Creating a Global Surface Solar Irradiance Product from OMI and SCIAMACHY

Joris van Lammeren
University of Utrecht & KNMI

Supervisors:

P. Wang & P. Stammes (KNMI)

M. Krol & W. van Sark (Univeristy of Utrecht)

July 6, 2018

Abstract

The solar energy market has grown rapidly over the last few years. Since solar energy depends strongly on weather and climate, solar resource assessment is essential for photovoltaic power plants. Accurate solar radiation climatology data sets are the foundation of photovoltaic power plant performance models. In this study data from two polar-orbiting satellite spectrometers (OMI and SCIAMACHY) was used to create a daily mean-monthly mean surface solar irradiance product. This product shows good agreement with surface measurements and a comparable surface solar irradiance product, CERES. The created surface solar irradiance product is used to analyze trends, time series and anomalies, globally and locally. It is found that a difference in aerosol optical depth input can affect the local surface solar irradiance up to 60 W/m^2, but globally differences aerosol optical depth input does not affect the surface solar irradiance. There is no clear trend in the created surface solar irradiance product in the overlapping time period of the two satellites.

Contents

1 Introduction

As the world is transitioning towards more sustainable ways of generating energy, the solar energy market has grown rapidly over the last few years. In order to predict the performance of photo-voltaic power plants an accurate solar resource assessment is crucial. In this research a surface solar irradiance daily mean-monthly mean product is derived from the surface solar irradiance (SSI) measurements from two satellites, OMI and SCIAMACHY.

OMI (Ozone Measuring Instrument) is a Dutch-Finnish instrument on board of the EOS-Aura satellite from NASA which is part of the A-train. The A-train consists of six satellites that are only a few minutes apart and crosses the equator each day approximately around 13:30. Aura (with OMI on board) crosses the equator as the last satellite in the constellation. OMI is a UV-visible spectrometer with a wavelenght range from 270 to 500 nm, the spectral resolution of OMI is 0.42 to 0.63 nm. This makes OMI not just capable of measuring ozone, but also other gases and aerosols. OMI has a wide enough swath to capture a global image in one day with a resolution of 13 x 24 km^2. OMI was launched in July 2004 and is still collecting data. SCIAMACHY (SCanning Imaging Absorption SpecroMeter for Atmospheric CHartographY) is an instrument onboard of ENVISAT (Environmental Satellite) and was active from 2002 until 2012. SCIAMACHY is a spectrometer with a wavelength range from 240 to 2380 nm with a spectral resolution of 0.2 to 1.5 nm. The goal of SCIAMACHY was to improve our knowledge of the amounts and distribution of several gases in the atmosphere. SCIAMACHY has a more narrow swath width than OMI and will reach global coverage after six days of measurements.

This report describes how these two satellite measurements can be combined into a daily mean-monthly mean SSI product. First both satellite SSI products will be individually validated using surface measurements from the Baseline Surface Radiation Network (BSRN). After this the combined product will be derived, first the daily mean and from the daily mean the monthly mean will be calculated, hence the daily mean-monthly mean. Furthermore the combined product is validated with surface measurements from BSRN-stations and using a comparable data-set from CERES (Clouds and the Earth's Radiant Energy System). After the validation, the created data-set is analyzed, looking for trends and anomalies. The research questions asked in this research are:

- Is it possible to create a reliable global surface solar irradiance product from two polar-orbitting spectrometers?
- What is the influence of clouds on the surface solar irradiance?
- What is the influence of aerosols on the surface solar irradiance?
- Are there significant trends in the surface solar irradiance, regional or global?

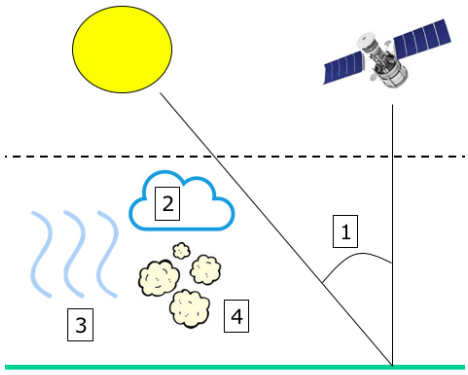


Figure 1: The 4 most important influences on SSI, 1: Solar Zenith Angle, 2: Clouds, 3: Atmospheric Gases, 4: Aerosols.

When researching the SSI there are four major components that influence the amount of radiation that reaches the surface of the earth. Figure 1 shows those four components. The first component is the solar zenith angle. This is the angle between the normal of the earth's surface and the position of the sun. If the solar zenith angle becomes smaller the SSI increases and vice versa. Clouds, atmospheric gases and aerosols affect the SSI by their reflecting, scattering and absorbing characteristics. Figure 2 shows two adjacent days of surface observations to demonstrate the impact of clouds on SSI. The first day, the blue line, is a cloud-free day and shows a parabola with the maximum at 720 minutes (12 hours) after midnight, just as expected. The second day, the red line, is a cloudy day. The graph shows how large the influence of clouds can be on the SSI, because it deviates significantly from the parabola from the day before. At the time where the cloud is present, the SSI is even a bit higher than the cloud-free SSI. This is due to the multiple scattering and reflection between scattered clouds.

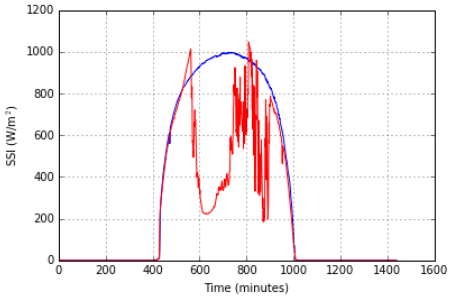


Figure 2: SSI of February 1, 2005 (cloud-free, blue) and February 2, 2005 (cloudy, red) at Alice Springs, Australia

2 Theoretical background

2.1 Solar irradiance and scattering

To understand the amount of solar irradiance that reaches the surface of the earth, the processes that the light will endure during the journey from the sun to the earth's surface have to be understood. What happens to the irradiance from the sun before it reaches the top of the atmosphere. The amount of irradiance that reaches the top of the atmosphere depends on two things. The first one is the distance between the sun and the earth. The distance between the sun and the earth has long term variabilities (Milankovich cycles) of thousands of years and short term variabilities within a year. This research does not contain the long term change in SSI, therefore the Milankovich cycles will not be discussed. The short term variability in the distance between the earth and the sun is created by the movement of the earth in an ellipse around the sun, see figure 3.

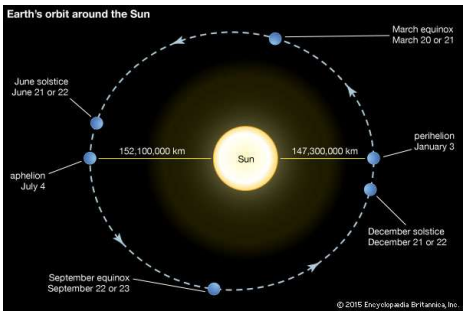


Figure 3: The orbit of the earth around the sun. From *Brittanica encyclopedia*, 2015

Figure 3 shows that the distance between the earth and the sun is minimal on January 3 and maximal on July 4. Therefore, the solar irradiance at the top of the atmosphere varies between 1415 W/m^2 on January 3 and 1321 W/m^2 on July 4 (Sengupta et al., 2015).

The second aspect that influences the magnitude of the solar irradiance at the top of the atmosphere on a short term is the appearance of regions of reduced surface temperature on the sun called sunspots. The larger the area covered by sunspots the smaller the solar irradiance is. The amount of sunspots on the sun has an 11-year periodicity and therefore induces an 11-year periodicity in the solar irradiance (Sengupta et al., 2015). The 11-year cycle is featured in figure 4

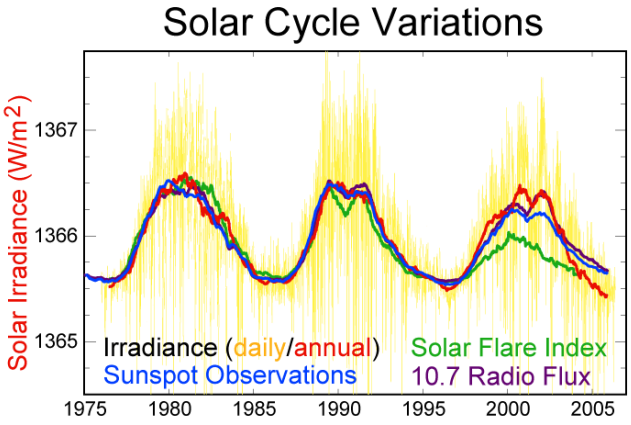


Figure 4: The relataion between solar irradiance and sunspots. From *R. A. Rhode*, 2008

The important feature from figure 4 is that the relation between the observed sunspots and the measured annual solar irradiance at the top of the atmosphere is very clear.

Now that it is clear what determines the solar irradiance that reaches the top of the atmosphere it is necessary to find out what happens when the solar irradiance enters the atmosphere. When the light enters the atmosphere it will stumble upon particles. These particles can absorb and scatter the sunlight. As a consequence the spectrum of the light that reaches the surface of the earth differs from the spectrum at the top of the atmosphere. Figure 5 shows the spectrum of sunlight at the top of the atmosphere and the spectrum at sea level, due to absorption and scattering.

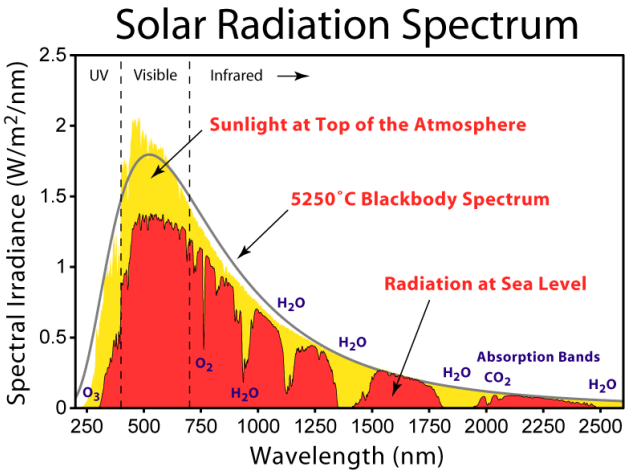


Figure 5: The relataion between solar irradiance and sunspots. From *R. A. Rhode*, 2007

Due to the interaction between light and a particle, light will scatter (Young and Freedman, 2015). The way of scattering is determined by the size of the particle and the wavelength of the light. There are three ways of scattering, Rayleigh scattering, Mie scattering and Geometric scattering. When the particle is much smaller than the wavelength the way of scattering will be Rayleigh scattering. When the size of the particle is in the same order of magnitude the way of scattering will be Mie scattering and when the particle is much larger than the wavelength Geometric scattering will occur (D.W. Hahn, 2009). In the atmosphere there will be Rayleigh scattering, because the particles of the gases in the atmosphere are much smaller than the wavelength of sunlight. Droplets in clouds have a similar size as the wavelength of sunlight, therefore when sunlight reaches a cloud, Mie scattering will occur.

If the sunlight collides with a much smaller particle than the wavelength, Rayleigh scattering will occur. Equation 1 shows the intensity I of the scattered light (D.W. Hahn, 2009).

$$I = I_0 \frac{1 + cos^2\theta}{2R^2} (\frac{2\pi}{\lambda})^4 (\frac{n^2 - 1}{n^2 + 2})^2 (\frac{d}{2})^6 \tag{1}$$

In this equation I_0 is the initial intensity, θ the scattering angle, n is the refractive index of the particle, d the diameter of the particle, R is the distance between the particle and the observer and λ is the wavelength of the light. Equation 1 shows that the intensity of the scattered light highly depends on the wavelength of the light. As can be seen from the equation, the shorter the wavelength, the stronger the scattering. This means for the visible light regime that blue light is scattered the strongest, hence a cloud-free sky is blue when not looking directly at the sun.

Mie scattering is the scattering that occurs when the size of the particle has the same order of magnitude as the wavelength. In the case of Mie scattering the scattering does not depend as heavily on the wavelength as Rayleigh scattering ($1/\lambda$). This means that every wavelength is about equally scattered and therefore the light observed from clouds is white (D.W. Hahn, 2009).

For this research the molecular absorption of light is used to determine the surface solar irradiance. OMI and SCIAMACHY are both spectrometers that measure the light spectrum returning from the atmosphere. In order to know how much light reaches the surface of the earth it is necessary to know whether there are clouds present. To measure the presence of clouds, the data from known oxygen absorption bands was used. This is because oxygen is known to be a well mixed gas and therefore the amount of absorption is linear with the distance that the solar irradiance travelled before reaching the spectrometer on the satellite (Stammes et al., 2008). Therefore the data used from OMI is the reflection at 477 nm (O_2-O_2 absorption) and from SCIAMACHY the reflection at 760 nm (O_2-A absorption). Using these reflectances a cloud fraction needs to be derived and from that cloud fraction a surface solar irradiance can be calculated.

2.2 Lambertian cloud model for OMI

The algorithm used to calculate SSI for OMI is based on a Lambertian cloud model. This is necessary because the pixels of OMI are too big to capture each single cloud. Therefore there are very little pixels that are completely without clouds. So it is assumed that pixel has a cloudy fraction c and a cloud-free fraction $(1-c)$. There are four cloud-parameters that have the most influence on the reflectance at the top of the atmosphere, the cloud fraction c, the cloud optical thickness τ, the thickness of the cloud h and the height of the cloud z_c. Because the satellites have a limited resolution and do not provide us with information on parameters like cloud vertical structure and particle shape and size the model is simplified. So therefore instead of a cloud with thickness h a Lambertian reflector is assumed with a fixed albedo. It is assumed that no light will be transmitted through the cloud. With these assumptions an effective cloud fraction c_{eff} can be calculated which agrees with measured reflected SSI.

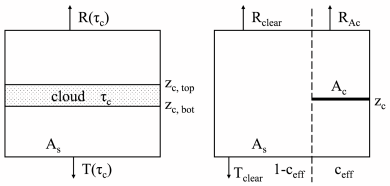


Figure 6: Left: a scattering cloud model, right: a Lambertian cloud model. From *Stammes et al.*, 2008

Figure 6 shows a scattering cloud model and a lambertian cloud model. It shows that in a scattering cloud model assumes either a cloudy pixel or a non-cloudy pixel. The cloud in the scattering cloud model has a certain optical thickness, a thickness and a height. In the simplified Lambertian model the fraction c_{eff} is covered by a Lambertian reflector with a fixed albedo.

The reasoning for the cloud albedo is as follows. There is assumed to be zero transmission through the Lambertian reflector and zero absorption. This means that the missing transmission from the Lambertian reflector has to be compensated by the part of the pixel without the reflector, $(1 - c_{eff})$. In order to calculate the Lambertian cloud albedo the case of a fully clouded pixel is studied. The goal is to find the reflection and transmission of a scattering cloud, R_S and T_S , using the Lambertian cloud model, $R_L(A_c)$. So when a fully clouded pixel is assumed this means $c = 1$ (not $c_{eff} = 1$). The equations that are obtained are:

$$R_s = R_L(A_c) = (1 - c_{eff})R_{clear} + c_{eff}R_{A_c} \tag{2}$$

$$T_s = T_L = (1 - c_{eff})T_{clear} \tag{3}$$

Note that the second equations does not have a component in the cloudy fraction of the pixel, because it is assumed that the transmission is zero in that fraction. Combining these equations yields:

$$R_{A_c} = \frac{R_s T_{clear} - T_s R_{clear}}{T_{clear} - T_s} \tag{4}$$

Now that R_{A_c} is isolated from c_{eff}. R_{A_c} can be calculated when R_s and T_s are obtained from a model run. In *Stammes et al.*, 2008 the DAK (Double-Adding KNMI) model was run for multiple cloud optical thicknesses.

