DIVISION OF GLOBAL RADIATION INTO DIRECT RADIATION AND DIFFUSE RADIATION

Master's thesis

Faculty of Science
University of Bern

presented by
Fabienne Lanini
2010

Supervisor:
Dr Stefan Wunderle
Institute of Geography, University of Bern

Co-Supervisor:
Dr. Michael Lehning
Institute for Snow and Avalanche Research WSL-SLF, Davos

Advisor:
Dr. Charles Fierz
Institute for Snow and Avalanche Research WSL-SLF, Davos

Aknowledgements
First, I thank Stefan Wunderle, Michi Lehning and Charles Fierz for supporting me in writing this thesis. They were always available for answering to my questions and providing interesting suggestions. I am especially thankful to Charles for his collaboration and the time and interest he took in my work.
Furthermore, many thanks go to Juliette Blanchet and Mathias Bavay, who provided great help for statistics and programming.

In addition, I would like to thank the people of SLF for the amazing atmosphere and the funny coffee breaks, which made of my stay in Davos an unique experience.

Thanks go to everybody who contributed to this work by providing data, information and suggestions, especially to Barbara Landl (MeteoSwiss) and Prof. Konrad Steffen (University of Colorado) who made at disposal radiation data for my work. Thanks also to Bruno Dürr for his cooperation.

I am extremely thankful to Christine Bollig for her patience by checking my English.
Finally, a special thanks to my family and my friends for their endless support. Grazie!

Abstract

In general, only measurements of the global radiation are known, but for many applications diffuse and direct fractions are also needed. For this reason, many models have been developed to establish correlations between the diffuse fraction and various predictors. However, at present, no reliable model which could be applied to worldwide stations, in particular to Swiss Alps, has been found.
The first purpose of this work is to gather radiation data from Swiss and worldwide stations with different geographical and climate conditions. The second goal is to present an overview of the different decomposition models and their limitations. The third goal, the most important part, is to test four models - Reindl-Helbig, Skartveit and Olseth, Boland-Ridley-Lauret and Maxwell - on a dataset with one-hour time steps.
The model proposed by Skartveit and Olseth is overall performing best, especially on data with a high clearness index.
It is shown that these models are poorly performing on three polar stations, where solar elevations are constantly low. This behavior reflects the limitation of the models on the estimation of the diffuse fraction for solar elevations lower than 30 °.
Furthermore, a model was developed by Perez on the basis of that of Maxwell. Applying a completely new method, Dürr proposed a further model. An analysis on those two additional models was conducted.
Potential predictors and adequate model structures are discussed: the clearness index k_t and solar altitude have been found to be significant predictors. A candidate for a universal model should include a variability index, because of the temporal weather evolution. Air mass and and aerosols have also to be considered, as they contain information about the atmospheric composition.
Finally, the models were performed on different time basis. For a smaller time interval, there is an increased variability of the data, but the performance still remains acceptable.

Contents

Appendix

List of Tables

List of Figures

Chapter 1

Introduction

Solar radiation is a crucial component of the global energy balance which drives different systems, such as the climate and hydrologic systems. In general, only measurements of the global radiation, I, are known, but for many applications diffuse and direct fractions, I_d and I_b, respectively, are also needed. For this reason, many models have been developed to establish correlations between the diffuse fraction and various predictors.

1.1 Background

The first studies on this subject were conducted by Liu and Jordan [1960]. They determined a relationship between daily diffuse and global radiation on clear days on a horizontal surface, by doing measurements on 98 sites in the US and Canada.
The portion of horizontal extraterrestrial radiation reaching the surface is defined as the clearness index

$$k_t = \frac{I}{I_0 \cos\theta_z},$$

where θ_z is the zenith angle. The diffuse fraction k_d is the portion of diffuse radiation composing the global radiation:

$$k_d = I_d/I.$$

Liu and Jordan found a correlation with k_t as predictor variable: $k_d = I_d/I_0 = f(k_t)$. The extraterrestrial value used is changing with latitude and time of year [Boland et al., 2001]. They also estimated the long term average hourly and daily sums of diffuse radiation and the daily sums of diffuse radiation for various categories of days for differing degrees of cloudiness. However, it was clear that the hourly diffuse fraction was not only a function of the clearness index.
Erbs et al. [1982] used a new database (from four US weather stations) composed of hourly direct normal radiation, $I_{b,norm}$, measured normal to the direction of the rays of the sun, global radiation to develop an estimation of the diffuse fraction of hourly, daily and monthly-average global radiation. They also determined the degree to which those relationships are dependent on season and location. A similar study has been conducted by Orgill and Hollands [1977], in Toronto (Canada).
On the other hand, Iqbal [1980] and Skartveit and Olseth [1987] used two predictors for their correlations, k_t and the solar elevation, ϕ. Skartveit and Olseth [1998] introduced then a new parameter: the hour-to-hour variability index, which considers the variation of the clearness index of the previous and the following time interval. They used hourly data from several European stations and avoided periods where a significant snow cover was present. Moreover, Garrison [1985] proposed a dependency of the diffuse fraction on the surface

albedo, atmospheric precipitable water, atmospheric turbidity, solar elevation and global horizontal radiation. But there was no further statistical analysis, so the importance of this study is unknown.
The atmospheric turbidity is the haziness in the atmosphere and the large size of aerosols is responsible for it. Aerosols can either absorb or scatter the radiation and, consequently, alter the energy balance of the Earth, especially under clear skies Gueymard [2005]. They also serve as condensation and freezing nuclei (see Appendix B.2).
Reindl et al. [1990] added two more significant predictors to the clearness index and the solar elevation: ambient temperature and relative humidity. So, they achieved to reduce the standard error of Liu- and Jordan- type models. They used stepwise regression to reduce a set of 28 potential predictor variables down to four significant predictors.

On the other side, Maxwell [1987] developed an exponential model for the estimation of the direct normal beam radiation from the global radiation. It is named the DISC model. This model was then improved by Perez et al. [1991]. They presented two different models: the first is a correction of the DISC-model and the second is composed of linear relationships. Surface temperature and relative humidity are used when available. Like Skartveit and Olseth, a hour-to-hour variability index was introduced. A four dimensional table of numerical constants to adjust the estimated values was introduced as well. The database was composed of records from 18 stations in the US and Europe. Other models were developed on the basis of the above mentioned ones. For example, Gonzalez and Calbo [1999] introduced a model based on the ratio between the diffuse and the extraterrestrial radiation, using data from Spanish stations.

Conversely, Boland et al. [2001] state that the predictability is not relevantly improved when taking solar elevation as predictor. They also show that the model for a 15 minutes interval is essentially the same as for hourly time intervals. In addition, they demonstrate that the apparent solar time is a good second predictor. They conducted the study in Australia, thus, it should be extended to other locations.
In 2009 a new model, named BRL-model (Boland-Ridley-Lauret-Model) was established, based on a logistic function (Ridley et al. [2010], Lauret et al. [2010]).

Dürr [2004] developed a new method to separate global shortwave incoming radiation into the diffuse and the direct fraction, in the context of the project APCADA (Automatic Partial Cloud Amount Detection Algorithm). This project estimates actual sky cloud cover with surface measurements of downward longwave radiation, temperature and relative humidity. Longwave downward radiation is measured by pyrgeometers, which would have to be permanently shaded to avoid influence of incoming shortwave radiation. The direct solar radiation has to be estimated in real time: it can be calculated from the difference between the measured global radiation and the estimated diffuse radiation. Marty [2000] developed a cubic spline method using the value measured at solar noon at cloud free situations to estimate the diffuse fraction. At present, the new method estimates the diffuse fraction using the transmissivity, q, of the incoming global radiation compared to the extraterrestrial radiation as the main parameter. It uses two-minutes interval data, which are then

averaged over 30 minutes.

Helbig [2009] presented a model for the shortwave surface radiation balance in complex terrain. For the diffuse radiation estimation, she applied a combination of two models proposed by Reindl et al. [1990]. Reindl I uses both the clearness index, k_t, and the solar elevation, ϕ, while Reindl II only considers k_t. The Reindl-Helbig model is an improvement on the Erbs model [Erbs et al., 1982]. However, it could clearly be seen that this decomposition model overestimates the measured sky-diffuse radiation. Therefore, Helbig stated that 'no reliable decomposition model was found that can be applied for the Eastern Swiss Alps. Thus, a statistical model valid at least for the Swiss Alps should be developed from locally measured diffuse sky and global radiation values.' [Helbig, 2009]
Nowadays, Reindl-Helbig is the model implemented in Alpine3D, which is a three dimensional model o surface processes. To know the amount of direct radiation reaching a slope o given exposition and incline, the horizontal global radiation measured by a nearby Automatic Weather Station must be decomposed into its diffuse and direct components. Then the direct radiation can be projected onto slopes [for snow and avalanche research), 2010].

The above mentioned models were compared to each other on different sites. As an example, Notton et al. [2004] compared various models on Ajaccio, a seaside Mediterranean site. A similar study was done by Gueymard [2005] and by Jacovides et al. [2006] in Cyprus.

1.2 Motivations

In order to have a complete overview on the existing decomposition models, this work is presenting a selection of models, which are thought to be the most representative. A detailed analysis and comparison of them is also reported.
To this end, this study is gathering radiation data from stations placed in different geographical and climatological sites.

This work is in particular focused on the answer of the following problems:

- What are the parameters, which are explaining the decomposition of the global radiation into direct and diffuse radiation?
- Is it possible to create a universal decomposition model? And what characteristics should a model candidate possess?

Furthermore, the performance of the model is investigated on data from Payerne for different time resolutions of the input data.

The work is organized as follows: in Chapter 2 the new dataset and the data quality control applied are described. Details of the stations can be found in Appendix A while a short overview of the most relevant solar equations is given in Appendix B. Different classes of models are presented and the most representative models selected in Chapter 3. Their performance is analyzed in Chapter 4. Appendix D summarizes the statistical methods

used while complementary figures and tables are in the Appendixes E, F and G. Results are discussed in Chapter 5 and Chapter 6 is drawing the conclusions and gives an outlook.

Chapter 2

Dataset

In order to develop an universal decomposition model, it is important to gather data of worldwide stations, covering different climate, geographical and meteorological conditions. For this reason, the dataset is composed by global and diffuse radiation for one-year time period (mostly for 2009) of 39 stations covering different regions of the globe.
Sunlight measurement values are available on different networks. For this study, measurements from 13 Swiss stations were offered by the Federal Office of Meteorology and Climatology MeteoSwiss and the Baseline Surface Radiation Network (BSRN) made data available data for 25 stations worldwide. Data from Greenland (Summit Station, BSRN candidate) were received from the University of Colorado.

In this chapter, the measurement data are first described and their representation is then explained. In order to have an homogeneous dataset, a data quality control is finally applied.

2.1 Data collection

Different organizations established a range of stations in different locations of the world. In Switzerland, the authorities require to record meteorological and climatological data throughout the Swiss territory [Seiz and Foppa, 2007]. For the worldwide stations, high quality data of solar radiation measurements are provided by the Baseline Surface Radiation Network (BSRN), initiated by the World Climate Research Programme (WCRP), jointly with the World Meteorological Organization (WMO), the International Council for Science (ICSU) and the Intergovernmental Oceanographic commission (IOC) of UNESCO.

2.1.1 Swiss stations

SwissMetNet is a project conducted by MeteoSwiss, aiming at the renewal and unification of ground measurements on the Swiss territory. 134 stations are going to be established by the year 2013, forming a network of automatic ground stations.
The dataset from this network is composed of 10-minutes interval measurements of 13 Swiss stations, which are listed in Table A.1. The period covered is one year, in general 2009, and the parameters at disposal are the global and diffuse radiation.
Stations with different climate and geographical characteristics were chosen. For instance, their altitude ranges from 413 m to 2 106 m. The stations are located in the Alps (e.g. Montana-Vermala), in urban areas (e.g. Geneva) or on the Swiss Plateau (e.g. Buchs).
Two different instruments are used to measure radiations. The SPN-1 Sunshine Pyranome-

ter [1] device measures diffuse and global radiation. Global radiation is detected by CM21 [2], as well.
In the ideal case, the two types of radiation should be recorded by the same instrument. However, the measure of global radiation used in the further analysis is from the CM21, because it covers the entire above mentioned time period. For the majority of the stations, SPN-1 started to measure global radiation only from the 2^{nd} July 2009 on.
Below, a comparison between the two instruments is made, in sort to justify the use of them.

Comparison instruments CM21 and SPN-1 for Global Radiation

Usually, CM21 is the official instrument for measuring the global radiation, while diffuse radiation is measured by SPN-1.
The accuracy between the two instruments have been tested, by comparing all the global radiation measurements for the entire period, as illustrated in Figure 2.1. The two measurement sets are clearly linearly correlated, with a coefficient of determination R^2 corresponding to 0.9846.
Then, the representation of the global radiation recorded with SPN-1 against the measurements with CM21 has been investigated for every station (see Figure 2.2, as an example). It has been seen that, in general, the data have a linear relationship. It means that the correlation is not location dependent. The station of Robiei has to be excluded from the dataset, having problems in the calibration of the instruments. Its measurements corresponds to the the line appearing parallel to the 1:1-representation in Figure 2.1.
Then, the same study has been done for the measurements comprised in every month. Figure 2.2 illustrates the results for a month, February, as an example. The same behavior is visible along all months, meaning no seasonal dependency.

It can be concluded that the two instruments are giving a similar result. For this reason, diffuse radiation measurements are hereafter from SPN-1 and global radiation from K&Z CM21.

2.1.2 Worldwide stations

The Baseline Surface Radiation Network (BSRN) is the global baseline network for monitoring the radiation field at the Earth's surface. It provides a high quality of solar radiation measurements. This network sets two principal objectives as follows [Lee, 2009]:

1. To provide data for calibration satellite-based estimates of the surface radiation budget (SRB) and radiation transfer through the atmosphere.

2. To monitor regional trends in radiation fluxes at the Earth's surface.

From the about 40 stations available, 26 where selected. Their description is listed in Table A.2. As an example, Figure 2.3 shows the station in Payerne.

[1] manifactured by Delta T
[2] manufactured by K&Z (Kipp & Zonen)

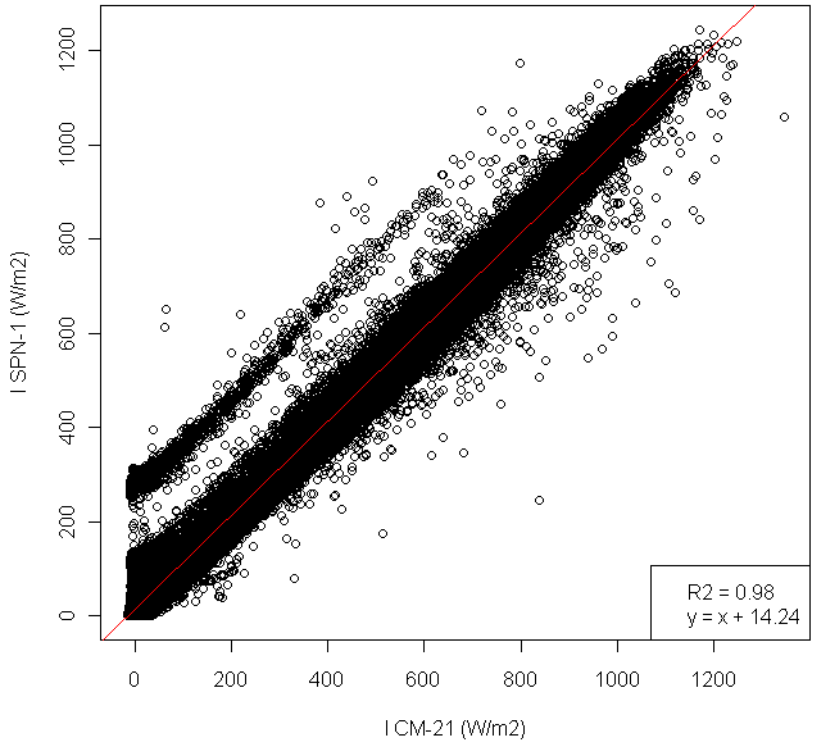


Figure 2.1: Diffuse fraction measured by SPN-1 against CM-21 measures. The red line represents the best fit. Its equation and the coefficient of determination R^2 are shown, as well. The line of measures parallel to the 1:1-line are from Robbia, because of calibrations problems. Those data were consequently excluded from the dataset.

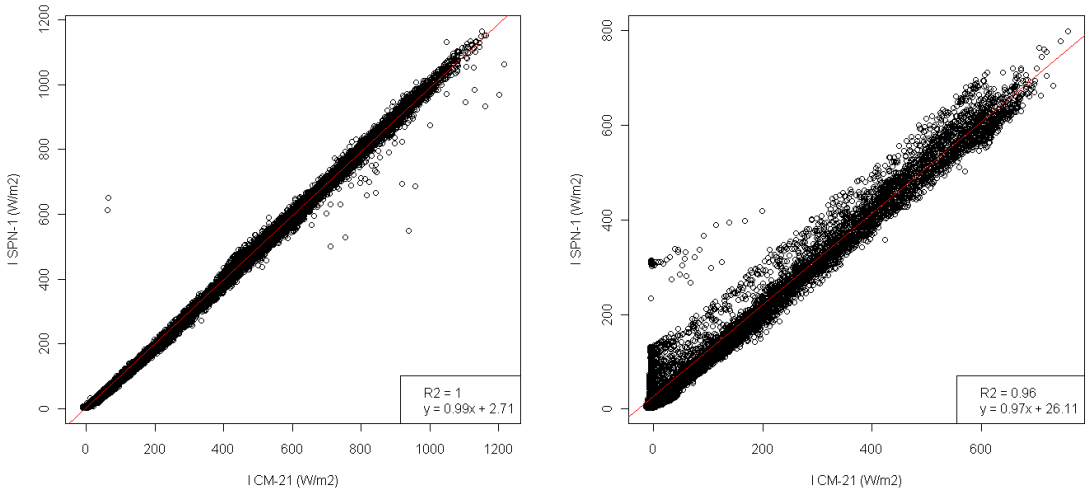


Figure 2.2: Scatterplot of global radiation SPN-1 against CM21 for the station Robbia (left) and for all stations measurements in February 2009 (right). The red lines represent the best fit. Its equation and the coefficient of determination R^2 are shown, as well.

Figure 2.3: BSRN Station in Payerne.

The dataset is composed of global and diffuse radiation measurements, for the 1-year time periods. The records are 1-minute intervals and detected by different instruments, depending on the individual choice of the station, observing BSRN standard requirements, which are listed in a manual edited by Kipp & Zonen [Lee, 2009]. For the global and diffuse solar radiation, the pyranometer must be ventilated and the body temperature monitored. Furthermore, both radiations have to be measured by the same type of pyranometer. The instruments have to be located on a place where the horizon is unobstructed in all directions. All obstructions (like mountains, buildings,...) should not exceed an elevation angle of 5 ° [Lee, 2009]. In addition, regarding the diffuse radiation, the detector has to be completely shaded by a ball or disk and the shading must correspond to a 5 ° field of view. The maximal uncertainty should not exceed 2 % or 3 W m^{-2} (diffuse radiation) and 5 W m^{-2} (global radiation).

2.2 Data quality control

The principal aim of this study is to determine a decomposition model where only the global radiation is measured. As a consequence, data for night periods are excluded. To have a positive global radiation I is, therefore, the first condition for the dataset:

$$I \geq 0. \tag{2.1}$$

Furthermore, for low solar elevations, instruments for the measurement of solar radiation have some limitations, because of the cosine response. For this reason, a threshold for the solar elevation ϕ has been set:

$$\phi \geq 5\,^{\circ}. \tag{2.2}$$

The radiation measurements were also checked against the quality controls listed by Jacovides et al. [2006] and proposed by the European Commission Daylight I, 1993:

$$\frac{I_d}{I} \leq 1.1 \tag{2.3a}$$

$$\frac{I}{I_{0,h}} \leq 1.2 \tag{2.3b}$$

$$\frac{I_d}{I_{0,h}} \leq 0.8 \tag{2.3c}$$

$$I \geq 5 \text{ W m}^2 \tag{2.3d}$$

$$I_b \leq I_{0,h}, \tag{2.3e}$$

where I is the global radiation, the sum of horizontal direct I_b and diffuse I_d radiation. $I_{0,h}$ is the horizontal extraterrestrial radiation.

Two additional conditions were introduced by Reindl et al. [1990] and lately used by Jacovides et al. [2006]:

$$\frac{I_d}{I} \geq 0.90 \text{ , for } k_t < 0.20 \tag{2.4}$$

$$\frac{I_d}{I} \leq 0.80 \text{ , for } k_t > 0.60 \tag{2.5}$$

For instance, during overcast weather situations , it can be thought that a large fraction of incoming radiation will be scattered by the clouds, resulting in a big diffuse fraction. By clear sky situation, a limit for diffuse radiation should also be placed.
As an example, 5.05 % of the total data of Altdorf (UR) were deleted because of Equations 2.1 and 2.2. Further 20.81 % and 2.74 % were excluded because of Equations 2.3a-2.3e and 2.4-2.5, respectively.

Horizon problem

As the Sun is obscured by mountains, the monitored radiation is entirely diffuse and useless for the aim of this study. For this reason, it can be omitted from the dataset. As a consequence, an additional condition for 10-minutes interval Swiss data was introduced: solar elevation has to be greater than the horizon angle,

$$\phi \geq \phi_{horiz}. \tag{2.6}$$

Horizon angle ϕ_{horiz} is the angle formed between the surface and the line connecting the meteorological station and the highest mountain located between the station and the Sun (see Figure 2.4).

This procedure was not applied to the worldwide data, because, due to standards requirements, BSRN stations are positioned in areas with no relevant problems of horizon (see Section 2.1.2).
For this purpose, an algorithm included in ALPINE-3D has been modified in order to determine the horizon angle for every solar azimuth. GIS is a set of tools permitting the selection of a desired area and to obtain its height profile for the chosen resolution. In this case, a resolution of 25 m was chosen and an area around each station including the surrounding mountains was selected. Then, for every station, the horizon angle ϕ_{horiz} was computed in each direction (1 ° interval) and defined as:

$$\tan \phi_{horiz} = \frac{h_{max} - h_0}{d} \tag{2.7}$$

where, as illustrated in Figure 2.4, h_0 is the altitude of the station, h_{max} the altitude of the highest mountain laying on the solar azimuth direction and d the horizontal distance between the meteorological instrument and the mountain peak.

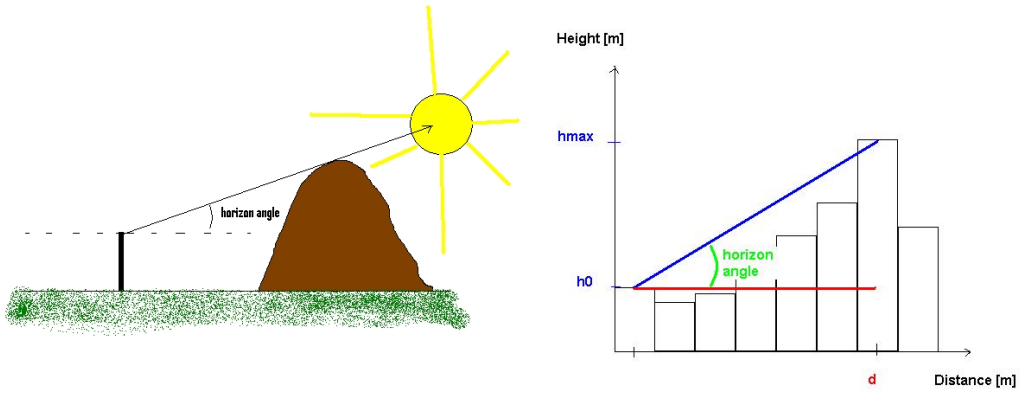


Figure 2.4: The horizon angle is the angle between the surface and the vector joining the meteo station and the highest mountain in the sun direction. For a given direction, it is calculated from a station situated at altitude h_0 and mountain with altitude h_{max}, situated at a distance d.

As predicted, only a small amount of data is deleted for stations located on the Swiss Plateu. On the other hand, a consistent number of measures is excluded for stations located at the bottom of a valley with determined characteristics. As an example, the stations of Visp (VS) and Altdorf (UR) are considered. Both stations are situated at the bottom of valleys, which have similar characteristics of depth but different orientations: Altdorf is situated in a North to South valley. Instead, Visp is in a East to West valley. Normally, the Sun is obscured when it has low elevation, i.e. in the morning and in the evening, when its position is in the East and in the West. For this reason, the mountains surrounding the stations have more influence for Altdorf than for Visp. Figure 2.5 illustrates the amount of data (in red) deleted by this last quality condition. It can clearly be seen that the station of Visp is barely perturbed by the surrounding mountains. Robbia (GR) had similar results as Altdorf, as the two valleys have similar orientations.

Hourly integrated data

After applying the above conditions, a dataset composed by 222 311 10-minutes interval Swiss data and 4 886 830 1-minute interval worldwide data is obtained. However, in order to conduct an homogeneous comparison and analysis of the models, the above mentioned dataset is transformed to a hourly integrated one. Figure 2.6 compares, for Payerne, the diffuse fraction against the clearness index for 1, 10, 30 and 60 minutes time steps.

It can clearly be seen how the data variation is decreasing with a higher time step. Furthermore, hourly data are obtained by the average of measures between the minutes 0 and 59. It is then recorded at the middle of the hour. The quality controls (Equations 2.3a-2.3e and 2.4-2.5) were then again applied to the integrated data.

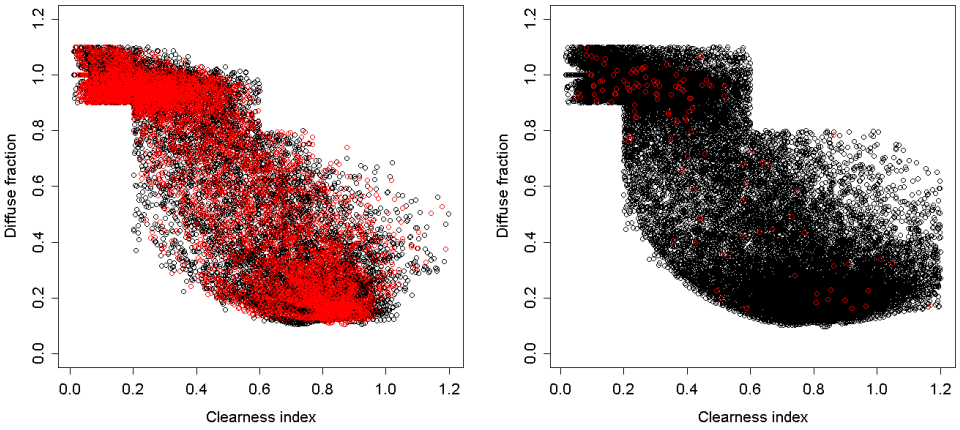


Figure 2.5: Representation of the radiation data of Altdorf (left) and Visp (right). In red, the measurements deleted by the horizon condition.

After the integration, the new dataset is composed by 135 694 hourly-integrated data.

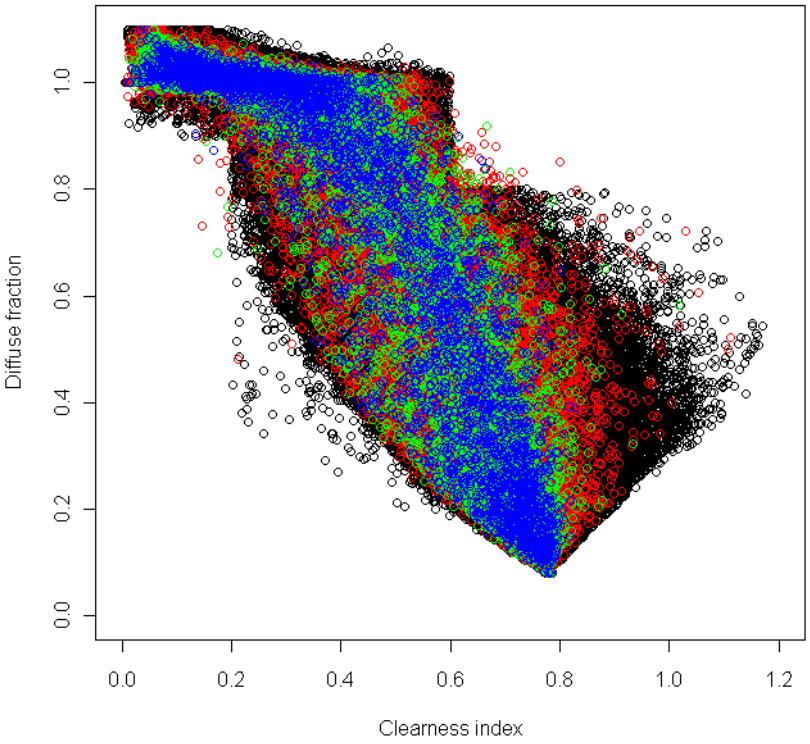


Figure 2.6: Comparison of diffuse fraction against clearness index with 1-minute (black), 10-minutes (red), 30-minutes (green) and hourly interval data (blue).

Chapter 3

Decomposition models

Nowadays, there is a general and continuous improvement of worldwide radiation networks and the data collection quality. However, for a desired location, more often only measurements of the global radiation I are known, but for many applications diffuse and direct fractions, I_d and I_b, respectively, are also needed. For instance, building engineering and solar thermal panels. For this reason, many models are available to establish correlations between the diffuse fraction and various predictors (e.g. Orgill and Hollands [1977], Boland and Ridley [2007] and Skartveit and Olseth [1998]).
Different methods have been developed in order to decompose the global radiation into the diffuse and direct fraction and several variables, as solar elevation and clearness index, are taken into account.

It is deduced that a great fraction of the scattered radiation is scattered by the water vapor molecules and it is shown that for the same direct radiation, the amount of diffuse radiation on cloudy days is higher than that for clear sky. This is due to additional scattering produced by clouds [Liu and Jordan, 1960].

In this chapter, four categories of models are established: polynomial, exponential, logistic and transmissivity models. Six models have been selected as representation of these groups and have been then analyzed. The models are listed in Table 3.1.

3.1 Polynomial models

First studies on decomposition models were conducted by Liu and Jordan [1960]. Using a data set of measurements from 98 sites in the US and Canada, they determined a relationship between daily diffuse and global radiation on clear days on a horizontal surface. They also estimated the long term average hourly and daily sums of diffuse radiation and the

Model	Type	Parameters
Reindl-Helbig	polynomial	k_t, Φ
Skartveit and Olseth	polynomial	k_t, Φ, ω_3
BRL	logistic function	k_t, ϕ, K_t, A_s, Ψ
Maxwell	exponential	k_t, ϕ, m_{air}
Perez	exponential	k_t, ϕ, m_{air}, $\chi(k_t', W, \Delta K_t', \theta_z)$
Dürr	regression tree	$q(m_{air}, \theta_z, k_t)$

Table 3.1: Summary of the models, their type and the parameters on which are based.

daily sums of diffuse radiation for various categories of days of differing degrees of cloudiness. They found a correlation with one predictor variable, the clearness index k_t, which is the proportion of extraterrestrial radiation I_0, reaching a location:

$$I_d/I = f\left(k_t\right).$$

This equation can be applied to different locations, where albedo and dust contamination are similar to those of the localities where it has been developed.
However, it was clear that the hourly diffuse fraction was not only a function of the clearness index.
Erbs et al. [1982] used a new database of hourly data from four US weather stations to develop an estimation of the diffuse fraction of hourly, daily and monthly-average global radiation. They also determined the degree to which those relationships depend on season and location. A similar study has been conducted by Orgill and Hollands [1977]. They developed a piecewise decomposition model, based on data from Toronto, Canada, depending on the clearness index k_t.
On the other hand, for their correlations, Iqbal [1980] and Skartveit and Olseth [1987] used two predictors, k_t and the solar altitude. Iqbal used data from three sites in Canada and two in France. Moreover, Garrison [1985] proposed a dependency of diffuse fraction on surface albedo, atmospheric precipitable moisture, atmospheric turbidity, solar elevation and global horizontal radiation. However, no further statistical analysis was made, so the significance of this study remains unknown.
Reindl et al. [1990] added other three significant predictors to the clearness index: solar altitude, surface temperature (2 m above the ground) and relative humidity. With this, they achieved to reduce the standard error of Liu and Jordan type models. They used stepwise regression on data from five locations in Europe and the US, to reduce a set of 28 potential predictor variables to four significant predictors.
In addition, seasonal correlations (wet and dry season correlations) have been developed by Jacovides et al. [2006] with a data set from a weather station in Cyprus.
Furthermore, Helbig et al. [2009] and Helbig [2009] presented a complet surface radiation balance model, where it was stated that 'no reliable decomposition model was found that can be applied for the Eastern Swiss Alps. Thus, a statistical model valid at least for the Swiss Alps should be developed from locally measured diffuse sky and global radiation values.' She developed a piecewise regression model, based on that of Reindl et al..

