Norwegian University of Life Sciences
Faculty of Enivironmental Science and Technology
Department of Mathematical Sciences and Technology

Master's Thesis 2015
30 credits

Quality Control of Global Solar Irradiation Measured at Four Stations in Eastern Norway

Kvalitetskontroll av global solinnstråling målt ved fire stasjoner på Østlandet

Sigbjørn Grini

Acknowledgements

First of all I want to thank my three advisers, Dr. Espen Olsen, Dr. Øyvind Byrkjedal and Dr. Mareile Astrid Wolff, for excellent guidance throughout the semester. Espen has contributed through weekly discussions, advice and with his insight in the solar energy industry. Øyvind, Head of Research and Developement at Kjeller Vindteknikk, initiated this thesis and has provided invaluable support and advice in every step of the way. His knowledge on meteorological data has been a substantial asset. Mareile has shared her experiences with solar irradiation measurements, answered questions and introduced me to people within the Norwegian Meteorological Institute.

I would also like to thank Halvard Hole from Bioforsk with information about the stations and general enthusiasm about my thesis, Øystein Ruud Hansen from ITAS, who provided information on the pyranometers at each station, Øystein Godøy for introducing me to libRadtran and rtmrun, Dr. Arne Auen Grimenes for helpful discussions on the quality control procedure and Dr. Heidi S. Nygård for pointers on the structure of this thesis.

All the employees at Kjeller Vindteknikk have provided me with a great work environment and encouragement. Everybody has been willing to answer my questions or give me advice. I am very grateful for my stay at KVT.

Lastly, I would like to thank my friends and family for moral support throughout this semester.

Ås, May 15th 2015

Sigbjørn Grini

Abstract

In this thesis, the quality of global solar irradiation measurements at four stations located in Eastern Norway has been examined. The stations are Ås, Lier, Ramnes and Tomb and the time series is from 1992 to 2012. The quality control procedure consisted of two parts; an automatic control and a visual control. The automatic control applied several tests, which flagged the data points that were either considered erroneous or suspicious. The visual control compared the solar irradiation between the four stations, checked whether there were any temporal changes to the measurements and determined if the suspicious data points were erroneous or not. The overall quality appeared to be good, however, one of the stations, Lier, had indications that the measurements had been excessively high in a large part of the time series. The percentage of erroneous data ranged from 5.23 % to 9.57 % and the impact of erroneous data on total measured solar irradiation was very low.

Quality controlled measurements have been compared with existing solar irradiation databases. The databases are Satel-Light, NASA SSE 6.0, the WRF model, Meteonorm 7, Classic PVGIS and CM-SAF PVGIS. Four comparison methods have been applied; comparison with daily values, and comparison with monthly, yearly and quarterly averages. For the first three methods, the databases have been compared with quality controlled solar irradiation from the four stations described above. For the last method, however, automatic quality controlled data from additional 12 stations in Eastern Norway have been included in the comparison.

The comparisons display how much each database deviates compared to quality controlled measurements with regard to the amount of daily solar irradiation and the time of year. This information may be used to achieve a more accurate estimation of solar irradiation in Eastern Norway.

Sammendrag

I denne oppgaven har kvaliteten på målt global solinnstråling ved fire stasjoner på Østlandet blitt undersøkt. Stasjonene er Ås, Lier, Ramnes og Tomb og tidsserien strekker seg fra 1992 til 2012. Kvalitetskontrollen besto av to deler; en automatisk kontroll og en visuell kontroll. Den automatiske kontrollen hadde flere tester, som flagget de datapunktene som ble vurdert feilaktig eller mistenkelig. Den visuelle kontrollen sammenliknet solinnstrålingen til de fire stasjonene med hverandre, sjekket om det var noen forandringer på målingene over tid og avgjorde om de mistenkelige datapunktene var feilaktige eller ikke. Den generelle kvaliteten virket god, men én av stasjonene, Lier, hadde indikasjoner om at målingene har vært for høye i store deler av tidsserien. Andelen feilaktige data strakk seg fra 5.23 % til 9.57 % og betydningen av feilaktige data på totalt målt verdi var veldig lav.

Kvalitetskontrollerte målinger har blitt sammenliknet med eksisterende databaser for solinnstråling. Databasene som ble brukt er Satel-Light, NASA SSE 6.0, WRF, Meteonorm 7, Classic PVGIS og CM-SAF PVGIS. Fire sammenlikningsmetoder ble anvendt; sammenlikning med daglige verdier, og med månedlige, årlige og kvartalvise gjennomsnitt. For de første tre metodene ble databasene sammenliknet med de fire stasjonene beskrevet tidligere. For den siste metoden ble også målinger ved 12 ekstra stasjoner på Østlandet lagt til i sammenlikningen. Disse målingene ble bare kvalitetskontrollert gjennom automatisk kontroll.

Sammenlikningene viser hvor mye hver database varierer i forhold til kvalitetskontrollerte målinger for daglige verdier og ved forskjellige tider på året. Disse resultatene kan bli brukt videre for å oppnå et mer nøyaktig bilde av hvor mye solinnstråling det er på Østlandet.

Contents

List of abbreviations

Acronyms / Abbreviations

AM	Air mass
EOT	Equation of time
MAE	Mean absolute error
MBD	Mean bias deviation
RMSD	Root-mean-square deviation

Chapter 1

Introduction

Investments in solar energy applications in Norway are rapidly increasing every year. Enova (public enterprise that is owned by the Ministry of Petroleum and Energy, www.enova.no) and the city of Oslo provide economic incentives to encourage investments in solar energy systems. Even though Norway is located at higher latitudes, there is potential for solar energy applications, especially on top of or integrated in buildings. Since the economic incentives still are limited, an accurate estimation of the future energy production for a solar energy system is essential. Andersen (2014) discovered that forecasted power production for a solar energy system in Norway underestimate the actual production. Hauge et al. (2014) report similar results for solar heating. Given that this is true for most solar energy systems in Norway, a higher income could be expected compared to what is currently estimated. One reason for the underestimation in power production is that solar cells produce more electricity for lower ambient temperatures compared to what was expected (Nordmann et al., 2014). Another reason could be that the solar energy resources are incorrectly estimated. The main issue with the latter is that the solar irradiation data from the existing databases vary considerably. The three master's theses from Størdal (2013), Aase (2013) and Romundstad (2014) found that the average yearly solar irradiation from the database with the highest value are respectively 17 %, 12 % and 15 % higher than the database with the lowest value for the given location.

There are in general two methods used to estimate the amount of solar irradiance at a given location and time. The first method is to use a ground measurement device which measures solar irradiance at a certain time interval. The second method is to use satellite images to estimate the ratio of clouds in the sky at a certain location and use a radiation model to estimate the amount of incoming solar irradiance. Ground measurements are considered

to be the most accurate and reliable of the two, even though there are several sources of errors (Younes et al., 2005). In order to take pictures from the same position several times a day, the satellites have to be geostationary, which means that they are restricted to orbit the earth above the equator at a fixed altitude. At higher latitudes, as in Norway, the acute angle between the surface and the satellite becomes increasingly low. The satellite estimate at higher latitudes is thus of lower resolution and higher uncertainty compared to locations closer to the equator. In addition, the satellites may interpret the snow-covered landscape as clouds (Hagen, 2011).

In Norway, three ground measurement stations are used in databases which provide data for solar energy applications. The three stations are located in Bergen on the west coast of Norway, and in Bodø and Tromsø which are both located in the northern part of Norway. There is, however, no ground measurement station in the vicinity of the largest city, Oslo. To estimate the solar irradiation for Oslo, databases apply an interpolation algorithm from the nearest measurement stations in addition to satellite images. This method is not ideal and introduces uncertainties.

Bioforsk, the Norwegian Institute of Agricultural and Environmental Research, measure hourly solar irradiation at 47 of their stations in Norway. The data are used for forecasting and information service developed for management of pests and diseases in cereals, vegetables and fruit crops (Hole, 2015). Many of the stations are located close to Oslo and have over 20 years of measured data. The solar irradiation data from Bioforsk are readily available for free online, but are not included in any databases commonly used today. The solar irradiation data from these stations could contribute to a much more accurate knowledge of solar energy in Norway if the quality of the data is sufficient. Thus, in this thesis, a quality control analysis of solar irradiation data at four Bioforsk stations, Ås, Lier, Ramnes and Tomb have been performed. Methodologies for quality controlling such data have been studied and applied. Furthermore, a comparison between existing solar irradiation databases and quality controlled data has been conducted.

Chapter 2

Theory

2.1 Solar radiation

Solar radiation is the radiation emitted by the sun and it spreads out in all directions. A tiny part of the solar radiation reaches the earth's atmosphere and this radiation is termed the extraterrestrial irradiance, I_{E}. The earth has an atmosphere which consists of several gases that absorb and reflect solar radiation. Among them are oxygen, ozone, carbon dioxide and water vapor. In addition, clouds, aerosols, fog and smoke scatter the solar beams. As a consequence, the solar irradiance that reaches the earth's surface, termed global horizontal irradiance, I_{global}, is always less than the extraterrestrial irradiance. In this thesis I_{global} is often noted as "solar irradiance" or "solar irradiation". Irradiance is the rate of solar energy per unit area, whilst irradiation is the amount of solar energy per unit area in a given time period. Other commonly used terms for I_{global} are global radiation, global irradiance and global irradiation. One part of the global horizontal irradiance passes through the atmosphere without noticeable molecular interaction and is arriving directly from the solar beam and is thus termed beam horizontal irradiance, I_{beam}. The remaining global horizontal irradiance are scattered in the atmosphere before eventually reaching the surface of the earth. This dispersed radiation is termed diffuse horizontal irradiance, I_{diffuse}. The relationship between these three is described in Coulson (1975) and shown in Eq. 2.1.1.

$$I_{\mathrm{global}} = I_{\mathrm{beam}} + I_{\mathrm{diffuse}} \left(\frac{\mathrm{W}}{\mathrm{m^2}} \right) \tag{2.1.1}$$

Solar luminosity

The sun's radiation is closely resembles black body radiation. A black body is a body which absorbs all incoming radiation and emits maximum possible radiation given its temperature. Black body radiation can be estimated using the *Stefan-Boltzmann* law shown in Eq. 2.1.2.

$$P = A\sigma T^4 \left(\frac{\mathrm{W}}{\mathrm{m^2}}\right) \tag{2.1.2}$$

Where P is the power radiated from the black body, A is the surface area of the object, T is its temperature and σ is the Stefan-Boltzmann constant. The Stefan-Boltzmann constant states that thermal radiation per area per temperature to the fourth power is a constant and $\sigma = 5.6697 \times 10^{-8} \frac{\mathrm{W}}{\mathrm{m^2T^4}}$ (Iqbal, 1983). The solar luminosity is defined as the total amount of energy emitted by the sun. Since the sun's effective temperature is 5778 K and its radius is 696000 km the solar luminosity is $L = 3.84 \times 10^{26}$ W (Williams, 2013).

Solar constant

The solar constant is often used in simple radiation models and as an upper limit of solar radiation on the earth's surface. It is the solar irradiance at the mean earth-sun distance, A_s (1 astronomical unit). The theoretical value for the solar constant is described in Iqbal (1983) and shown in Eq. 2.1.3.

$$S = \frac{L}{4\pi A_s^2} = 1367 \frac{\mathrm{W}}{\mathrm{m^2}} \tag{2.1.3}$$

Solar spectral radiation

The solar radiation is a function of wavelength and temperature, which is true for any surface with a temperature above zero. The radiative distribution of a black body was shown by Max Planck in 1900 and is shown in Fig. 2.1.1. As described above, the sun emits close to black body radiation, and the solar spectral radiation at the surface of the earth is shown in Fig. 2.1.2. This, however, is not the spectral distribution that is observed at the surface of the earth due to the ability of atmospheric molecules to absorb solar irradiance of certain wavelengths.

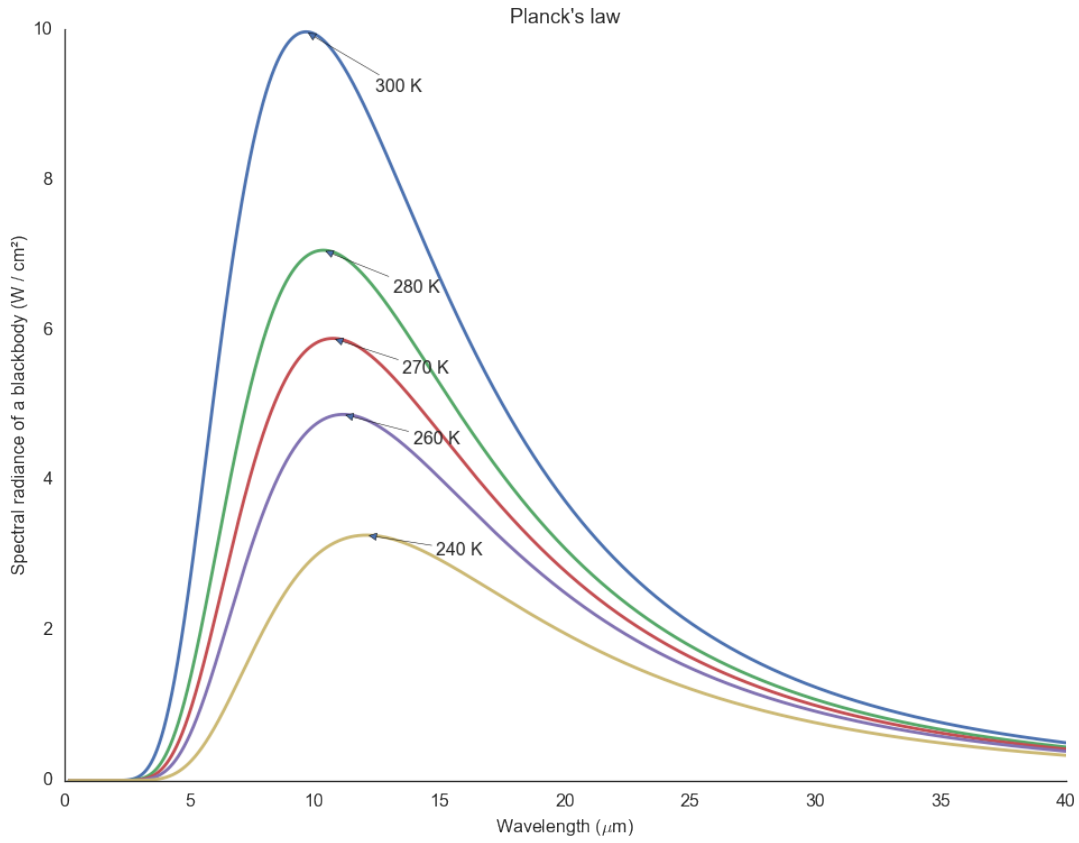


Figure 2.1.1 – Planck's law displayed for different temperatures.

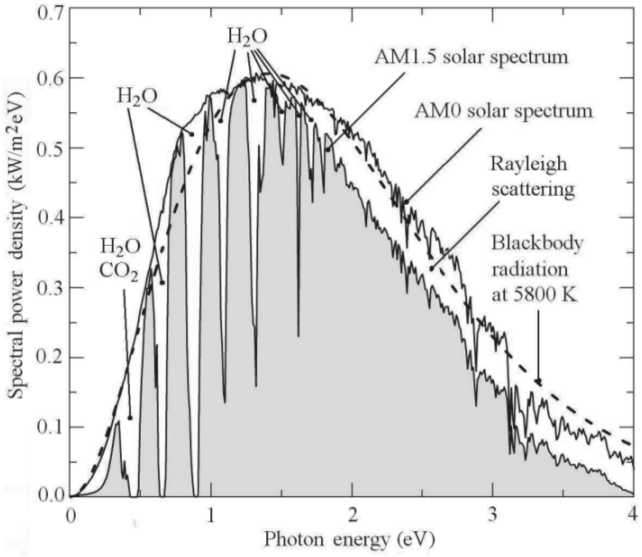


Figure 2.1.2 – The absolute solar spectral radiation at $AM = 0$ and $AM = 1.5$. The figure adopted from Chen (2011) and used with permission from John Wiley & Sons.

Air mass (AM)

As described in section above, the solar radiation interacts with the atmosphere before eventually arriving at the surface of the earth. If the sun is positioned at zenith, which is the point directly above an observer, the amount of interacting atmospheric molecules are at a minimum, since it is the shortest distance from the top of the atmosphere to the surface of the earth. Thus, when the solar zenith angle is zero and with a reasonably clear sky, the air mass (AM) has been defined equal to 1. The air mass increases with increased solar zenith angle, $AM = \frac{ds}{dz} = \frac{1}{\cos\theta} > 0$ as shown in Fig. 2.1.3. Air mass is used as a standardization parameter in solar energy applications. $AM = 1.5$ is a part of the standard test condition. In Norway, the lowest air mass is around $AM = 1.22$.

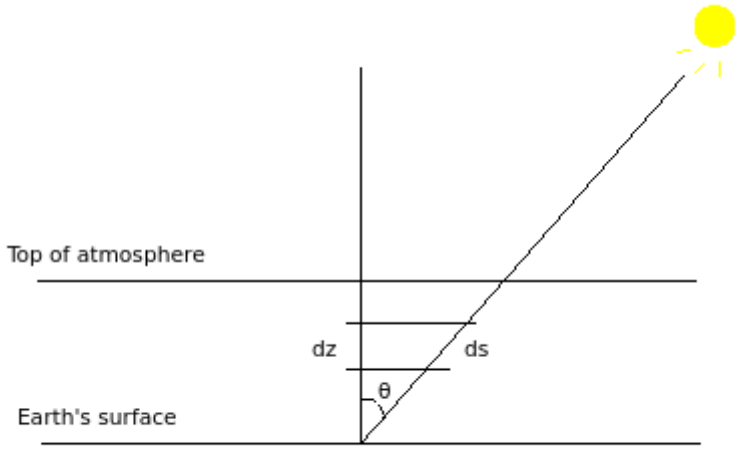


Figure 2.1.3 – Simplified figure of air mass (AM). dz is an incremental distance in the direction of zenith, ds is an incremental distance in the direction of the sun and θ is the solar zenith angle.

2.2 Solar geometry

The following theory of solar geometry have been inspired mainly by Chen (2011); Coulson (1975); Iqbal (1983); Muneer (1997) and additional sources are otherwise specified.

In order to estimate the potential amount of solar irradiance at a certain location, an understanding of the solar geometry is required. The earth orbits the sun in an elliptic path, and its plane is known as the ecliptic plane. In addition, the earth rotates around its own axis and its plane is called the equatorial plane. The angle between the planes is known as the obliquity of the ecliptic, ε, and its mean value is derived by the Astronomical Almanac (2014) to be $\varepsilon = 23.439°$. The two planes are shown in Fig. 2.2.1.

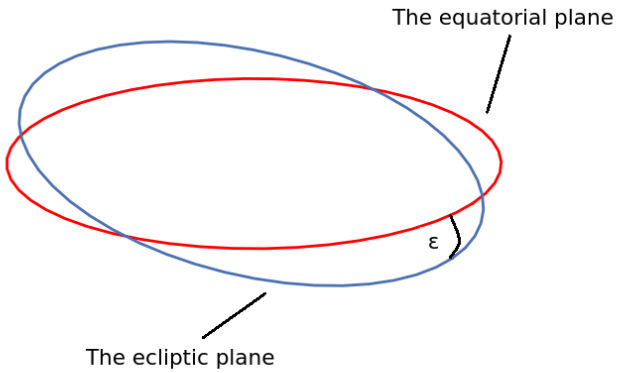


Figure 2.2.1 – The equatorial plane and the ecliptic plane. The angle between the planes is known as the obliquity of the ecliptic, ε. The equatorial plane is the plane the earth rotates in and the ecliptic plane is the plane in which the earth passes around the sun.

As a consequence there is a constant change in the sun's position for an observer on the surface of the earth. During the summer, the earth's northern hemisphere is tilted towards the sun, while during the winter, the northern hemisphere is tilted away from the sun. This results in significantly lower solar irradiance during the winter in the northern hemisphere and likewise for the southern hemisphere in the summer. This can be described by the angle between the sun - earth vector and the equatorial plane and is known as the solar declination, δ. A simple approximation for calculating the solar declination in radians is presented in Chen (2011) and shown in Eq. 2.2.1.

$$\delta \approx \varepsilon \sin(\frac{2\pi(N-80)}{365.2422} \tag{2.2.1}$$

Where N is the number of day in the year starting with January 1st and ε is the obliquity of the ecliptic. The solar declination is at a minimum at winter solstice, which is at Dec 21st/22nd and at a maximum at summer solstice, which is at Jun 20th/21th. When the solar declination is zero it is either vernal or autumnal equinox. This occurs at March 20th/21th and September 22nd/23rd. An illustration adopted from Chen (2011) is shown in Fig. 2.2.2. The models on the earth - sun geometry are using time based on the sun's position in the sky (true solar time), while the solar irradiance measurements are accumulated in local standard time (LST), which is GMT+1. Hence, a function is required to easily convert from true solar time to local standard time for a given location. An immediate problem is that the length of

a day in true solar time, defined as the time from one solar noon to another, varies with each day. Thus, the equation of time has to be calculated.

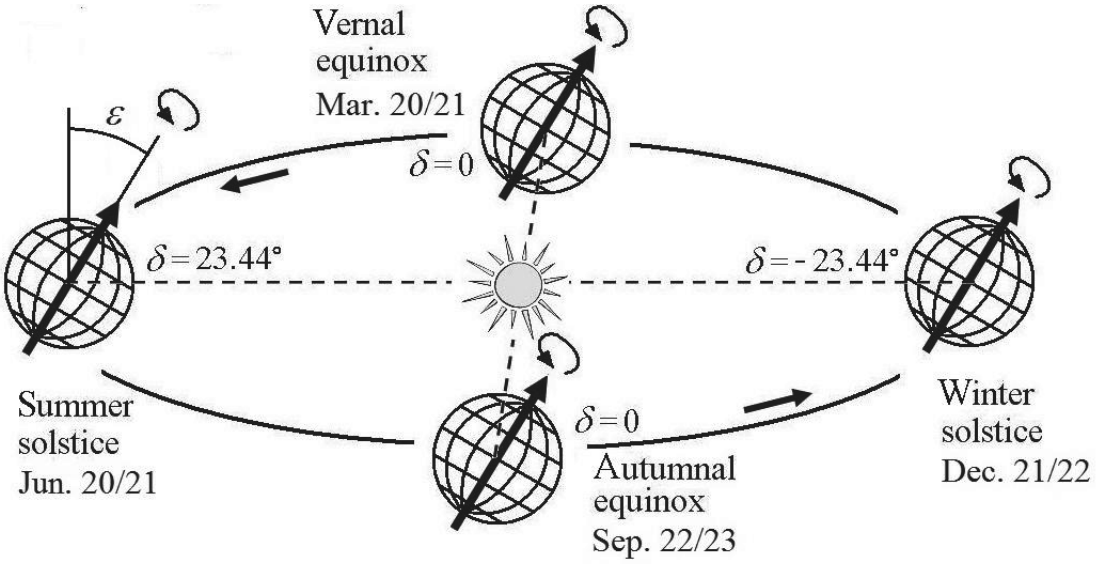


Figure 2.2.2 – The solar declination varies throughout the year. δ is the solar declination and ε is the obliquity of the ecliptic. $\delta = 0$ at vernal and autumnal equinoxes and $\delta = \pm 23.44^{\circ}$ at summer and winter solstices. The figure is adopted from Chen (2011) and used with permission from John Wiley & Sons.

Equation of time

The equation of time is the time difference between true solar time and mean solar time. The mean solar time is defined from the mean solar noon and 24 hour difference between each noon. The equation of time varies throughout the year and is caused by two events:

- While the earth rotates around its own axis at a constant angular velocity, it also has forward motion with various speed due to the eccentricity of its elliptic orbit (Kepler's second law - the line between the earth and the sun always covers the same area at a certain time interval). This results in a sinusoidal time difference since in mean solar time, the earth's orbit is circular.

- As a result of the obliquity of the ecliptic, ε, the apparent sun motion differs from the mean sun motion. The mean sun moves in the equatorial plane, while the apparent sun moves in the ecliptic plane. A sinusoidal effect occurs with zero difference in the equinoxes and solstices.

The equation of time using radians is estimated in Chen (2011) and shown in Eq. 2.2.2.

$$EOT = \left[9.85\sin\left(\frac{4\pi\,(N-80)}{365.2422}\right) - 7.65\sin\left(\frac{2\pi\,(N-3)}{365.2422}\right)\right] \tag{2.2.2}$$

Where N is the number of day in the year starting with January 1st and the day of the perihelion (the day the earth is closest to the sun) is assumed to be January 3rd. The first component is a result of the obliquity of the ecliptic and the second component is due to the eccentricity of the earth's elliptic orbit. The two components and the summation are shown in Fig. 2.2.3.

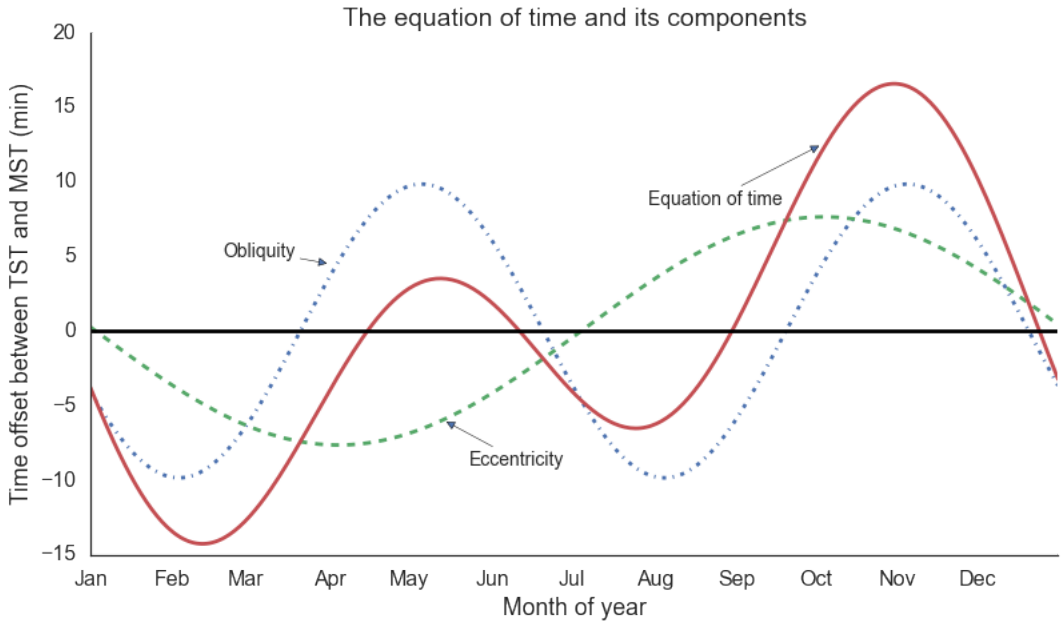


Figure 2.2.3 – The obliquity and eccentricity components and their summation, the equation of time. TST is true solar time and MST is mean solar time.

In addition to the equation of time, the longitude is used to convert solar mean time to local standard time. For all international standard time zones, 15° longitude corresponds to one hour, which implies that 1° longitude correspond to 4 minutes. Finally, the conversion from true solar time (TST) to local standard time (LST) can be calculated as in Iqbal (1983) and is shown in Eq. 2.2.3.

$$LST = TST - 4(L_s - L_e) - EOT \tag{2.2.3}$$

Where LST is the local standard time, TST is the true solar time, L_s is the standard longitude for the given timezone, L_e is the local longitude and EOT is the equation of time. With Eq. 2.2.3 true solar time can easily be converted to local standard time.

