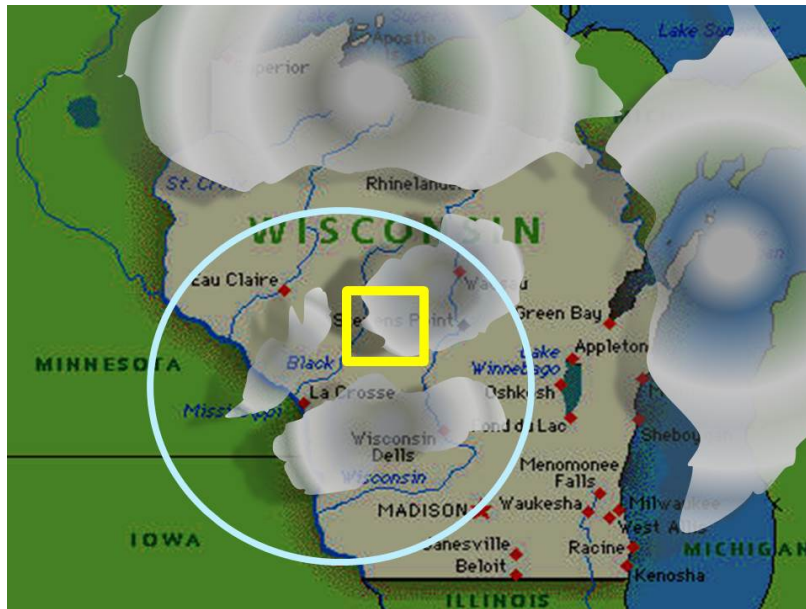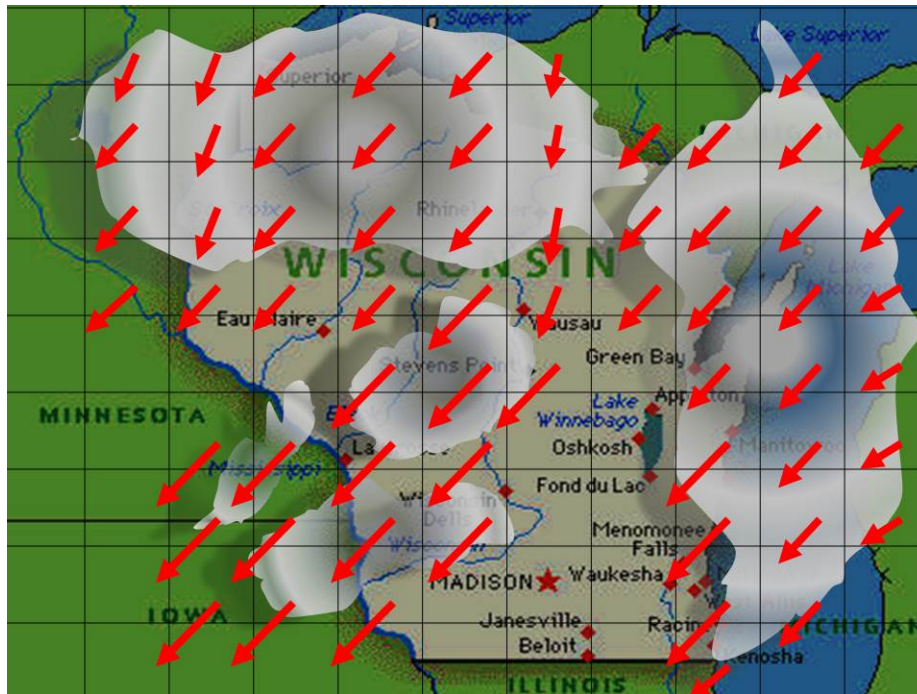
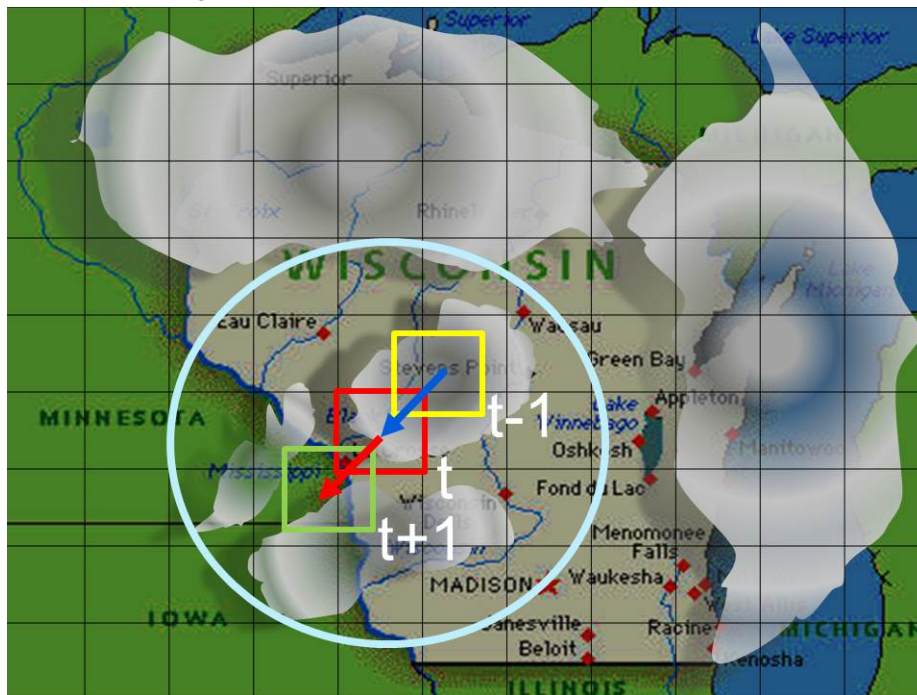(a) An illustration of a satellite based image. This image represent the cloud condition for time t-1
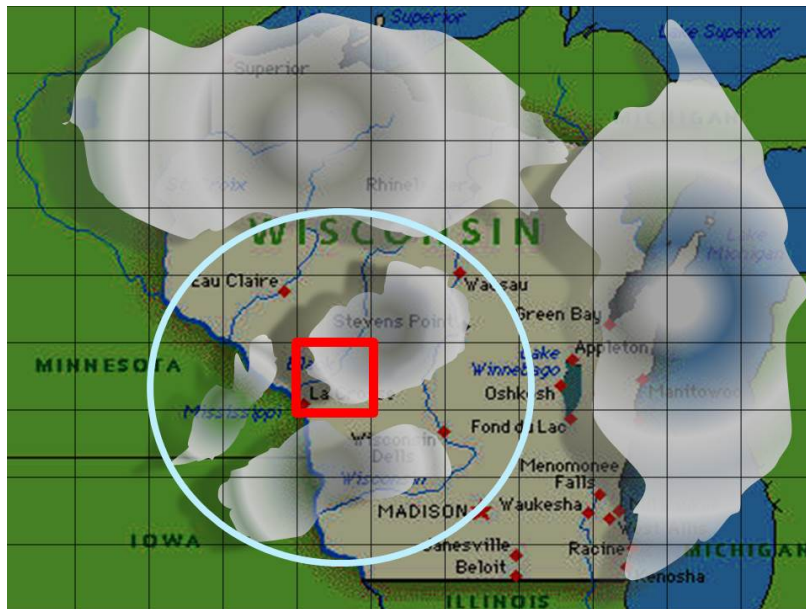
(b) An illustration of a satellite based image. This image represent the cloud condition for time t. The grid seen on the image represents the pixels blocks on which the cross-correlation algorithm is applied. For the grid with red surrounding box, the blue circle represents the area with which the red cell is compared against, to find the maximum correlation.