3.1.1 Reindl-Helbig

In addition to the four-variables decomposition model, Reindl et al. [1990] developed one additional model depending on solar altitude and clearness index and another depending on the clearness index only. These models are purely statistical and based on linear regressions.
Helbig [2009] developed a new decomposition model modifying the last two models. She then compared it to two other models, Erbs et al. and Reindl et al., demonstrating that Reindl*, as the new model is named, is better representing the observed diffuse fractions.
The dataset is composed of measurements taken by two weather stations in Davos and at

the Weissfluhjoch (Switzerland).
The decomposition model Reindl* is a piecewise regression model and is defined as it follows:

$$\begin{array}{ll} I_d/I = 0.1020 - 0.248k_t & \text{for } k_t \leq 0.3,\ I_d/I \leq 1.0 \\ I_d/I = 1.400 - 1.749k_t + 0.177\sin(\pi/2 - \theta_z) & \text{for } 0.3 < k_t < 0.78,\ 0.1 \leq I_d/I \leq 0.97 \\ I_d/I = 0.147 & \text{for } k_t \geq 0.78 \end{array} \tag{3.1}$$

It is composed of three intervals. For overcast situations, where the clearness index is low, it has a maximal diffuse fraction value and decreases linearly with an increasing clearness index; clouds are decreasing, corresponding to less reflection and consequently less diffuse radiation. It reaches a constant value for clear sky weather (high clearness index), where the diffuse radiation is at minimum. In the middle interval (partly cloudy sky), an additional variable is affecting the diffuse fraction: the solar elevation

$$\phi = \pi/2 - \theta_z.$$

A high solar elevation (with a corresponding high sinus value) contributes to the increase of the diffuse fraction. The diffuse fraction are plotted against the clearness index for seven different solar elevation ϕ in Figure 3.1.

A great disadvantage of this model is its composition of three piece-wise fits, which are not continue at the boundaries (for $k_t = 0.3$ and $k_t = 0.78$). Moreover, Reindl et al. [1990] found the best four variables out of 28, by an elaborate regression analysis. Several models were developed out of those variables. But Helbig was able to obtain a model, from a combination of those of Reindl. This new model, Reindl*, resulted to perform better than those of Reindl.

3.1.2 Skartveith and Holseth

In 1987, another polynomial model was developed by Skartveit and Olseth, also on the basis of Liu and Jordan. However, this model tended to overestimate the diffuse fraction under cloudless sky.

Lately, in 1998, a new model was established, based on hourly data and adding a new parameter: a hour-to-hour variability index σ_3. A variation of the model with the regional surface albedo as further parameter was developed, as well [Skartveit and Olseth, 1998].
The model is developed by using data measured in Bergen, Norway over a period of 32 years. The data are for the season from April to November, where no snow cover is present and a surface albedo of 0.15 is assumed. The model then was tested against other samples of data measured in four European stations.
The hourly variability index σ_3 is defined as the root mean squared deviation between the clear sky index of a period $\theta(t)$ and those of the former $\theta(t-1)$ and latter period $\theta(t+1)$:

$$\sigma_3 = \left\{ \frac{[\theta(t) - \theta(t-1)]^2 + [\theta(t) - \theta(t+1)]^2}{16^2} \right\}^{0.5} \tag{3.2}$$

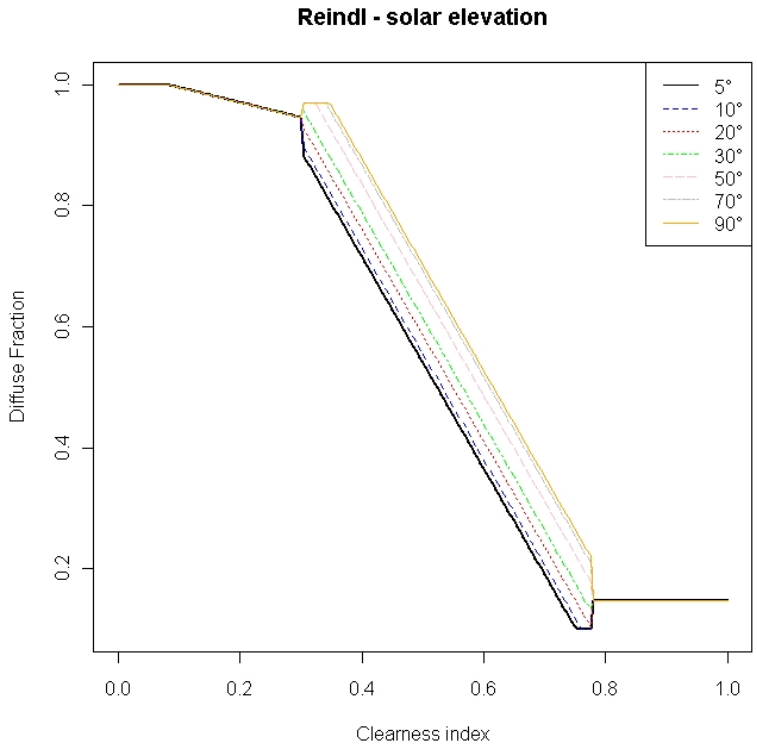


Figure 3.1: Diffuse radiation estimated by Reindl* against clearness index for seven solar elevations ϕ.

$$\sigma_3 = |\theta(t) - \theta(t \pm 1)|\,, \tag{3.3}$$

where the clear sky index is defined as it follows:

$$\theta = \frac{k_t}{k_{t,1}}, \tag{3.4}$$

and $k_{t,1}$ is the cloudless clearness index:

$$k_{t,1} = 0.83 - 0.56 \cdot exp(-0.06 \cdot \phi). \tag{3.5}$$

A low σ_3 corresponds to stable sky condition and is associated with overcast ($\theta \approx 0$) or cloudless sky situations ($0.9 < \theta < 1.0$). In this case, the model is separated in four situations, depending on the value of the clearness index k_t:

1. $k_t \leq 0.22$: solar elevation independent, no relevant direct radiation is observed,

$$k_d = 1. \tag{3.6}$$

2. $0.22 \leq k_t \leq k_{t,2}$: presence of broken clouds which substantially obscure the Sun. The diffuse fraction depends on the clearness index k_t and the solar elevation ϕ; it is defined as it follows:

$$k_d = f(k_t, \phi) = 1 - (1 - k_{d,1})(0.11\sqrt{K} + 0.15K + 0.74K^2) \tag{3.7}$$

where:

$$K = 0.5\left\{1 + sin\left[\pi\frac{k_t - 0.22}{k_1 - 0.22} - \frac{\pi}{2}\right]\right\}, \tag{3.8}$$

$$k_{t,2} = 0.95k_{t,1}, \tag{3.9}$$

$$k_{d,1} = 0.07 + 0.046\left(\frac{90 - \phi}{\phi + 3}\right). \tag{3.10}$$

3. $k_{t,2} \leq k_t \leq k_{t,max}$, where it is assumed that the direct radiation does not exceed a maximal value $k_{b,max}$:

$$k_{t,max} = \frac{k_{b,max} + k_{d,2}k_{t,2}}{1 - k_{t,2}}, \tag{3.11}$$

where:

$$k_{b,max} = 0.81^{\alpha}, \tag{3.12}$$

with

$$\alpha = (\frac{1}{sin(\phi)})^{0.6}, \tag{3.13}$$

is fitted by an extreme beam transmittance. It was modeled by an Arctic winter atmosphere with fixed precipitable water vapor and aerosols optical depth. It is assumed that clear sky prevails. The diffuse radiation is fixed:

$$k_d = k_{d,2}k_{t,2}\frac{1 - k_t}{k_t(1 - k_{t,2})}, \tag{3.14}$$

where $k_{d,2}$ is obtained from Equation 3.7:

$$k_{d,2} = f(k_{t,2}, \phi) \tag{3.15}$$

4. $k_t \geq k_{t,max}$: the diffuse radiation is assumed to be caused by clouds not obscuring the Sun, while the direct radiation remains fixed by $k_{t,max}$:

$$k_d = 1 - k_{t,max}\frac{1 - k_{d,max}}{k_t} \tag{3.16}$$

The diffuse fraction from Skartveit and Olseth against the clearness index is illustrated for a null variability index and for seven solar elevation in Figure 3.2.

On the other hand, high σ_3 is associated with scattered clouds (intermediate or extremely high θ). Consequently, for $\sigma_3 > 0$, an empirical term representing the variability due to inhomogeneous clouds is added to Equations 3.6, 3.7, 3.14 and 3.16:

$$\Delta(k_t, \phi, \sigma_3) = -3K_L^2(1 - k_L)\sigma_3^{1.3} \text{ , for } 0.14 \leq k_t \leq k_x \tag{3.17a}$$

$$\Delta(k_t, \phi, \sigma_3) = 3k_R(1 - k_R)^2\sigma_3^{0.6} \text{ , for } k_x \leq k_t \leq (k_x + 0.71) \tag{3.17b}$$

$$\Delta(k_t, \phi, \sigma_3) = 0 \text{ , for } k_t \leq 0.14 \text{ and } k_t \geq k_x + 0.71 \tag{3.17c}$$

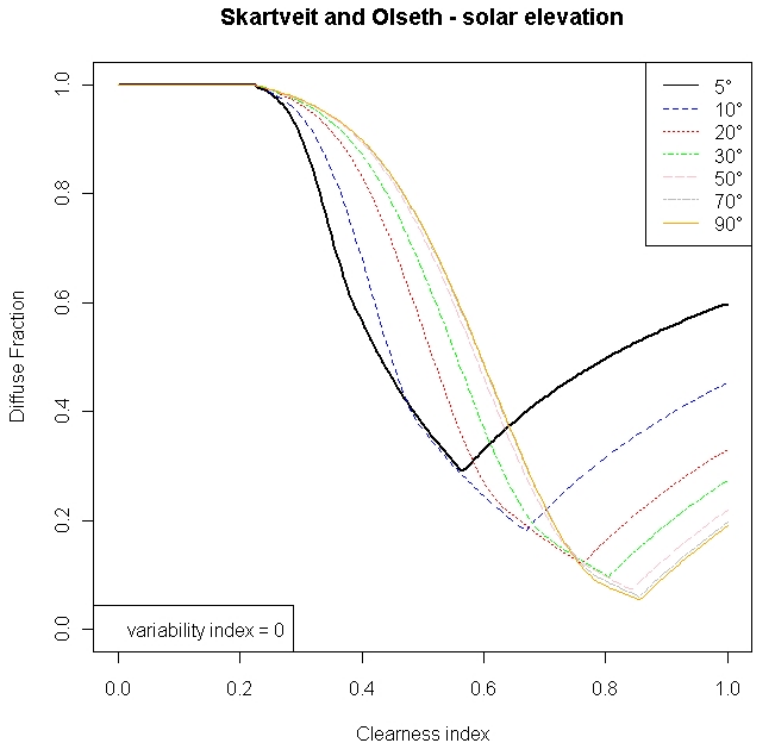


Figure 3.2: Diffuse fraction estimated,by Skartveit and Olseth in function of the clearness index, for seven solar elevations and a null variability index.

where:

$$k_x = 0.56 - 0.32e^{-0.06\phi} \tag{3.18a}$$

$$k_L = \frac{k_t - 0.14}{k_x - 0.14} \tag{3.18b}$$

$$k_R = \frac{k_t - k_x}{0.71} \tag{3.18c}$$

The estimated diffuse fraction is plotted against the clearness index for different variability indexes and solar elevations in Figure 3.3.
For increasing σ_3 and a low clearness index, the diffuse fraction is decreasing. In contrast, it is increasing for high clearness indexes. Furthermore, for clear sky situations (high clearness index), the variation of σ_3 is having an enormous influence on the diffuse fraction. Moreover, for partly cloudy and clear skies, the solar elevation is in general very important for the diffuse fraction for an altitude of less than 30°.

This new model takes into account the presence of clouds by including a variability index as variable, which is derived from the global radiation only. However, as already mentioned, this model was developed and then tested on data for a snow-free ground and a surface albedo of 0.15 was set. Because of multiple reflections caused by a bigger surface

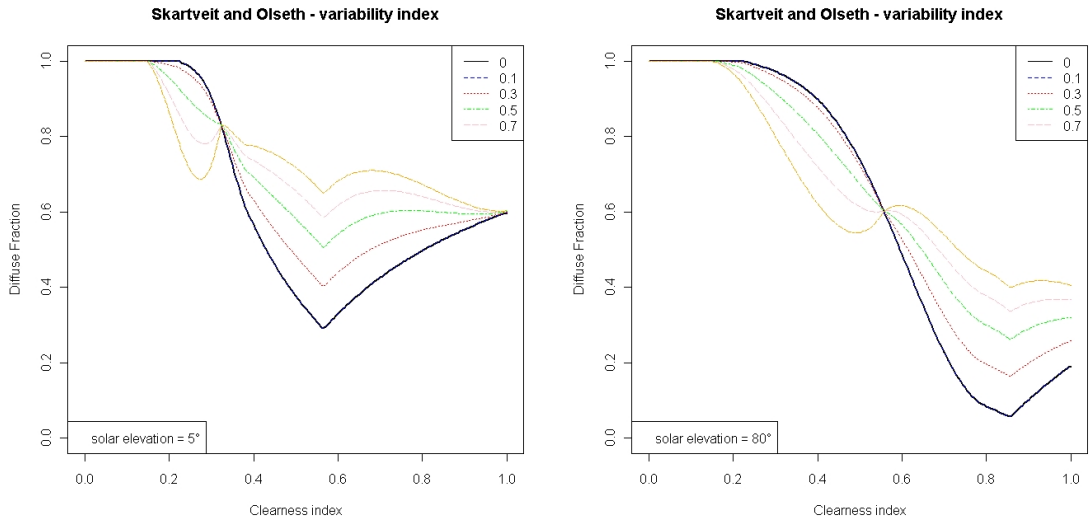


Figure 3.3: Diffuse fraction from Skartveit and Olseth for four index of variability and a solar elevation of 5 ° (left) and 80 ° (right).

albedo, a correction term was proposed. However, the correction was not verified, since it is extremely difficult to determine a precise value of albedo [Skartveit and Olseth, 1998].

3.2 Logistic function models

Models developed in North America and Europe seemed to be inefficient for the Southern Hemisphere. For this reason, in Australia, a completely different approach has been taken. Boland et al. [2001] developed a model for Australian conditions, using a logistic function. The dataset was composed of hourly data from a station in Geelong, Victoria.
It is one of the first models developed from data of stations with a latitude greater than 40 °. A line is fitted on the transformed data, by using a statistical method to find an estimation of the slope [Sen, 1968]. The fitted line is then back-transformed. Finally, a logistical function is fitted on the back-transformed line of best fit (Boland et al. [2008], Boland and Ridley [2007]). In 2008, statistical techniques were used to justify the use of a logistic function as a model.
This model was also tested in Cyprus by Jacovides et al. [2006].
In addition, a method for identifying incorrect values was constructed, using a quadratic programming, where outliers are eliminated [Boland and Ridley, 2007]. The algorithm has been applied to various localities, in various parts of the world, showing satisfactory results. Furthermore, the effect of adding further parameters to explain the data was investigated. Consequently, in 2010 a multiple predictor model, was proposed as an universal model, named Boland-Ridley-Lauret (BRL) Model (Ridley et al. [2010], Lauret et al. [2010]). It is simpler than the models proposed previously by the same authors.

3.2.1 Boland-Ridley-Lauret (BRL)

First introduced by Ridley et al. [2010], the BRL model is based on hourly data from stations of Australia and other worldwide stations. Hourly clearness index k_t, daily clearness index

$$K_t = \frac{\sum_{j=1}^{24} I_j}{\sum_{j=1}^{24} I_{0,j}}, \tag{3.19}$$

the solar elevation ϕ, apparent solar time A_s and a measure of persistence of global radiation level Ψ are the variables used in this model. It was assumed that the Apparent Solar Time (AST) A_s, which is asymmetric with regard to the solar noon, could explain differences between morning and afternoon.
Moreover, the persistence factor is defined as it follows :

$$\Psi = \begin{cases} \frac{k_{t-1}+k_{t+1}}{2} & \text{sunrise} < \text{t} < \text{sunset} \\ k_{t+1} & \text{t} = \text{sunrise} \\ k_{t-1} & \text{t} = \text{sunset} \end{cases} \tag{3.20}$$

Lauret et al. [2010] completed the work by determining the coefficients in a Bayesian framework.
The extension of the single predictor model results then as follows [Lauret et al., 2010]:

$$k_d = \frac{1}{1 + e^{-5.32+7.28k_t-0.03A_s-0.0047\phi+1.72K_t+1.08\Psi}} \tag{3.21}$$

The use of K_t, the daily clearness index, is necessary, because the whole day may have a particular pattern.

Significant predictors should be independent from each other. However, even being highly correlated, k_t, K_t and Ψ are considered as good predictors. The Apparent Solar Time has a small coefficient and is only slighly correlated to k_t and is, therefore, considered as a relevant parameter. The solar elevation ϕ is correlated with the other variables, but was nonetheless considered as a significant predictor [Boland et al., 2008].

The estimated diffuse fraction is plotted against the clearness index in Figure 3.4, for different values of the parameters. The influence of solar elevation and Apparent Solar Time are relatively small and correspond to the magnitude of the coefficients in the model. For a constant clearness index, the diffuse fraction is larger for small solar elevations. This is physically possible, as it corresponds to sunset and sunrise times, where radiation is known to have a great diffuse fraction. A similar behavior is observable for the Apparent Solar Time, which, at constant clearness indexes, is decreasing for a smaller diffuse fraction. The daily clearness index and the persistence factor prevalently have a great sensibility for the middle interval of the clearness index, where the variation of the diffuse fraction is relatively large and decreasing for an increasing daily clearness index and persistence factor.

An advantage of this model is that it is formed by a single equation, in contrast to the previous model category. However, the time of computation of this model is extremely long for 1-minute interval data.

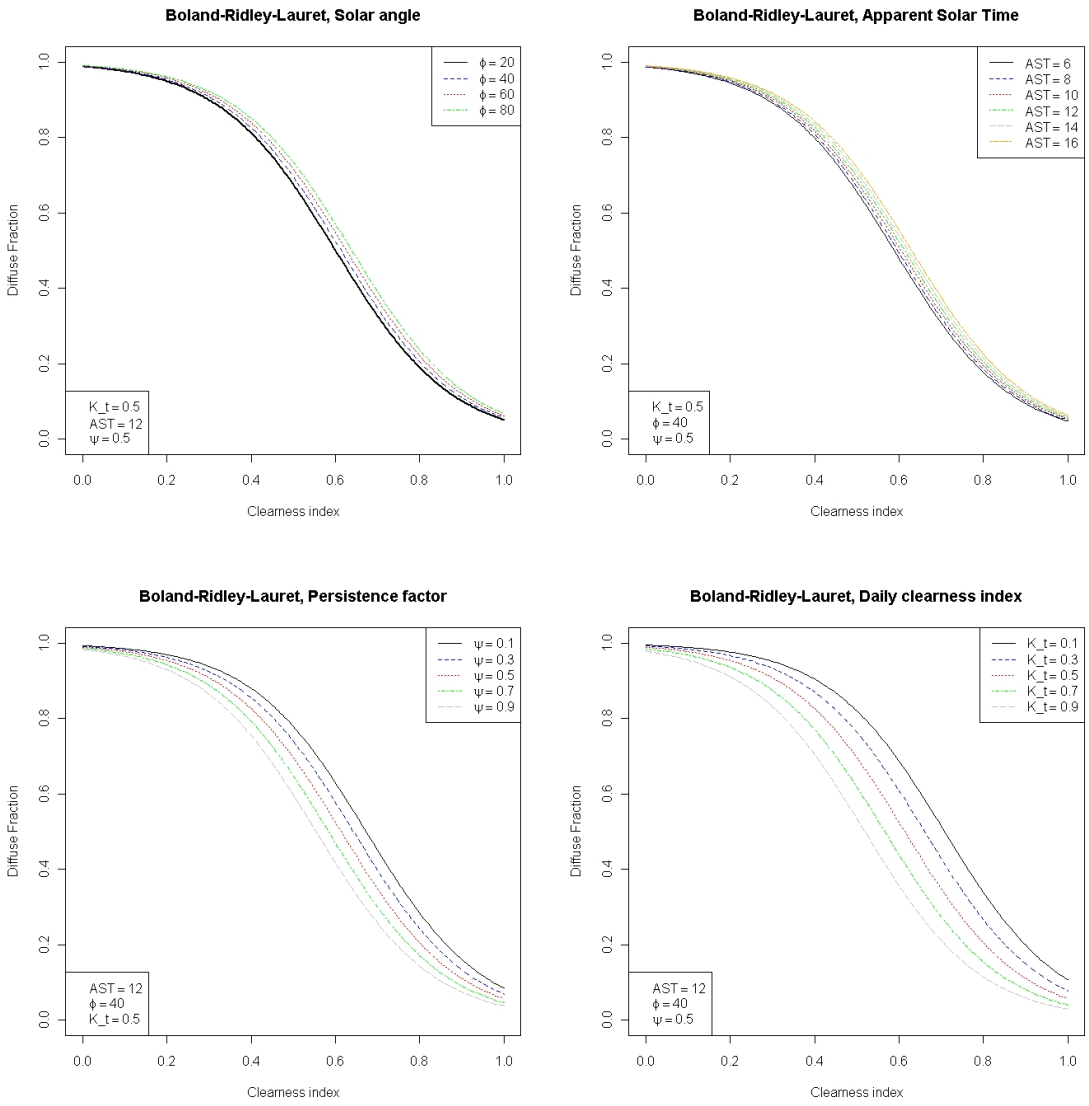


Figure 3.4: Diffuse fraction from BRL-model for four 4 solar elevations ϕ , 6 AST values, 5 persistence factors and 5 daily clearness index.

3.3 Exponential Model

In contrast to the other statistical models, Maxwell [1987] tried to propose a quasi-physical model to estimate the normal direct radiation I_b. Physical principles are used to develop the model and this algorithm has been obtained by the development of a computer program named the Direct Insolation Simulation Code (DISC) [Maxwell, 1987]. Later, two models were developed by Perez et al.. The first has coefficients correcting the DISC-model; the second is air mass dependent and includes a solar position-independent direct transmittance. When available, the surface dew temperature T_d is used. Moreover, two limitations of the clearness index as a variable were emphasized: its solar elevation dependency and the impossibility to differentiate two weather conditions with the same global radiation [Perez et al., 1990b].

3.3.1 Maxwell

Maxwell [1987] developed a model which predicts hourly normal beam irradiation I_b from hourly global irradiation values. It combines a physical clear sky model with empirical fits for other weather conditions. An exponential function is used and the data are from four US stations.

Air mass m_{air} is the principal parameter, in order to detect diurnal variations in insolation. The clearness index is the second parameter.

At first, a maximum for the clearness index for cloudless sky $K_{n,c}$ is calculated (Maxwell [1987], Perez et al. [1990c]):

$$K_{n,c} = 0.866 - 0.122 m_{air} + 0.0121 m_{air}^2 - 0.000653 m_{air}^3 + 0.000014 m_{air}^4 \tag{3.22}$$

where m_{air} is the air mass:

$$m_{air} = \frac{1}{sin(\phi) + \frac{0.50572}{(\phi + 6.07995)^{1.6364}}}, \tag{3.23}$$

depending on the solar elevation ϕ.
Secondly, a reduction ΔK_n of this maximum is applied:

$$\Delta K_n = a + b e^{c\, m_{air}}, \tag{3.24}$$

where a, b, c are determined for two ranges of clearness index:

$$\text{if } k_t \leq 0.6 \begin{cases} a = 0.512 - 1.560 k_t + 2.286 k_t^2 - 2.222 k_t^3 \\ b = 0.370 + 0.962 kt \\ c = -0.280 + 0.932 k_t - 2.048 k_t^2 \end{cases} \tag{3.25}$$

$$\text{if } k_t > 0.6 \begin{cases} a = -5.743 + 21.77 k_t - 27.49 k_t^2 + 11.56 k_t^3 \\ b = 41.4 - 118.5 k_t + 66.05 k_t^2 + 31.9 k_t^3 \\ c = -47.01 + 184.2 k_t - 222 k_t^2 + 73.81 k_t^3 \end{cases} \tag{3.26}$$

The direct radiation transmittance is then calculated:

$$K_n = K_{n,c} - \Delta K_n \tag{3.27}$$

Finally, the direct normal radiance is determined and the diffuse radiation can be derived:

$$I_{b,DISC} = I_0 \cdot K_n \tag{3.28a}$$

$$I_d = I - I_{b,DISC} \tag{3.28b}$$

3.3.2 Perez

Perez et al. [1991] proposed two different models, developed with dataset composed of hourly measurements from stations located in the US and Europe (Geneva included).
The first model is a correction of Maxwell [1987] and is defined as it follows:

$$I_b = I_{b,DISC} \cdot \chi(k'_t, \theta_z, W, \Delta k'_t) \tag{3.29}$$

where χ is a 6x6x5x7 matrix (see Appendix C) which corrects the DISC-model and is based on the bins listed in Table 3.2 for each one of the four parameters: the solar zenith angle θ_z, the normalized clearness index k'_t, the precipitable water W and the stability index $\Delta k'_t$.
The normalized clearness index is defined as it follows:

$$k'_t = \frac{k_t}{1.031 \cdot \exp^{\frac{-1.4}{0.9+\frac{9.4}{m_{air}}}+0.1}} \tag{3.30}$$

where m_{air} is the air mass defined in Equation 3.23.
In Equation 3.30, the clearness index becomes zenith angle-independent, by dividing it by a simple direct irradiance attenuation formula [Perez et al., 1990b]. The term 1.4 in Equation 3.30 corresponds to the Linke turbidity.
Similarly to Skartveit and Olseth [1998], a variability index is introduced. In fact, the third dimension, the stability index $\Delta k'_t$, has low values for stable conditions and high values for unstable conditions:

$$\Delta k'_t = \frac{\left[\left|k'_{t,i} - k'_{t,i+1}\right| + \left|k'_{t,i} - k'_{t,i-1}\right|\right]}{2}, \tag{3.31}$$

where $(i+1)$ and $(i-1)$ are the former and latter hour interval. If one of the two points is missing, Equation 3.31 becomes:

$$\Delta k'_t = \left|k'_{t,i} - k'_{t,i\pm1}\right|. \tag{3.32}$$

This term should permit make a difference between partly cloudy conditions, where there are jumps from one hour to the other, and stable conditions (such as overcast weather situations or haze), where the jumps are subtiles.
The precipitable water W is expressed as it follows:

$$W = \exp_{24}^{0.07 \cdot T_d - 0.075} \tag{3.33}$$

Bin nr.	k'_t	θ_z (°)	W (cm)	$\Delta k'_t$
1	0.00-0.24	00-25	0-1	0.000-0.015
2	0.24-0.40	25-40	1-2	0.015-0.035
3	0.40-0.56	40-55	2-3	0.035-0.070
4	0.56-0.70	55-70	3-∞	0.070-0.150
5	0.70-0.80	70-80	0-∞[1]	0.150-0.300
6	0.80-1.00[2]	80-90[3]		0.399-1.000
7				0.000-1.000[4]

Table 3.2: Bins used for the function $\chi(k'_t, \theta_z, W, \Delta k'_t)$ [Perez et al., 1991].

where T_d is the surface dew temperature. It has small time or space gradients. Therefore, data from neighboring weather stations could be applied. It is an estimator of precipitable water, which influences absorption (direct vs. diffuse) and aerosol growth (scattering and direct vs. diffuse) [Perez et al., 1991]. The influence of precipitable water is less marked but not to be neglected.
In the absence of measurements, the dew temperature T_d can be obtained from the relative humidity R_h [WMO, 2008]:

$$T_d = \frac{243.12\ H}{17.62 - H} \tag{3.34}$$

with:

$$H = ln\left(R_h + e^{\frac{17.62 T_{air}}{243.12 + T_{air}}}\right) \tag{3.35}$$

where T_{air} is the surface temperature, which is the air temperature measured 2 m above ground.

A second model is a set of linear relationships, expressed in terms of solar position-independent direct transmittance index K'_b:

$$I_b = \frac{I_0\ K'_b\ \exp^{\frac{-1.4}{0.9 + \frac{9.4}{m_{air}}}}}{0.87291} \tag{3.36}$$

where, for $k'_t < 0.2$, $K'_b = 0$, otherwise:

$$K'_b = a(Kt', Z, W, \Delta Kt') \cdot k'_t + b(k'_t, Z, W, \Delta k'_t) \tag{3.37}$$

The coefficients a and b are determined for different bins. The diffuse radiation can then be derived from Equation 3.28b.

Both models are considering the former and the latter time interval data, which permits to have an overview on the weather stability. However, the models are complicated and not

[1] when W is not available.
[2] values of k'_t superior to 0.87 should be treated like suspected for elevations above 1 000 m.
[3] the model was validated just for $\theta_z < 85\,°$.
[4] when $\Delta k'_t$ is not available.

easily transportable. Furthermore, not every station monitors T_d and R_h with the same frequency as solar radiation data.
For the present study, the first model is taken into consideration.

3.4 Transmissivity Model

The diffuse radiation can be measured by pyranometers, where the domes were previously shadowed for 10 minutes at solar noon. Then, the diffuse radiation for the entire day was estimated by applying a cubic spline interpolation on the measured values at solar noon at cloud free situations. In 2000, a model was developed by Marty where internal and external perturbations were considered. It was stated that solar transmittance varies with water vapor content and solar altitude, which are season and location dependent. For instance, there is a high solar transmittance for low water vapor content and high solar altitude.
Since the project APCADA (Automatic Partial Cloud Amount Detection Algorithm) needs a real-time correction of the direct solar radiation, Dürr [2004] developed a new method to separate the global shortwave incoming radiation into the diffuse and the direct fraction, by estimating the diffuse fraction using the transmissivity q of the incoming global solar radiation to solar radiation at the top of the atmosphere.

3.4.1 Dürr

APCADA needs real time correction of the direct solar radiation and allows to detect the effect of direct solar radiation on Longwave Downward Radiation (LDR) at noon. A fixed shadow band should be installed on PIR pyrgeometers. However, it is impossible to permanently shade ASRB PIRs, because of the difficult weather conditions in the Alps.
This new approach estimates diffuse radiation I_d from 2-minutes interval data of global radiation. It is a regression model and was developed out of data from Swiss stations located in Payerne, Davos and Versuchsfeld (Davos). It is based on the analysis of the transmissivity q, which is defined as:

$$q = (\frac{I}{I_0 \cdot cos\theta_z})^{1/m_{air}} \tag{3.38}$$

The distinction of different cloud situations is made, by comparing the transmissivity average q_{avg} over 10 minutes to the lower limit of q for cloud-free conditions:

$$q_{diff} = q_{avg} - q_{lim} \tag{3.39}$$

where:

$$q_{lim} = \Delta q + (0.98 - \Delta q)(\frac{t_i - \bar{t}}{\bar{t} - t_0})^4. \tag{3.40}$$

t_i is the time between the sunrise (t_0) and the sunset (t_n) and $\bar{t}$ is the time at solar noon:

$$\bar{t} = t_0 + \frac{t_n - t_0}{26 \quad 2}.$$

The limit of transmissivity is depending on a transmissivity offset Δq which is altitude and season dependent. In fact, the diffuse radiation is increasing in winter because of the low solar elevation and the increased surface albedo. With the albedo offset Δa, the transmissivity offset is for every station separately calculated. Every parameter is station specific and is empirically fitted on measurement data from free-clouds days (Dürr, personal communication, 2010).
Furthermore, a nearly zero transmissivity difference (Equation 3.40) with a low variability reflects a blue sky situation. Whereas, a difference going to infinity and a low variability correspond to overcast weather situations.
The variability (within the past 30 minutes) is expressed with a standard deviation q_{std} and a range q_r. In order to determine the standard deviation, the previous 15 2-minutes interval data of q are fitted and then the residuals are fitted as well. To reduce small fluctuations of the global radiation, the difference between the minimal and maximal residual is considered, as well.
Afterwards, the model has been simplified (Dürr, personal communication, 2010).