Sun's position relative to an observer on earth

Normally, for solar energy applications, the sun's position is expressed relative to the observer on earth and thus the trigonometric relationship has to be explained. A celestial sphere, as shown in Fig. 2.2.4, is used to illustrate this.

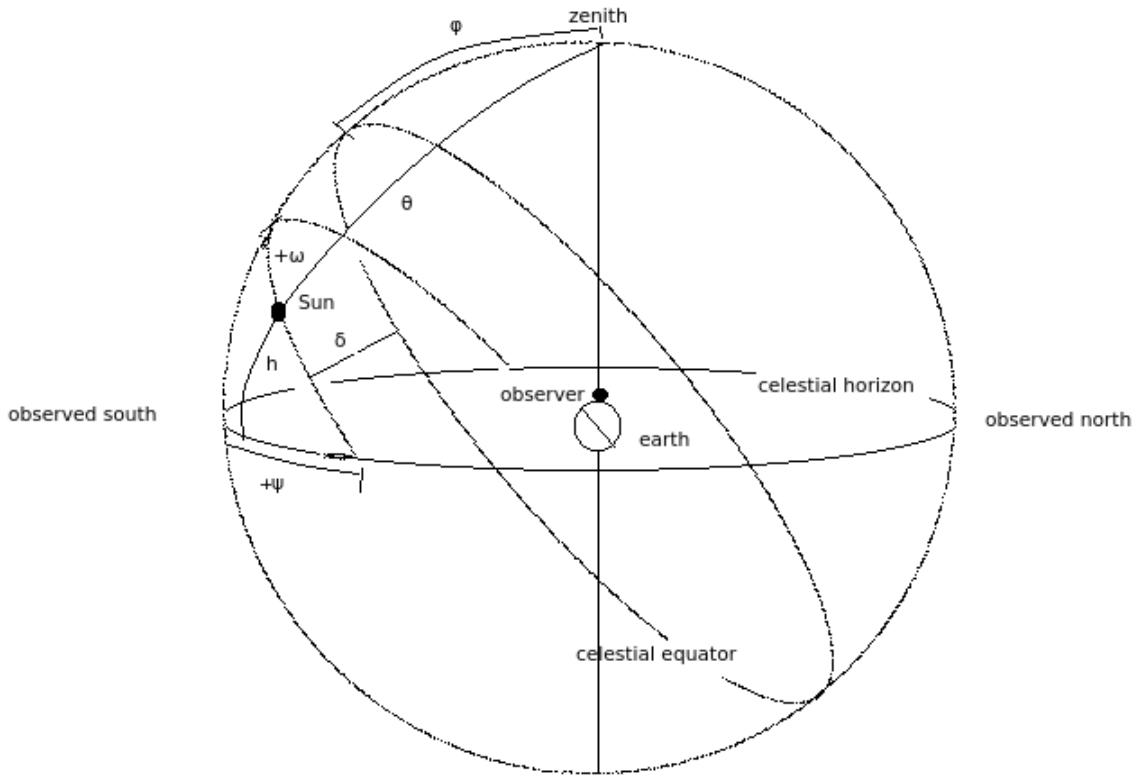


Figure 2.2.4 – A celestial sphere with the earth as its center. δ is the solar declination, ϕ is the latitude, ψ is the azimuth angle, ω is the hour angle, h is the solar elevation and θ is the zenith angle. The figure is inspired by Iqbal (1983).

The observer's position on earth is defined by its zenith, which is directly vertical above the observer on the celestial sphere. The angle from zenith to the position of the sun is known as the solar zenith angle, θ, and is 0° in zenith. When the solar zenith angle is 90°, it intersects the plane of the horizon. Along the plane of the horizon is the azimuth angle, ψ, which is defined as 0° in observer's south and is east positive. The angle between the horizon and the position of the sun is known as the solar elevation, h, which is essentially only $90° - \theta$. Along the path of the sun is the hour angle, ω, which is usually defined as positive before noon and negative after noon. As described by Chen (2011), the expression

for the solar elevation and solar zenith angle is shown in Eq. 2.2.4.

$$\sin h = \sin\delta \sin\phi + \cos\delta \cos\phi \cos\omega = \cos\theta \tag{2.2.4}$$

During sunrise and sunset is $\omega = \omega_s$ and $\sin h = \cos\theta = 0$. Thus, follows the expression shown in Eq. 2.2.5 from Eq. 2.2.4.

$$\omega_s = \cos^{-1}(-\tan\phi \tan\delta) \tag{2.2.5}$$

As described earlier in this chapter, for every hour, ω moves 15°, which means that the relationship between time and hour angle in degrees can be described in Eq. 2.2.6.

$$t = 12 - \frac{\omega}{15} \tag{2.2.6}$$

Where t is true solar time in hours and solar noon is at 12.

2.3 Climate in Eastern Norway and its influence on solar irradiance

The Norwegian climate is heavily influenced by the warm North Atlantic Current and the Westerlies. Hot and humid air currents from the south encounter the cold air currents from the north and create a polar front. On the boundary, low-pressure areas emerge and move east (Utaaker, 1991). These air currents are generally humid and are forced to move upwards by the high mountains and condensates. This results in rain on the windward side (west) and dry air on the leeward side (east). As a consequence, the eastern part of Norway has dry weather to a greater extent compared to the western part of Norway. Yearly mean precipitation values for Norway are shown in Fig. 2.3.1. Since the four stations evaluated in this thesis are located in proximity of each other, the main factor is local variations in climate. The frequency of cloud formation is a good indicator of the rate of incoming solar radiation that reaches the surface of the earth. Thus, observations of *sky octas* from nearby stations were collected from eKlima (2015) and are shown in Fig. 2.3.2. Sky octas is a unit used in the measurement of cloud cover. It ranges from 0 (completely clear sky) to 8 (completely cloudy). The measurements are usually carried out by a trained observer. The figure shows that Torp and Melsom, which are close to Ramnes, have a lower amount of cloud cover, especially in the summer months.

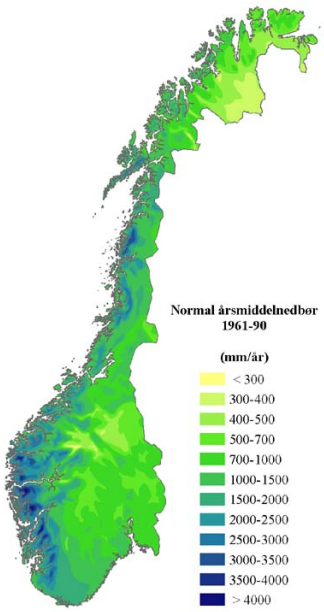


Figure 2.3.1 – Yearly mean precipitation values for Norway made by the Norwegian Meteorological Institute (met.no, 2015). Displayed with permission.

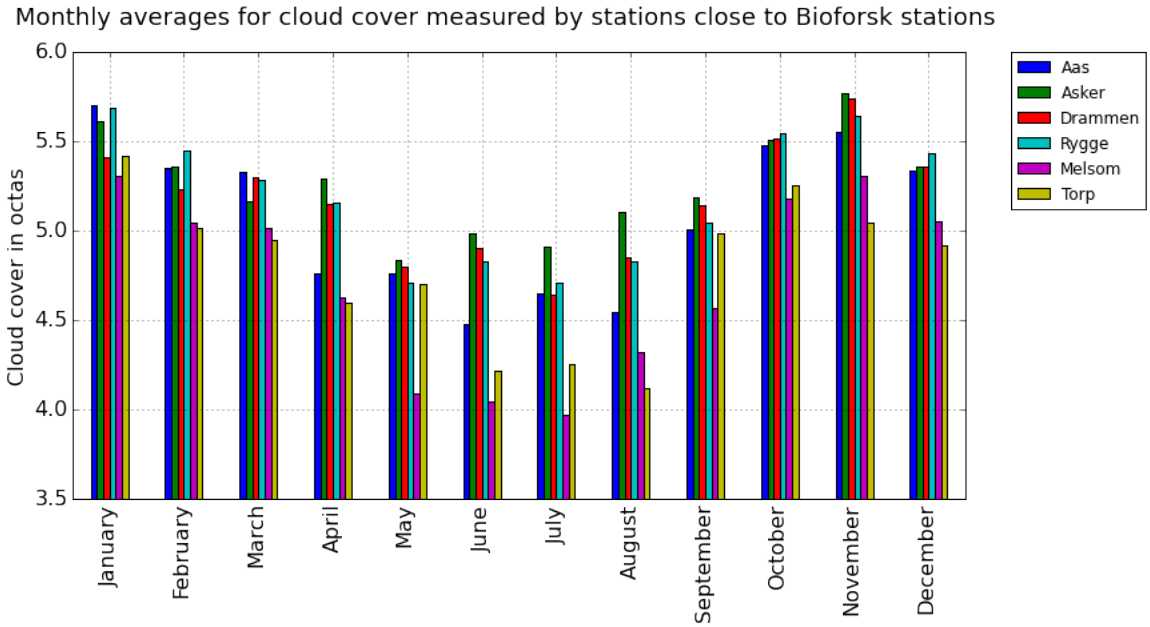


Figure 2.3.2 – Average monthly sky octas for relevant stations. Drammen and Asker are in proximity to Lier, Rygge is nearby Tomb and Melsom and Torp are south of Ramnes. Data is gathered from eKlima (eKlima, 2015).

2.4 Measurement techniques

In order to study solar irradiation, measurement instruments have been invented and applied. In general, two types of instruments are commonly used to calculate solar irradiation, the pyrheliometer and the pyranometer. The pyrheliometer is used to measure beam irradiance, I_{beam}, while the pyranometer is used to measure global irradiance, I_{global}. Diffuse irradiance, I_{diffuse}, can be calculated with a pyranometer using a shading device, which prevents incoming beam irradiance to reach the pyranometer. Each instrument has a sensor which converts the solar radiation input to a number. Iqbal (1983) classifies four types of radiation sensors.

1. Calorimetric sensors estimate the heat generated by solar radiation. The added heat can be measured with a flowing fluid or enthalpy, or in more recent years, with a transducer, calculating the electrical heating needed to keep a constant temperature difference.

2. Thermomechanical sensors are based on the principle that metals expand with an increase in temperature. If one strip is exposed to solar radiation and another isolated from it, a difference in expansion between the two represents the amount of incoming solar radiation.

3. Thermoelectric sensors are very commonly used today. These sensors exploit the thermoelectric effect, which is the fact that any conducting material affected by a temperature gradient generates voltage. The temperature gradient causes charge carriers to move from the hot side to the cold side. Two dissimilar metals which have their ends connected are called a thermocouple. If there is a temperature difference between the two ends of a thermocouple, an output electric voltage is generated. For thermoelectric sensors, sensitivity is a very important variable. It indicates the voltage output for a given amount of solar irradiance. The sensitivity may change over time due to sun exposure or poor maintenance.

4. Photoelectric sensors utilize the photovoltaic effect and are made from a doped semiconductor. While a photoelectric sensor has a number of applications, its spectral response varies greatly with the material used.

2.4.1 Pyrheliometer

The pyrheliometer is used to measure beam irradiance, I_{beam}. It has a telescopic shape with a narrow opening facing the sun. A pyrheliometer requires a sun tracker to follow the motion of the sun. Early versions used calorimetric sensors with water or a silver disk

(Iqbal, 1983). Knut Ångström invented an electrical compensated pyrheliometer, where a material is electrically heated to a temperature equal to the temperature of the material which was heated by solar radiation. In the 1960s, cavity pyrheliometers were invented. These instruments have a cone shaped receiver in order to increase the absorption of the material. Cavity pyrheliometers are often used as reference pyrheliometers. For field measurements, however, thermoelectric sensors are often used.

2.4.2 Pyranometer

The pyranometer is used to measure global irradiation, I_{global}, which is a vital parameter to the estimation of power production from solar energy. Pyranometers have a hemispherical field of view and is usually installed in a horizontal position. At most of the Bioforsk stations, the Kipp & Zonen CM11 pyranometer is used.

Kipp & Zonen CM11 pyranometer The Kipp & Zonen CM11 pyranometer is commonly used as a field instrument to measure solar irradiation. A schematic figure of the device is shown in Fig. 2.4.1. These pyranometers consist of a black painted disk sealed by two glass domes, which prevents incoming thermal radiation and isolates the sensor. The heat generated in the disk flows through a thermal resistor to a heat sink and creates a thermoelectric voltage, which is interpreted by a measuring device. The disk contains 100 thermocouples, which becomes a thermopile. Furthermore, the pyranometer has a desiccant, which absorbs water molecules to prevent humidity damage on the sensor. The CM11 pyranometer comply with the requirements of a high quality pyranometer by WMO (2008) and is classified as a secondary standard pyranometer by the International Standard Organization, which is the highest quality class ISO (1990). In addition, it is regarded as the standard reference pyranometer due to its accuracy, stability and quality of construction (Muneer and Fairooz, 2002). An adopted table by Kipp & Zonen is shown in Table 2.4.1 (Kipp & Zonen, 2000).

Table 2.4.1 – Adopted table from the Kipp & Zonen CM11 pyranometer manual showing specifications related to errors and operation, (Kipp & Zonen, 2000).

Characteristics	CM 11	High quality	Good quality	Moderate quality
ISO 9060 classification		Secondary standard	First class	Second class
Response time (95 percent response)	12 s	< 15 s	< 30 s	< 60 s
Zero offset: (a) Response to 200 W/m^2 net thermal radiation (ventilated) (b) Response 5 K/h change in ambient temperature	 <± 7 W/m^2 (with CV 2) <± 2 W/m^2	 <± 7 W/m^2 <± 2 W/m^2	 <± 15 W/m^2 <± 4 W/m^2	 <± 30 W/m^2 <± 8 W/m^2
Resolution (smallest detectable change)	<± 1 W/m^2	± 1 W/m^2	± 5 W/m^2	± 10 W/m^2
Stability (change per year, percentage of full scale)	< 0.5	< 0.8	< 1.5	< 3.0
Directional response of beam radiation (The range of errors caused by assuming that the normal incidence responsivity is valid for all directions when measuring, from any direction, a beam radiation whose normal incidence irradiance is 1000 W/m^2)	<± 10 W/m^2	<± 10 W/m^2	<± 20 W/m^2	<± 30 W/m^2
Temperature response (percentage of maximum due to any change of ambient temperature within an interval of 50 K)	< 1 -10^{0}C..+40^{0}C	< 2	< 4	< 8
Non-linearity (percentage deviation from the responsivity at 500 W/m^2 due to any change of irradiance within the range 100 to 1000 W/m^2)	< 0.5	< 0.5	< 1	< 3
Spectral sensitivity (percentage of deviation of the product of spectral absorptance and spectral transmittance from the corresponding mean within the range of 0.3 to 3 µm)	< 2	< 2	< 5	< 10
Tilt response (percentage deviation from the responsivity at 0° tilt, horizontal, due to change in tilt from 0° to 90° at 1000 W/m^2 irradiance)	< 0.2	< 0.5	< 2	< 5
Achievable uncertainty, 95 percent confidence level Hourly totals Daily totals	 3 % 3 %	 3% 2%	 8% 5%	 20% 10%

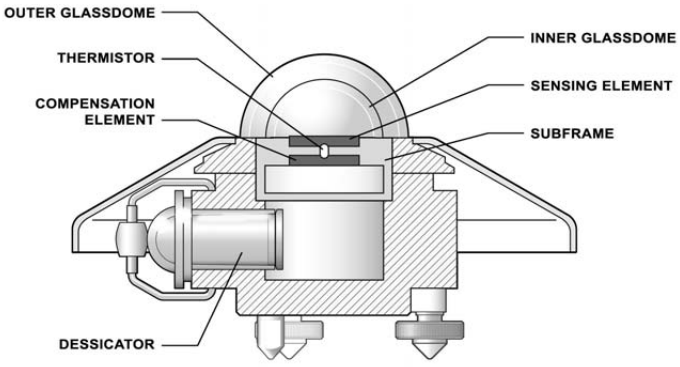


Figure 2.4.1 – A schematic diagram of the CM11 Pyranometer, (Kipp & Zonen, 2000).

2.4.3 Sources of errors

Younes et al. (2005) present several sources to errors and uncertainties related to solar radiation measurements. They can be categorized in two categories.

Equipment error and uncertainty Younes et al. (2005) state that the largest source of error is the *cosine response*. The cosine response is the fact that the response of the pyranometer varies with the angle of incidence (Iqbal, 1983). At solar altitudes below 6°, the error of the cosine response becomes increasingly high. In addition, the response of the pyranometer varies with the azimuthal change of the position of the sun, called *azimuth response*. This is a result of imperfections in the glass domes and the angular reflection properties of the black paint (Younes et al., 2005). Both angular responses are often described as directional response (WMO, 2008).

The pyranometer also has a *temperature response* which is an error caused by large or rapid changes in external temperature (Clarke, 2009). The glass domes and the black paint of the sensing material have a spectral selectivity and the output from the solar radiation may not be linear with intensity and thus could cause a small error. During a clear sky, the inner dome radiates due to a temperature difference, which causes a temperature drop inside the dome and could be recorded as negative radiation (zero offset) by the thermopile. Another potential error is the non-stability of the instrument over time, which may result in a very different sensitivity from its initial calibration.

Operation related problems and errors Numerous errors might be caused by the operation of a pyranometer. Snow, dew, dust and bird droppings are not uncommon problems with the Bioforsk stations (Kroken, 2015). To prevent a substantial impact of such errors, it is important to clean the pyranometer dome frequently. Other factors stated by Younes et al. (2005) are incorrect sensor leveling, shading caused by building structures, electric field in the close by cables, mechanical loading of cables and station shutdown.

2.5 Uncertainty calculation

Uncertainty calculation is a very important part of measurements. Unfortunately, no instrument measures perfectly and thus a calculation of its uncertainty is required. In this thesis it is natural to base the calculations on the given uncertainty in the manual of the measuring instruments. For the Kipp & Zonen CM 11 pyranometer, an uncertainty of 3 % is given for hourly and daily totals. In addition, more factors needs to be considered. Change in sensitivity, maintenance and replacement of pyranometers will influence the accuracy. This

consideration may be rather subjective and in this thesis, an effort to estimate exact accuracy of the measurements has not been made.

2.6 Statistics

To compare a model or database with observed values, statistical methods are commonly used. In this thesis, three methods are applied; mean bias deviation, mean absolute error and root-mean-square deviation. In addition, the relative change between two values is explained.

2.6.1 Mean bias deviation (MBD)

Mean bias deviation describes the overall bias of a model. In general, if the MBD is negative, the model underestimates compared to the observed value. The MBD is defined in Eq. 2.6.1.

$$MBD = \frac{\sum_{i=1}^{n} (x_{\text{model}_i} - x_{\text{measurement}_i})}{\sum_{i=1}^{n} x_{\text{measurement}_i}} \tag{2.6.1}$$

Where x_{model_i} is the estimated value from a model, $x_{\text{measurement}_i}$ is the corresponding measured value and n is the number of values used in the series.

2.6.2 Mean absolute error (MAE)

Mean absolute error describes the absolute bias of a model. Compared to MBD, MAE does not factor in whether the bias is negative or positive and thus any error adds to the total error. The MAE is defined in Eq. 2.6.2.

$$MAE = \frac{\sum_{i=1}^{n} |x_{\text{model}_i} - x_{\text{measurement}_i}|}{\sum_{i=1}^{n} x_{\text{measurement}_i}} \tag{2.6.2}$$

Where x_{model_i} is the estimated value from a model, $x_{\text{measurement}_i}$ is the corresponding measured value and n is the number of values used in the series.

2.6.3 Root-mean-square deviation (RMSD)

Root-mean-square deviation estimates the standard deviation of the differences between the model and measurement. It is a good indicator on how accurate the model estimates the

measured values. The RMSD is defined in Eq. 2.6.3.

$$RMSD = \frac{\sqrt{\frac{1}{\mathrm{n}}\sum_{i=1}^{n}\left(x_{\mathrm{model}_i} - x_{\mathrm{measurement}_i}\right)^2}}{\frac{1}{\mathrm{n}}\sum_{i=1}^{n} x_{\mathrm{measurement}_i}} \tag{2.6.3}$$

Where x_{model_i} is the estimated value from a model, $x_{\mathrm{measurement}_i}$ is the corresponding measured value and n is the number of values used in the series.

2.6.4 Relative change

The relative change is a common method to compare how much percentage a value differs from a reference value. The relative change is defined in Eq. 2.6.4.

$$\text{Relative change} = \frac{x - x_{\mathrm{ref}}}{x_{\mathrm{ref}}} \tag{2.6.4}$$

Where x is the value that is compared with the reference value, x_{ref}.

Chapter 3

Data

3.1 Bioforsk (Landbruksmeterologisk tjeneste)

Bioforsk is the Norwegian Institute for Agricultural and Environmental Research. Their objective is to provide industries, governments and consumers with new knowledge, services and solutions within multifunctional agriculture, plant sciences, environmental protection and natural resource management (Bioforsk, 2015). As an important part of their research, "Landbruksmeteorologisk Tjeneste" (lmt.bioforsk.no), is a agro meteorological measuring service which provides measurements of temperature, precipitation, wind speed and direction, soil temperature and moisture, leaf wetness, and global solar irradiation. A few stations also provide sunshine hours. Today, 47 Bioforsk stations nationwide measure solar irradiation (Byrkjedal et al., 2013). The first 10 stations were installed during 1987 and most of the remaining stations were installed during the early 1990s. The measurement device used for most of the stations are the Kipp & Zonen CM11 Pyranometer (Kipp & Zonen, 2000), which is described in detail in subsection 2.4.2 on page 14. The measured values are sent through a GPRS system and stored in a central database as hourly data. The length and frequency of the time series satisfy the criteria set by Myers (2005) for "Performance and economics, system lifetime applications". The regularly maintenance of the stations are performed by station hosts who work and live nearby. Their tasks are to clean the dome of the pyranometer frequently, once a week if possible, change desiccant whenever needed and mowing the lawn regularly during growing season (ITAS, 2013). In addition, a yearly calibration routine is performed by ITAS Eierdrift to ensure no offsets.

There are 21 Bioforsk stations with over 15 years of measured solar irradiation data located in Eastern Norway. The data is registered as the mean of the measured solar irradiance

within the hour. A preliminary quality control procedure was performed to filter out stations which obviously contain too much erroneous data. Afterwards, 16 stations remained and four stations close to Oslo were chosen to be examined further in this thesis. These are Ås, Lier, Ramnes and Tomb and they are described further in subsections 3.1.1, 3.1.2, 3.1.3 and 3.1.4 respectively. The time series evaluated in this thesis is from 1992 to 2012. The locations are shown in Fig. 3.1.1 and described in Table 3.1.1.

Figure 3.1.1 – Map of the Oslofjord region with the four stations evaluated in this thesis marked with blue circles. The region is located in the red box in map of Norway in the right corner of the figure. The map is generated through the web service www.norgeskart.no.

Table 3.1.1 – Location info for the four Bioforsk stations examined in detail in this thesis. Latitude is north positive and longitude is east positive. Elevation is the height above sea level.

Name	Latitude(°)	Longitude(°)	Elevation (m)
Ås	59.660468	10.781989	94
Lier	59.79005	10.2604	39
Ramnes	59.38081	10.23923	38
Tomb	59.31893	10.81449	12

3.1.1 Ås

Ås is the northeast most station of the four. It is positioned at Sørås field, where the meteorological field station, FAGKLIM, maintained by the Norwegian University of Life Sciences (NMBU), is located. The Bioforsk station is not a part of FAGKLIM, but is still maintained by the station host of FAGKLIM. The pyranometer dome is cleaned and checked every weekday (Monday to Friday), which is more often than expected for the Bioforsk stations. As a consequence, an assumption could arguably be made that Ås should have the least operational related errors of the four stations. In addition, the Sørås field has the lowest horizon of the four stations, which is shown in Fig. 3.1.2. The pyranometer was changed in 2003.

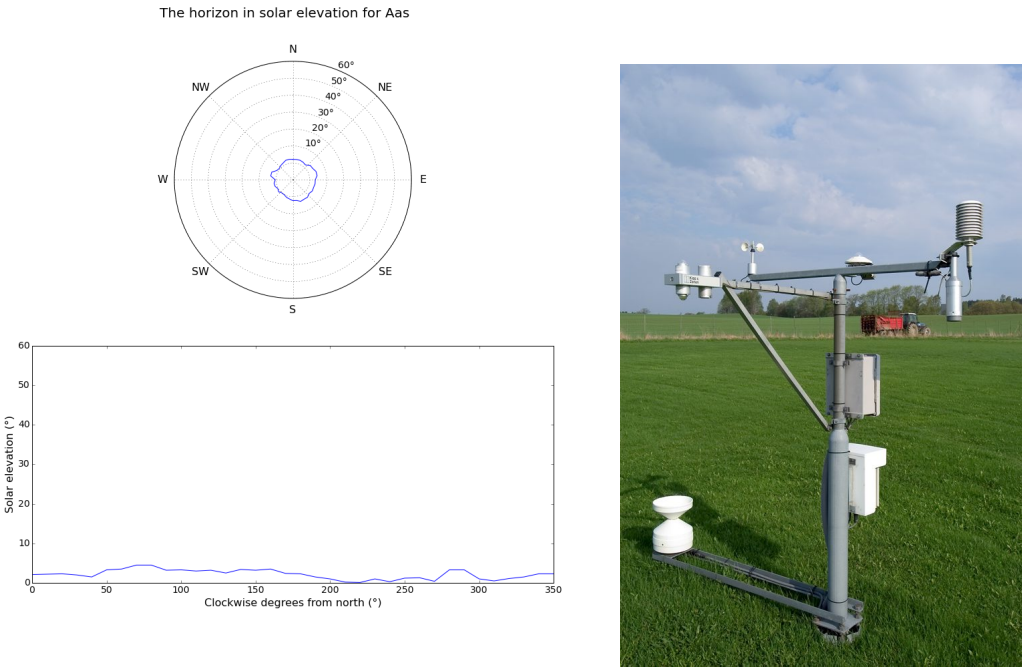


Figure 3.1.2 – To the left are two figures which describe the horizon for Ås, measured in elevation from the ground. The upper figure shows the horizon in polar coordinates, where a larger value indicate a higher elevation in degrees above the horizon. In the lower figure is the horizon from north plotted against the solar elevation in degrees. To the right is a picture of the station facing northwest. Photo by ITAS/Bioforsk.

3.1.2 Lier

Lier is the northern most station of the four. It is unique since it is the only station positioned in a valley. Since it is placed in the eastern part of the valley, its horizon to the east is thus limited as shown in Fig. 3.1.3. The pyranometer was changed in 2010.