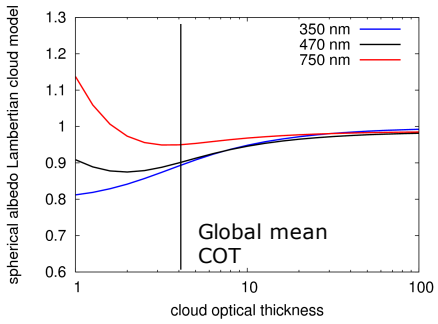


Figure 7: The output of the DAK model run in order to find the cloud albedo for the Lambertian cloud model, From *Stammes et al.*, 2008

The result is shown in figure 7. The O_2-O_2-method uses the wavelength of 477 nm. So when observing figure 7 and following the line of the global mean cloud optical thickness to where it intersects with the black line of 470 nm, it shows that the albedo of the Lambertian reflector in the cloud model has to be 0.9.

2.3 FRESCO for SCIAMACHY

In the case of SCIAMACHY, FRESCO (Fast REtrieval Scheme for Clouds from the Oxygen A-band) is used to retrieve an effective cloud fraction and cloud pressure. In FRESCO it is assumed that reflectance only occurs at the top of the cloud and at the surface, both are assumed to be Lambertian reflectors. The reflectance at the top of the atmosphere is then given by:

$$R_{TOA} = cT_cA_c + cR_c + (1-c)T_sA_s + (1-c)R_s \tag{5}$$

The factor cT_cA_c is the reflectance at the top of the cloud and $(1-c)T_sA_s$ is the reflectance at the surface. cR_c and $(1-c)R_s$ represent the single Rayleigh scattering above the cloud and above the surface. T_c, R_c, T_s and R_s are obtained from look-up tables. The surface albedo A_s is also known. The algorithm can not calculate both the cloud fraction c and the cloud albedo A_c and therefore A_c is assumed to be a fixed value. This fixed value was found by finding the cloud albedo for which the SCIAMACHY SSI performs best (see figure 7). This is a cloud albedo of 0.95. So now c can be calculated by measuring R_{TOA} .

2.4 Heliosat Method for both satellites

The satellite measures the reflectance of the clouds, which is used to calculate the effective cloud fraction using the lambertian cloud model or FRESCO. Once the effective cloud fraction is obtained it is used to obtain a full sky SSI from a clear sky SSI. For this the Heliosat method is used. The heliosat method assumes that te cloud cover index is the dominant factor in blocking the radiation from

the sun. From Wang et al., (2014) the definition of the clear sky index is used. Which is defined as:

$$SSI = SSI_{clr} \cdot k \tag{6}$$

SSI_{clr} is calculated using the MAGIC algorithm (Mueller et al., 2004) and is the clear sky surface solar irradiance. The clear sky surface solar irradiance is the SSI without the effects of clouds. So the full sky surface solar irradiance is the SII with the effects of clouds taken into account. In order to calculate the clear sky index k the relation with the cloud cover index is needed. Hammer et al., (2003) and Rigollier et al., (2004) found it to be:

$$-0.2 < c_{eff} < 0 \rightarrow k = 1.2 \tag{7}$$

$$0 \leq c_{eff} \leq 0.8 \rightarrow k = 1 - c_{eff} \tag{8}$$

$$0.8 < c_{eff} \leq 1.1 \rightarrow k = 2.0667 - 3.6667 c_{eff} + 1.6667 c_{eff}^2 \tag{9}$$

$$1.1 < c_{eff} \rightarrow k = 0.05 \tag{10}$$

These equations were found by comparing surface measurements with sattelite observations at several locations. Once k is determined the full sky SSI can be calculated by using equation 6.

3 Data

3.1 OMI

The data that is used from the OMI instrument is the clear sky SSI and the full sky SSI, which are calculated as explained in the theory section. These values are used to calculate the transmission coefficient as explained in the method section. An example of the data used from OMI for this research is shown in figure 8. OMI has an overpass time around 13:30 and orbits measures from south to north. The orbid width is 2600 kilometers. The grid size is 0.25 x 0.25 degrees, which is 13 x 24 km^2 at nadir, and only the data from 60 degrees north until 60 degrees south is used. This is because the measurements of the cloud fraction does not function optimally over an area with a high albedo.

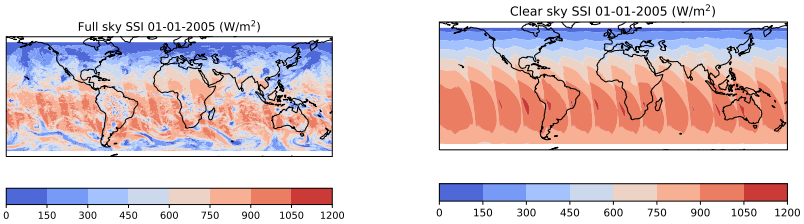


Figure 8: Global image of daily measurement of OMI, with on the left the full sky SSI and on the right the clear sky SSI

Figure 8 shows that OMI provides a full global image every day. OMI does so in fifteen overpasses, which can be seen in the figure by counting the bands.

3.2 SCIAMACHY

The same data output as OMI is used from SCIAMACHY. So the full sky SSI and the clear sky SSI. SCIAMACHY has an overpass time at 10:00 in the morning local time and orbits from north to south while measuring. SCIAMACHY has nadir and limb measurements during an orbit and the gaps in an orbit are due to limb measurements. The orbit width is 960 km and the pixel size is 30x 60 kilometers. The obtained grid size is also 0.25 x 0.25 degrees and again the data from 60 degrees north to 60 degrees south is used. Figure 9 shows an example of the daily data from SCIAMACHY.

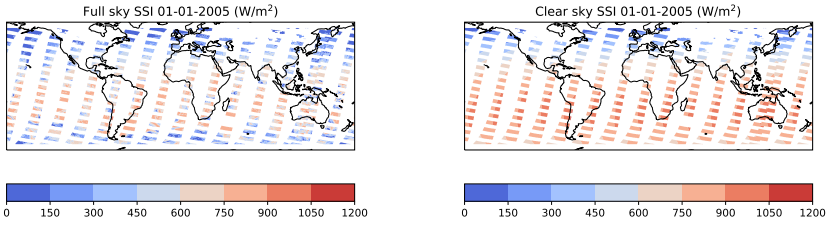


Figure 9: Global image of daily measurement of SCIAMACHY, with on the left the full sky SSI and on the right the clear sky SSI

In contrast to OMI, SCIAMACHY does not provide a full global image every day. It takes about 6 days of data to obtain a full global image from SCIAMACHY. By counting the bands it can be seen that SCIAMACHY also has fifteen overpasses each day.

3.3 CLARA

CLARA is a dataset created by CM SAF (The Satellite Application Facility on Climate Monitoring). The method that is used to derive this dataset uses the MAGIC radiative transfer model to include the properties of the atmosphere. Which is the same model used for deriving the OMI and SCIAMACHY SSI. Figure 10 shows the flowscheme of the CLARA model for the daily mean clear sky SSI. This is the product that will be used in this research to calculate the full sky daily mean SSI based on OMI and SCIAMACHY.

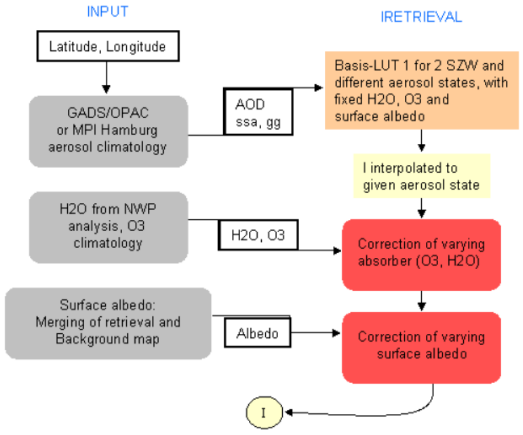


Figure 10: Flowchart of the CLARA clear sky SSI product, from *Mueller et al.*, 2009

The input data for the radiative transfer model is partly from satellites, like the TOA albedo and the surface albedo. Aerosol and ozone input is taken from climatological datasets and the water vapour is analysis data by the German weather service. For more information on the CLARA algorithm see *Mueller et al*, 2009 or *CM SAF*, 2016. Figure 11 shows an example of the CLARA daily mean clear sky SSI. The data has been reduced to 60 degrees north to 60 degrees south, because that is range of data that will be used from OMI and SCIAMACHY.

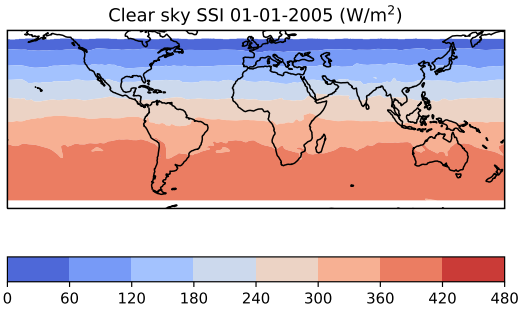


Figure 11: Global clear sky daily mean SSI from CLARA on January 1 2005

Figure 11 shows obviously that the SSI in january is higher on the southern hemisphere, than on the northern hemisphere, as to be expected. Compared to figures 9 and 8 the CLARA dataset does not contain the bands that the datasets created by the satellite data contain, due to the orbits of the satellites.

3.4 CERES

The dataset of CERES is used to validate the full sky SSI product that are created in this research. Just like the CLARA dataset, the CERES dataset is based on a radiative transfer model. In figure 12 the flowscheme of CERES is shown.

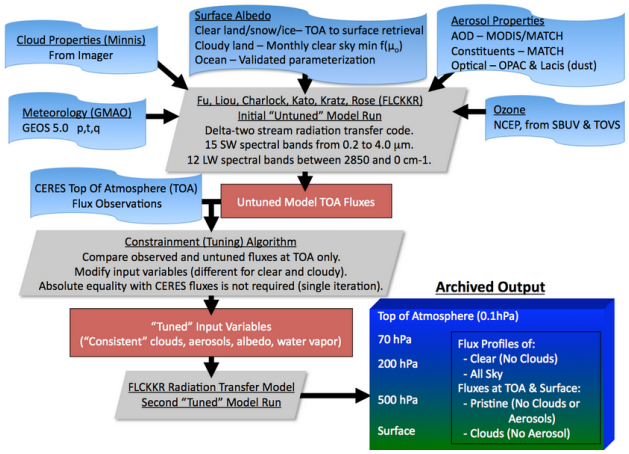


Figure 12: Flowchart of the CERES full sky SSI product, From *Loeb et al.*, 2001

The biggest difference between the input of CERES and CLARA is that CERES uses measurements for the aerosol input from MODIS (Moderate-Resolution Imaging Spectroradiometer) and ozone analysis data from NCEP (National Centers for Environmental Predictions). By using data from multiple satellites, also geostationary satellites, instead of climatology, CERES can assumed to be more exact. Therefore CERES is useful for validating the combined product.

In figure 13 an example of a full sky monthly mean SSI from CERES is shown. Because SCIAMACHY does not provide a full global dataset on a daily basis, the combined product will be compared to CERES by using the monthly means.

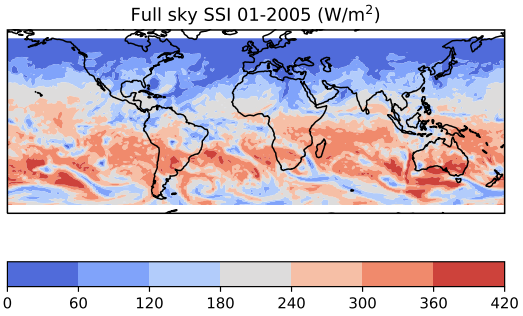


Figure 13: Global full sky daily mean SSI from CERES on January 1 2005

3.5 BSRN

For the validation of the obtained products in this research the data is compared to the data from fifteen measurement stations from the BSRN-database (Baseline Surface Radiation Network). The BSRN-measurments are a project of GEWEX (Global Energy and Water Cycle Experiment) to detect changes in climate. The locations of the used measurement stations are shown in figure 14.

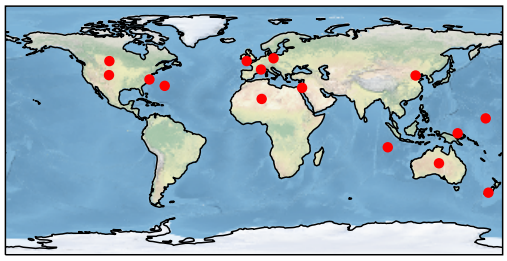


Figure 14: The locations of the 15 BSRN measurement stations

From the BSRN-stations the data at the overpass time of each satellite will be used to validate the data from OMI and SCIAMACHY. Furthermore, a daily and monthly mean will be calculated to validate the monthly mean dataset developed by in this research.

3.6 MODIS

To evaluate the climatology dataset that is used in the OMI and SCIAMACHY SSI, data from MODIS is used. MODIS is a measurement instrument on board of the Terra/Aqua satellites and can create a global map in one day. Figure 15 shows an example of a monthly mean aerosol optical thickness and figure 16 shows the time series of the aerosol optical depth climatology used by OMI and SCIAMACHY and the measured aerosol optical depth by MODIS.

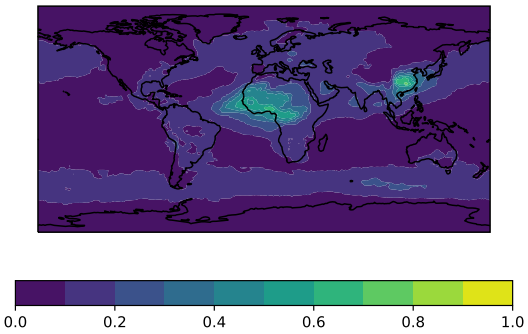


Figure 15: Global aerosol optical thickness from MODIS on January 1, 2005

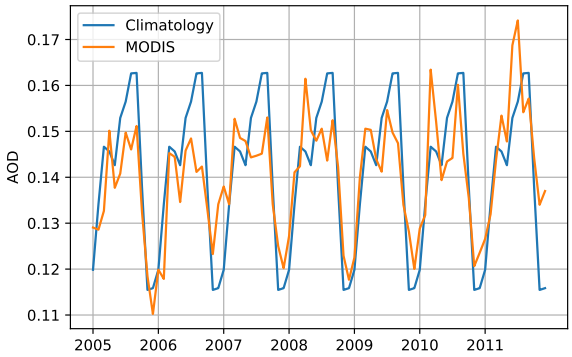


Figure 16: Monthly mean of the aerosol optical depth climatology and the aerosol optical depth measured by MODIS

4 Method

4.1 Gridding OMI and SCIAMACHY

In order to obtain a global image the data from the satellites has to be gridded to a grid with grid cells of the size of 0.25 ° x 0.25 °, lattitude x longitude. To do that the algorithm runs through all the data from one day and assigns each datapoint to a grid cell based on the lattitude and longitude information of that datapoint. There are several filters that remove negative values or unrealistic values, to prevent outliers. These outliers for instance occur a lot above areas with snow or ice at the surface. The cloud pressure is only selected and used if the cloud fraction at that location is above 0.05. At the end of the daily dataset the averages SSI and SSI_{cls} will be calculated for each grid cell. From there the averaged daily global data was stored per satellite. So a dataset of OMI SSI and a dataset of SCIAMACHY SSI was created. It is important to mention that during the course of 2007 a row anomaly developed in the measurements of OMI. This means that under a certain angle the spectrometer was blocked by a different instrument which can be seen as a row on the map. Therefore the term 'row anomaly'. The data having an anomaly are removed in the analysis.