At first sight, this regression model appears complicated. It is covering all weather situations considering the previous 30 minutes. However, parameters have to be calculated for every single stations separately, which leads to a difficult transportability of the model. Furthermore, it is demanding to apply such a model to data with a time interval greater then that used for the development (e.g. 30-minutes, hourly data).

Chapter 4

Results

Figure 4.1 shows the global and the diffuse radiation over one day for Chasseral (NE). If the estimated diffuse radiation were represented on the same graphic, it would be difficult to assess either qualitatively or quantitatively its ability to reproduce measurements.

The representation of the diffuse fraction as a function of the clearness index is thus used in the literature (e.g. Boland and Ridley [2007], Helbig et al. [2009], Orgill and Hollands [1977] and Perez et al. [1990a]) and gives a clearer and better understanding of the data. An example is illustrated in Figure 4.2.

The diffuse fraction k_d is the ratio of diffuse to global radiation:

$$k_d = \frac{I_d}{I}, \tag{4.1}$$

and the clearness index k_t the portion of extraterrestrial radiation reaching the surface:

$$k_t = \frac{I}{I_0 \cos\theta_z}. \tag{4.2}$$

Both parameters range between 0 and 1 and their combinations describe different weather situations. A low clearness index implies, for instance, a small portion of radiation reaching the surface, reflecting an overcast weather situation and, therefore, a high diffuse fraction. Whereas, a high clearness index represents a clear sky, with small diffuse radiation and, for instance, a low diffuse fraction. The intermediate values of clearness index represents a partly-cloudy sky. For this interval of the clearness index, the corresponding possible values of the diffuse fraction is extremely hard to determine, because of possible multiple reflections between the surroundings and the clouds. Erbs et al. [1982] stated that, normally, 90% of the data are included in the interval $0.2 \leq k_t \leq 0.8$. The dataset used for this work only has the 78%.

Furthermore, clouds change the diffuse fraction. For example, extensive clouds (like in an overcast weather situation) reduce the amount of beam radiation reaching the surface. Broken clouds, instead, may enhance the diffuse fraction without affecting the intensity of the beam radiation [Skartveit and Olseth, 1998]. The stratospheric sulphate aerosols content is higher in the Northern than in the Southern Hemisphere. This fact produces, in the Northern Hemisphere, a higher proportion of values with high clearness index and corresponding high values of diffuse fraction. Gueymard [2005] says that turbidity is the most important extinction affecting the direct and diffuse radiation under clear sky [Boland et al., 2008].

In 1991, there was an enormous eruption event of Mt Pinatubo. An important amount of sulphate was injected into the stratosphere, which resulted in a could covering most of the Earth, affecting the amount of diffuse fraction [Molineaux and Ineichen, 1996]. Moreover, an

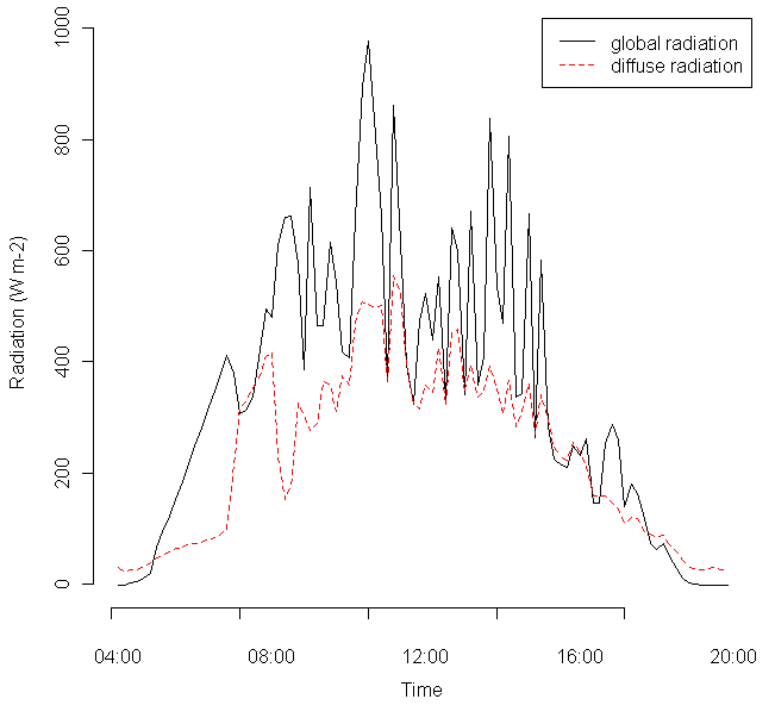


Figure 4.1: Representation of the global radiation and its diffuse part, on 1^{st} May 2009 in Chasseral (NE).

extremely high clearness index with a corresponding intermediate diffuse fraction could be produced by unobstructed Sun and the presence of scattered clouds [Skartveit and Olseth, 1998].

The data do not present any seasonal dependency. As an example, Figure 4.3 shows the usual data representation for Pilatus (LU) for the different seasons. For an increasing solar elevation, high and extremely high clearness indexes are becoming rare (see Figure 4.4).

The majority of the data with a high clearness index ($k_t > 0.8$) were measured at morning times, showing a pronounced forenoon-afternoon dependence (see Figure 4.5). In the morning, there is less water vapor content in the atmosphere and lower turbulent mixing. The surface needs time to become warmer and drive turbulences. Moreover, the boundary layer is usually shallower in the morning. As a consequence, the air is clearer at this time of the day, the path through through the boundary layer can be shorter and the scattering is then reduced.

4.1 Clear days representation

In order to investigate clear day situations, five stations with different climate and geographical characteristics were selected: Geneva (Switzerland, urban area), Georg von Neumayer

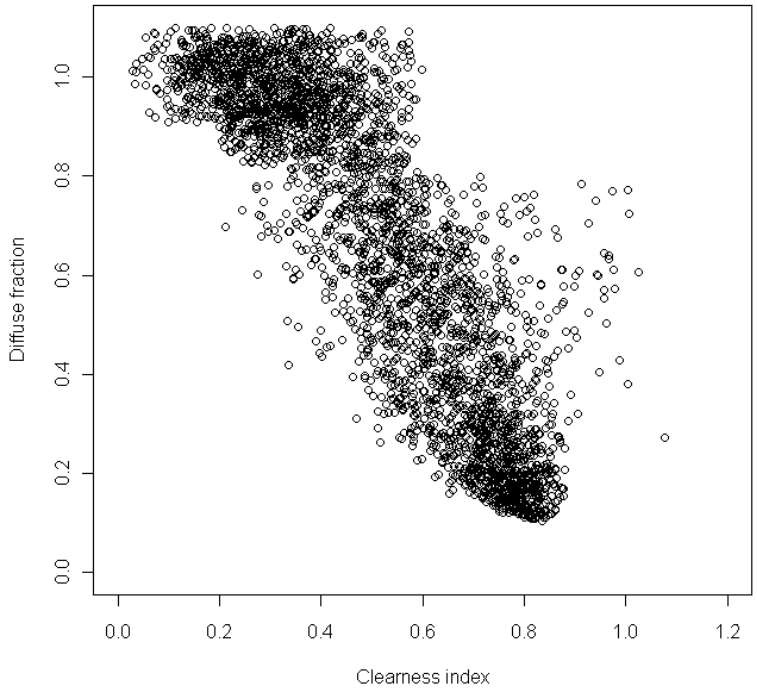


Figure 4.2: diffuse fraction in function of clearness index for Chasseral (NE), for the period 1^{st} February 2009 to 31^{st} January 2010, for a zenith angle ranging from 0° to 85°.

and Concordia Station (Antarctica, iceshelf and glacial area, respectively), Summit Station (Arctic, glacial area) and Alice Springs (Australia, desert). Clear days were manually selected, looking to the daily evolution of the direct radiation: 36 days in Geneva, 23 days in Georg von Neumayer, 20 days in Concordia Station (Dome C), 7 days in Summit Station and 110 days in Alice Springs. The low number of clear days in Geneva is due to the often unstable cloud conditions.

Looking at clear days only, multiple reflection might be avoided. In fact, the graphical representation over these days is mostly composed by high clearness indexes with a low diffuse fraction, which is typical for this day type.

In Figure 4.6, the diffuse fraction on clear days for the five stations is illustrated. It can clearly be noticed that Geneva presents a quite large portion of high diffuse fraction even for clear days. This could be due to the urban location of the station, which implies a higher content of aerosol in the atmosphere (due to the pollution of the city). Alice Springs has a lower diffuse fraction for the same clearness index when compared to Geneva. It could be attributed to the dry atmosphere (due to the desert area where the station is located), which corresponds to less water vapor content in the atmosphere. In addition, pollution is not relevant in that area.

A similar behavior was stated by Ridley et al. [2010] for stations from both Hemispheres. The tropospheric part enhances the backscatter of solar radiation. The anthropogenic in-

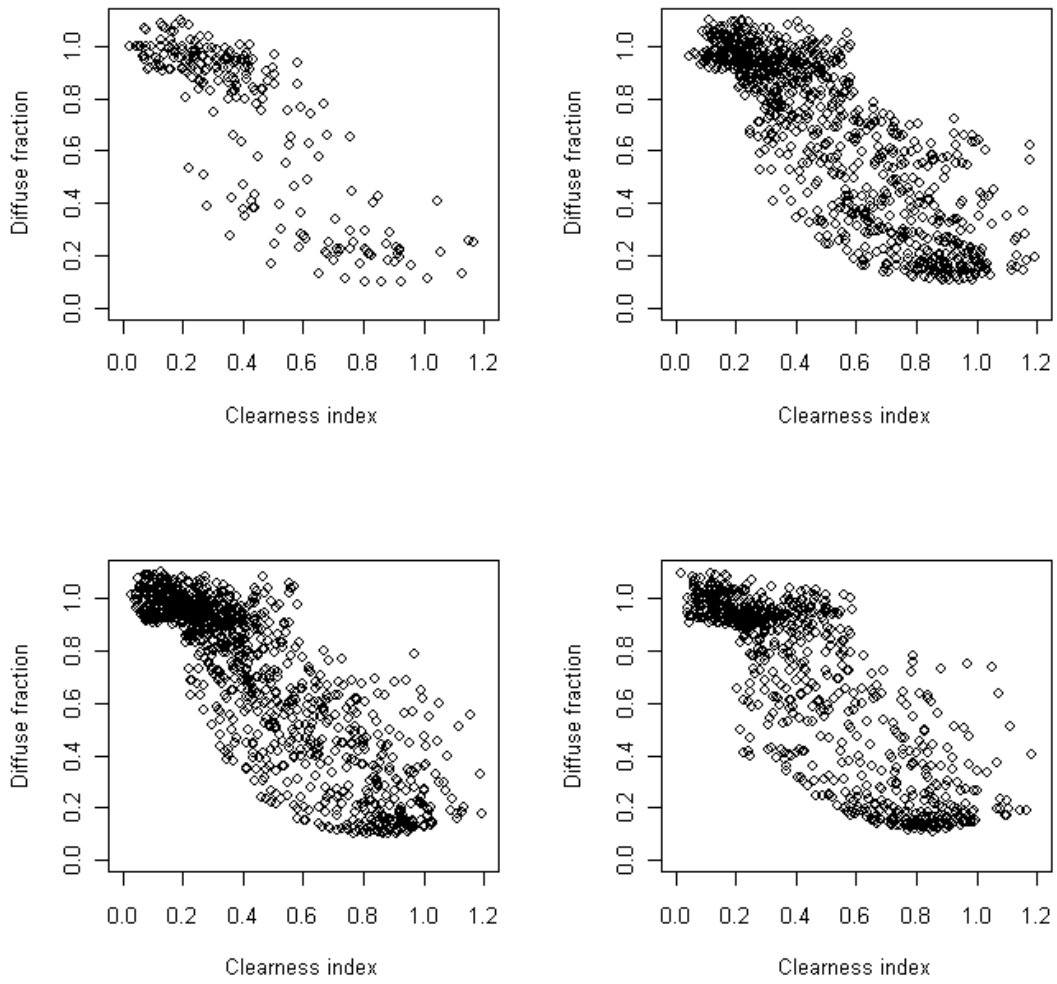


Figure 4.3: Diffuse fraction against clearness index of Pilatus (LU), for the different seasons: winter (top-left), spring (top-right), summer (bottom-left) and autumn (bottom-right).

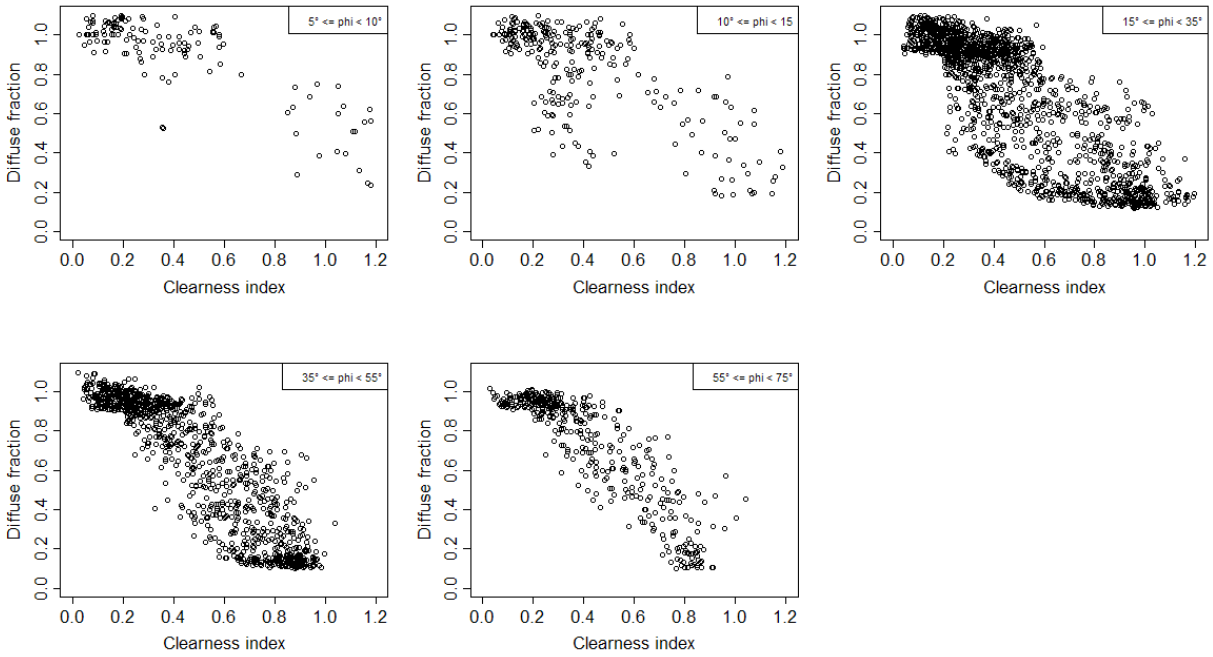


Figure 4.4: Diffuse fraction against clearness index of Pilatus (LU),for different solar elevations.

put goes into the lower part of the atmosphere: it has an enormous dispersion potential (because of the wind) and can reach an elevation of 3 100 m. There is a declination of the stratospheric sulfate. However, there is still an increase of the tropospheric sulfate, which is almost confined in the Northern Hemisphere.

The two antarctic stations, Concordia Station and Georg von Neumayer, have a similar radiation pattern compared to Summit Station. First, their maximal clearness index is higher than the two stations previously compared. This could be associated to the extreme clean atmosphere, which permits a greater amount of radiation to reach the surface. However, this radiation is more scattered than in Geneva and Alice Springs. When the solar elevation is always low, the solar radiation has a longer path through the atmosphere, which implies a higher diffuse radiation production. In addition, the elevated content of ice crystals in the atmosphere produces further multiple reflections between the snowy and icy surface and the atmosphere. For this reason, the two Antarctic and the Arctic stations have a higher diffuse fraction compared to the others.

4.2 Models performance evaluation

Six models representing four classes were presented in Chapter 3. They were tested on data of 39 stations, which are listed in Tables A.1 and A.2. A statistical analysis was conducted

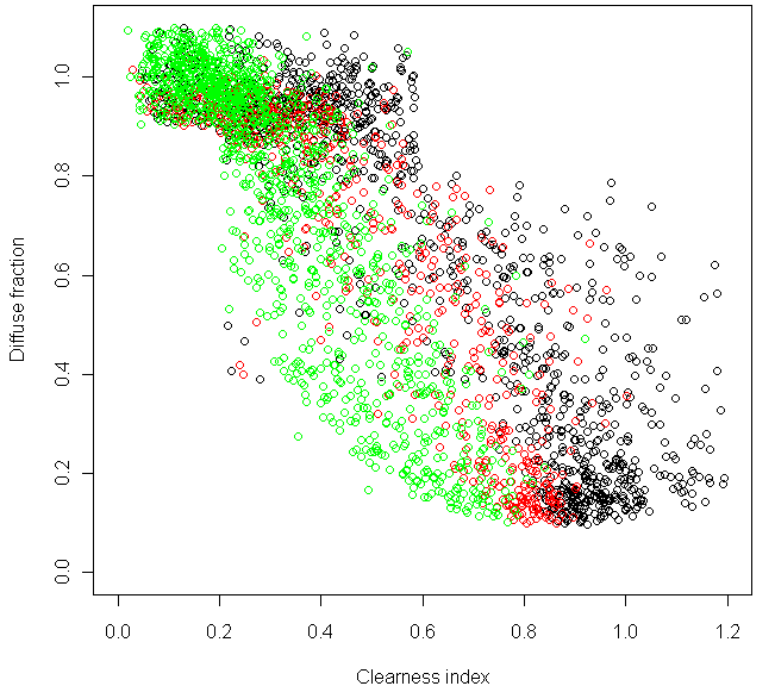


Figure 4.5: Diffuse fraction against clearness index of Pilatus (LU), for the different daytimes: morning (black), midday (red) and afternoon (green).

in order to evaluate the performance of each single model. Four statistical instruments were used to quantify the performance of the models: the coefficient of determination R^2, the coefficient of efficiency E, the Mean Absolute Error and the Mean Bias Error. They are described in Appendix D.

Models of Reindl-Helbig, Skartveit and Olseth, BRL and Maxwell presented some similar pattern and are illustrated in Appendix G. As Table 4.1 shows, the model of Skartveit and Olseth overall performs better than the other models. The model proposed by Maxwell seems to have more limitations than the others. Furthermore, the two Antarctic and the Arctic stations seems to have an enormous influence on the limited performance of the models. In fact, Table 4.1 also shows the results without accounting those three stations and the performance appears definitively improved. P_d is the number of data N_d whose estimated diffuse fraction $\hat{k}_d$ is within ± 10 % of the observed value k_d:

$$P_d = \frac{N_d[\left|\hat{k}_d - k_d\right| \leq 0.1]}{N_d} \cdot 100\ \% \tag{4.3}$$

The effect of different geographical conditions was investigated. No relevant dependency on station altitude is visible. No particular dependency on latitude and longitude is present (Figures G.1 - G.6). As already mentioned by Miguel et al. [2001], there is a seasonal

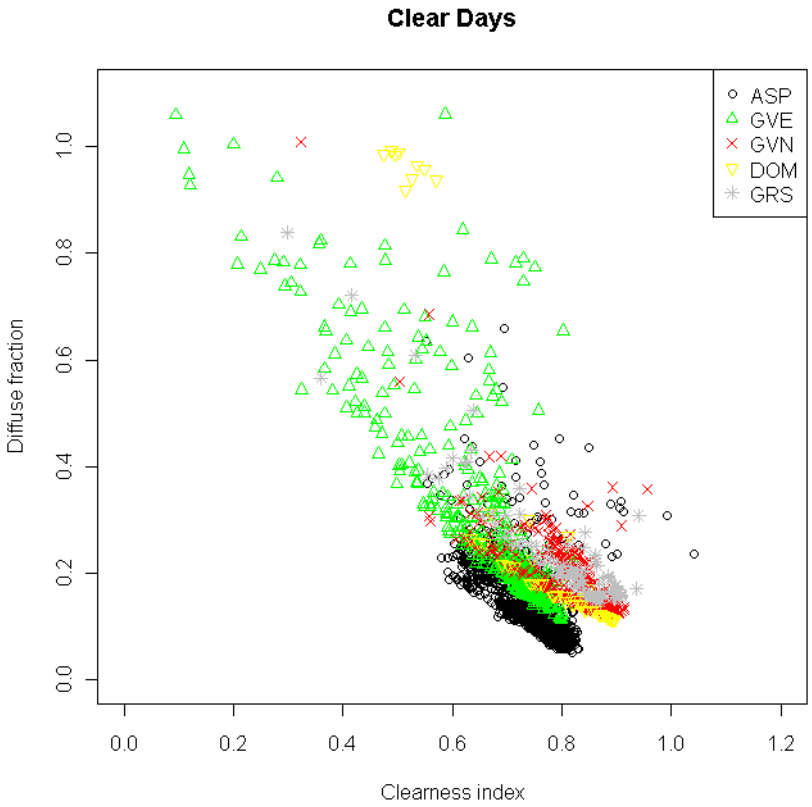


Figure 4.6: Representation of the data for clear sky days in Alice Springs (Australia), Geneva (Switzerland), Georg von Neumayer (Antarctica) and Concordia Station (Antarctica)

variation of the error. For this type of analysis, data have been divided into two groups depending on the hemisphere where the meteo station is located. Data have also divided in four groups corresponding to the four seasons: winter (December, January and February), spring (March, April and May), summer (June, July and August), autumn (September, October and November). For the Southern Hemisphere, seasons were inverted. Spring presents an increased error, with respect to the winter period (see Figures G.7 and G.8) . It can be attributed to a clearer atmosphere in winter months. Furthermore, the four models present a limitation in the estimation of the data for a solar elevation smaller than 30 ° (see Figures G.12 and G.13).

Previous studies underlined the necessity of a model able to especially give a better estimation for intermediate [1] values of clearness index. It is the interval which is more affected by atmospheric and solar effects, as solar elevation and weather variability. However, the models present a more pronounced limitation on the estimation of data for extremely clearness index values.

[1]Notation for the different clearness index intervals is used: *low* ($k_t \leq 0.30$), *intermediate* ($0.30 < k_t \leq 0.80$), *high* ($0.80 < k_t \leq 1.00$) and *extreme high* ($k_t > 1.00$).

Model		MBE	MAE	R^2	E	P_d
Reindl*	*all*	0.2078	0.2485	0.6861	-0.6820	7.83
	no pol.stat.	0.0489	0.0983	0.8735	0.8416	35.36
S&O	*all*	0.2014	0.2304	0.5188	-1.2902	10.28
	no pol.stat.	0.0386	0.0987	0.8773	0.8159	34.70
BRL	*all*	0.2910	0.2239	0.6576	-1.1790	1.23
	no pol.stat.	0.0431	0.0991	0.8672	0.8402	35.47
Maxwell	*all*	0.2406	0.2747	0.3764	-2.4930	7.31
	no pol.stat.	0.0541	0.0913	0.8694	0.8446	38.14

Table 4.1: Statistical performance of the models Reindl-Helbig (Reindl*), Skartveit and Olseth (S&O), BRL and Maxwell for all stations (all) and without the polar stations (Dome C, von Neumayer Station, Summit Station).

Here below, a description of the performance of every model, their advantages and limitations is presented.

4.2.1 Reindl - Helbig

Initially, Reindl et al. [1990] developed different regression models.

This piecewise-regression pattern of the model is still clearly visible in the Reindl-Helbig model. The diffuse fraction is not supposed to exceed a maximal value of unity, but this is not the case. The model assumes that, for a low clearness index (overcast weather situation) the incoming radiation is entirely diffuse and then the ratio linearly decreases. For a high clearness index (clear sky situation), the diffuse fraction is constant for clearness indexes included in the interval. However, these two assumptions are too simple and, as shown in Figure 4.7, a relevant number of data is not efficiently estimated by the model. In particular, for high clearness index, the diffuse fraction is clearly not constant and presents many higher values than those estimated.
In the intermediate interval (partly cloudy sky), the diffuse fraction decreases with an increasing clearness index and depends on the solar elevation. This explains the band structure. However, the model is underestimating the measurements. Furthermore, as illustrated in Figure 4.8, the pattern of the model is the same for every location (Toravere and von Neumayer Station, as examples), but measurements do not present this behavior. This demonstrates that additional parameters have to be considered and that the diffuse fraction is not only clearness index and solar elevation dependent. In the representation of the diffuse fraction against the clearness index for clear days shown in Figure 4.6, it was pointed out that the minimum of diffuse radiation has a different correspondent clearness index depending on the location of the station. This is also visible in Figure 4.8, where von Neumayer Station presents a higher minimal diffuse fraction than Toravere and a higher corresponding clearness index. Nevertheless, the Reindl-Helbig model is not able to reproduce this behavior.

Scatter plots permit to have a better understanding of the performance of the model. They are the representation of the estimated values against the observations. In an ideal

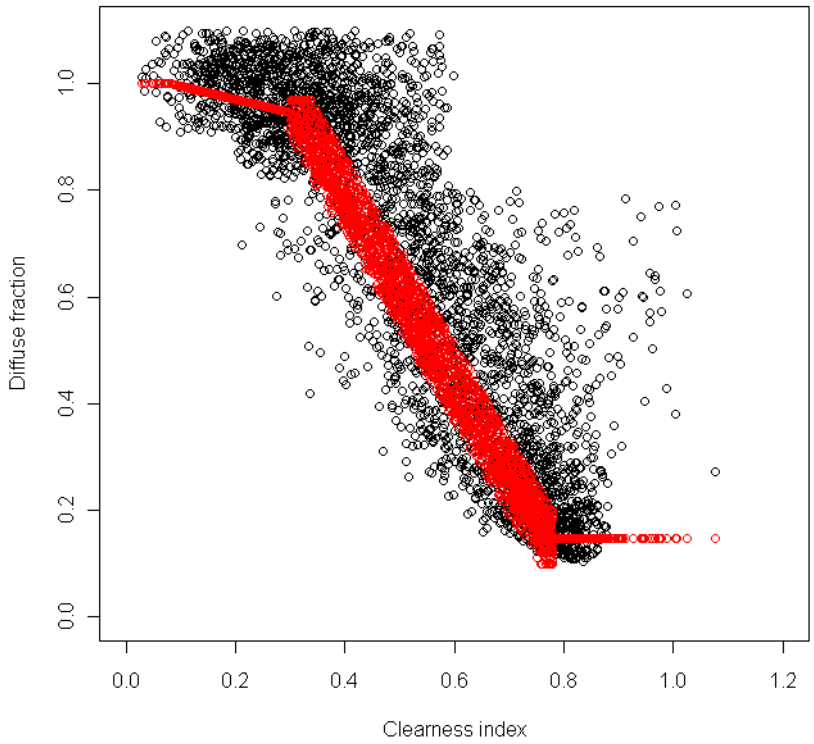


Figure 4.7: Representation of the measured (in black) and estimated from Reindl-Helbig (in red) diffuse fraction against the clearness index for Chasseral (NE).

case, all points would lay on a 1:1 line. A scatter plot was established for the data of every station. As an example, Figure 4.9 illustrates the modeled against the measured diffuse fraction for Altdorf (UR).

The coefficient of determination R^2 corresponds to 0.7530, which can represent the values of the other Swiss stations, which are listed in Table E.1. BSRN stations have in average a R^2 10 % higher than Swiss stations. This could be attributed to the different procedure to calculate the hourly integrated data (see Chapter 2).
In that table, the number of hourly data N_d is reported, and the percentage of well represented data P_d, as well. Generally, less than 50 % of the measured points are well estimated by the model having a maximal absolute error of 10 %. 55.31 % of the measurements at Cabauw are well represented by the model. At the opposite case, Izana has only 14.97 %"¿of the measurements covered. An illustration of both stations is shown in Figure 4.10 and Figure 4.11.

The coefficient of efficiency is giving the goodness of the performance of a model. Statistical results for a selection of stations are listed in Table 4.2. The Dome C station has a negative E (-2.6282), which means that the model is not appropriate to the estimation of the points. The other Antarctic station, Neumayer Station, also has a low value of E (-0.1732). A negative value is obtained for the Arctic station at the Summit (-0.6820). The

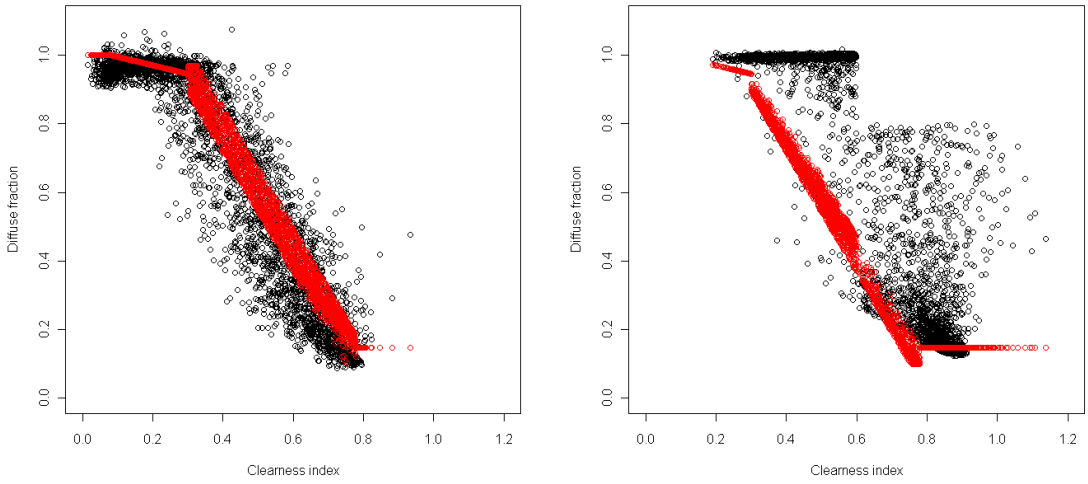


Figure 4.8: Representation of the measured (in black) and estimated from Reindl-Helbig (in red) diffuse fraction against the clearness index for Toravere (left) and von Neumayer Station (right).

number of points covered by the model is poor and the MBE and MAE are more elevated compared to the others.

Additionally, an investigation of the performance of the Reindl-Helbig model was made in function of the clearness index. The data of all stations were separated in three intervals (as chosen for the model), depending on the clearness index values and the results are listed in Table 4.3. The intermediate interval contains the majority of the number of points (about 60%). The choice of the fit, depending on the solar elevation, seems to be appropriate because of an coefficient of efficiency E of 0.644. However, the enormous spread of data results in a poor explanation of the points and elevated values of MBE and MAE. It can be deduced that the choice of the solar elevation is appropriate, but should not be taken as single explanatory parameter.

On the other hand, for low clearness indexes ($k_t \leq 0.30$), the majority of the data is explained by the model within an error of ± 10 %. However, the coefficient of determination R^2 is only 0.0730, which corresponds to a non-correlation between the modeled and observed diffuse fractions. As already observed in the studied graphics above, in the last interval of clearness indexes, for clear sky situations, the number of data explained is extremely poor (14.44 %). Consequently, the choice of a constant diffuse fraction is inappropriate for clear sky situations: other phenomena are contributing to the more elevated values of diffuse fraction.

As support of the above mentioned explanation, Figure F.1 is showing the residuals for the first two clearness index intervals.