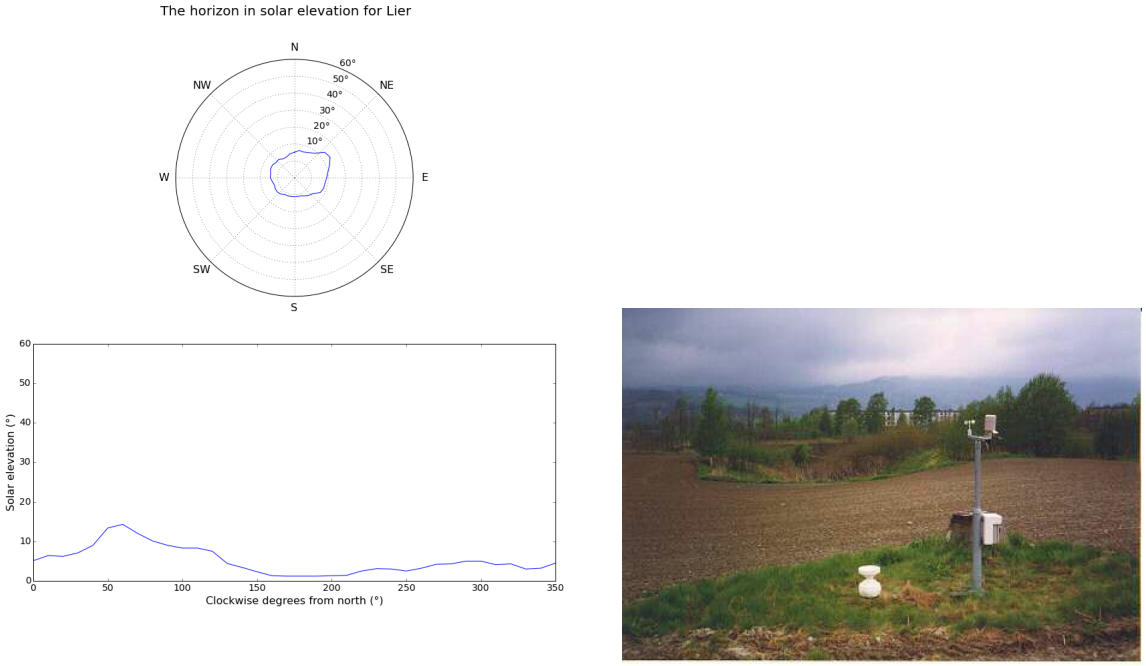


Figure 3.1.3 – To the left are two figures which describe the horizon for Lier, measured in elevation from the ground. The upper figure shows the horizon in polar coordinates, where a larger value indicate a higher elevation in degrees above the horizon. In the lower figure is the horizon from north plotted against the solar elevation in degrees. To the right is a picture of the Lier station facing west. Photo by ITAS/Bioforsk.

3.1.3 Ramnes

Ramnes is located southwest of Oslo, close to the cities Horten and Tønsberg. Its horizon is shown in Fig. 3.1.4. From Fig. 2.3.2 on page 12, Ramnes is expected to have the least cloud cover of the four stations. The pyranometer was changed in 2003.

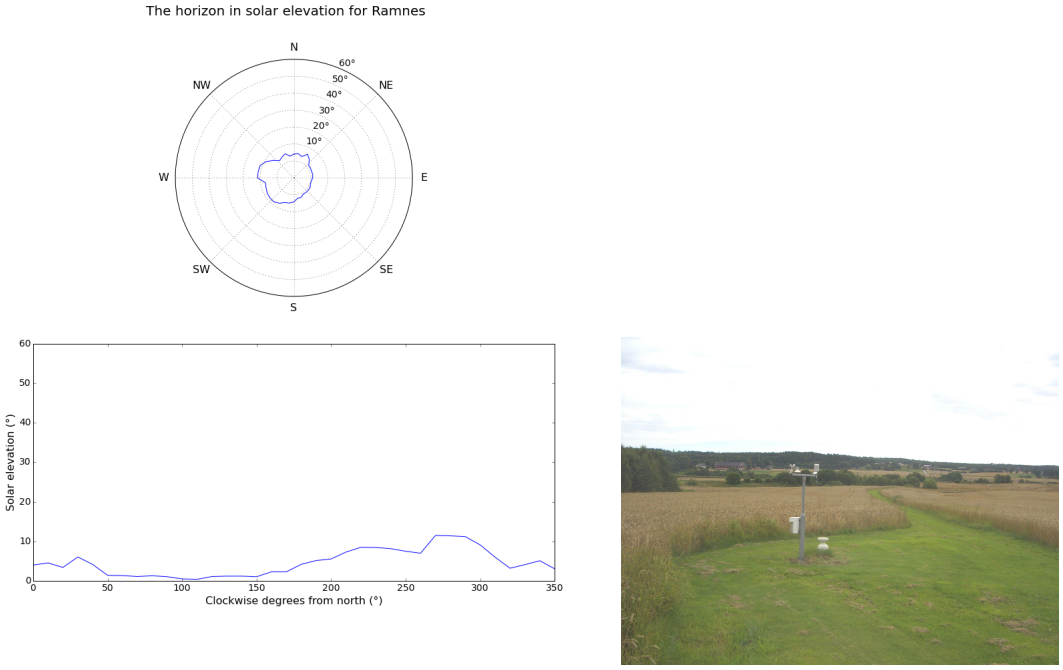


Figure 3.1.4 – To the left are two figures which describe the horizon for Ramnes, measured in elevation from the ground. The upper figure shows the horizon in polar coordinates, where a larger value indicate a higher elevation in degrees above the horizon. In the lower figure is the horizon from north plotted against the solar elevation in degrees. To the right is a picture of the Ramnes station facing east. Photo by ITAS/Bioforsk.

3.1.4 Tomb

Tomb is located southeast of Oslo, close to the cities Moss and Fredrikstad. The Tomb station is closest of the four to the Oslofjord. This could introduce formation of fog in some parts of the year. Its horizon is shown in Fig. 3.1.5. The pyranometer was changed in 2013.

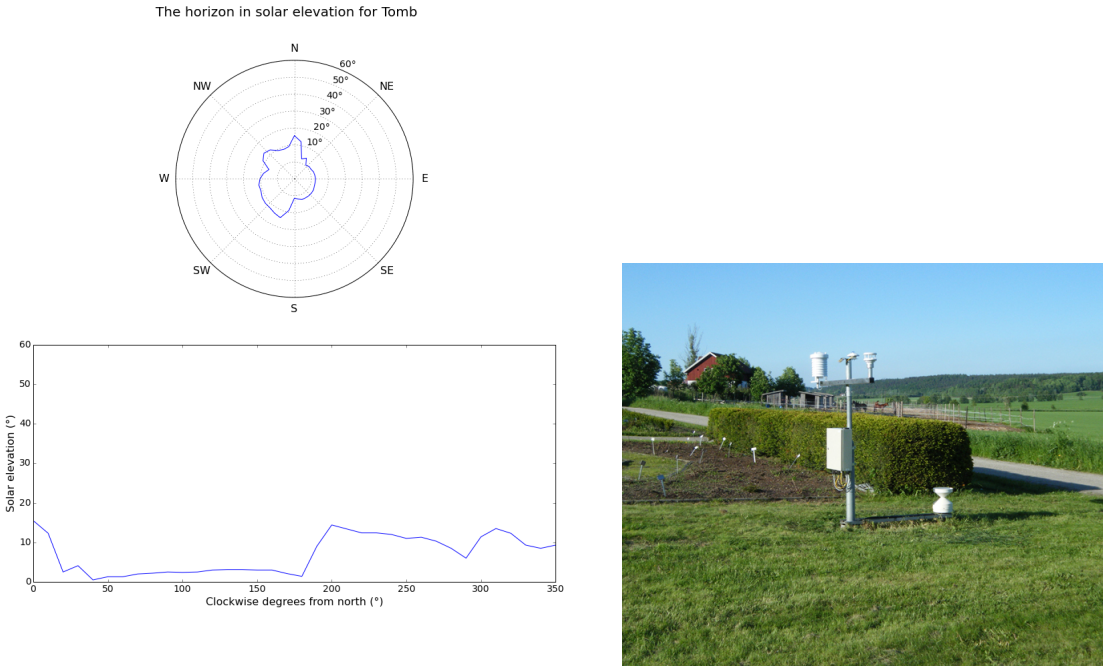


Figure 3.1.5 – To the left are two figures which describe the horizon for Tomb, measured in elevation from the ground. The upper figure shows the horizon in polar coordinates, where a larger value indicate a higher elevation in degrees above the horizon. In the lower figure is the horizon from north plotted against the solar elevation in degrees. To the right is a picture of the Tomb station facing northeast. Photo by ITAS/Bioforsk.

3.1.5 Replacement of pyranometers

If a pyranometer stops working, the pyranometer dome shatters or there is a poor maintenance, the pyranometer is replaced with a new one (Ruud Hansen, 2015). The new pyranometer has a different sensitivity compared to the previous one and the data logger has to be initialized with a new sensitivity. If possible, the old pyranometer should be recalibrated in order to find out how much the sensitivity has changed in the period it has been operated. This has unfortunately not been performed for the four stations evaluated in this thesis. During the time period from 1992 to 2012, three pyranometers have been replaced. Ås and Ramnes changed their pyranometer in 2003 and Lier changed its pyranometer in 2010. General information on each pyranometer is shown in Table 3.1.2. The serial numbers for the originally installed pyranometers are unfortunately not known.

Table 3.1.2 – The serial number, sensitivity, calibration date and installation date for the pyranometers used in the four Bioforsk stations. N/A is shown when the information is not available. Beginning of time series varies (Ruud Hansen, 2015).

Station	Sensor type	Serial nr.	Sensitivity	Calibrated	Installed
Ås	Kipp&Zonen CM11	N/A	$4.47\frac{\mu V}{W/m^2}$	N/A	1991/08/30
Ås	Kipp&Zonen CM11	924073	$4.42\frac{\mu V}{W/m^2}$	2003/03/24	2003/05/09
Lier	Kipp&Zonen CM11	N/A	$4.47\frac{\mu V}{W/m^2}$	N/A	1991/12/31
Lier	Kipp&Zonen CM11	986512	$5.17\frac{\mu V}{W/m^2}$	2010/01/22	2010/04/20
Ramnes	Kipp&Zonen CM11	N/A	$4.42\frac{\mu V}{W/m^2}$	N/A	1991/01/02
Ramnes	Kipp&Zonen CM11	882367	$4.36\frac{\mu V}{W/m^2}$	2003/11/04	2003/11/07
Tomb	Kipp&Zonen CM11	N/A	$4,47\frac{\mu V}{W/m^2}$	N/A	1991/01/02
Tomb	Kipp&Zonen CMP11	139350	$7.68\frac{\mu V}{W/m^2}$	2013/02/22	2013/06/06

3.2 Modeled irradiance

3.2.1 Modeled extraterrestrial irradiance

Extraterrestrial irradiance, also known as solar irradiance at the top of the atmosphere, is a very important boundary used in quality control of measured solar irradiance. The extraterrestrial irradiance as estimated in Duffie and Beckman (1980) is shown in Eq. 3.2.1.

$$I_E = S(1 + 0.033\cos(\frac{360}{365}N)\cos\theta \tag{3.2.1}$$

Where I_E is the extraterrestrial irradiance, S is the solar constant, N is the number of day in the year starting with January 1st and θ is the solar zenith angle in degrees. The modeled value for the extraterrestrial radiation is in the middle of each hour of the data set.

3.2.2 Modeled clear sky irradiance

The model used for this thesis is *libRadtran* (Mayer and Kylling, 2005). *libRadtran* is a library for radiative heat transfer. The version used for this thesis is version 1.7. The clear sky time series has been estimated with the use of *rtmrun*, a Perl wrapper around *libRadtran*, designed by Øystein Godøy (Godøy, 2013). Equal to modeled extraterrestrial irradiance, the modeled clear sky is modeled for the middle of each hour of the data set. An example configuration file for Ramnes is shown in Appendix A. The advantage of libRadtran is that it does not require additional parameters for the calculation.

3.3 Existing solar irradiation databases

3.3.1 Satel-Light

The Satel-Light project was funded by the European Union (Directorate General XII) from 1996 to 1998 and provides satellite derived data from 1996 to 2000 (Fontoynont et al., 1998). Its data is stored in a web server, www.satel-light.com. The website provides 30 min data for all stations. The solar irradiation is derived from satellite images using the Heliosat procedure first presented in (Cano et al., 1986). The spatial resolution is dependent on the latitude since there is a limited pixel size from the satellite images. For Scandinavia a pixel represent 5 km by 16 km (Hagen, 2011).

3.3.2 NASA SSE 6.0

“These data were obtained from the NASA Langley Research Center Atmospheric Science Data Center Surface meteorological and Solar Energy (SSE) web portal supported by the NASA LaRC POWER Project" (Stackhouse and Whitlock, 2008). The NASA SSE data have a spatial resolution limited to 1 latitude by 1 longitude, which is around 50 km by 100 km in the southern part of Norway. This means that the data provided for all stations that are studied in this thesis are equal, since the stations are located within 59 to 60 degrees north and 10 to 11 degrees east. The data is estimated from satellite images from 1983 to 2004.

3.3.3 WRF model

WRF (The Weather Research and Forecasting Model, www.wrf-model.org) is “a next-generation mesoscale numerical weather prediction system designed to serve both atmo-

spheric research and operational forecasting needs" (Michalakes et al., 1998). The data in this thesis are provided by Kjeller Vindteknikk. The time period is from 1979 to 2012 and the spatial resolution is 6 km by 6 km.

3.3.4 Meteonorm 7

The global climatological database Meteonorm (www.meteonorm.com) is often used for planning of solar energy systems or buildings (Remund, 2008). The data gathered is from the time period 1991 to 2010 and the spatial resolution is 1 km by 1 km.

3.3.5 PVGIS

PVGIS is a web-based solar radiation database developed with a GIS-based methodology (Šúri et al., 2005). In this thesis two databases developed by PVGIS are used. The data is readily available for free online through http://re.jrc.ec.europa.eu/pvgis/.

3.3.5.1 Classic PVGIS

The Classic PVGIS estimates solar irradiation based on ground measurements that were originally part of the European Solar Radiation Atlas (Scharmer and Greif, 2000) and the r.sun radiation model within the GRASS GIS software. The spatial resolution is 1 km by 1 km and the time period is from 1981 to 1990.

3.3.5.2 CM-SAF PVGIS

The CM-SAF PVGIS estimates solar irradiation based on satellite images performed by CM-SAF (The Satellite Application Facility on Climate Monitoring, http://www.cmsaf.eu), which represents 12 years of data. The data is gathered from the first generation of Meteosat satellite images from 1998 to 2005 and the second generation of Meteosat satellite images from June 2006 to December 2011. The spatial resolution around 2.5 km x 2.5 km. (Huld et al., 2012; PVGIS, 2015). The CM-SAF PVGIS database was originally only available at latitudes below 58°N due to high uncertainty, but have since the release in 2012 expanded to 60°N. This implies that the database is available for all four stations evaluated in this thesis, however, not for all Bioforsk stations in Eastern Norway.

Chapter 4

Methods

4.1 Automatic quality control

Large data sets require an automatic control of data as a first step in the quality control procedure. The data processed through the automatic control will have a flag on each test it failed. The manual part of the quality control procedure will thus be easier to handle, since much of the erroneous data is now identified. The automatic quality control procedure is described in the subsections below.

4.1.1 Offset test

As explained in subsection 2.4.3 on page 16, the pyranometers have a tendency to output negative values during clear sky nights due to zero offset. According to the CM11 Pyranometer manual, the output should not fall lower than $-12\frac{\mathrm{W}}{\mathrm{m}^2}$ (Kipp & Zonen, 2000). In addition, data points which exceed a certain positive value during nighttime due to offset should be flagged. As explained in section 3.1 on page 19, the data is gathered as the mean of one hour. The solar zenith angle for each hour is set as the midpoint of the hour. Therefore, the measurement may register some solar irradiation even though the time step show a solar zenith angle which exceeds 90°. To account for this, only values with solar zenith angle more than 93° are considered. To sum up, the Offset tests consists of two conditions which have to be fulfilled in order to avoid a flag. The tests are shown in Eq. 4.1.1 and 4.1.2.

$$I_{\text{global}} \geq -12\frac{\mathrm{W}}{\mathrm{m}^2} \tag{4.1.1}$$

$$I_{\text{global}} \leq 6\frac{\mathrm{W}}{\mathrm{m}^2} \text{ if } \theta > 93^\circ \tag{4.1.2}$$

Where I_{global} is the measured global horizontal irradiation and θ is the solar zenith angle. If any data point does not satisfy both of these criteria, that data point is flagged for visual control. The offset test is unique, since it tests the data which is recorded during the night. Its purpose is to recognize patterns of an incorrect offset of the measured data.

Set nighttime values to zero For further quality control, it is not needed to keep the measured nighttime values. While they are not removed, the logic move is to set the nighttime values to zero, since there is obviously no solar irradiance at night. The procedure utilized in this thesis is to set $I_{\text{global}} = 0$ where $I_{\text{E}} = 0$.

4.1.2 Upper limit tests

An advantage of solar irradiation quality control is the regular motion of the sun. At any given time, the upper limit of solar irradiance is known. Shi et al. (2008), however, described that the diffusive effect of clouds not in the way of solar beam, may result in a solar radiation at the surface of the earth which exceeds this upper limit. Nonetheless, the phenomenon does not last over longer periods. It is thus an error if the pyranometer measure an hourly average amount of solar irradiance which is greater than extraterrestrial irradiance. The test have been presented in Geiger et al. (2002), Younes et al. (2005), Shi et al. (2008), Moradi (2009), Tang et al. (2010) and Journée and Bertrand (2011). In this thesis, the test is called U_1 and is shown in Eq. 4.1.3.

$$I_{\text{global}} \leq I_{\text{E}} \tag{4.1.3}$$

Many consider I_{E} to be a too high of a limit. The solar beam is always attenuated by the atmosphere and a clear sky model can be estimated. In Geiger et al. (2002), Younes et al. (2005) and Shi et al. (2008), the clear sky irradiance, I_{cs}, is suggested to be the upper limit. In this thesis the more liberal $1.1I_{cs}$ is used as in Moradi (2009), Tang et al. (2010) and Journée and Bertrand (2011). An explanation of the increased limit is that the test is applied on hourly data, while conventional use is daily data. As described in subsection 3.2.2 on page 26, the clear sky model used in this thesis is libRadtran, (Mayer and Kylling, 2005), and the configuration files are presented in Appendix A on page 81. The test used in this thesis is called U_2 and is suggested by Journée and Bertrand (2011). It is shown in Eqs. 4.1.4 and 4.1.5.

$$I_{\text{global}} \leq 1.1I_{cs} \text{ if } \theta < 88^\circ \tag{4.1.4}$$

$$I_{\text{global}} \leq 2I_{cs} \text{ if } \theta \geq 88 \tag{4.1.5}$$

Where I_{global} is the measured global horizontal irradiance, I_{cs} is the modeled clear sky irradiance and θ is the solar zenith angle.

4.1.3 Lower limit tests

Modeled extraterrestrial irradiance is also useful for estimating a lower limit of solar irradiance. Despite at heavily overcast conditions dense clouds may reflect most of the solar radiation, at least 3% eventually reaches the surface of the earth (Chen, 2011). Thus, for a given day, as presented by Journée and Bertrand (2011), the daily mean of the ratio between measured global horizontal irradiance and modeled extraterrestrial irradiance should never be less than 3%. In this thesis the test is called L_1 and is shown in Eq. 4.1.6.

$$\mu\left(\frac{I_{\text{global}}}{I_{\text{E}}}\right) \geq 0.03 \tag{4.1.6}$$

Where I_{global} is the measured global horizontal irradiance, I_{E} is the modeled extraterrestrial irradiance and μ is the daily mean. The test could arguably be executed with hourly data, however, it would probably result in a drastic increase of rejected data for lower solar elevation. As a substitute, Journée and Bertrand (2011) proposed a test for sub-hourly data that is more restrictive the lower the solar zenith angle. In this thesis the test is applied on hourly data, called L_2 and is shown in Eq. 4.1.7.

$$I_{\text{global}} \geq 10^{-4}(80-\theta)I_{\text{E}} \text{ if } \theta \leq 80^{\circ} \tag{4.1.7}$$

Where I_{global} is the measured global horizontal irradiance, I_{E} is the modeled extraterrestrial irradiance and θ is the solar zenith angle. The test is only applied to solar zenith angles less than 80°.

4.1.4 Difference in time steps

The amount of solar radiation reaching the surface of the earth varies throughout the day. Sky conditions may change suddenly, however, the absolute change of the ratio $\frac{I_{\text{global}}}{I_{\text{E}}}$ between two time steps should not exceed an upper limit. In Journée and Bertrand (2011), that limit is set to 0.75 and it is also chosen in this thesis. Since the time resolution in this thesis is hourly data, the limit for largest solar zenith angle in this thesis is chosen to be 80°. The

test is called Difference and is given in Eq. 4.1.8.

$$\left| \frac{I_{\text{global}}(t)}{I_{\text{E}}(t)} - \frac{I_{\text{global}}(t-1)}{I_{\text{E}}(t-1)} \right| < 0.75 \text{ if } \theta < 80^{\circ} \tag{4.1.8}$$

Where I_{global} is the measured global horizontal irradiance, I_{E} is the modeled extraterrestrial irradiance and θ is the solar zenith angle.

4.1.5 Daily consistency

The ratio between measured global horizontal irradiance and modeled extraterrestrial irradiance varies throughout the day. However, if it deviates too much, there might be a temporary erroneous value in part of the day. In addition, if there is a very low deviation, there might be a constant value output from the system. If the standard deviation of the daily values are less than or larger than certain limits, the day should be flagged as suspicious and marked for visual control. Journée and Bertrand (2011) proposed a persistence test which is adopted in this thesis, however, due to the low solar heights in Norway, the upper limit is set to 0.80 and the lower limit is set to $\frac{1}{16}$. The test is called Consistency and is shown in Eq. 4.1.9.

$$\frac{1}{16}\mu\left(\frac{I_{\text{global}}}{I_{\text{E}}}\right) \leq \sigma\left(\frac{I_{\text{global}}}{I_{\text{E}}}\right) \leq 0.80 \tag{4.1.9}$$

Where I_{global} is the measured global horizontal irradiance, I_{E} is the modeled extraterrestrial irradiance, θ is the solar zenith angle, μ is the mean of the data points from sunrise to sunset and σ is the standard deviation of the data points from sunrise to sunset. A summary of the automatic quality control tests are shown in Table 4.1.1.

Table 4.1.1 – An overview of the automatic quality control tests used in this thesis. If a data point fails the quality requirement, it is either flagged as erroneous or reviewed in the visual control.

Name	Quality requirement	Quality procedure
Offset	$I_{\text{global}} \geq -12\frac{\text{W}}{\text{m}^2}$ $I_{\text{global}} < 6\frac{\text{W}}{\text{m}^2}$ if $\theta > 93°$	The flagged data are reviewed in the visual control
U1	$I_{\text{global}} < I_{\text{E}}$	Data points are flagged as erroneous
U2	$I_{\text{global}} \leq 1.1 I_{cs}$ if $\theta < 88°$ $I_{\text{global}} \leq 2 I_{cs}$ if $\theta \geq 88°$	Data points are flagged as erroneous
L1	$\mu\left(\frac{I_{\text{global}}}{I_{\text{E}}}\right) \geq 0.03$	All data from sunrise to sunset are flagged as erroneous
L2	$I_{\text{global}} \geq 10^{-4}(80-\theta) I_{\text{E}}$ if $\theta \leq 80°$	Data points are flagged as erroneous
Difference	$\left\lvert \frac{I_{\text{global}}(t)}{I_{\text{E}}(t)} - \frac{I_{\text{global}}(t-1)}{I_{\text{E}}(t-1)} \right\rvert < 0.75$ if $\theta < 80°$	Data points are flagged as erroneous
Consistency	$\frac{1}{16}\mu\left(\frac{I_{\text{global}}}{I_{\text{E}}}\right) \leq \sigma\left(\frac{I_{\text{global}}}{I_{\text{E}}}\right) \leq 0.80$	All data from sunrise to sunset are reviewed in the visual control

4.2 Visual quality control

In addition to the automatic quality control procedure, a visual control of the time series is desirable. Factors such as temporal change, changes related to the operation of the pyranometers and periods of inconsistent data may be difficult for the automatic quality control to define. Ideally, visual quality control should be limited, since it is time consuming. However, for a thorough analysis as is performed in this thesis, it is required. The visual control is carried out in three steps. The first step is to compare measured data from the four stations with each other to evaluate whether there are significant discrepancies between the stations over the time series. A method is to compare moving averages. The second step is to find out whether there are any temporal changes in the data such as change in sensitivity or after a pyranometer has been replaced. The last step is to evaluate the Offset

and Consistency flagged data, and performing a visual control of each day to determine whether the data should be flagged as erroneous.

4.3 Calculation of monthly averages

In this thesis, average monthly values are estimated from long time series of hourly data. The procedure is very similar to the one used for CM-SAF PVGIS (Huld et al., 2012) and thus a similar notation is used. The first step is to average each hour of the year for each year as in Eq.

$$G_{hdm} = \frac{1}{N} \sum_{y=1}^{N} G_{hdmy} \tag{4.3.1}$$

Where N is the number of years of the time series where there is not a missing value or a value is flagged as erroneous for the given hour and G_{hdmy} is the hourly value in a year. Different from Huld et al. (2012) is that there are certain hours that have either missing value or data is erroneous for all years. In this unusual case, the value is linearly interpolated between the two nearest hourly values. In this thesis, the hourly values for February 29th are excluded in the calculations of monthly averages. For easier handling the length of February, the average daily values for each month are found.

$$G_m = \frac{1}{D_m} \sum_{h=1}^{24} \sum_{d=1}^{D_m} G_{hdm} \tag{4.3.2}$$

Where D_m is the number of days in the month. To find total monthly global horizontal irradiation, the value is scaled with average number of days in each month. Yearly average is the summation of all months.

4.3.1 Use of Typical Meteorological Year

Typical Meteorological Year (TMY) is often used for PV system simulations (Honsberg and Bowden, 2014). A TMY is a year consisting of hourly solar irradiation and other parameters that describes the typical conditions at a location. Different from hourly averages, TMY attempts to remove the variability that atypical years provide. The most common method to estimate TMY is to utilize the Filkenstein-Schafer statistical method, which involves estimating cumulative distribution functions (CDF) for each month and for several parameters with each parameter having a weighing factor (Chan et al., 2006; Petrakis et al., 1998). The month with the CDF closest to the average of a sample is chosen to be the month in the TMY. Since the TMY calculation often require either measured direct or diffuse horizontal

irradiation which is not measured at Bioforsk stations, TMY have not been estimated in this thesis.

4.4 The Python quality control code

The automatic quality control procedures have been applied with the use of the Python programming language, Python (Python Software Foundation. Python Language Reference, version 2.7. Available at http://www.python.org), and the SciPy ecosystem Jones et al. (2001–) with special use of the Python based statistical package, Pandas (McKinney, 2010). The code for the automatic quality control procedure is described in Appendix B with an example file for Ås. The code is also available on https://github.com/sigbjorngrini/solqc. The code for all plots in this thesis is also available on request.

Chapter 5

Results and discussions

This chapter is divided into four main parts; the quality control procedure and selection of quality controlled data, an evaluation of the impact of the quality control on total measured value, the comparison of quality controlled data with existing solar irradiation databases and an overall discussion.

5.1 Automatic quality control

The results of the automatic quality control for the four stations are shown in Fig. 5.1.1 as the percentage of flagged data points for each test at each station. Missing values, U2 and Consistency are the tests which produce the most flags. The Offset flags occur mainly for Tomb, while L1, L2 and the Difference flag have a tiny impact. A good sign is that for every test, the percentage of flagged data is less than 4 %.

