

# Cloud gazing and catching the sun's rays:

## Quantifying cloud-caused variability in solar irradiance

Esther Peerlings

March 9, 2019

M.SC. THESIS





# Cloud gazing and catching the sun's rays:

Quantifying cloud-caused variability in solar irradiance

Esther Peerlings

March 9, 2019

*Supervised by:*

dr.ir. Chiel van Heerwaarden

*Examined by:*

prof.dr. Bert Holtslag

M.Sc. THESIS

METEOROLOGY AND AIR QUALITY GROUP

DEPARTMENT OF ENVIRONMENTAL SCIENCES

WAGENINGEN UNIVERSITY & RESEARCH





## Abstract

The global energy transition transforms the world's energy sector from fossil-based towards based on renewable energies. As each cloud condition causes different solar irradiance characteristics, quantifying cloud-caused variability in solar irradiance is important for the successful grid integration of weather-dependent decentralized solar energy systems. This study quantifies the cloud-caused variability in the 1-minute averaged solar irradiance measurements from the Veenkampen weather station near Wageningen, the Netherlands, for the period June 2011 to December 2017. We expand on previous studies and create a classification scheme with nine classes that represent certain cloud conditions. The variability in the solar irradiance magnitude and temporal variability are quantified per cloud condition with: (1) the mean diffuse and direct solar irradiance, (2) the probability density function of the clear-sky index, (3) the temporal autocorrelation of the solar irradiance, and (4) the probability density function of the ramp rates in solar irradiance. The low-variability clear-sky class mostly experiences variability in solar irradiance due to the solar cycle in the sky. The low-variability overcast class experiences low variability in magnitude, but this variability has a random nature after about 10 minutes. The high-variability mixed-sky class has high variability in magnitude, within one minute the solar irradiance can drop with  $670 \text{ W m}^{-2}$ . The temporal variability in this class is high as well, as the variability becomes random already after 4 minutes. The classification scheme and belonging results can be used to reduce the uncertainty in solar irradiance forecasts, thereby helping grid operators with balancing the electricity load on current and future electricity grids.



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Theoretical background</b>	<b>5</b>
<b>3</b>	<b>Data and methods</b>	<b>7</b>
3.1	Site description . . . . .	7
3.2	Data collection . . . . .	7
3.2.1	Solar irradiance observations . . . . .	7
3.2.2	Solar irradiance under clear sky . . . . .	8
3.3	Data processing . . . . .	8
3.3.1	Solar irradiance observations . . . . .	8
3.3.2	Solar irradiance under clear sky . . . . .	9
3.4	Variability quantification . . . . .	9
3.4.1	Variability indices . . . . .	9
3.4.2	Classification scheme of cloud conditions . . . . .	10
3.4.3	Probability density function of solar irradiance per cloud condition . . . . .	11
3.4.4	Temporal autocorrelation of solar irradiance per cloud condition . . . . .	12
3.4.5	Ramp rates in clear-sky index per cloud condition . . . . .	13
<b>4</b>	<b>Results</b>	<b>15</b>
4.1	Seasonal variability in solar irradiance . . . . .	15
4.2	Classification scheme of cloud conditions . . . . .	17
4.3	Diffuse and direct components of solar irradiance per cloud condition . . . . .	18
4.4	Probability density function of solar irradiance per cloud condition . . . . .	20
4.5	Temporal autocorrelation of solar irradiance per cloud condition . . . . .	22
4.6	Ramp rates in clear-sky index per cloud condition . . . . .	23

<b>5</b>	<b>Discussion</b>	<b>31</b>
<b>6</b>	<b>Conclusions</b>	<b>35</b>
<b>7</b>	<b>Acknowledgements</b>	<b>37</b>







# 1 | Introduction

The global energy transition transforms the world's energy sector from fossil-based towards based on renewable energies. The combined share of variable solar and wind power in the energy sector is expected to increase to 58% in 2050 (IRENA, 2018). Almost half of the solar capacity growth over 2018-2023 will consist of decentralized energy generated by commercial and industrial projects and residential applications (IRENA, 2018). However, traditional neighbourhood electricity grids are designed for centralized energy generated by non-variable, often non-renewable, energy sources like coal fired power plants. In case of a high photovoltaic (PV) power system density in a neighbourhood, all these PV power systems generate about their maximum amount of power when the sun shines and generate only little power when the sky is clouded.

On a sunny day with fair weather clouds, solar irradiance fluctuates within minutes between values as low as clouded sky conditions and as high as clear-sky conditions. On a day like in Figure 1.1, this solar irradiance difference can be in 1 minute as large  $670 \text{ W m}^{-2}$  or a clear-sky index difference of 0.73. Furthermore, measured solar irradiance also frequently exceeds the values of clear-sky conditions, because clouds also reflect sunlight thereby adding diffuse radiation to direct radiation (e.g., Schade et al., 2007). As PV systems transform intermittent solar irradiance into solar energy, the passing of clouds causes similar fast and large intensity fluctuations in solar energy production. These fluctuations can result in an electricity load on the electricity grid beyond acceptable limits. This exceedance can cause local voltage-flicker issues, damage electromechanical equipment and even lead to a blackout (Lave et al., 2012). Furthermore, fluctuations accelerate the fatigue of PV modules materials (Tomson, 2010). Weather-dependent decentralized solar energy generation thus challenges grid operators to balance electricity supply and demand in order to to guarantee the stability and quality of the electricity grid.

For the successful grid integration of decentralized PV power systems, it is important to quantify cloud-variability in solar irradiance (e.g., Woyte et al., 2007). Each cloud condition causes different solar irradiance characteristics, therefore a complete quantification of variability should

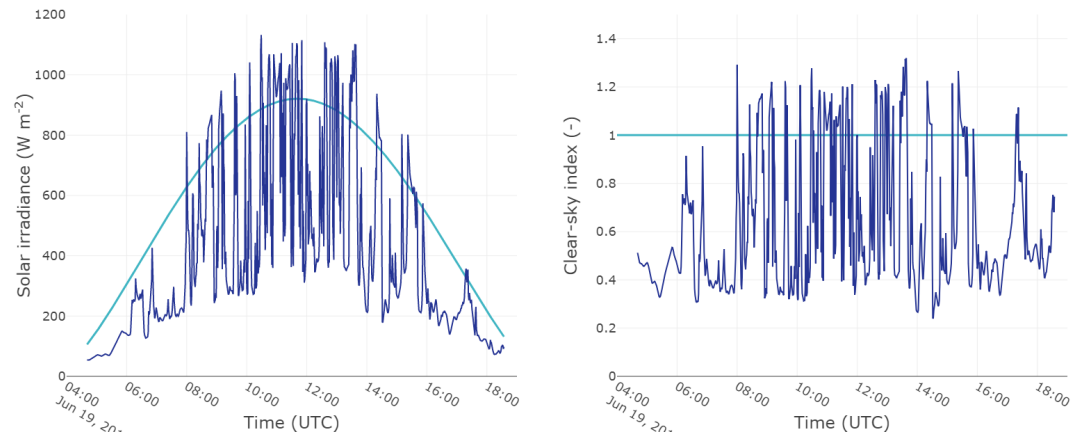


Figure 1.1: (a) Solar irradiance and (b) the clear-sky index on a day with fair-weather clouds (19 June 2016). The temporal resolution of the time series is 1 minute. The dark blue line in panel (a) is the observed solar irradiance and in (b) the clear-sky index. The light blue line is in panel (a) the modelled clear-sky irradiance and in (b) the clear-sky index of 1.

describe per cloud condition: (1) the physical quantity that varies, (2) the variability time scale, and (3) the time period over which the variability is assessed (Perez et al., 2016). So point (1) and (2) are respectively the variability in the solar irradiance quantity and in solar irradiance ramp rates (i.e., changes over specified intervals of time) (e.g., Lave et al., 2012). Quantifying the variability in solar irradiance is primarily relevant for forecasting the yield of PV systems and the capacity of energy storage (Lohmann, 2018). Quantifying the ramp rates in solar irradiance is primarily relevant for the balancing of the power supply and electricity load on an electricity grid, and for guaranteeing the stability and quality of the electricity grid (Lohmann, 2018). The quantification of solar irradiance variability can also help to decrease the forecast uncertainty of solar irradiance and solar energy forecast models. Precisely forecasting solar energy generation is even highlighted as one of the key challenges to enable the extensive integration of PV systems needed to succeed in the global energy transition (European Photovoltaic Industry Association and others, 2012; Barth et al., 2014).

All solar irradiance measurements deviating from the modelled clear-sky irradiance are caused by clouds (Figure 1.1). As solar irradiance can fluctuate within minutes between clouded and clear-sky values, solar forecast models need to be able to capture this high temporal variability. Clouds, and therefore solar irradiance, can be seen as the final (but ever changing) stage of complex dynamics between (1) land surface (providing heat and water that form clouds), (2) atmospheric

flow (transporting heat and water), and (3) microphysics (processes that lead to cloud development, movement and dissipation). Of all standard meteorological variables, solar irradiance is the most affected by clouds. Clouds absorb, block, and scatter solar irradiance, however, these properties depend on the clouds' position, structure and composition (e.g. condensation nuclei, water vapour and ice crystals) relative to the point of measurement. Although the physics of the complex dynamics and cloud properties are known, the turbulence processes give clouds a stochastic nature and therefore clouds become difficult to parameterise and model (Kleissl, 2013). Hence, more research on quantification of the stochastic cloud-caused variability in solar irradiance is important for electricity grid operation (Lohmann, 2018). Especially, variability quantification on time series at (sub)minute scale are essential as these time series are quite scarce, but highly relevant for quantifying the stochastic nature of solar irradiance (Lohmann, 2018).

Our aim is to quantify cloud-caused in solar irradiance. There is much literature available on quantifying variability in solar irradiance and its underlying processes in space and time. Three recently published papers, Blaga and Paulescu (2018), Lohmann (2018) and Schroedter-Homscheidt et al. (2018) have reviewed a comprehensive set of literature on quantifying variability in solar irradiance. Different methods exist in capturing the variability in solar irradiance time series and up to this moment there is no method that is found most suitable (Blaga et al., 2019).

Although some studies quantify the actual magnitude in solar irradiance expressed in  $\text{W m}^{-2}$ , most studies make use of the clear-sky index. This index is the normalization of measured solar irradiance to the modelled clear-sky irradiance and therefore removes the daily and seasonal trends in solar irradiance caused by the sun's position (e.g., Liu and Jordan, 1960). To quantify the variability in the quantity of solar irradiance, typically studies use the probability density function (e.g., Olseth and Skartveit, 1984), the cumulative distribution/fractional time distribution (e.g., Suehrcke and McCormick, 1988), and/or a (variation on) the standard deviation (e.g., Skartveit and Olseth, 1992).

The previous methods, however, are independent of time and thus do not quantify how the clear-sky index changes over time. Several methods exist to define the temporal variability, often it consists of some kind of statistics on the ramp rate (e.g., Hoff and Perez, 2010). Several studies first want to classify the time series into classes of a classification scheme. The classification scheme is divided according to some measure(s) of variability in solar irradiance and can represent for example cloud conditions (e.g., Stein et al., 2012). As the variability in solar irradiance is depended on clouds, quantifying the temporal variability in solar irradiance per cloud condition would therefore be more representative. Although the temporal autocorrelation is found to be a good measure of variability over time according to Skartveit and Olseth (1992), to our knowledge no study has applied the

temporal autocorrelation on the clear-sky index time series per class of a classification scheme.

The objectives of this study are to quantify the variability in solar irradiance's quantity and ramp rates per cloud condition. We use the solar irradiance data set as measured per minute at the Veenkampen weather station near Wageningen, the Netherlands, for the period June 2011 to December 2017. We produce a cloud condition classification scheme with nine classes based on the hourly mean clear-sky index and the hourly standard deviation of the 1-minute ramp rate (Section 3.4.2). Per cloud condition class we quantify the variability quantity via a probability density function and the temporal variability via a temporal autocorrelation and a PDF on the ramp rates.