The standardized residuals plotted against the standardized modeled diffuse fraction shows that low and intermediate diffuse fraction values are overestimated and high values of dif-

Station	R^2	E	MAE	MBE	P_d	N_d
Altdorf	0.7531478	0.7404422	0.12105326	0.04928409	42.129246	2414
Chasseral	0.7598021	0.6600968	0.13118180	0.09341299	32.940830	3397
Cabauw	0.8850711	0.8688919	0.07143916	0.04867081	55.313634	3858
Dome C	0.5653018	-2.6281581	0.11898770	0.09185393	21.090132	2807
Neumayer Station	0.7485279	-0.1731673	0.23317960	0.20953988	15.601266	3160
Toravere	0.9047302	0.8776950	0.07155062	0.02471932	56.541678	3707
Izana	0.8314110	0.7230235	0.08531861	0.04424831	14.968999	3387

Table 4.2: Table listing the statistical results for the Reindl-Helbig model for a station selection: the coefficient of determination (R^2), the coefficient of efficiency (E), the mean absolute error (MAE), the mean bias error (MBE), the number of hourly integrated data (N_d) and the percentage of relevant estimated data (P_d).

k_t interval	R^2	E	MAE	MBE	N_d	P_d
$k_t \leq 0.3$	0.072966564	0.0695353	0.04854994	0.01903916	27710	90.71093
$0.3 < k_t \leq 0.78$	0.658582212	0.6445940	0.13125325	0.05973583	88152	22.94446
$k_t > 0.78$	0.001373691	-0.2092628	0.09751887	0.01598817	20219	14.43692

Table 4.3: Statistical results of the performance of the Reindl-Helbig model for all stations, for the three different clearness index-intervals.

fuse fraction are mainly underestimated. Furthermore, the histograms of the standardized residuals show that they are not normally distributed for a low clearness index. A better normal distribution is present in the intermediate interval.
For high clearness index, the modeled values are underestimating the observed diffuse fraction and the residuals are not normally distributed. However, a standardized residual representation was not possible because of the null standard deviation value obtained because of the constant modeled value of diffuse fraction.

A summary of the statistical parameters for the Reindl-Helbig model is reported in Table E.1.

4.2.2 Skartveit and Olseth

This model has also a piece-wise regression pattern. However, contrary to the Reindl-Helbig model, a totally new variable is introduced: the variability index σ. This new parameter considers the former and the latter time-period, in an attempt to detect the presence of changing cloudiness.

A graphical representation of the estimated data overlaid on the measured data for Chasseral (NE) is shown in Figure 4.12. A relative good fit is visible, especially for low diffuse fraction values. For low clearness indexes, the global radiation is assumed to be completely diffuse. For an increasing clearness index, the diffuse fraction is decreasing, but it has a wider interval of possible values.

However, the model is generally underestimating the diffuse fraction and seems to be

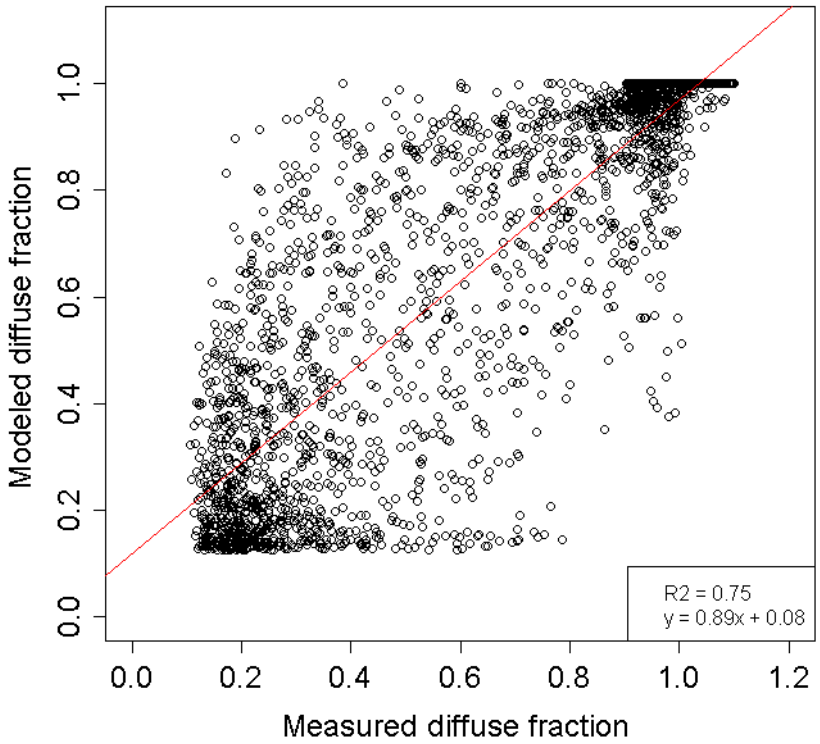


Figure 4.9: Scatter plot of the estimated values from Reindl-Helbig against the measured values for Altdorf (UR). The red line represents the best fit. Its equation is shown with R^2, as well.

unable to account for multiple reflections, where the clearness index has intermediate values and a high corresponding diffuse fraction.
Only by looking the graphics, the Skartveit and Olseth model has a great capacity to reproduce the points for high clearness index, i.e. $k_t \geq 0.8$. However, it still presents some limitations in the estimation of the minimal k_d for every station. As an example, Figure 4.13 shows the data for Toravere. On the other hand, the data of the Georg von Neumayer station, as illustrated in Figure 4.13, underestimated both the diffuse fraction minimum and the corresponding clearness index.
The corresponding scatter plot of the modeled against the observed values of the diffuse fraction for the data of Chasseral, illustrated in Figure 4.14, shows a good correlation between the two variables. In fact, the R^2 underlies the presence of a linear dependency. It is worthy to note that the low diffuse fraction is visibly well estimated.

A summary of all statistic results are listed in Table E.2.
Furthermore, BSRN stations have again a higher R^2 than Swiss stations. This could be due to the more elevated number of points, because of the smaller interval of measure. However, the number of data covered (with a 10 % of error) are oscillating between 4.06 % (Dome

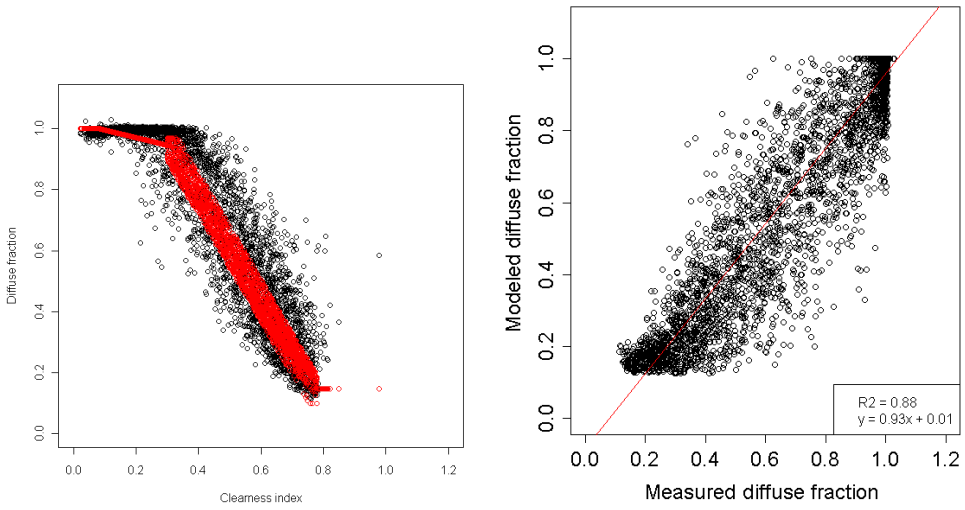


Figure 4.10: For Cabauw, measured (in black) and estimated by Reindl-Helbig (in red) diffuse fraction plotted against the clearness index, on the left. On the right, scatter plot of the modeled against the measured diffuse fraction, with the best fit (red line). The equation of the best fit and R^2 are also shown.

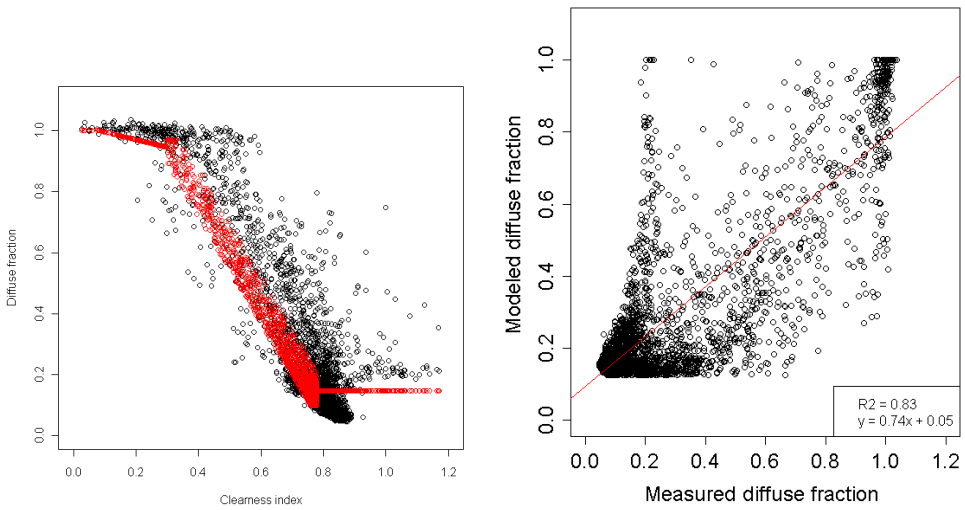


Figure 4.11: For Izana, measured (in black) and estimated by Reindl-Helbig (in red) diffuse fraction plotted against the clearness index, on the left. On the right, scatter plot of the modeled against the measured diffuse fraction, with the best fit (red line). The equation of the best fit and R^2 are also shown.

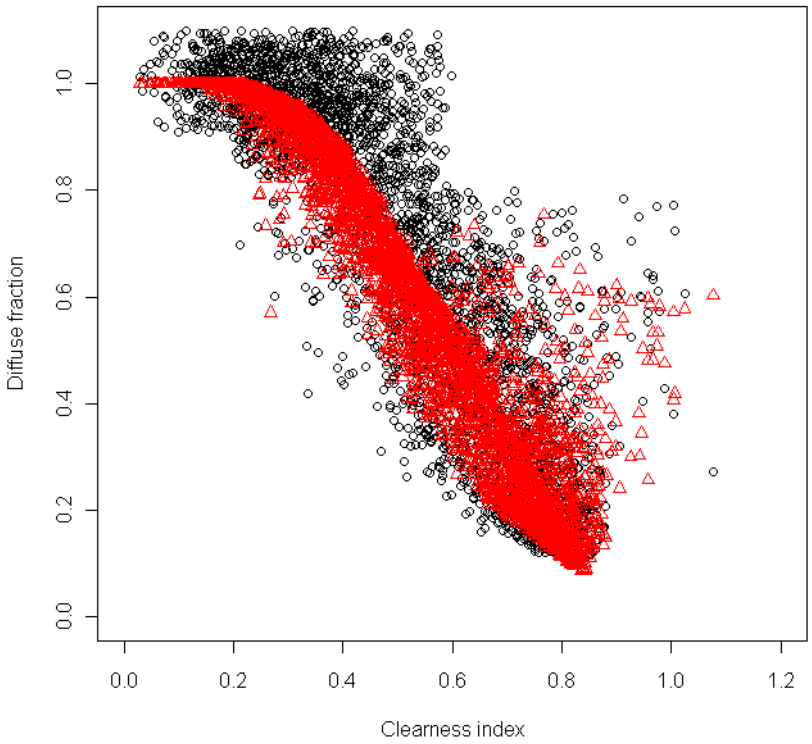


Figure 4.12: Representation of the observed (black) and estimated by Skartveit and Olseth (red) diffuse fraction against the clearness index, for Chasseral (NE).

C) and 16.92 % (Tamanrasset) to 57.24 % (Torovere).
In general, the coefficient of determination R^2 has elevated values, except for Dome C. This reflects the negative coefficient of efficiency, which corresponds to -2.4589. The second Antarctic station, Georg von Neumayer Station, has a negative E, as well.
The low performance of the model on the Antarctic data could be correlated to the low solar elevation present at the two stations.
For this purpose, the performance of the Skartveit and Olseth model was studied for different solar elevations. The results summary is shown in Table F.1.
In Figure 4.15, the R^2 and E are illustrated in function of the solar elevation. It can be observed that both parameters are increasing for an increasing solar angle, which results in an enhancement of the performance of the model. This supposition is confirmed by looking to the MAE and the MBE illustrated in Figure 4.16, where they both decrease for an increasing solar elevation. The Skarveit and Holseth model results in being a good choice of model, especially for solar elevations greater than 30 °. Figure G.12 shows that the other models have the same behavior.
This is nevertheless contradictory to the number of points covered (see Table F.1), which is decreasing from 35.74 % ($\phi \leq 10\,°$) to 31.23 % ($85\,° < \phi \leq 90\,°$). The scatter plot for these two extremes are illustrated in Figure 4.17 and show that for a high solar elevation, there is

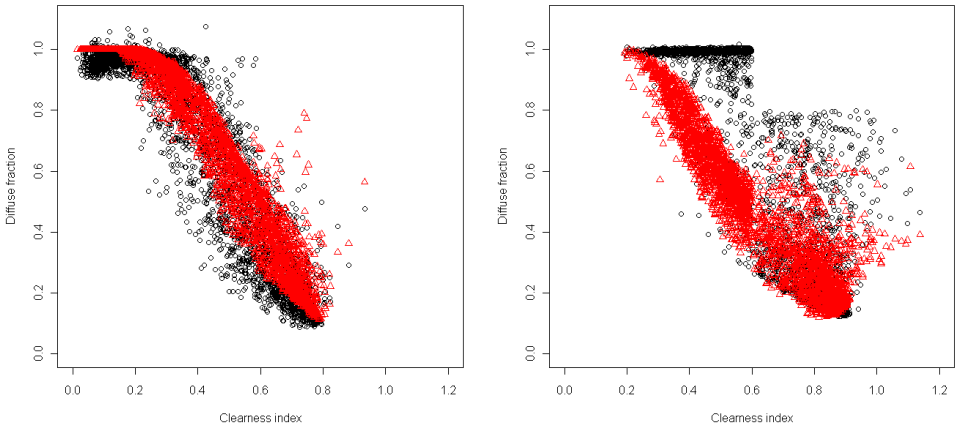


Figure 4.13: For Toravere, measured (black) and estimated by Skartveit and Olseth (red) diffuse fraction plotted against the clearness index, on the left. On the right, for Georg von Neumeyer station.

k_t interval	R^2	E	MAE	MBE	N_d	P_d
$k_t \leq 0.8$	0.763717911	0.7938526	0.1027369	0.04782207	120908	37.863500
$0.8 < k_t \leq 1.0$	0.002208062	-0.7507364	0.1266305	0.01078300	14481	8.638906
$k_t > 1.0$	0.017805841	-3.2627993	0.2731075	0.02286230	692	0.000000

Table 4.4: Statistical results of the performance of the BRL model for all stations, for three different clearness index-intervals.

a low amount of data which is well modeled. In contradiction, for a low solar elevation, the number of data is high and homogeneously distributed over all the graphic, which results in a large number of good modeled data, but also a relevant number of poorly estimated data, which produces a low value of R^2.

4.2.3 BRL

The model developed by Boland, Ridley and Lauret is based on a logistic function, which is characterizing the shape of the estimated data. As an example, Figure 4.18 is illustrating the modeled diffuse fraction, for Chasseral (NE), against the clearness index. The estimated diffuse fraction is ranging from 0 to 1, which corresponds to a good agreement in the fitting for low and intermediate clearness indexes. As the other models, BRL supposes that for overcast situations (low clearness index), the radiation is completely diffuse. However, for a high clearness index, the model assumes that the diffuse radiation is completely disappearing and the incoming radiation is entirely direct. This is, however, not reflecting the observations, which show that after having reached a minimum point the diffuse radiation is increasing again. To reinforce this supposition, Table 4.4 is showing the statistics for

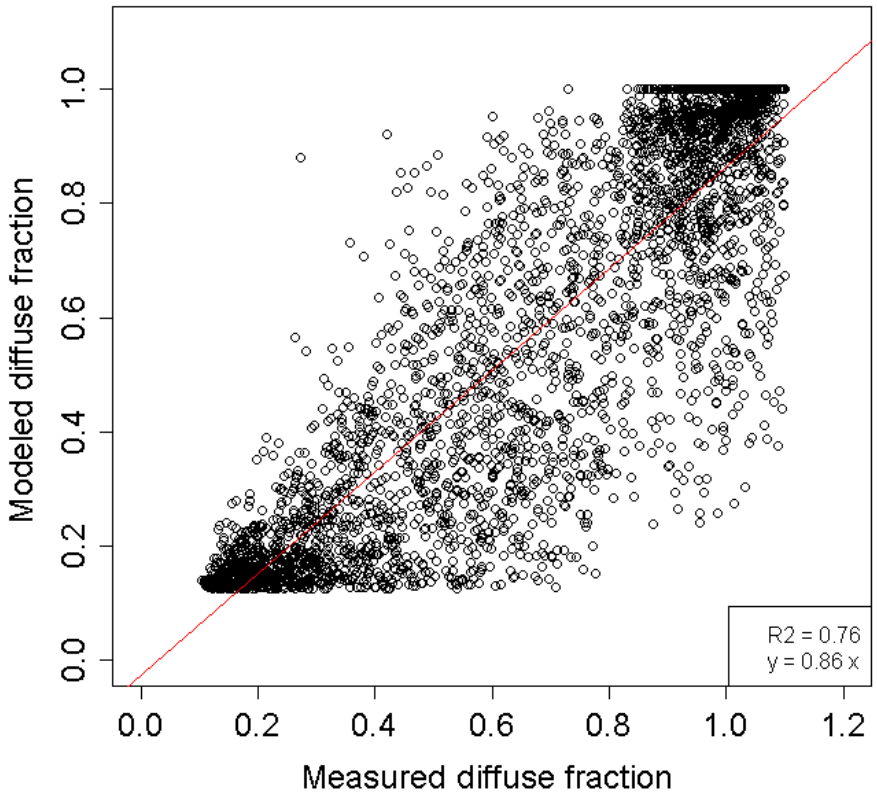


Figure 4.14: Scatter plot of the estimated diffuse fraction from Skartveit and Olseth vs. the measured diffuse fraction for Chasseral (NE). The red line represents the best fit. Its equation is shown with R^2, as well.

low and intermediate clearness indexes, for a high clearness index and for a extremely high clearness index. For a clearness index belonging to the first interval, the performance of the model is good. The coefficient of efficiency is 0.7938. Whereas, it has a negative value for the second and third interval, meaning that the model is not adequate for points belonging to those intervals. Furthermore, for $k_t \leq 0.8$, 37.86 % of the points are covered by the model with an error of ± 10 %, in contrast to the 8.64 % for the other clearness index. Studies of Boland and Ridley [2007], Boland et al. [2008] and Ridley et al. [2010] were not showing this problem and were showing graphics with data belonging only to the first interval (low and intermediate clearness index). It might be possible that their use of hourly data canceled data with an extremely high clearness index. In fact, also in this study the number of points belonging to an extremely high clearness index represent 0.51 % of the total amount of data, only.

The residual analysis for the two clearness index intervals is illustrated in Figure F.2.

In the case the error would be random, in the representations of the standardized residuals against the standardized estimated diffuse fractions, the points would randomly

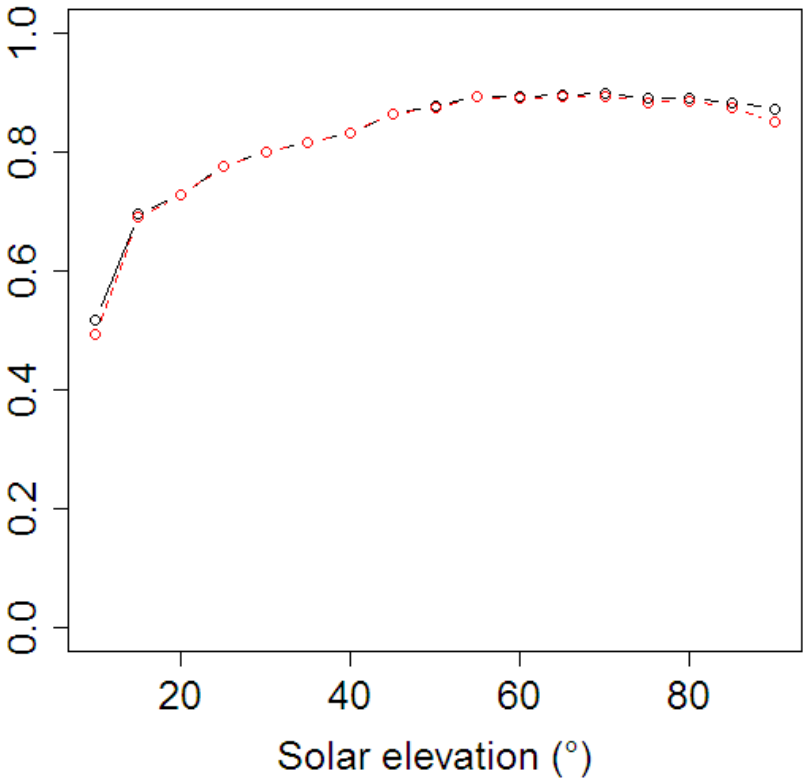


Figure 4.15: Coefficient of determination R^2 (black) and coefficient of efficiency E (red) of Skartveit and Olseth for different solar elevations, for all stations.

be situated next to the horizontal line. However, this is clearly not the case for the data of both clearness index intervals, as shown by the two graphics on the left of Figure F.2. The residuals of both intervals are clearly not normally distributed, especially those for an extremely high clearness index. The model overestimates the diffuse fraction for both intervals, especially for high clearness indexes.
In addition, the correlation between the estimated and the observed diffuse fraction can be visualized by a dispersion plot, as illustrated in Figure 4.19. This is for the station of Chasseral (NE) and is based on 3 397 pairs of points. A linear correlation between the estimated and the observed diffuse fraction is visible and reflected in a high coefficient of determination, which corresponds to 0.7096. Furthermore, the coefficient of efficiency is 0.6904, which reflects a good choice of the model.
This correlation was investigated for all stations and can be seen in Table E.3. The results for a selection of stations is reported in Table 4.6. In general, the model performs better on BSRN stations. However, the worst results are given by Georg von Neumayer Station and Concordia Station, both located in Antarctica. As already mentioned, there might be some problems with the representation of radiation data, because of the local time calculation

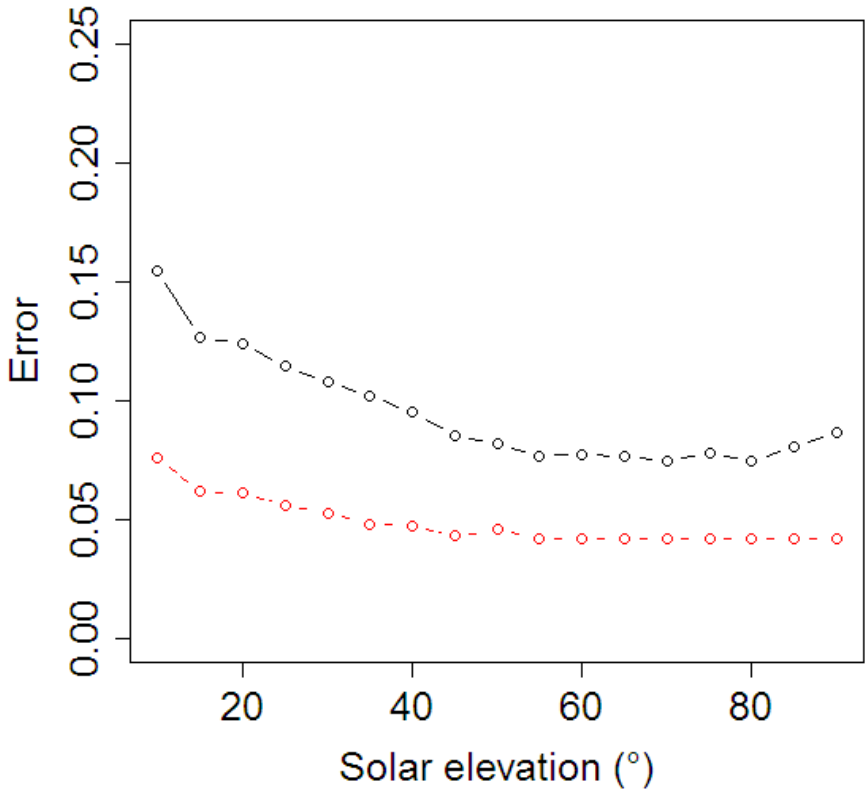


Figure 4.16: Mean Absolute Error (black) and Mean Bias Error (red) of Skartveit and Olseth for different solar elevations, for all stations.

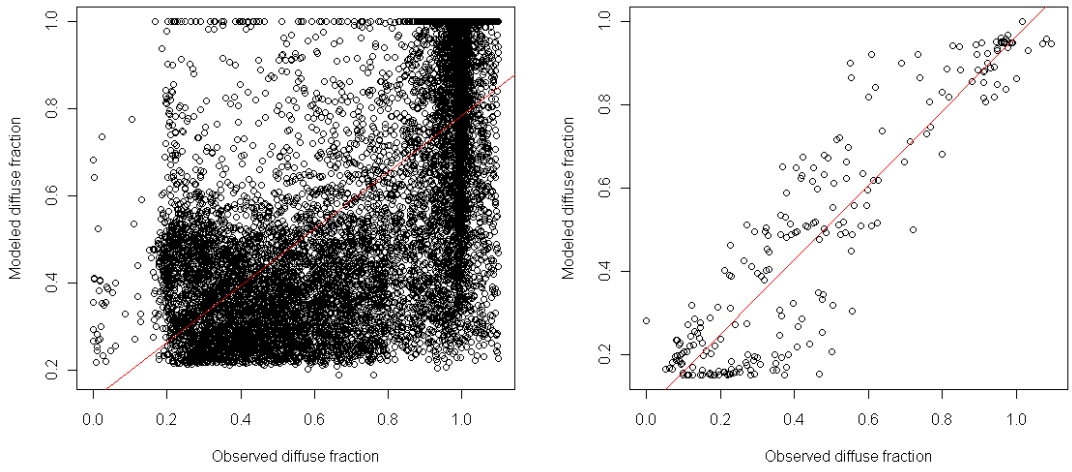


Figure 4.17: Scatter plot of the estimated values from Skartveit and Olseth against the observed diffuse fraction for a solar elevation smaller than 10 ° (left) and between 85 ° and 90 ° (right).

Hemisphere	R^2	E	MAE	N_d	P_d
Northern	0.7555864	0.7705345	0.1060534	103087	35.18969
Southern	0.8000939	0.7657510	0.1064351	32994	29.34776
Northern (no Arctic)	0.7681762	0.7876107	0.1021697	100967	35.90282
Southern (no Antarctica)	0.8849301	0.8658314	0.0830090	27027	34.82813

Table 4.5: Statistical results of the performance of the BRL model for station of Southern and Northern Hemisphere, respectively.

and low solar elevations.

The purpose of the development of this new category of models was based on the need of models applicable to stations of the Southern Hemisphere. Table 4.5 lists the statistics for data divided into the two hemispheres. Ridley et al. [2010] stated that the newly developed model was better performing in the Southern Hemisphere than in the Northern. However, the model tested in this study is covering the 29.35 % of the Southern Hemisphere data, against the 35.19 % of the Northern Hemisphere with a maximal error of ±10 %. Considering that the two Antarctic and the one Arctic stations are giving poor performances of the model and eliminating the data of these two stations, the amount of covered points increases to 34.83 % for the Southern Hemisphere, but still remains inferior to the performance on Northern Hemisphere stations (35.90 %). On the other hand the model performes qualitatively better on the Southern Hemisphere, for which stations the coefficients of dermination and efficiency are more elevated than those for the other Hemisphere.

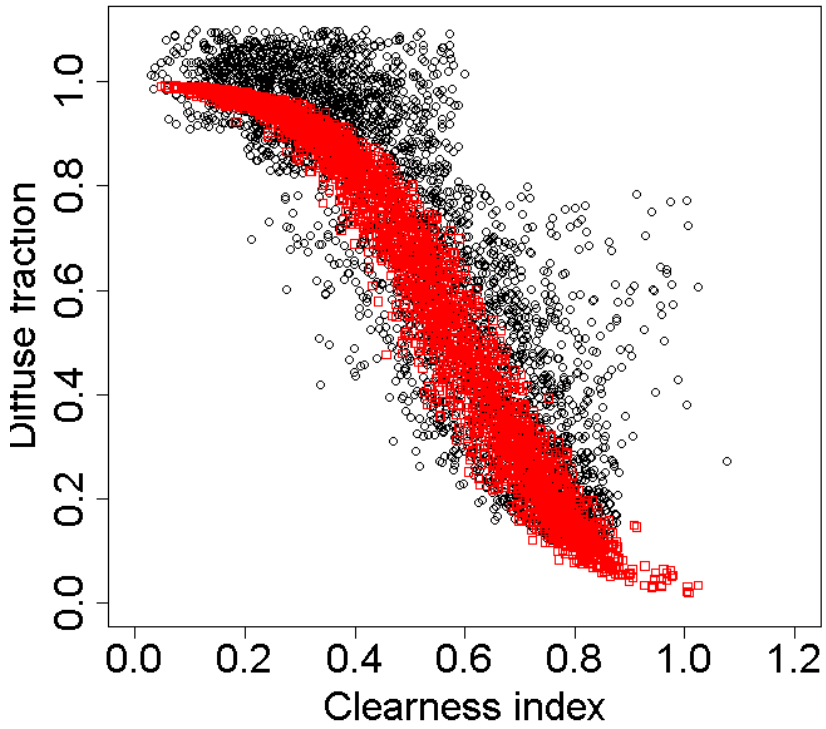


Figure 4.18: Representation of the observed diffuse radiation in function of the clearness index, for Chasseral (NE). The estimated diffuse radiation from BRL is overlaid.

The analysis of the residuals confirms the above observations. Figure F.3 shows a non-normality of the errors, especially for the Northern Hemisphere. Furthermore, the errors are not randomly distributed.

4.2.4 Maxwell

Exponential models is a further approach, initially developed in the context of the DISC project by Maxwell [1987]. As an example, Figure 4.20 shows, for the station Chasseral (NE), the estimated diffuse fraction against the clearness index. An upper limit (set to 1) of the diffuse fraction is set: it is assumed that for low clearness indexes the diffuse radiation can not exceed the global radiation. Then, the diffuse fraction is decreasing for an increasing clearness index. It reaches a minimum, which partly corresponds to the observation. For extremely clear sky situations (high clearness index), the diffuse fraction is increasing again, similarly to the observations. However, this model, as the others, is limited in the modeling of high diffuse fraction for intermediate clearness indexes. Moreover, Chasseral does not present a relevant amount of data with a clearness index exceeding the unity and the model seems to be a good fit. The scatter plot of the estimated against the observed diffuse fraction confirms these observations and is shown in Figure 4.21.

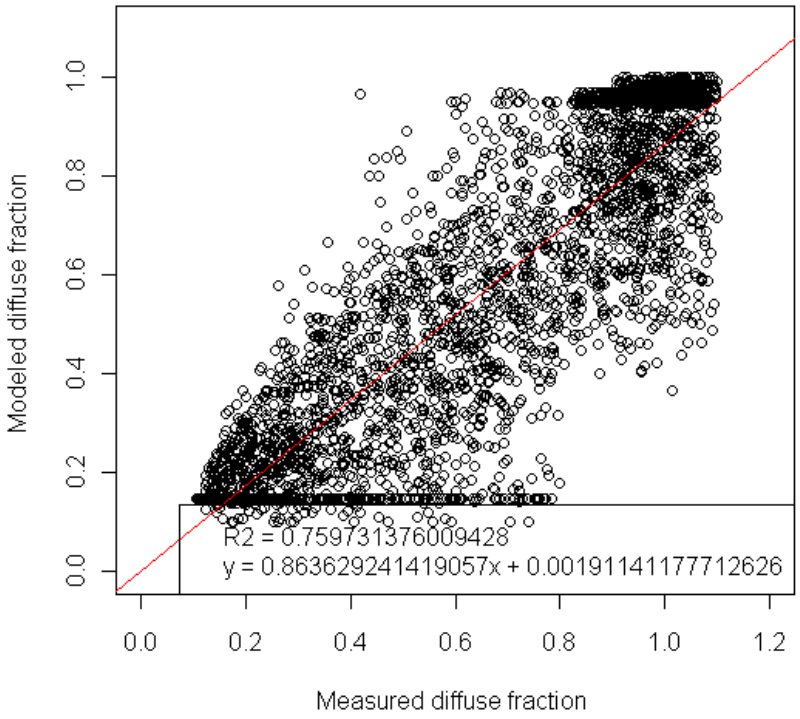


Figure 4.19: Dispersion diagram between the diffuse fraction estimated by the BRL model and the observed, for the station of Chasseral (NE). The line represents the best linear fit. The equation of this fit and the coefficient of determination are illustrated, as well.