4.2 Validating OMI and SCIAMACHY

After gridding the data for both the OMI and the SCIAMACHY satellite, the data has been validated with the use of BSRN-surface measurements. The product of SCIAMACHY and OMI will be compared to the CERES-dataset and the CLARA-dataset.

Due to the fact that the satellite measures every location only once per day, the data from the surface measurement has to be filtered to obtain the measurement that coincides with the satellite measurement. For this the solar zenith angle is used, see figure 17. The solar zenith angle is the angle between the normal on the earth's surface and the line from the earth's surface to the centre of the sun. Both OMI and SCIAMACHY store the solar zenith angle at the time of the measurement. For the BSRN-measurements the solar zenith angle has to be calculated using the longitude, lattitude, the date and the time.

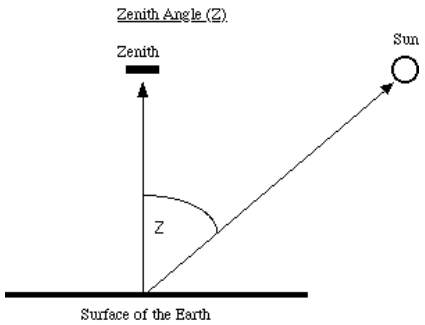


Figure 17: The Solar Zenith Angle

After the solar zenith angles of the BSRN-dataset are calculated they are compared to the measured solar zenith angles from the satellites. This has to be done carefully because the solar zenith angle is symmetric in the morning and afternoon, therefore for SCIAMACHY the solar zenith angles in the afternoon have to be ignored (overpass time in the morning) and for OMI the solar zenith angles in the morning have to be ignored (overpass time in the afternoon). After the removal the time of the day is found where the difference between the solar zenith angle of the satellite and the solar zenith angle of the surface measurement is the smallest. It is assumed that the surface measurement at that time is the closest possible time to the overpass time. To avoid the interference of outliers and to match the pixel size of the satellite the value is averaged with the surrounding values of plus and minus thirty minutes. So we obtain an hourly average of the surface solar irradiance around the overpass time of the satellite measured at the surface. These values can then be compared to the satellite observations.

4.3 Combining OMI and SCIAMACHY

In order to create a daily mean full sky SSI derived from OMI and SCIAMACHY equation 11(Moser & Raschke, 1984) is used:

$$SSI_{dm} = SSI_{clr,dm} * \frac{\sum_i SSI_i}{\sum_i SSI_{clr,i}} \tag{11}$$

In this equation $SSI_{clr,dm}$ is the clear sky daily mean product from CLARA, SSI_i are the full sky observations of OMI and SCIAMACHY, $SSI_{clr,i}$ are the clear sky observations of OMI and SCIAMACHY and i is the number of satellites. The clear-sky daily mean from CLARA is used because it is created by the same algorithm as the SSI products from OMI and SCIAMACHY. What this equation in fact does, is that it calculates an average transmission coefficient and multiplies that with the clear sky daily mean to obtain a full sky daily mean. In theory the average transmission coefficient should get better by increasing the number of satellites. To test this, the daily mean clear sky

SSI, the daily mean full sky SSI using OMI, the daily mean full sky SSI using SCIAMACHY and the daily mean full sky SSI using both satellites are all compared to the CERES dataset and the BSRN measurement stations. For the CERES dataset this is done by comparing global means of both datasets. For the surface measurements a lattitude and a longitude is extracted from the data and the corresponding grid cell from the gridded SSI datasets is used to compare the SSI with the surface measurements. To validate the data set with the BSRN measurement stations a daily mean has to be calculated from the BSRN data. This is done by calculating the average from sunrise until sunset and by discarding a day once there is over an hour of data missing, because that will highly effect the daily mean (Roesch et al., 2011).

4.4 Statistical methods

To intercompare the datasets with each other a few statistical methods are used. For starters the correlation coefficient is often calculated to measure the correlation between two datasets. The correlation coefficient ($\mathbf{r}$) is defined as:

$$\mathbf{r} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{12}$$

Here n is the number of datapoints, x and y are respectively the satellite measurements and the surface measurements, where $\bar{x}$ and $\bar{y}$ are the averages. The correlation coefficient is a measure of the linear relatedness between two datasets. The closer it is to one the better the linear relation is between the datasets.

To calculate the difference between two datasets the root mean square error (RMSE) and the mean absolute error (MAE) are used. They are defined as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(x_i - y_i)^2}{n}} \tag{13}$$

$$MAE = \frac{\sum_{i=1}^{n}|x_i - y_i|}{n} \tag{14}$$

In these equations x, y and n mean the same as in equation 12.

5 Results

5.1 Validating OMI and SCIAMACHY

In figure 18 the scatterplot between OMI and the surface measurements are shown and the scatterplot between SCIAMACHY and the surface measurement are shown for the location ALice Springs in Australia. Each datapoint represents the SSI at overpass time of the satellite. For OMI this means that there is a datapoint every day and for SCIAMACHY there is a datapoint every six days. The data shown in figure 18 is of 2005 and 2006.

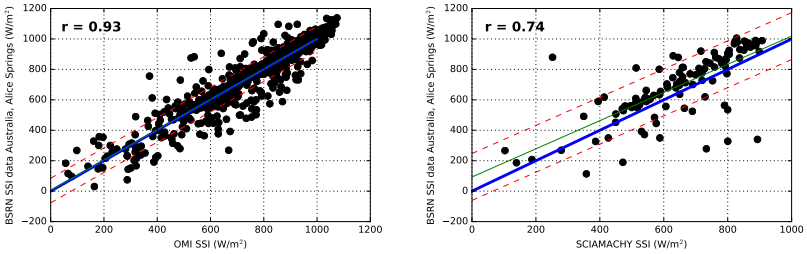


Figure 18: Scatter plots of total SSI at Alice Springs (Australia). The blue line is the diagonal, the green line is the best fit and the dashed red lines are the best fit plus or minus the standard deviation. **r** is the Pearson correlation coefficient. The data is from 2005 and 2006

The correlation between the surface measurements and the satellite observations when averaging all 15 surface measurement locations is for the OMI-instrument 0.92 and for the SCIAMACHY-instrument 0.86. Tables 1 & 2 show the statistics of this validation of the OMI and SCIAMACHY instruments

BSRN-site	lat	lon	OMI SSI	BSRN SSI	Bias (%)	RMSE	MAE	correlation
ASP	-23.6	133.9	719.6	720.7	-0.15	80.8	51.6	0.93
BER	32.3	-64.7	578.1	567.4	1.88	138.3	95.5	0.86
BOU	40.1	-105.2	572.8	585.4	-2.15	127.5	89.0	0.87
CAR	44.1	5.1	555.2	554.6	0.12	67.1	44.4	0.97
CLH	36.9	-75.7	583.6	583.6	0.01	80.9	53.9	0.95
COC	-12.2	96.8	669.3	637.7	4.96	112.8	79.5	0.87
KWA	8.7	167.7	703.5	584.5	20.3	223.6	167.9	0.71
LAU	-45.0	169.7	402.8	408.8	-1.47	108.8	71.9	0.90
LIN	52.1	14.1	367.5	383.5	-4.18	81.5	55.8	0.95
MAN	-2.1	147.4	561.6	555.4	1.13	157.0	115.0	0.82
REG	50.2	-104.7	414.0	468.0	-11.5	150.9	100.9	0.87
SBO	30.9	34.8	747.0	737.8	1.24	84.6	51.7	0.93
TAM	22.8	5.5	750.3	782.3	-4.09	113.1	78.5	0.88
XIA	39.8	117.0	533.4	506.1	5.38	89.6	64.0	0.93

Table 1: OMI-BSRN comparison, bias is (OMI-BSRN)/BSRN, RMSE is the root mean squared error and MAE is the mean absolute error

BSRN-site	lat	lon	SCIA SSI	BSRN SSI	Bias (%)	RMSE	MAE	correlation
ASP	-23.6	133.9	646.5	690.9	-6.42	153.8	114.9	0.74
BER	32.3	-64.7	582.8	581.7	0.19	141.1	102.8	0.84
BOU	40.1	-105.2	569.2	639.6	-11.0	127.2	99.6	0.89
CAR	44.1	5.1	524.1	542.0	-3.31	100.5	69.5	0.92
CLH	36.9	-75.7	534.3	542.4	-1.48	69.5	48.7	0.97
COC	-12.2	96.8	642.1	614.4	4.51	135.7	85.8	0.80
KWA	8.7	167.7	668.9	675.8	-1.03	114.1	80.2	0.86
LAU	-45.0	169.7	382.1	366.5	4.26	118.6	78.7	0.88
LIN	52.1	14.1	393.0	394.4	-0.36	79.2	57.9	0.95
MAN	-2.1	147.4	538.8	579.9	-7.08	310.5	238.1	0.19
REG	50.2	-104.7	486.8	492.3	-1.11	102.6	67.5	0.92
SBO	30.9	34.8	645.1	728.6	-11.45	142.1	110.3	0.84
TAM	22.8	5.5	783.7	822.6	-4.73	91.9	67.4	0.85
XIA	39.8	117.0	531.9	486.8	9.27	111.8	82.0	0.90

Table 2: SCIA-BSRN comparison, bias is (SCIA-BSRN)/BSRN, RMSE is the root mean squared error and MAE is the mean absolute error

Not just the total surface solar irradiance is measured by the BSRN-locations and the satellite-instruments, but also the direct irradiance. The direct radiance is the radiation that reaches the surface of the earth without being scattered or reflected. And using the same method as with the total surface solar irradiance by finding corresponding solar zenith angles the correlation was calculated, the results are shown in Appendix A.1. The distance between a datapoint and the one-to-one line is the absolute error and the mean of this error of all datapoints

is the mean absolute error. To check whether this absolute error for the total surface solar irradiance is seasonally dependent or dependent on the measured cloud fraction the correlation between the MAE is compared to the seasonality and the cloud fraction. The results are shown in Appendices A.2 and A.3. The seasonal mean absolute errors are normalized in this case with the mean of that season.

5.2 Daily mean clear sky product from CLARA

5.2.1 Comparison to BSRN

To see what the results are of using the satellite data from OMI and SCIAMACHY, first the agreement between the clear sky daily mean and the BSRN-observations is looked at. The results of the comparison between the BSRN-measurements and the CLARA daily mean - monthly mean are shown in figure 19 and table 3.

BSRN-site	lat	lon	CLARA cs	BSRN SSI	Bias (%)	RMSE	MAE	correlation
ASP	-23.6	133.9	281.1	262.7	7.03	26.9	19.3	0.96
BER	32.3	-64.7	257.5	194.5	32.21	65.1	63.0	0.98
BOU	40.1	-105.2	244.9	197.0	24.32	51.9	47.9	0.99
CAM	50.2	-5.3	196.1	128.1	53.09	76.1	68.0	0.98
CAR	44.1	5.1	217.0	182.8	18.73	37.5	34.2	0.99
CLH	36.9	-75.7	240.3	185.6	29.43	58.1	54.6	0.98
COC	-12.2	96.8	294.2	219.0	34.35	69.9	67.9	0.84
KWA	8.7	167.7	293.4	227.2	29.14	67.3	66.2	0.83
LAU	-45.0	169.7	225.8	163.2	38.37	72.9	62.6	0.99
LIN	52.1	14.1	186.2	130.7	42.40	63.8	55.4	0.98
MAN	-2.1	147.4	293.8	198.4	48.09	96.9	95.4	0.47
REG	50.2	-104.7	204.3	157.6	29.58	56.3	46.6	0.98
SBO	30.9	34.8	258.9	272.8	-5.10	44.8	30.2	0.94
TAM	22.8	5.5	278.9	261.5	6.6	22.6	18.6	0.96
XIA	39.8	117.0	229.6	162.3	41.49	77.2	67.4	0.95

Table 3: CLARA clear sky daily mean, BSRN comparison of the years 2005 and 2006, bias is (CLARA - BSRN)/BSRN, RMSE is the root mean squared error and MAE is the mean absolute error

The first thing that stands out is that the mean of the clear sky is significantly higher than the mean of the BSRN-measurements stations. This makes sense, because there are no clouds reflecting the sunlight. This means that the calculated sunlight that reaches the surface of the earth is higher than it should be. This is also represented by the high bias. The second important feature is that the correlation coefficient is already very high (0.9 for most cases). This means that the transmission coefficient of the atmosphere does not influence the trend significantly, it mostly influences the absolute value of the surface solar irradiance. The large correlation coefficients are due to the effects of the solzar zenith angle on the SSI, which is the same for the clear sky SSI and the full sky SSI.

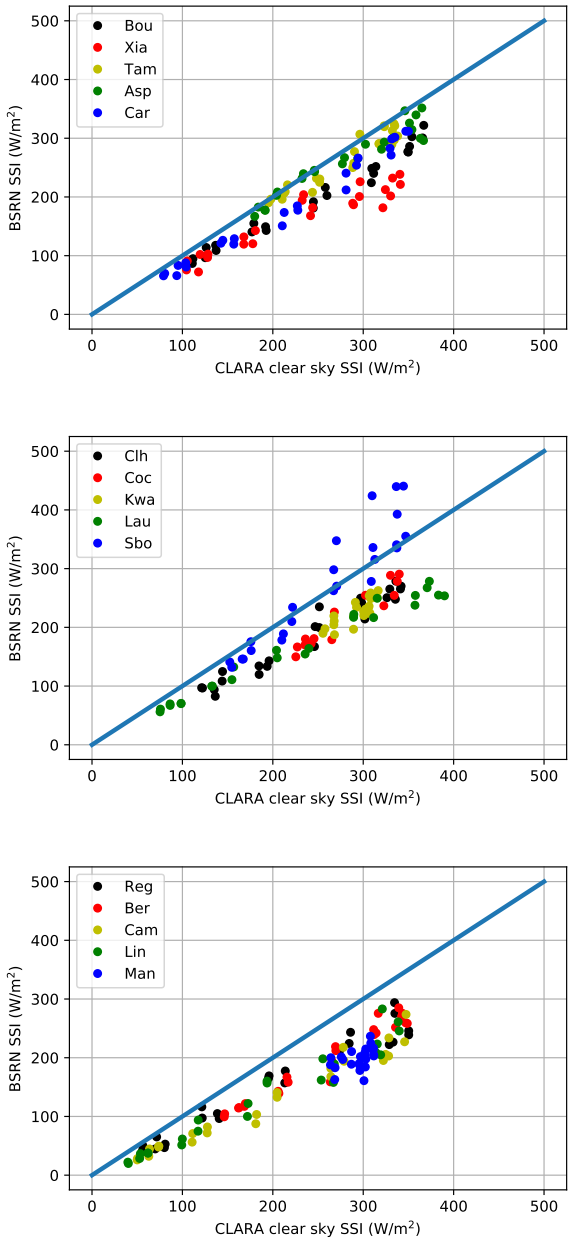


Figure 19: Scatter plots of the bsrn locations monthly means compared to the CLARA monthly means of the years 2005 and 2006.