A theoretical background on the interaction between clouds and solar irradiance is provided in Chapter 2. Chapter 3 reports on the data handling and the methods of classifying the solar irradiance time series according a cloud condition classification scheme. The chapter also reports on the methods of quantifying cloud-caused variability in solar irradiance per cloud condition. The results are presented in Chapter 4 and discussed in Chapter 5. Finally, conclusions are drawn in Chapter 6.

## 2 | Theoretical background

The flux of the radiation emitted by the sun that reaches the world's atmosphere is called the solar constant and is equal to  $1376 \text{ W m}^{-2}$  (World Meteorological Organization, 1982). The solar constant can vary slightly due to for example sun spots and solar flares, but these generally do not affect the solar constant by more than 0.1% (Kipp & Zonen, 2018b). When solar radiation passes through the atmosphere, solar radiation is scattered, absorbed and transmitted by atmospheric molecules, aerosol particles, trace gases and clouds. The amount of shortwave incoming radiation (UV, visible and near infrared) eventually received on a horizontal plane at the Earth's surface is called the global horizontal irradiance (*GHI*). The SI unit for the measure of solar irradiance (radiative flux) is  $\text{W m}^{-2}$ .

Variability in solar irradiance has three causes: (1) the solar cycle, so the sun's position in the sky, which causes precisely predictable seasonal and diurnal variations; (2) atmospheric molecules, aerosol particles and trace gases (with an exception of clouds) which cause deterministic variability that is most pronounced on a daily temporal resolution (Schepanski et al., 2015); (3) the cloud's position, structure, composition and velocity which are responsible for the stochastic and intermittent variability in solar irradiance. As the cloud-caused variability in solar irradiance is difficult to parameterise and model (Kleissl, 2013), we specifically want to quantify this variability.

The *GHI* consists of two components: the diffuse horizontal irradiance (*DHI*) and the direct beam normal irradiance (*BNI*). The *DHI* is all solar irradiance that indirectly reaches the point of measurement via scattering on atmospheric molecules, aerosol particles, trace gases and clouds. The *BNI* is the received direct irradiance from the sun and its aureole (not the from the surrounding sky) on a plane perpendicular to the *BNI*. As the *BNI* always reaches the horizontal plane in an angle, the *GHI* can be calculated following Equation 2.1:

$$GHI = \cos(\theta) \cdot BNI + DHI \quad (2.1)$$

, where  $\theta$  is the local solar zenith angle (*SZA*) (vertical above the location is  $0^\circ$ , horizontal is  $90^\circ$ ).

In concentrating solar or thermal power plants, the solar irradiance is focused onto a receiver with mirrors or lenses. These mirrors or lenses face the sun and track the movement of the sun

throughout the day by changing their east west orientation. For these applications, the global normal irradiance ( $GNI$ ) and  $BNI$  are especially of interest (e.g., Olseth and Skartveit, 1987).  $GNI$  is the solar irradiance received at a plane directed normal to the sun that tracks the sun's movements and is calculated by means of Equation 2.2:

$$GNI = BNI + DHI \quad (2.2)$$



## 3 | Data and methods

This chapter first presents a site description of the weather station in Section 3.1. Section 3.2 focuses on the collection of the solar irradiance time series. Section 3.3 reports on the data handling of the solar irradiance time series. Finally, Section 3.4 describes the methods of quantifying cloud-caused variability in the solar irradiance and clear-sky index time series.

### 3.1 Site description

This study uses the in situ *GHI*, *BNI* and *DHI* 1-minute averaged measurements from the Veenkampen weather station. The weather station is located northwest of the city Wageningen in the province Gelderland of the Netherlands (51.9814 N, 5.6217 E) and is owned by Wageningen University & Research. The local landscape is characterized by flat grasslands and the elevation of the weather station is 5.0 m above sea level. The region has a prevailing southwest wind direction and is classified as a temperate oceanic climate (Cfb, Köppen climate classification).

### 3.2 Data collection

#### 3.2.1 Solar irradiance observations

The *GHI* and *DHI* are measured respectively with a Kipp & Zonen CMP11 pyranometer (Kipp & Zonen, 2018a) and a Hukseflux SR11 pyranometer (Hukseflux Thermal Sensors B.V., 2015) that store 1 minute averages. To measure the *DHI*, a shading ball on a tracker blocks the *BNI*, so only indirect irradiance via the atmosphere reaches the pyranometer. The measurement uncertainty (95% confidence interval) of the *GHI* pyranometer is  $\pm 10 \text{ W m}^{-2}$  (Kipp & Zonen, 2018a). The achievable uncertainty (95% confidence level) of the minute totals of the *DHI* pyranometer is  $\pm 6.5\%$  (Hukseflux Thermal Sensors B.V., 2015). The *BNI* is measured with a Hukseflux Thermal Sensor DR01 pyrheliometer that stores 1 minute averages (Hukseflux Thermal Sensors, 2018). The achievable uncertainty (95% confidence level) of the minute totals is  $\pm 1.8\%$  (Hukseflux Thermal Sensors B.V.,

2016). Our investigation period is from June 2011 to December 2017, as the pyranometers and pyrliometer that provide measurement data sets with an 1-minute temporal resolution are installed in June 2011.

### 3.2.2 Solar irradiance under clear sky

The clear-sky irradiance is the solar irradiance under a cloud-free sky. It includes the sun's seasonal and diurnal cycle and also the composition of the atmosphere like the aerosol optical depth (Lefèvre et al., 2013). So the difference between measured solar irradiance and modelled clear-sky irradiance is caused by clouds. We collected modelled clear-sky irradiance with an 1-minute temporal resolution for the investigation period June 2011 to December 2017 for the same coordinates as the Veenkampen weather station. The modelled  $GHI$ ,  $BNI$  and  $DHI$  under cloud-free conditions ( $GHI_{clear}$ ,  $BNI_{clear}$  and  $DHI_{clear}$ ) were obtained from the Copernicus Atmosphere Monitoring Service (CAMS) McClear Clear Sky Irradiance service version 3.1 (Lefèvre et al., 2013). The bias in the CAMS McClear  $GHI_{clear}$  comprises between  $-6$  and  $25 \text{ W m}^{-2}$  (Copernicus Atmosphere Monitoring Service, 2016). The bias in the CAMS McClear  $BNI_{clear}$  comprises between  $-48$  and  $33 \text{ W m}^{-2}$  (Copernicus Atmosphere Monitoring Service, 2016).

## 3.3 Data processing

### 3.3.1 Solar irradiance observations

After careful inspection, solar irradiance measurements are being omitted at timestamps when one or more of the measurement instruments had an instrument failure. This failure could express as: (1) a data gap, (2) a double timestamp, (3) a non-realistic or reliable measurement value. The time series timestamps with a failure of type (3) were identified after careful personal interpretation and were mostly found around periods with data gaps or double timestamps.

The measured solar irradiance and clear-sky irradiance data are selected for timestamps when  $SA$  was between  $0^\circ$  and  $80^\circ$ . This  $SA$  range is to eliminate (1) night time values; (2) early morning and late evening periods when the pyranometers are subject to errors in cosine response; and (3) moments when measurements are potentially contaminated by the solar reflection and beam blocking due to trees, buildings and topography near the horizon (e.g., Mathiesen and Kleissl, 2011; Aguiar et al., 2015). The  $SA$  per minute has been calculated following Stull (2000).

### 3.3.2 Solar irradiance under clear sky

To quantify cloud-caused variability, the first step is to separate the effect of clouds from the deterministic seasonal and diurnal variability in solar irradiance time series. To achieve this, some studies focus on the actual magnitude of solar irradiance, but most studies normalize the measured solar irradiance  $GHI(t)$  to the modelled clear-sky irradiance  $GHI_{clear}(t)$  (e.g., Lohmann, 2018). This normalization results in the clear-sky index ( $k^*$ ). We calculated the clear-sky index for each minute in the investigation period, when  $SZA < 80^\circ$ , by means of Equation 3.1:

$$k^*(t) = \frac{GHI(t)}{GHI_{clear}(t)} \quad (3.1)$$

Figure 1.1 presents in panels (a) and (b) the  $GHI$  and clear-sky index for 19 June 2016. The clear-sky index shows some typical features. For example, the lowest values are never zero, as even the darkest clouds still do not absorb all solar irradiance. Furthermore, the highest values can exceed a clear-sky index of 1.0. This phenomenon is called cloud enhancement and is primarily due to short-term reflections of clouds and for a small degree due to limitations in the clear-sky models (Lohmann, 2018). Schade et al. (2007) investigated cloud enhancement events and found the strongest cloud enhancement of  $522 \text{ W m}^{-2}$  during nearly overcast altocumulus clouds and fractional cumulus clouds.

As mentioned in Section 3.2.2, the  $GHI_{clear}$  can have a bias of  $-6$  and  $25 \text{ W m}^{-2}$ . This bias is especially influential when investigating moments that the clear-sky index exceeds 1.0. We have excluded all timestamps in the time series when the period that the clear-sky index exceeded 1.0 took longer than 20 minutes. It is impossible to determine the real bias in the  $GHI_{clear}$  time series for each minute, but we wanted to exclude unrealistic values. The probability density was 0.003 for the time periods that the clear-sky index exceeded 1.0 took 20 minutes. Schade et al. (2007) stated that the durations of cloud enhancement events depend on their strength and range from 20 seconds (enhancements  $> 400 \text{ W m}^{-2}$ ) up to 140 seconds (enhancements  $> 200 \text{ W m}^{-2}$ ). Inman et al. (2016) found that cloud enhancement events have a time scale of 4 minutes. Therefore, the two studies confirm our choice of 20 minutes.

## 3.4 Variability quantification

### 3.4.1 Variability indices

Different methods exist for quantifying the variability in solar irradiance time series, but up to this moment there is no general consensus on which method is most suitable (Blaga et al., 2019). A

straightforward method to express the variability in a subset of a clear-sky index time series of length  $T$ , is by calculating the subset's mean ( $\overline{k^*}$ ) and standard deviation ( $\sigma^{k^*}$ ). However, the mean and standard deviation are independent on the measurements' ordering in time. In other words, they do not take into account how the clear-sky index changes over time. A useful measure for intermittency is a measure that determines the change of  $k^*$  over a specified time interval  $\tau$ , in other words the increment or ramp rate ( $\Delta k_\tau^*$ ) (e.g., Friedrich et al., 2011). The ramp rate is calculated via Equation 3.2:

$$\Delta k_\tau^* = k^*(t + \tau) - k^*(t) \quad (3.2)$$

### 3.4.2 Classification scheme of cloud conditions

Some studies have applied a classification scheme to group subsets of solar irradiance or clear-sky index time series according to the subset's variability characteristics (e.g., Stein et al., 2012). These variability classes are also indicators for cloud conditions. To obtain a quantification of variability in solar irradiance per cloud condition, we first grouped hours of the *GHI* and clear-sky index time series in a classification scheme. As cloud conditions can change heavily during the day, more accurate quantifications of variability in solar irradiance per cloud condition can be found when grouping hours instead of days. Our classification scheme is based on two statistics: the hourly mean clear-sky index ( $\overline{k^*}$ ) and the hourly nominal variability ( $\sigma^{\Delta k^*}$ ). In these equations,  $t$  is 1 minute,  $\tau$  is 1 minute and  $T$  is 60 minutes. The hourly mean clear-sky index divides the data set on the variability in the quantity. The hourly nominal variability divides the data set on the temporal variability, as it is the standard deviation of the 1-minute ramp rate in the clear-sky index time series. Some studies express variability by calculating the maximum or mean ramp rate over a time interval (e.g., Hoff and Perez, 2010). However, we define variability as the nominal variability which is in line with most studies (e.g., Perez et al., 2016). This method takes into account the whole spectrum of variability, not just the maximum and mean.