Figure 2-8: An illustration of how the key components of the cross-correlation method works. The cell with red surrounding in 2-8b is compared against the area in blue circle in 2-8a (the previous image). As a result, for this case the maximum correlation is found for the cell with yellow surrounding in 2-8a. The motion vector for the red cell in 2-8b is generated using the difference its position difference to the yellow cell in 2-8a

(a) An illustration of a satellite based image. As can be seen, for all the blocks, a motion vector is calculated, resulting in a vector field representing cloud movement on the image.

(b) An illustration of a satellite based image. Pixel block in red box, is found to be correlated with the pixel block in yellow box in the previous image. This results in the motion vector between red and yellow boxes. To generate predictions for the next time-step, the red box is moved, according to the motion vector, to the location designated by the green box in the prediction image.

Figure 2-9: An illustration of how the forecast mechanism in cross-correlation method works. Generation of motion vectors using the maximally correlated pixel positions is shown with an example in Figure. 2-9b for one grid cell, in Figure. 2-9a the result of
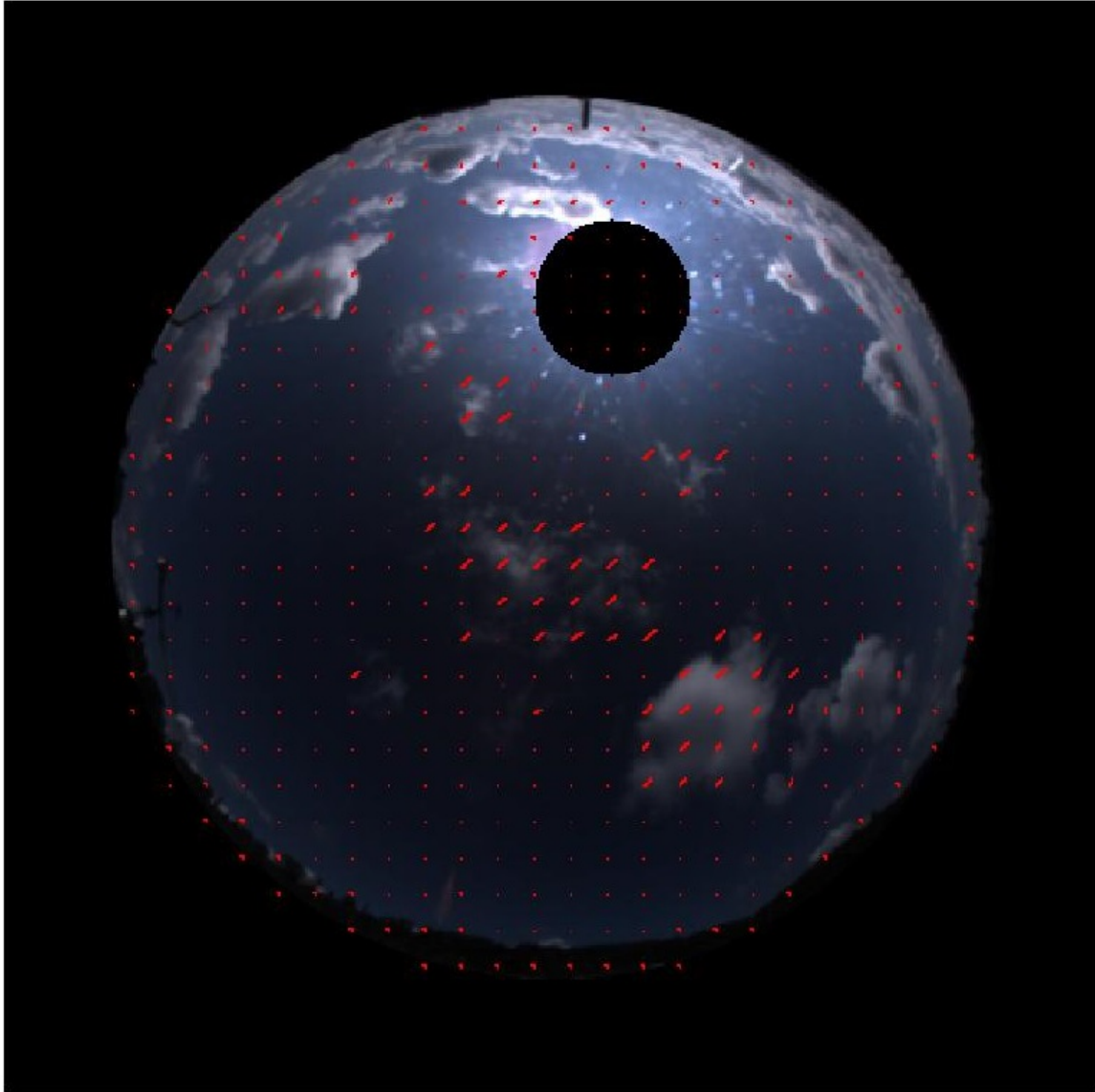
Figure 2-10: As can be seen, there are periodic red marks on the image, these correspond to the motion vectors obtained using correlation with the previous image. Motion vectors are used to place corresponding pixel blocks in the prediction scenario. This image is obtained using a sample implementation of the method on the sky images obtained during this project.
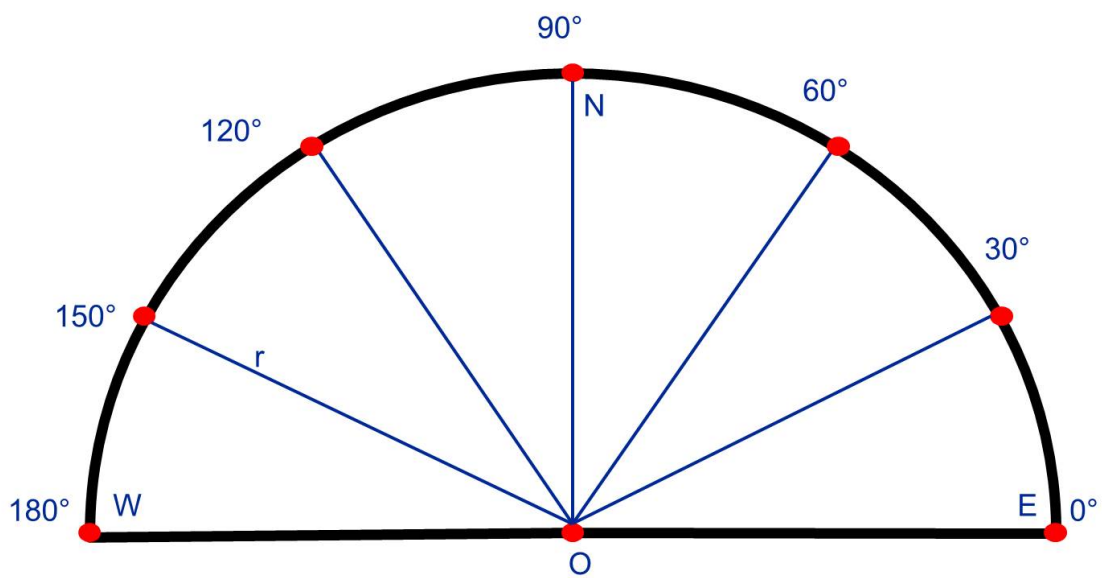
Figure 2-11: The sensor system is as can be seen in this image. Sensors are the red dots on the perimeter of the circle. The radius of the circle is 6m.