5.2.2 Comparison to CERES

In order to compare the dataset to the dataset from CERES, a monthly global mean is created for the years 2005 until 2011 for both the CLARA dataset and the CERES dataset. After this the statistics, just like the comparison with the BSRN-data, several statistics are calculated to assess the comparison on a yearly basis. The results of the comparison are shown in table 4. It is clear that the statistics hardly vary on a yearly basis and that the CLARA SSI is about 32% per cent higher than the CERES SSI, this is off course due to the fact that clouds are not included in the CLARA dataset.

Year	CLARA cs dm SSI	CERES SSI	Bias (%)	RMSE	MAE	correlation
2005	265.04	195.76	32.60	74.88	65.87	0.94
2006	264.30	196.88	32.16	73.81	64.95	0.94
2007	264.37	196.60	32.37	74.51	65.31	0.94
2008	264.57	197.19	31.99	73.85	64.78	0.94
2009	263.92	197.29	31.89	73.27	64.43	0.94
2010	264.26	196.12	32.53	74.63	65.56	0.94
2011	264.42	196.70	32.31	74.44	65.23	0.94
Average	264.41	196.65	32.26	74.20	65.16	0.94

Table 4: CLARA daily means compared with CERES, bias is (CLARA - CERES)/CERES, RMSE is the root mean squared error and MAE is the mean absolute error.

5.3 Daily mean product using SCIAMACHY

5.3.1 Comparison to BSRN

After comparing the CLARA clear sky daily mean with the BSRN measurements and the CERES dataset, a sky daily mean product was created by using only data from SCIAMACHY and CLARA. This was done by using equation 11. With the daily mean results the monthly means were calculated. Table 5 shows the statistics of the comparison of the monthly means from the product of equation 11 using only SCIAMACHY and the BSRN data. For these statistics the monthly means from the years 2005 and 2006 were used. Figure 20 shows the scatter plots of this data. Compared to the case where only the CLARA data was used, the bias has decreased significantly due to the implementation of a transmission coefficient. The correlation coefficients are comparable to the correlation coefficients in table 3.

BSRN-site	lat	lon	SCIA dm SSI	BSRN SSI	Bias (%)	RMSE	MAE	correlation
ASP	-23.6	133.9	233.8	262.7	-11.01	33.9	28.9	0.95
BER	32.3	-64.7	198.5	194.5	2.07	25.7	18.7	0.92
BOU	40.1	-105.2	189.0	197.0	-4.07	20.2	16.3	0.97
CAM	50.2	-5.3	121.0	128.1	-5.56	20.9	16.3	0.97
CAR	44.1	5.1	169.2	182.8	-7.46	19.5	15.0	0.99
CLH	36.9	-75.7	176.3	185.6	-5.04	33.1	27.0	0.90
COC	-12.2	96.8	238.2	219.0	8.77	26.0	22.5	0.83
KWA	8.7	167.7	231.2	227.2	1.76	28.2	22.3	0.64
LAU	-45.0	169.7	170.8	163.2	4.67	21.0	15.9	0.97
LIN	52.1	14.1	127.7	130.7	-2.34	14.3	12.1	0.99
MAN	-2.1	147.4	191.7	198.4	-3.38	26.1	20.5	0.44
REG	50.2	-104.7	155.7	157.6	-1.27	17.8	14.7	0.98
SBO	30.9	34.8	206.0	272.8	-24.51	81.8	66.9	0.93
TAM	22.8	5.5	260.6	261.5	-0.33	20.2	16.6	0.93
XIA	39.8	117.0	170.8	162.3	5.26	21.9	14.7	0.96

Table 5: SCIAMACHY daily mean, BSRN comparison of the years 2005 and 2006, bias is (SCIA - BSRN)/BSRN, RMSE is the root mean squared error and MAE is the mean absolute error

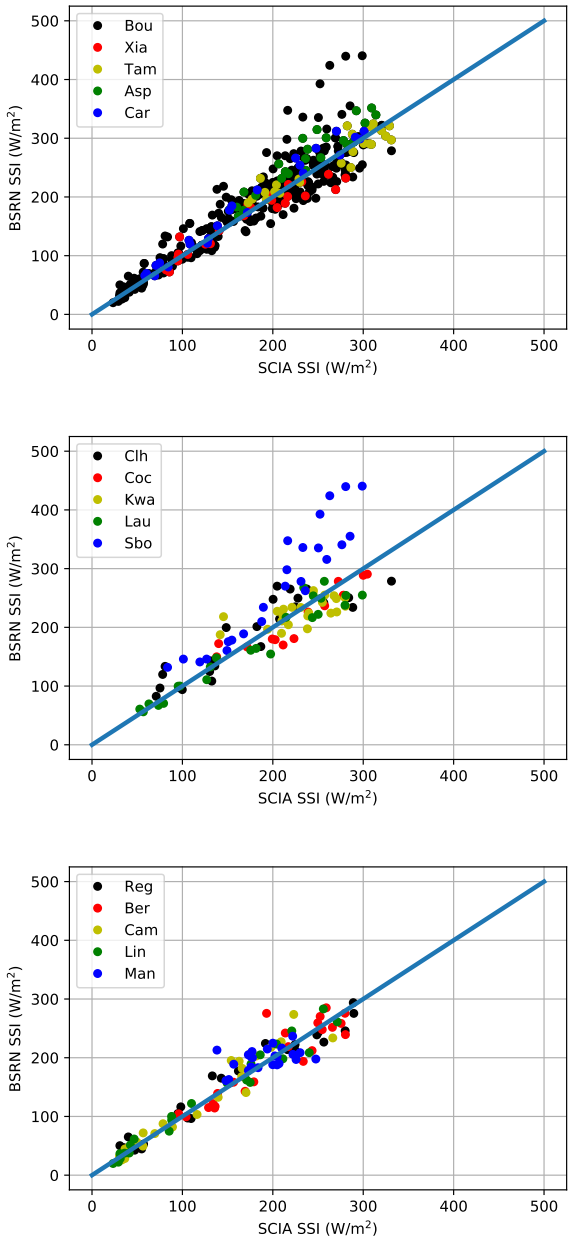


Figure 20: Scatter plots of the bsrn locations daily mean - monthly means compared to the Sciamachy daily mean - monthly means of the years 2005 and 2006.

5.3.2 Comparison to CERES

The generated SSI dataset by using SCIAMACHY to create a fullsky SSI product is compared to the CERES dataset in the same way as the CLARA dataset was compared to the CERES dataset. This means that the monthly global means are compared per year. This provides the numbers shown in table 6. The numbers are still stable from year to year. The important difference compared to the CLARA dataset is that the bias between the datasets is smaller than 1 %. Also 6 shows that the correlation coefficient has increased compared to the case without satellite data.

Year	SCIA dm SSI	CERES SSI	Bias (%)	RMSE	MAE	correlation
2005	196.71	195.82	-0.62	22.43	16.45	0.97
2006	197.29	196.50	-0.63	22.67	16.66	0.97
2007	196.75	196.21	-0.77	22.01	16.15	0.97
2008	197.39	196.82	-0.74	21.40	15.76	0.97
2009	197.42	196.85	-0.76	22.01	16.20	0.97
2010	196.18	195.77	-0.82	22.14	16.27	0.97
2011	197.19	196.30	-0.55	21.43	15.81	0.97
Average	196.99	196.32	-0.70	22.01	16.19	0.97

Table 6: SCIAMACHY daily mean compared with CERES, bias is (SCIA - CERES)/CERES, RMSE is the root mean squared error and MAE is the mean absolute error.

5.4 Daily mean product using OMI

For the OMI satellite the exact same thing has been done as in the previous section with the SCIAMACHY satellite. So a full sky daily mean SSI products was produced, using the transmission coefficient of OMI. This will of course provide a different dataset than the one using SCIAMACHY for several reasons. The most important reason is that OMI has a different overpass time than SCIAMACHY, therefore the measured cloud fraction will be different and therefore the transmission coefficient will not be the same as the one obtained from the SCIAMACHY data. This new dataset is then compared to the BSRN measurement stations and the results are shown in table 7 and figure 21.

5.4.1 Comparison to BSRN

BSRN-site	lat	lon	OMI dm SSI	BSRN SSI	Bias (%)	RMSE	MAE	correlation
ASP	-23.6	133.9	249.5	262.7	-5.02	15.2	13.7	0.99
BER	32.3	-64.7	195.7	194.5	0.59	8.3	6.1	0.99
BOU	40.1	-105.2	184.0	197.0	-6.56	16.2	13.1	0.99
CAM	50.2	-5.3	129.8	128.1	1.36	7.0	4.9	1.00
CAR	44.1	5.1	179.2	182.8	-1.96	7.0	5.9	1.00
CLH	36.9	-75.7	187.2	185.6	0.85	8.2	6.3	0.99
COC	-12.2	96.8	238.7	219.0	8.98	19.2	14.4	0.87
KWA	8.7	167.7	233.6	227.2	2.81	8.1	6.8	0.97
LAU	-45.0	169.7	156.8	163.2	-3.92	13.0	9.6	0.99
LIN	52.1	14.1	122.9	130.7	-5.98	10.5	8.1	1.00
MAN	-2.1	147.4	190.3	198.4	-4.08	11.3	9.2	0.90
REG	50.2	-104.7	140.3	157.6	-11.03	31.2	18.1	0.96
SBO	30.9	34.8	236.4	272.8	-13.35	52.1	36.4	0.95
TAM	22.8	5.5	243.1	261.5	-7.03	20.1	18.4	0.98
XIA	39.8	117.0	172.3	162.3	6.22	12.9	10.7	0.99

Table 7: OMI daily mean, BSRN comparison of the years 2005 and 2006, bias is (OMI - BSRN)/BSRN, RMSE is the root mean squared error and MAE is the mean absolute error

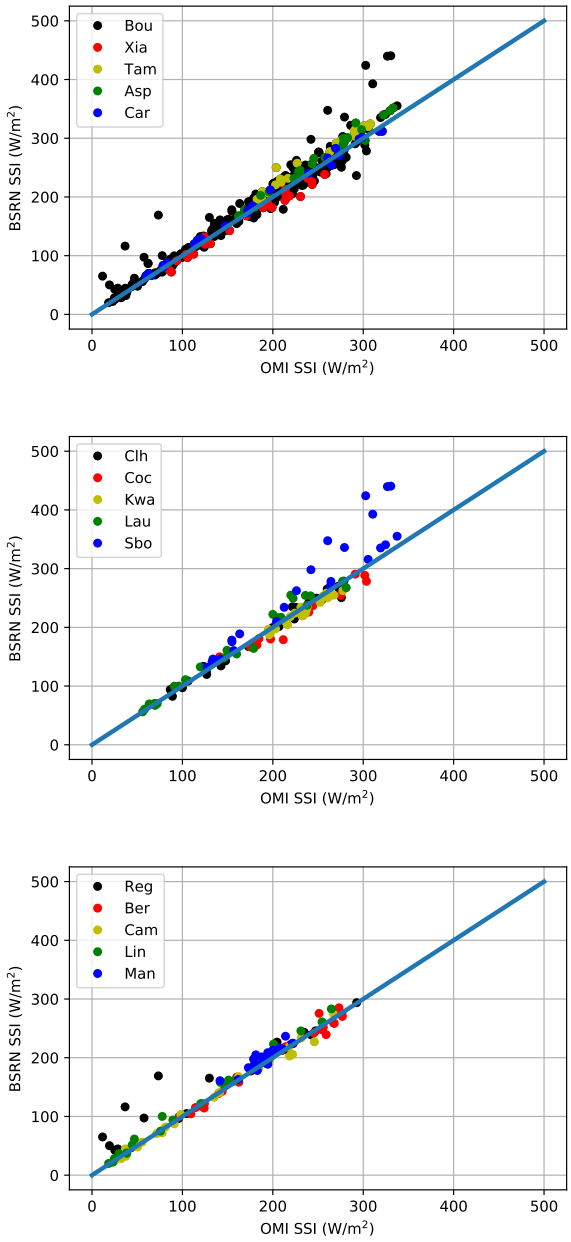


Figure 21: Scatter plots of the bsrn locations daily means - monthly means compared to the OMI daily means - monthly means of the years 2005 and 2006.

Table 7 shows that the correlation between the OMI full sky daily mean SSI and the surface measurements is very high, in three occasions the correlation coefficient is even equal to 1.00. This can also be seen in figure 21, because for most surface measurement locations the scattering is very much aligned with the 1- to 1-line.

5.4.2 Comparison to CERES

Of course the dataset created with the OMI satellite has also been compared to the CERES dataset and the results are shown in table 8. The first thing that stands out is that the correlation coefficient is again consistently high (0.99). The second thing that stands out is that the bias is consistently positive. So the created dataset is consistently higher than the CERES dataset. Whereas for the dataset created with the SCIAMACHY transmission coefficient this was the other way around. The RMSE and the MAE are smaller than for the SCIAMACHY daily mean product.

Year	OMI dm SSI	CERES SSI	Bias (%)	RMSE	MAE	correlation
2005	201.15	196.49	2.16	13.08	9.14	0.99
2006	201.44	197.14	1.98	12.59	8.94	0.99
2007	201.25	196.85	2.04	12.93	8.96	0.99
2008	201.95	197.46	2.07	12.82	8.96	0.99
2009	202.10	197.50	2.11	12.95	9.14	0.99
2010	202.02	196.40	2.66	13.92	9.89	0.99
2011	202.91	196.96	2.84	14.29	10.13	0.99
Average	201.83	196.97	2.27	13.22	9.31	0.99

Table 8: OMI daily mean compared with CERES, bias is (OMI - CERES)/CERES, RMSE is the root mean squared error and MAE is the mean absolute error.

5.5 The OMI-SCIAMACHY product

5.5.1 Comparison to BSRN

At this point both satellites were used to obtain a daily mean SSI, by using both SCIAMACHY and OMI in equation 11. Theoretically, the transmission coefficient should improve when an average of two overpass times is used. To test this, the dataset is compared to surface measurements in the same way as the single-satellite datasets. The results are shown in table 9 and figure 22.

BSRN-site	lat	lon	OMI-SCIA SSI	BSRN SSI	Bias (%)	RMSE	MAE	correlation
ASP	-23.6	133.9	235.0	262.7	-10.53	37.1	27.8	0.89
BER	32.3	-64.7	200.5	194.5	3.07	27.1	19.7	0.91
BOU	40.1	-105.2	189.3	197.0	-3.92	20.0	16.0	0.97
CAM	50.2	-5.3	125.4	128.1	-2.10	15.3	11.3	0.98
CAR	44.1	5.1	176.9	182.8	-3.25	15.2	12.0	0.99
CLH	36.9	-75.7	177.8	185.6	-4.23	30.2	25.3	0.91
COC	-12.2	96.8	238.1	219.0	8.71	23.5	19.3	0.86
KWA	8.7	167.7	230.4	227.2	1.41	18.9	16.8	0.71
LAU	-45.0	169.7	163.0	163.2	-0.10	13.6	11.0	0.98
LIN	52.1	14.1	124.0	130.7	-5.13	14.8	11.1	0.99
MAN	-2.1	147.4	188.5	198.4	-4.96	25.4	20.8	0.63
REG	50.2	-104.7	148.7	157.6	-5.66	23.3	17.7	0.97
SBO	30.9	34.8	221.6	272.8	-18.80	64.8	51.3	0.95
TAM	22.8	5.5	251.2	261.5	-3.93	17.7	14.4	0.96
XIA	39.8	117.0	170.6	162.3	5.15	20.21	15.6	0.95

Table 9: OMI-SCIA, BSRN comparison, bias is (OMI-SCIA - BSRN)/BSRN, RMSE is the root mean squared error and MAE is the mean absolute error

These results are very similar to the comparison between the one-satellite datasets and the BSRN measurement stations. The comparison between the CERES dataset and surface measurement has also been executed, the results are shown in Appendix A.5.