There is a correlation between the two variables, which is demonstrated by the coefficient of determination R^2 equal to 0.7495.

However, when stations present a large amount of data with a high clearness index, the performance of that interval is limited. The statistical results for the Maxwell model are represented in Table E.4 and it can be seen that in many stations the performance of the model is poor. A selection of those stations is represented in Table 4.7, where a Swiss station and three BSRN stations are listed.

The Mean Absolute Error of Buchs and Ilorin is double with respect to De Aar and Toravere. Furthermore, the coefficient of determination and the coefficient of efficiency would suggest a limited performance of the model on these data. The representations of the observed and modeled data for Buchs (AG) and Ilorin (Nigeria) are shown in Figure 4.22. It can be seen that Buchs, similarly to the other Swiss stations, presents a large number of data with an extremely high clearness index. Ilorin, instead, presents a large number of data with an high value of diffuse fraction for intermediate clearness index.

For both stations, the minimum of diffuse fraction is well estimated, but, being an exponential function, for extremely high clearness indexes, the estimations strongly increase, overestimating the observations. Such overestimation is also visible in the corresponding

Station	R^2	E	MAE	MBE	P_d	N_d
Alice Springs	0.8810923	0.8566982	0.07197961	0.02994504	21.111678	3922
Comprovasco	0.6707144	0.7620895	0.10455593	0.03317614	34.805195	2695
Cabauw	0.8676571	0.8642308	0.07042890	0.03326044	59.020218	3858
Dome C	0.5235350	-2.7406773	0.18478684	0.10095282	2.137513	2807
Neumayer Station	0.7543687	0.1033020	0.23719585	0.19002296	6.645570	3160
Darwin	0.8754014	0.8358246	0.08755820	0.03309197	33.843910	4113

Table 4.6: Statistical results for the BRL model on a selection of stations.

Station	R^2	E	MBE	MAE	P_d	N_d
Buchs	0.61794373	0.57058849	0.14621858	0.06641948	41.058079	3478
Ilorin	0.49141223	0.06707967	0.12717591	0.06037992	37.055678	3179
Summit Station	0.37641493	-2.49301482	0.27466443	0.24026858	7.456347	2119
De Aar	0.90145686	0.87390448	0.07771343	0.03512430	32.202881	3332
Toravere	0.90078986	0.90007445	0.07082736	0.03344563	57.971405	3707

Table 4.7: Statistical results for the Maxwell model, for some stations.

scatter plots, which show a slight correlation between the modeled and the measured diffuse fraction, except for high diffuse fractions.

The representation of the errors MAE and MBE in function of the clearness index in Figure 4.23 reflects such a limitation. In fact, an error extremely high for extremely high clearness indexes is visible. Moreover, the error at intermediate clearness indexes is relatively high as well, as observed for the data of Ilorin. The error for overcast situations is reflected on the high Mean Bias Error of some stations, as Table 4.7 shows.

In order to have a better overview of the model performance for high clearness indexes, the model was applied to three clearness index intervals: low and intermediate k_t, high k_t and extremely high k_t. The statistical results are shown in Table 4.8. The Coefficient of efficiency E is well representing the poorness of the model performance for high clearness index values. In fact, it decreases reaching negative values for a high clearness index interval, meaning that the model has limitations to represent those data.

The residuals are not normally distributed for either of the three intervals (see Figure F.4).

The model developed by Maxwell has also some limitations for low solar elevations. In fact, Figure 4.24 shows the R^2 and E increasing with the increasing of the solar elevation. The performance of the model increases with the solar elevation, reaching a constant value for

Interval	R^2	E	MBE	MAE	N_d	P_d
$k_t \leq 0.80$	0.76050177	0.73005472	0.1161381	0.04734667	120908	38.315910
$0.80 < k_t \leq 1$	0.13748537	-0.03350034	0.1049007	0.04049715	14481	11.870727
$k_t > 1$	0.01695666	-7.60217890	0.4005152	0.38062133	692	3.468208

Table 4.8: Statistical results for the Maxwell model, for three clearness index intervals.

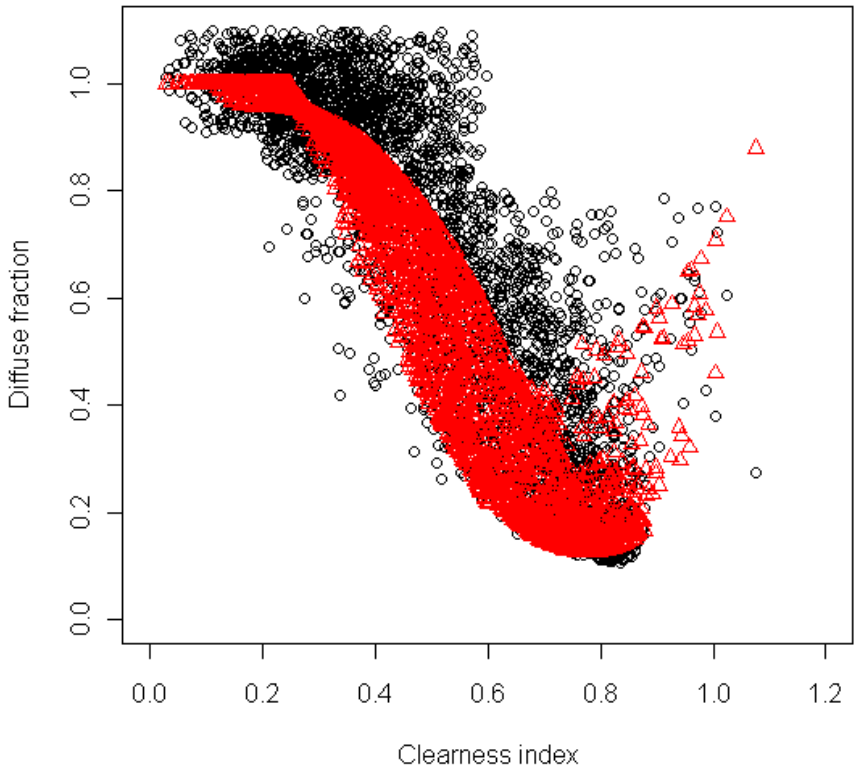


Figure 4.20: Representation of the measured (in black) and estimated from Maxwell (in red) against the clearness index for Chasseral (NE).

a solar elevation greater than 40 °, where the coefficient of efficiency E and the coefficient of determination are about 0.85. For greater solar elevations, the model is not depending on the solar elevation any more.

4.2.5 Perez

A correction matrix is applied to the diffuse fraction estimated by Maxwell, which results are described in Section 4.2.4. It should be an improvement on the previous model. However, the graphics in Figure 4.25 show a model which has limitations for a clearness index smaller than 0.3.
Table 4.9 shows the statistics representing the performance of the model for three stations: Chasseral, Andeer and Payerne. The coefficient of determination E clearly indicates that the model is not adequate.

For this reason, an individual analysis and no comparison with the other models have been made.
In order to have a better understanding on model of Perez, data from Payerne (a BSRN

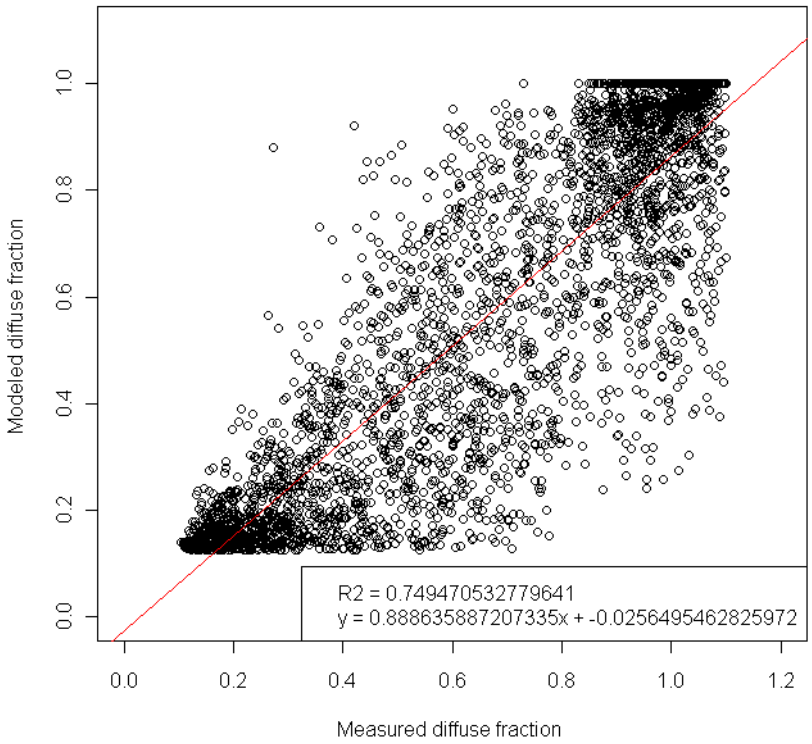


Figure 4.21: Scatter plot of the estimated values from Maxwell against the measured values for Chasseral (NE). The best fit (red line), its equation and the coefficient of determination are shown, as well.

station) are analyzed. In contrast to the other models, Perez seems not to have a solar elevation dependency on this site, in terms of coefficient of determination R^2. However, the MAE is decreasing for an increasing solar elevation. The coefficient of determination R^2 and the MAE are represented in function of the solar altitude in Figure 4.26. Furthermore, the MAE in function of the clearness index (Figure 4.27), after reaching a peak for a clearness equivalent to 0.3, is then decreasing. This reflects the points in Figure 4.25 where $k_t < 0.4$ and $k_d < 0.6$.
Being the model a correction of the Maxwell one, it suggests that the correction matrix used is inappropriate. Further investigations are necessary.

4.2.6 Dürr

This last model is based on a regression tree and the transmissivity q is the principal parameter. The combination of the transmissivity and its behavior in the former 30-minutes interval defines different weather situations, and , as a consequence, attributes a value to the diffuse fraction.
Furthermore, the model requires station specific parameters. Dürr kindly provided those

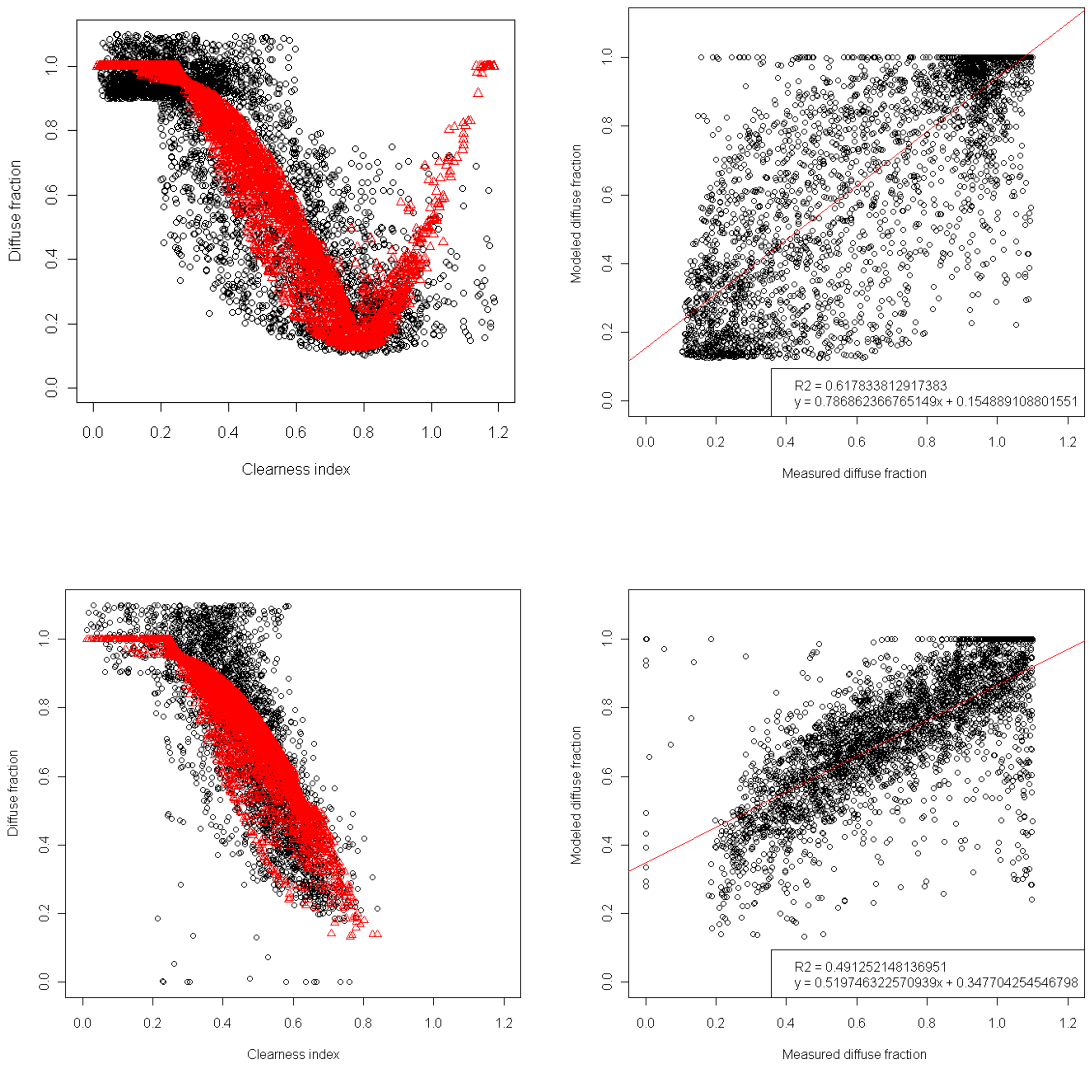


Figure 4.22: For the stations Buchs (top) and Ilorin (bottom), representation of the modeled by Maxwell and observed diffuse fraction against the clearness index (left) and the estimated against the observed diffuse fraction (right). The best fit (red line), its equation and the coefficient of determination are shown, as well.

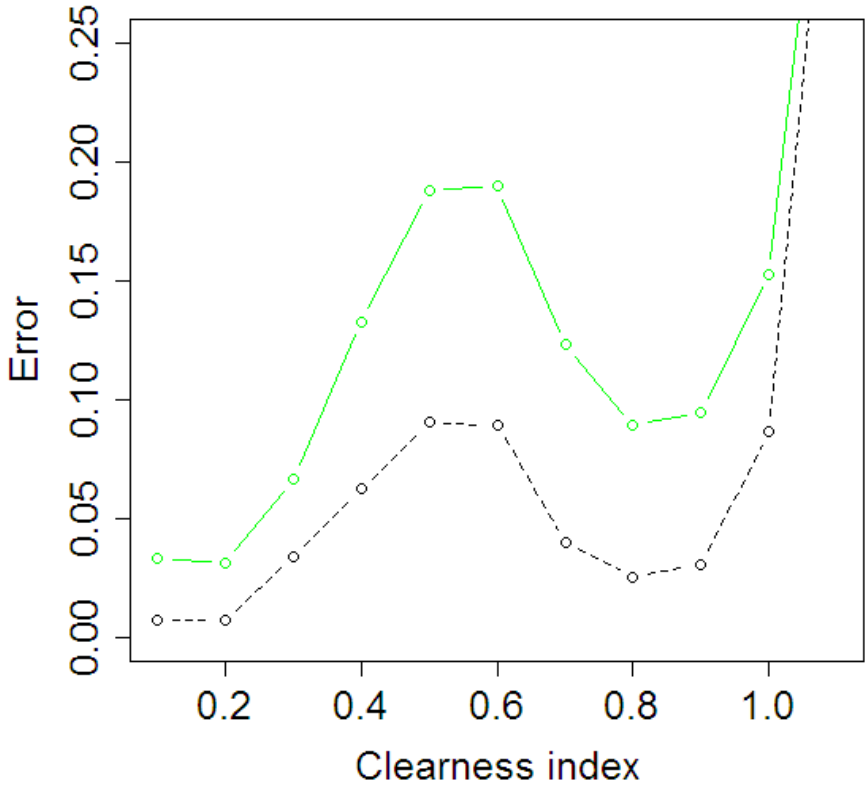


Figure 4.23: Representation of the MBE (in black) and MAE (in green) in function of the clearness index of the Maxwell model.

for Payerne, so that a graphical representation of the data for this station was possible. Figure 4.28 represents the diffuse fraction against clearness index for 1-minute data from Payerne on the 2^{nd} May 2009. There is a good estimation for clear sky situations. On the other hand, for a big variability of the transmissivity, the models is not able to correctly estimate the diffuse fraction. In fact, it principally attributes to the diffuse fraction the value of the previous period, which produces the horizontal lines visible in Figure 4.28. Further investigations have, therefore, to be conducted.

4.3 Time resolution

The above comparison and analysis is based on hourly data. This time step is the most frequently used. However, SNOWPACK and ALPINE 3D can run at shorter ones, 30 minutes as an example.

For this reason, an investigation of the performance of the models Reindl-Helbig, Skartveit and Olseth, BRL and Maxwell was conducted on smaller time intervals. The analysis is

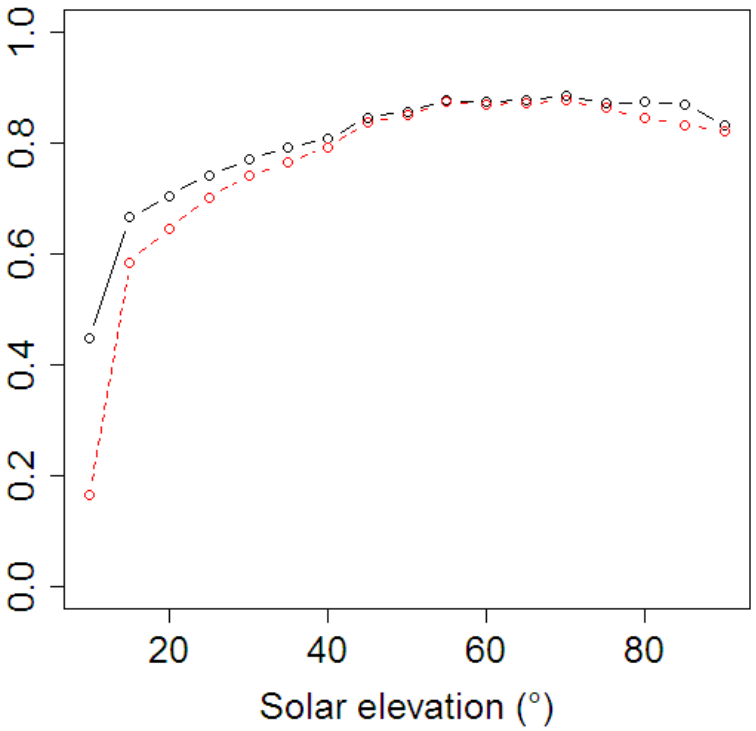


Figure 4.24: Coefficient of determination R^2 (black) and coefficient of efficiency E (red) in function of solar elevation of Maxwell model.

Station	R^2	E	MAE	MBE	P_d	N_d
Chasseral	0.3150178	-0.96590792	0.2832100	0.2716697	15.69031	3397
Andeer	0.3288448	-0.05224847	0.2254009	0.1742493	20.90985	2396
Payerne	0.4260680	-0.31779856	0.2421085	0.2352026	22.90076	3144

Table 4.9: Table listing the statistical results for the Perez model for a station selection: the coefficient of determination (R^2), the coefficient of efficiency (E), the mean absolute error (MAE), the mean bias error (MBE), the number of hourly integrated data (N_d) and the percentage of relevant estimated data (P_d).

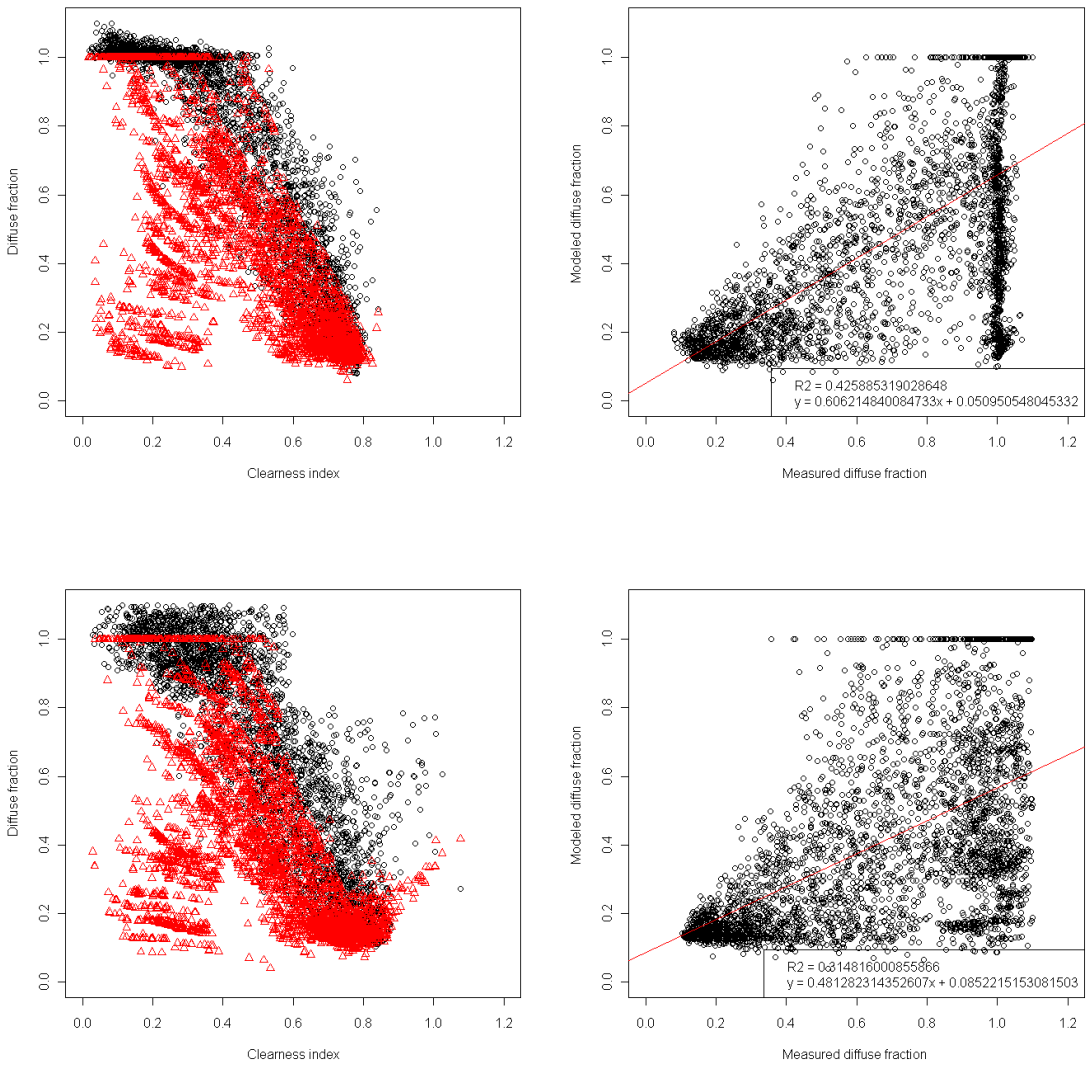


Figure 4.25: For the stations Payerne (top) and Chasseral (bottom), representation of the modeled by Perez (in red) and observed (in black) diffuse fraction against the clearness index (left) and the estimated against the observed diffuse fraction (right). The best fit (red line), its equation and the coefficient of determination are shown, as well.

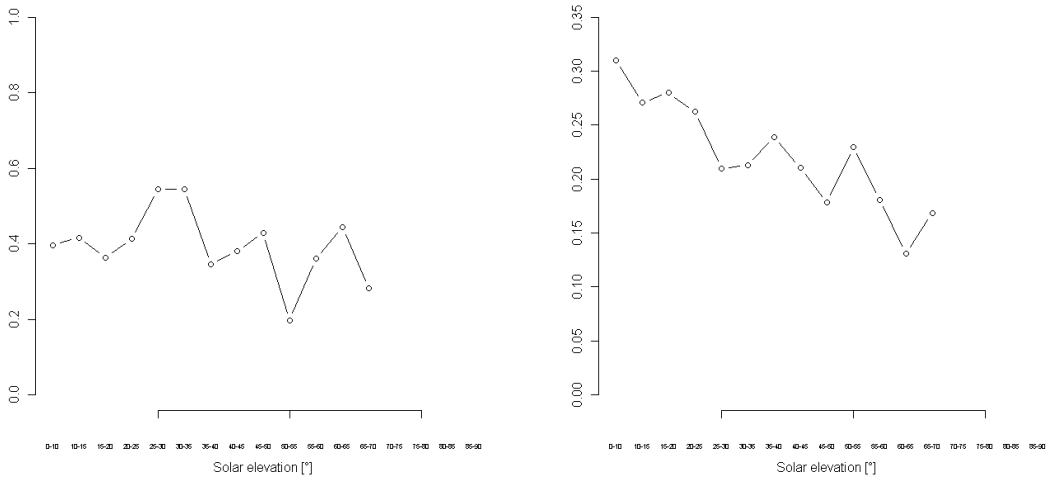


Figure 4.26: For Payerne, R^2 (left) and MAE (right) in function of the solar elevation.

Model		N_d	P_d	MBE	MAE	R^2	E
Reindl*	*1-min.*	185338	49.51	0.0605	0.0989	0.8675	0.8164
	10-min.	23826	46.98	0.0296	0.0982	0.8632	0.8376
	30-min.	7960	47.32	0.0282	0.0945	0.8732	0.8481
	hourly	4000	47.15	0.0271	0.0901	0.8840	0.8609
S&O	*1-min.*	185338	54.72	0.0573	0.0838	0.8663	0.9029
	10 min.	23826	53.52	0.0217	0.0761	0.9129	0.8850
	30 min.	7960	52.67	0.0222	0.0757	0.9160	0.8901
	hourly	4000	53.22	0.0230	0.0737	0.9170	0.8952
BRL	*1-min.*	185338	51.77	0.0501	0.0929	0.8847	0.8485
	10 min.	23826	48.01	0.0336	0.0942	0.8677	0.8590
	30 min.	7960	49.45	0.0333	0.0898	0.8739	0.8681
	hourly	4000	50.37	0.0327	0.0853	0.8817	0.8793
Maxwell	*1-min.*	185338	33.48	0.1029	1.1611	0.6698	0.6543
	10 min.	23826	47.19	0.0221	0.0971	0.8746	0.8162
	30 min.	7960	47.80	0.0209	0.0945	0.8792	0.8213
	hourly	4000	36.62	0.0642	0.1392	0.6859	0.3854

Table 4.10: Statistical performance of the models Reindl-Helbig (Reindl*), Skartveit and Olseth (S&O), BRL and Maxwell for different time intervals.

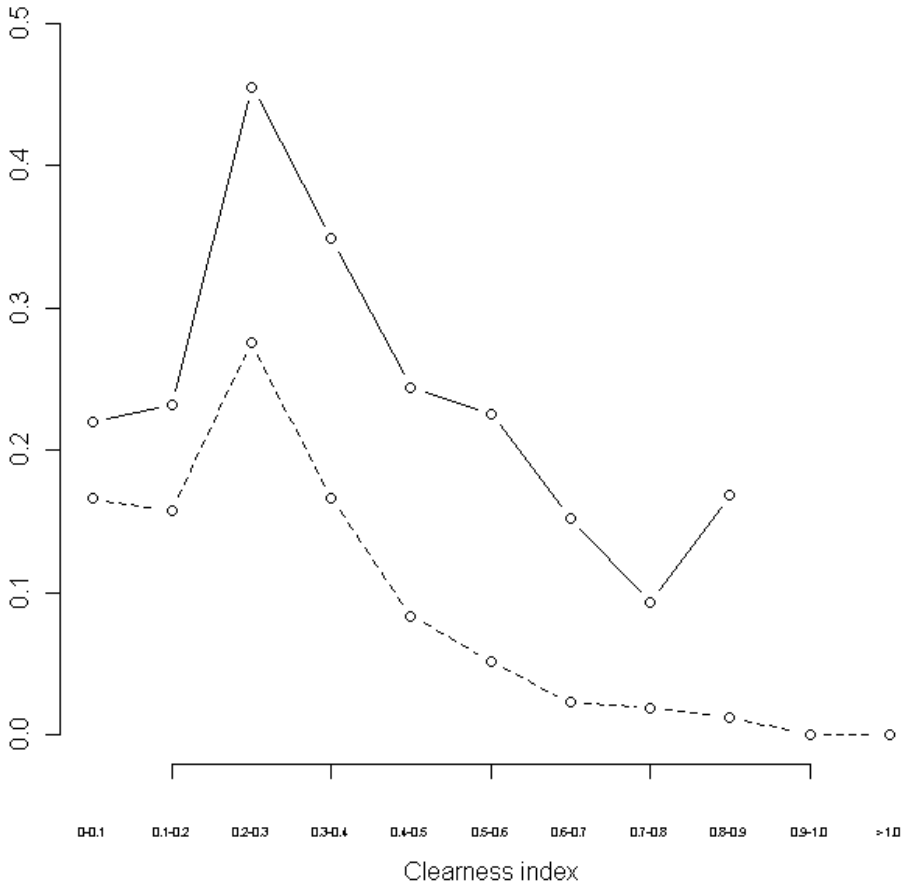


Figure 4.27: For Payerne, MAE (lined) and MBE (dashed) against the clearness index.

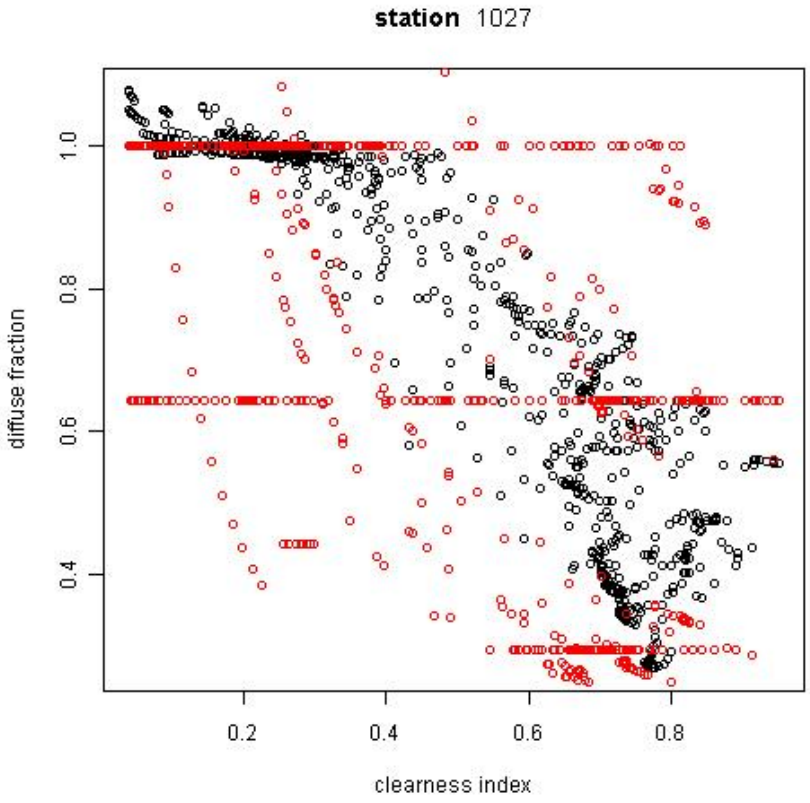


Figure 4.28: Diffuse fraction (in black) against the clearness index for Payerne on the 2^{nd} May 2009. The estimation of Dürr (in red) is shown, as well.