Bibliography

[1] The GFS model, Global Forecast System by Environmental Modeling Center, NOAA Washington, DC, 2003. URL http://www.emc.ncep.noaa.gov/index.php?branch=GFS. 20, 21

[2] Global Forecast System Documentation, 2003. URL http://www.emc.ncep.noaa.gov/GFS/doc.php. 20, 21

[3] ECMWF 2010, european center for medium-range weather forecasts, 2010. URL http://www.ecmwf.int. 20

[4] The weather research and forecasting model, 2010. URL http://www.wrf-model.org/index.php. 20

[5] Global Horizontal Irradiance Clear Sky Models : Implementation and Analysis. Technical Report March, Sandia National Laboratories, Albuquerque, New Mexico and Livermore, California, 2012. 23, 24

[6] M. A. Atwater and J. T. Ball. A Surface Solar Radiation Model for Cloudy Atmosphere. Monthly Weather Review, 109(April):878–888, 1981. doi: http://dx.doi.org/10.1175/1520-0493(1981)109<0878:ASSRMF>2.0.CO;2. 22

[7] V. Badescu. Verification of some very simple clear and cloudy sky models to evaluate global solar irradiance. Solar Energy, 61(4):251–264, 1997. 23

[8] F. Batlles, M. Rubio, J. Tovar, F. Olmo, and L. Alados-Arboledas. Empirical modeling of hourly direct irradiance by means of hourly global irradiance. Energy, 25(7):675–688, July 2000. ISSN 03605442. doi: 10.1016/S0360-5442(00)00007-4. URL http://linkinghub.elsevier.com/retrieve/pii/S0360544200000074. 14

[9] A. Bianco and H.-C. Huang. 44

[10] R. Bird and C. Riordan. Simple Solar Spectral Model for Direct and Diffuse Irradiance on Horizontal and Tilted Planes at the Earth's Surface for Cloudless Atmospheres. Technical report, Solar Energy Research Institute, Golden, Colorado, 1984. 24

[11] J. L. Bosch and J. Kleissl. CLOUD VELOCITY ESTIMATION FROM AN ARRAY OF SOLAR RADIATION MEASUREMENTS. In World Renewable Energy Forum (WREF), Denver, Colorado, 2012. American Solar Energy Society. 51

[12] J. Calbó and J. Sabburg. Feature Extraction from Whole-Sky Ground-Based Images for Cloud-Type Recognition. Journal of Atmospheric and Oceanic Technology, 25(1): 3–14, Jan. 2008. ISSN 0739-0572. doi: 10.1175/2007JTECHA959.1. URL http:// journals.ametsoc.org/doi/abs/10.1175/2007JTECHA959.1. 15, 16, 39, 40, 42

[13] J. Cao and S. Cao. Study of forecasting solar irradiance using neural networks with preprocessing sample data by wavelet analysis. Energy, 31(15):3435–3445, Dec. 2006. ISSN 03605442. doi: 10.1016/j.energy.2006.04.001. URL http://linkinghub. elsevier.com/retrieve/pii/S0360544206001009. 26

[14] J. Cao and X. Lin. Application of the diagonal recurrent wavelet neural network to solar irradiation forecast assisted with fuzzy technique. Engineering Applications of Artificial Intelligence, 21(8):1255–1263, Dec. 2008. ISSN 09521976. doi: 10.1016/j.engappai.2008.02.003. URL http://linkinghub.elsevier. com/retrieve/pii/S0952197608000092. 26

[15] S. Cao and J. Cao. Forecast of solar irradiance using recurrent neural networks combined with wavelet analysis. Applied Thermal Engineering, 25(2-3):161–172, Feb. 2005. ISSN 13594311. doi: 10.1016/j.applthermaleng.2004.06.017. URL http:// linkinghub.elsevier.com/retrieve/pii/S1359431104001814. 26

[16] A. Cazorla, F. J. Olmo, and L. Alados-Arboledas. Development of a sky imager for cloud cover assessment. Journal of the Optical Society of America. A, Optics, image science, and vision, 25(1):29–39, Jan. 2008. ISSN 1084-7529. URL http: //www.ncbi.nlm.nih.gov/pubmed/18157209. 37, 50

[17] 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. Solar Energy, 85(11):2881–2893, Nov. 2011. ISSN 0038092X. doi: 10.1016/j.solener.2011.08.025. URL http://linkinghub.elsevier.com/ retrieve/pii/S0038092X11002982. 13, 15, 16, 29, 30, 33, 36, 43, 48

[18] C. Chupong and B. Plangklang. Forecasting power output of PV grid connected system in Thailand without using solar radiation measurement. Energy Procedia, 9(May):230–237, Jan. 2011. ISSN 18766102. doi: 10.1016/j.egypro. 2011.09.024. URL http://linkinghub.elsevier.com/retrieve/pii/ S1876610211017772. 13, 14, 17

[19] S. Clough, M. Shephard, E. Mlawer, J. Delamere, M. Iacono, K. Cady-Pereira, S. Boukabara, and P. Brown. Atmospheric radiative transfer modeling: a summary of the AER codes. Journal of Quantitative Spectroscopy and Radiative Transfer, 91(2):233–244, Mar. 2005. ISSN 00224073. doi: 10.1016/j.jqsrt.2004.05.058. URL http:// linkinghub.elsevier.com/retrieve/pii/S0022407304002158. 21