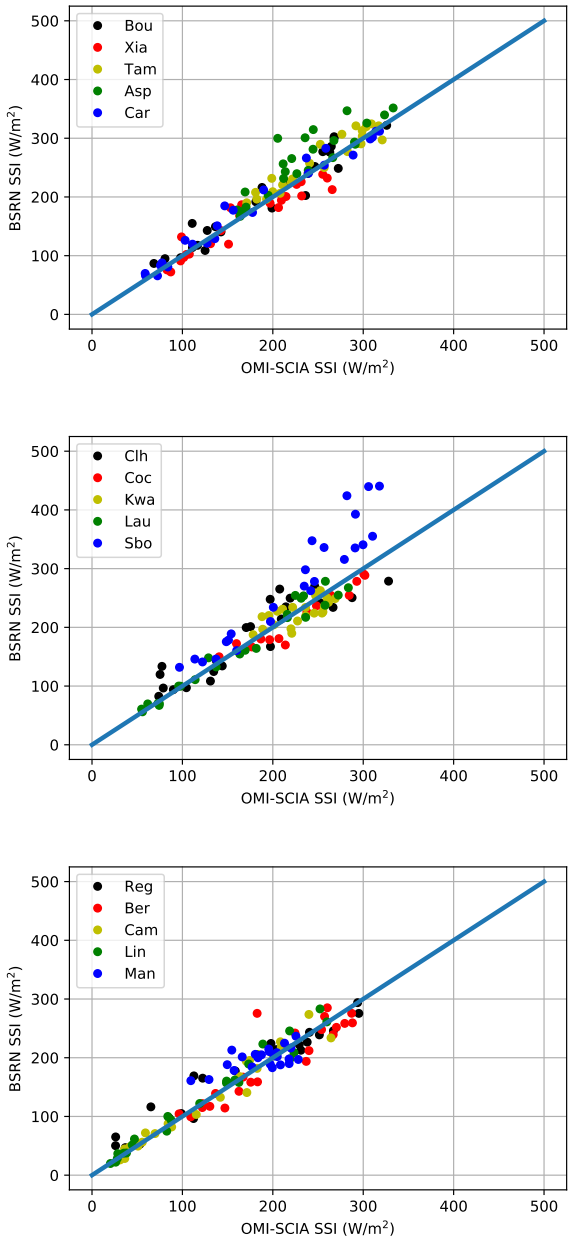


Figure 22: Scatter plots of the bsrn locations daily monthly means compared to the SCIA-OMI daily monthly means.

5.5.2 Differences between OMI-SCIAMACHY and CERES

The next step is to compare the dataset with the dataset of CERES. First this is done with the same method as was used for the one-satellite datasets. These results are shown in table 10 and a density plot is shown in figure 23.

Year	OMI-SCIA SSI	CERES SSI	Bias (%)	RMSE	MAE	correlation
2005	199.68	195.81	0.87	19.24	14.14	0.98
2006	200.13	196.49	0.79	19.58	14.42	0.98
2007	199.72	196.21	0.72	18.80	13.82	0.98
2008	200.34	196.80	0.73	18.50	13.64	0.98
2009	200.48	196.84	0.77	20.38	14.90	0.98
2010	199.81	195.74	1.00	21.26	15.54	0.98
2011	200.80	196.27	1.26	21.53	15.64	0.98
Average	200.14	196.31	0.88	19.90	14.59	0.98

Table 10: OMI-SCIA compared with CERES, bias is (OMI-SCIA - CERES)/CERES, RMSE is the root mean squared error and MAE is the mean absolute error.

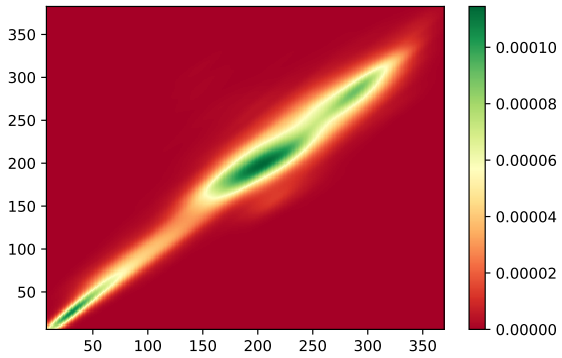
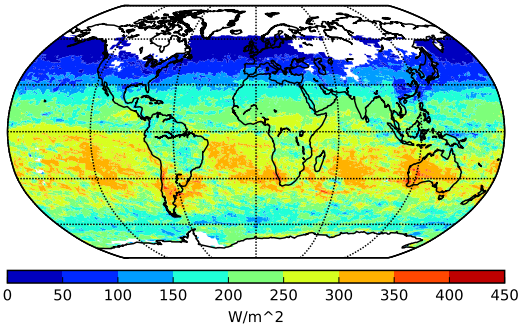


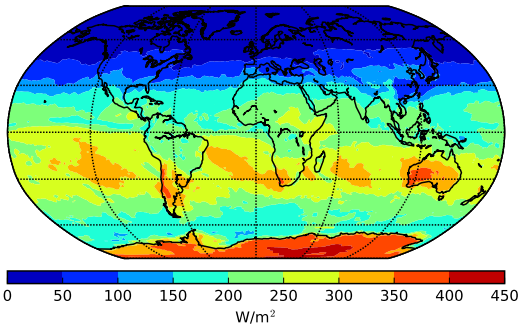
Figure 23: Density plot of each grid cell January 2005

Figure 23 shows that the scattering is concentrated around the one- to one-diagonal, which explains why the correlation coefficient in table 10 is as high as 0.98. Table 9 also shows that the bias is on average lower than 1%.

Figure 24 shows the monthly mean of January 2005 for both the OMI-SCIAMACHY daily monthly mean and the CERES daily monthly mean. Again it can be seen that the data is not complete over the poles, therefore in the calculation of the averages only the data between 60 degrees north and 60 degrees south was used. Furthermore the OMI-SCIAMACHY map looks very similar to the map created with CERES data.



(a) Monthly mean January 2005 SCIA-OMI



(b) Monthly mean January 2005 CERES

Figure 24: Global SSI

To see if there are large regional differences the difference between the OMI-SCIAMACHY dataset and the CERES dataset is also plotted. This is done in figure 25.

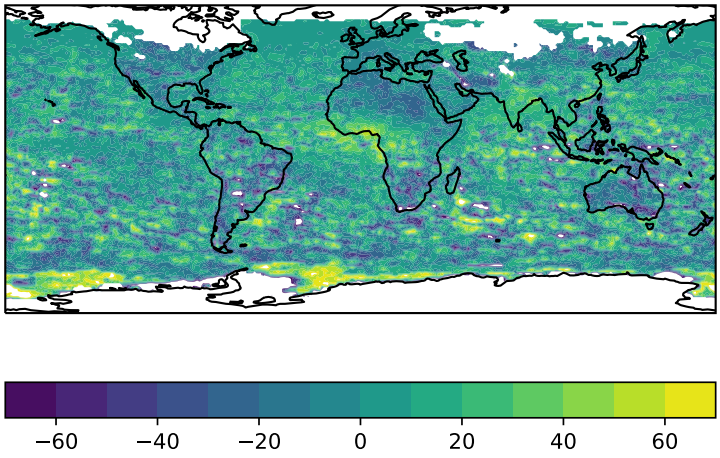


Figure 25: Monthly mean difference January 2005 (SCIA-OMI - CERES)

Figure 25 shows that the difference is quite homogenous, apart from the area near the poles (where the satellite detection is not fully functioning) and an area over Western Africa. There is a difference up to 50 W/m^2 over Western Africa. In the data section it is stated that the CLARA, OMI and SCIAMACHY algorithms use aerosol climatology to calculate the SSI. CERES on the other hand uses MODIS aerosol observations in its algorithm to obtain the SSI. Therefore the average Aerosol Optical Depth from MODIS of January 2005 is plotted in figure 26.

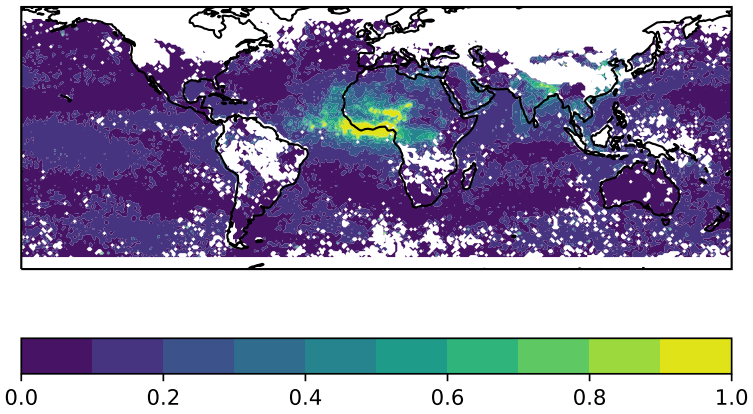


Figure 26: Aerosol Optical Depth January 2005

When comparing figures 25 and 26 it is clear that there is a relation between the aerosol concentration and the difference between the two SSI datasets. To investigate this relation further a scatter plot of the difference between the datasets and the aerosol optical depth was created. To do this every grid cell is plotted as a dot on the scatter plot. The result is shown in figure 27.

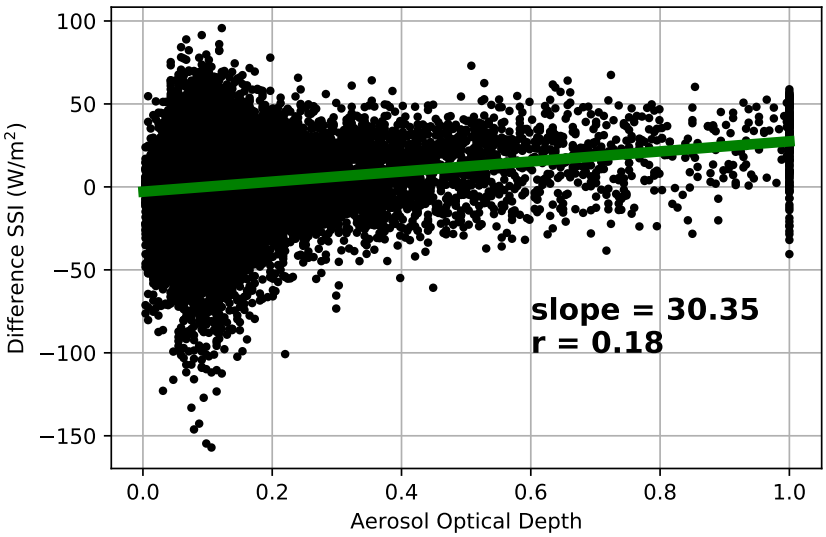


Figure 27: Scatterplot of January 2005

In figure 27 the aerosol optical depth is plotted horizontally and the difference (OMI-SCIAMACHY - CERES) is plotted vertically. The green line is the best linear fit of this dataset, which has a slope of 30.35. r is the correlation coefficient and is 0.18 in this case, which is not very convincing. The reason that the correlation coefficient is that low is because most data has a very small aerosol optical depth and there is quite some scattering around that small value. Table 11 shows the slopes of each linear fit of each month. The most important thing about this table is that almost every value is positive, which suggests that a relation between the difference and the aerosol optical depth is valid.

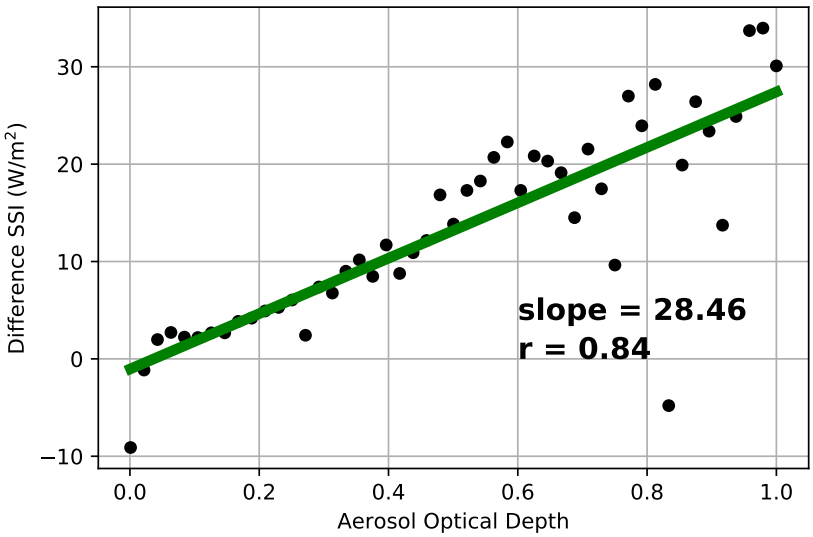


Figure 28: Binned scatterplot of January 2005

	2005	2006	2007	2008	2009	2010	2011	Average
January	30.35	32.5	32.02	31.53	32.77	19.18	26.08	29.20
February	24.15	17.8	26.04	26.01	24.56	20.86	18.74	22.59
March	16.67	18.17	17.21	17.71	18.47	20.26	12.78	17.32
April	15.11	10.4	20.93	10.07	6.45	15.84	15.02	13.40
May	11.06	17.3	13.64	14.21	12.72	12.62	12.28	13.40
June	8.19	5.87	3.77	15.68	7.33	10.38	16.56	9.68
July	10.12	8.39	10.98	14.55	13.1	7.04	15.41	11.37
August	14.03	9.86	11.87	9.07	14.79	18.7	13.1	13.06
September	21.91	13.55	28.48	13.55	7.35	22.19	12.34	17.05
October	-0.94	10.91	9.85	2.96	3.44	14.76	1.47	6.06
November	11.63	6.55	7.19	7.68	2.09	7.66	8.37	7.31
December	19.19	21.54	26.24	26.77	22.52	20.02	24.79	23.01
Average	15.12	14.40	17.35	15.82	13.80	15.79	14.75	15.29

Table 11: Monthly slopes from the 60N-60S data

To see if there is really a strong relation between the aerosol optical depth and the difference is to bin the data in steps of 0.02 in the horizontal direction. So 50 bins are created and averaged, the results of this binning are shown in 28.