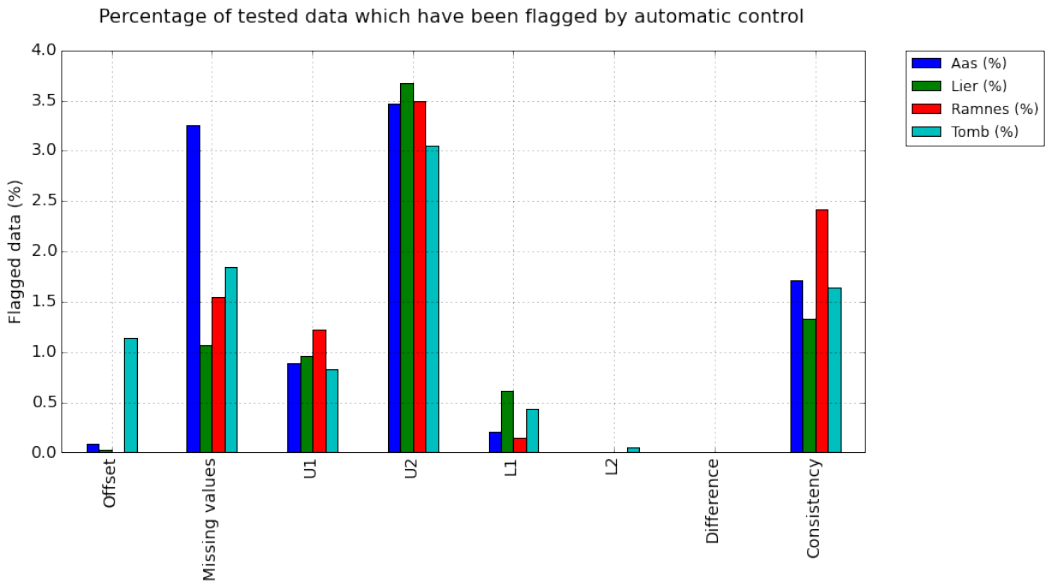


Figure 5.1.1 – The percentage of tested data which have been flagged by the automatic control procedure. Offset flag is tested on all data point, while the rest is only tested on data from sunrise to sunset. U1 and U2 are the upper limit tests, while L1 and L2 are the lower limit tests. The Offset and Consistency flags are marked for visual control and the rest are flagged as erroneous.

Since each data point may have several flags, it is interesting to evaluate how many data points in the time series that are flagged as erroneous and for visual control according to the procedure described in section 4.1 on page 29. An overview is shown in Table 5.1.1.

Table 5.1.1 – The percentage of data from sunset to sunrise that are flagged as erroneous and the percentage of all data points that are flagged for visual control.

	Aas (%)	Lier (%)	Ramnes (%)	Tomb (%)
Erroneous	6.94	5.35	5.23	5.34
Visual control	0.97	0.72	1.24	1.98

The percentage of erroneous data ranges from 5.23 % to 6.94 %, which is higher compared to the result in Shi et al. (2008). However, this thesis applies more tests and for hourly solar irradiation compared to daily values. The measured solar irradiation plotted against solar elevation for all stations are shown in Fig. 5.1.2.

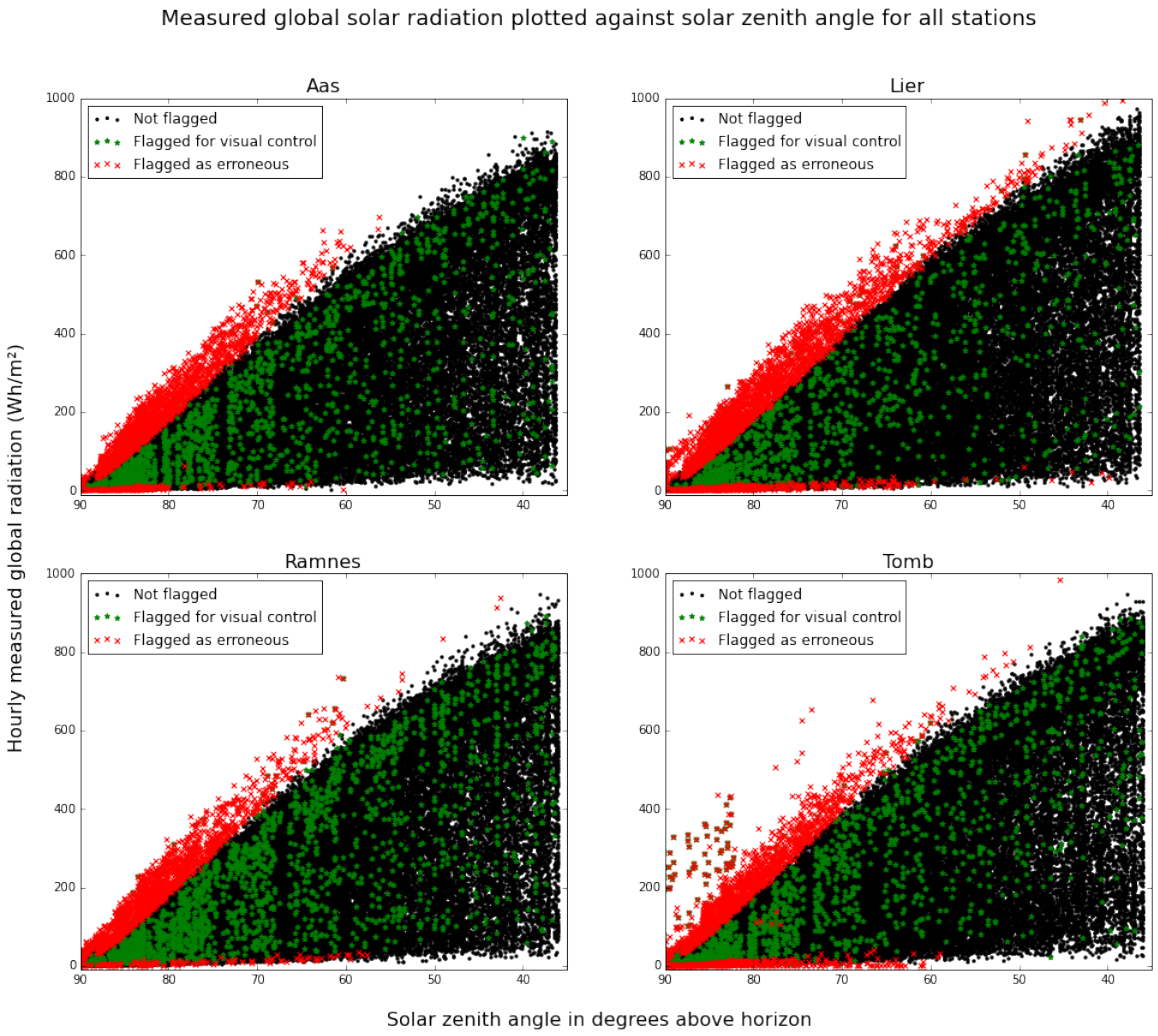


Figure 5.1.2 – Measured solar irradiation plotted against solar zenith angle for all stations. Black dot indicates that the data point is not flagged, green star indicates that the data point is flagged for visual control and red cross indicates that the data point is flagged as erroneous. The missing value flags are not shown.

For all stations exist several data points which exceed the U1 and U2 tests and are flagged as erroneous. In addition, some data points fail the L1 and L2 tests and are also flagged as erroneous. There are some distinctions between the stations, though. Lier has several data points flagged as erroneous for solar zenith angles lower than 50° and Tomb have some very high values at high solar zenith angles that are flagged as erroneous. Ås is the station which appears to have flags limited to high solar zenith angles and at close to the boundary of the tests, which are the type of errors related to uncertainty of the measurements. This appears to be consistent with the fact that Ås has the most frequent maintenance, as described in subsection 3.1.1 on page 21. Overall, it appears that the upper and lower limit flags occur consistently for all stations. Thus, it is interesting to evaluate how the flags are distributed

for each year of the time series. This distribution is shown in Fig. 5.1.3. It is important to note that the vertical axis indicates a cumulative amount. This means that for Tomb in 1993, 6% of all data points tested for Offset test were flagged, 3 % of all data from sunrise to sunset were missing values and 1% of all data from sunrise to sunset failed the U1 test.

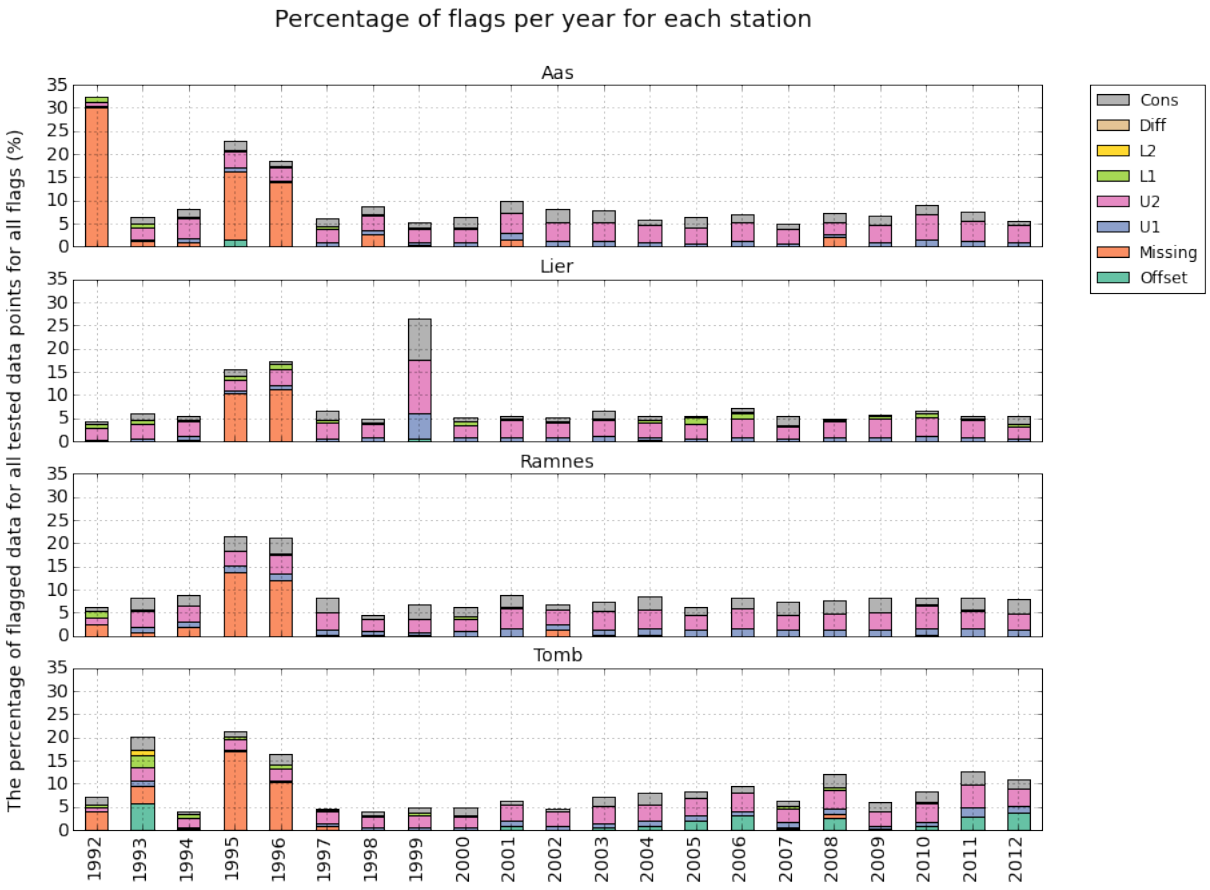


Figure 5.1.3 – The percentage of all tested data points that was flagged for all flags for each year for all stations. The vertical axis indicates a cumulative amount. This means that for Tomb in 1993, 6 % of all data points tested for Offset test were flagged, 3 % of all data from sunrise to sunset were missing values and 1 % of all data from sunrise to sunset failed the U1 test. The height of the bars does not indicate the total of all data points, since each data point can have several flags.

Most notable is that the years 1995 and 1996 have many missing values for all stations. The missing values occurred due to a data logging issue which happened almost every Tuesday and Wednesday. A figure displaying missing values in 1995 and 1996 is shown in Fig. 5.1.4.

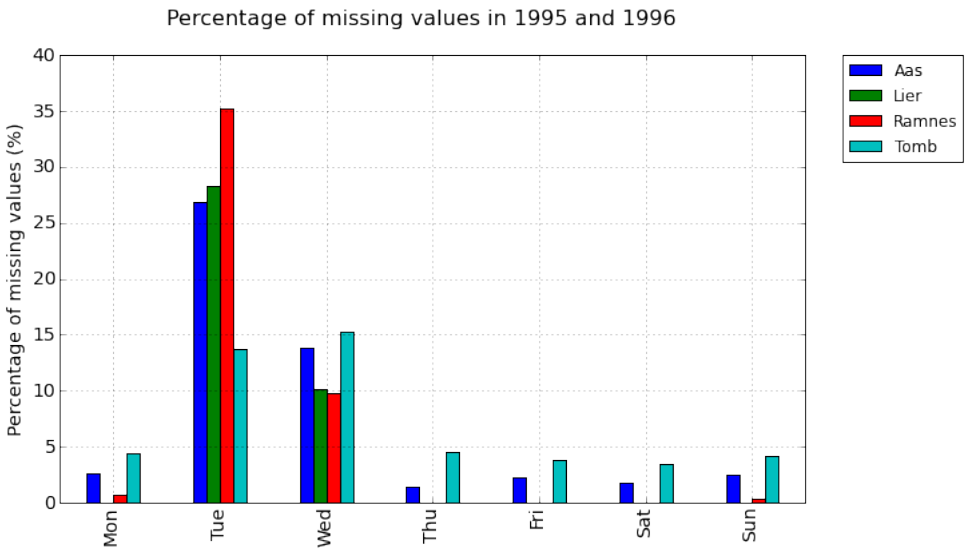


Figure 5.1.4 – The percentage of missing values per weekday in 1995 and 1996. Displays that there was a logging issue for Tuesdays and Wednesdays in this time period.

Since this was a frequent trend, there is not a longer period without missing values in 1995 and 1996. Other notable periods with many missing values are for Ås in 1992 and Ramnes and Tomb in 1992 to 1994. However, in the 16 year period from 1997 to 2012, there is only a small amount of missing values for all four stations. The Offset flag only occurs for Tomb, however most are in the latter part of the time series, which might pose a problem. This issue is explored in further detail in subsection 5.2.4 on page 46. Another issue is for Lier in 1999, where several flags have occurred to a much greater extent compared to other years and stations. This issue is explored in further detail in subsection 5.2.5 on page 47. The U2 flag and the Consistency flag appear to occur a certain amount each year. As shown in Fig. 5.1.2 on page 39, the upper limit flags appear mostly for high solar zenith angles, where the uncertainty in the measurements is higher and thus it is expected that some values exceed the upper limit. The Consistency flag aims to find days where there are suspicious values and those days are marked for visual control. The fact that there are suspicious days every year is not unexpected. In addition to the distribution of flags per year, the distribution of flags per month for the four stations is shown in Fig. 5.1.5. In this figure the vertical axis also indicates a cumulative amount.

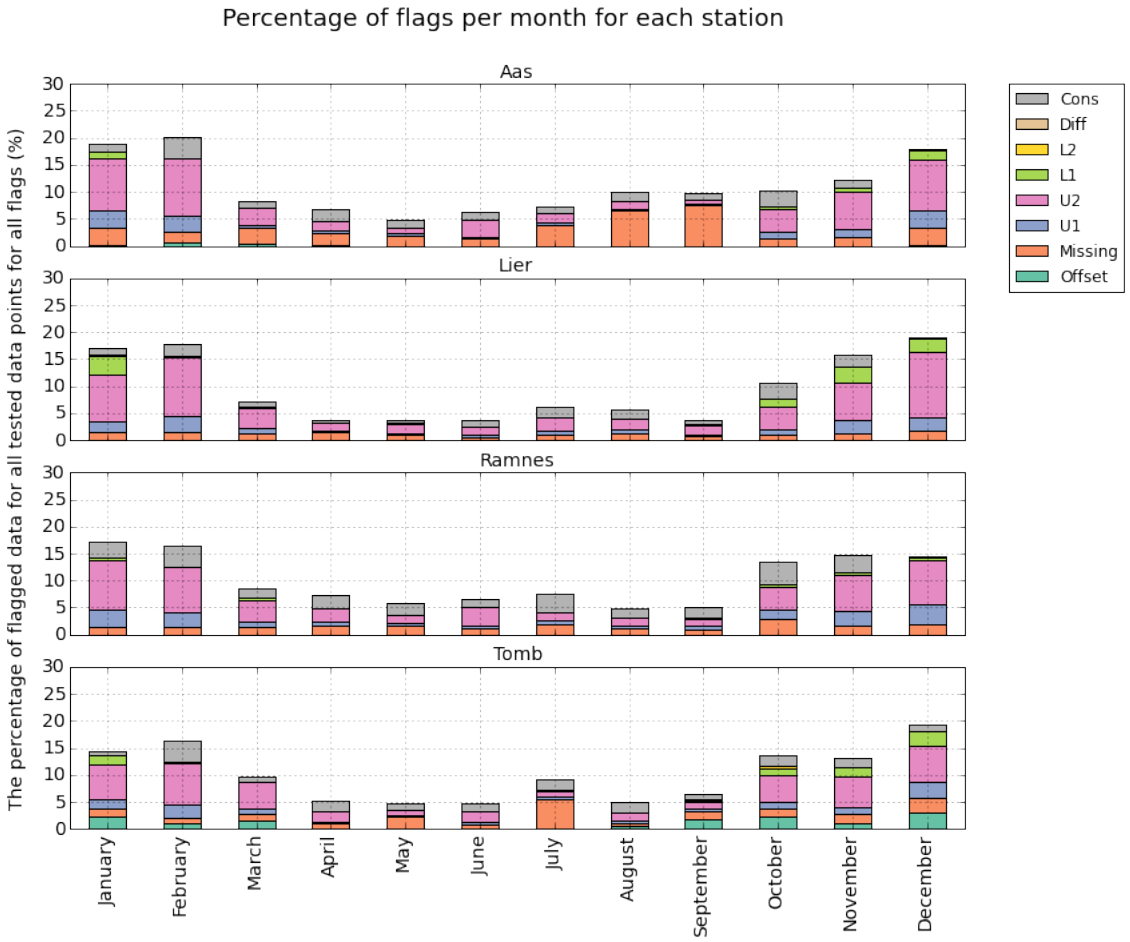


Figure 5.1.5 – The percentage of all tested data points that were flagged for all flags for each month for all years for all stations. The vertical axis indicates a cumulative amount. This means that for Tomb in February for all years, 5% of all data from sunrise to sunset were missing values and 3\:% of all data from sunrise to sunset failed the Consistency test. The height of the bars does not indicate the total of all data points, since each data point can have several flags.

This illustrates a significant distinction between which type of flag occurs in the summer and the winter. There are more flagged data points in the winter compared to the summer. The L1, U1 and the Offset flags occur almost exclusively from September to April. The U2 flag occurs most in the winter, but is also significant in the summer months. The Consistency flag and missing values are fairly evenly distributed throughout the year. The number missing values has a peak in July for many stations, which is probably due to the fact that errors are not fixed as quickly during summer vacation.

5.2 Visual control

In addition to an automatic quality control, a visual control procedure is necessary to ensure that the erroneous data have been filtered by the automatic control and to evaluate whether the data flagged for visual control is erroneous or not. In addition, temporal changes in measured value such as sensitivity change and operational issues need to be investigated.

5.2.1 Comparison between the four stations

Although the local climate differs for the four stations, over time, there is often a trend of which location has the most solar irradiation. Therefore, it is relevant to examine the measured total solar irradiation for the four stations to observe if the trend changes over time or in certain years. A procedure to analyze meteorological data is to estimate the trend of the time series is by the use of moving average (Takezawa, 2005). In this thesis a moving average of 10 days was used, since it is in the order of a synoptic timescale. Even though there were some differences between the stations in the 10 day moving average in the time period from 1992 to 2012, the differences were not consistent enough to draw any conclusions. Small differences are often due to contrasts in the local weather.

5.2.2 Replacement of pyranometers

During the time period from 1992 to 2012, the pyranometers for three stations have been replaced. The reason for a change of pyranometer could be that the pyranometer stopped working, the pyranometer dome shattered or poor maintenance (Ruud Hansen, 2015). As shown in Table 3.1.2 on page 25, Ås changed its pyranometer on May 9th 2003, Ramnes changed its pyranometer on November 7th 2003 and Lier changed its pyranometer on April 20th 2010. It is important to perform a visual control and compare the data before and after the replacement to ensure that there was no offset change. A method is to observe the average maximum daily solar irradiation for each month for the time period before and after the replacement. This is shown in figures 5.2.1, 5.2.2 and 5.2.3 for Ås, Lier and Ramnes respectively.

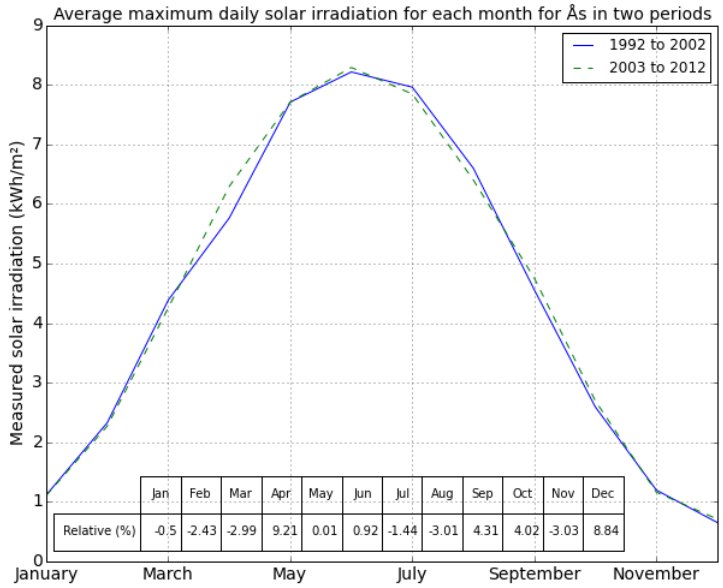


Figure 5.2.1 – The average maximum daily values for each month for Ås for the time periods 1992 to 2002 and 2003 to 2012. The table displays the relative change for each month with the older time period as reference.

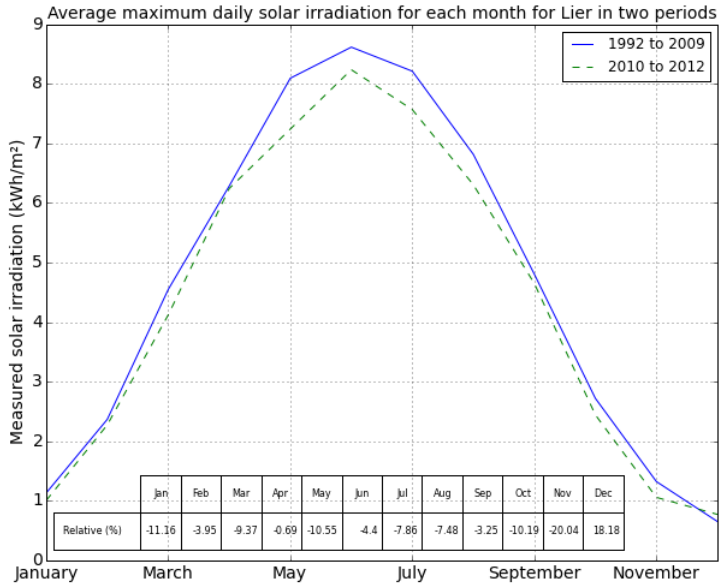


Figure 5.2.2 – The average maximum daily values for each month for Lier for the time periods 1992 to 2009 and 2010 to 2012. The table displays the relative change for each month with the older time period as reference.

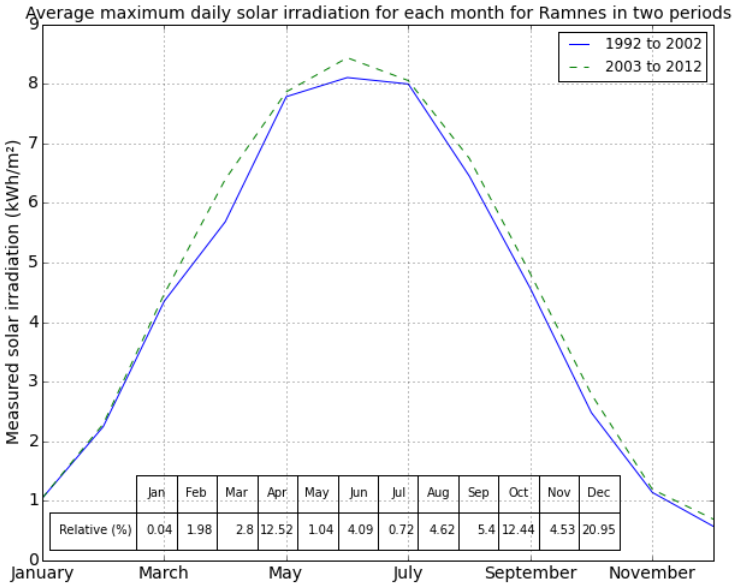


Figure 5.2.3 – The average maximum daily values for each month for Ramnes for the time periods 1992 to 2002 and 2003 to 2012. The table displays the relative change for each month with the older time period as reference.

The average maximum values of daily solar irradiation for Lier in the summer months before the replacement are consistently higher compared to after the replacement. The maximum solar irradiation before 2010 is higher than the values for Ås and Ramnes. This could indicate that the solar irradiation measured before 2010 was excessively high. This observation contributes to an uncertainty to the measurements at Lier. Ramnes also has a difference in average maximum daily values for each month before and after the replacement of its pyranometer in 2003. The difference, however, is not as consistent as for Lier. Thus, a conclusion cannot be made whether the new pyranometer measure higher solar irradiation compared to the previous. Ås appears to have nearly equal average maximum daily solar irradiation before and after the replacement of its pyranometer in 2003.

5.2.3 Change in sensitivity

As explained in subsection 2.4.3 on page 16, the sensitivity of a pyranometer may change over time. The Kipp & Zonen CM11 pyranometer is expected to have a less than 0.5 % change in sensitivity per year. 0.5 % change is for maximum sun exposure, and consequently it should be less for the cloudier Norwegian climate. Since there is too much variation in weather between the four locations, it is difficult to discover the exact value of this change.

An effort was made to evaluate clear sky days for the same station over several years, but no clear correlation was found. There has not been performed any recalibration of the pyranometers from the four stations evaluated in this thesis. As an alternative, an two certificates for another pyranometer were shown by Ruud Hansen (2015) (are shown in Appendix D), where the change of sensitivity for the pyranometer in 10 years of use was 0.23 %.

5.2.4 Offset flag issues for Tomb

As shown in section 5.1 on page 37, Tomb has many data points which have an Offset flag. With a quick analysis of the data, it is obvious that there are certain parts of the year which have most of the flagged data. Especially December and January are the two months where station at Tomb suddenly starts to measure high values during the night. A yearly distribution is shown in Fig. 5.2.4. Following this conclusion, if there is an offset issue, Tomb should thus have higher values during the day and over time, have more U2 flags compared to the other stations.

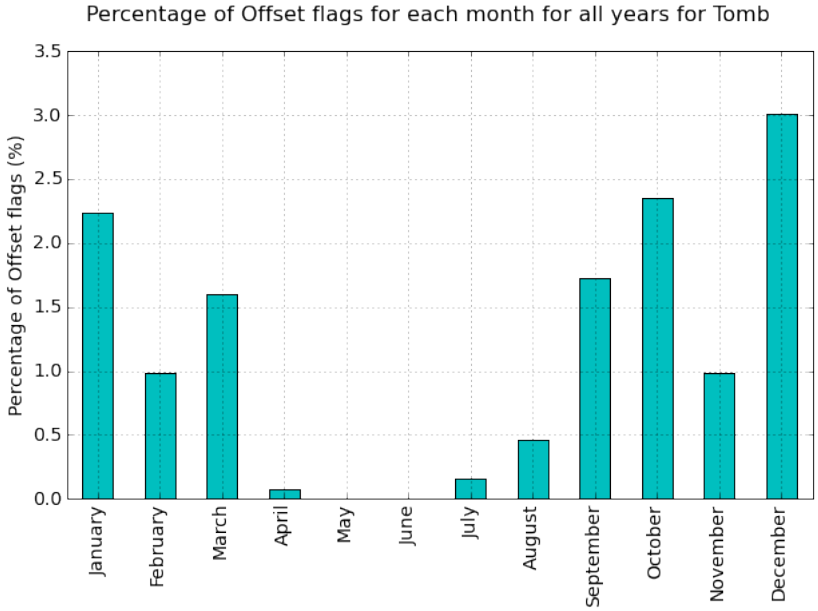


Figure 5.2.4 – The percentage of Offset flags per month for all years for Tomb for all data from 1992 to 2012.