Aguiar et al. (2015) performed a similar study on 1 minute *GHI* time series when  $SZA < 80^\circ$ . However, they defined a classification scheme on the same statistics, but classified days ( $T$  is 1 day) instead of hours. Aguiar et al. (2015) divided the classification scheme in nine classes (three categories per statistic) which we used as well. The mean clear-sky index is divided into the classes overcast, mixed-sky and clear-sky, with limits of respectively 0 - 0.5, 0.5 - 0.9 and  $> 0.9$ . The nominal variability is divided into the classes low-variability, medium-variability and high-variability, with limits of respectively 0 - 0.05, 0.05 - 0.15 and  $> 0.15$ . The limits of the classes were defined by Aguiar et al. (2015) after an empirical observation of the local weather.

Another study worth mentioning is Lohmann et al. (2016). They have performed a very similar study, but defined their classification scheme on  $\overline{k^*}$  and  $\sigma^{k^*}$ , instead of  $\overline{k^*}$  and  $\sigma^{\Delta k^*}$ . Another difference is that the clear-sky index data set of Lohmann et al. (2016) consists of measurements from 99 synchronized silicon photodiode pyranometers operating at 1 Hz temporal resolution. Therefore, this high temporal resolution allowed them to classify non-overlapping windows of 900 s (15 minutes) and the high amount of measurements allowed them to make 25 classes. Our pyranometer is thermal load based and therefore cannot achieve this high temporal resolution, but the measurement accuracy is higher than silicon photodiode pyranometers. As our data set has a temporal resolution of 1 minute, we classify solar irradiance hours in 9 classes.

To define what an hour is, we do not simply choose for  $T$  whole hours (e.g. 08.00 - 08.59 hour) to calculate  $\overline{k^*}$ . This is because: (1) moments of sunrise and sunset shift throughout the year, thus some hours will not have 60 measurements; and (2) the correlation of solar irradiance over time will be lost to some extent, especially for the first and last values in an hour. Therefore, we first calculate the centred moving average clear-sky index ( $k_{ma}^*(t)$ ) as in Equation 3.3 and then the hourly mean of the centred moving average ( $\overline{k_{ma}^*}$ ) as in Equation 3.4:

$$k_{ma}^*(t) = \frac{1}{T} \sum_{i=0}^{T-1} k^*(t + \frac{1}{2}(T-1) - i) \quad (3.3)$$

$$\overline{k_{ma}^*} = \frac{1}{T} \sum_{t=1}^T k_{ma}^*(t) \quad (3.4)$$

, where  $T$  is the amount of minutes in each whole hour per day in the time series (e.g. 2017/01/01 08.00 - 2017/01/01 08.59 hour). Also, by performing this centred moving average, each day loses the first and last 30 minutes in the time series for which  $SZA < 80^\circ$  applied.

To calculate the nominal variability, we first calculate the centred moving average ramp rate ( $\Delta k_{ma}^*(t)$ ) and hourly mean of the centred moving average ramp rate ( $\overline{\Delta k_{ma}^*}$ ) in a similar way as to Equations 3.3 and 3.4. The hourly nominal variability we used for making the classification scheme is calculated as in Equation 3.5:

$$\sigma^{\Delta k_{ma}^*} = \sqrt{\frac{1}{T} \sum_{t=1}^T (\Delta k_{ma}^*(t) - \overline{\Delta k_{ma}^*})^2} \quad (3.5)$$

### 3.4.3 Probability density function of solar irradiance per cloud condition

The classification scheme has nine classes, with each class representing a cloud condition. To quantify the cloud-caused variability in the solar irradiance's quantity, we calculate per class' clear-sky index

time series the kernel density estimate (KDE) of the probability density function (PDF) (Waskom, 2018). The PDF provides the information on the range of possible values and the probability associated to those values. As mentioned before, the variability in the clear-sky index time series is only caused by clouds and not by, e.g., the sun's position in the sky.

#### 3.4.4 Temporal autocorrelation of solar irradiance per cloud condition

The first way we quantify the temporal variability in the solar irradiance time series, is by making a temporal autocorrelation on the classes' *GHI* time series. An temporal autocorrelation shows the autocorrelation between the observed solar irradiance at the initial moment *lag 0* and the solar irradiance in the future *lag x*. As the temporal resolution of the solar irradiance time series is 1 minute, *lag x* represents *x minutes* in the future. A high autocorrelation indicates that there is a well predicted trend in the solar irradiance time series, for example,  $t = 0$  has clear-sky conditions and  $x$  minutes in the future also has clear-sky condition. An temporal autocorrelation of zero indicates that the variability in the solar irradiance time series is completely random.

We first made a temporal autocorrelation per class of the classification scheme on the clear-sky index time series, but these gave impractical outcomes. For example, the variability was very minimal in absolute values in the clear-sky index time series of the low-variability clear-sky class, but relatively large and random. As a result the temporal autocorrelation went to zero in only 2 minutes, whereas a temporal autocorrelation of close to 1.0 would be expected for a much longer period. Therefore we choose to make a temporal autocorrelation per class on the *GHI* time series. However, the cloud-caused temporal variability is thus not isolated from the temporal variability caused by the sun's position in the sky.

The time series per class consist of non-consecutive hours. Therefore, we can only first calculate the temporal autocorrelations per separate hour per class and then average these per lag into one temporal autocorrelation that is representative for that class. We use the `pandas.series.autocorr` function that calculates the temporal autocorrelations via the Pearson correlation (Pandas-Dev, 2019). We have calculated the 95% confidence interval as:  $\pm \frac{1.96}{\sqrt{N}}$ , where 1.96 is the 95<sup>th</sup> quantile of the normal distribution and  $N$  is the sample size. In our case,  $N$  is equal to 60, as an hour contains 60 minutes. The confidence interval is thus  $0 \pm 0.253$  and we assume the time series to be random when the temporal autocorrelation lies within this range.

The temporal autocorrelation provides information on the decorrelation time of the solar irradiance time series and thus also on how random the time series is (Barnett et al., 1998). However, to our knowledge only Skartveit and Olseth (1992) used temporal autocorrelations to quantify the

temporal change of solar irradiance. They have calculated the lag one autocorrelation of the short-time clearness index time series for different averaging times. Skartveit and Olseth (1992) stated that autocorrelations with larger lags would only be reliable when the data records are sufficiently long, but do not state how long this period should be.

#### **3.4.5 Ramp rates in clear-sky index per cloud condition**

The second way we quantify the temporal variability in solar irradiance time series, is by making PDFs on the ramp rates in the classes' *GHI* time series. Many studies quantify the temporal variability by making PDFs or cumulative probability distribution functions (CDFs) on the ramp rates in the clear-sky index or solar irradiance time series (e.g., Lave and Kleissl, 2010). We have calculated per class the ramp rates for the following four time intervals (the  $k^*$  in Equation 3.2 is replaced with *GHI*): 1, 5, 10 and 30 minutes. The PDFs are calculated similar to the PDFs of the solar irradiance data set, namely with a kernel density estimate (KDE) of the probability density function (PDF) (Waskom, 2018). An advantage over the temporal autocorrelation is that the PDFs of the ramp rates show the probability of quantities. PDFs on the ramp rates are also useful because they can be made on the clear-sky index and do not face the problem of relative variability that forced us to use temporal autocorrelations on solar irradiance instead of on the clear-sky index.





## 4 | Results

This chapter presents the results. Section 4.1 presents the monthly means of solar irradiance variables. Section 4.2 shows the results of the classification scheme with nine classes representing different cloud conditions. Section 4.3 presents the classes' mean diffuse and direct solar irradiance components. The results of the quantification of the variability in the quantities of the clear-sky index per class is presented in Section 4.4. Sections 4.5 and 4.6 present the results of the quantification of the temporal variability in the solar irradiance per class.

### 4.1 Seasonal variability in solar irradiance

Cloud-caused variability in solar irradiance has a (sub)minute temporal scale, but this variability translates into cloud-caused variability on a seasonal temporal scale. Monthly means of solar irradiance components show how clouds affect solar irradiance per month (Figure 4.1) (Liu and Jordan, 1960). The variability on a seasonal scale is affected, next to clouds, also by the daily and seasonal solar cycle. The daily solar cycle gives  $GHI_{clear}$  a parabolic shape: irradiance increase after sunrise until solar noon and decreases it until sunset. Due to the seasonal solar cycle,  $GHI_{clear}$  also has a seasonal parabolic shape with lowest values on winter solstice (21 to 22 December) and highest on summer solstice (21 to 22 June). The seasonal parabolic shape in monthly mean  $GHI_{clear}$  is not completely symmetric, because of two factors: (1) the solar solstices are not half-way the months, and (2) the scattering and absorption of solar irradiance by atmospheric molecules, aerosols and trace gases differs per month.

The monthly means of  $GHI$  also form a parabolic shape similar to  $GHI_{clear}$ . The highest monthly means of  $GHI$  are found for the months May, June, July and August, and do not deviate more than  $10 \text{ W m}^{-2}$  from each other ( $351 \pm 5 \text{ W m}^{-2}$ ). Suggesting that the cloudiness in these months dampen the effect of the seasonal solar cycle on solar irradiance. December has the lowest monthly mean  $GHI$  ( $89.6 \text{ W m}^{-2}$ ). Per month, the difference between the monthly mean  $GHI_{clear}$  and  $GHI$  is caused by scattering and absorption by clouds. The largest difference between monthly mean  $GHI_{clear}$  and  $GHI$

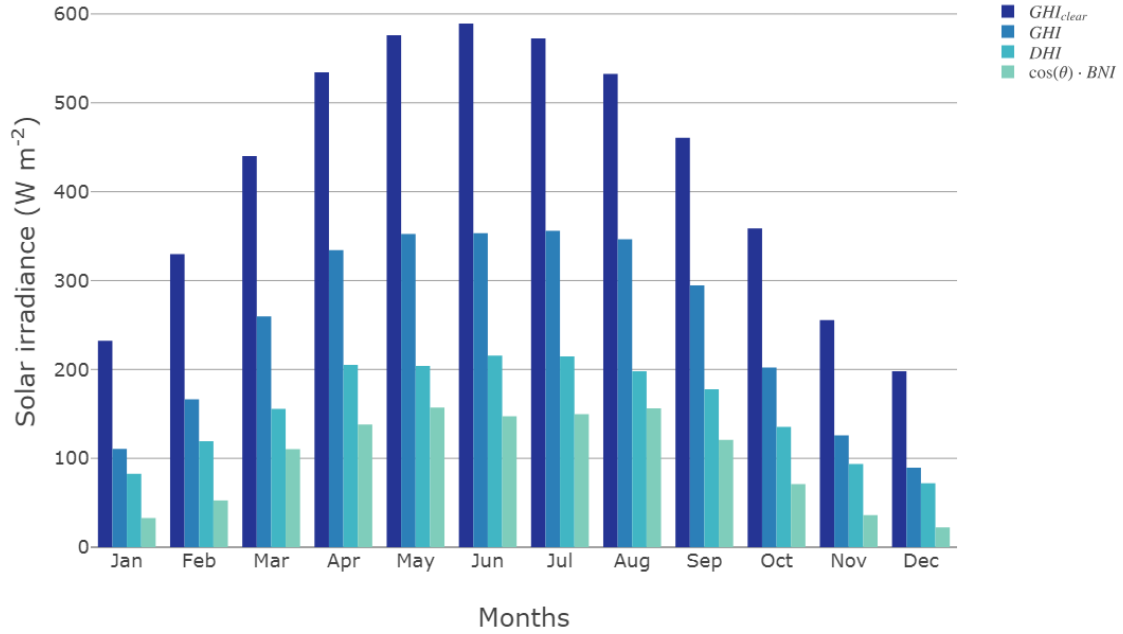


Figure 4.1: The monthly mean of the clear-sky global horizontal irradiance ( $GHI_{clear}$ ), observed global horizontal irradiance  $GHI$ , diffuse horizontal irradiance ( $DHI$ ) and the horizontal component of the beam normal irradiance ( $\cos(\theta) \cdot BNI$ ).

is found in June and is  $235.9 W m^{-2}$ . The smallest difference is found in December and is  $108.5 W m^{-2}$ . However, these results do not indicate that June is the most overcast month and December the most clear-sky month. The absolute differences still include the seasonal and daily solar cycle.