[20] J. Davies and D. McKay. Evaluation of selected models for estimating solar radiation on horizontal surfaces. Solar Energy, 43(3):153–168, Jan. 1989. ISSN 0038092X.

doi: 10.1016/0038-092X(89)90027-3. URL http://linkinghub.elsevier.com/retrieve/pii/0038092X89900273. 24

[21] J. a. Davies and D. C. McKay. Estimating solar irradiance and components. Solar Energy, 29(1):55–64, Jan. 1982. ISSN 0038092X. doi: 10.1016/0038-092X(82)90280-8. URL http://linkinghub.elsevier.com/retrieve/pii/0038092X82902808. 24

[22] H. M. Diagne, M. David, P. Lauret, and J. Boland. Solar Irradiation Forecasting : State-of-the-Art and Proposition for Future Developments for Small-Scale Insular Grids. In World Renewable Energy Forum, WREF 2012, Including World Renewable Energy Congress XII and Colorado Renewable Energy Society (CRES) Annual Conference Colorado 2012-05-13, pages 1–8, 2012. 24, 25, 26, 28

[23] J. A. Duffie and W. A. Beckman. Solar Engineering of Thermal Processes, volume 53. New York : Wiley, c1991 xxiii, 919 p. : ill. ; 25 cm., New York, 2nd ed. edition, 1985. ISBN 0471698679. doi: 10.1119/1.14178. 14, 17

[24] G. A. Grell, J. Dudhia, and D. R. Stauffer. A description of the fifth-generation penn state/ncar mesoscale model (mm5), 1995. URL http://www.mmm.ucar.edu/mm5/documents/mm5-desc-doc.html. 20, 21, 22

[25] T. M. Hamill and T. Nehrkorn. A Short-Term Cloud Forecast Scheme Using Cross Correlations. Weather and Forcasting, 8(4):401–411, 1993. 15, 33, 42, 45

[26] A. Hammer, D. Heinemann, E. Lorenz, and B. Lückehe. Short-Term Forecasting of Solar Radiation: A Statistical Approach Using Satellite Data. In EUMETSAT Meteorological Satellite Data Users Conference, number 1, pages 331–337, 1999. 15, 46

[27] A. Hammer, D. Heinemann, C. Hoyer, and E. Lorenz. Satellite based short- term forecasting of solar irradianceâŁ"comparison of methods and error analysis, 2001. 14

[28] T. Hashimoto and Y. Nagakura. Prediction of Output Power Variation of Solar Power Plant by Image Measurement of Cloud Movement. 2(2):1–6, 2011. 15, 16, 30, 38, 42, 49, 50

[29] B. Haurwitz. Insolation in relation to cloudiness and cloud density. Journal of Meteorology, 2(September):154–166, 1945. 23

[30] B. Haurwitz. Iinsolation in relation to Cloud Type. Journal of Meteorology, 5(June):110–113, 1948. 23

[31] D. Heinemann, J. Hurka, E. Lorenz, and H. G. Beyer. Irradiance Forecasting for the Power Prediction of Grid-Connected Photovoltaic Systems, volume 2. 2009. doi: 10.1109/JSTARS.2009.2020300. 13

[32] a. Heinle, a. Macke, and a. Srivastav. Automatic cloud classification of whole sky images. Atmospheric Measurement Techniques, 3(3):557–567, May 2010. ISSN 1867-8548. doi: 10.5194/amt-3-557-2010. URL http://www.atmos-meas-tech.net/3/557/2010/. 35, 39, 40

[33] H. Huang, S. Yoo, D. Yu, D. Huang, and H. Qin. Correlation and local feature based cloud motion estimation. In Proceedings of the Twelfth International Workshop on Multimedia Data Mining, MDMKDD '12, pages 1–9, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1556-2. doi: 10.1145/2343862.2343863. URL http://doi.acm.org/10.1145/2343862.2343863. 43, 46, 47

[34] J. Huo and D. Lu. Comparison of Cloud Cover Using All-sky Imager and Meteorological Observer. Journal of Atmospheric and Oceanic Technology, 29(August 2012): 1093–1101, 2012. 15

[35] P. Ineichen and R. Perez. A new airmass independent formulation for the Linke turbidity coefficient. Solar Energy, 73:151–157, 2002. 23, 29, 30

[36] F. M. Janeiro, P. M. Ramos, F. Wagner, and a.M. Silva. Developments of low-cost procedure to estimate cloud base height based on a digital camera. Measurement, 43(5):684–689, June 2010. ISSN 02632241. doi: 10.1016/j.measurement.2010.01.007. URL http://linkinghub.elsevier.com/retrieve/pii/S0263224110000138. 16

[37] V. T. Jayadevan, J. J. Rodriguez, A. D. Cronin, and V. P. A. Lonij. Forecasting Solar Power Intermittency Using Ground-Based Cloud. In World Renewable Energy Forum (WREF), Denver, Colorado, 2012. American Solar Energy Society. 14, 15, 29, 33, 46, 49

[38] W. Ji and K. C. Chee. Prediction of hourly solar radiation using a novel hybrid model of ARMA and TDNN. Solar Energy, 85(5):808–817, May 2011. ISSN 0038092X. doi: 10.1016/j.solener.2011.01.013. URL http://linkinghub.elsevier.com/retrieve/pii/S0038092X11000259. 25