This is surprisingly not the case. As shown in Fig. 5.2.5, the opposite appears to be true. The only months where Tomb has the most U2 flags compared to the other stations are March and October. For the other months Tomb has lower amounts of these errors. While

there appears to be a positive offset of the measured values during the night, this cannot be verified for measured values from sunrise to sunset.

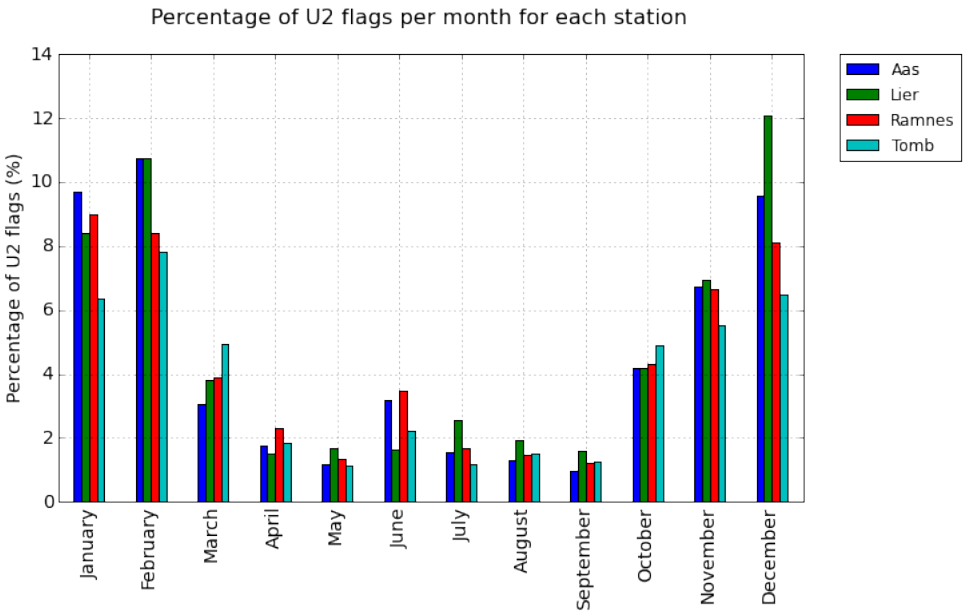


Figure 5.2.5 – The percentage of U2 flags from sunrise to sunset averaged per month for all years for each station. Displays that Tomb has the least percentage of U2 flags in December and January.

5.2.5 High amounts of flags for Lier in 1999

As shown in section 5.1, Lier has a higher amount of flags in the year 1999 compared to other years and other stations. The flags consist of mostly Consistency and U2 flags. The distribution of flags across the year of 1999 for Lier is shown in Fig. 5.2.6. The abnormal amount of flags appears to begin in May. A method to analyze solar irradiation data, which gives a clear image of what is happening, is to look at days with a clear sky. In Fig. 5.2.7a), 9 days for all four stations are shown. On days with a clear sky for all stations it is obvious that Lier measures too high amounts of solar irradiation in the middle of the day. Taking a closer look at July 11, 1999 in Fig. 5.2.7b), there appears to also be a time displacement for Lier's values, which is why the Consistency test fails. The issues appear to be solved towards the end of the year.

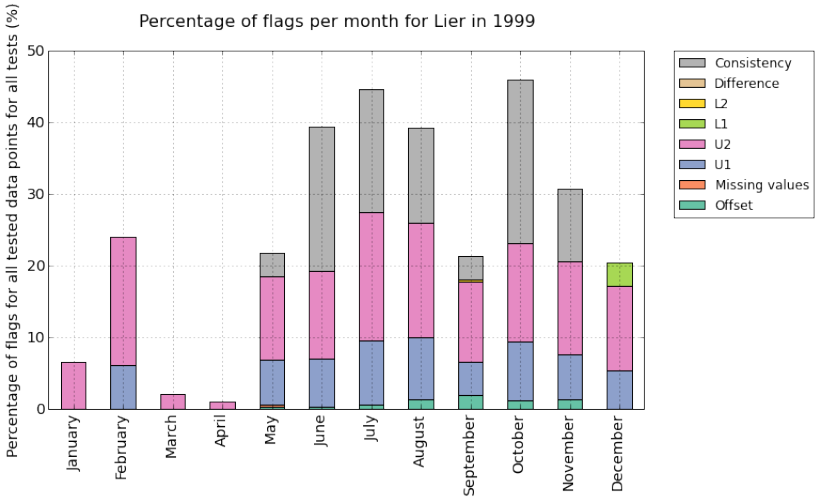


Figure 5.2.6 – The percentage of all flags per month for Lier in 1999. The vertical axis indicates a cumulative amount. This means that 6 % of all data points tested for the U1 test in December were flagged, 12 % of all tested for U2 were flagged and 3 % of all data tested for L1 test were flagged. The height of the bars does not indicate the total of all data points, since each data point can have more than one flag.

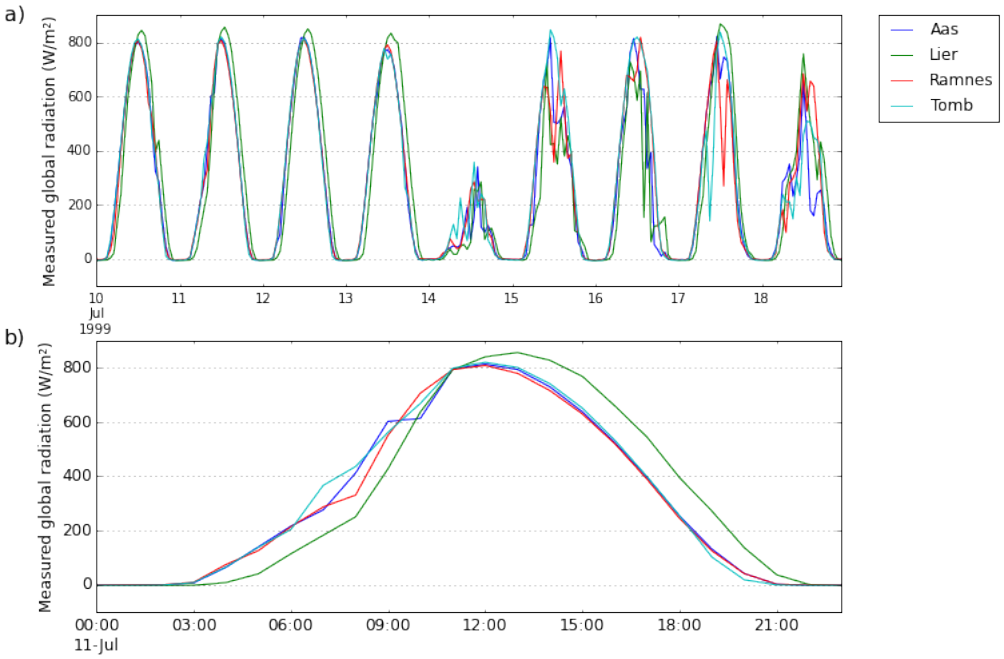


Figure 5.2.7 – a) The measured solar irradiation for all stations for 9 days in July 1999. The days from July 10th to July 13th are clear sky days and the days from July 15th to July 18th are cloudier days. b) The measured solar irradiation for all stations in July 11, 1999 for all stations. This displays the time displacement issue for Lier which took place in 1999.

5.2.6 Consideration of Consistency flags

The Consistency flag is designed to capture days where there are constant values or days with an abnormal change in the ratio between measured solar irradiation and modeled extraterrestrial irradiation. Since this test may also flag days where the change is normal and due to natural changes in weather, it is important to have a visual control of the flagged days. Table 5.2.1 lists the days that were considered erroneous after the visual control. The visual controls were carried out by observing each individual day and subjectively choose which days that appeared erroneous. The remaining days were considered valid and likely caused by natural high variation during the day. Lier is the only station with a problem of time displacement. Apart from the problem in 1999 discussed in subsection 5.2.5, the Consistency flag also discovered two days in 2007 where the time displacement issue occurred. The reason is unknown. As discussed in subsection 5.1 on page 37, Tomb has many data points which measure substantial amount of solar irradiation at high solar zenith angle. Several of these occur in the end of 2007 and 2011 and the set of days are shown in Fig. 5.2.8. This could be due to electric fields in the vicinity of cables or mechanical loading of cables (Younes et al., 2005), which are both operational errors. Most of these data points are already flagged by U1 and U2 flags, however, all data points in both periods are flagged as erroneous.

Table 5.2.1 – Days that have been flagged as erroneous by the visual control of Consistency flags. Ramnes did not have any days that were considered erroneous by the visual control.

Station	Date(s)	Reason
Ås	1995/05/25	Too high values in part of the day
Lier	2007/01/08 - 2007/01/09	Time displacement
Lier	1999/01/01 - 1999/12/31	Time displacement and too high values
Tomb	2007/12/30 - 2008/01/03	High values, probably operational error
Tomb	2011/12/10 - 2011/12/15	High values, probably operational error

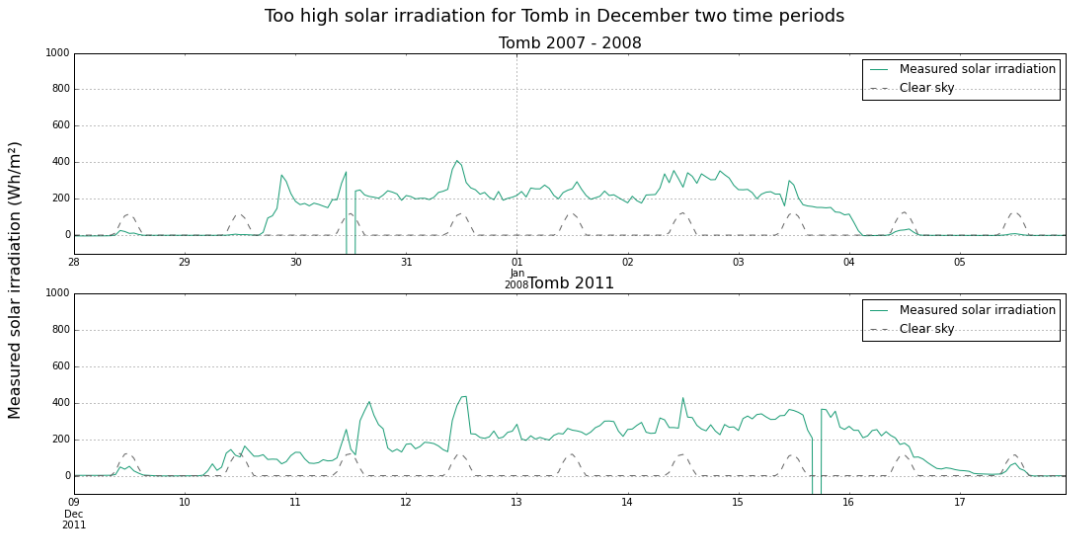


Figure 5.2.8 – Too high solar irradiation for Tomb two time periods. The dark green lines indicate measured solar irradiation and the dotted gray lines indicate modeled clear sky.

5.2.7 Summary of visual quality control

The visual quality control is an additional method to automatic quality control which purpose is to discover erroneous data points that was missed by the automatic quality control. In addition, it serves as quality check of the automatic quality control. After automatic and visual quality control has been undertaken, the percentage of erroneous data for each station is as shown in Table 5.2.2.

Table 5.2.2 – The percentage of erroneous data from sunrise to sunset for each station after the automatic control and the visual control have been applied.

	Aas (%)	Lier (%)	Ramnes (%)	Tomb (%)
Erroneous	6.96	9.57	5.23	5.35

The percentage of erroneous data ranges from 5.23 % for Ramnes to 9.57 % for Lier. The solar irradiation measured at Lier is also less trustworthy due to the difference in the level of measured solar irradiation discussed in subsection 5.2.2. Despite the problems at Lier, it is interesting to keep the quality controlled value for comparison with existing solar irradiation databases.

5.3 Impact of erroneous data on total measured solar irradiation

In many cases a quality control analysis of measured data is unavailable. Therefore, it is interesting to analyze how large of an impact the erroneous data have on the total measured global solar irradiation for the time series. In addition, it is a good indication of the overall trustworthiness of the measured data compared to quality controlled data. In Table 5.3.1, the relative change between total measured solar irradiation for all years and its replacement alternatives is shown.

Table 5.3.1 – The relative change of solar irradiation with the use of a replacement method for values flagged by upper limit tests instead of measured values for all years. Negative values indicate that the replaced values are lower than measured values. *Clear sky* and *set to zero* are used as reference.

	Aas (%)	Lier (%)	Ramnes (%)	Tomb (%)
Replace method for upper values				
Clear sky	-0.49	-0.69	-0.47	-0.49
Mean value	-1.21	-1.72	-1.05	-1.16
Set to zero	-1.91	-2.90	-1.75	-1.81

The mean value is the average value for that hour for all years in the time series that are not flagged as erroneous. In this thesis, only the mean value replacement option is explored, since that implies that flagged values have no impact on the calculation of average years. As shown in Table 5.3.1, the relative change varies between $-2.90\,\%$ and $-0.47\,\%$. The *set to zero* option is a good indicator of how large of an impact the upper limit flagged values could possibly have on the total measured value. The relative changes between measured solar irradiation and the replacement options for the lower limit flags are shown in Table 5.3.2.

Table 5.3.2 – The relative change of solar irradiation with the use of a replacement method for values flagged by lower limit tests instead of measured values. Positive values indicate that the replaced values are higher than measured values. *Clear sky* and *Lowest possible accepted* are used as reference.

	Aas (%)	Lier (%)	Ramnes (%)	Tomb (%)
Replace method for lower values				
Clear sky	0.12	0.40	0.14	0.32
Mean value	0.05	0.17	0.06	0.14
Lowest possible accepted	0.00	0.00	0.00	0.01

As expected from the results in section 5.1 on page 37, the replacement for lower limit tests have less impact on total measured solar irradiation, since it has fewer flagged data points and lower values. The relative change varies between 0.00 % and 0.40 %, both for Lier. In Table 5.3.3 are the relative changes between replacing missing values with different alternatives and keeping them at zero shown.

Table 5.3.3 – The relative change of solar irradiation with the use of a replacement method for missing values. *Clear sky* and *Top of Atmosphere accepted* are used as reference. Since there was a lot of missing values before 1997 compared to after 1997, the time series from 1997 are also displayed.

	Aas (%)	Lier (%)	Ramnes (%)	Tomb (%)
Replacement for missing values				
Mean value	3.47	1.02	1.54	2.03
Clear sky	5.88	1.70	2.60	3.33
Top of Atmosphere	7.53	2.20	3.34	4.22
Mean value (from 1997)	0.45	0.04	0.15	0.13
Clear sky (from 1997)	0.73	0.07	0.24	0.23
Top of Atmosphere (from 1997)	0.94	0.09	0.30	0.30

For the whole time series the relative change varies between 1.02 % and 7.53 %, however, if only the values from 1997 to 2012 are accounted for, the relative change varies between 0.04 % and 0.94 %. For the Difference flag, there are not enough flags for any replacement to have an impact on the total measured solar irradiation. For the Consistency flag, as shown in Table 5.2.1, there are only a few distinct periods which are flagged as erroneous after visual control. An overview of the relative changes between measured total solar irradiation from 1997 to 2012 and the total value when erroneous values are replaced with mean values are

shown in Table 5.3.4. The period from 1997 to 2012 has been chosen since the years before 1997 have much more missing values compared to the rest of the time series. Lier has the whole year of 1999 flagged and thus its relative change is calculated from 2000 to 2012.

Table 5.3.4 – The relative change of total solar irradiation when erroneous values are replaced with mean values. The time period is from 1997 to 2012, except for Lier which the time period is from 2000 to 2012. Negative values indicate that the replaced values are lower than measured values.

	Aas (%)	Lier (from 2000) (%)	Ramnes (%)	Tomb (%)
Mean value	-0.78	-1.26	-0.9	-1.09

Table 5.3.4 shows that the relative change for the four stations between measured total solar irradiation and the total solar irradiation when values flagged as erroneous are replaced with a mean value ranges from $-1.26\,\%$ to $-0.78\,\%$. In other words, measured values could be expected to overestimate the total solar irradiation with roughly $1\,\%$. Given that the uncertainty of each daily measurement is $3\,\%$, the difference is considered to be not significant.

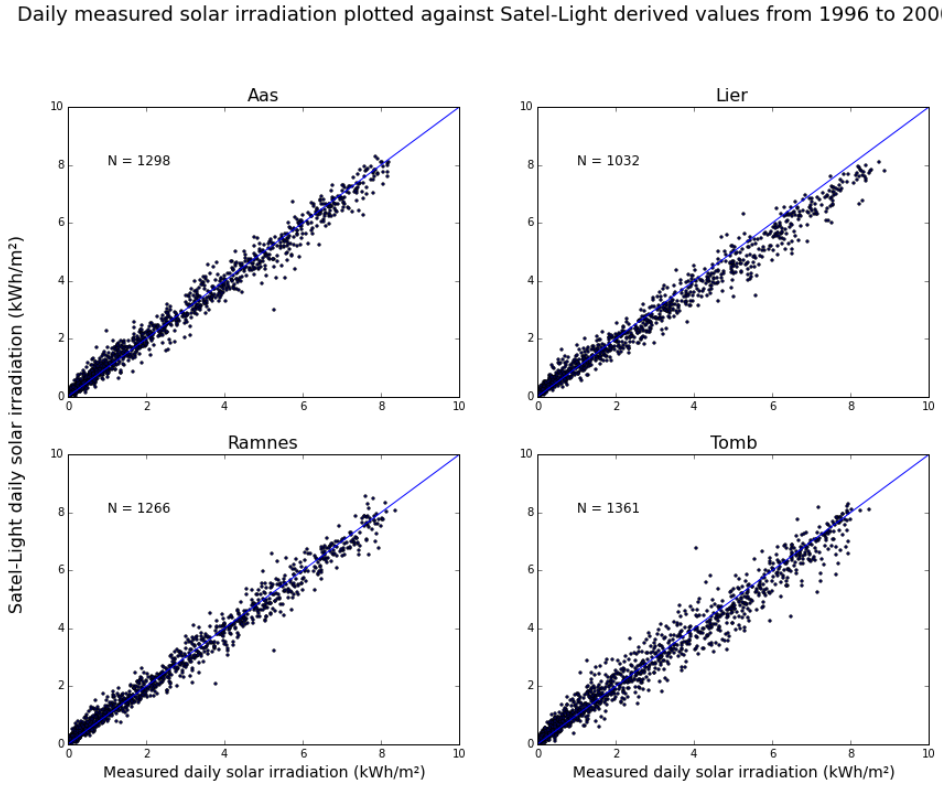
5.4 Comparison of quality controlled solar irradiation with existing databases

A set of quality controlled data has been achieved through sections 5.1 and 5.2. The data points which have been flagged as erroneous are not included in this comparison between quality controlled measured solar irradiation and existing solar irradiation databases. There are several databases readily available either for free on the web or through PVsyst (Mermoud, 2012). The databases acquired for this thesis are Meteonorm 7, Classic PVGIS, CM-SAF PVGIS, Satel-Light, NASA SSE 6.0 and WRF. These are described in section 3.3 on page 26.

5.4.1 Comparison with time series

Some of the databases offer time series data of solar irradiation in a given time period. It is thus interesting to compare the difference between quality controlled measured solar irradiation and the solar irradiation from existing databases for each day. This has been done for Satel-Light in the time period from 1996 to 2000, for NASA SSE in the time period from 1992 to 2005 and for WRF in the time period from 1992 to 2012. The comparisons are visualized with daily solar irradiation from each database plotted against the corresponding

quality controlled measured values. In addition, the mean bias deviation (MBD), mean absolute error (MAE) and root-mean-square deviation (RMSD) are added to the figures. These statistical methods are explained in section 2.6 on page 17. For Satel-Light, the comparisons with all stations are shown in Fig. 5.4.1.

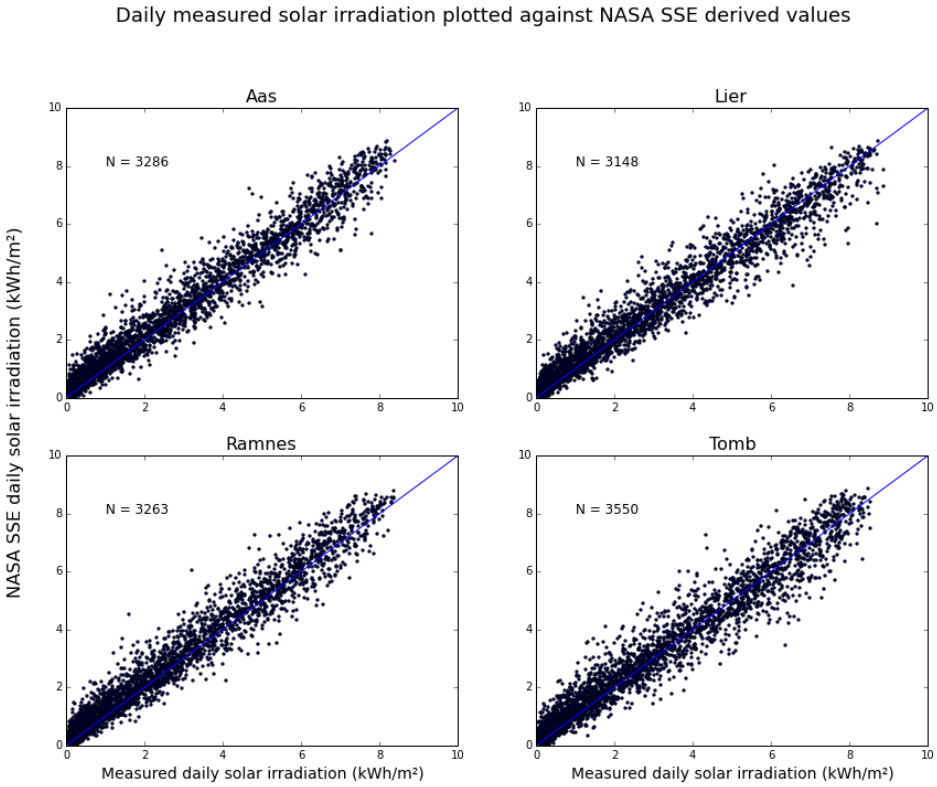


	Aas	Lier	Ramnes	Tomb
MBD (%)	0.12	-5.74	0.08	-0.41
MAE (%)	8.62	11.42	8.56	11.4
RMSD (%)	12.22	15.89	12.14	16.55

Figure 5.4.1 – Daily measured solar irradiation plotted against the corresponding daily values derived from the Satel-Light website for all stations in the period from 1996 to 2000. Days where at least one value has been flagged during quality control have been ignored. Statistics are added to the figure where, MBD is mean bias deviation, MAE is mean absolute error and RMSD is root-mean-square deviation.

The MBD is close to zero for all stations except Lier. The Satel-Light database shows lower daily solar irradiation compared to measured values for Lier at lower solar zenith angles. As discussed in subsection 5.2.2 on page 43, Lier appears to measure too high solar

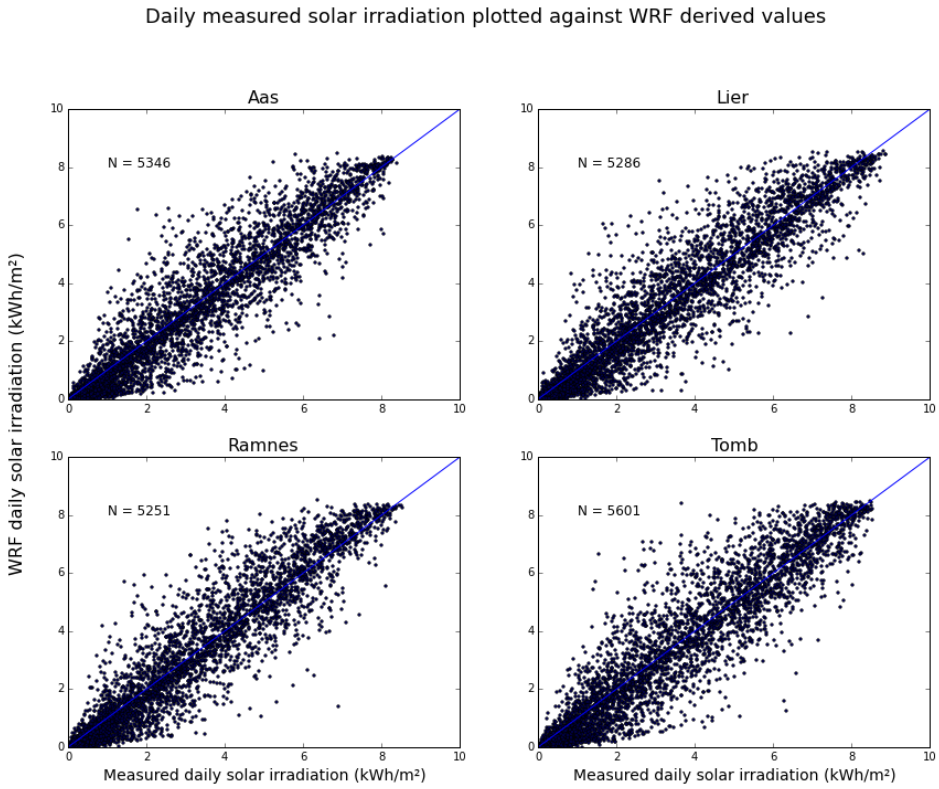
irradiation until 2010. Since the comparison in Fig. 5.4.1 is from before 2010, it supports this hypothesis. Tomb has a higher MAE and RMSD compared to Ås and Ramnes. This implies that Tomb has a higher deviation and could relate to its offset issues described in subsection 5.2.4 on page 46. Ramnes has the lowest MBD with 0.08 % and the lowest RMSD with 12.14 % and Ås has similar values. In total, daily solar irradiation from Satel-Light appears to have a good correlation with the quality controlled solar irradiation, however, appears to overestimate at very low daily values. For NASA SSE, the comparison with all stations are shown in Fig. 5.4.2.



	Aas	Lier	Ramnes	Tomb
MBD (%)	6.21	2.83	7.5	4.64
MAE (%)	13.58	13.47	14.41	14.74
RMSD (%)	19.02	19.07	20.1	20.71

Figure 5.4.2 – Daily measured solar irradiation plotted against the corresponding daily values derived from the NASA SSE database for all stations in the period from 1992 to 2005. Days where at least one value has been flagged during quality control have been ignored. Statistics are added to the figure, where MBD is mean bias deviation, MAE is mean absolute error and RMSD is root-mean-square deviation.