The ratio between  $GHI_{clear}$  and  $GHI$  eliminates the seasonal and daily solar cycle, and is able to provide the monthly mean cloudiness. This ratio is the monthly mean clear-sky index. The order of months from highest clear-sky index to lowest clear-sky index is different than the order from highest to lowest absolute difference between  $GHI_{clear}$  and  $GHI$ . June is only the sixth clearest month (clear-sky index of 0.59) and December is actually the most overcast month (clear-sky index of 0.45). The months clearer than June are: August, with on average the highest clearest sky (clear-sky index of 0.64), then September (0.63), April (0.62), July (0.61) and lastly May (0.60). This shows that quantifying cloud-caused variability in solar irradiance expressed in  $W m^{-2}$  provides different information than the variability expressed in the clear-sky index.

Clouds do not only vary the  $GHI$ , they also vary the two components of  $GHI$ : the  $DHI$  and

$\cos(\theta) \cdot BNI$ . The  $DHI$  is the received solar irradiance on a horizontal plane that is first scattered on atmospheric particles and clouds. Clouds thus increase the relative contribution of  $DHI$  to  $GHI$ . On the opposite, the  $\cos(\theta) \cdot BNI$  is the received solar irradiance on a horizontal plane from the sun and its aureole, which means that clouds only have a blocking effect on the relative contribution of  $\cos(\theta) \cdot BNI$  to  $BNI$ . For each month, monthly mean  $DHI$  is higher than monthly mean  $\cos(\theta) \cdot BNI$ . Indicating that in all months, clouded conditions prevail over clear-sky conditions. If a month is more clouded than another, the relative contribution of  $DHI$  should be higher and  $\cos(\theta) \cdot BNI$  should be lower. This, however, does not always hold when we compare the monthly clear-sky index with the relative contributions of  $DHI$  and  $\cos(\theta) \cdot BNI$ . The order from highest to lowest monthly mean clear-sky index and relative contribution of  $\cos(\theta) \cdot BNI$  to  $GHI$  are for eight months the same or deviate one month. But the rank in these orders for the months March, April, May and September do not coincide. The months' relative contribution of  $\cos(\theta) \cdot BNI$  is higher and the relative contribution of  $DHI$  to  $GHI$  is lower than expected. An explanation could be is that the sun's position in the sky also the  $DHI$  and  $\cos(\theta) \cdot BNI$  affects. A lower sun elevation means that solar irradiance needs to travel a longer distance through the atmosphere, so it is more likely to scatter on atmospheric molecules, thereby increasing the relative contribution of  $DHI$  and decreasing the relative amount of  $\cos(\theta) \cdot BNI$  to the  $GHI_{clear}$ .

## 4.2 Classification scheme of cloud conditions

The nature of cloud-caused variability is dependent on the cloud conditions, for example, stratus clouds affect solar irradiance differently compared to cumulus clouds. Therefore, to quantify cloud-caused variability in solar irradiance, it is useful to group similar cloud conditions and quantify variability in solar irradiance per cloud condition (Stein et al., 2012). To group similar cloud conditions, we have made a classification scheme extending on other studies (e.g. Stein et al., 2012). Figure 4.2 shows the classification scheme with nine cloud condition classes (low-variable overcast conditions AI) to high-variable clear-sky conditions CIII). The shades of blue indicate the joint probability density of the hourly mean clear-sky index (Equation 3.4) and the hourly nominal variability (Equation 3.5). The joint probability density figure is also called an Arrow Head figure due to it's shape (e.g., Stein et al., 2012), because the classes AI, BII, BIII and CI have the highest joint probabilities. In summary, at the Veenkampen, the Netherlands, on hourly basis, low-variable overcast conditions occur most often (33.8%). Medium- and high-variable mixed-sky conditions occur 18.3% and 12.8% respectively of the time. Low-variable clear-sky conditions occur 12.6% of the time. The classification scheme shows

that all medium- and high-variable cloud conditions occur about 50% of the time, highlighting the relevance to quantify the stochastic cloud-caused variability in solar irradiance. Fractional cumulus clouds are most likely the main cloud conditions responsible for the medium- and high-temporal variability in the clear-sky index (Inman et al., 2016).

The joint probability density in the low-variable clear-sky class (CI) is high ( $>> 120$  hours) around the hourly mean clear-sky index between 0.9 - 1.0, and hourly nominal variability between 0.0 - 0.01. However, these limits are narrow. The integrated joint probability densities are higher for the medium- and high-variable mixed-sky conditions than for the low-variable clear-sky class, even though this is perhaps not immediately expected from the figure. The joint probability density is relatively low for the low-variable mixed-sky class (BI: 6.8%). Thus in situations with mixed-skies, the probability is 81.2% higher to have at least medium-temporal variability in the clear-sky index. Cloud conditions for the BI class could be, e.g., constant cirrostratus clouds or a change from a clear to overcast sky in case of the passing of a pressure front. The high-variable overcast (AIII) and clear-sky (CIII) classes consist of the least amount of hours, respectively 0.8% and 1.8%.

### 4.3 Diffuse and direct components of solar irradiance per cloud condition

The classes of the classification scheme represent cloud conditions. To provide a general quantification of how these cloud conditions affect solar irradiance, we have plotted the classes' mean  $DHI$  and  $\cos(\theta) \cdot BNI$ . The sum of these two per class is the class' mean  $GHI$ . The mean  $GHI$  increases from class AI ( $109.9 \text{ W m}^{-2}$ ) to CIII ( $561.0 \text{ W m}^{-2}$ ). The BI class is the only exception, its mean  $GHI$  is  $10.4 \text{ W m}^{-2}$  lower than the AIII class. The increase from classes A to B and then to C is logic as the classes are defined from overcast to mixed-sky and then to clear-sky. The mean  $GHI$  also increases from the variability classes I, via II to III. For the A and B classes this is probably due to: (1) that clouded moments alternate more often into clear-sky moments, and/or (2) that the cloud become thinner. Note, the mean values are slightly higher than expected from the seasonal variability in solar irradiance, because the first and last half hour of all days are removed when classifying hours into the classification scheme.

The  $\cos(\theta) \cdot BNI$  measures the solar irradiance only from the sun and its aureole, which means that clouds only have a blocking effect on the  $\cos(\theta) \cdot BNI$ . Clouds cause the highest difference between  $GHI$  and  $GHI_{clear}$  when they lie in the direct path between the sun and the pyrheliometer. In those cases  $\cos(\theta) \cdot BNI$  changes from its maximum to a lower value. This lower value depends on mostly the thickness and composition of the cloud. We see indeed that  $\cos(\theta) \cdot BNI$  increases in

the A and B classes from the variability classes I to III. Although the C classes' mean  $GHI$  increase from variability classes I to III, the  $\cos(\theta) \cdot BNI$  does not. The CI has the highest mean  $\cos(\theta) \cdot BNI$  ( $360.7 \text{ W m}^{-2}$ ), than the CIII class ( $350.7 \text{ W m}^{-2}$ ), and the CII has the lowest mean  $\cos(\theta) \cdot BNI$  ( $327.2 \text{ W m}^{-2}$ ). The CI is the most clear-sky class of all, so it is logic that this class has the highest mean  $\cos(\theta) \cdot BNI$ . However, apparently the two reasons why in the A and B classes the  $\cos(\theta) \cdot BNI$  increases with increasing variability class, do not apply for the CII and CIII classes. In these two classes the mean  $DHI$  increases even more (from  $175.2 \text{ W m}^{-2}$  for CII to  $212.1 \text{ W m}^{-2}$  for CIII). This could suggest that in the CII class, the  $GHI$  varies relatively more between clear-sky moments and clouds blocking  $GHI$ . Whereas in the CIII class, the  $GHI$  varies between clear-sky moments and cloud enhancement events.

The classes' mean  $DHI$  have a slightly different pattern compared to the classes' mean  $GHI$  and  $\cos(\theta) \cdot BNI$ . It also increases from classes A to B and from variability classes I to III. But the C classes have a lower mean  $DHI$  than the A and B classes. As mentioned before clouds absorb and block solar irradiance. Therefore, the overcast classes receive less solar irradiance than the mixed-sky classes, which is why the mean  $DHI$  is lower for the A classes ( $180.0 \text{ W m}^{-2}$ ) than the B classes ( $239.0 \text{ W m}^{-2}$ ). However, clouds also scatter solar irradiance. This is why the relative contribution of  $DHI$  to  $GHI$  is higher for the A classes than for the B classes. As the C classes have a relatively low cloud amount, scattering is low, and thus have a lower relative contribution of  $DHI$  to  $GHI$ . Within the A, B and C classes, the  $DHI$  always increases from class I to III. This is too some extent explained by that the  $GHI$  also increases. But to another extent this is explained by the fact that apparently, scattering is stronger on the type of clouds that cause high variability in solar irradiance.

However, the previous interpretations on cloud-caused variability would be completely valid if the classes' mean  $GHI_{clear}$  would be the same, but this is not the case. Within the A, B and C classes, the mean  $GHI_{clear}$  increases from low- to high-variability conditions. Thus the low-variability classes consist on average more out of time stamps in early and/or late daytime and/or in winter month. The high-variability classes thus consist on average most out of time stamps around noon and/or in summer months. Furthermore, each individual class just has a different mean  $GHI_{clear}$ . However, the mean  $GHI_{clear}$  also support results like that solar irradiance is the most absorbed and blocked by clouds for the A classes.

The differences in mean  $GHI_{clear}$  show that the solar cycle thus also influences the identification of the classes in the classification scheme, and therefore that the clear-sky index is dependent on the sun's position in the sky (e.g. Kleissl, 2013). This has two reasons. First, solar irradiance consists relatively more out of  $DHI$  when the sun's position is low in the sky as it needs to travel a larger

distance in the atmosphere and is thus more likely to get absorbed and scattered by atmospheric molecules (see also Section 4.1). This causes the range of the clear-sky index to reduce when the sun's elevation in the sky decreases (Kleissl, 2013) because of two possibilities: (1)  $DHI$  has such a large contribution that clouds have less effect in reducing the  $GHI$ , and (2) if the CAMS McClear model performs not great on the relationship between the sun's elevation and the  $DHI$ , the  $GHI$  can never reach the  $GHI_{clear}$ . Second, the clear-sky index is affected by the relative positions between the sun, cloud and point of measurement, because of the:  $SZA$ , cloud base height, cloud structure and cloud composition. In other words, the same cloud causes a different variability in the clear-sky index at, e.g., noon and sunrise. This having stated, for the interpretation of the coming results we assume that the clear-sky index is independent on the solar cycle which we believe is a relatively fair assumption to make.

#### 4.4 Probability density function of solar irradiance per cloud condition

The previous section showed the classes' mean  $DHI$ ,  $\cos(\theta) \cdot BNI$ ,  $GHI$  and  $GHI_{clear}$ . However, a mean value does not provide the information on the variability around that mean. Furthermore, the previous section showed that it is difficult to split the influence of the daily and seasonal solar cycle to the influence of clouds on solar irradiance. Therefore, to quantify the variability in the quantity of solar irradiance, we have made per class of the classification scheme a PDF of the clear-sky index time series (Figure 4.4). The PDFs of the classes are ordered in the same layout as in the classification scheme (Figure 4.2), with an increase in the hourly mean clear-sky index of classes A to C and an increase in the nominal variability of classes I to III. Logically, the overcast classes have relatively low clear-sky index values with modes between a probability density of 0.15 and 0.35. Also logically, the clear-sky classes have relatively high clear-sky index values with modes between 0.95 and 1.05. The PDFs of the clear-sky classes are much narrower and thus higher than the other classes, with the modes between a probability density of 3.5 and 14.5. The differences between all PDFs are caused by clouds and how the classes are defined. Firstly, clouds increase the variability in clear-sky index values and overcast conditions lower the hourly mean clear-sky index. Furthermore, overcast conditions are defined for a broader range of hourly mean clear-sky index (0.0 - 0.5) than the clear-sky conditions (0.9 - 1.0). Finally, enhancement values (clear-sky index values higher than 1.0) are caused by clouds as they add  $DHI$  to  $\cos(\theta) \cdot BNI$  (e.g., Schade et al., 2007; Inman et al., 2016). These cloud enhancement events can therefore also shift the class' mean  $GHI$  upwards. The results of these PDFs coincide with what we found in the previous section. The relative contribution of  $DHI$  is relatively high for relative high

probabilities of low clear-sky index values. The relative contribution of  $\cos(\theta) \cdot BNI$  is relatively high for relative high probabilities of high clear-sky index values.