[39] S. Kaplanis and E. Kaplani. A model to predict expected mean and stochastic hourly global solar radiation I(h;nj) values. Renewable Energy, 32(8):1414–1425, July 2007. ISSN 09601481. doi: 10.1016/j.renene.2006.06.014. URL http://linkinghub.elsevier.com/retrieve/pii/S0960148106001649. 26

[40] F. Kasten and G. Czeplak. Solar and Terrestrial Radiation Dependent on the Amount and Type of Cloud. Solar Energy, 24:177–189, 1980. 23

[41] M. L. Kavvas and Z. Chen. the radar based short term prediction of the time space evolution of the rain fields. Technical report, Geophysics Laboratory Air Force Systems Command United States Air Force, Masschusetts, 1989. 44

[42] Y. Kemmoku, S. Orita, S. Nakagawa, and T. Sakakibara. Daily Insolation Forecasting Using a Multi-Stage Neural Network. Solar Energy, 66(3):193–199, June 1999. ISSN

0038092X. doi: 10.1016/S0038-092X(99)00017-1. URL http://linkinghub.elsevier.com/retrieve/pii/S0038092X99000171. 26

[43] T. L. Koehler, R. W. Johnson, and J. E. Shields. Status of the whole sky imager database.pdf. In Cloud Impacts on DOD Operations and Systems, 1991. 34, 35

[44] S. Kuszamaul, A. Ellis, J. Stein, and L. Johnson. LANAI HIGH-DENSITY IRRA-DIANCE SENSOR NETWORK FOR CHARACTERIZING SOLAR RESOURCE VARIABILITY OF MW-SCALE PV SYSTEM. In IEEE Photovoltaic Specialists Conference, pages 4–9, Honolulu , HI, 2010. 51

[45] A. A. Lacis and J. E. Hansen. A Parametrization for the Absorption of Solar Radiation in the Earth's Atmosphere. Journal of Atmospheric Sciences, 31(January):118–133, 1974. 22

[46] V. Lara-Fanego, J. A. Ruiz-Arias, D. Pozo-Vázquez, F. J. Santos-Alamillos, and J. Tovar-Pescador. Evaluation of the {WRF} model solar irradiance forecasts in Andalusia (southern Spain). Solar Energy, 86(8):2200–2217, 2012. ISSN 0038-092X. doi: http://dx.doi.org/10.1016/j.solener.2011.02.014. URL http://www.sciencedirect.com/science/article/pii/S0038092X11000582. 14, 15

[47] Q. Li, J. Yang, and W. Lu. A Hybrid Thresholding Algorithm for Cloud Detection on Ground-Based Color Images. Journal of Atmospheric and Oceanic Technology, 28(10):1286–1296, 2011. ISSN 0739-0572. doi: 10.1175/JTECH-D-11-00009.1. 36

[48] F. Linke. Transmissions-Koeffizient und Trubungsfaktor, volume 10. Beitr. Phys. fr. Atmos., 1922. 30

[49] C. N. Long. Correcting for Circumsolar and Near-Horizon Errors in Sky Cover Retrievals from Sky Images. The Open Athmospheric Science Journal, 4:45–52, 2010. 15

[50] C. N. Long, J. M. Sabburg, J. Calbó, and D. Pagès. Retrieving Cloud Characteristics from Ground-Based Daytime Color All-Sky Images. Journal of Atmospheric and Oceanic Technology, 23(5):633–652, May 2006. ISSN 0739-0572. doi: 10.1175/JTECH1875.1. URL http://journals.ametsoc.org/doi/abs/10.1175/JTECH1875.1. 15, 34, 35

[51] V. P. a. Lonij, V. T. Jayadevan, A. E. Brooks, J. J. Rodriguez, K. Koch, M. Leuthold, and A. D. Cronin. Forecasts of PV power output using power measurements of 80 residential PV installs. 2012 38th IEEE Photovoltaic Specialists Conference, pages 003300–003305, June 2012. doi: 10.1109/PVSC.2012.6318280. URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6318280. 13, 15, 29, 32, 46

[52] E. Lorenz, J. Remund, S. C. Müller, W. Traunmüller, G. Steinmaurer, D. Pozo, J. Antonio, V. L. Fanego, L. Ramirez, M. G. Romeo, C. Kurz, L. M. Pomares, and C. G.