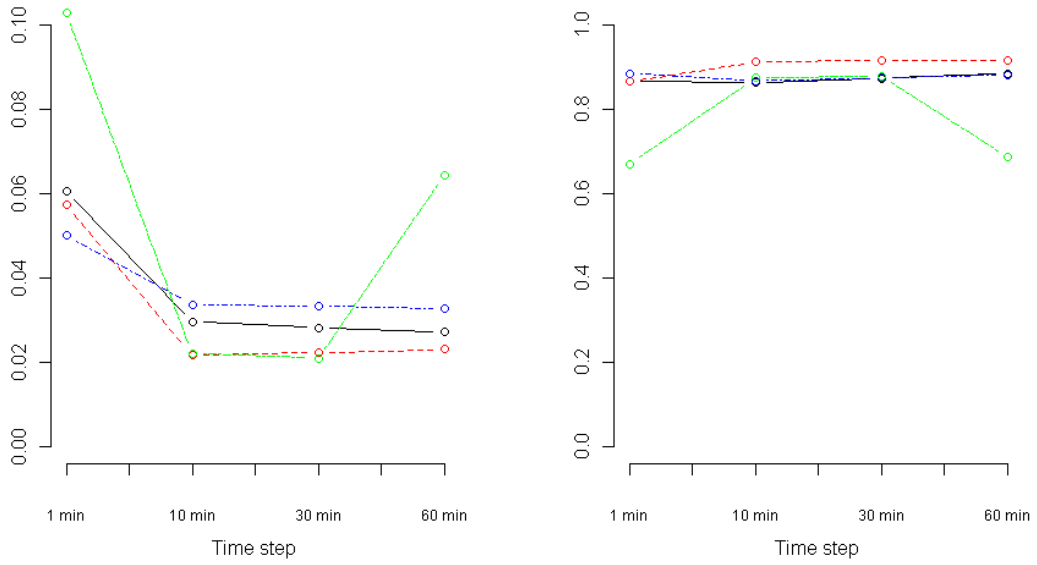


Figure 4.29: For Payerne, MBE (left) and R^2 (right) for different time steps for Reindl-Helbig (in black), Skartveit & Olseth (in red), BRL (in blue) and Maxwell (in green).

done on data of Payerne at 1, 10, 30 and 60 minutes intervals. The statistical results are reported in Table 4.10. It can be seen, that the error tends to increase as the time resolution is smaller (see Figure 4.29) However, these variations are acceptable. The precision of the model can be represented by the coefficient of determination and its evolution is shown in Figure 4.29. However, more pronounced variations are present for Maxwell.

Similar results for Maxwell and Perez were observed by Ineichen on data from Geneva [Ineichen, 2007].

Chapter 5

Discussion

Some previous studies were testing models on data from different locations. Others were verifying different models on a desired location. In this work, there is a need of an universal model, which could be applicable to data from any location. For this reason, the work attend a selection of the most relevant models, which were then applied on data from different locations situated worldwide. With difference to previous studies the model comparison was done on a dataset with several station conditions, as polar and alpine stations and stations located in desert, on the ocean and on ice shelves, as well. The station selection is covering latitudes ranging from $-78.100\,^\circ$ to $78.925\,^\circ$, longitudes from $-237.617\,^\circ$ to $169.689\,^\circ$ and altitudes from 0 m to 3233 m.

5.1 Data representation

International network and weather institutions often offer weather and climatological data, such as incoming shortwave radiation and surface air temperature. The dataset gathered for this study is composed of 10-minute interval data from MeteoSchweiz (Swiss stations except for Payerne), 1-minute interval data from the University of Colorado (Summit station) and BSRN. In order to have an homogeneous dataset, measurements where averaged over one hour, which is the time interval on which the models listed in Chapter 3 are based. An adequate comparison with other studies can then be made. Furthermore, a data quality control was applied before and after the average process. However, there is a clear difference between the representations of Swiss and worldwide hourly integrated data. As an example, Figure 5.1 shows the diffuse fraction against the clearness index for Bermuda and Andeer. It can be observed that Andeer presents a large number of data with a high clearness index. Data of Swiss stations had an additional control: only data with a solar elevation greater than the horizon angle were considered, in order to eliminate data where the global radiation is composed of diffuse radiation only.
Iqbal [1980] stated that there were no data for $k_t > 0.90$. Later, Erbs et al. [1982] stated that data in the interval $k_t \geq 0.80$ are not well understood to justify fitting a curve to them [Helbig, 2009]. Knowing that BSRN and MeteoSchweiz are applying data quality controls and have rigorous instrumental requirements [Lee, 2009], the presence of high and extremely high clearness index data should be attributed to other causes. It could be thought that this is due to the difference in the data averaging. In fact, for BSRN data, 60 1-minute measures are averaged. Whereas, for Swiss data, only 6 10-minutes values build up the hourly mean. As a consequence, the variation of the data is less for a longer time interval. This is further discussed in Section 5.3 In addition, it is unknown whether the values reported by the two institutions are averages or instantaneous values. This fact could affect the final average.

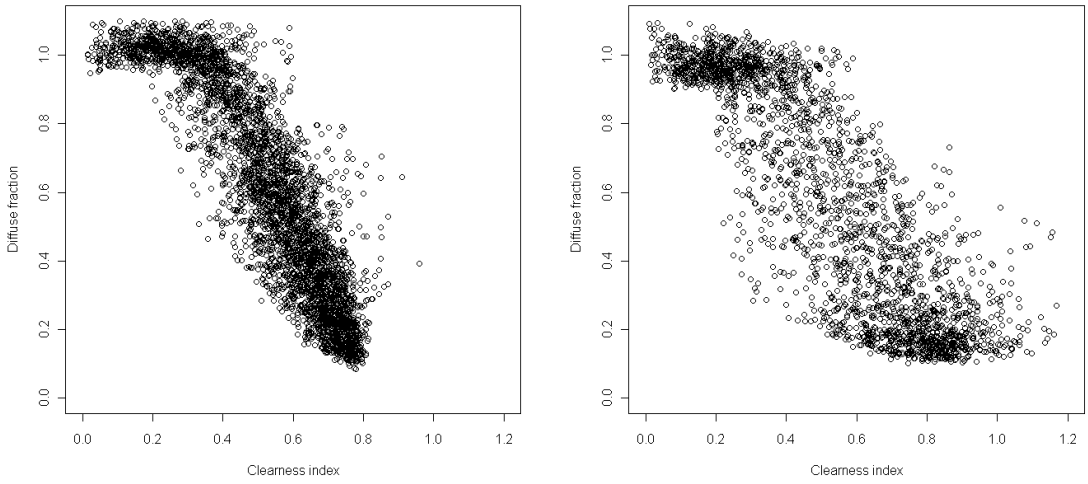


Figure 5.1: Diffuse fraction against clearness index for Bermuda (left) and Andeer (right).

5.2 Model's performance

Four models (Reindl-Helbig, Skartveit & Olseth, BRL and Maxwell) were compared and their performance analyzed in function of different parameters and conditions.
At first sight, looking to Table 4.1, Skartveit and Olseth seem to have developed the best model. The coefficient of determination corresponds to 0.8773 and the Mean Absolute Error to 0.0386. However, the other models obtained a slightly poorer performance, so that a more detailed analysis has to be conducted.

Clearness index

In order to visualize the data, the usual representation of the hourly diffuse fraction against the hourly clearness index was applied. As shown in Figure 5.2, all four models present limitations for intermediate and extremely high clearness index situations, where the MAE and MBE have higher values. The intermediate values, where the MAE reaches a maximum of about 0.3, corresponds to partly cloudy situations. The poor performance of all models for this interval is due to the inability to reproduce broken clouds weather situations, where the global radiation is principally composed of diffuse radiation (see Figure 5.3). Solar radiation is reaching the ground and is then reflected by the surface and clouds. Unfortunately, this behavior (represented schematically in Figure 5.4) is hard to reproduce. In addition, models present a limitation for extremely high clearness index values. For these values, the diffuse fraction reaches a minimal value and tends then to increase again. As stated by Skartveit and Olseth [1998], direct radiation reaches a maximal value. For this reason, in the case of extremely high clearness indexes, the direct radiation has reached its maximal value and the additional radiation detected is completely diffuse. It could be produced by the surroundings, as ice and snow surface, and clouds which are not obscuring

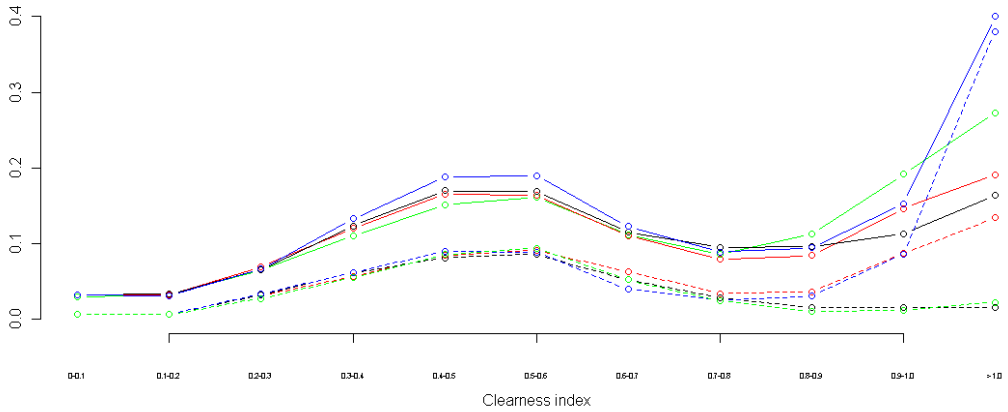


Figure 5.2: MAE (solid line) and MBE (dashed line) against the clearness index, for Reindl-Helbig (black), S&O (red), BRL (blue) and Maxwell (green).

the Sun [Miguel et al., 2001]. This lack in estimation of extremely high clearness index data, is particularly visible in Figure 5.3. A logistic function tends to decrease to zero, but it is not able to reproduce the real behavior of the data. Using a constant value, Reindl* has also an inappropriate choice of fit for this last k_t interval. S&O and Maxwell seem to have a model reproducing this behavior best. However, Maxwell (having an exponential function) tends to overestimate the diffuse fraction for the last clearness index interval, which corresponds to a MAE value of 0.4, which is almost the double compared to the other models (see Figure 5.2).

Solar elevation

Already in the 90's, it was clear (i.e. Reindl et al. [1990], Skartveit and Olseth [1987]), that the clearness index would not be the only significant predictor. Solar elevation has a great influence on the composition of global radiation, as well.
In this study, only data with a solar elevation greater than 5 degrees were considered. Instrumental imprecisions are present at low solar elevations, especially due to the cosine response [Miguel et al., 2001].
As Figure 5.5 shows, models perform poorly for a solar elevation smaller than 30°. It corresponds to the limit set by Skartveit and Olseth [1987], who stated that, for partly cloudy and clear skies, this solar elevation range has an enormous influence on the diffuse fraction.

Furthermore, because of their geographical position, the solar elevation at the two stations located in Antarctica and at the one in Arctic is low all over the year. For instance, the maximal solar angle in Georg von Neumayer Station is 42.79 °. At Dome C it reaches 38.30 ° and 40.85 ° at Summit Station. All models seem to not well estimate the data of those three stations: the coefficient of efficiency has negative values, which corresponds to

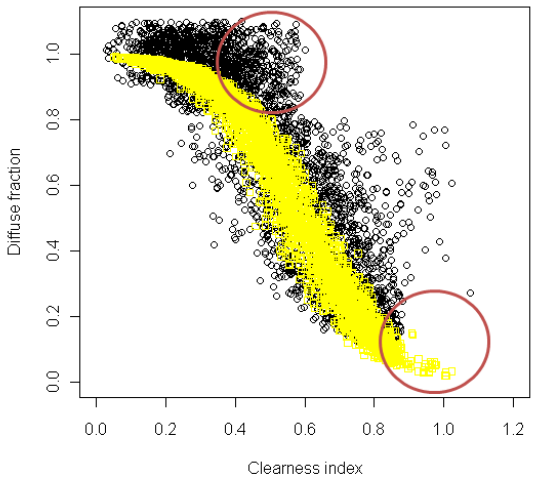


Figure 5.3: Hourly diffuse radiation against hourly clearness index for Chasseral (NE). The estimation of BRL is also shown (in yellow). The red circles represent the main modeling problems.

an inadequate model choice. In addition, Table 4.1 shows a clear improvement of the global performance of the models, when the three stations are omitted from the study. These results would then confirm the limitations of all models in the estimation of data for low solar elevations.

Variability index

The clearness index k_t is the measure of global radiation reaching the surface compared to the extraterrestrial radiation. A low k_t combined with a high diffuse fraction k_d corresponds to a overcast situation. Whereas a high k_t with low k_d indicates clear sky. For this reason, an investigation on the time evolution of the clearness index could give more information about the weather stability: a stable k_t would correspond to an overcast or clear sky situation; on the other hand, an inconstant k_t would show unstable cloud conditions, i.e. broken clouds.
Some of the models analyzed in this study introduced a parameter trying to reproduce this phenomena. Skartveit and Olseth introduced a variability index ω_3, described by Equation 3.2, Perez a stability index $\Delta K_t'$ defined by Equation 3.31 and Equation B.9 is the definition of the persistence factor Ψ introduced in BRL. These three parameters compare the clearness index at time t with those of the previous $(t-1)$ and following $(t+1)$ hour. A small value corresponds to stable weather.

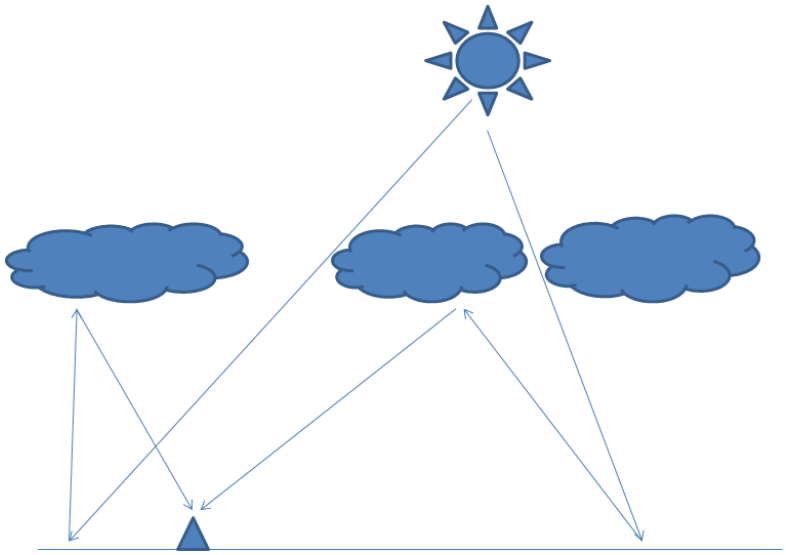


Figure 5.4: Scheme representing a partly cloudy weather situation, where no direct radiation is detected by the instrument (the triangle). However, solar radiation, after reaching the surface, can be reflected and reaches the instrument after back-scattering by clouds.

In addition, BRL has the daily clearness index K_t (Equation 3.19) as parameter, in order to consider daily variances.
In my opinion, it is essential to introduce a parameter reproducing better cloud conditions. This could be a reason why the models of S&O and BRL generally perform better than Reindl-Helbig and Maxwell.
However, these models are not appropriate for an operational use, i.e. estimation of diffuse fraction in nowcasting mode, because it needs data of future time intervals or even of a complete day. Reindl-Helbig and Maxwell would then be more adequate for this purpose. The new method developed by Dürr would be even more adequate. In fact, his model has the transmissivity (Equation 3.38) as principal parameter and considers its variation over the previous 30-minutes interval, in terms of range and standard deviation.

Atmospheric composition

In Chapter 4.1, the difference in diffuse fraction during clear days was described in detail for a selection of five stations. Figure 4.6 shows that the three polar stations have a similar behavior, compared to Geneva and Alice Springs.
In Alice Springs, the atmosphere is clearer, dryer and has no relevant presence of pollution. For these reasons, it has the lowest minimal diffuse fraction $k_{d,min}$ compared to the other stations. Geneva is a polluted city, with consequently a high aerosol concentration in the at-

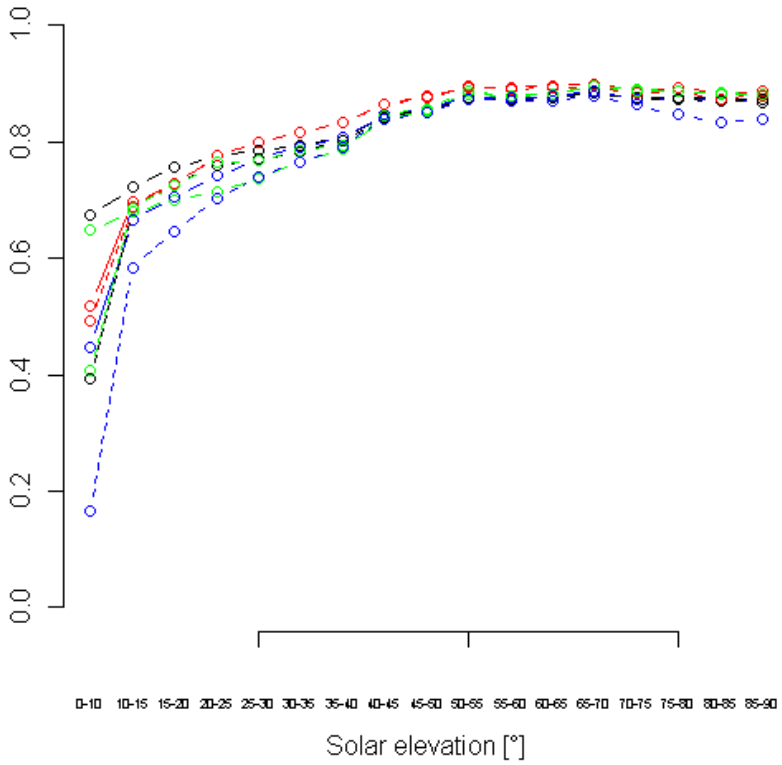


Figure 5.5: The coefficient of determination R^2 (lined) and the coefficient of efficiency E (dashed) against the solar elevation for Reindl* (black), S&O (red), BRL (blue) and Maxwell (green).

mosphere. Furthermore, the anthropogenic input of tropospheric sulfate, which attenuates incoming radiation, is increasing [Balling and Idso, 1991].It is principally confined in the Northern Hemisphere and could be the reason why $k_{d,min}$ of Summit Station is higher than those of the Antarctic stations. Moreover, the polar atmosphere has a high concentration of ice crystals, which may enhance multiple reflections between the icy surface and the atmosphere, and, as a consequence, a high $k_{d,min}$. In addition, they present a higher maximal clearness index $k_{t,max}$, because of a clearer atmosphere and, consequently, less back scatter or attenuation by atmospheric components.
As Gueymard [2005] stated, turbidity plays an important role under clear sky: it characterizes the extinction effect of aerosols on solar radiation.
Aerosols are either from terrestrial or marine origin. The amount of dust is greater on land than on water and greater in drier seasons. Rain reduces the number of particles, but increases the size of those that remain. Therefore, turbidity remains unchanged after rain [Iqbal, 1983]. Turbidity is an optical parameter of the atmosphere, related to horizontal visibility [Iqbal, 1983].

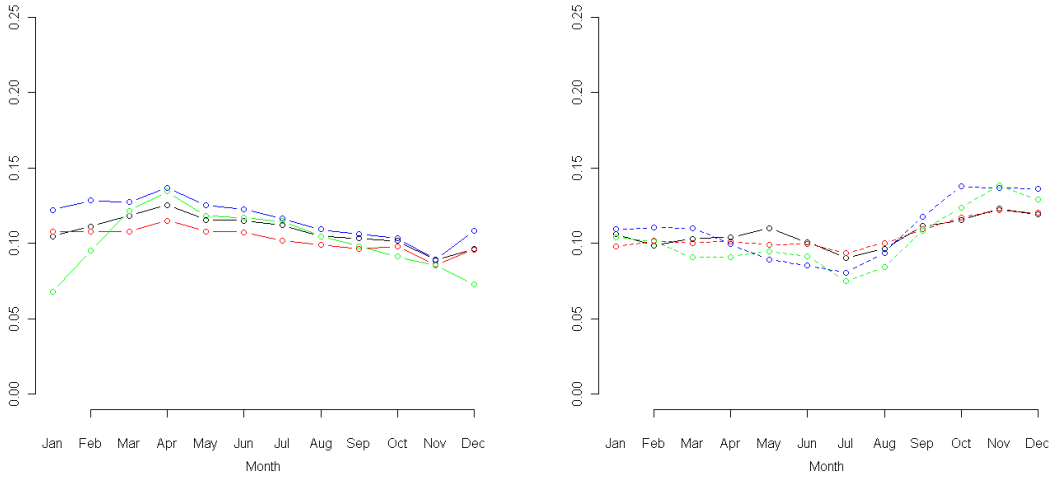


Figure 5.6: Mean Absolute Error (MAE) in function of the months for stations of the Northern (left) and Southern Hemisphere (right) for Reindl-Helbig (black), S&O (red), BRL (blue) and Maxwell (green).

In Figure 5.6, the variation of MAE through the months is represented: a clear seasonal dependency is visible. There is an increased error during the spring. In fact, in winter the atmosphere is clearer than in and the turbidity is then affected.

Skartveit and Olseth [1998] also shared the opinion that turbidity information would have a potential of improving the performance of a diffuse fraction model.
Gonzalez and Calbo [1999] stated that the diffuse fraction depends on aerosols and water vapor, by producing a variability on longer time periods. The Linke's turbidity considers the effect of aerosols and water vapor [Ineichen, 2007]. Perez includes this parameter in its model as a constant. It should be considered for the development of a future model.

Returning to Figure 4.6, the three polar stations present a relatively high $k_{d,min}$, because of a permanently low solar elevation. At low solar elevations, the incoming radiation has to cross a longer atmosphere column and consequently, a relevant portion of radiation is forward scattered by the atmospheric particles. The air mass reflects the atmospheric column, which the incoming radiation crosses before reaching the surface. It is solar elevation dependent and considers the effect of permanent gases.
This air mass is included in the DISC-(Maxwell) and Perez models. Dürr also included this parameter, in the definition of the transmissivity (see Equation 3.38).

5.3 Temporal Resolution

Averaging 1-minute (or 10-minutes, for Meteo Schweiz stations) interval data to hourly data, the extreme values disappear. Figure 5.7 illustrates the results on averaging 1-minute interval data to 10-minutes, 30-minutes interval and finally hourly data for Payerne. It can

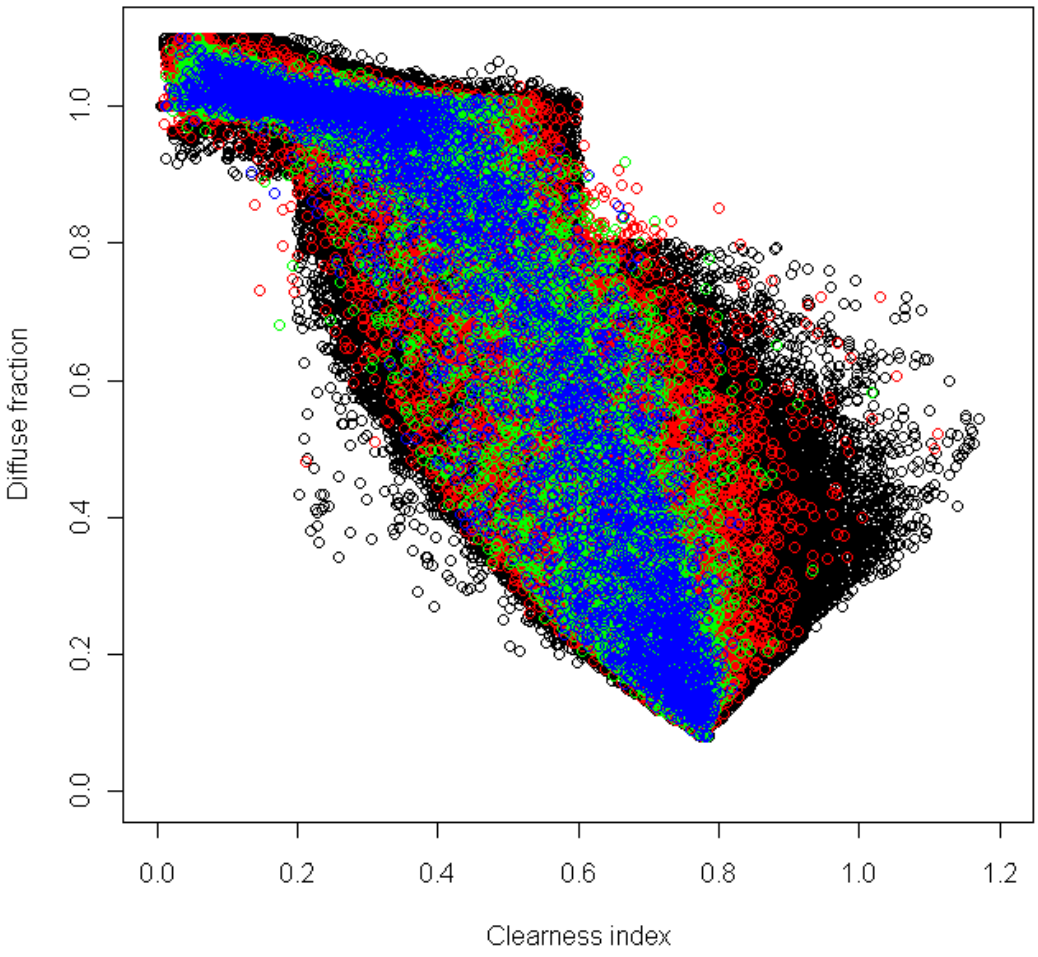


Figure 5.7: Comparison of diffuse fraction against clearness index with 1-minute (black), 10-minutes (red), 30-minutes (green) and hourly interval data (blue).

be observed that hourly data have less variability and do not present any clearness index greater than 0.80. As a consequence, the hourly estimated data have a smaller error than data for a smaller time interval. Table 4.10 summarizes the statistical results for models applied on such data. For example, BRL has a MBE which increases from 0.0327 (hourly data) to 0.0501 (1-minute interval data). However, the improvement is not that evident. For instance, for the S&O, the coefficient of efficiency is decreasing from 0.9029 (1-minute interval data) to 0.8952 (hourly data). These fluctuations are due to the length of the integration path. In fact, an increased integration length tends to cancel the variability in the data. In general, the models present satisfactory results even when applied to data with shorter time steps. Skartveit and Olseth [1998] were having a poorer performance on 10-minutes intervals data compared to hourly data. Furthermore, Boland et al. [2001] showed that their model applied on 15-minutes interval data is performing as well as when applied on hourly data. Ineichen [2007] obtained similar results as in this study while applying the models of Perez and Maxwell to data with different time step lengths.

5.4 Universal model

The previous sections listed different reasons why the determination of a model applicable to data from stations worldwide with different time steps is not trivial. Reindl-Helbig was included in the analysis as being the model used nowadays in SNOWPACK and ALPINE 3D. However, it results to be inappropriate: it only considers the clearness index and the solar altitude as parameters. Furthermore, this model is an improvement of four models proposed by Reindl et al. [1990]. In addition, it assumes that the diffuse fraction is constant for high clearness indexes, which is not reflecting the measured values.
The suitable model should consider cloud conditions variability, like S&O, BRL and Dürr do. The first two models, however, are not appropriate for an operational use: clearness indexes of future intervals are needed. Moreover, BRL uses a logistic function which is not appropriate for high clearness indexes, because the diffuse fraction is not tending to zero as the model would estimate.
The universal model should also include the air mass, as Maxwell, Perez and Dürr do. Being based on an exponential function, Maxwell and Perez are inappropriate to estimate values for high clearness indexes because modeling diffuse fractions, which values are physically impossible.
Therefore the approach of Dürr seems to be the most suitable. However, 1-minute interval data have to be at disposal. Furthermore, the need of station-specific parameters (to calculate the transmissivity and albedo offsets) impedes an easy transportability of the model.
On the other hand, Perez should be an improvement of Maxwell, but it seems to have problems with the correction matrix. As a consequence, erroneous values are modeled. However, the error should be located in the code, as previous analyses of the model were not presenting this problem.
Finally, not every station is measuring atmospheric parameters (like relative humidity and surface temperature) with a so short time interval as for solar radiation. An ideal model should require as few parameters as possible.

Chapter 6

Conclusions and Outlook

Different studies have been conducted to determine a correlation between the global incoming radiation and its diffuse fraction. However, the findings of those studies are restricted to local application. The present work has focused on the analysis and comparison of a selection of four models: Reindl-Helbig [Helbig, 2009], Skartveit and Olseth [Skartveit and Olseth, 1998], Boland-Ridley-Lauret [Ridley et al., 2010] and Maxwell [Maxwell, 1987]. The models of Perez et al. [1990b] and Dürr [2004] were considered, as well.
An exaustive worldwide dataset of hourly integrated global and diffuse radiation data were used to test the above models.
In general, the first four models are well estimating the measured diffuse fraction, with exception to data from polar stations, where solar elevations are permanently low. In facts, those models have limitations for solar elevations smaller than 30 °.
In addition, at intermediate clearness indexes, the models are unable to correctly estimate high diffuse fractions. Radiation is reflected between the ground and clouds and it is difficult to reproduce this behavior. Overall, the model proposed by Skartveit and Olseth is statistically performing the best. Furthermore, it is able to reproduce diffuse radiation for a high clearness index, in contrast to the other compared models.
The models were tested for different time resolutions (1, 10, 30 and 60 minutes) on data from Payerne. A larger dispersion of the diffuse fraction was visible for 1-minute data. However, no relevant differences in terms of performance were detected.

Because of lack of time, it was, however, not possible to develop a new universal model. It could be obtained by improving an existing model or by developing a completely new one.
Clearness index, solar elevation are significant predictors. A variability index should be included to account for the variability due to cloudiness. Air mass and turbidity would describe the atmospheric composition.
Further investigations on Perez should be conducted. Furthermore, it would be advantageous to apply the model of Dürr on 1-minute interval data and consequently average the modeled values over one hour. His approach appears to be the most appropriate and should be compared to the other results, at least for Payerne.
The introduction of albedo, as parameter, could be considered, as suggested by Skartveit and Olseth [1998].
In addition, the performance of the existing models could be improved only by adapting the indexes of variability used [Ineichen, 2007]. By the development of a complete new model, the principal component analysis could be applied to determine the explanatory parameters. Moreover, as Dürr seems to be the most adequate method, a regression tree

could be applied to determine the new model.
To conclude, it would be suitable to work with a dataset covering a time interval longer than one year.

Bibliography

Sun position during the day. "http://pda.88000.org/12_Sun_Position_During_The_Day.html", December 2010.

AMS. Glossary. "http://amsglossary.allenpress.com", November 2010.

ArcGis. Calculating solar radiation. "http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?TopicName=Calculating%20solar%20radiation", November 2010.

R.C. Balling and S.B. Idso. Sulfate aerosols of the stratosphere and troposphere: combined effects on surface air temperature. *Theoretical and applied climatology*, 44:239–241, 1991.

J. Boland and B. Ridley. Recent advances in modelling diffuse radiation. In *Proceedings of ISES Solar World Congress 2007: Solar energy and human settlement*, 2007.

J. Boland, L. Scott, and M. Luther. Modeling the diffuse fraction of global solar radiation on a horizontal surface. *Environmetric*, 12:103–116, 2001.

J. Boland, B. Ridley, and B. Brown. Models of diffuse solar radiation. *Renewable Energy*, 33:575–584, 2008.

B. Bourges. Improvement in solar declination computation. *Solar Energy*, 35:367–369, 1985.

BSRN. Bsrn stations map. "http://www.bsrn.awi.de/fileadmin/user_upload/Home/Maps/BSRN-Station-Global.png", November 2010.

B. Dürr. *The greenhouse effect in the alps by models and observations*. PhD thesis, ETH Zurich, 2004.

D.G. Erbs, S.A. Klein, and J.A. Duffie. Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation. *Solar Energy*, 28:293–302, 1982.

WSL SLF (Institute for snow and avalanche research). Erforschung der schneedeckenvariabilität. "http://www.slf.ch/ueber/organisation/schnee_permafrost/projekte/Measuring_and_modelling_spatial_variability_of_snowpack_stability/index_DE", November 2010.

C. Fröhlich. Solar irradiance variability since 1978. *Space Science Reviews*, 25(23):4377–4380, 2006.

C. Fröhlich and J. Lean. The sun's total irradiance: Cycles, trends and related climate change uncertainties since 1976. *Geophysical research letters*, 25:4377–4380, 1998.

J.D. Garrison. A study of the division of global irradiance into direct and diffuse irradiance at thirty three u.s. sites. *Solar Energy*, 35:341–351, 1985.

J.A. Gonzalez and J. Calbo. Influence of the global radiation variability on the hourly diffuse fraction correlations. *Solar Energy*, 65:119–131, 1999.