The mixed-sky classes have a bimodal shape. The two modes correspond to respectively clouded and clear-sky conditions (e.g., Suehrcke and McCormick, 1988). The difference in clear-sky index between the two modes per class increases from low-variability (I) to high-variability (III) with an increase in difference from 0.33 to 0.64. This indicates that the BIII class has the highest probability to expect the largest quantity in solar irradiance fluctuations. For this class the transitions between clouded and clear-sky conditions are the sharpest (from very low to very high clear-sky index). Remember that the B classes are defined on having an hourly mean clear-sky index of 0.5 - 0.9, whereas the PDFs of the B classes have two modes that lie respectively lower and higher than those values. Thus taking an hourly mean smooths out the variability and therefore high frequency data is required for analysing sub-hour variability (Skartveit and Olseth, 1992).

The PDF of the total clear-sky index time series has a bimodal shape just like the B classes (Figure 4.5). However, the PDF shape is different. The first mode is quite broad and lies more to the left than for the first mode of the B classes. Probably because the A classes shift that mode more to overcast values. This also shows that overcast and clouded conditions can vary per minute and per hour on their cloud thickness and composition. The second mode lies just under a clear-sky index of 1.0 (0.97) and the probability is almost twice as high as the first mode. However, the peak is narrower than the first peak. The PDFs of all classes show how this PDF is divided, as combining the PDFs of the classes would result in this PDF. Literature has investigated, among others, the effects of optical air mass (e.g., Tovar et al., 1998), daily mean cloud conditions (e.g., Soubdhan et al., 2009) and temporal averaging (e.g., Lohmann and Monahan, 2018) on the bimodal shape of the total clear-sky index time series. However, PDFs for similar defined cloud condition classes as our study are not to our knowledge presented before. Although literature quantifies variability in solar irradiance differently, their results have a general agreement with our results (e.g., Blaga and Paulescu, 2018).

For the A, B and C classes, an increase in hourly nominal variability (classes I to III) causes a broadening and lowering of the PDF. Furthermore, also the probability density of cloud enhancement events increases. The classes with the highest probability of cloud enhancements events are the CIII (probability of 0.557), the BIII (0.285) and then the CII (0.321) class (Table 4.1). The mean time periods of these events are between the 3 and 5 minutes for classes that have a probability of enhancement events above 0.1. The mean clear-sky index when it is higher than 1. is for all data 1.095. More results are found in the table. Schade et al. (2007) and Inman et al. (2016)

Table 4.1: Statistics on cloud enhancement events ( $k^* > 1$ .) per class of the classification scheme. The probability of the cloud enhancement events are presented, as well and the mean time period of such events and what their mean clear-sky index value is.

Class	Probability $k^* > 1$ .	Mean time period $k^* > 1$ . (minutes)	Mean $k^*$ in time periods $k^* > 1$ .
AI	0.000	2.705	1.111
BI	0.020	4.031	1.064
CI	0.072	4.287	1.020
AII	0.026	2.549	1.140
BII	0.133	3.790	1.109
CII	0.321	4.543	1.059
AIII	0.080	2.224	1.149
BIII	0.285	3.321	1.117
CIII	0.557	4.685	1.100
All data	0.105	3.892	1.095

investigated enhanced solar irradiance and found that largest enhancements are caused by nearly overcast altocumulus clouds and by fractional cumulus clouds. These cloud types would probably fit in respectively the CII and CIII, and the BIII classes. It is interesting that some classes have relatively high probabilities of cloud enhancement events, as many solar irradiance forecasting models do not take enhancement values into account or do not have a high enough temporal resolution to capture this enhancement (e.g., Inman et al., 2016).

#### 4.5 Temporal autocorrelation of solar irradiance per cloud condition

For solar design methods it is important to determine whether the PDFs of the clear-sky index, especially the bimodal PDFs, are maintained for longer time intervals (Suehrcke and McCormick, 1988). Longer time intervals are better to forecast, and high frequency fluctuations have different complications than low frequency fluctuations. The temporal autocorrelation sub-figures show per class the correlation between the observed *GHI* at the initial moment *lag 0* and *GHI* in the future *lag x* (Figure 4.6). Within the A, B and C classes, the temporal autocorrelation at each lag is always lower going from variability class I to III. The variability classes are defined on the standard deviation of the 1-minute ramp rate in the solar irradiance time series. This 1-minute ramp rate thus increases in quantity going from classes I to III, which is in line with previous results. The results from the temporal autocorrelations show that the decorrelation over time coincides with increasing ramp rate.



The III classes therefore have the most random nature and the ramp rates are the highest in quantity.

From the PDFs per class of the clear-sky index time series, we would expect to find that the temporal autocorrelations of the B classes would cross the confidence interval the quickest per variability class compared to the A and C classes. The PDFs of the B classes have a bimodal shape, indicating that the quantity of the clear-sky index fluctuates between values similar to clouded conditions and values similar to clear-sky conditions. However, this expected higher decorrelation time is not (clearly) visible in the temporal autocorrelations compared to those of the other classes. The AII and AIII classes are almost identical to respectively the BII and BIII classes. Furthermore, the CIII class has a lower temporal autocorrelation per lag than the B classes. The temporal autocorrelation of the CII class first decreases faster than the BII and AII classes. However, it levels off and crosses the confidence interval around lag 8, which also applies for the BII and AII classes.

For the I classes the order is different, the CI has the highest temporal autocorrelation per lag, than the BI and the AI has the lowest temporal autocorrelation. A high temporal variability in time, however, does not also mean high variability in quantity. A temporal autocorrelation is defined on the autocorrelation normalized to the covariance. This makes that variability of low absolute values could still have a high decorrelation if the relative variability is large and random. This could explain the results on the temporal autocorrelations per class.

Two factors need to be taken into account when interpreting the results. Firstly, the PDFs were made on the clear-sky index time series and the temporal autocorrelations on the *GHI* time series. The temporal autocorrelations therefore also include the effect of the solar cycle in the sky. This affect may not be so large as the temporal autocorrelations were made on hourly subsets of the classes' *GHI* time series. We found in previous sections that the low-variability classes probably have relative more early/late daytime time stamps. In those time, the *SZA* changes fastest over time. This could also explain why these classes have a relatively higher decorrelation time than the other classes. Secondly, Skartveit and Olseth (1992) stated that temporal autocorrelations would only be reliable when data records are sufficiently long, but do not state how long this period should be. It is possible that our temporal autocorrelation taken for subsets of one hour are not sufficiently long for reliable results.

## 4.6 Ramp rates in clear-sky index per cloud condition

The previous section showed the decorrelation of solar irradiance over time. However, it did not provide information about the absolute values of the fluctuations. The PDFs of the ramp rates in

solar irradiance per class, with different time intervals, provide information with how much  $\text{W m}^{-2}$  solar irradiance changes over time (Figure 4.7). Per class, a higher probability for large ramp rates is found for increasing time interval, which corresponds to the results on the temporal autocorrelation. This is best visible in the sub-figures by the decrease in the peaks of the 4 PDFs of the solar irradiance (time intervals of 1, 5, 10 and 30 minutes) around a ramp rate of  $0 \text{ W m}^{-2}$ . This illustrates that cloud conditions thus have the highest change to be different for increasing time interval after the initial moment of observation.

The ramp rates also increase from the low- to the medium and then to the high-variability classes, whereby the ramp rates increase around the order of respectively  $\pm 100 \text{ W m}^{-2}$ ,  $\pm 200 \text{ W m}^{-2}$  and  $\pm 300 \text{ W m}^{-2}$ . Furthermore, the ramp rates increase as well within each variability class from the overcast to the clear and then to the mixed-sky class. The differences are, however, not always very distinct. The low-variability clear-sky class has apparent shapes of its PDFs in the ramp rates which are caused by the solar cycle in the sky (Lohmann et al., 2016). The low-variability overcast class has a random nature in the time series after around 11 minutes. Its belonging ramp rates in solar irradiance are constrained to about  $\pm 100 \text{ W m}^{-2}$ . The high-variability mixed-sky class has a random nature in the time series after 4 minutes and has the largest ramp rates (around  $\pm 800 \text{ W m}^{-2}$  for time intervals of higher than 5 minutes). Lohmann et al. (2016) quantified ramp rates in clear-sky index for different cloud conditions and found similar results compared to ours. These results suggest that solar irradiance forecast models will have the highest difficulty, and therefore forecast uncertainty, for forecasting solar irradiance variability under high-variability cloud conditions and high-variability mixed-sky conditions in particular.

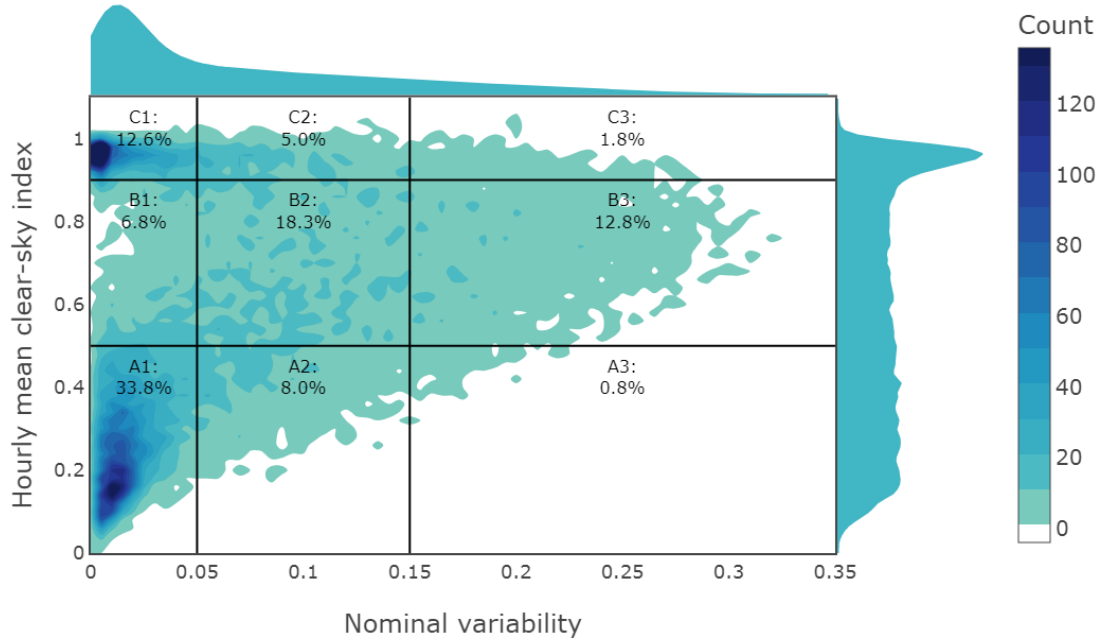


Figure 4.2: The cloud condition classification scheme. The y-axis shows the hourly mean of the centred moving averaged clear-sky index (Equation 3.4) and is divided into three classes: overcast (A), mixed-sky (B) and clear-sky (C), with limits of respectively 0.0 - 0.5, 0.5 - 0.9 and  $> 0.9$ . The x-axis shows the hourly nominal variability (Equation 3.5) and is divided into three classes: low-variability (I), medium-variability (II) and high-variability (III), with limits of respectively 0.0 - 0.05, 0.05 - 0.15 and  $> 0.15$ . The shades of blue indicate the occurrences on hourly basis as an joint probability density, 2D kernel density estimate, with a bandwidth of 0.01. The sum of all hourly data points in the period June 2011 - December 2017 is 21,418. Additionally, placed at the top and right-side of figure are the probability density functions of respectively the hourly mean clear-sky index and the nominal variability.