The linear relation between the aerosol optical depth and the difference between the datasets is much clearer in the binned scatter plot, figure 28. This

is also represented by increased correlation coefficient of 0.84. The slopes of all months of data using this method are shown in table 12

	2005	2006	2007	2008	2009	2010	2011	Average
January	28.46	9.92	29.47	24.74	18.16	16.12	26.2	21.87
February	12.23	2.52	13.05	30.73	18.15	21.87	10.92	15.64
March	20.67	19.77	27.27	14.77	21.64	21.89	5.30	18.76
April	17.37	11.80	19.33	12.08	17.00	18.24	22.9	16.96
May	11.03	14.86	20.12	16.65	6.89	8.62	9.78	12.56
June	12.43	5.6	9.68	20.12	6.98	8.63	18.32	11.68
July	18.42	22.81	21.86	19.74	20.64	12.8	16.55	18.97
August	28.69	15.66	21.22	19.56	18.59	25.98	20.21	21.42
September	40.13	27.21	42.22	23.12	11.19	43.42	23.59	30.13
October	10.96	20.64	22	5.55	7.46	21.14	2.24	12.86
November	3.49	6.60	10.51	16.94	12.15	15.02	17.12	11.69
December	18.07	12.82	13.93	21.16	17.60	13.74	21.44	16.97
Average	18.50	14.17	20.89	18.76	14.70	18.96	16.21	17.46

Table 12: Monthly slopes from the 60N-60S binned data

5.5.3 Correcting using MODIS data

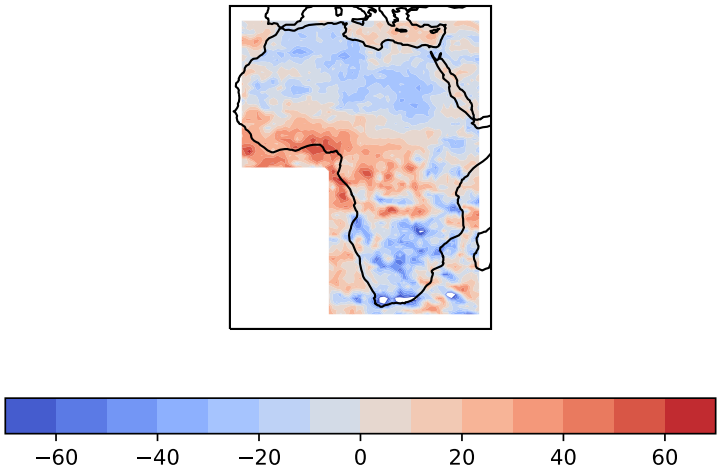
Using these slopes a correction can be performed on the data to account for the aerosol data. To see the effects of this correction, Africa is a useful case-study, because it has high aerosol concentrations and large difference with the CERES dataset. To highlight this, figure 29 shows the difference in SSI and the aerosol optical depth over Africa.

It is very clear from 29 that the difference in SSI over Africa is high in the same area that has a high aerosol optical depth. This can be corrected for by using the slopes from table 12. The average monthly slopes are used in this instance, because during the course of the year, different causes of high aerosol concentrations occur. Because different aerosols have different effects on the SSI, they cause a different slope. Therefore the monthly average slopes are used to correct the data by the following equation:

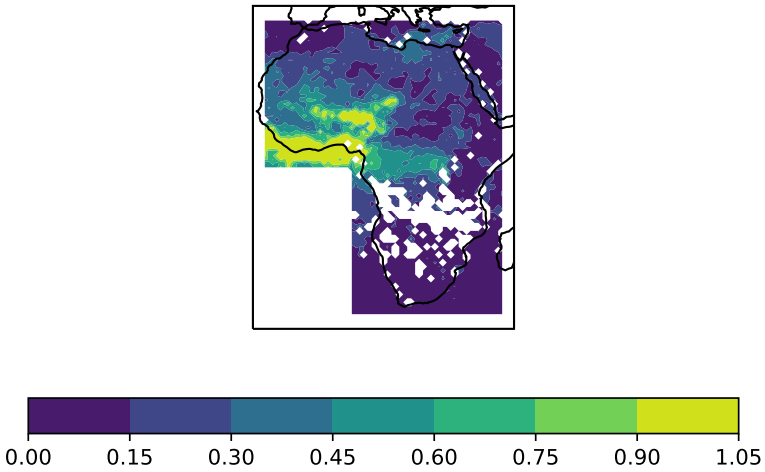
$$SSI_{aer} = SSI_{SCIA-OMI} - slope \cdot AOD \tag{15}$$

Using this equation, again the difference between SCIA-OMI and CERES can be plotted and this is featured in figure 30.

It is clear that the large differences over Western Africa have dissapeared by applying this method.



(a) Original monthly mean differences (SCIA-OMI - CERES)



(b) Aerosol optical depth

Figure 29: The SSI difference and aerosol optical depth over Africa in January 2005

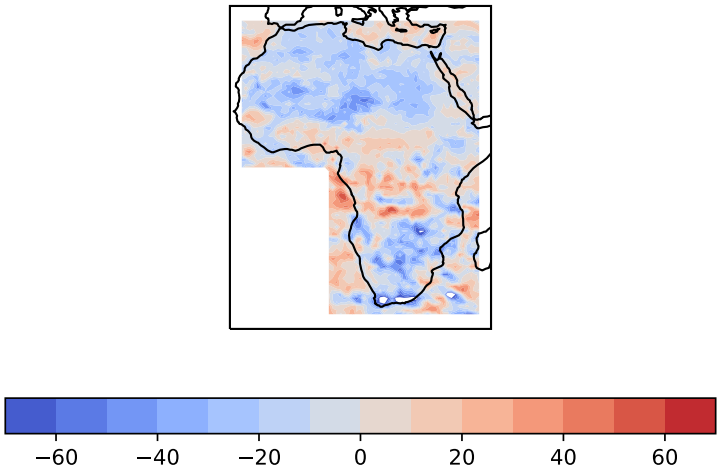


Figure 30: Aerosol optical thickness in Africa in January 2005

5.6 Trends

Now that the OMI-SCIAMACHY is created and validated it can be used to detect trends in SSI. Therefore the monthly mean of global (60°N to 60°S) SSI is plotted in figure 31.

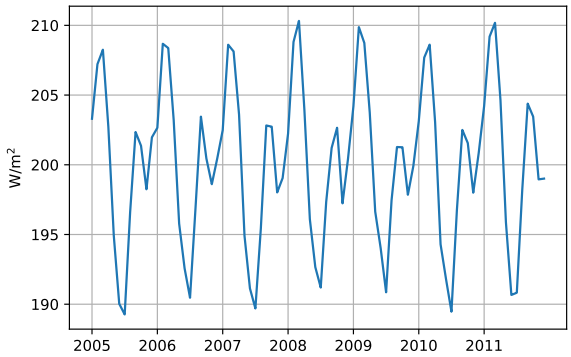


Figure 31: Global monthly mean SSI

When observing this monthly mean, it is clear that the SSI is periodic and peaks every year in February. Furthermore, there is no clear trend observable in the monthly mean SSI.

In order to be able to explain the shape of the time series of the SSI, it is necessary to look at the cloud fraction. In figures 32 and 33 the time series of the measured cloud fractions are featured.

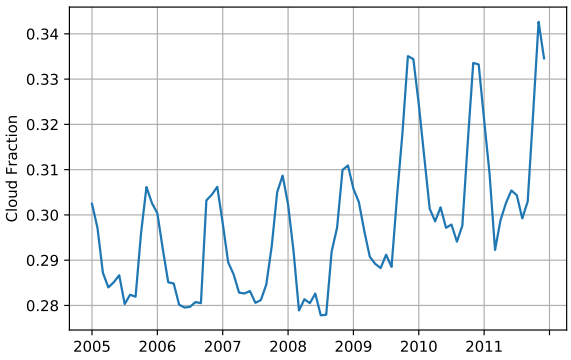


Figure 32: Global monthly mean cloud fraction from OMI

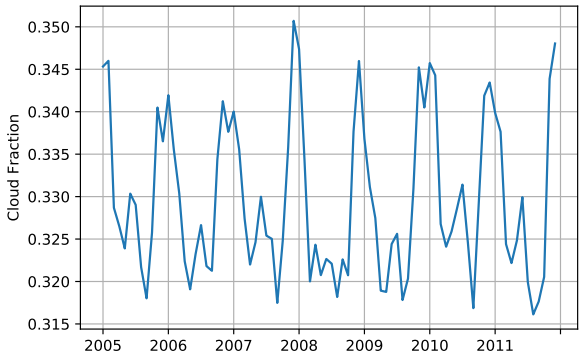


Figure 33: Global monthly mean cloud fraction from SCIAMACHY

The cloud fractions of OMI and SCIAMACHY show that the peaks and troughs of the SSI are opposite to the peaks and troughs of the cloud fraction. Which makes sens, because more clouds mean less sunlight reaching the surface of the earth. There also seems to be a shift in the OMI cloud fractions in the course of 2009. This is also the case when plotting the full time series of OMI, figures 34 and 35 show those.

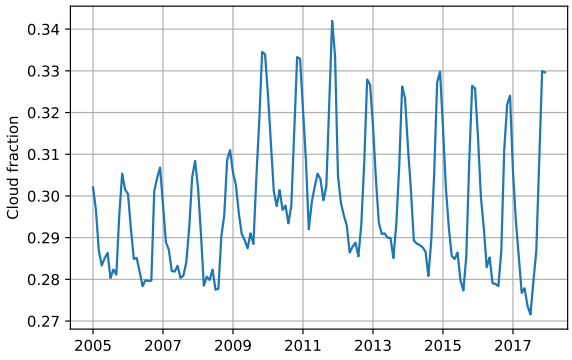


Figure 34: Complete time series of OMI cloud fraction

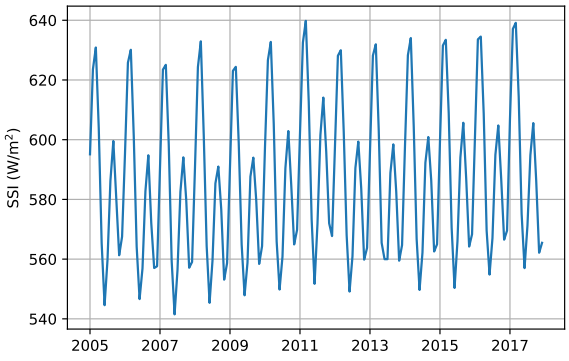


Figure 35: Complete time series of OMI SSI, at the overpass time of OMI

Figure 34 makes the shift in 2009 even clearer. The cause of this shift could be the OMI row anomaly which started in 2007 and expanded at the end of 2008. A row anomaly is an anomaly which has an effect on the data, but only at a certain measurement angle of the satellite. This means that this anomaly will show on your map as a row, therefore the term. In the calculation of the SSI a correction was performed for the row anomaly, therefore the row anomaly does not show up in figure 35. It is clear from figures 34 and 35 and the theory section that the cloud fraction and the full sky SSI are highly related. To see in what manner these quantities are related figure 36 features the scatter plot of the cloud fraction and SSI at overpass time from OMI. Because the cloud fraction experiences the row anomaly from 2010 and on wards, the scatter plot is divided into two groups. The blue dots represent the monthly means from 2005 until 2009 and the red dots the monthly means from 2010 until 2017. The correlation coefficients are calculated for each group seperated because there is a clear shift in the cloud fraction from 2010 on wards. The cloud fractions of -0.96 and -0.98 show that the cloud fraction is highly anti-correlated with the full sky SSI, as expected.

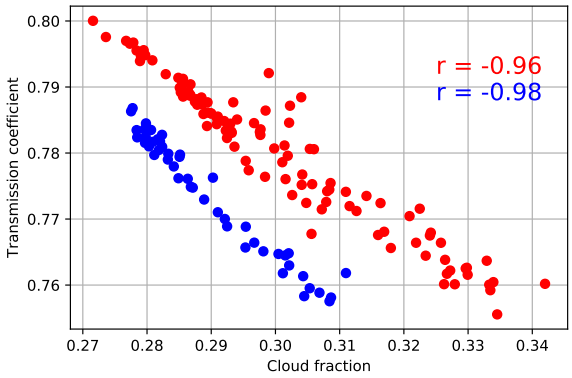


Figure 36: Scatter plot of the monthly means of the Transmisscion coefficient and the cloud fraction, 2005-2009 (blue) and 2010-2017 (red)

Previously it was shown that the bias between the created SSI datasets and the CERES dataset was minimal, to see if this bias has a certain periodicity the time series of the bias (monthly means) is plotted in figure 37.

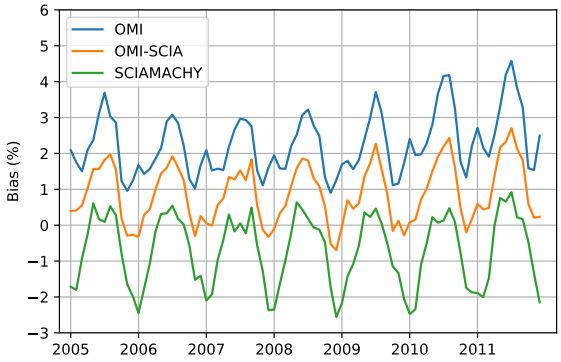


Figure 37: Time series of the monthly mean biases of the created products

Figure 37 shows that the bias is yearly periodic with peaks in the middle of the years and troughs at the end/beginning of a year. Also there seems to be a slight trend upwards. A possible cause of this periodicity could be the varying aerosol concentration during the year. To check if this could be the case the time series of the SCIAMACHY-OMI product is plotted together with the time series of the aerosol optical depth from MODIS (figure 38).

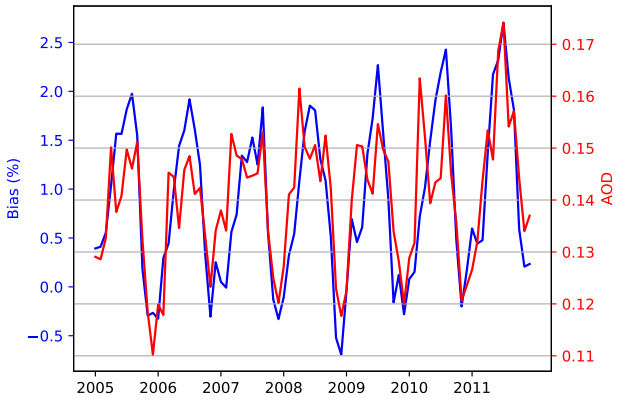


Figure 38: Time series of the monthly mean (SCIA-OMI - CERES)/CERES bias and the aerosol optical depth monthly mean

Figure 38 shows that the aerosol optical depth has similar peaks and troughs as the bias. To confirm that these two quantities are correlated a scatter plot was created with the monthly means from 2005 up and until 2011. This scatter plot is shown in figure 39

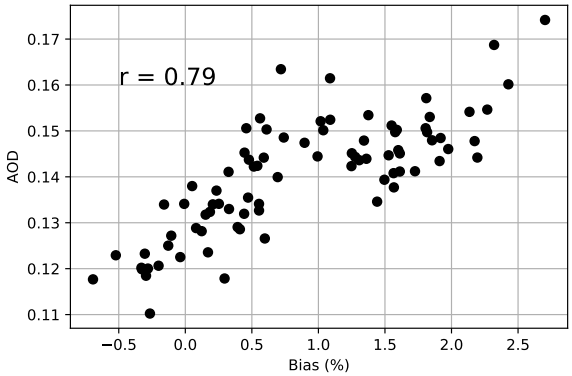


Figure 39: Scatter plot of the monthly mean (SCIA-OMI - CERES)/CERES bias and the aerosol optical depth monthly mean

The correlation coefficient of 0.79 shows that the bias between the two datasets is correlated with the amount of aerosols in the atmosphere. In this case the correlation was calculated by comparing the bias to the aerosol optical depth measured by MODIS, the correlation with the aerosol optical depth climatology is 0.83, the plots are shown in Appendix A.4. To check whether the bias is caused by the difference between the MODIS aod and the climatology

aod, the time series of the difference between MODIS aod and the climatology aod is plotted with the bias in figure 40.