Guerrero. Benchmarking of Different Approaches to Forecast Solar Irradiance. In 24th European Photovoltaic Solar Energy Conference, pages 4199 – 4208, Hamburg, Germany, 2009. 15

[53] P. Lynch. The origins of computer weather prediction and climate modeling. Journal of Computational Physics, 227(7):3431–3444, Mar. 2008. ISSN 00219991. doi: 10.1016/j.jcp.2007.02.034. URL http://linkinghub.elsevier.com/retrieve/pii/S0021999107000952. 20

[54] R. Marquez and C. F. Coimbra. Intra-hour DNI forecasting based on cloud tracking image analysis. Solar Energy, 91:327–336, May 2013. ISSN 0038092X. doi: 10.1016/j.solener.2012.09.018. URL http://linkinghub.elsevier.com/retrieve/pii/S0038092X1200343X. 14, 15, 16, 31, 36, 48

[55] R. Marquez and C. F. M. Coimbra. A Novel Metric for Evaluation of Solar Forecasting Models. In 5th National Conference on Energy Sustainability, Washington, DC, 2011. ASME. 25, 26

[56] R. Marquez, A. Mechanics, C. F. M. Coimbra, and L. Jolla. COMPARISON OF CLEAR-SKY MODELS FOR EVALUATING. In C. Fellows, editor, World Renewable Energy Forum, WREF, Denver, Colorado, 2012. American Solar Energy Society. 26

[57] L. Martín, L. F. Zarzalejo, J. Polo, A. Navarro, R. Marchante, and M. Cony. Prediction of global solar irradiance based on time series analysis: Application to solar thermal power plants energy production planning. Solar Energy, 84(10):1772–1781, Oct. 2010. ISSN 0038092X. doi: 10.1016/j.solener. 2010.07.002. URL http://linkinghub.elsevier.com/retrieve/pii/S0038092X10002379. 13, 15, 24, 25, 27

[58] A. Mellit and A. M. Pavan. A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy. Solar Energy, 84(5):807–821, May 2010. ISSN 0038092X. doi: 10.1016/j.solener.2010.02.006. URL http://linkinghub.elsevier.com/retrieve/pii/S0038092X10000782. 28

[59] a. Mellit, M. Benghanem, and S. Kalogirou. An adaptive wavelet-network model for forecasting daily total solar-radiation. Applied Energy, 83(7):705–722, July 2006. ISSN 03062619. doi: 10.1016/j.apenergy.2005.06.003. URL http://linkinghub.elsevier.com/retrieve/pii/S0306261905000875. 26

[60] a. Mellit, H. Eleuch, M. Benghanem, C. Elaoun, and a. M. Pavan. An adaptive model for predicting of global, direct and diffuse hourly solar irradiance. Energy Conversion and Management, 51(4):771–782, Apr. 2010. ISSN 01968904. doi: 10.1016/j.enconman. 2009.10.034. URL http://linkinghub.elsevier.com/retrieve/pii/S0196890409004427. 13, 14, 15, 27, 28, 29

[61] T. Nehrkorn. Nowcasting Methods for Satellite Imagery. Phillips Laboratory, Directorate of Geophysics, Air Force Materiel Command, 1993. URL http://books.google.ch/books?id=r3BXHAAACAAJ. 42, 43, 44, 45

[62] J. Nijs, A. Woyte, R. Belmans, and V. Thong. Voltage Fluctuations on Distribution Level Introduced by Photovoltaic Systems, volume 21. 2006. doi: 10.1109/TEC.2005. 845454. 13

[63] J. C. Nova, J. B. Cunha, and P. B. D. Moura. Solar Irradiation Forecast Model Using Time Series Analysis and Sky Images. In 5th Conference of the European Federation for Information Technology in Agriculture, Food and Environment (EFITA/WCCA 2005), number July, pages 1408–1415, Vila Real (Portugal), 2005. 15, 31, 37

[64] W. M. Organization. International Cloud Atlas, Vol. 2. International Cloud Atlas Vol. 2. Secretariat of the World Meteorological Organization, 1987. ISBN 9263L24078. URL http://library.wmo.int/pmb_ged/wmo_407_en-v2.pdf. 39, 40

[65] J. K. Page. Prediction of Solar Radiation on Inclined Surfaces, volume 3. Solar Energy R&D in the European Community, Series F: Solar Radiation Data, 1986. 15

[66] G. W. Paltridge, A. Physics, P. O. Box, M. Engineering, and P. O. B. Highett. Monthly Mean Solar Radiation Statistics for Australia. Solar Energye, 18(3):235–243, 1976. 23

[67] R. Perez, S. Kivalov, J. Schlemmer, K. Hemker, D. Renné, and T. E. Hoff. Validation of short and medium term operational solar radiation forecasts in the US. Solar Energy, 84(12):2161–2172, Dec. 2010. ISSN 0038092X. doi: 10.1016/j.solener. 2010.08.014. URL http://linkinghub.elsevier.com/retrieve/pii/S0038092X10002823. 14, 21

[68] R. Perez, E. Lorenz, S. Pelland, M. Beauharnois, G. Van Knowe, K. Hemker, 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. Solar Energy, 94:305–326, Aug. 2013. ISSN 0038092X. doi: 10.1016/j.solener.2013.05.005. URL http://linkinghub.elsevier.com/retrieve/pii/S0038092X13001886. 25

[69] G. Petty. A first course in atmospheric radiation. Sundog Publishing, 2006. ISBN 9780972903318. URL http://books.google.ch/books?id=YpspAQAAMAAJ. 34, 53

[70] M. H. Rahman and S. Yamashiro. Novel Distributed Power Generating System of PV-ECaSS Using Solar Energy Estimation. IEEE Transactions on Energy Conversion, 22(2):358–367, June 2007. ISSN 0885-8969. doi: 10.1109/TEC.2006.870832. URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4207466. 21