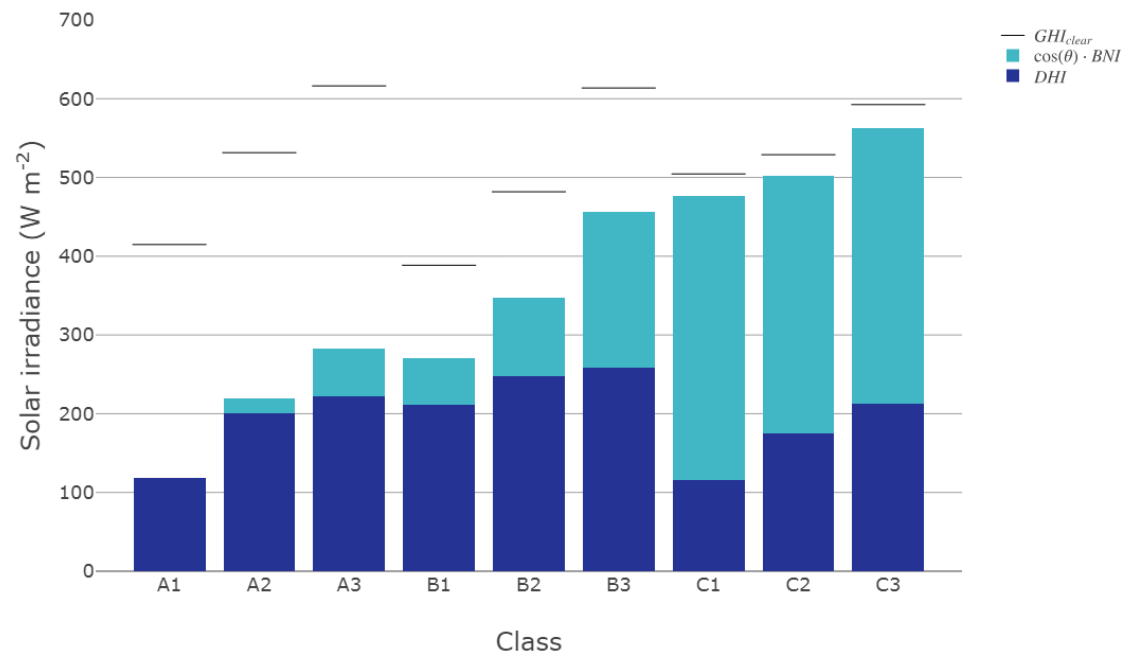


Figure 4.3: The means of the  $DHI$  and  $\cos(\theta) \cdot BNI$  per class of the classification scheme. The sum of the means per class is the class' mean of the  $GHI$ . The black horizontal lines are the class' mean  $GHI_{clear}$ .

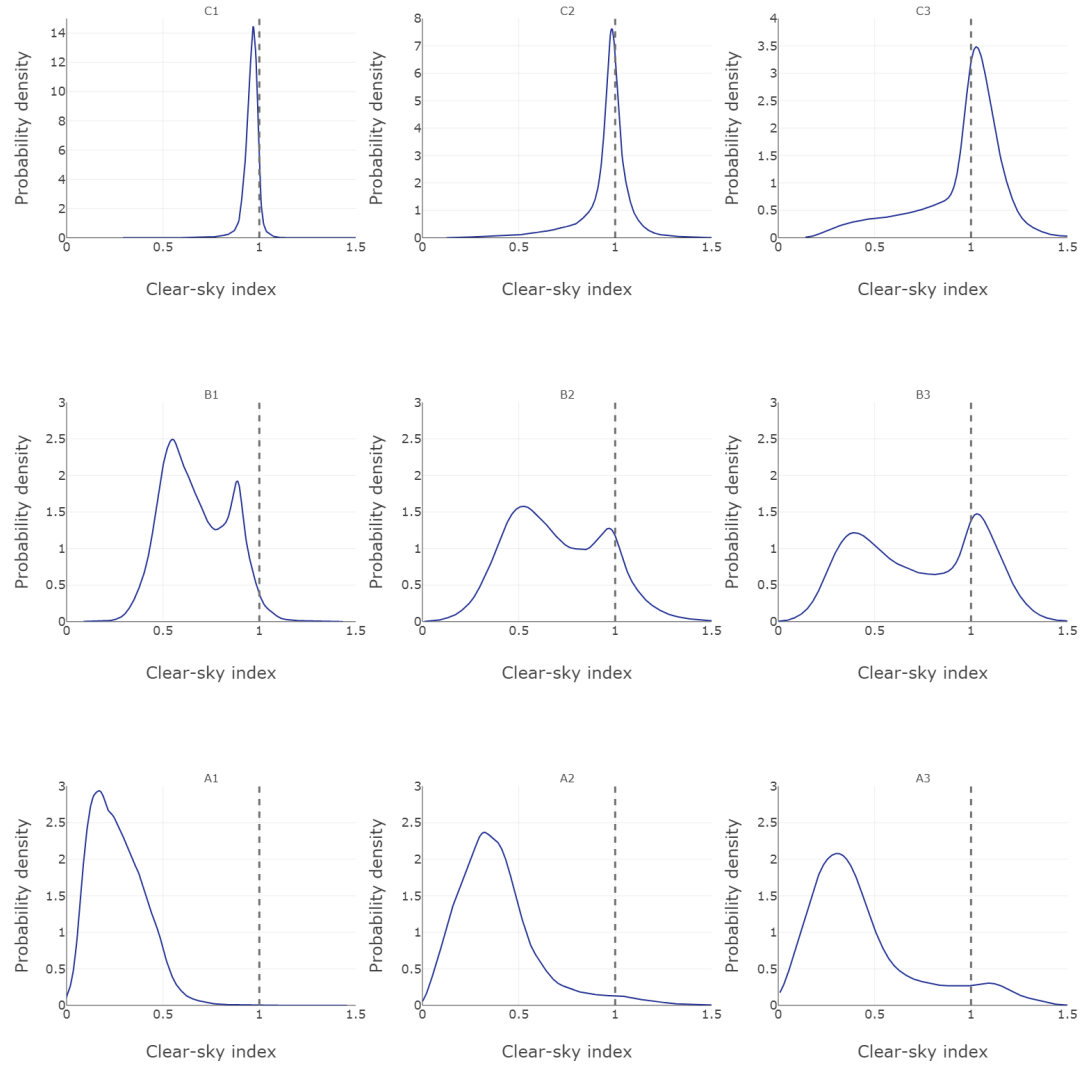


Figure 4.4: PDFs of the clear-sky index time series with an 1-minute temporal resolution per class of the classification scheme. The PDFs sub-figures are ordered in the same layout as in the classification scheme (Figure 4.2), with an increase in the hourly mean clear-sky index from classes A to C and an increase in the nominal variability from classes I to III. The dashed lines at a clear-sky index of 1.0 are to visualize the PDFs of cloud enhancement events per class.

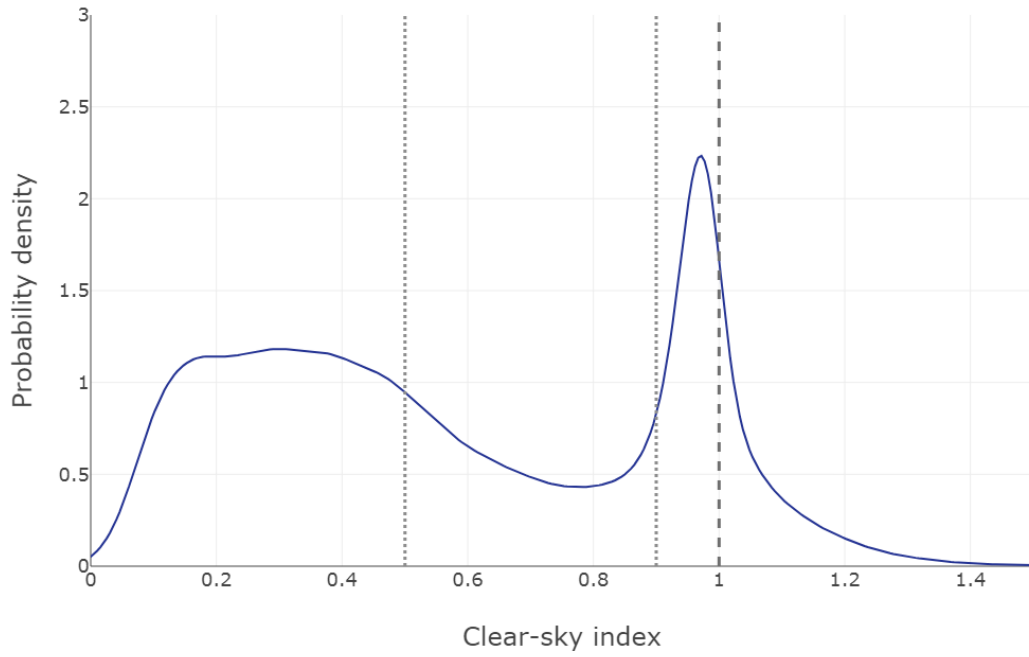


Figure 4.5: The PDF of the total clear-sky index time series with an 1-minute temporal resolution. The sum of all data points in the period June 2011 - December 2017 is 1,365,829. The dashed line at a clear-sky index of 1.0 is to visualize the PDFs of cloud enhancement events. The dotted lines at a clear-sky index of 0.5 and 0.9 are to visualize the limits used for the hourly mean clear-sky index in the cloud condition classification scheme.

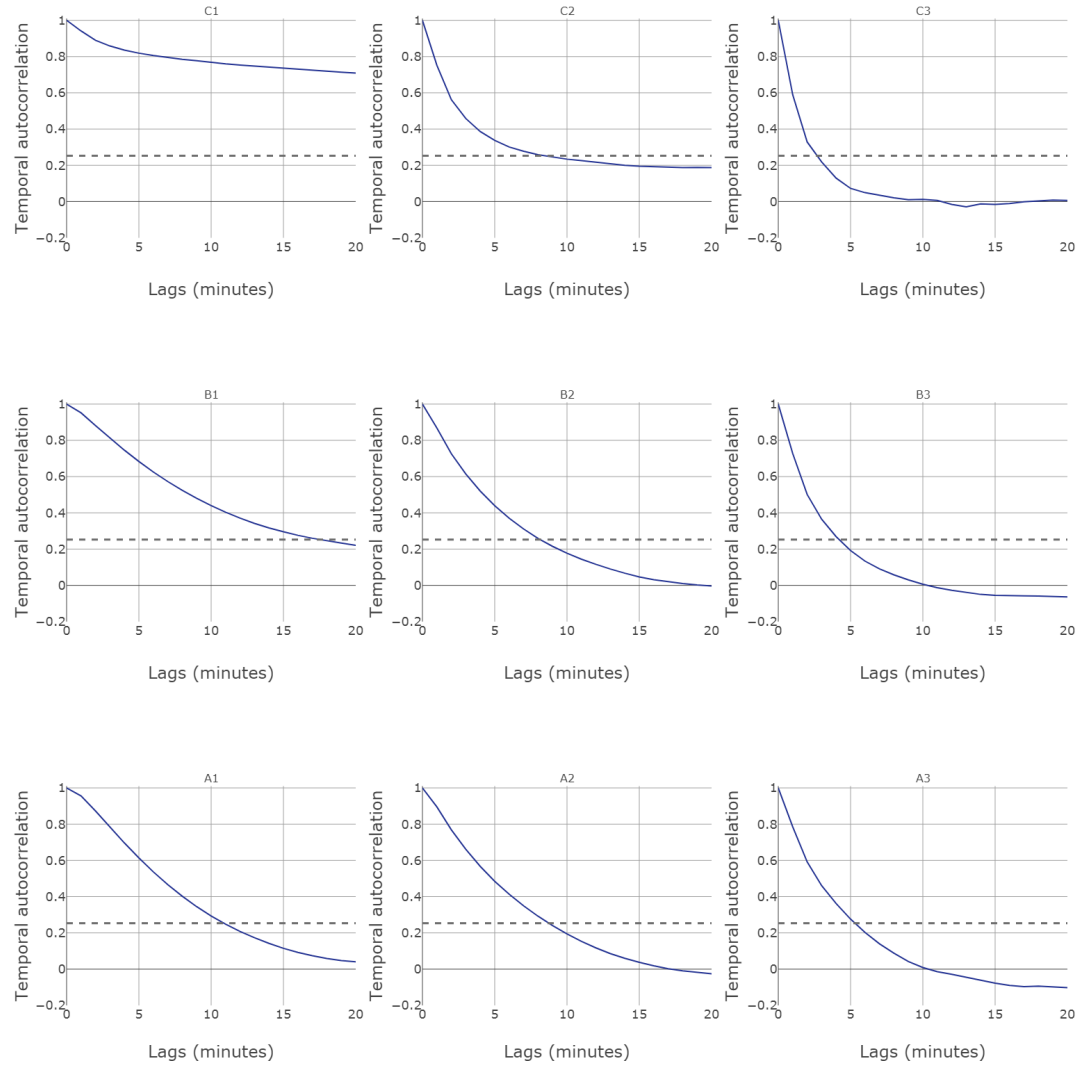


Figure 4.6: Temporal autocorrelations of the *GHI* time series with an 1-minute temporal resolution per class of the classification scheme. The dotted line is the top border of the confidence interval of  $0 \pm 0.253$ . The temporal autocorrelations sub-figures are ordered in the same layout as in the classification scheme (Figure 4.2), with an increase in the hourly mean clear-sky index from classes A to C and an increase in the nominal variability from classes I to III.