The NASA SSE daily solar irradiation in Fig. 5.4.2 deviate much more compared to Satel-Light in Fig. 5.4.1. This is expected since the NASA SSE values have a significantly lower resolution and provides equal daily values for all stations. Lier has the lowest MBD with 2.83 % and Ramnes has the highest MBD with 7.5 %. The MAE and RMSD are fairly equal for all stations with a MAE of $13-14$ % and a RMSD of $19-20$ %. Looking closer at Fig. 5.4.2, the NASA SSE daily solar irradiation appears to have highest correlation with high daily solar irradiation for Lier. This is the solar irradiation values where Lier probably overestimates as discussed in subsection 5.2.2 on page 43. Furthermore, for all other stations is most of the daily solar irradiation for NASA SSE higher compared to the corresponding quality controlled value. For WRF, the comparisons with all stations are shown in Fig. 5.4.3.



	Aas	Lier	Ramnes	Tomb
MBD (%)	-4.1	-2.63	-1.36	-3.71
MAE (%)	20.75	19.78	20.26	21.09
RMSD (%)	31.18	29.75	30.92	31.67

Figure 5.4.3 – Daily measured solar irradiation plotted against the corresponding daily values derived from the WRF model for all stations in the period from 1992 to 2012. Days where at least one value has been flagged during quality control have been ignored. Statistics are added to the figure, where MBD is mean bias deviation, MAE is mean absolute error and RMSD is root-mean-square deviation.

The WRF model is the database that has the highest deviation from quality controlled measured values. The MAE is around 20 % for all stations and the RMSD is around 30 %. However, the MBD is reasonably low, indicating that the total solar irradiation throughout each year is not as inaccurate as daily values. This is also shown in subsection 5.4.2. The WRF model has the best correlation with the highest daily solar irradiation In addition, it appears that WRF tend to underestimate for low daily solar irradiation and overestimate for higher daily solar irradiation. This is the reason WRF why has a low MBD for all stations.

5.4.2 Comparison with monthly averages

In order to apply solar irradiation data on solar energy applications such as PV systems and solar heating for a basic system analysis, monthly averages are commonly used (Honsberg and Bowden, 2014). In this subsection, the monthly averages from each database is compared with the quality controlled measured values from the Bioforsk stations, which is calculated according to the method presented in section 4.3 on page 34. The raw data for the monthly averages are shown in Appendix C. It is important to keep in mind that the monthly averages are made from different time periods and as a consequence, the comparisons are not meant to be perfectly equal. However, these averages are used as an estimation of present solar energy resource and the comparisons with quality controlled monthly averages are therefore valid. The results of the comparisons are shown in figures 5.4.4, 5.4.5, 5.4.6 and 5.4.7.

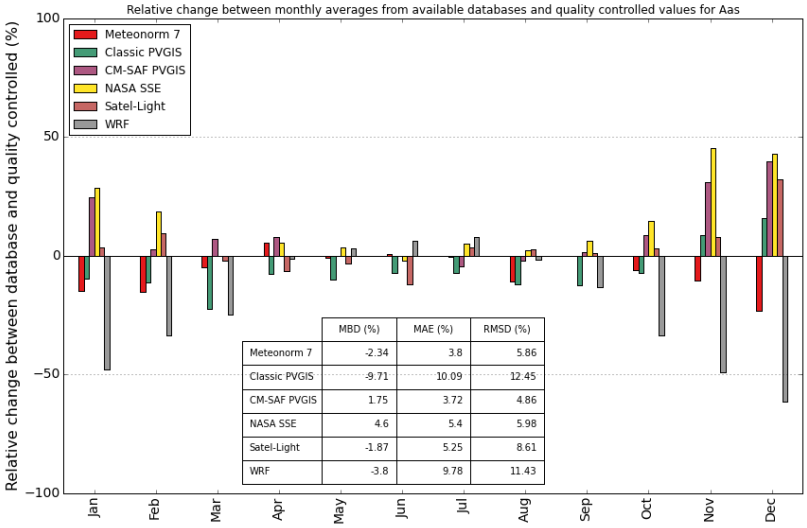


Figure 5.4.4 – Relative change between monthly average solar irradiation from databases and quality controlled values for Ås. Statistics for the comparison with all databases based on monthly values are shown in the table in the figure. MBD is mean bias deviation, MAE is mean absolute error and RMSD is root-mean-square deviation.

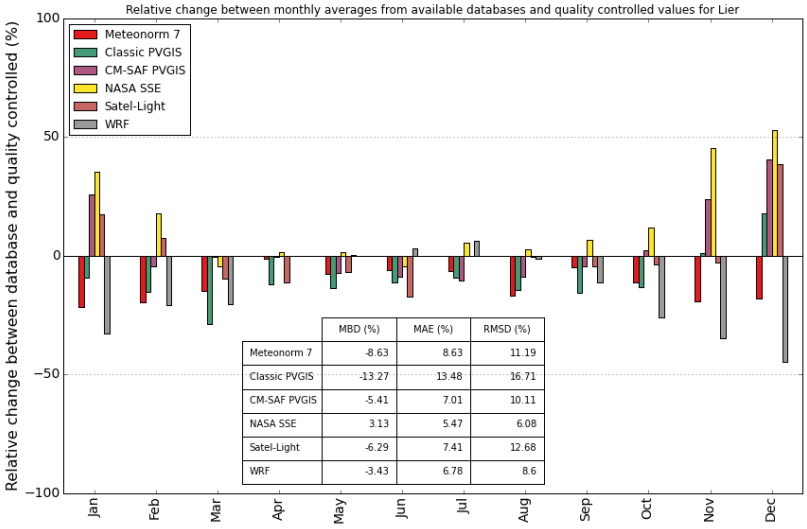


Figure 5.4.5 – Relative change between monthly average solar irradiation from databases and quality controlled values for Lier. Statistics for the comparison with all databases based on monthly values are shown in the table in the figure. MBD is mean bias deviation, MAE is mean absolute error and RMSD is root-mean-square deviation.

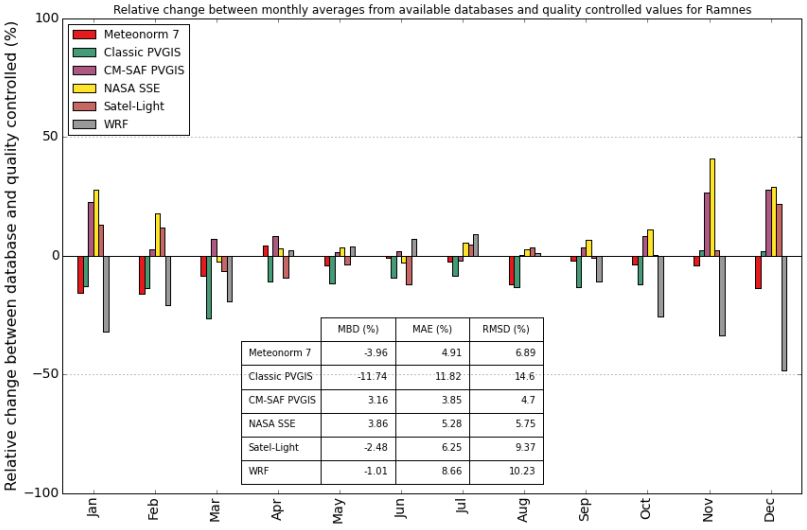


Figure 5.4.6 – Relative change between monthly average solar irradiation from databases and quality controlled values for Ramnes. Statistics for the comparison with all databases based on monthly values are shown in the table in the figure.

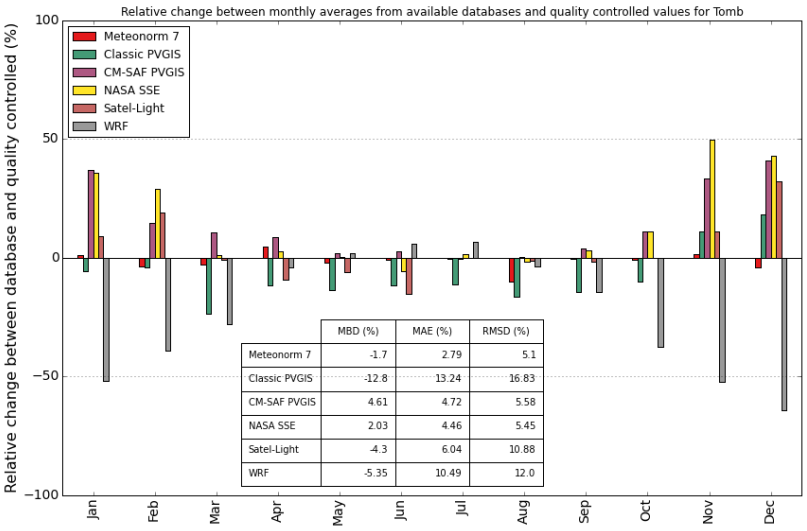


Figure 5.4.7 – Relative change between monthly average solar irradiation from databases and quality controlled values for Tomb. Statistics for the comparison with all databases based on monthly values are shown in the table in the figure.

In general, Meteonorm 7, CM-SAF PVGIS, and NASA SSE tend to have the lowest RMSD. Classic PVGIS has the highest RMSD for all stations due to heavy underestimation in the months from January to October. Meteonorm 7, Classic PVGIS, Satel-Light and WRF have a negative MBD for all stations. CM-SAF PVGIS, however, has a positive MBE for all stations except Lier and NASE SSE has a positive MBD for all stations. The relative change is higher during the winter compared to the summer. This is mainly due to a higher uncertainty of solar irradiation measurements at higher solar zenith angles, however, it is important to note that most of the solar irradiation is measured during the summer and thus the difference in solar irradiation is often higher in the summer compared to the winter. The databases appear to underestimate much more for Lier and this supports the hypothesis that Lier has measured too high values until the change of pyranometer in 2010 as discussed in section 5.2 on page 43. If the comparisons for Lier are excluded, the lowest MBD and RMSD for the stations are $-1.01\,\%$ for WRF and $4.70\,\%$ for CM-SAF PVGIS, both at Ramnes. The highest MBD and RMSD are $-12.80\,\%$ and $16.83\,\%$, both for Classic PVGIS at Tomb.

Looking at the monthly averages in figures 5.4.4, 5.4.5, 5.4.6 and 5.4.7, there is a very consistent pattern between the stations (except for Lier) for which months each database either overestimates or underestimates. For instance, for the months from October to March and August, Meteonorm 7 underestimates for all stations (except for two months for Tomb). From May to July and September, Meteonorm 7 appears to estimate very close to the quality controlled values. Furthermore, April is the only month where Meteonorm 7 consistently overestimates. Since Meteonorm 7 is based on data from 1991 to 2010, the comparisons between the database and quality controlled data are highly relevant. Despite overestimation and underestimation, Meteonorm 7 is one of two databases which fit the quality controlled data the most. The other database with low deviation from quality controlled monthly averages, CM-SAF PVGIS, consistently overestimates from October to April. However, in the rest of the year CM-SAF PVGIS appears to fit the quality controlled data. Similar conclusions can be made for the other databases. This result could provide an improvement to the existing databases and assist in a more correct estimation of the expected amount of solar radiation in a given location close to the Oslo area.

A database that separates from most of the other databases is WRF. From September to March, the WRF model heavily underestimates compared to the quality controlled measured solar irradiation. In, the summer, however, WRF does the exact opposite and overestimates.

From the figures 5.4.4, 5.4.5, 5.4.6 and 5.4.7, there are not a month where there are a consensus between the databases on whether to overestimate or underestimate. The closest to a consensus is for June, where most databases underestimate, except for WRF, Meteonorm 7 for Ås, and CM-SAF PVGIS for Ramnes and Tomb.

5.4.3 Comparison with yearly averages

Despite seasonal variations, it is interesting to compare a database's yearly average value with the quality controlled solar irradiation. The yearly averages have been calculated by a summation of the monthly averages, hence not accounting for missing values. These are shown in Fig. 5.4.8.

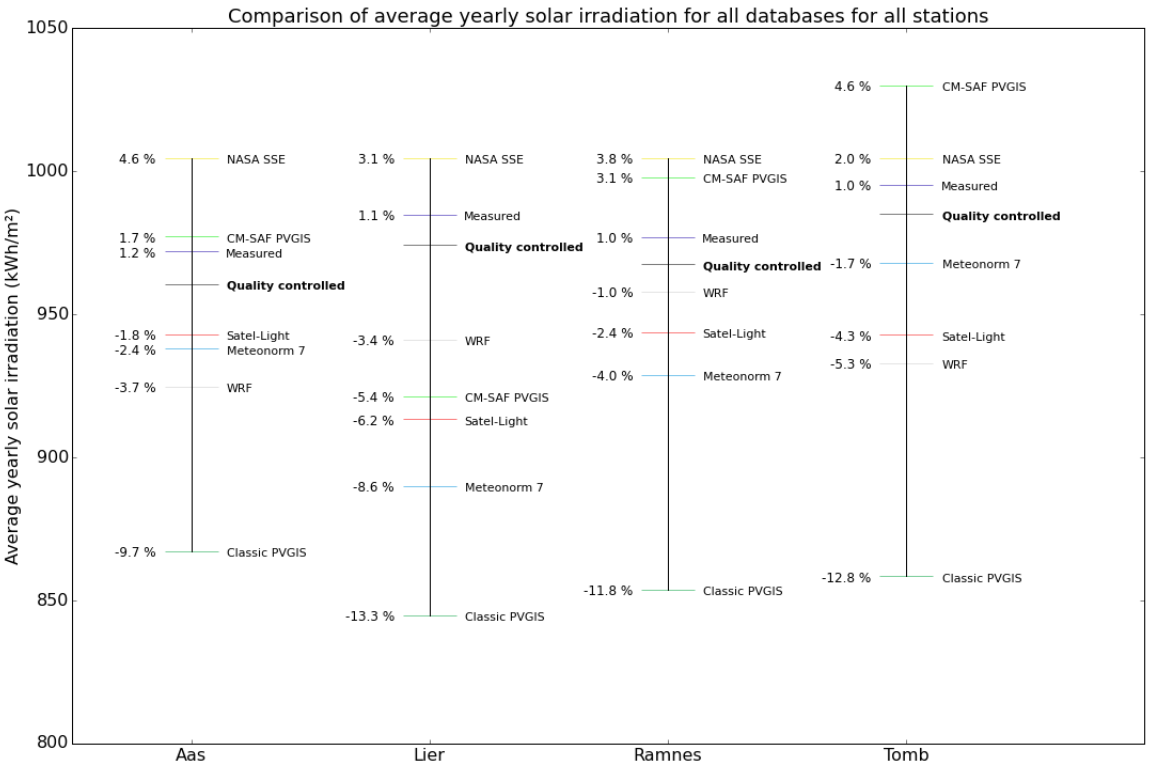


Figure 5.4.8 – Comparison of average yearly solar irradiation for all databases with quality controlled values for all stations. Classic PVGIS shows the lowest average values for all stations, while NASA SSE shows the highest average values for all stations except Tomb. The relative change is to the left of the database name and given in percent. Measured indicate yearly averages if no quality control has been performed except for ignoring missing values.

Classic PVGIS has the lowest average values for all stations, while NASA SSE has the highest average values for all stations except Tomb. For Ås, CM-SAF PVGIS and Satel-Light have the closest to quality controlled value. The closest to quality controlled value for all stations is WRF for Ramnes with $-1.0\,\%$ or $-9.50\,\frac{\mathrm{kWh}}{\mathrm{m^2}}$. Lier is the station where the deviation is highest, which is expected due to probably too high measured solar irradiation before 2010. Classic PVGIS underestimates with a relative change which ranges from $-13.3\,\%$ to $-9.7\,\%$, which correspond to $-129.4\,\frac{\mathrm{kWh}}{\mathrm{m^2}}$ and $-93.5\,\frac{\mathrm{kWh}}{\mathrm{m^2}}$ respectively. This underestimation is sufficient to conclude that Classic PVGIS underestimates to a great extent in the Oslo area and should not be used as an accurate source of present solar irradiation.

It is noteworthy to identify that the quality controlled data have less variation compared to most of the databases. Surprisingly, Meteonorm 7 and CM-SAF PVGIS are the two databases which differ the most in average yearly solar irradiation between each station. The relative change for Meteonorm 7 ranges from $-8.6\,\%$ to $-1.7\,\%$ and the relative change for CM-SAF PVGIS ranges from $-5.4\,\%$ to $4.6\,\%$. This could be due to an oversensitive compensation for the latitude, local climate or topography.

Fig. 5.4.8 illustrate to a great extent the issue that occurred for Størdal (2013), Aase (2013) and Romundstad (2014), that the relative difference in yearly solar irradiation between the solar irradiation databases were high and it was difficult to estimate which database was the most accurate. For the four stations evaluated in this thesis, the relative difference between the highest yearly average value and the lowest yearly average value ranges from $16\,\%$ for Ås to $20\,\%$ for Tomb. In Størdal (2013) and Aase (2013), were also Classic PVGIS utilized and displayed the lowest yearly value of the databases in both cases. If the Classic PVGIS yearly solar irradiation is ignored, the relative difference would improve to $8.7\,\%$ for Ås and $10\,\%$ for Tomb.

Satel-Light provided a good estimation for daily solar irradiation for the years from 1996 to 2000 as discussed in subsection 5.4.1. Since yearly solar irradiation for Satel-Light from 1996 to 2000 underestimates compared to quality controlled solar irradiation based on measurements from 1992 to 2012, it is a good indication that the average yearly solar irradiation has increased since 2000.

5.4.4 Comparison with quarterly and yearly averages from 16 Bioforsk stations

As described in 3.1 on page 19, 16 stations which measure solar irradiation in Eastern Norway appeared to have good quality after a preliminary quality control. Only 4 stations have been evaluated in detail in this thesis due to time constraints. However, once the automatic control procedure was created, the automatic quality control described in 4.1 on page 29 on the remaining stations was easily performed. The name, location, time series and quality control of each station are described in table 5.4.1.

Table 5.4.1 – The stations that are used in average quarterly and yearly comparisons with solar irradiation databases. The *Quality control* column indicate which quality control procedure has been conducted on each station. Only the four station evaluated in this thesis have had both automatic and visual control.

Name of station	Latitude (°)	Longitude (°)	Time series	Quality control
Ås	59.660468	10.781989	1992 - 2012	Automatic and visual
Lier	59.79005	10.2604	1992 - 2012	Automatic and visual
Ramnes	59.38081	10.23923	1992 - 2012	Automatic and visual
Tomb	59.31893	10.81449	1992 - 2012	Automatic and visual
Årnes	60.1268	11.39342	1999 - 2012	Automatic
Apelsvoll	60.70024	10.86952	1988 - 2012	Automatic
Bø	59.4175	9.02859	1992 - 2012	Automatic
Fåvang	61.45822	10.1872	1993 - 2012	Automatic
Gausdal	61.22468	10.25878	1993 - 2012	Automatic
Hokksund	59.76152	9.89166	1992 - 2012	Automatic
Hønefoss	60.14032	10.2661	1992 - 2012	Automatic
Ilseng	60.80264	11.20298	1992 - 2012	Automatic
Kise	60.77324	10.80569	1987 - 2012	Automatic
Løken	61.12183	9.06302	1988 - 2012	Automatic
Rakkestad	59.38824	11.39042	1992 - 2012	Automatic
Roverud	60.25378	12.09144	1992 - 2012	Automatic

The 12 new stations could have quality issues that are not detected in the automatic quality control, however, for the four stations (except Lier) only a tiny portion was flagged from visual control. In addition, this gives a larger statistical sample for the comparisons with existing databases. Thus, the relative change has been calculated between average quarterly and yearly solar irradiation from existing databases and the corresponding quality

controlled value. Note that CM-SAF PVGIS has only been compared with the 7 stations located on latitudes below 60°N The average relative change for all 16 stations have been found and is shown in Fig. 5.4.9.

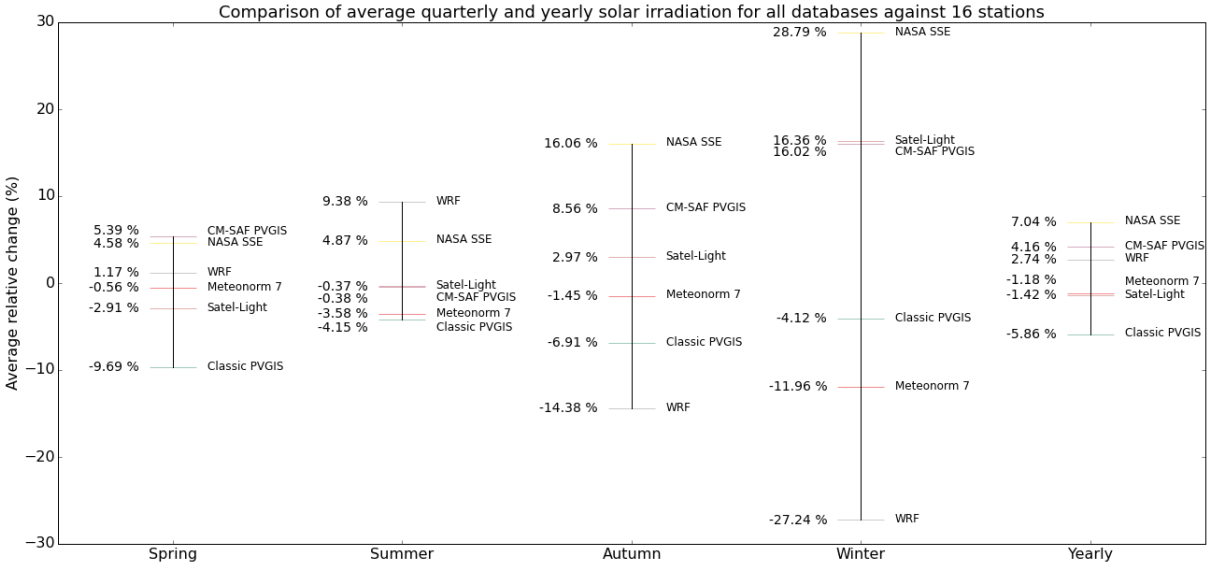


Figure 5.4.9 – The average relative change between quality controlled data and databases for average quarterly and yearly solar irradiation at 16 stations including the four stations evaluated in detail in this thesis. The stations are located in the eastern part of Norway and the exact location, time series and quality control of the 16 stations are described in table 5.4.1 on the facing page. Spring consists of March, April and May, summer consists of June, July and August, Autumn consists of September, October and November and winter consist of December, January and February. CM-SAF has only been compared with the seven stations that have a latitude south of 60°N.

The databases have the lowest average relative change in the spring and the summer. The highest deviations from quality controlled values are found in the winter. Meteonorm 7 has the average relative change closest to zero with an average of $-4.39\,\%$ for all quarters. Satel-Light and CM-SAF PVGIS have the least average relative change in the summer months, when the total solar irradiation is largest. NASA SSE has the average relative change farthest from zero with an average of $13.58\,\%$ for all quarters.

Fig. 5.4.9 gives similar results to the ones found in section 5.4.2 on page 58. The quarterly averages have good correlation with monthly averages shown in figures 5.4.4, 5.4.5, 5.4.6 and 5.4.7 on page 60. For the yearly averages, WRF has an overestimation for the 16 stations, which is a different trend compared to the results for the initial four stations. In addition, Classic PVGIS displays a lower deviation compared to the yearly averages in 5.4.8,

however, still underestimates to a great extent.

An interesting note is that the three databases which are solely based on satellite data tend to overestimate in the winter months. This is counter intuitive to the fact that satellite may interpret snow as clouds and underestimate as discussed in Hagen (2011). An explanation could be that the algorithm in the calculations of solar irradiation tries to account for this and interpret clouds as snow or that the measurements underestimate due to snow on the sensor. Furthermore, most of the solar irradiation in the winter is registered at high solar zenith angles. Since satellite images do not take the horizon into account, that may be the reason for the overestimation. Low solar zenith angles are also where the instrument has most sources of error.

The average relative change, however, may be misleading since the relative change often is positive and negative for different stations and the average calculation cancels these out. Therefore, it is interesting to examine the average *absolute* relative change between average quarterly and yearly solar irradiation from existing databases and the corresponding quality controlled value. The average absolute relative change for all 16 stations has been found and is shown in Fig. 5.4.10.

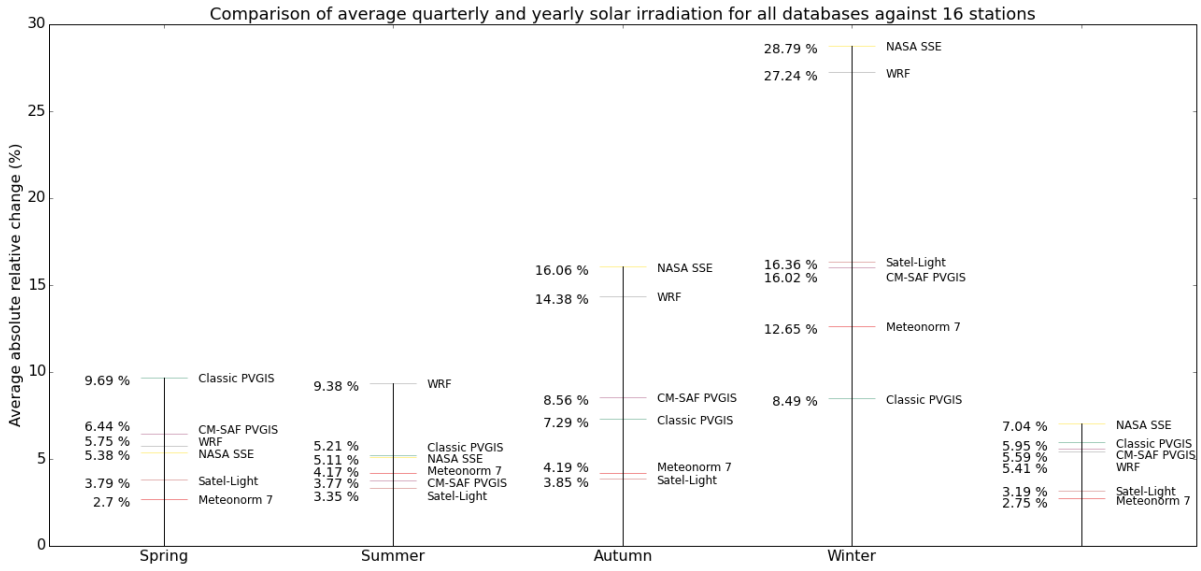


Figure 5.4.10 – The average absolute relative change between quality controlled data and databases for quarterly and yearly solar irradiation at 16 stations including the four stations evaluated in detail in this thesis. The stations are located in the eastern part of Norway and the exact location, time series and quality control of the 16 stations are described in table 5.4.1 on page 64. Spring consists of March, April and May, summer consists of June, July and August, Autumn consists of September, October and November and winter consist of December, January and February. CM-SAF has only been compared with the seven stations that have a latitude south of 60°N.

Most of the databases have an increased average relative change compared to the values in Fig. 5.4.9. Some exceptions occur for quarters where the relative change is far from zero. None of the databases has an average quarterly absolute relative change less than 2.70 %. Meteonorm 7 has the lowest average absolute relative change with an average of 5.9 % for all quarters and the lowest average yearly absolute relative change with 2.75 %. The WRF model has the highest average absolute relative change with 14.20 % for all quarters and NASA SSE has the highest average yearly absolute relative change with 7.04 %.

5.4.5 Summary of the comparisons with existing databases

The quality controlled measured values have been compared with existing solar irradiation databases. An overview of the performance of each database is shown below.

NASA SSE 6.0 NASA SSE overestimates in general compared to the quality controlled measured solar irradiation, especially at low daily values. Monthly averages from NASA SSE overestimate for all months except for June and the yearly averages are in

general higher than quality controlled values. For 16 stations, the average yearly absolute relative change is 7.04 %.

Satel-Light Satel-Light appears to have good correlation with quality controlled measured daily solar irradiation. For monthly averages, Satel-Light overestimates in the winter months and underestimates in June. The yearly averages are in general lower than quality controlled values. Satel-Light only has data from 1996 to 2000 and the changes in monthly and yearly averages may be due to climatic changes in solar irradiation since 2000. For 16 stations, the average yearly absolute relative change is 3.19 %.