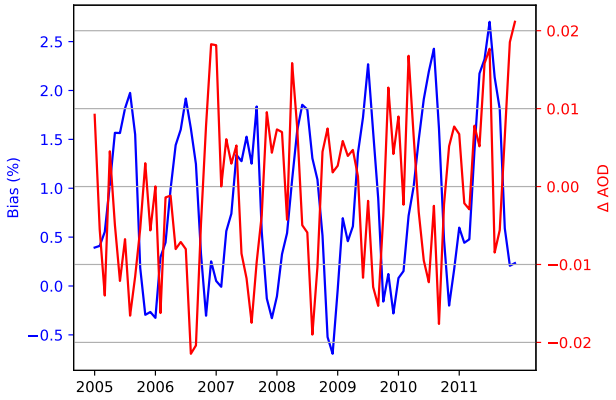


Figure 40: Time series of the monthly mean (SCIA-OMI - CERES)/CERES bias and the difference in aerosol optical depth (MODIS - climatology) monthly mean

Figure 40 clarifies that the bias is not caused by the absolute difference between the aerosol optical depth input of the two datasets. This is also demonstrated by the scatter plot of every monthly mean between 2005 and 2011 of the difference in aerosol optical depth and the bias in figure 41. The correlation coefficient for this scenario is -0.37.

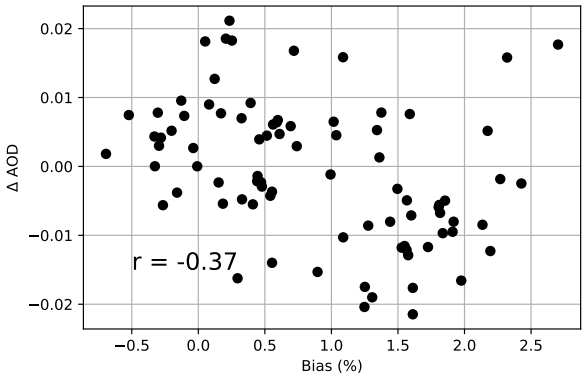


Figure 41: Scatter plot of the monthly mean (SCIA-OMI - CERES)/CERES bias and the difference in aerosol optical depth (MODIS - climatology) monthly mean

6 Conclusion

A broadband daily mean and monthly mean surface solar irradiance product has been derived from the effective cloud fractions from OMI and SCIAMACHY and the clear sky data from CLARA. The surface solar irradiances at overpass time have been calculated using the Lambertian cloud model and the Heliosat method. In order to obtain daily means and monthly means equation 11 was used. The individual satellite products at overpass time were validated using 15 BSRN surface stations. The correlation for the OMI data being on average 0.89 and for SCIAMACHY 0.83. The mean absolute biases were respectively 4.18 % and 4.73 %. Comparing the clear sky daily mean monthly mean SSI from CLARA to the surface stations and the CERES dataset showed that without the effect of clouds incorporated the average global SSI is 32.26 % to high. On the other hand the correlation coefficient with the CERES dataset is 0.94 on average, so this means that clouds highly affect the absolute value of the SSI, but have a much smaller effect on the periodicity of the SSI. By including the transmission coefficient of SCIAMACHY or OMI in equation 11 the bias of the daily mean monthly mean decreases significantly. The mean absolute bias for the daily mean monthly mean SSI using only SCIAMACHY compared to the surface measurements is 5.83 % and compared to the CERES measurements it is 0.70 %. For OMI this is respectively 5.31 % and 2.27 %. This is significantly improved compared to the results from using just the clear sky dataset. The average correlation coefficient of the SCIAMACHY daily mean SSI compared to the surface stations is 0.89 and compared to the CERES dataset it is 0.97. For OMI this is 0.97 and 0.99. Especially the correlation coefficients retrieved from the OMI daily mean monthly mean SSI are exceptionally high, regarding the fact that only the cloud fraction at 13:30 PM each day is taken into account.

After this both the data from OMI and the data from SCIAMACHY were taken into account in equation 11 to derive a daily mean monthly mean SSI. The bias of this combined dataset compared to the surface measurements is 5.40 % and compared to the CERES dataset it is 0.88 %. The correlation coefficients are on average 0.91 compared to BSRN data and 0.98 compared to CERES data. These statistics show that the data is in good agreement with on one hand measurements at the surface of the earth and on the other hand with a dataset created by different satellite data. It is clear from the results that regional differences in aerosol optical depth input can create regional differences of up to 60 W/m^2 in SSI between the OMI-SCIA product and the CERES dataset. This difference can be corrected for by using equation 15, because OMI-SCIA products uses climatology and therefore does not include local events like forest fires which could influence the regional aerosol optical depth quite severely. These events are captured by using data from MODIS. Figure 40 shows that the difference in aerosol input does not influence the global bias compared to the CERES dataset. Figure 36 features the expected relation between SSI and the cloud fraction. This relation is quite obvious, because the cloud fraction is used to calculate the transmission coefficient, but this figure also shows that anomalies like the row anomaly have to be understood to know that this scatter

plot has to be plotted in two parts.

This research has proved that it is possible to create a global SSI product from the OMI and SCIAMACHY satellites which is in good agreement with surface measurements and comparable datasets. There is no clear trend in the 7 years of overlapping data from OMI and SCIAMACHY, which can be expected, because a lot of climatologies were used in deriving this dataset and global yearly mean SSI does not vary a lot in the time frame of a few years. For further research it would be interesting to look further into the periodicity in the bias between the OMI-SCIA SSI and the CERES SSI. The cause of this periodic bias might improve the product. It would also be interesting to obtain surface measurements in areas with high aerosol optical depths, because these areas contain large differences between the satellite products.

7 Acknowledgements

First of all I would like to thank my daily supervisor Ping Wang from the KNMI. Her door was always open whenever I had a question or ran into trouble with my research. She steered me in the right direction, but also let this research be my own work.

Second of all I would like to thank my other supervisors, Piet Stammes from the KNMI and Maarten Krol and Wilfried van Sark from the University of Utrecht. Without them I would not have been able to complete this thesis and therefore obtain my master's degree.

I also would like to thank the KNMI for the opportunity and the resources they provided me with, necessary to conduct this research. Thank you to all the researchers from the department for the input and advice during the monthly meetings.

Finally, I must express my very profound gratitude to my family and friends for supporting and encouraging me during the course of my study. A special thanks to André, Jetse and Marlous for correcting my grammar and my typos.

References

[1] J.R. Acarreta et al., *Cloud pressure retrieval using the O_2-O_2 absorption band at 477 nm*, Journal of Geophysical Research, Volume 109, 2004.

[2] H. Bovensmann et al., *SCIAMACHY: Mission Objectives and Measurement Modes*, Journal of the Atmospheric Sciences, Volume 56, 1999.

[3] D. Cano et al., *A method for the determination of the global solar radiation from meteorological satellites data*, Solar Energy, Volume 37, 1986.

[4] E. Dribbsa et al., *A modification of the Heliosat method to improve its performance*, Solar Energy, Volume 65, 1998.

[5] S. Gehlot et al., *Impact of Sahara dust on solar radiation at Cape Verde Islands derived from MODIS and surface measurements*, Remote Sensing of Environment, Volume 166, 2015.

[6] D.W. Hahn, *Light Scattering Theory*, University of Florida, 2009.

[7] H.C. van de Hulst, *Light Scattering by Small Particles*, Dover publications, 1957.

[8] D.P. Kratz et al., *Validation of the CERES Edition 2B Surface-Only Flux Algorithms*, Journal of applied Meteoroly and Climatology, Volume 49, 2009.

[9] G. Louis Smith et al., *Clouds and the Earth's Radiant Energy System (CERES) Algorithm Theoretical Basis Document*, Journal of Geophysical Research, 1997

[10] R.W. Mueller et al., *The CM-SAF operational scheme for the satellite based retrieval of solar surface irradiance - A LUT based eigenvector approach*, Remote Sensing of Environment, Volume 113, 2009.

[11] R.W. Mueller et al., *Towards Optimal Aerosol Information for the Retrieval of Solar Surface Radiation Using Heliosat*, Atmosphere, Volume 6, 2015.

[12] A. Roesch et al., *Assesment of BSRN radiation records for the computation of monthly means*, Atmospheric Measurement Techniques, Volume 4, 2011.

[13] H. Senghor et al., *Seasonal cycle of desert aerosols in western Africa: analysis of the coastal transition with passive and active sensors*, Atmospheric Chemistry and Physics, Volume 17, 2017.

[14] M. Sengupta et al., *Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications*, National Renewable Energy Laboratory, 2nd edition, 2015.

[15] M. Sneep et al., *Three-way comparison between OMI and PARASOL cloud pressure products*, Journal of Geophysical Research, Volume 112, 2008.

[16] P. Stammes et al., *Effective cloud fractions from the Ozone Monitoring Instrument: Theoretical framework and validation*, Journal of Geophysical research, Volume 113, 2008.

[17] P. Wang et al., *Evaluation of broadband surface solar irradiance derived from the Ozone Monitoring Instrument*, Remote Sensing of Environment, Volume 149, 2014.

[18] P. Wang et al., *Surface solar irradiance from SCIAMACHY measurements: algorithm and validation*, Atmospheric Measurement Techniques, Volume 4, 2011.

[19] H.D. Young & R.A. Freedman, *University Physics wit Modern Physics*, Pearson, 14th edition, 2015.

[20] T. Zhang et al., *The validation of the GEWEX SRB surface shortwave flux data products using BSRN measurements: A systematic quality control, production and application approach*, Journal of Quantitative Spectroscopy & Radiative Transfer, Volume 122, 2013.

8 Appendix

8.1 A.1: Comparing direct SSI measurements to BSRN measurements

Figure 42 shows that the correlation between the direct irradiances is much lower than the correlation between the total irradiances. This is because the direct irradiance is much harder to measure due to the fact that you have to divide the incoming light into scattered light and direct light.

8.2 A.2: Linking the MAE to the cloud fraction

Although it is possible to draw a positive linear relation in each scenario. There is no clear relation between the MAE and the cloud fraction. The reason that it is possible to obtain a positive slope each occasion is that the MAE is close to 0 at a cloud fraction of 0, which pulls the best-fit down to (0,0).

8.3 A.3: Relating the MAE to seasonality

Figures 44 & 45 show that there is no clear connection between the mean absolute error and the time of the year.

8.4 A.4: Correlating climatology AOD to the SCIA-OMI and CERES bias

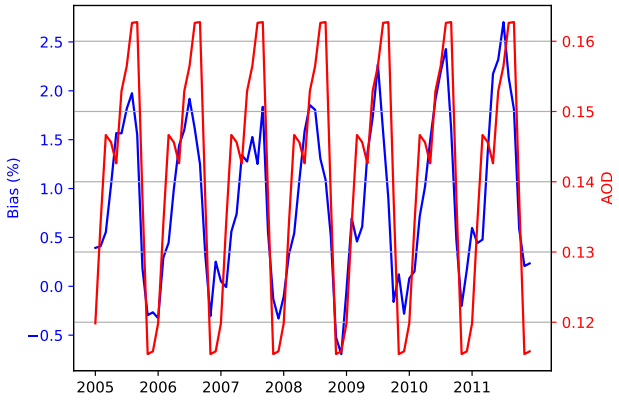


Figure 46: Time series of the monthly mean (SCIA-OMI - CERES)/CERES bias and the aerosol optical depth monthly mean

Figure 46 shows that the aerosol optical depth has similar peaks and troughs as the bias. To confirm that these two quantities are correlated a scatter plot was created with the monthly means from 2005 up and until 2011. This scatter plot is shown in figure 47. The correlation coefficient is 0.83.

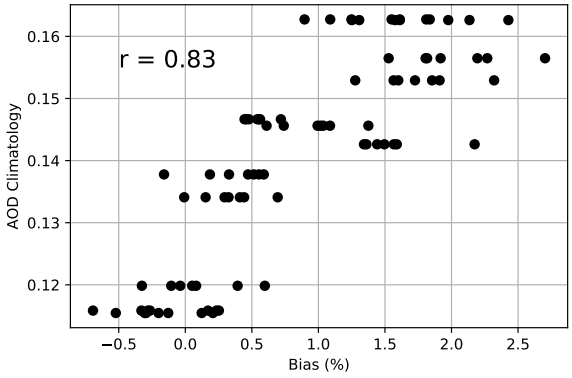


Figure 47: Scatter plot of the monthly mean (SCIA-OMI - CERES)/CERES bias and the aerosol optical depth monthly mean

8.5 A.5: Comparing CERES with BSRN data

BSRN-site	lat	lon	CERES mm SSI	BSRN SSI	Bias (%)	RMSE	MAE	correlation
ASP	-23.6	133.9	186.0	262.7	-29.2	84.0	76.7	0.83
BER	32.3	-64.7	148.1	194.5	-23.9	56.4	47.5	0.89
BOU	40.1	-105.2	178.3	197.0	-9.5	27.3	22.3	0.96
CAM	50.2	-5.3	98.9	128.1	-22.7	47.9	32.2	0.93
CAR	44.1	5.1	116.6	182.8	-36.2	79.8	66.2	0.96
CLH	36.9	-75.7	189.8	185.6	2.3	23.2	20.0	0.95
COC	-12.2	96.8	204.4	219.0	-6.7	48.1	41.8	0.58
KWA	8.7	167.7	208.6	227.2	-8.2	36.0	28.9	0.10
LAU	-45.0	169.7	129.8	163.2	-20.5	38.8	33.4	0.97
LIN	52.1	14.1	103.7	130.7	-20.7	39.9	28.2	0.97
MAN	-2.1	147.4	256.4	198.4	29.2	61.9	58.0	0.36
REG	50.2	-104.7	152.4	157.6	-3.4	21.7	16.1	0.97
SBO	30.9	34.8	197.4	272.8	-27.6	87.4	75.4	0.95
TAM	22.8	5.5	241.3	261.5	-7.7	29.6	23.9	0.90
XIA	39.8	117.0	152.3	162.3	-6.1	42.5	38.4	0.86

Table 13: CERES monthly mean, BSRN comparison, bias is (CERES - BSRN)/BSRN, RMSE is the root mean squared error and MAE is the mean absolute error

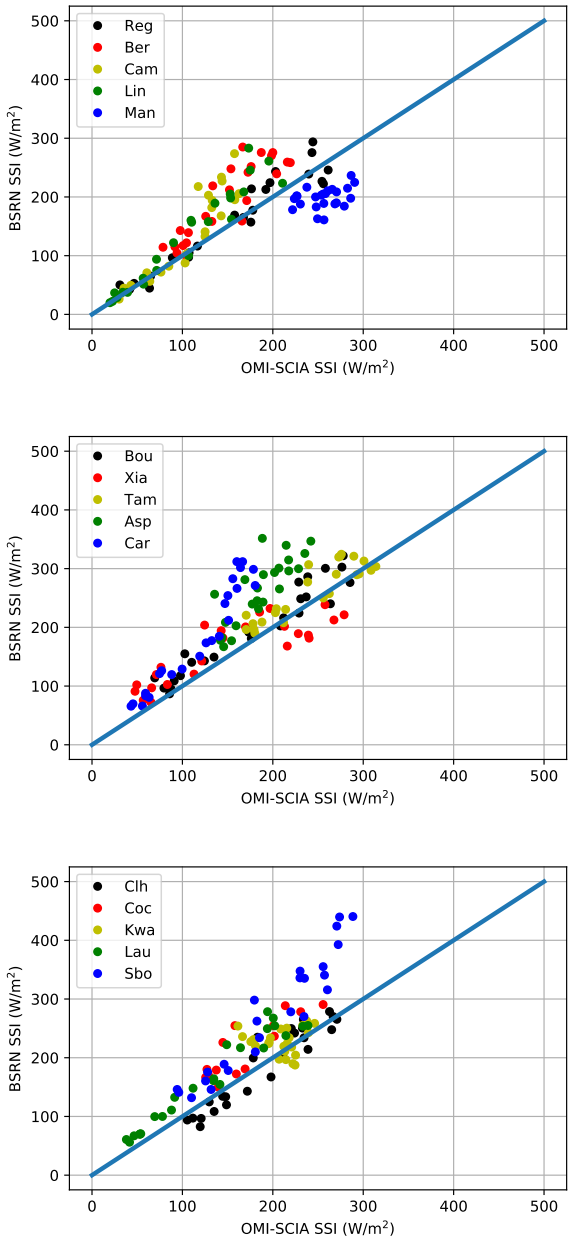


Figure 48: Scatter plots of the bsrn locations daily monthly means compared to the CERES monthly means.