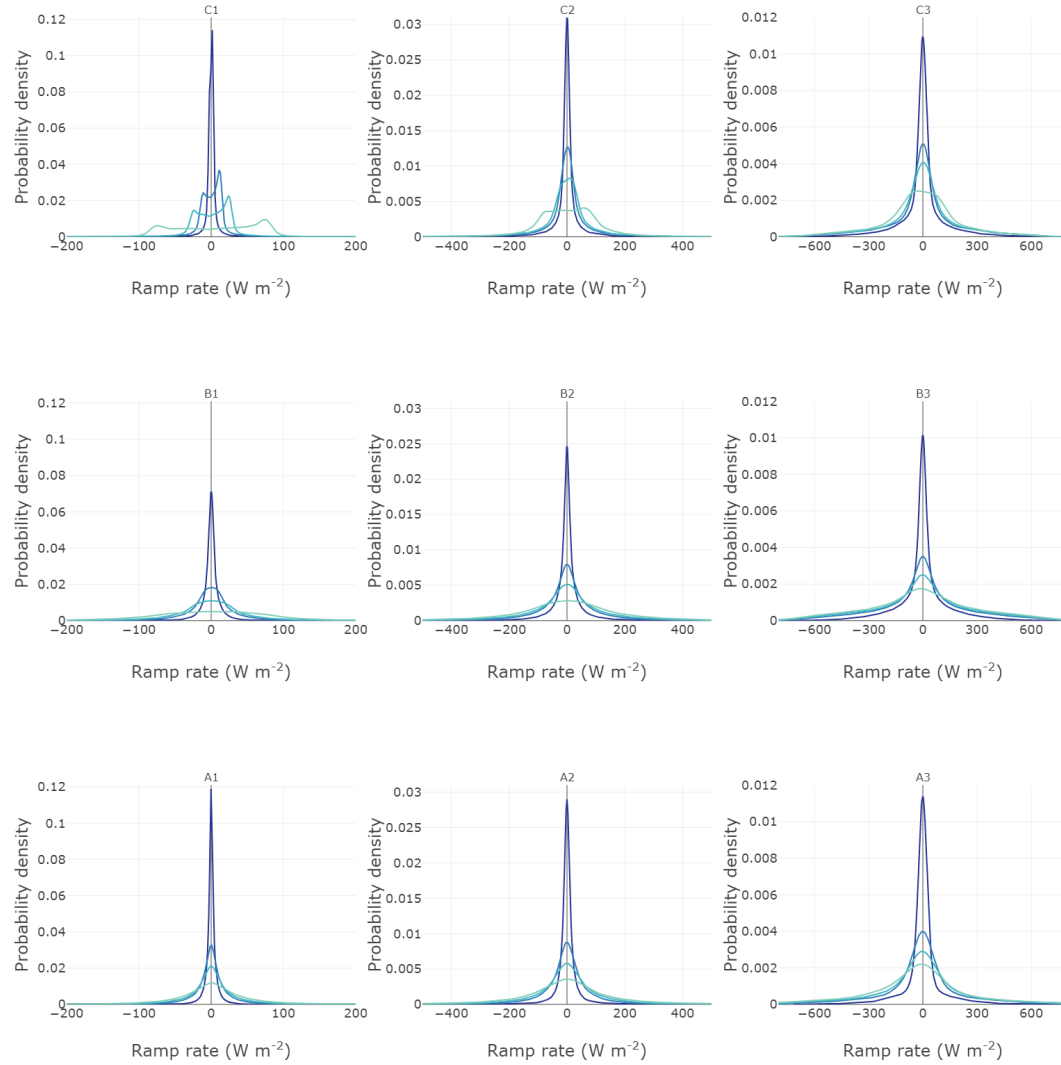


Figure 4.7: PDFs of the ramp rates in the *GHI* time series with an 1-minute temporal resolution per class of the classification scheme. The four coloured lines per sub-figure represent from dark blue to light blue respectively the 1, 5, 10 and 30 minute ramp rates. The PDFs of the ramp rates sub-figures are ordered in the same layout as in the classification scheme (Figure 4.2), with an increase in the hourly mean clear-sky index from classes A to C and an increase in the nominal variability from classes I to III.



## 5 | Discussion

The aim of this study is to quantify cloud-caused variability in surface solar irradiance. We extend on previous literature to make a classification scheme with nine classes, with each class representing a certain cloud condition (e.g. Stein et al., 2012). The classes are defined on the hourly mean clear-sky index and the hourly nominal variability, and are representative for certain cloud conditions. The objectives of this study are to quantify the solar irradiance's quantity and temporal variability per class of the classification scheme. We reach these objectives by means of calculating the monthly means of the solar irradiance components, PDFs on the clear-sky index time series, temporal autocorrelations on the *GHI* time series, and PDFs on the ramp rates in the *GHI* time series.

The results of the classification scheme show the usefulness of classifying a solar irradiance data set into different cloud condition classes and using our used methods to quantify cloud-caused variability in solar irradiance. A classification scheme like this can also be used to compare solar irradiance characteristics as measured at different weather stations, both nationally and internationally (e.g., Kang and Tam, 2013). Furthermore, a similar classification scheme is suitable for, e.g., assessing the solar irradiance forecast accuracy as function of the solar irradiance variability and can therefore be used to improve forecast models (e.g., Remund et al., 2015). Finally, a classification scheme on solar irradiance variability with a (sub)second temporal resolution can be used to downscale solar irradiance forecasts with a lower temporal resolution (e.g., Mathiesen and Kleissl, 2011). This avoids the need for forecast models to have a temporal resolution of 1 minute or less, which is computationally costly and may not be feasible in forecast time.

Throughout the study we have stated that the classes in our classification scheme represent certain cloud conditions, but the classification scheme is based on statistics of the clear-sky index time series. Therefore, physical properties of clouds per class cannot directly be provided. However, literature have made similar classification schemes and found with, e.g., satellite and sky imagery which cloud conditions could be representative for which class (e.g., Elsinga and van Sark, 2017).

As we are interested to quantify the cloud-caused variability, the clear-sky index is used for

isolating the effects of clouds on variability in solar irradiance. Although the clear-sky index is widely used in literature to investigate the cloud-caused variability and eliminate the effects of the sun's position in the sky on solar irradiance, the effects of the sun's position are not completely eliminated (e.g., Kleissl, 2013). A full explanation has been provided in Section 4.3. Although the effects of the sun's position on solar irradiance cannot completely be eliminated, the clear-sky index remains useful. To really grasp how clouds and solar irradiance interact, further research is required to investigate this on cloud level.

Some issues are left unresolved by our results, e.g. the quantification of the: spatial variability in solar irradiance, (sub)second variability in solar irradiance, and variability in global normal irradiance (*GNI*). The next paragraphs will elaborate on these unresolved issues.

To quantify spatial variability in solar irradiance is of course also crucial for getting a better understanding of cloud-caused variability in solar irradiance. Furthermore, solar forecast models have to include both the temporal and spatial variability in solar irradiance and solar power production. Different solar forecast techniques exist which depend on the purpose of the forecast and therefore the spatiotemporal scale of the technique. Due to spatial smoothing, the variability in solar power load on a regional electricity grid will be damped compared to the solar power load in a electricity grid of a street (e.g. Lorenz et al., 2009). Therefore, the operation of a regional electricity grid also does not need the same forecast accuracy as the operation on street level (e.g. Lorenz et al., 2009). More literature on solar energy forecasting can be found in the literature review papers by Antonanzas et al. (2016) and Blaga et al. (2019).

A data set with a (sub)second temporal resolution would be even better for quantifying the stochastic variability in solar irradiance (e.g., Lohmann and Monahan, 2018). Solar irradiance can be highly variable within a minute and the duration of cloud enhancement events is in the order of seconds to a few minutes (e.g., Schade et al., 2007). Yordanov et al. (2013) therefore argue that the optimal temporal resolution is around 10 Hz to not miss the extreme short and high magnitude fluctuations. As the variability can be better captured with a (sub)second time series, some results per class of classification scheme would likely change. The fluctuations would not be averaged over 1 minute anymore, so the extreme values in the clear-sky index are measured as well. Therefore the PDFs of the clear-sky index time series will likely broaden. Especially the peaks in the clear-sky index and cloud enhancements will probably increase in probability density when going from the classes I to III. The temporal autocorrelations will likely change as well. The temporal autocorrelations are probably quicker going to zero, as the high temporal variability is better measured so the time series get a more stochastic nature. The PDFs of the ramp rates in solar irradiance time series will, similar

to the PDFs on the clear-sky index, get a higher probability for higher ramp rates. An 1-second solar irradiance data set is available from the weather station at Cabauw, the Netherlands, which is owned by the Royal Netherlands Meteorological Institute (KNMI). Future research is planned to quantify the cloud-caused variability in this 1-second solar irradiance data set.

Variability in solar irradiance and the clear-sky index is reported to be higher on titled planes than horizontal planes (e.g., Schroedter-Homscheidt et al., 2018). Solar panels could thus produce solar power with a larger variability than the variability what we found in this study. A titled plane excludes a part of the sky from which it cannot receive *DHI* wherefore variability in the *BNI* becomes relatively more important. Therefore, it is useful to also quantify the cloud-caused variability in *GNI* and *BNI* time series. Especially for applications such as concentrating solar or thermal power plants (e.g., Olseth and Skartveit, 1987) and PV power stations and neighbourhood PV systems which have solar panels with a static inclination and orientation (e.g., Lohmann et al., 2016).



## 6 | Conclusions

Due to the global energy transition, quantifying cloud-caused variability in solar irradiance for different cloud conditions is important for the successful integration of weather-dependent decentralized PV power systems. We quantified cloud-caused variability in the 1-minute averaged solar irradiance measurements from the Veenkampen weather station near Wageningen, the Netherlands, for the period June 2011 to December 2017. We extended on previous studies (e.g. Stein et al., 2012) and created a classification scheme based on clear-sky index statistics. The nine classes are identified as overcast, mixed-sky and clear-sky conditions and are subdivided on low-, medium- and high-variability conditions. For a complete quantification of cloud-caused variability, we quantified per cloud condition first the variability in the solar irradiance quantity and second the temporal variability.

The variability in solar irradiance quantity is characterized for mixed-sky conditions as having a large spread and intermittent values corresponding to clouded and clear-sky conditions. Cloud enhancement events result in solar irradiance quantities higher than clear-sky conditions and for high-variability mixed-sky conditions they have a probability density of 0.29 to occur and can reach clear-sky index values of 1.5. Of the the nine classes, the mixed-sky conditions, and its high-variability in particular, thus likely cause the largest spread in solar power production.