WRF The WRF model underestimates for low daily solar irradiation and overestimates for higher values, resulting in a slight underestimation in yearly average solar irradiation. The general deviation from quarterly quality controlled averages is very high. For 16 stations, the average yearly absolute relative change is 5.41 %.

Meteonorm 7 Monthly averages from Meteonorm 7 underestimate from October to March and in August, have good estimation from May to July and September and overestimate in April. For average yearly values, Meteonorm 7 underestimates for all four stations. However, for 16 stations, Meteonorm 7 has the lowest average yearly absolute relative change with 2.75 %.

Classic PVGIS Monthly averages from Classic PVGIS underestimate from January to October and overestimate in November and December. This results in a substantial underestimation in yearly averages compared to quality controlled values from the four stations and suggests that Classic PVGIS should not be used to represent present yearly solar irradiation in the Oslo area. The average yearly absolute relative change for 16 stations, however, is 5.95 %, which indicate that Classic PVGIS underestimates to a lesser extent for the 12 other stations.

CM-SAF PVGIS Monthly averages from CM-SAF PVGIS overestimate from October to April and have good estimation in the summer. For average yearly values, CM-SAF PVGIS overestimates for all stations except Lier. The variation in overestimation and underestimation results in an average yearly absolute relative change of 5.59 % for seven stations.

5.5 Overall discussion

5.5.1 Overall quality of the Bioforsk data

After an extensive quality control of the measured solar irradiation at the four Bioforsk stations it is indispensable to discuss the overall quality of the data. A common starting point is to discuss the sources of error in a set of measured data. As described in subsection 2.4.3 on page 16, the errors are either related to equipment error and uncertainty, or operation related problems and errors (Younes et al., 2005). The equipment used at the four Bioforsk stations, the CM11 Pyranometer, is of highest quality class and it is regarded as the standard reference pyranometer due to its accuracy, stability and quality of construction (Muneer and Fairooz, 2002). It is, however, recommended by Kipp & Zonen to perform a periodic recalibration of the instrument at least every two years to ensure no change in sensitivity. This procedure has not been followed for the pyranometers at the four Bioforsk stations, however, from the two calibration certificates shown from another pyranometer (in Appendix D), the indication was that the change in sensitivity is low. Furthermore, yearly maintenance routines and frequent observations of the outputs are performed. In addition, the station hosts perform regular maintenance. Since the data are stored in a central database, missing values due to station shutdown are easily registered and fixed quickly.

For the automatic quality control, the percentage of flagged data for the different tests ranges from 0 % to 3.8 % as shown in Fig. 5.1.1. The highest percentage of flags appears in the period from 1992 to 1996, mostly due to missing values. In section 5.3, the results show that if the data points that are flagged as erroneous are replaced with the mean value for that hour, the total change ranges from -1.26 % for Lier and -0.78 % for Ås. In other words, an extensive quality control, where erroneous values are changed with the mean value for that hour, does not change the measured values to a great extent and the measured data appear to have good quality. The visual control, however, provides another aspect to the quality control process. There are errors that are difficult to detect through an automatic control, such as operational issues and temporal change. Through the visual control it has become apparent that there are several issues with the time series from the station at Lier. The issue appears to be that until the replacement of the pyranometer in 2010, the station at Lier has measured too high values. This is shown in Fig. 5.2.2 and confirmed when compared to Satel-Light values in Fig. 5.4.1. This was not discovered through the automatic quality control procedure and indicates that even the data points from Lier that have passed the quality control are still not trustworthy.

An unfortunate limitation of the quality control procedure of the measurements is that the Bioforsk stations only measure global irradiation and do not have an additional measurement device that measures beam or diffuse irradiation. Several additional quality control procedures could be applied if one of the two was provided (Journée and Bertrand, 2011; Shi et al., 2008; Younes et al., 2005). There are, however, a few Bioforsk stations (not the four stations chosen for this thesis) that measure sunshine duration, which could also be used in a quality control procedure as in Moradi (2009). Since sunshine duration is not commonly measured at Bioforsk stations, its possibilities have not been investigated in this thesis. All Bioforsk stations do, however, measure other meteorological parameters such as relative humidity and surface temperature. Even though these parameters could be used in a model for estimating the solar irradiation as in Yang et al. (2006), this is beyond the scope of this thesis. In addition, the use of only one parameter does provide simplicity for future applications.

The uncertainty of each daily and hourly measurement is set to be 3 % from the Kipp & Zonen CM 11 pyranometer manual (Kipp & Zonen, 2000). Furthermore, other factors have to be taken into account, which may influence the accuracy of the measurements. This is important to note when comparing with other solar databases.

5.5.2 Usefulness of quality controlling several stations simultaneously

In this thesis, four stations have been quality controlled through an automatic and a visual control. Having more than one station has proven advantageous in many situations, since some erroneous data often initially appear to be valid. The fact that Lier had a time displacement issue, as discussed in subsection 5.2.5, was more easily identifiable while comparing with the three other stations. In addition, since Tomb was the only station to have many Offset flags, this was examined more detail. The comparisons between stations are more reliable the closer each station are located each other due to more similar climate and weather. Even though the four stations are all located around the Oslofjord, not many conclusions could be made. If two stations measure differently in a hourly or daily time frame there is not enough knowledge on whether the difference is caused by natural high variation in meteorological conditions or erroneous measurements. This is the reason why direct comparisons of measured solar irradiation between the stations are limited. For that purpose, two pyranometers at the same location or very close by have to be utilized.

5.5.3 Further work

As described in section 3.1 on page 19, there are 47 Bioforsk stations which measure global solar irradiation. All of these stations should be quality controlled thoroughly with similar methods as used in this thesis. Furthermore, the quality controlled values should be compared with existing databases to determine whether there is a general agreement on which parts of the year databases overestimate or underestimate. Quality controlled data from the Bioforsk stations should be incorporated in solar irradiation databases. This would provide to a more accurate estimation of solar irradiation in Norway. In addition, comparisons between quality controlled measurements at several Bioforsk stations could provide valuable information on topographical and climatological effects on solar irradiation.

The meteorological station at NMBU, FAGKLIM, measure global and diffuse solar irradiation. Since these instruments are located at the same location as the Bioforsk station in Ås, the measurements should be compared to note and deduce discrepancies in the measurements. The limitations described in subsection 5.5.2 would not be an issue.

To improve the quality control and easier conversion for solar energy applications, diffuse or beam irradiation should be measured at the Bioforsk stations. As described in section 5.5.1, a beam or diffuse component would open possibilities for more tests, which easier reveal erroneous data. The diffuse component is also important for solar energy use and has to be estimated if only global irradiation is available.

To prevent changes in the sensitivity over time, pyranometers should be calibrated regularly. If a pyranometer is replaced with a newer one, both should be operative at a longer time period to ensure that they measure close to equal amount of solar irradiance.

In this thesis, the deviations for monthly averages for different databases compared to quality controlled measurements have been mapped. A study should be made where a correction for these deviations is made for a location close to the four Bioforsk stations and afterwards compare estimated power production with actual power production to analyze whether this contributes to a more accurate result. That would greatly improve the knowledge on how useful the measurements at the Bioforsk stations are.

Chapter 6

Conclusions

A thorough quality control analysis of measured solar irradiation from 1992 to 2012 at four Bioforsk stations in the Oslo area has been performed. The stations are Ås, Lier, Ramnes, and Tomb. The quality control consisted of an automatic control and a visual control. The automatic quality control was mostly based on previous academic work and utilized modeled extraterrestrial solar irradiation and irradiation under clear sky as parameters for upper and lower limit of valid measurements. If a data point failed an automatic test, the data point was either flagged as erroneous or marked for visual control. The visual quality control consisted of a comparison between the four stations, an evaluation whether there were any temporal changes caused to the measurements and determining if data points flagged for visual control were erroneous.

The percentage of erroneous data from sunrise to sunset after the quality control was 6.96 % for Ås, 9.57 % for Lier, 5.23 % for Ramnes, and 5.35 % for Tomb. Most of the erroneous data points were due to solar irradiation that exceeded the upper limits set by the tests and missing values. In addition to the erroneous data, Lier appeared to measure higher solar irradiation before the replacement of the pyranometer in 2010 compared to after the replacement. From the comparison with existing solar irradiation databases, the solar irradiation for Lier before 2010 appeared to be excessively high. For the other three stations, the overall quality of the data appeared to be good. After the quality control, an estimation of the impact of the quality control on total measured value has been performed. This indicates that if the time period from 1992 to 1996 (2000 for Lier) is ignored due to missing values, measured values are expected to overestimate quality controlled values by roughly 1 %. Given that the uncertainty for the measurements is at least 3 %, the quality of measured solar irradiation is sufficient.

Quality controlled measurements have been compared with existing solar irradiation databases. The databases are Satel-Light, NASA SSE 6.0, the WRF model, Meteonorm 7, Classic PVGIS and CM-SAF PVGIS. For the first three databases, daily solar irradiation has been compared with quality controlled values. Satel-Light had the best correlation with an average root-mean-square deviation (RMSD) of 12.14 %. NASA SSE and WRF had a RMSD of 19.73 % and 30.88 %, respectively. Monthly averages from the six databases were compared with quality controlled data. Most databases underestimated or overestimated in the same months for all stations. This knowledge could be used to improve the databases for solar resources in the Oslo area.

Furthermore, yearly averages from the databases were compared with quality controlled data. In general, the six databases were divided into three groups. NASA SSE and CM-SAF PVGIS overestimated for all four stations in almost all cases, where CM-SAF in general was closer to quality controlled value. Meteonorm 7, WRF and Satel-Light underestimated for all four stations, however, not by a substantial amount. Classic PVGIS heavily underestimated for all four stations.

Lastly, the automatic quality control was performed for 12 additional Bioforsk stations located in Eastern Norway. Existing databases were compared with average quarterly and yearly solar irradiation from the original four stations and the 12 new stations. The average relative change and the absolute average relative change were calculated. The comparisons appeared to display mostly similar results as for the four stations.

Bibliography

Aase, G. K. F.: 2013, Master's thesis, Norwegian University of Life Sciences. Rekåa Hybrid Power Station, Yield Simulations and Recommended Regulation of a PV System Combined with a Hydroelectric Power Station.

Andersen, M.: 2014, Master's thesis, Norwegian University of Life Sciences. Analysis of actual and forecasted power production for a solar energy system in Norway.

Astronomical Almanac: 2014, *Astronomical Almanac for the Year 2015 and Its Companion, The Astronomical Almanac Online*, Dept. of the Navy.

Bioforsk: 2015, Bioforsk About Web Page, http://www.bioforsk.no/ikbViewer/page/en/about. Accessed: 2015-02-18.

Byrkjedal, Ø., Løvholm, A. L. and Liléo, S.: 2013, Resource mapping of solar energy - An overview of available data in Norway, *Technical report*.

Cano, D., Monget, J.-M., Albuisson, M., Guillard, H., Regas, N. and Wald, L.: 1986, A method for the determination of the global solar radiation from meteorological satellite data, *Solar Energy* **37**(1), 31–39.

Chan, A. L., Chow, T.-T., Fong, S. K. and Lin, J. Z.: 2006, Generation of a typical meteorological year for hong kong, *Energy Conversion and Management* **47**(1), 87–96.

Chen, C. J.: 2011, *Physics of solar energy*, John Wiley & Sons.

Clarke, P.: 2009, *Mathematical modelling of BIPV-micro wind system: production, storage and usages*, PhD thesis, Napier University.

Coulson, K. L.: 1975, *Solar and Terrestrial Radiation: Methods and Measurements*, Academic Press.

Duffie, J. A. and Beckman, W. A.: 1980, *Solar engineering of thermal processes*, Vol. 3, Wiley New York etc.

eKlima: 2015, Free access to weather- and climate data from Norwegian Meteorological Institute from historical data to real time observations, eklima.met.no. Accessed: 2015-03-23.

Fontoynont, M., Dumortier, D., Heinnemann, D., Hammer, A., Olseth, J., Skarveit, A., Ineichen, P., Reise, C., Page, J., Roche, L. et al.: 1998, Satellight: a WWW server which provides high quality daylight and solar radiation data for Western and Central Europe, *9th Conference on Satellite Meteorology and Oceanography*, number 2, American Meteorological Society Ed., Boston, Massachusetts, USA, pp. 434–437.

Geiger, M., Diabaté, L., Ménard, L. and Wald, L.: 2002, A web service for controlling the quality of measurements of global solar irradiation, *Solar energy* **73**(6), 475–480.

Godøy, Ø.: 2013, Rtmrun, a Perl wrapper around libRadtran , https://github.com/steingod/rtmrun. Accessed: 2015-02-09.

Hagen, L.: 2011, Master's thesis, University of Bergen. Measured, Modelled and Satelite derived solar radiation in Scandinavia.

Hauge, Å. L., Sørnes, K., Godbolt, Å. L., Kristjansdottir, T., Sørensen, Å. L. and Fredriksen, E.: 2014, Suksessfaktorer for økt bruk av solvarme, *Technical report*.

Hole, H.: 2015, Personal communication with Halvard Hole, Bioforsk.

Honsberg, C. and Bowden, S.: 2014, Pv education, *ORG.(access May 2015) http://www.pveducation.org/pvcdrom/properties-of-sunlight/average-solar-radiation* .

Huld, T., Müller, R. and Gambardella, A.: 2012, A new solar radiation database for estimating PV performance in Europe and Africa, *Solar Energy* **86**(6), 1803–1815.

Iqbal, M.: 1983, *Introduction to Solar Radiation*, Academic Press Inc.

ISO: 1990, *ISO 9060:1990: Solar energy – Specification and classification of instruments for measuring hemispherical solar and direct solar radiation*.

ITAS: 2013, *Driftsrutine LMT værstasjon 2013*.

Jones, E., Oliphant, T., Peterson, P. et al.: 2001–, SciPy: Open source scientific tools for Python. [Online; accessed 2015-04-20].

Journée, M. and Bertrand, C.: 2011, Quality control of solar radiation data within the RMIB solar measurements network, *Solar Energy* **85**(1), 72–86.

Kipp & Zonen: 2000, *CM 11 Pyranometer / CM 14 Albedometer - Manual*.

Kroken, S.: 2015, Personal communication with Signe Kroken, the station host of Ås station.

Mayer, B. and Kylling, A.: 2005, Technical note: The libRadtran software package for radiative transfer calculations - description and examples of use, *Atmospheric Chemistry and Physics* **5**(7), 1855–1877.

McKinney, W.: 2010, Data Structures for Statistical Computing in Python, *in* S. van der Walt and J. Millman (eds), *Proceedings of the 9th Python in Science Conference*, pp. 51–56.

Mermoud, A.: 2012, Pvsyst: Software for the study and simulation of photovoltaic systems, *ISE, University of Geneva, www.pvsyst.com* .

met.no: 2015, Kart med nedbørnormal for Norge i perioden 1961-1990. http://met.no/Klima/Klimastatistikk/Klimanormaler/Nedbor/.

Michalakes, J., Dudhia, J., Gill, D., Klemp, J. and Skamarock, W.: 1998, Design of a next-generation regional weather research and forecast model, *Towards teracomputing* pp. 117–124.

Moradi, I.: 2009, Quality control of global solar radiation using sunshine duration hours, *Energy* **34**(1), 1–6.

Muneer, T.: 1997, *Solar Radiation and Daylight Models for the Energy Efficient Design of Buildings*, Architectural Press.

Muneer, T. and Fairooz, F.: 2002, Quality control of solar radiation and sunshine measurements–lessons learnt from processing worldwide databases, *Building Services Engineering Research and Technology* **23**(3), 151–166.

Myers, D. R.: 2005, Solar radiation modeling and measurements for renewable energy applications: data and model quality, *Energy* **30**(9), 1517–1531.

Nordmann, T., Clavadetscher, L., van Sark, W. G. and Green, M.: 2014, Report IEA: Analysis of Long-Term Performance of PV Systems, *Technical report*, International Energy Agency.

Petrakis, M., Kambezidis, H., Lykoudis, S., Adamopoulos, A., Kassomenos, P., Michaelides, I., Kalogirou, S., Roditis, G., Chrysis, I. and Hadjigianni, A.: 1998, Generation of a typical meteorological year for nicosia, cyprus, *Renewable Energy* **13**(3), 381–388.

PVGIS: 2015, PVGIS radiation databases, http://re.jrc.ec.europa.eu/pvgis/apps4/databasehelp_en.html. Online, accessed 22.04.2015.

Remund, J.: 2008, Quality of meteonorm version 6.0, *Europe* **6**(1.1).

Romundstad, R. M.: 2014, Master's thesis, Norwegian University of Life Sciences. The Possibility of Integrating Solar Collectors in Hafslund's District Heating System at Gardermoen.

Ruud Hansen, Ø.: 2015, Personal communication with Øystein Ruud Hansen, head in charge of maintenance of all Bioforsk stations.

Scharmer, K. and Greif, J.: 2000, *The European solar radiation atlas*, Vol. 1, Presses des MINES.

Shi, G.-Y., Hayasaka, T., Ohmura, A., Chen, Z.-H., Wang, B., Zhao, J.-Q., Che, H.-Z. and Xu, L.: 2008, Data quality assessment and the long-term trend of ground solar radiation in China, *Journal of Applied Meteorology and Climatology* **47**(4), 1006–1016.

Stackhouse, P. and Whitlock, C.: 2008, Surface meteorology and Solar Energy (SSE) release 6.0, NASA SSE 6.0, *Earth Science Enterprise Program, National Aeronautic and Space Administration (NASA), Langley, http://eosweb.larc.nasa.gov/sse* .

Størdal, A. M.: 2013, Master's thesis, Norwegian University of Life Sciences. PV System Design and Yield Simulations for a Farm in Rygge Municipality.

Šúri, M., Huld, T. A. and Dunlop, E. D.: 2005, Pv-gis: a web-based solar radiation database for the calculation of pv potential in europe, *International Journal of Sustainable Energy* **24**(2), 55–67.

Takezawa, K.: 2005, *Introduction to nonparametric regression*, Vol. 606, John Wiley & Sons.

Tang, W., Yang, K., He, J. and Qin, J.: 2010, Quality control and estimation of global solar radiation in China, *Solar Energy* **84**(3), 466–475.

Utaaker, K.: 1991, *Mikro og lokalmeteorologi*, Alma Mater.

Williams, D. R.: 2013, Sun Fact Sheet, http://nssdc.gsfc.nasa.gov/planetary/factsheet/sunfact.html. Online, accessed 29.01.2015.

WMO: 2008, *Guide to Meteorological Instruments and Methods of Observation 7th edition*.

Yang, K., Koike, T. and Ye, B.: 2006, Improving estimation of hourly, daily, and monthly solar radiation by importing global data sets, *Agricultural and Forest Meteorology* **137**(1), 43–55.

Younes, S., Claywell, R. and Muneer, T.: 2005, Quality control of solar radiation data: present status and proposed new approaches, *Energy* **30**(9), 1533–1549.

Appendix A

libRadtran configuration file

```
# Specify libRadtran version to use
libradtran => /home/sigbjorn/software/libRadtran-1.7
# libRadtran applications that are used
uvspec => /home/sigbjorn/software/libRadtran-1.7/bin/uvspec
zenith => /home/sigbjorn/software/libRadtran-1.7/bin/zenith -q -t1
# Atmospheric composition and extraterrestric specturm
atmosphere_file => /home/sigbjorn/software/libRadtran-1.7/data/atmmod/afglss.dat
solar_file => /home/sigbjorn/software/libRadtran-1.7/data/solar_flux/atlas_plus_modtran
# File to store results on
outfile => libradtran_simulation_Ramnes.txt
# Temporary file to use during calculations
tmpfile => tmpfile_Ramnes.txt
# File to store error messages in
errfile => errfile_Ramnes.txt

# Specify position
latitude => 59.38081
longitude => -10.23923 # positive westwards for libRadtran
# Specify time period
startdate => 1992-01-01T00:30:00
enddate => 2012-01-01T23:59:59
# Specify timestep/recursion in minutes
timestep => 60
```

```
# Specify model setup
rte_solver => cdisort # Radiative transfer equation solver
deltam => on # delta-M scaling on
nstr=> 6 # Number of streams
wavelength => 300.0 3000.0 # Wavelength range [nm]
correlated_k => kato2
output => sum
output_user => sza eglo
albedo => 0.20 # Surface albedo
dens_column => O3 250 DU
h2o_precip => 5.
cdisort_pseudospherical => NOVALUE
```