[71] L. Rasmussen, H. Conway, and S. Ferguson. Estimation of atmospheric transmittance from upper-air humidity. Solar Energy, 62(5):359 – 368, 1998. ISSN 0038-092X. doi: http://dx.doi.org/10.1016/S0038-092X(98) 00013-9. URL http://www.sciencedirect.com/science/article/pii/S0038092X98000139. 22

[72] S. Rehman. Empirical model development and comparison with existing correlations. Applied Energy, 64(1-4):369–378, Sept. 1999. ISSN 03062619. doi: 10.1016/S0306-2619(99)00108-7. URL http://linkinghub.elsevier.com/retrieve/pii/S0306261999001087. 14

[73] G. Reikard. Predicting solar radiation at high resolutions: A comparison of time series forecasts. Solar Energy, 83(3):342–349, Mar. 2009. ISSN 0038092X. doi: 10.1016/j.solener.2008.08.007. URL http://linkinghub.elsevier.com/retrieve/pii/S0038092X08002107. 26

[74] J. Remun, R. Perez, and E. Lorenz. Comparison of Solar Radiation Forecasts for the USA. In European PV Conference, volume 2, pages 3–5, Valencia, Spain, 2008. 15

[75] L. Robledo and A. Soler. Luminous efficacy of global solar radiation for clear skies. Energy Conversion and Management, 41:1769–1779, 2000. 23

[76] M. G. Romeo, L. Ramirez, P. Fernandes, S. Lozano, and X. Nicuesa. A system of direct radiation forecasting based on numerical weather predictions , satellite image and machine learning. In International Symposium on Forecasting, pages 154–155, 2011. URL http://www.forecasters.org/submissions/RamirezLourdesISF2011.pdf. 15

[77] G. Seiz and M. Baltsavias. CLOUD MAPPING USING GROUND-BASED IM-AGERS. XXXIII:1349–1356, 2000. 15

[78] G. Seiz, E. P. Baltsavias, and A. Gruen. Cloud Mapping from the Ground : Use of Photogrammetric Methods. 68(9):941–951, 2002. 15

[79] G. Seiz, J. Shields, U. Feister, E. P. Baltsavias, and A. Gruen. Cloud Mapping With Ground Based Photogrammetric Cameras. International Journal of Remote Sensing, 28(9):2001–2032, 2007. 15

[80] A. Sfetsos and A. H. Coonick. UNIVARIATE AND MULTIVARIATE FORECAST-ING OF HOURLY SOLAR. 68(2):169–178, 2000. 26, 27, 28

[81] T. Stoffel, D. Renné, D. Myers, S. Wilcox, M. Sengupta, R. George, and C. Turchi. CONCENTRATING SOLAR POWER Best Practices Handbook for the Collection and Use of Solar Resource Data. Technical Report September, National Renewable Energy Laboratory, 2010. 17, 18

[82] C. Tao, D. Shanxu, and C. Changsong. Forecasting power output for grid-connected photovoltaic power system without using solar radiation measurement.

The 2nd International Symposium on Power Electronics for Distributed Generation Systems, pages 230–237, June 2010. doi: 10.1109/PEDG.2010.5545754. URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5545754. 17, 21

[83] F. Wang, Z. Mi, S. Su, and H. Zhao. Short-Term Solar Irradiance Forecasting Model Based on Artificial Neural Network Using Statistical Feature Parameters. Energies, 5(12):1355–1370, May 2012. ISSN 1996-1073. doi: 10.3390/en5051355. URL http://www.mdpi.com/1996-1073/5/5/1355/. 13, 14, 20, 27, 28

[84] M. Xue, K. K. Droegemeier, V. Wong, A. Shapiro, and K. Brewster. The Advanced Regional Prediction System (ARPS), 1995. URL http://twister.ou.edu/ARPS/Index.htmlv. 20, 21, 22

[85] M. Xue, K. K. Droegemeier, V. Wong, A. Shapiro, and K. Brewster. ARPS Version 4.0 User's Guide. Technical report, Center for Analysis and Prediction of Storm[Available from CAPS, Univ. of Oklahoma, 100 E. Boyd St., Norman OK 73019], 1995. 22

[86] J. Yang, W. Lu, Y. Ma, and W. Yao. An Automated Cirrus Cloud Detection Method for a Ground-Based Cloud Image. Journal of Atmospheric and Oceanic Technology, 29(April 2012):527–536, 2012. 15, 16, 21, 34, 36, 37, 39

[87] K. Yang and T. Koike. Estimating surface solar radiation from upper-air humidity. Solar Energy, 72(2):177–186, Feb. 2002. ISSN 0038092X. doi: 10.1016/S0038-092X(01)00084-6. URL http://linkinghub.elsevier.com/retrieve/pii/S0038092X01000846. 21, 22

[88] K. Yang, G. Huang, and N. Tamai. A hybrid model for estimating global solar radiation. Solar Energy, 70(1):13–22, Jan. 2001. ISSN 0038092X. doi: 10.1016/S0038-092X(00)00121-3. URL http://linkinghub.elsevier.com/retrieve/pii/S0038092X00001213. 22