The temporal variability in solar irradiance has the shortest time scale of fluctuations and largest ramp rates for high-variability conditions. The time series are decorrelated after less than 5 minutes and has ramp rates of  $\pm 800 \text{ W m}^{-2}$ . The mixed-sky high-variability cloud condition has the highest probability to have such high ramp rates compared to the two other high-variability classes. These results suggest that solar irradiance forecast models will have the highest difficulty, and therefore forecast uncertainty, for forecasting solar irradiance variability under high-variability cloud conditions.

Although it was not the aim of this study, we also have quantified the affect of the solar cycle in the sky on the solar irradiance variability. Low-variability conditions are more likely to occur when the sun's elevation is low in the sky, thus during winter months or in early/late hours of the day. High-variability conditions are more likely to occur when the sun's elevation is high in the sky, thus

during summer months or in hours around noon. So, the sun's position in the sky not only affect the variability in the quantity of both solar irradiance and the clear-sky index, but likely also the temporal variability in the two variables. Furthermore, this study shows that the clear-sky index is not completely able to eliminate the effect of the sun's elevation on cloud-caused variability in solar irradiance.

The complete quantification of cloud-caused variability in solar irradiance concludes that the high-variability mixed-sky cloud condition is potentially the most problematic for grid operators to balance electricity load on the local grid in terms of short-term fluctuations. The classification scheme and belonging methods presented in this study can be used by future studies to reduce the uncertainty in solar irradiance forecasts, which is a stepping stone to enable large-scale installation of PV power systems in current and future electricity grids.

## 7 | Acknowledgements

I want to acknowledge everyone who helped and supported me during my MSc thesis research. In particular I want to thank my supervisor Chiel van Heerwaarden for all his support and to have helped me to get enthusiastic to continue in the topic of solar energy meteorology as a PhD candidate. The group meetings on Thursdays were also useful and a good moment to hang out with new people. Furthermore, I want to thank Frank Kreuwel for inviting us at Alliander and for our discussions on how to shape our research on solar energy meteorology. I also want to thank Kees van den Dries for showing me around at the Veenkampen weather station and to make sure that the solar irradiance data set is of the highest quality. In addition, I want to thank Caroline Wiltink-de Bruijn and Peter Kalverla for our small project to make posters of all MSc thesis students and members of the Meteorology and Air Quality group. I hope the posters will be updated also after I have left the university. Furthermore, I want to thank all students participating in the Meteorology and Air Quality group's thesis rings. Your feedback on my texts and our discussions have always helped me to improve my scientific writing skills and I also hope to have helped others with their thesis. Finally, I want to thank my friends and everyone who wrote their thesis in the same period as me in the thesis room. I enjoyed our extensive lunches, drinking loads of coffee with you, our walks in the breaks and to hang out with you. We definitely had fun during our 'thesis life'.





## Bibliography

- Aguiar, L. M., Pereira, B., David, M., Diaz, F., and Lauret, P. (2015). Use of satellite data to improve solar radiation forecasting with Bayesian Artificial Neural Networks. *Solar Energy*, 122:1309–1324.
- Antonanzas, J., Osorio, N., Escobar, R., Urraca, R., Martinez-de Pison, F., and Antonanzas-Torres, F. (2016). Review of photovoltaic power forecasting. *Solar Energy*, 136:78–111.
- Barnett, T., Ritchie, J., Foat, J., and Stokes, G. (1998). On the space–time scales of the surface solar radiation field. *Journal of Climate*, 11(1):88–96.
- Barth, B., Concas, G., Zane, E. B., Franz, O., Frias, P., Hermes, R., Lama, R., Loew, H., Mateo, C., Reking, M., et al. (2014). PV Grid-Final Project Report. *Berlin: PV GRID Project*.
- Blaga, R. and Paulescu, M. (2018). Quantifiers for the solar irradiance variability: A new perspective. *Solar Energy*, 174:606–616.
- Blaga, R., Sabadus, A., Stefu, N., Dughir, C., Paulescu, M., and Badescu, V. (2019). A current perspective on the accuracy of incoming solar energy forecasting. *Progress in Energy and Combustion Science*, 70:119–144.
- Copernicus Atmosphere Monitoring Service (2016). *The Copernicus Atmosphere Monitoring Service (CAMS) Radiation Service in a nutshell*. Issued by: M. Schroedter-Homscheidt, DLR.
- Elsinga, B. and van Sark, W. G. (2017). Short-term peer-to-peer solar forecasting in a network of photovoltaic systems. *Applied Energy*, 206:1464–1483.
- European Photovoltaic Industry Association and others (2012). Connecting the Sun: Solar photovoltaics on the road to large-scale grid integration. *Brussels, Belgium*.
- Friedrich, R., Peinke, J., Sahimi, M., and Tabar, M. R. R. (2011). Approaching complexity by stochastic methods: From biological systems to turbulence. *Physics Reports*, 506(5):87–162.

- Hoff, T. E. and Perez, R. (2010). Quantifying PV power output variability. *Solar Energy*, 84(10):1782–1793.
- Hukseflux Thermal Sensors (2018). Hukseflux Thermal Sensors DR01 Pyrheliometer. <https://www.hukseflux.com/products/solar-radiation-sensors/pyrheliometers/dr01-pyrheliometer>. Accessed: 2018-11-15.
- Hukseflux Thermal Sensors B.V. (2015). *User manual SR11 v1608*.
- Hukseflux Thermal Sensors B.V. (2016). *User manual DR-Series Pyrheliometers, DR01 DR02 DR03 manual v1609*.
- Inman, R. H., Chu, Y., and Coimbra, C. F. (2016). Cloud enhancement of global horizontal irradiance in California and Hawaii. *Solar Energy*, 130:128–138.
- IRENA (2018). Power System Flexibility for the Energy Transition, Part 1: Overview for policy makers. Technical report, International Renewable Energy Agency, Abu Dhabi.
- Kang, B. O. and Tam, K.-S. (2013). A new characterization and classification method for daily sky conditions based on ground-based solar irradiance measurement data. *Solar Energy*, 94:102–118.
- Kipp & Zonen (2018a). Kipp & Zonen CMP11 Pyranometer. [https://www.kippzonen.com/Product/13/CMP11-Pyranometer#.W\\_QIizhKi71](https://www.kippzonen.com/Product/13/CMP11-Pyranometer#.W_QIizhKi71). Accessed: 2018-11-15.
- Kipp & Zonen (2018b). *Solar radiation measurements for solar energy applications: Precision monitoring of solar radiation for photovoltaic and thermal concentrating solar energy systems*.
- Kleissl, J. (2013). *Solar energy forecasting and resource assessment*. Academic Press.
- Lave, M. and Kleissl, J. (2010). Solar variability of four sites across the state of Colorado. *Renewable Energy*, 35(12):2867–2873.
- Lave, M., Kleissl, J., and Arias-Castro, E. (2012). High-frequency irradiance fluctuations and geographic smoothing. *Solar Energy*, 86(8):2190–2199.
- Lefèvre, M., Oumbe, A., Blanc, P., Espinar, B., Gschwind, B., Qu, Z., Wald, L., Homscheidt, M. S., Hoyer-Klick, C., Arola, A., et al. (2013). McClear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions. *Atmospheric Measurement Techniques*, 6:2403–2418.

- Liu, B. Y. and Jordan, R. C. (1960). The interrelationship and characteristic distribution of direct, diffuse and total solar radiation. *Solar energy*, 4(3):1–19.
- Lohmann, G. (2018). Irradiance variability quantification and small-scale averaging in space and time: A short review. *Atmosphere*, 9(7):264.
- Lohmann, G. M. and Monahan, A. H. (2018). Effects of temporal averaging on short-term irradiance variability under mixed sky conditions. *Atmospheric Measurement Techniques*, 11(5):3131–3144.
- Lohmann, G. M., Monahan, A. H., and Heinemann, D. (2016). Local short-term variability in solar irradiance. *Atmospheric Chemistry and Physics*, 16(10):6365–6379.
- Lorenz, E., Hurka, J., Heinemann, D., and Beyer, H. G. (2009). Irradiance forecasting for the power prediction of grid-connected photovoltaic systems. *IEEE Journal of selected topics in applied earth observations and remote sensing*, 2(1):2–10.
- Mathiesen, P. and Kleissl, J. (2011). Evaluation of numerical weather prediction for intra-day solar forecasting in the continental united states. *Solar Energy*, 85(5):967–977.
- Olseth, J. A. and Skartveit, A. (1984). A probability density function for daily insolation within the temperate storm belts. *Solar Energy*, 33(6):533–542.
- Olseth, J. A. and Skartveit, A. (1987). A probability density model for hourly total and beam irradiance on arbitrarily orientated planes. *Solar Energy*, 39(4):343–351.
- Pandas-Dev (2019). Github pandas-dev pandas, autocorr function. <https://github.com/pandas-dev/pandas/blob/v0.24.1/pandas/core/series.py#L2196-L2239>.
- Perez, R., David, M., Hoff, T. E., Jamaly, M., Kivalov, S., Kleissl, J., Lauret, P., Perez, M., et al. (2016). Spatial and temporal variability of solar energy. *Foundations and Trends® in Renewable Energy*, 1(1):1–44.
- Remund, J., Calhau, C., Perret, L., and Marcel, D. (2015). Characterization of the spatio-temporal variations and ramp rates of solar radiation and pv. *Report IEA-PVPS T14-05*.
- Schade, N. H., Macke, A., Sandmann, H., and Stick, C. (2007). Enhanced solar global irradiance during cloudy sky conditions. *Meteorologische Zeitschrift*, 16(3):295–303.
- Schepanski, K., Klüser, L., Heinold, B., and Tegen, I. (2015). Spatial and temporal correlation length as a measure for the stationarity of atmospheric dust aerosol distribution. *Atmospheric Environment*, 122:10–21.

- Schroedter-Homscheidt, M., Kosmale, M., Jung, S., and Kleissl, J. (2018). Classifying ground-measured 1 minute temporal variability within hourly intervals for direct normal irradiances. *Meteorologische Zeitschrift*.
- Skartveit, A. and Olseth, J. (1992). The probability density and autocorrelation of short-term global and beam irradiance. *Solar Energy*, 49(6):477–487.
- Soubdhan, T., Emilion, R., and Calif, R. (2009). Classification of daily solar radiation distributions using a mixture of Dirichlet distributions. *Solar energy*, 83(7):1056–1063.
- Stein, J. S., Hansen, C. W., and Reno, M. J. (2012). The variability index: A new and novel metric for quantifying irradiance and PV output variability. In *World Renewable Energy Forum*, pages 13–17. Denver CO.
- Stull, R. (2000). *Meteorology for scientists and engineers*. Brooks/Cole.
- Suehrcke, H. and McCormick, P. (1988). The frequency distribution of instantaneous insolation values. *Solar energy*, 40(5):413–422.
- Tomson, T. (2010). Fast dynamic processes of solar radiation. *Solar Energy*, 84(2):318–323.
- Tovar, J., Olmo, F., and Alados-Arboledas, L. (1998). One-minute global irradiance probability density distributions conditioned to the optical air mass. *Solar energy*, 62(6):387–393.
- Waskom, M. (2018). Visualizing the distribution of a dataset. <https://seaborn.pydata.org/tutorial/distributions.html>. Accessed: 2018-10-01.
- World Meteorological Organization (1982). *Annual report of the World Meteorological Organization (WMO) 1981*.
- Woyte, A., Belmans, R., and Nijs, J. (2007). Fluctuations in instantaneous clearness index: Analysis and statistics. *Solar Energy*, 81(2):195–206.
- Yordanov, G. H., Saetre, T. O., and Midtgard, O.-M. (2013). Optimal temporal resolution for detailed studies of cloud-enhanced sunlight (Overirradiance). In *Photovoltaic Specialists Conference (PVSC), 2013 IEEE 39th*, pages 0985–0988. IEEE.