Appendix B

Python code

B.1 Bioforskstation class with all tests as functions

```
1  import pandas as pd
2  import numpy as np
3  import ConfigParser
4  from datetime import datetime
5  import os
6
7
8  class BioforskStation(object):
9      """
10     A container for all applications applied to the data from the
11     Bioforsk Stations.
12
13     :param name_of_station: The name of the location of the station.
14     :type name_of_station: string
15     :param remove_partial_years: If you want your data to not show
16       the first and last year if they are incomplete.
17     :type remove_partial_years: boolean, default True
18
19     """
20
21     conf = ConfigParser.RawConfigParser()
22
23     def __init__(self, name_of_station, remove_partial_years=True, path='..'):
24         """Sets up an instance"""
25         self.conf.read(os.path.join(path, 'config', 'stations.cfg'))
26         self.name = name_of_station
27         self.data = pd.read_csv(os.path.join(path, 'data', 'raw_data',
28                                              '{}.csv'.format(
29                                                  name_of_station)),
30                                 sep=';', parse_dates=True,
31                                 index_col='time_measured', dayfirst=True)
32         self.longitude = self.conf.getfloat(self.name, 'lon')
33         self.latitude = self.conf.getfloat(self.name, 'lat')
```

```
34          self.altitude = self.conf.getfloat(self.name, 'hgt')
35          self.station_id = self.conf.getfloat(self.name, 'id')
36
37          # Makes sure to fill in NaN data where there are none
38          # registered in the data file
39          self.data = self.data.asfreq('H')
40
41          self.raw = self.data.copy() # Keeps the raw files
42
43          if remove_partial_years:
44              self.remove_partial_year()
45
46          # Sometimes the instrument records -6999 og 6999 instead of NaN
47          self.data.loc[(self.data['qo'] == -6999) |
48                        (self.data['qo'] == 6999), 'qo'] = np.nan
49
50          # Import extraterrestrial irradiation
51          self.data.loc[:, 'toa'] = pd.read_csv(os.path.join(path, 'data', 'toa',
52                                                             '{}toa.csv'.format(
53                                                                 self.name)),
54                                                header=None).values
55
56          # Import clear sky irradiation
57          clear_sky = pd.read_csv(os.path.join(path, 'data', 'clear_sky',
58                                               '{}clear.txt'.format(
59                                                   self.name)),
60                                  header=None, delim_whitespace=True,
61                                  parse_dates=True, index_col=0, dayfirst=True)
62          clear_sky.columns = ['sza', 'qo']
63          self.data.loc[:, 'sza'] = clear_sky['sza'].values
64          self.data.loc[:, 'clear_sky'] = clear_sky['qo'].values
65
66          # Initializes the flags DataFrame
67          self.flags = pd.DataFrame(index=self.data.index)
68
69      def remove_partial_year(self):
70          """Returns removed start and end year if not fully complete"""
71          if not self.data.index[0].is_year_start:
72              self.data = self.data[self.data.index.year
73                                    != self.data.index[0].year]
74          if not self.data.index[-1].is_year_end:
75              self.data = self.data[self.data.index.year
76                                    != self.data.index[-1].year]
77
78      def count_flags_per_year(self, pesd=False):
79          """
80          Returns a count of all flags registered from every test in yearly sum
81
82          :param pesd: If true the data will be percent of errorneous data.
83          :type pesd: boolean, default False
84          :returns: Pandas DataFrame
85
86          """
87          if pesd:
88              flagged = self.flags[self.data['toa'] > 0].groupby(self.flags[
```

```
89                  self.data['toa'] > 0].index.year).aggregate(np.mean) * 100
90
91              # Offset flag is a percentage of every data
92              flagged.loc[:, 'Offset'] = self.flags.loc[:, 'Offset'].groupby(
93                  self.flags.index.year).aggregate(np.mean) * 100
94
95
96          else:
97              flagged = self.flags.groupby(self.flags.index.year).aggregate(sum)
98              flagged['Sum'] = flagged.sum(axis=1)
99
100         # Change column names for easier read. Changed some of the names according
101         # to thesis.
102         flagged.name = self.name
103         flagged.index.name = 'Year'
104
105         return flagged
106
107     def count_flags_per_month(self, pesd=False):
108         """
109         Returns a count of all flags registered from every test in yearly sum
110
111         :param pesd: If true the data will be percent of errorneous data.
112         :type pesd: boolean, default False
113         :returns: Pandas DataFrame
114
115         """
116
117         if pesd:
118             flagged = self.flags[self.data['toa'] > 0].groupby(self.flags[
119                 self.data['toa'] > 0].index.month).aggregate(np.mean) * 100
120
121             # Offset flag is a percentage of every data
122             flagged.loc[:, 'Offset'] = self.flags.loc[:, 'Offset'].groupby(
123                 self.flags.index.month).aggregate(np.mean) * 100
124
125
126         else:
127             flagged = self.flags.groupby(
128                 self.flags.index.month).aggregate(sum)
129             flagged.loc[:, 'Offset'] = self.flags.loc[:, 'Offset'].groupby(
130                 self.flags.index.month).aggregate(sum)
131             flagged['Sum'] = flagged.sum(axis=1)
132
133         flagged.index = ['January', 'February', 'March', 'April', 'May',
134                          'June', 'July', 'August', 'September', 'October',
135                          'November', 'December']
136
137         flagged.name = self.name
138         flagged.index.name = 'Month'
139
140         return flagged
141
142     def zero_out(self):
143         """
```

```
144         Sets all of the global irradiance values to zero if the corresponding
145         toa value is zero or the registered value is negative
146 
147         """
148         self.data.loc[(self.data['toa'] == 0) |
149                       (self.data['qo'] < 0), 'qo'] = 0
150 
151     def flag_info(self, pesd=True, start_date='', end_date=''):
152         """
153         Returns info about the flagged data
154 
155         :param pesd: If true the data will be percent of errorneous data.
156         :type pesd: boolean, default True
157         :param start_date: Start date on the format 'yyyy-mm-dd hh:mm',
158          recursive need from years.
159         :type start_date: string, default ''
160         :param end_date: End date on the format 'yyyy-mm-dd hh:mm',
161          recursive need from years.
162         :type end_date: string, default ''
163         :returns: Pandas DataFrame
164 
165         """
166         if start_date == '':
167             start_date = self.flags.index[0]
168         if end_date == '':
169             end_date = self.flags.index[-1]
170 
171         if pesd:
172             offset = np.mean(self.flags[start_date:
173                                         end_date].iloc[:, :1]) * 100
174             flagged = offset.append(np.mean(
175                 self.flags[start_date:end_date][self.data[
176                     'toa'] > 0].iloc[:, 1:]) * 100)
177         else:
178             flagged = np.sum(self.flags[start_date:end_date])
179 
180         flag_table = pd.DataFrame(flagged.values, flagged.index)
181         flag_table.columns = ["Flagged data (%)"]
182         flag_table.name = self.name
183         flag_table.index.name = 'Flag type'
184         return flag_table
185 
186     def nan_periods(self, start_date='', end_date=''):
187 
188         """
189         Returns the periods of the timeseries that contains NaN values only
190 
191         :param start_date: Start date on the format 'yyyy-mm-dd hh:mm',
192          recursive need from years.
193         :type start_date: string, default ''
194         :param end_date: End date on the format 'yyyy-mm-dd hh:mm',
195          recursive need from years.
196         :type end_date: string, default ''
197         """
198         if start_date == '':
```

```
199             start_date = self.flags.index[0]
200         if end_date == '':
201             end_date = self.flags.index[-1]
202 
203         result = pd.DataFrame(columns=['Start Date', 'End Date'])
204         df_nan = self.raw[start_date:end_date][
205             self.raw['qo'][start_date:end_date].isnull()]
206         start = True
207         end = False
208         idx = 0
209         for i in range(len(df_nan)):
210             if start:
211                 start_date = df_nan.iloc[i].name
212                 start = False
213             if not df_nan.iloc[i].name == df_nan.iloc[-1].name:
214                 if not (df_nan.iloc[i + 1].name -
215                         df_nan.iloc[i].name == pd.Timedelta('1h')):
216                     end = True
217             else:
218                 end = True
219             if end:
220                 end_date = df_nan.iloc[i].name
221                 result.loc[idx] = [start_date, end_date]
222                 idx += 1
223                 start = True
224                 end = False
225         return result
226 
227     def flag_offset(self):
228         """
229         Tests whether there are negative values less than -12 and
230         if any nightly values are alrger than 6.
231 
232         """
233         self.flags['Offset'] = False
234         test_condition = (((self.data['qo'] > 6) & (self.data['sza'] > 93))
235                           | (self.data['qo'] < -12))
236         self.flags.loc[test_condition, 'Offset'] = True
237 
238     def flag_U1(self):
239         """
240         Tests the upper boundry of the data using
241         Top of Atmosphere modelled data
242         Adds a boolean flag for every data point that fails the test
243 
244         """
245 
246         self.flags['U1'] = False
247         test_condition = self.data['qo'] > self.data['toa']
248         self.flags.loc[test_condition, 'U1'] = True
249 
250     def flag_U2(self):
251         """
252         Tests the upper boundry of the data using Clear Sky modelled data
253         Adds a boolean flag for every data point that fails the test
```

```
254
255         """
256         self.flags['U2'] = False
257         test_condition = (((self.data['qo'] > 1.1 * self.data['clear_sky']) &
258                            (self.data['sza'] < 88))
259                           | ((self.data['qo'] > 2 * self.data['clear_sky']) &
260                              (self.data['sza'] >= 88)))
261         self.flags.loc[test_condition, 'U2'] = True
262
263     def flag_L1(self):
264         """
265         Tests the lower boundry of the data using
266         Top of Atmosphere modeled data and the mean of sunlight data
267         Adds a boolean flag for every data point that fails the test
268
269         """
270         self.flags['L1'] = False
271         sunlight = self.data['toa'] > 0
272         key = [lambda x: x.day, lambda x:  x.month, lambda x: x.year]
273         relative_qo = self.data[sunlight]['qo'] / self.data[sunlight]['toa']
274         day_means = relative_qo.groupby(key).transform(lambda x: x.mean())
275         flag = day_means < 0.03
276
277         # flag and sunlight have different length, but this works since they
278         # both have date index
279         self.flags.loc[sunlight & flag, 'L1'] = True
280
281     def flag_L2(self):
282         """
283         Tests the lower boundry of the data using
284         Top of Atmosphere modelled data
285         Adds a boolean flag for every data point that fails the test
286
287         """
288         self.flags['L2'] = False
289         flag = ((self.data['qo'].values
290                 < (1e-4 * (80 - self.data['sza'].values)
291                    * self.data['toa'].values))
292                 & (self.data['sza'] <= 80))
293         self.flags.loc[flag, 'L2'] = True
294
295     def flag_difference(self):
296         """
297         Tests the difference between two timesteps to ensure no extreme changes
298         Adds a boolean flag for every data point that fails the test
299
300         """
301         self.flags['Difference'] = False
302
303         # Calculate the ratio between measured solar irradiation and
304         # extraterrestrial irradiation
305         relative_qo = self.data['qo'].values / self.data['toa'].values
306
307         difference_qo = np.empty(len(relative_qo))
308         difference_qo[0] = 0
```

```
309             for i in range(len(relative_qo) - 1):
310                 difference_qo[i + 1] = abs(relative_qo[i + 1] - relative_qo[i])
311             flag = (difference_qo >= 0.75) & (self.data['sza'].values < 80)
312             self.flags.loc[flag, 'Difference'] = True
313
314         def flag_consistency(self):
315             """
316             Tests the consistency of a daily value.
317             Adds a boolean flag to every day that does not pass the test.
318
319             """
320             self.flags['Consistency'] = False
321             sunlight = self.data['toa'] > 0
322             key = [lambda x: x.day, lambda x:  x.month, lambda x: x.year]
323             relative_qo = self.data[sunlight]['qo'] / self.data[sunlight]['toa']
324             day_means = relative_qo.groupby(key).transform(lambda x: x.mean())
325             day_stds = relative_qo.groupby(key).transform(lambda x: x.std())
326             flag = (day_stds < 1. / 16. * day_means) | (day_stds > 0.80)
327
328             # This does work even sunlight and flag have different length,
329             # since index is a time series with same frequency
330             self.flags.loc[sunlight & flag, 'Consistency'] = True
331
332         def missing_values(self):
333             """
334             Checks whether there are missing values
335             Adds a boolean flag for every missing value
336
337             """
338             self.flags['Missing values'] = False
339             missing = self.data['qo'].isnull()
340             self.flags.loc[missing, 'Missing values'] = True
341
342         def get_average_year(self, visual_control_dates = [], start_date='', end_date='',
                quality_control=True, leap_day=False):
343             """
344             Returns an average year of the time series.
345
346             The average year procedure:
347
348             - Group all the same hour for all years together
349             - Take mean of the values that are not missing
350             - If all values are missing, take mean of neighbouring values
351             - The first two are same procedure as for CM-SAF PVGIS
352
353             :param visual_control_dates: python list of of all dates that are removed by visual
                    control,
354              write as nested list: [[start_date1, end_date1], [start_date2, end_date2]]
355             :type visual_control_dates: nested list default []
356             :param start_date: Start date on the format 'yyyy-mm-dd hh:mm',
357              recursive need from years.
358             :type start_date: string, default ''
359             :param end_date: End date on the format 'yyyy-mm-dd hh:mm',
360              recursive need from years.
361             :type end_date: string, default ''
```

```
362         """
363
364         if start_date == '':
365             start_date = self.data.index[0]
366         if end_date == '':
367             end_date = self.data.index[-1]
368
369
370         df = self.data['qo'][start_date:end_date].copy()
371
372         if quality_control:
373             df.loc[self.flags.drop(
374                 ['Offset', 'Consistency'], axis=1).any(axis=1)] = np.nan
375
376             for dates in visual_control_dates:
377                 df.loc[dates[0]:dates[1]] = np.nan
378         # Find mean for all years
379         grouped = df.groupby([df.index.month, df.index.day, df.index.hour])
380         avg = pd.DataFrame(data=grouped.aggregate(lambda x: np.nanmean(x)))
381
382         # Need to do this to convert from MultiIndex to DateTimeIndex
383         avg['date'] = avg.index.map(lambda x: datetime(2040, x[0], x[1], x[2]))
384         avg = avg.set_index('date').asfreq('H')
385
386         # Some hours do not have any quality controlled values and are therefore
387         # replaced with average between the two points.
388         ffill = avg.fillna(method='ffill')
389         bfill = avg.fillna(method='bfill')
390         avg['ffill'] = ffill
391         avg['bfill'] = bfill
392         avg['avg'] = avg.mean(axis=1)
393
394         # Remove leap day
395         if not leap_day:
396             avg.drop(avg['2040-02-29'].index, inplace=True)
397
398         # returns only the average
399         return avg['avg']
400
401     def get_pesd(self):
402         """Returns the percentage of erroneous data from automatic control"""
403         return self.flags[self.data['toa'] > 0].drop(
404             ['Offset', 'Consistency'], axis=1).any(axis=1).mean() * 100
405
406     def get_visual(self):
407         """
408         Returns the percentage of data from automatic control flag for
409         visual control
410
411         """
412         return self.flags.loc[:, ['Offset', 'Consistency']].any(axis=1).mean() * 100
413
414
415 if __name__ == '__main__':
416     station = BioforskStation('Aas')
```

B.2 Statisitics

```
1  import numpy as np
2
3  def mbd(model, measure):
4      return (model - measure).mean() / measure.mean() * 100
5
6  def mae(model, measure):
7      return abs((model - measure)).mean() / measure.mean() *
           100
8
9  def rmsd(model, measure):
10     return np.sqrt(((model - measure) ** 2).mean()) /
           measure.mean() * 100
11
12 def relative_change(x, ref):
13     return (x - ref) / ref * 100
```

B.3 Example run with tests and average year calculation for Ås

```
1  # Import class
2  from solqc.bioforskstation import BioforskStation
3
4  # Create Bioforsk station
5  station = BioforskStation('Aas')
6
7  # Automatic tests
8  station.flag_offset()
9  station.zero_out()
10 station.missing_values()
11 station.flag_U1()
12 station.flag_U2()
13 station.flag_L1()
14 station.flag_L2()
```

```
15 station.flag_difference()
16 station.flag_consistency()
17
18 # Perform visual control
19 visual_control_dates = [['1995-05-25', '1995-05-25']]
20
21 # Get average year (hourly data)
22 avg_year = station.get_average_year(visual_control_dates)
23
24 # Get monthly averages
25 avg_year.resample('D', how='sum').resample('M')
```

B.4 Stations config file

```
[Aas]
lon = 10.781989
lat = 59.660468
hgt = 94
id = 5

[Alvdal]
lon = 10.62687
lat = 62.10944
hgt = 478
id = 10

[Apelsvoll]
lon = 10.86952
lat = 60.70024
hgt = 255
id = 11

[Boe]
lon = 9.02859
lat = 59.4175
hgt = 105
```

```
id = 13

[Faavang]
lon = 10.1872
lat = 61.45822
hgt = 184
id = 17

[Fagklim]
lon = 10.781989
lat = 59.660468
hgt = 94
id = 999

[Gausdal]
lon = 10.25878
lat = 61.22468
hgt = 375
id = 18

[Gjerpen]
lon = 9.57805
lat = 59.22684
hgt = 41
id = 19

[Gran]
lon = 10.55906
lat = 60.35575
hgt = 245
id = 20

[Gvarv]
lon = 9.21189
lat = 59.38223
hgt = 46
```

```
id = 21

[Hokksund]
lon = 9.89166
lat = 59.76152
hgt = 15
id = 23

[Honefoss]
lon = 10.2661
lat = 60.14032
hgt = 126
id = 25

[Ilseng]
lon = 11.20298
lat = 60.80264
hgt = 182
id = 26

[Kise]
lon = 10.80569
lat = 60.77324
hgt = 129
id = 27

[Lier]
lon = 10.2604
lat = 59.79005
hgt = 39
id = 30

[Loeken]
lon = 9.06302
lat = 61.12183
hgt = 527
```

```
id = 33

[Rakkestad]
lon = 11.39042
lat = 59.38824
hgt = 102
id = 37

[Ramnes]
lon = 10.23923
lat = 59.38081
hgt = 39
id = 38

[Roverud]
lon = 12.09144
lat = 60.25378
hgt = 150
id = 40

[Sande]
lon = 10.22339
lat = 59.6162
hgt = 35
id = 42

[Tjoelling]
lon = 10.12513
lat = 59.04641
hgt = 19
id = 50

[Tomb]
lon = 10.81449
lat = 59.31893
hgt = 12
```

```
id = 52

[Aarnes]
lon = 11.39342
lat = 60.1268
hgt = 162
id = 53

[Oesaker]
lon = 11.04221
lat = 59.31936
hgt = 45
id = 118
```

Appendix C

Monthly averages

The average monthly solar irradiation for quality controlled measured values and solar irradiation databases for each station are shown in tables C.0.1, C.0.2, C.0.3, and C.0.4. The unit is in Wh/m²/day.

Table C.0.1 – The monthly averages for quality controlled measured values and solar irradiation databases for Ås. The unit is in Wh/m²/day.

	Quality controlled	Satel-Light	NASA SSE	WRF	Meteonorm 7	Classic PVGIS	CM-SAF PVGIS
Jan	303	314	390	158	258	274	378
Feb	919	1007	1090	612	779	815	944
Mar	2309	2266	2310	1737	2194	1790	2470
Apr	3505	3275	3700	3466	3700	3240	3780
May	5138	4980	5320	5309	5097	4630	5140
Jun	5666	4976	5540	6032	5700	5260	5670
Jul	5227	5403	5500	5632	5194	4840	5000
Aug	4123	4232	4220	4053	3677	3630	4040
Sep	2600	2628	2760	2262	2600	2280	2640
Oct	1133	1170	1300	753	1065	1050	1230
Nov	371	401	540	189	333	404	486
Dec	168	222	240	65	129	195	235

Table C.0.2 – The monthly averages for quality controlled measured values and solar irradiation databases for Lier. The unit is in $Wh/m^2/day$.

	Quality controlled	Satel-Light	NASA SSE	WRF	Meteonorm 7	Classic PVGIS	CM-SAF PVGIS
Jan	288	338	390	194	226	261	362
Feb	925	993	1090	734	743	784	883
Mar	2419	2191	2310	1926	2065	1720	2410
Apr	3643	3233	3700	3646	3600	3210	3620
May	5240	4878	5320	5269	4839	4530	4860
Jun	5805	4812	5540	5982	5467	5160	5290
Jul	5204	5194	5500	5530	4871	4730	4660
Aug	4108	4093	4220	4050	3419	3520	3750
Sep	2590	2473	2760	2296	2467	2190	2480
Oct	1161	1121	1300	861	1032	1010	1190
Nov	371	361	540	242	300	376	460
Dec	157	218	240	87	129	185	221

Table C.0.3 – The monthly averages for quality controlled measured values and solar irradiation databases for Ramnes. The unit is in $Wh/m^2/day$.

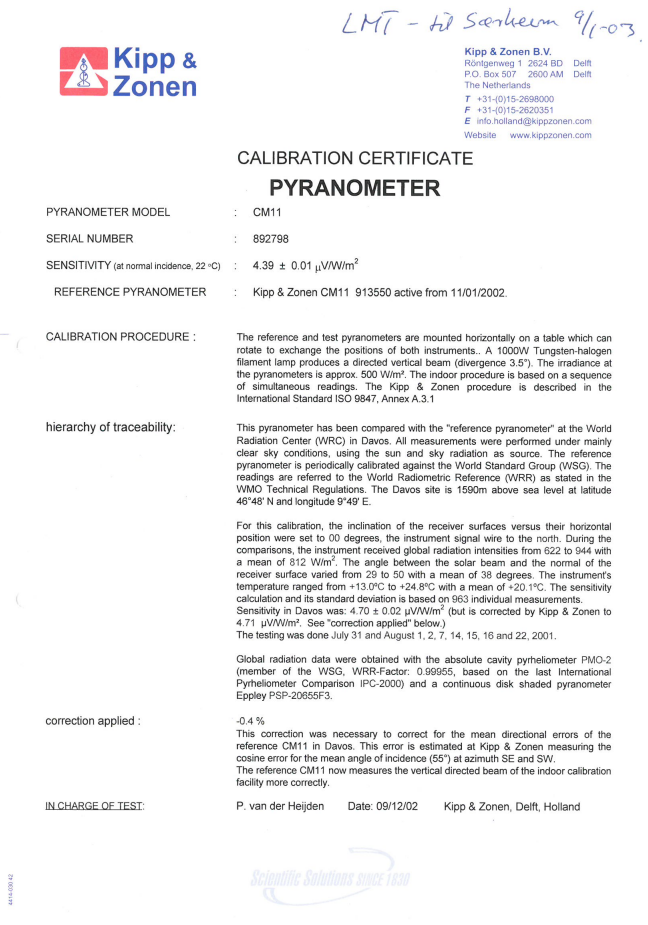
	Quality controlled	Satel-Light	NASA SSE	WRF	Meteonorm 7	Classic PVGIS	CM-SAF PVGIS
Jan	305	345	390	208	258	266	374
Feb	926	1038	1090	733	779	800	951
Mar	2364	2210	2310	1914	2161	1740	2530
Apr	3582	3256	3700	3666	3733	3200	3880
May	5141	4947	5320	5350	4935	4550	5220
Jun	5712	5028	5540	6110	5667	5190	5830
Jul	5220	5471	5500	5694	5097	4790	5110
Aug	4107	4247	4220	4154	3613	3570	4120
Sep	2590	2565	2760	2312	2533	2250	2680
Oct	1172	1177	1300	875	1129	1030	1270
Nov	383	391	540	255	367	392	485
Dec	186	227	240	96	161	190	238

Table C.0.4 – The monthly averages for quality controlled measured values and solar irradiation databases for Tomb. The unit is in $Wh/m^2/day$.

	Quality controlled	Satel-Light	NASA SSE	WRF	Meteonorm 7	Classic PVGIS	CM-SAF PVGIS
Jan	287	313	390	138	290	271	393
Feb	846	1007	1090	516	814	813	969
Mar	2287	2265	2310	1650	2226	1750	2530
Apr	3601	3274	3700	3450	3767	3190	3910
May	5300	4979	5320	5408	5194	4570	5410
Jun	5858	4975	5540	6216	5800	5190	6030
Jul	5408	5402	5500	5771	5387	4810	5390
Aug	4293	4231	4220	4131	3871	3590	4310
Sep	2676	2627	2760	2286	2667	2290	2780
Oct	1169	1169	1300	731	1161	1050	1300
Nov	361	401	540	173	367	401	481
Dec	168	222	240	60	161	199	237

Appendix D

Calibration certificates



LMT - Hd Saarbrem 9/1-03.

Kipp & Zonen

Kipp & Zonen B.V.
Röntgenweg 1 2624 BD Delft
P.O. Box 507 2600 AM Delft
The Netherlands
T +31-(0)15-2698000
F +31-(0)15-2620351
E info.holland@kippzonen.com
Website www.kippzonen.com

CALIBRATION CERTIFICATE

PYRANOMETER

PYRANOMETER MODEL	:	CM11
SERIAL NUMBER	:	892798
SENSITIVITY (at normal incidence, 22 °C)	:	4.39 ± 0.01 μV/W/m²
REFERENCE PYRANOMETER	:	Kipp & Zonen CM11 913550 active from 11/01/2002.

CALIBRATION PROCEDURE :

The reference and test pyranometers are mounted horizontally on a table which can rotate to exchange the positions of both instruments.. A 1000W Tungsten-halogen filament lamp produces a directed vertical beam (divergence 3.5°). The irradiance at the pyranometers is approx. 500 W/m². The indoor procedure is based on a sequence of simultaneous readings. The Kipp & Zonen procedure is described in the International Standard ISO 9847, Annex A.3.1

hierarchy of traceability:

This pyranometer has been compared with the "reference pyranometer" at the World Radiation Center (WRC) in Davos. All measurements were performed under mainly clear sky conditions, using the sun and sky radiation as source. The reference pyranometer is periodically calibrated against the World Standard Group (WSG). The readings are referred to the World Radiometric Reference (WRR) as stated in the WMO Technical Regulations. The Davos site is 1590m above sea level at latitude 46°48' N and longitude 9°49' E.

For this calibration, the inclination of the receiver surfaces versus their horizontal position were set to 00 degrees, the instrument signal wire to the north. During the comparisons, the instrument received global radiation intensities from 622 to 944 with a mean of 812 W/m². The angle between the solar beam and the normal of the receiver surface varied from 29 to 50 with a mean of 38 degrees. The instrument's temperature ranged from +13.0°C to +24.8°C with a mean of +20.1°C. The sensitivity calculation and its standard deviation is based on 963 individual measurements.
Sensitivity in Davos was: 4.70 ± 0.02 μV/W/m² (but is corrected by Kipp & Zonen to 4.71 μV/W/m². See "correction applied" below.)
The testing was done July 31 and August 1, 2, 7, 14, 15, 16 and 22, 2001.

Global radiation data were obtained with the absolute cavity pyrheliometer PMO-2 (member of the WSG, WRR-Factor: 0.99955, based on the last International Pyrheliometer Comparison IPC-2000) and a continuous disk shaded pyranometer Eppley PSP-20655F3.

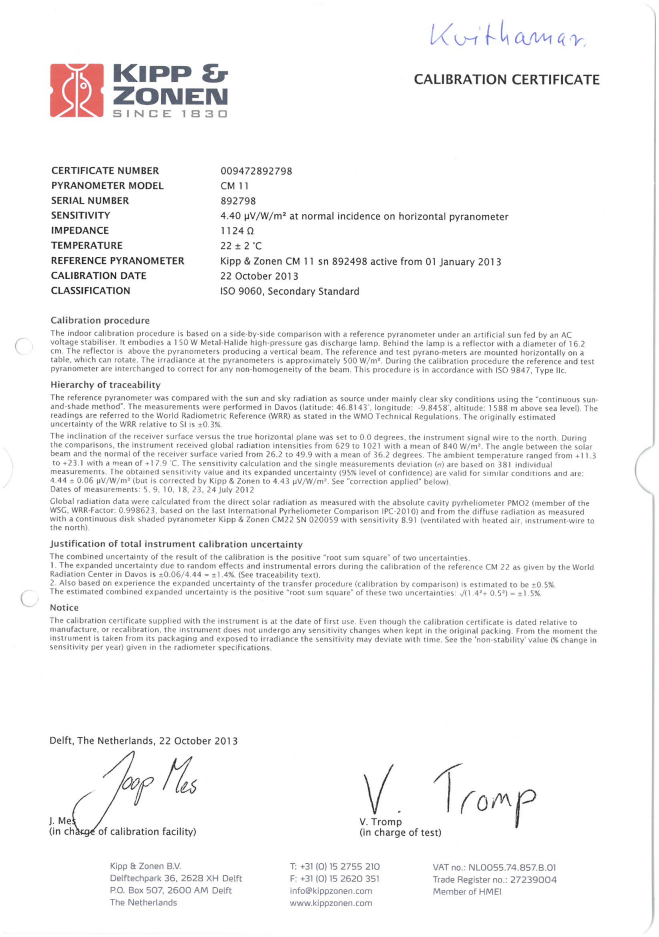
correction applied :

-0.4 %
This correction was necessary to correct for the mean directional errors of the reference CM11 in Davos. This error is estimated at Kipp & Zonen measuring the cosine error for the mean angle of incidence (55°) at azimuth SE and SW.
The reference CM11 now measures the vertical directed beam of the indoor calibration facility more correctly.

IN CHARGE OF TEST: P. van der Heijden Date: 09/12/02 Kipp & Zonen, Delft, Holland

Scientific Solutions since 1830

Figure D.0.1 – The first calibration certificate for the Kipp & Zonen CM11 pyranometer with serial number 892798. Displays a sensitivity of $4.39 \pm 0.01\,\mu\mathrm{V/W/m^2}$. The calibration was performed on 2002/12/09.



Kwithaman

KIPP & ZONEN
SINCE 1830

CALIBRATION CERTIFICATE

CERTIFICATE NUMBER	009472892798
PYRANOMETER MODEL	CM 11
SERIAL NUMBER	892798
SENSITIVITY	4.40 µV/W/m² at normal incidence on horizontal pyranometer
IMPEDANCE	1124 Ω
TEMPERATURE	22 ± 2 °C
REFERENCE PYRANOMETER	Kipp & Zonen CM 11 sn 892498 active from 01 January 2013
CALIBRATION DATE	22 October 2013
CLASSIFICATION	ISO 9060, Secondary Standard

Calibration procedure

The indoor calibration procedure is based on a side-by-side comparison with a reference pyranometer under an artificial sun fed by an AC voltage stabiliser. It embodies a 150 W Metal-Halide high-pressure gas discharge lamp. Behind the lamp is a reflector with a diameter of 16.2 cm. The reflector is above the pyranometers producing a vertical beam. The reference and test pyrano-meters are mounted horizontally on a table, which can rotate. The irradiance at the pyranometers is approximately 500 W/m². During the calibration procedure the reference and test pyranometer are interchanged to correct for any non-homogeneity of the beam. This procedure is in accordance with ISO 9847, Type IIc.

Hierarchy of traceability

The reference pyranometer was compared with the sun and sky radiation as source under mainly clear sky conditions using the "continuous sun-and-shade method". The measurements were performed in Davos (latitude: 46.8143°, longitude: -9.8458°, altitude: 1588 m above sea level). The readings are referred to the World Radiometric Reference (WRR) as stated in the WMO Technical Regulations. The originally estimated uncertainty of the WRR relative to SI is ±0.3%.

The inclination of the receiver surface versus the true horizontal plane was set to 0.0 degrees, the instrument signal wire to the north. During the comparisons, the instrument received global radiation intensities from 629 to 1021 with a mean of 840 W/m². The angle between the solar beam and the normal of the receiver surface varied from 26.2 to 45.9 with a mean of 36.2 degrees. The ambient temperature ranged from +11.3 to +23.1 with a mean of +17.9 °C. The sensitivity calculation and the single measurements deviation (σ) are based on 381 individual measurements. The obtained sensitivity value and its expanded uncertainty (95% level of confidence) are valid for similar conditions and are: 4.44 ± 0.06 µV/W/m² (but is corrected by Kipp & Zonen to 4.43 µV/W/m². See "correction applied" below).
Dates of measurements: 5, 9, 10, 18, 23, 24 July 2012

Global radiation data were calculated from the direct solar radiation as measured with the absolute cavity pyrheliometer PMO2 (member of the WSG, WRR-Factor: 0.998623, based on the last International Pyrheliometer Comparison IPC-2010) and from the diffuse radiation as measured with a continuous disk shaded pyranometer Kipp & Zonen CM22 SN 020059 with sensitivity 8.91 (ventilated with heated air, instrument-wire to the north).

Justification of total instrument calibration uncertainty

The combined uncertainty of the result of the calibration is the positive "root sum square" of two uncertainties.
1. The expanded uncertainty due to random effects and instrumental errors during the calibration of the reference CM 22 as given by the World Radiation Center in Davos is ±0.06/4.44 = ±1.4%. (See traceability text).
2. Also based on experience the expanded uncertainty of the transfer procedure (calibration by comparison) is estimated to be ±0.5%.
The estimated combined expanded uncertainty is the positive "root sum square" of these two uncertainties: √(1.4²+ 0.5²) = ±1.5%.

Notice

The calibration certificate supplied with the instrument is at the date of first use. Even though the calibration certificate is dated relative to manufacture, or recalibration, the instrument does not undergo any sensitivity changes when kept in the original packing. From the moment the instrument is taken from its packaging and exposed to irradiance the sensitivity may deviate with time. See the 'non-stability' value (% change in sensitivity per year) given in the radiometer specifications.

Delft, The Netherlands, 22 October 2013

Joop Mes

J. Mes
(in charge of calibration facility)

V. Tromp

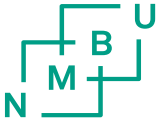
V. Tromp
(in charge of test)

Kipp & Zonen B.V.
Delftechpark 36, 2628 XH Delft
P.O. Box 507, 2600 AM Delft
The Netherlands

T: +31 (0) 15 2755 210
F: +31 (0) 15 2620 351
info@kippzonen.com
www.kippzonen.com

VAT no.: NL0055.74.857.B.01
Trade Register no.: 27239004
Member of HMEI

Figure D.0.2 – The second calibration certificate for the Kipp & Zonen CM11 pyranometer with serial number 892798. Displays a sensitivity of $4.40\,\mu\mathrm{V/W/m^2}$. The calibration was performed on 2013/10/22.



Norwegian University
of Life Sciences

Postboks 5003
NO-1432 Ås, Norway
+47 67 23 00 00
www.nmbu.no