C.A. Gueymard. Importance of atmospheric turbidity and associated uncertainties in solar radiation and luminous efficacy modelling. *Energy*, 30:16031621, 2005. doi: doi:10.1016/j.energy.2004.04.040.

N. Helbig. *Aplication of the radiosity approach to the radiation balance in complex terrain.* PhD thesis, University Zurich, 2009.

N. Helbig, H. Löwe, and M. Lehning. Radiosity approach for the shortwave surface radiation balance in complex terrain. 2009.

Faculty of Geology Illinois Central College. Sun radiation and heating of the atmosphere. "`http://faculty.icc.edu/easc111lab/labs/labi/prelab_i.html`".

P. Ineichen. Comparision and validation of three global-to-beam irradiance models against ground measurements. 2007.

M. Iqbal. Prediction of hourly diffuse solar radiation from measured hourly global solar radiation on a horizontal surface. *Solar Energy*, 24:491–03, 1980.

M. Iqbal. *An introduction to solar radiation.* Academic Press, 1983.

C.P. Jacovides, F.S. Tymvios, V.D. Assimakopoulos, and N.A. Kaltsounides. Comparative study of various correlations in estimating hourly diffuse fraction of global solar radiation. *Renewable Energy*, 31:2492–2504, 2006.

N. Kämpfer. Lecture "atmospheric physics", institute of applied physics, university of bern. *Unpublished*, 2009.

P. Lauret, J. Boland, and B. Ridley. Derivation of a solar diffue fraction model in a bayesian framework. *Case Studies in Business, Industry and Government Statistics*, 3(1):108–122, 2010.

C. Lee. *An overview of the requirements and operational issues, Derived from the BSRN Operations Manual Version 2.1.* Kipp & Zonen, 2009.

D.R. Legates and G.J. McCabe. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water resources research*, 35(1):233–241, 1999.

B.Y.H. Liu and R.C. Jordan. The interralationship and characteristic distribution of direct, diffuse and total solar radiation. *Solar Energy*, 4:1–9, 1960.

Ch. Marty. *Surface radiation, cloud forcing and greenhouse effect in the Alps.* PhD thesis, Swiss Federal Institut of Technology Zurich, 2000.

Wolfram MathWorld. Coefficient of correlation. "`http://mathworld.wolfram.com/CorrelationCoefficient.html`", November 2010.

E. L. Maxwell. A quasi-physical model for converting hourly global horizontal to direct normal insolation. *Report SERI/TR-215-3087, Solar Energy Institute, Golden, CO*, 1987.

A. De Miguel, J. Bilbao, R. Aguiar, H. Kambezidis, and E. Negro. Diffuse solar irradiation model evaluation in the north mediterranean belt area. *Solar Energy*, 70:143–153, 2001.

B. Molineaux and P. Ineichen. Impact of pinatubo aerosols on the seasonal trends of global, direct and diffuse irradiance in two northern mid-latitude sites. *Solar Energy*, 58:91–101, 1996.

J.E. Nash and J.V. Suttcliffe. River flow forecasting through conceptual models, part i - a discussion of principles. *Journal of Hydrology*, 10:282–290, 1970.

G. Notton, C. Cristofari, M. Muselli, and P. Poggi. Calculation on an hourly basis of solar diffuse irradiations from global data for horizontal surfaces in ajaccio. *Energy conversion and management*, 45:2849–2866, 2004.

A. Ohmura, E.G. Dutton, B. Forgan, C. Fröhlich, H. Gilgen, H. Hegner, A. Heimo, G. König-Langlo, B. McArthur, G. Müller, R. Philipona, R. Pinker, C.H. Whitlock, K. Dehne, and M. Wild. Baseline surface radiation network (bsrn/wcrp): New precision radiometry for climate research. *Bulletin of the American Meteorological Society*, 1998.

T.R. Oke. *Boundary Layer Climates*. Routledge, 1978.

J.F. Orgill and K.G. Hollands. Correlation equation for hourly diffuse radiation on a horizontal surface. *Solar Energy*, 19:357–359, 1977.

M.C. Peel, B.L. Finlayson, and T.A. McMahon. Updated world map of the köppen-geiger climate classification. *Hydrology and Earth System Sciences*, 11:1633–1644, 2007.

R. Perez, P. Ineichen, R. Seals, J. Michalsky, and R. Stewart. Modeling daylight availability and irradiance components from direct and global irradiance. *Solar*, 44:271–289, 1990a.

R. Perez, P. Ineichen, R. Seals, and A. Zelenka. Making full use of the clearness index for parameterizing hourly insolation conditions. *Sola energy*, 45(2):111–114, 1990b.

R. Perez, R. Seals, A. Zelenka, and P. Ineichen. Climatic evolution of models that predict hourly direct irradiance from hourly global irradiance: prospects for performance improvements. *Solar Energy conversion and management*, 44 (2):99–108, 1990c.

R. Perez, P. Ineichen, E. Maxwell, R. Seals, and A. Zelenka. Dynamic global-to-direct irradiance conversion models. *ISES Solar World Congress, Denver USA*, pages 951–956, 1991.

D.T. Reindl, W.A. Beckman, and J.A. Duffie. Diffuse fraction correlations. *Solar Energy*, 45:1–7, 1990.

B. Ridley, J. Boland, and P. Lauret. Modelling of diffuse solar fraction with multiple predictors. *Renewable Energy*, 35:478–483, 2010. doi: 10.1016/j.renene.2009.07.018.

R.A. Rohde. Atmospheric transmission. "http://www.globalwarmingart.com/wiki/File:Atmospheric_Transmission_png", December 2010.

G. Seiz and N. Foppa. *National Climate Observing System (GCOS Switzerland)*. Publication of the Federal Office of Meteorology and Climatology MeteoSwuiss and ProClim, 2007.

K. Sen. Estimates of the regression coefficient based on kendall's tau. *Journal of the American Statistical Association*, 63(324):1379–1389, 1968. URL http://www.jstor.org/stable/2285891.

A. Skartveit and J.A. Olseth. A model for the diffuse fraction of hourly global radiation. *Solar Energy*, 38:271–274, 1987.

A. Skartveit and J.A. Olseth. An hourly diffuse fraction model with correction for variability and surface albedo. *Solar Energy*, 63(3):173–183, 1998.

J.W. Spencer. Fourier series representation of the position of the sun. *Search*, 2(5):172, 1971.

Tingilinde. Latest sunset, earliest sunrise. "http://tingilinde.typepad.com/starstuff/2006/06/latest_sunset_e.html", December 2010.

M. Willis. Propagation tutorial. "http://www.mike-willis.com/Tutorial/PF10.htm", November 2010.

WMO. *Guide to Meteorological Instruments and Methods of Observation*. World Meteorological Organization, 2008.

M. Zappa, F. Pos, U. Strasser, P. Warmerdam, and J. Gurtz. Seasonal water balance of and alpine catchment as evaluated by different methods for spatially distributed snowmelt modelling. *Nordic Hydrology*, 34(3):179–202, 2003.

Appendix A

Stations details

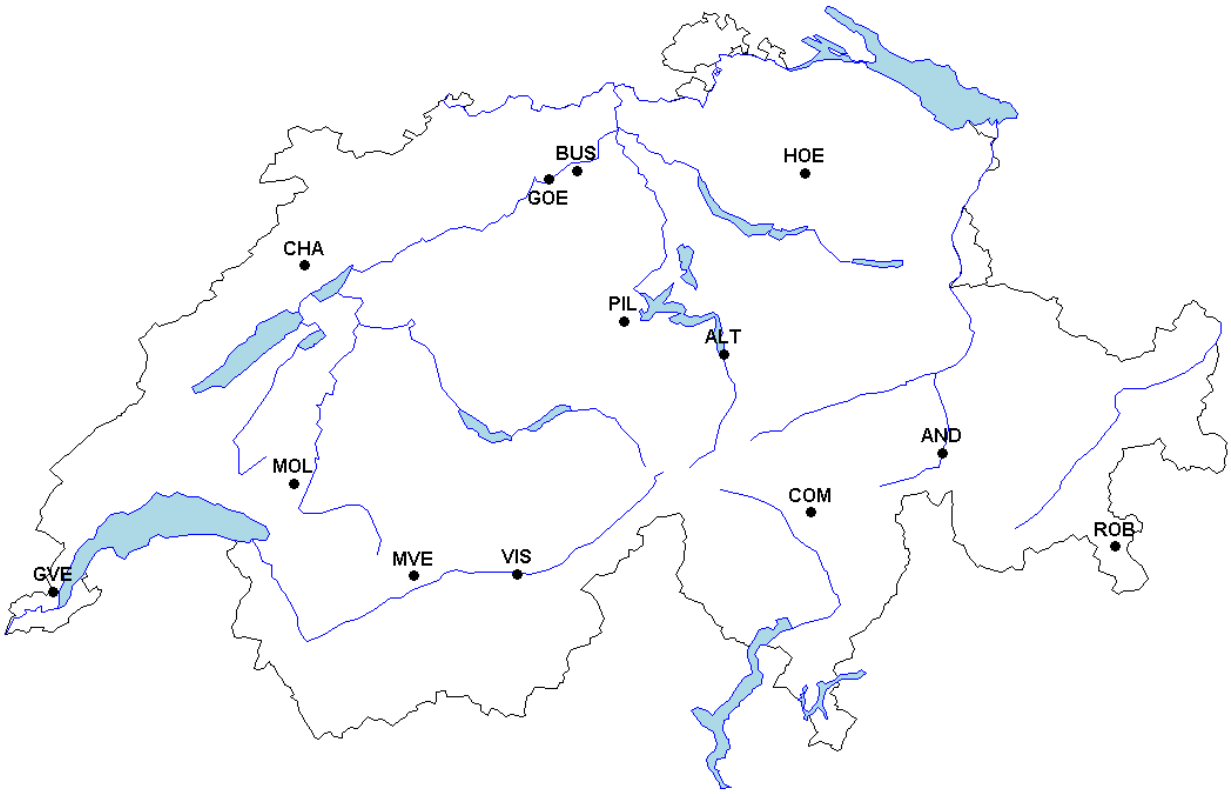


Figure A.1: Graphical illustration of the Swiss stations.

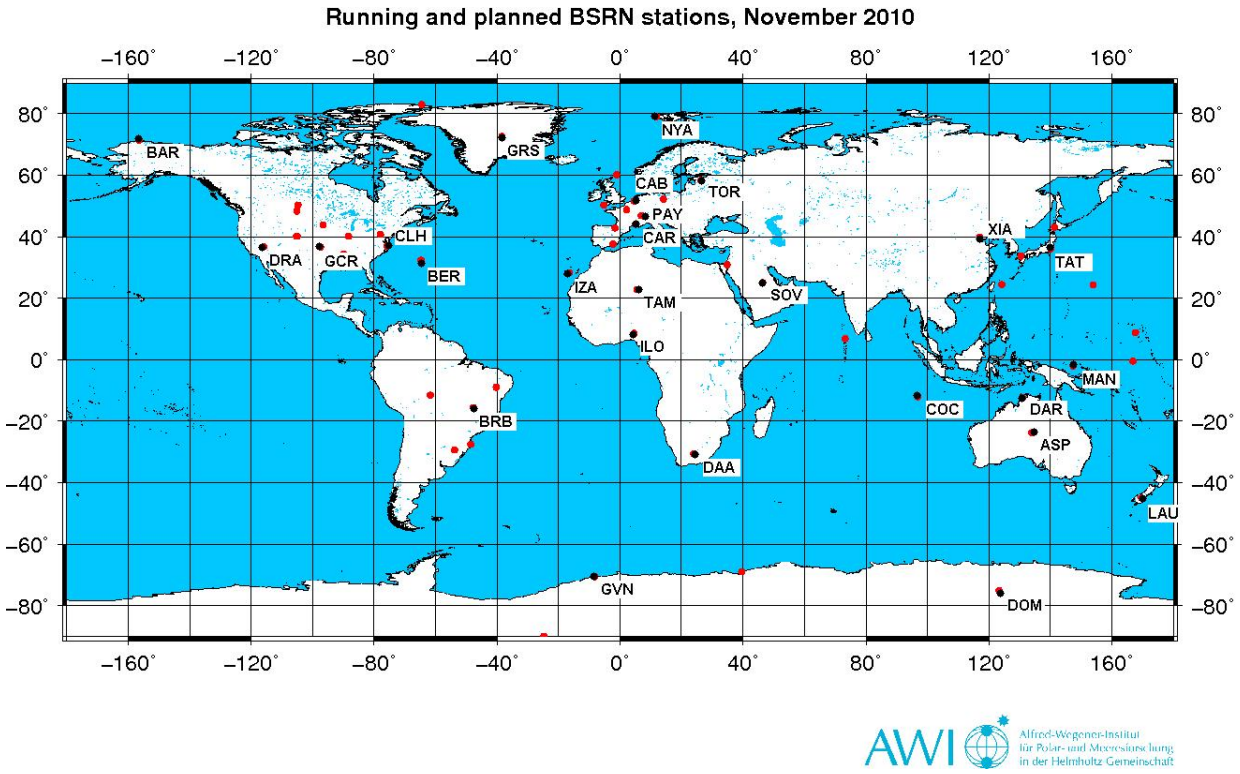


Figure A.2: World map with BSRN stations (red points). The black points represents the selected stations BSRN [2010].

Table A.1: Description of Swiss Stations (ASRB Network)

Station	Label	Nr	Region	Coordinates	Elevation (m)	Period	Climate [a]
Robbia	ROB	28	Graubünden	46.35 °N 10.06 °E	1078	Feb09-Jan10	Cfb
Le Moléson	MLS	38	Fribourg	46.55 °N 7.02 °E	1974	Feb09-Jan10	Dfb
Montana	MVE	40	Wallis	46.30 °N 7.46 °E	1427	Mar09-Feb10	Dfb
Chasseral	CHA	42	Neutschatel	47.13 °N 7.05 °E	1594	Feb09-Jan10	Dfb
Pilatus	PIL	43	Luzern	46.98 °N 8.25 °E	2106	Mar09-Feb10	Dfb
Altdorf	ALT	44	Uri	46.89 °N 8.62 °E	438	Feb09-Jan10	Cfb
Geneva	GVE	58	Geneva	46.25 °N 6.13 °E	413	Jun09-Jun10	Cfb
Visp	VIS	64	Visp	46.30 °N 7.84 °E	639	Feb09-Jan10	Cfb
Buchs	BUS	67	Aarau	47.38 °N 8.08 °E	387	Feb09-Jan10	Cfb
Goesgen	GOE	80	Aarau	47.36 °N 7.97 °E	380	Feb09-Jan10	Cfb
Acquarossa-Comprovasco	COM	92	Ticino	46.46 °N 8.93 °E	575	Feb09-Jan10	Cfb
Hoernli	HOE	94	Zurich	47.37 °N 8.94 °E	1132	Feb09-Jan10	Cfb
Andeer	AND	405	Graubünden	46.61 °N 9.43 °E	987	Feb09-Jan10	Cfb

[a] Köppen Classification (Peel et al. [2007], Ohmura et al. [1998]):
Af: tropical rainforest, *Aw*: savanna, *Bs*: steppe, *Bw*: desert, *Cfa*: humid subtropical, *Cfb*: dry-summer subtropical, *Cs*: mediterranean, *Dfb*: warm summer continental, *Dfc*: boreal (taiga), *Dwa*: hot summer continental, *Ef*: ice cap, *ET*: tundra.

Table A.2: Description of Worldwide Stations (BSRN Network)

Station	Label	Nr	Country	Coordinates	Elevation (m)	Period	Climate [a]
Alice Springs	ASP	1001	Australia	23.80 °S 133.89 °E	547	Jan08-Dec08	Bw
Barrow	BAR	1002	Alaska, USA	71.32 °N 156.61 °W	8	Jan09-Dec08	ET
Bermuda	BER	1003	Bermuda	32.27 °N 64.67 °W	8	Jan08-Dec08	Af
Cabauw	CAB	1008	The Netherlands	51.97 °N 4.93 °E	0	Jan08-Dec08	Cfb
Carpentras	CAR	1009	France	44.08 °N 5.06 °E	100	Jan08-Dec08	Cs
Chesapeake Light	CLH	1010	North Atlantic Ocean	36.90 °N 75.71 °W	37	Jan08-Dec08	Cfa
Cocos Island	COC	1011	Cocos Islands	12.15 °S 96.83 °E	0	Mar07-Feb08	Af
Concordia Station, Dome C	DOM	1012	Antarctica	75.10 °S 123.38 °E	3233	Jan08-Dec08	Ef
Desert Rock	DRA	1013	Nevada, USA	36.63 °N 116.02 °W	1007	Jan08-Dec08	Bw
Goodwin Creek	GCR	1016	Mississippi, USA	34.25 °N 89.87 °W	98	Jan08-Dec08	Cfa
Georg von Neumayer	GVN	1017	Antarctica	70.65 °S 8.25 °W	42	Jan08-Dec08	Ef
Momote	MAN	1019	Papua New Guinea	2.06 °S 147.42 °E	6	Jan08-Dec08	Af
Ny-Alesung	NYA	1020	Ny-Alesund, Spitzbergen	78.92 °*N* 11.95 °E	11	Jan09-Dec09	ET
Tamanrasset	TAM	1024	Algeria	22.78 °N 5.51 °E	1385	Jan08-Dec08	Bw
Tateno	TAT	1025	Japan	36.05 °N 140.13 °E	25	Jan08-Dec08	Cfa
Toravere	TOR	1026	Estonia	58.25 °N 26.46 °E	70	Jan08-Dec08	Dfc
Payerne	PAY	1027	Switzerland	46.81 °N 6.94 °E	491	Jan09-Dec09	Cfb
Brasilia	BRB	1029	Brazil	15.60 °S 47.71 °W	1023	Sep06-Oct07	Aw
De Aar	DAA	1031	South Africa	30.67 °S 23.99 °W	1297	Jan04-Dec04	Bs
Darwin	DAR	1032	Australia	12.42 °S 130.89 °E	30	Jan07-Dec07	Aw
Xianghe	XIA	1033	China	39.75 °N 116.96 °E	33	Sep07-Aug08	Dwa
Solar Village	SOV	1035	Saudi Arabia	24.91 °N 46.41 °E	650	Jan02-Dec02	Dw
Summit Station	GRS	1042	Greenland	72.58 °N 38.50 °W	3208	Jan09-Dec09	Ef
Ilorin	ILO	1038	Nigeria	8.53 °N 4.57 °E	350	Jan04-Dec04	Aw
Izana	IZA	1039	Spain	28.31 °N 16.50 °W	2372.9	Mar09-Feb10	Bw
Lauder	LAU	1040	New Zealand	45.04 °S 169.69 °E	350	Jan07-Dec07	Cfa

[a] Köppen Classification (Peel et al. [2007], Ohmura et al. [1998]):
Af: tropical rainforest, *Aw*: savanna, *Bs*: steppe, *Bw*: desert, *Cfa*: humid subtropical, *Cfb*: dry-summer subtropical, *Cs*: mediterranean, *Dfb*: warm summer continental, *Dfc*: boreal (taiga), *Ef*: ice cap, *ET*: tundra.

Appendix B

Astronomical equations and radiation definitions

Being the principal source of energy for the Earth, the Sun is crucial for the life on Earth. Its radiation is an important component for the global energy balance which drives the weather and climate systems.
Knowing the exact position of the Sun, the maximal exploitation of its incoming radiation can be achieved. This is, for example, important for the positioning of photovoltaic and thermal panels.

In this chapter we provide a brief summary of the principal astronomical equations that are necessary to describe the solar positions and the different types of radiation.

B.1 The solar position

The Earth is rotating around the Sun with a frequency of one day. The position of the Sun is varying as illustrated in Figure B.1.

Figure B.1: Solar postions during a day [sun, 2010]

Furthermore, for a fixed day time, the apparent sun position is also changing in the

course of the year. For this reason, astronomical relationships have to be determined in order to compute the position of the Sun during the entire year.

Equation of time

The solar day is defined as the time that is needed by the Sun to achieve a complete tour of the Earth [Iqbal, 1983]. This does not necessarily correspond to 24 hours and it varies from year to year. Figure B.2 illustrates how the relative Sun position is moving across the sky: pictures of the Sun taken by an immobile photographer at the same time of the day have been superimposed. It can be seen that after one whole year of observations, the Sun is computing a eight-shape circuit.

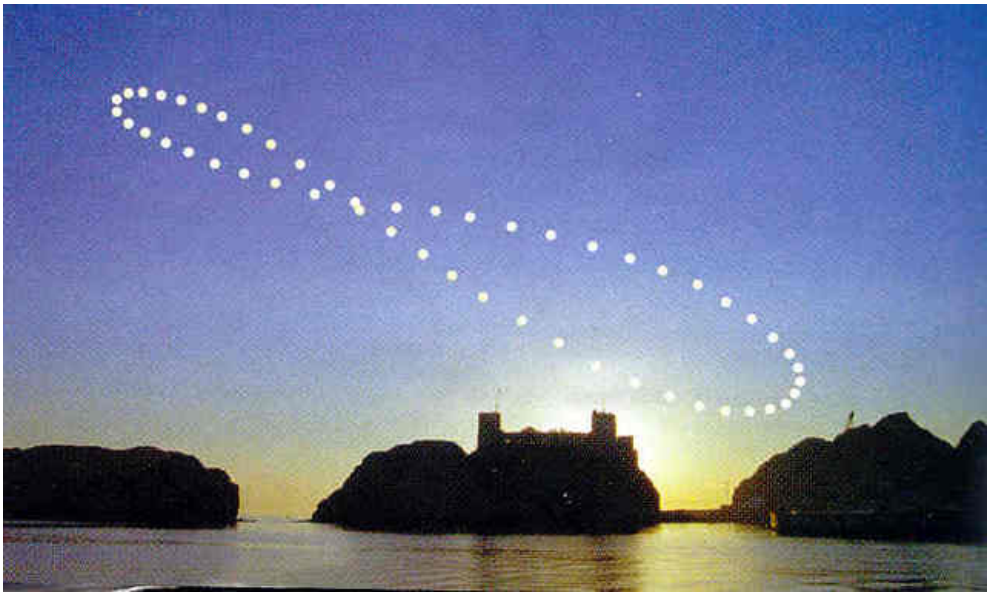


Figure B.2: Analemma [Tingilinde, 2010]

The principal causes of this phenomena are the elliptical shape of the terrestrial orbit around the Sun and the tilt of the Earth in relation to the plane of its orbit.
As a consequence, at 12 noon the Sun does not have the same position in the different months. The curve is described by the so called *equation of time* E_t and can reach up to 16 minutes approximately [Helbig, 2009]. The following formula has been developed by Spencer [1971] and then truncated by Iqbal [1983] [Helbig, 2009]:

$$E_t = 229.18(0.000075 + 0.01868\cos\Gamma - 0.032077\sin\Gamma - 0.014615\cos 2\Gamma - 0.04089 sin 2\Gamma) \quad \text{(B.1)}$$

where Γ is the **day angle** between the radius vector and the radius vector on 1^{st} of January in the ecliptic plane [Iqbal, 1983].
Assuming a circular orbit, the day angle can be defined as:

$$\Gamma = \frac{2\pi(d_n - 1)}{365.2425} \quad \text{(B.2)}$$

where d_n is the **day number**: 1 for 1^{st} of January and 365 for 31^{st} of December.
Moreover, the mean year length of 365.2425 days of the Gregorian calendar, replacing 365, was introduced in Spencer's formula (Spencer [1971] and Helbig [2009]).
The coefficient of Equation B.1 converts the expression from radians to minutes [Iqbal, 1983]:

$$229.18 = 24 \cdot \frac{60}{2 \cdot \pi}.$$

Apparent solar time

Data such as wind speed and temperature are often recorded in terms of local time, whereas the solar radiation data is recorded in terms of apparent solar time, also named true solar time. In fact, it is necessary to have data recorded at the same period of time in order to compute the angle of incidence of solar radiation.
For this reason, in addition to the equation of time (Equation B.1), a longitude correction L_{corr} is needed [Iqbal, 1983]:

$$L_{corr} = 4(L_{loc} - L_{st}) \tag{B.3}$$

where L_{loc} is the local longitude and L_{st} is the standard longitude, defined as:

$$L_{st} = 15 \left\lfloor \frac{L_{loc}}{15} \pm 0.5 \right\rfloor$$

"'+"' applies to positive (east direction) and "'-"' to negative (west direction) L_{loc}.
In Equation B.3, a correction factor of 4 minutes for every degree is applied. It corresponds to the time needed by the Sun to cross 1° degree of longitude and is positive if the local meridian is East of the standard and is negative West of the standard meridian.
As a result, the standard time is usually converted in local apparent time t_{loc}, affected by these two corrections:

$$\begin{aligned} t_{loc} &= t_{mean} + E_t \\ &= t_{sd} + L_{corr} + E_t \\ &= t_{sd} + 4(L_{st} - L_{loc}) + E_t \end{aligned} \tag{B.4}$$

where t_{mean} and t_{sd} are the local mean time and the local standard time, respectively. It is important to note that the summer time is not accounted for [Helbig, 2009].

Hour angle

At apparent solar noon, the Sun crosses the meridian. The angular displacement of the Sun East or West of this local meridian is named the hour angle ω. In other words, it corresponds to the angle measured at the celestial pole between the observer's meridian and the solar meridian and it varies by 15 ° every hour (Iqbal [1983],Helbig [2009]):

$$\omega = 15 * (12 - t_{loc}) \tag{B.5}$$

The hour angle is zero at noon, positive before solar noon and negative afterwards.

Solar declination angle

The Sun has a diurnal variation of the incident radiation, which is due because of the rotation of the Earth around its axis. Furthermore, it has a seasonal variation, which is attributed to the rotation of the Earth around the Sun Helbig [2009]. This rotation is counterclockwise and the plane of the Equator is inclined of approximately 23.5 ° with respect to the equatorial plane. An ecliptic plane is the plane of revolution of the Earth around the Sun and the axis around which the earth rotates is named polar axis. The angle between the two remains unchanged [Iqbal, 1983].
The position of the Sun during its variations is described by two coordinates: the hour angle ω (Equation B.5) and the solar declination angle δ. They constitute the equatorial system. The solar declination angle is the angle between a line joining the centers of the Sun and the Earth to the equatorial plane. It has a daily variation and, in the Northern Hemisphere, its maximum is equivalent to $+23.45\,^{\circ}$ at the summer solstice and its minimum to $-23.45\,^{\circ}$ at the winter solstice. The angle is zero at the spring and the autumn equinoxes (Bourges [1985], Helbig [2009]), where the equinox is defined as either of the two points of intersection of the Sun's apparent annual path and the plane of the earth's equator, that is, a point of intersection of the ecliptic and the celestial equator AMS [2010]. The spring equinox (called also vernal equinox) occurs approximately on $21st$ March. The dates of the equinoxes slightly change every year and the four season cited are referred to the Northern Hemisphere. For the Southern Hemisphere, they have to be reversed.
Spencer [1971] proposed a formula, which presents maximum of error of 0.0006 rad. Then, Bourges [1985] introduced a new formula (in radian) that reduced the error of Spencer to 0.0003 rad:

$$\begin{aligned}\delta = (&0.3723 + 23.2567 \sin\Gamma_{ne} - 0.758 \cos\Gamma_{ne} \\ &+ 0.1149 \sin 2\Gamma_{ne} + 0.3656 \cos 2\Gamma_{ne} \\ &- 0.1712 sin3\Gamma_{ne} + 0.0201 cos3\Gamma_{ne})\pi/180 \end{aligned} \tag{B.6}$$

where Γ_{ne} is the *day angle* in radians relative to spring equinox time, defined similarly to Equation B.2. It is relative to the *day number* d_{ne} of the year at noon:

$$d_{ne} = d_n - 0.5 - n_0, \tag{B.7}$$

where 0.5 is subtracted to return to noon. If d_{ne} is negative, 365.2425 is added.
The spring equinox time is expressed in days departing from the beginning of the year y [Bourges, 1985]:

$$n_0 = 78.801 + 0.2422(y - 1969) - \lfloor 0.25 * (y - 1969) \rfloor . \tag{B.8}$$

For the period 2000-2050, Helbig [2009] developed a new formula where the first term is recalculated and corresponds to 9.3542. It starts at 1^{st} January 2000 with the spring equinox day on 20 March 2000 at 8:30 UTC+1.

Solar azimuth

The apparent position of the Sun in the sky for a given location and for a given time is determined by the solar azimuth and the solar zenith angle [Oke, 1978]. **Solar azimuth** Ψ is the first Sun coordinate and is defined as the angle between the projection of the Sun-Earth vector and the true north projection on the horizontal plane of the observer's horizon:

$$\cos\Psi = (\sin\delta\cos\phi - \cos\delta\sin\phi\cos\omega)/\sin\theta_z \tag{B.9}$$

where ϕ is the latitude and θ_z the solar zenith angle.
It is anticlockwise from South and clockwise from North and is positive in the East direction and negative in the West direction. It varies between $0\,^\circ$ and $180\,^\circ$ [Iqbal, 1983].
The solar **zenith** θ_z is the angle of an observer on the earth's surface corresponding to the celestial sphere. The opposite point is called nadir.
It is also called zenith distance, is comprised between $0\,^\circ$ and $90\,^\circ$ and it can be defined as a trigonometric relationship [Iqbal, 1983]:

$$\begin{aligned}\cos\theta_z &= \sin\delta\sin\phi + \cos\delta\cos\delta\cos\omega \\ \Longleftrightarrow \theta_z &= \arccos(\sin\delta\sin\phi + \cos\delta\cos\delta\cos\omega)\end{aligned}$$

In this case, the mean solar time is exceeding twelve, therefore the azimuth angle needs to be subtracted from 2π.

B.2 Solar radiation

Essential energy is provided to the Earth by the Sun in the form of radiation. It is necessary for the life on the Earth. For example, it is the principal constituent of photosynthesis.

Radiation describes the process by which the energy is transmitted through the medium by photons. Photons have similar properties as particles and waves. The oscillations can be considered as waves traveling with a wavelength λ. Furthermore, radiation is able to travel in vacuum and moves at the speed of light ($3 \cdot 10^8 m/s$).
All bodies possessing energy are emitting radiation. For instance, the Sun emits radiation with wavelengths from 0.15 μm (ultra-violet) to 3.0 μm (near infra-red) mostly, whereas the Earth-Atmosphere system emits radiations ranging from 3.0 μm to 100 μm (see Figure B.3). For this reason, radiations in the range 0.15-3.0 μm are designed as *short-wave* or *solar* radiation, and radiation in the range 3.0-100 μm are named *long-wave* radiation [Oke, 1978].

Extraterrestrial radiation

The quantity of short-wave radiation entering the atmosphere is called *extraterrestrial radiation* (or *top of the atmosphere radiation*) I_0. It is a calculated diurnal value and is defined as follows [Oke, 1978]:

$$I_0 = E_0 \cdot I_{SC} \tag{B.10}$$

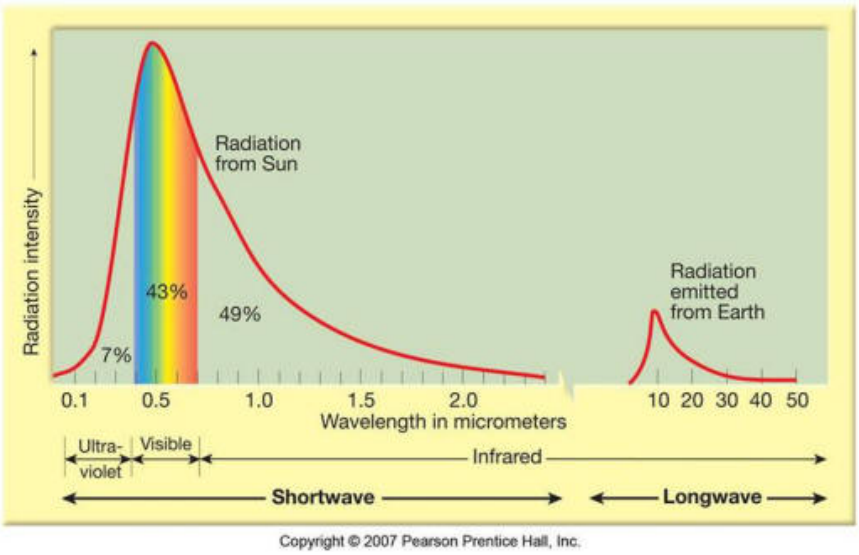


Figure B.3: Short-wave and long-wave radiation spectrum Illinois Central College

where E_0 is the Eccentricity correction and TSI the solar constant. The extraterrestrial value used is changing with latitude and time of year [Boland et al., 2001].