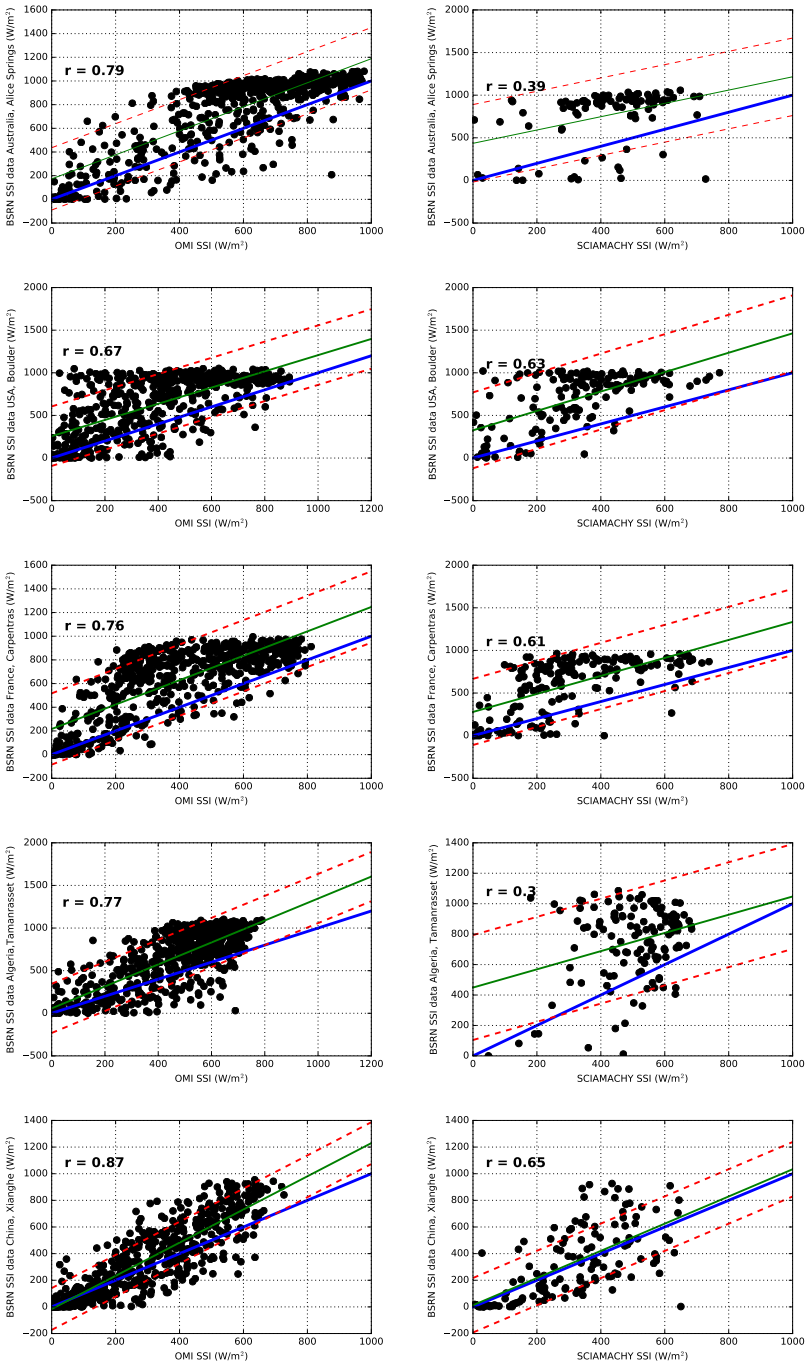


Figure 42: Scatter plots of direct SSI

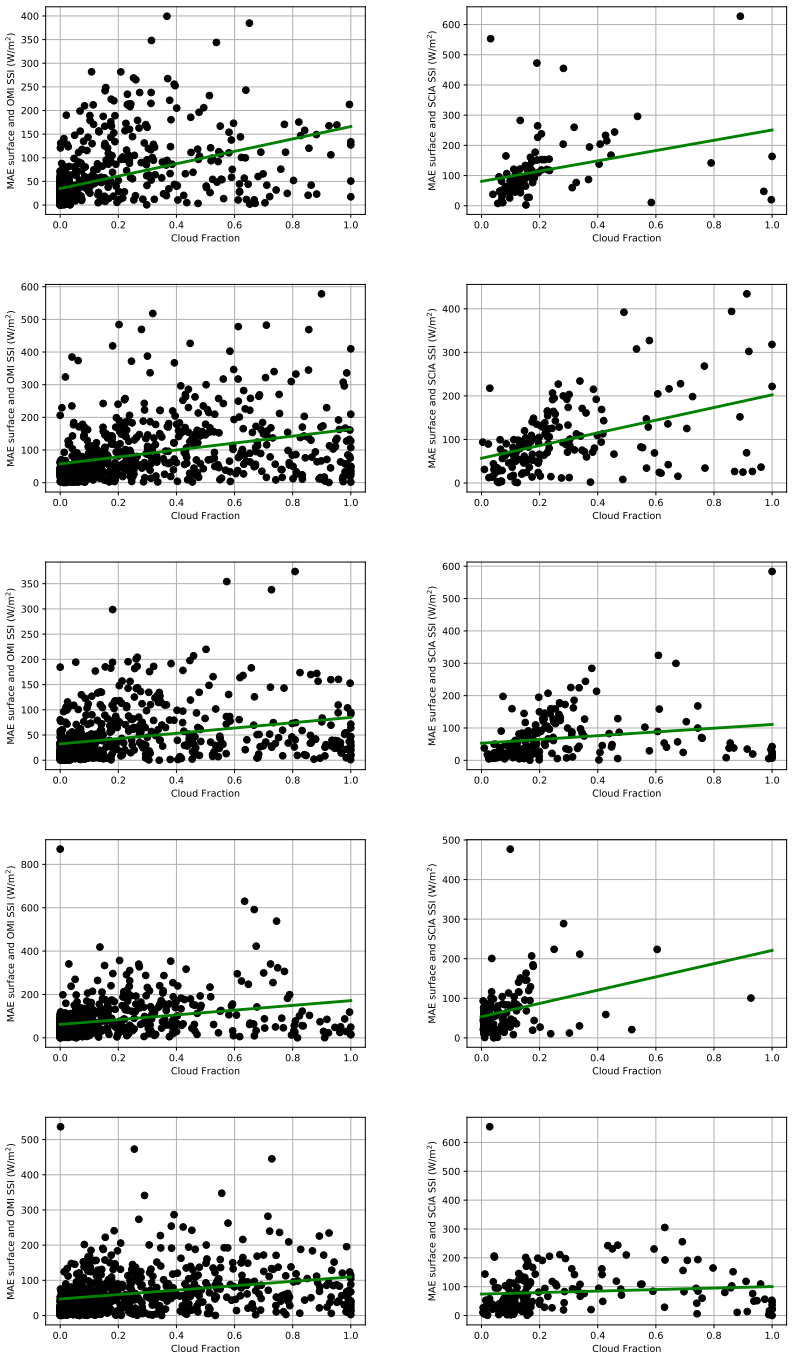


Figure 43: Scatter plots of MAE as a function of cloud fraction per location, left OMI-BSRN MAE, right SCIA-BSRN MAE

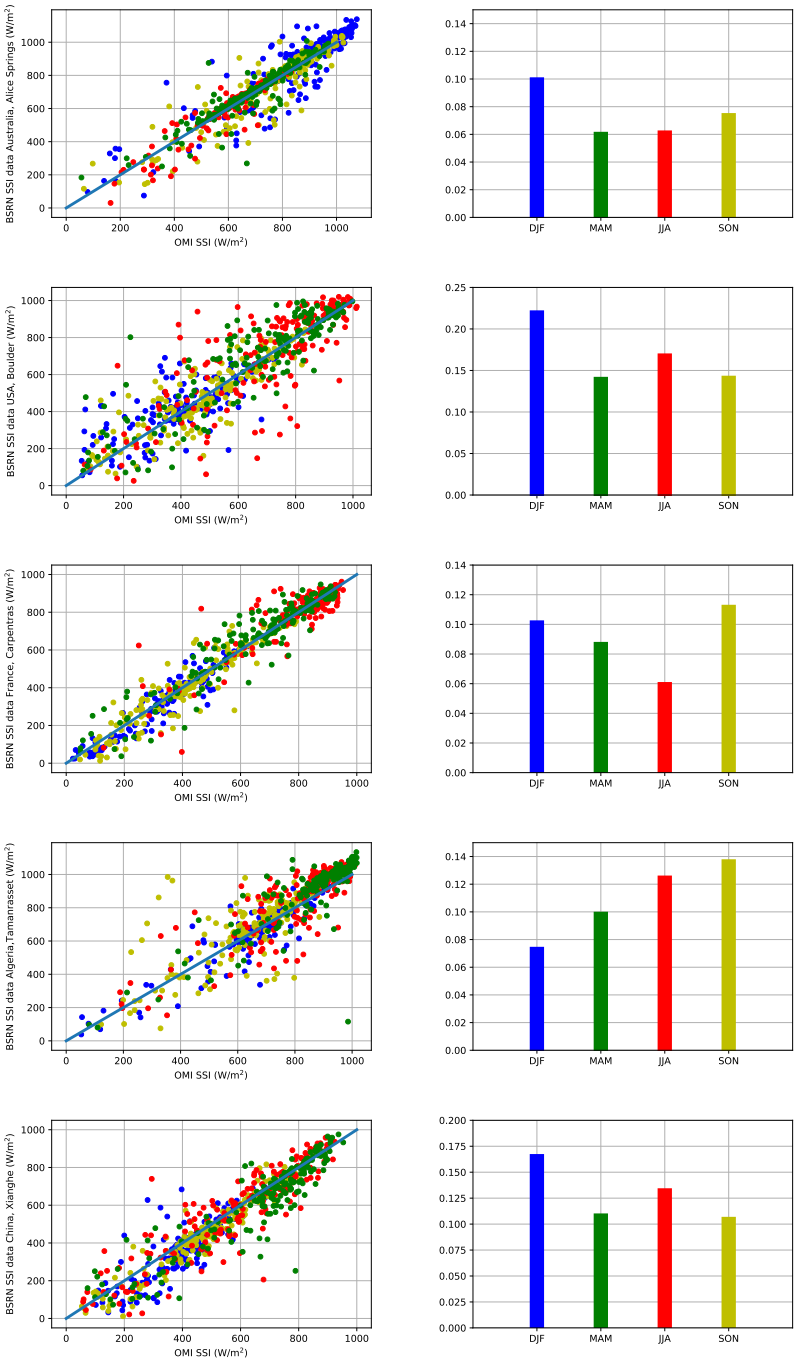


Figure 44: Scatter plots of seasonal OMI - BSRN SSI and the MAE/SSImean per location per season

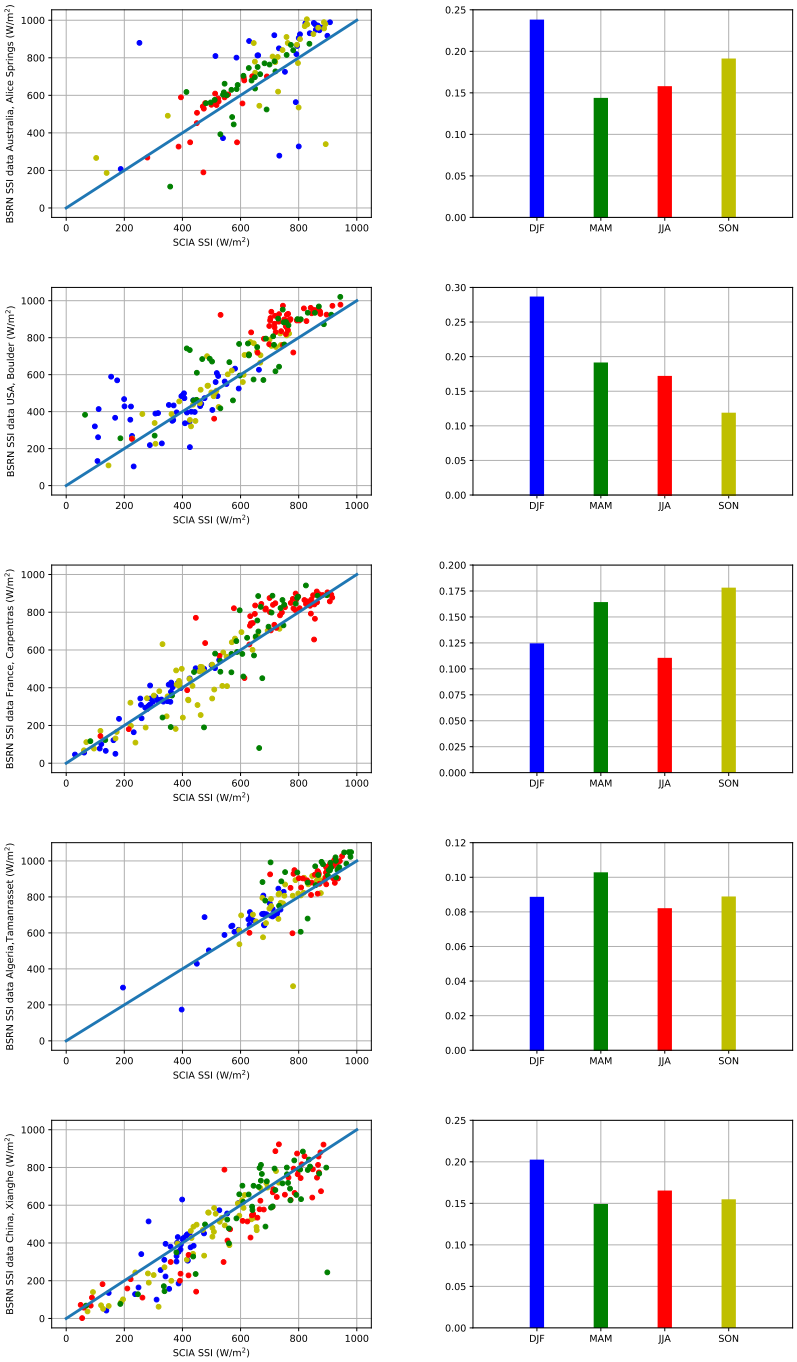


Figure 45: Scatter plots of seasonal SCIA - BSRN SSI and the MAE/SSImean per location per season