Often called the solar constant (SOLCON), the *Total Solar Irradiance* I_{SC} is defined as rate of incident solar energy at all wavelengths on an unita area, which is normally exposed to rays at one astronomical distance [Iqbal, 1983] ($1.496 \cdot 10^{11}$ m). This measure is not completely constant. From space measurements, small temporal fluctuations have been observed (because of solar variability fluctuations) and its average value is $1366.1 W/m^{-2}$ (Fröhlich and Lean [1998],Fröhlich [2006], Helbig [2009]), which is used in this study.
As already mentioned, the Earth has an elliptical orbit and its distance from the Sun is varying. Furthermore, the Sun has a highly variable temperature and energy emission. In addition, sunspots have live cycles of about 11 years . With other solar activities (such as solar flares), it produces variations in the solar spectrum distribution. These phenomena are affecting the TSI, which is then varying and a correction to the eccentricity is needed. Therefore, an *eccentricity correction factor* E_0 is applied [Helbig, 2009]; Spencer (with a maximum error of 0.0001) developed an expression for the eccentricity; it is dimensionless and represents the orbital diurnal changes:

$$\begin{aligned} E_0 &= (\frac{r_0}{r})^2 \\ &= 1.00011 + 0.034221 \cos\Gamma + 0.00128 \sin\Gamma + 0.000719 \cos 2\Gamma + 0.000077 \sin 2\Gamma \end{aligned} \quad \text{(B.11)}$$

In general, the distance r is expressed by a Fourier series with a number of coefficients. r_0 corresponds to the mean distance between the Sun and the Earth.

Solar radiation and processes

As solar radiation enters the atmosphere, it is either absorbed or scattered by particles in the atmosphere (Figure B.5). These interactions are produced by clouds, water vapor, aerosols

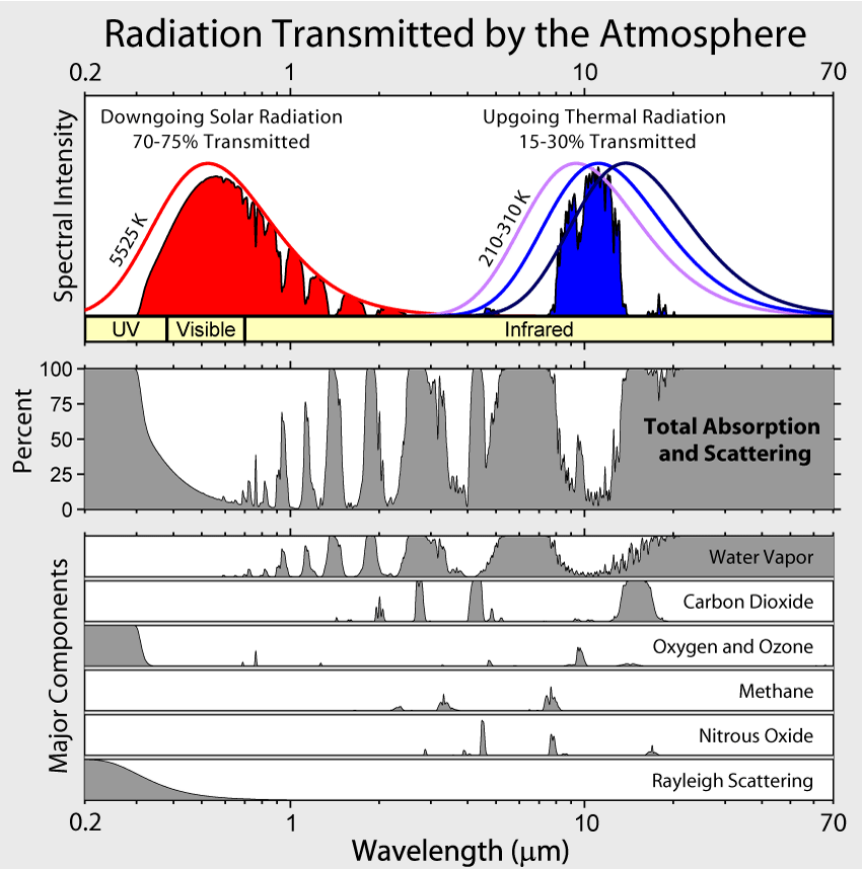


Figure B.4: Spectrum of transmission and scattering of shortwave radiation by different trace gases Rohde [2010].

(as salt crystals and dust particles) and other constituents of the atmosphere. **Aerosols** are defined as suspensions of liquid or solid particles in the air. Their diameter range from 10^{-4} μm to 10 μm. They can either absorb or scatter radiation and, as a consequence, alter the Earth energy balance. Aerosols also play a crucial role in the cloud formation process and precipitation, where water vapor molecules are coagulating to tiny drops. In fact, they also serve as condensation and freezing nuclei Kämpfer [2009].

Absorption Solar radiation can be absorbed by atmospheric gases (as O_3) and aerosols. This absorption yield to oscillations in the spectral radiation curve. Figure B.4 represents the spectrum of shortwave absorption and scattering by gases, as water vapor and methane. Atmospheric gases, as water vapor (H_2O), ozone (O_3) and carbon dioxide (CO_2), have the most important impacts on the solar spectra. Aerosols causes a intensity reduction dependent on the path length through the atmosphere that the solar radiation has to accomplish.

This phenomena affects the portion of extraterrestrial radiation reaching the surface and is described by the clearness index k_t.

Scattering In addition to the absorption, light can be also scattered by air molecules, cloud droplets and aerosols. A fraction of the scattered beam reaches the ground surface in the form of diffuse radiation.
The amount of radiation reaching the ground depends on the thickness of the atmosphere that it had to cross. Such a thickness is determined by the elevation of the Sun. In fact, when the Sun is at the zenith, the path that the radiation has to cover through the atmosphere is minimal and maximal at the horizon.
Diffuse radiation can be defined as the short wavelength radiation coming from all

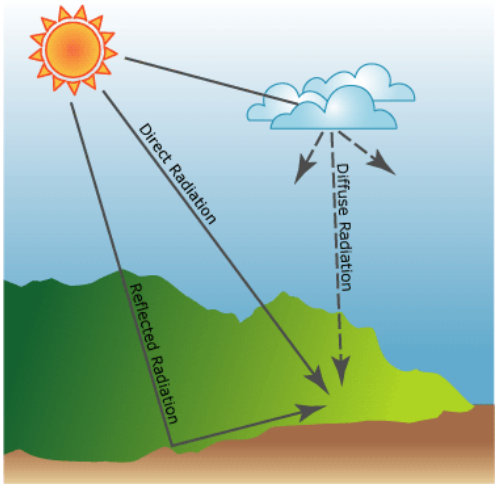


Figure B.5: Composition of the solar radiation reaching the terrestrial surfaceArcGis [2010]

parts of the sky [Liu and Jordan, 1960]. It is reflected and scattered and includes multiple reflections between the atmosphere and the surface. Clouds are one of the principal factors responsible for diffuse radiation.
There are two main types of scattering (Figure B.6b): the Rayleigh scattering and the Mie scattering. The first one is elastic scattering of light by particles, such as atoms and molecules, whose diameter d is much smaller than the light wavelength λ. Its intensity I is wavelength dependent ($I \propto \lambda^{-4}$) and is equal for forward and backward scattering. The minimum of its intensity is at $90\,^{\circ}$ of its line of incidence. It explains the blue sky and the red sunsets and sunrises. On the other hand, the Mie scattering is caused by particles, such as dust, pollen and smoke, whose diameter has the same magnitude as the incident light radiation. It often occurs in the lower portion of the atmosphere, where larger particles are more abundant, and is mainly in the forward direction. In a clean and dry atmosphere, half of the radiation is scattered back to the space and the other half reaches the surface as diffuse radiation. With more dust in the atmosphere, more scattered radiation reaches the surface because of a forward direction scattering. For this reason, there is a big portion that appears to be from a small annular area around the solar disk: circumsolar radiation. The diffuse radiation generated by the primary scattering is dominant. Ozone absorbs in the ultraviolet, visible and infrared light. The attenuation is very strong in the far-ultraviolet

(a) Mie (b) Rayleigh

Figure B.6: Mie and Rayleigh scattering Willis [2010]

region [Iqbal, 1983]. The amount of radiation reaching the surface without any scattering or reflection is defined as **direct radiation**. The **global radiation** is the sum of the two components, diffuse and direct.

The diffuse fraction k_d is affected by the scattering, as it represents the fraction of diffuse radiation compared to the global radiation reaching a location.

Appendix C

Look-up tables for Perez model

1	0.385230	0.385230	0.385230	0.462880	0.317440
2	0.338390	0.338390	0.221270	0.316730	0.503650
3	0.235680	0.235680	0.241280	0.157830	0.269440
4	0.830130	0.830130	0.171970	0.841070	0.457370
5	0.548010	0.548010	0.478000	0.966880	1.036370
6	0.548010	0.548010	1.000000	3.012370	1.976540
7	0.582690	0.582690	0.229720	0.892710	0.569950
1	0.131280	0.131280	0.385460	0.511070	0.127940
2	0.223710	0.223710	0.193560	0.304560	0.193940
3	0.229970	0.229970	0.275020	0.312730	0.244610
4	0.090100	0.184580	0.260500	0.687480	0.579440
5	0.131530	0.131530	0.370190	1.380350	1.052270
6	1.116250	1.116250	0.928030	3.525490	2.316920
7	0.090100	0.237000	0.300040	0.812470	0.664970
1	0.587510	0.130000	0.400000	0.537210	0.832490
2	0.306210	0.129830	0.204460	0.500000	0.681640
3	0.224020	0.260620	0.334080	0.501040	0.350470
4	0.421540	0.753970	0.750660	3.706840	0.983790
5	0.706680	0.373530	1.245670	0.864860	1.992630
6	4.864400	0.117390	0.265180	0.359180	3.310820
7	0.392080	0.493290	0.651560	1.932780	0.898730
1	0.126970	0.126970	0.126970	0.126970	0.126970
2	0.810820	0.810820	0.810820	0.810820	0.810820
3	3.241680	2.500000	2.291440	2.291440	2.291440
4	4.000000	3.000000	2.000000	0.975430	1.965570
5	12.494170	12.494170	8.000000	5.083520	8.792390
6	21.744240	21.744240	21.744240	21.744240	21.744240
7	3.241680	12.494170	1.620760	1.375250	2.331620

Continued on Next Page...

1	0.126970	0.126970	0.126970	0.126970	0.126970
2	0.810820	0.810820	0.810820	0.810820	0.810820
3	3.241680	2.500000	2.291440	2.291440	2.291440
4	4.000000	3.000000	2.000000	0.975430	1.965570
5	12.494170	12.494170	8.000000	5.083520	8.792390
6	21.744240	21.744240	21.744240	21.744240	21.744240
7	3.241680	12.494170	1.620760	1.375250	2.331620
1	0.126970	0.126970	0.126970	0.126970	0.126970
2	0.810820	0.810820	0.810820	0.810820	0.810820
3	3.241680	2.500000	2.291440	2.291440	2.291440
4	4.000000	3.000000	2.000000	0.975430	1.965570
5	12.494170	12.494170	8.000000	5.083520	8.792390
6	21.744240	21.744240	21.744240	21.744240	21.744240
7	3.241680	12.494170	1.620760	1.375250	2.331620
1	0.337440	0.337440	0.969110	1.097190	1.116080
2	0.337440	0.337440	0.969110	1.116030	0.623900
3	0.337440	0.337440	1.530590	1.024420	0.908480
4	0.584040	0.584040	0.847250	0.914940	1.289300
5	0.337440	0.337440	0.310240	1.435020	1.852830
6	0.337440	0.337440	1.015010	1.097190	2.117230
7	0.337440	0.337440	0.969110	1.145730	1.476400
1	0.300000	0.300000	0.700000	1.100000	0.796940
2	0.219870	0.219870	0.526530	0.809610	0.649300
3	0.386650	0.386650	0.119320	0.576120	0.685460
4	0.746730	0.399830	0.470970	0.986530	0.785370
5	0.575420	0.936700	1.649200	1.495840	1.335590
6	1.319670	4.002570	1.276390	2.644550	2.518670
7	0.665190	0.678910	1.012360	1.199940	0.986580
1	0.378870	0.974060	0.500000	0.491880	0.665290
2	0.105210	0.263470	0.407040	0.553460	0.582590
3	0.312900	0.345240	1.144180	0.854790	0.612280
4	0.119070	0.365120	0.560520	0.793720	0.802600
5	0.781610	0.837390	1.270420	1.537980	1.292950
6	1.152290	1.152290	1.492080	1.245370	2.177100

Continued on Next Page...

7	0.424660	0.529550	0.966910	1.033460	0.958730
1	0.310590	0.714410	0.252450	0.500000	0.607600
2	0.975190	0.363420	0.500000	0.400000	0.502800
3	0.175580	0.196250	0.476360	1.072470	0.490510
4	0.719280	0.698620	0.657770	1.190840	0.681110
5	0.426240	1.464840	0.678550	1.157730	0.978430
6	2.501120	1.789130	1.387090	2.394180	2.394180
7	0.491640	0.677610	0.685610	1.082400	0.735410
1	0.597000	0.500000	0.300000	0.310050	0.413510
2	0.314790	0.336310	0.400000	0.400000	0.442460
3	0.166510	0.460440	0.552570	1.000000	0.461610
4	0.401020	0.559110	0.403630	1.016710	0.671490
5	0.400360	0.750830	0.842640	1.802600	1.023830
6	3.315300	1.510380	2.443650	1.638820	2.133990
7	0.530790	0.745850	0.693050	1.458040	0.804500
1	0.597000	0.500000	0.300000	0.310050	0.800920
2	0.314790	0.336310	0.400000	0.400000	0.237040
3	0.166510	0.460440	0.552570	1.000000	0.581990
4	0.401020	0.559110	0.403630	1.016710	0.898570
5	0.400360	0.750830	0.842640	1.802600	3.400390
6	3.315300	1.510380	2.443650	1.638820	2.508780
7	0.204340	1.157740	2.003080	2.622080	1.409380
1	1.242210	1.242210	1.242210	1.242210	1.242210
2	0.056980	0.056980	0.656990	0.656990	0.925160
3	0.089090	0.089090	1.040430	1.232480	1.205300
4	1.053850	1.053850	1.399690	1.084640	1.233340
5	1.151540	1.151540	1.118290	1.531640	1.411840
6	1.494980	1.494980	1.700000	1.800810	1.671600
7	1.018450	1.018450	1.153600	1.321890	1.294670
1	0.700000	0.700000	1.023460	0.700000	0.945830
2	0.886300	0.886300	1.333620	0.800000	1.066620
3	0.902180	0.902180	0.954330	1.126690	1.097310
4	1.095300	1.075060	1.176490	1.139470	1.096110
5	1.201660	1.201660	1.438200	1.256280	1.198060

Continued on Next Page...

6	1.525850	1.525850	1.869160	1.985410	1.911590
7	1.288220	1.082810	1.286370	1.166170	1.119330
1	0.600000	1.029910	0.859890	0.550000	0.813600
2	0.604450	1.029910	0.859890	0.656700	0.928840
3	0.455850	0.750580	0.804930	0.823000	0.911000
4	0.526580	0.932310	0.908620	0.983520	0.988090
5	1.036110	1.100690	0.848380	1.035270	1.042380
6	1.048440	1.652720	0.900000	2.350410	1.082950
7	0.817410	0.976160	0.861300	0.974780	1.004580
1	0.782110	0.564280	0.600000	0.600000	0.665740
2	0.894480	0.680730	0.541990	0.800000	0.669140
3	0.487460	0.818950	0.841830	0.872540	0.709040
4	0.709310	0.872780	0.908480	0.953290	0.844350
5	0.863920	0.947770	0.876220	1.078750	0.936910
6	1.280350	0.866720	0.769790	1.078750	0.975130
7	0.725420	0.869970	0.868810	0.951190	0.829220
1	0.791750	0.654040	0.483170	0.409000	0.597180
2	0.566140	0.948990	0.971820	0.653570	0.718550
3	0.648710	0.637730	0.870510	0.860600	0.694300
4	0.637630	0.767610	0.925670	0.990310	0.847670
5	0.736380	0.946060	1.117590	1.029340	0.947020
6	1.180970	0.850000	1.050000	0.950000	0.888580
7	0.700560	0.801440	0.961970	0.906140	0.823880
1	0.500000	0.500000	0.586770	0.470550	0.629790
2	0.500000	0.500000	1.056220	1.260140	0.658140
3	0.500000	0.500000	0.631830	0.842620	0.582780
4	0.554710	0.734730	0.985820	0.915640	0.898260
5	0.712510	1.205990	0.909510	1.078260	0.885610
6	1.899260	1.559710	1.000000	1.150000	1.120390
7	0.653880	0.793120	0.903320	0.944070	0.796130
1	1.000000	1.000000	1.050000	1.170380	1.178090
2	0.960580	0.960580	1.059530	1.179030	1.131690
3	0.871470	0.871470	0.995860	1.141910	1.114600
4	1.201590	1.201590	0.993610	1.109380	1.126320

Continued on Next Page...

5	1.065010	1.065010	0.828660	0.939970	1.017930
6	1.065010	1.065010	0.623690	1.119620	1.132260
7	1.071570	1.071570	0.958070	1.114130	1.127110
1	0.950000	0.973390	0.852520	1.092200	1.096590
2	0.804120	0.913870	0.980990	1.094580	1.042420
3	0.737540	0.935970	0.999940	1.056490	1.050060
4	1.032980	1.034540	0.968460	1.032080	1.015780
5	0.900000	0.977210	0.945960	1.008840	0.969960
6	0.600000	0.750000	0.750000	0.844710	0.899100
7	0.926800	0.965030	0.968520	1.044910	1.032310
1	0.850000	1.029710	0.961100	1.055670	1.009700
2	0.818530	0.960010	0.996450	1.081970	1.036470
3	0.765380	0.953500	0.948260	1.052110	1.000140
4	0.775610	0.909610	0.927800	0.987800	0.952100
5	1.000990	0.881880	0.875950	0.949100	0.893690
6	0.902370	0.875960	0.807990	0.942410	0.917920
7	0.856580	0.928270	0.946820	1.032260	0.972990
1	0.750000	0.857930	0.983800	1.056540	0.980240
2	0.750000	0.987010	1.013730	1.133780	1.038250
3	0.800000	0.947380	1.012380	1.091270	0.999840
4	0.800000	0.914550	0.908570	0.999190	0.915230
5	0.778540	0.800590	0.799070	0.902180	0.851560
6	0.680190	0.317410	0.507680	0.388910	0.646710
7	0.794920	0.912780	0.960830	1.057110	0.947950
1	0.750000	0.833890	0.867530	1.059890	0.932840
2	0.979700	0.971470	0.995510	1.068490	1.030150
3	0.858850	0.987920	1.043220	1.108700	1.044900
4	0.802400	0.955110	0.911660	1.045070	0.944470
5	0.884890	0.766210	0.885390	0.859070	0.818190
6	0.615680	0.700000	0.850000	0.624620	0.669300
7	0.835570	0.946150	0.977090	1.049350	0.979970
1	0.689220	0.809600	0.900000	0.789500	0.853990
2	0.854660	0.852840	0.938200	0.923110	0.955010
3	0.938600	0.932980	1.010390	1.043950	1.041640

Continued on Next Page...

4	0.843620	0.981300	0.951590	0.946100	0.966330
5	0.694740	0.814690	0.572650	0.400000	0.726830
6	0.211370	0.671780	0.416340	0.297290	0.498050
7	0.843540	0.882330	0.911760	0.898420	0.960210
1	1.054880	1.075210	1.068460	1.153370	1.069220
2	1.000000	1.062220	1.013470	1.088170	1.046200
3	0.885090	0.993530	0.942590	1.054990	1.012740
4	0.920000	0.950000	0.978720	1.020280	0.984440
5	0.850000	0.908500	0.839940	0.985570	0.962180
6	0.800000	0.800000	0.810080	0.950000	0.961550
7	1.038590	1.063200	1.034440	1.112780	1.037800
1	1.017610	1.028360	1.058960	1.133180	1.045620
2	0.920000	0.998970	1.033590	1.089030	1.022060
3	0.912370	0.949930	0.979770	1.020420	0.981770
4	0.847160	0.935300	0.930540	0.955050	0.946560
5	0.880260	0.867110	0.874130	0.972650	0.883420
6	0.627150	0.627150	0.700000	0.774070	0.845130
7	0.973700	1.006240	1.026190	1.071960	1.017240
1	1.028710	1.017570	1.025900	1.081790	1.024240
2	0.924980	0.985500	1.014100	1.092210	0.999610
3	0.828570	0.934920	0.994950	1.024590	0.949710
4	0.900810	0.901330	0.928830	0.979570	0.913100
5	0.761030	0.845150	0.805360	0.936790	0.853460
6	0.626400	0.546750	0.730500	0.850000	0.689050
7	0.957630	0.985480	0.991790	1.050220	0.987900
1	0.992730	0.993880	1.017150	1.059120	1.017450
2	0.975610	0.987160	1.026820	1.075440	1.007250
3	0.871090	0.933190	0.974690	0.979840	0.952730
4	0.828750	0.868090	0.834920	0.905510	0.871530
5	0.781540	0.782470	0.767910	0.764140	0.795890
6	0.743460	0.693390	0.514870	0.630150	0.715660
7	0.934760	0.957870	0.959640	0.972510	0.981640
1	0.965840	0.941240	0.987100	1.022540	1.011160
2	0.988630	0.994770	0.976590	0.950000	1.034840

Continued on Next Page...

3	0.958200	1.018080	0.974480	0.920000	0.989870
4	0.811720	0.869090	0.812020	0.850000	0.821050
5	0.682030	0.679480	0.632450	0.746580	0.738550
6	0.668290	0.445860	0.500000	0.678920	0.696510
7	0.926940	0.953350	0.959050	0.876210	0.991490
1	0.948940	0.997760	0.850000	0.826520	0.998470
2	1.017860	0.970000	0.850000	0.700000	0.988560
3	1.000000	0.950000	0.850000	0.606240	0.947260
4	1.000000	0.746140	0.751740	0.598390	0.725230
5	0.922210	0.500000	0.376800	0.517110	0.548630
6	0.500000	0.450000	0.429970	0.404490	0.539940
7	0.960430	0.881630	0.775640	0.596350	0.937680
1	1.030000	1.040000	1.000000	1.000000	1.049510
2	1.050000	0.990000	0.990000	0.950000	0.996530
3	1.050000	0.990000	0.990000	0.820000	0.971940
4	1.050000	0.790000	0.880000	0.820000	0.951840
5	1.000000	0.530000	0.440000	0.710000	0.928730
6	0.540000	0.470000	0.500000	0.550000	0.773950
7	1.038270	0.920180	0.910930	0.821140	1.034560
1	1.041020	0.997520	0.961600	1.000000	1.035780
2	0.948030	0.980000	0.900000	0.950360	0.977460
3	0.950000	0.977250	0.869270	0.800000	0.951680
4	0.951870	0.850000	0.748770	0.700000	0.883850
5	0.900000	0.823190	0.727450	0.600000	0.839870
6	0.850000	0.805020	0.692310	0.500000	0.788410
7	1.010090	0.895270	0.773030	0.816280	1.011680
1	1.022450	1.004600	0.983650	1.000000	1.032940
2	0.943960	0.999240	0.983920	0.905990	0.978150
3	0.936240	0.946480	0.850000	0.850000	0.930320
4	0.816420	0.885000	0.644950	0.817650	0.865310
5	0.742960	0.765690	0.561520	0.700000	0.827140
6	0.643870	0.596710	0.474460	0.600000	0.651200
7	0.971740	0.940560	0.714880	0.864380	1.001650
1	0.995260	0.977010	1.000000	1.000000	1.035250

Continued on Next Page. . .

2	0.939810	0.975250	0.939980	0.950000	0.982550
3	0.876870	0.879440	0.850000	0.900000	0.917810
4	0.873480	0.873450	0.751470	0.850000	0.863040
5	0.761470	0.702360	0.638770	0.750000	0.783120
6	0.734080	0.650000	0.600000	0.650000	0.715660
7	0.942160	0.919100	0.770340	0.731170	0.995180
1	0.952560	0.916780	0.920000	0.900000	1.005880
2	0.928620	0.994420	0.900000	0.900000	0.983720
3	0.913070	0.850000	0.850000	0.800000	0.924280
4	0.868090	0.807170	0.823550	0.600000	0.844520
5	0.769570	0.719870	0.650000	0.550000	0.733500
6	0.580250	0.650000	0.600000	0.500000	0.628850
7	0.904770	0.852650	0.708370	0.493730	0.949030
1	0.911970	0.800000	0.800000	0.800000	0.956320
2	0.912620	0.682610	0.750000	0.700000	0.950110
3	0.653450	0.659330	0.700000	0.600000	0.856110
4	0.648440	0.600000	0.641120	0.500000	0.695780
5	0.570000	0.550000	0.598800	0.400000	0.560150
6	0.475230	0.500000	0.518640	0.339970	0.520230
7	0.743440	0.592190	0.603060	0.316930	0.794390

Table C.1: Correction matrix $\chi(k_t', \theta_z, W, \Delta k_t')$.

Appendix D

Statistical methods

Different statistical instruments are available in the literature. In this study, the Mean Bias Error and the Mean Absolute Error are used to quantify the performance of the model described in Chapter 3. Furthermore, in order to have a qualitative representation of the model performance, two further statistical instruments are used: the coefficient of determination R^2 and the coefficient of efficiency E.

In this chapter, a brief description of the above mentioned statistical instruments is illustrated.

Mean Bias Error

There are different instruments that can quantify the error between the observed and the measured values. The first is the Mean Bias Error (MBE) is the average deviation of the estimated values from the measured values. It is defined as:

$$MBE = \frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)}{N} \tag{D.1}$$

where $\hat{y}_i$ is the i_{th} predicted value, y_i the i_{th} observed value, and N the number of observations.
A model would perfectly fit the observed data if the MBE is equal to zero, i.e. the sum of pairwise identical negative and positive deviations would also be zero [Iqbal, 1983].

Mean Absolute Error

An additional instrument to quantify the errors is the Mean Absolute Error (MAE), which is defined as follows:

$$MAE = \frac{\sum_{i=1}^{N}|\hat{y}_i - y_i|}{N} \tag{D.2}$$

Similarly to MBE, it describes the difference between the modeled data and the observations in the units of the variables. Furthermore, MAE is preferred over the Root Mean Square Error [Legates and McCabe, 1999].

Coefficient of determination

The linear relationship between the estimated and the measured values is tested by the correlation coefficient r. This statistical measure verifies its linear correlation and is defined

as follows:

$$r = \frac{\sum_{i=1}^{N}(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\left\{\left[\sum_{i=1}^{N}(\hat{y}_i - \bar{\hat{y}})^2\right]\left[\sum_{i=1}^{N}(y_i - \bar{y})^2\right]\right\}^{0.5}} \tag{D.3}$$

where $\bar{\hat{y}}$ is the predicted mean value and $\bar{y}$ the measured mean value.
Ideally, r should correspond to 1 or -1 Iqbal [1983].
The numerator of Equation D.3 is a measure how the observed and modeled values varie together using their deviation from the mean. The denominator measures how the measured and the estimated values vary separately.
With r, the coefficient of determination R^2 describes the degree of collinearity between the observed and the computed variates. R^2 is the proportion of the total variance in the observed data that can be explained by a model. It ranges from 0 to 1 and is defined as follows:

$$R^2 := r^2 = \left\{ \frac{\sum_{i=1}^{N}(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\left\{\left[\sum_{i=1}^{N}(\hat{y}_i - \bar{\hat{y}})^2\right]\left[\sum_{i=1}^{N}(y_i - \bar{y})^2\right]\right\}^{0.5}} \right\}^2 \tag{D.4}$$

It is the ratio between the unnormalized covariance and the unnormalized variances. The coefficient of determination explains how well the observed data are reproduced by the modeled data. In other words, it represents the percentage of the explained to the total of variations.
It only evaluates linear relationships between the variables and is insensitive to additive and proportional differences between model and observations. This type of measure is more sensitive to outliers than observations near the mean and leads to a bias toward the extreme events. In addition, its statistical distribution is well known [Legates and McCabe, 1999].
However, R^2 and r both have limitations. Moreover, the coefficient of determination does not account for differences in the mean of the two time series and assumes that the observed variance is less than the model-predicted variance [Legates and McCabe, 1999]. It is demonstrated MathWorld [2010] that the coefficient of determination is the portion of variance of the modeled values which is accounted by the 1:1 line.

Coefficient of efficiency

In 2003, three temperature-index based models and of one energy balance based snowmelt model were investigated by Zappa et al. [2003]. For this purpose, the efficiency coefficient E_2 was applied, which is a generic version, based on Legates and McCabe [1999] and Nash and Suttcliffe [1970] studies:

$$E := E_1 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{D.5}$$

where E is the ratio of the mean square error

$$MSE = \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N} \tag{D.6}$$

to the variance in the observed data, subtracted from unity.
It ranges from $-\infty$ to 1. A zero value reflects a good estimation by the model. Whereas, negative values signifies a poor performance of the model.
E is sensitive to differences in the observed and model simulated means and variances Legates and McCabe [1999].

Nash and Sutcliffe introduce a third statistical instrument, the index of agreement:

$$d = 1 - \frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{N} (|\hat{y}_i - \bar{y}| + |y_i - \bar{y}|)^2} = 1 - N\frac{MSE}{PE} \tag{D.7}$$

where PE is the potential error

$$PE = \sum_{i=1}^{N} (|\hat{y}_i - \bar{y}| + |y_i - \bar{y}|)^2 . \tag{D.8}$$

It ranges from 0 to 1, which is for the best fit and is sensitive to extreme values.
Furthermore, E and d are very similar methods. They both are sensible to outliers.
However, in my opinion, d is not sensible to the variation in the performance of the models.
For this reason, only the coefficient of efficiency E is considered as statistical measure in this study.

Residuals

A complementary tool to the model investigation is the analysis of the residuals r:

$$e = y - \hat{y}, \tag{D.9}$$

which is the difference between the estimated and the observed value, $\hat{y}$ and y, respectively.
At first, the plot of the standardized residuals against the standardized estimated diffuse fraction is formed. An standardized variable is defined as follows:

$$x_{std} = \frac{x - \bar{x}}{sd(x)} \tag{D.10}$$

where x is the unstandardized value, $\bar{x}$ its mean and $sd(x)$ the standard deviation.
The plotted values generally are close to the zero horizontal line, which means that the points are well predicted. Values above are underpredicted and values below are overpredicted. The linearity assumption is confirmed whether the same number of points is present above and below the horizontal line.
Then a histogram of the standardized residuals can be plotted and a curve of the normal distribution can be overlaid. It permits to test the normal distribution of the residuals.
A QQ-Plot ("'Q"' stands for quantile) is then used to reinforce the assumption of a normal distribution of the residuals. They are normal distributed if the points lay on the line.

Appendix E

Statistical results

... Continued on Next Page

Station	R^2	E	MAE	MBE	P_d	N_d
Robbia	0.7384671	0.7181619	0.12797762	0.04443873	30.496151	2338
Moleson	0.6981069	0.6652712	0.13725952	0.06791140	30.875456	3838
Montana	0.8111276	0.7413688	0.11187587	0.07869043	31.312018	3628
Chasseral	0.7598021	0.6600968	0.13118180	0.09341299	32.940830	3397
Pilatus	0.7097518	0.6876686	0.12918435	0.07024258	45.562481	3369
Altdorf	0.7531478	0.7404422	0.12105326	0.04928409	42.129246	2414
Geneva	0.8437190	0.8390440	0.08922130	0.05269562	45.692389	3784
Visp	0.6984615	0.6702561	0.13885696	0.06395248	32.223580	3274
Buchs	0.6816737	0.6619043	0.13675572	0.06278430	40.138010	3478
Goesgen	0.6629828	0.6438390	0.14329333	0.06605141	37.422037	3367
Comprovasco	0.7734217	0.7538922	0.11416177	0.04232742	42.894249	2695
Hoernli	0.7311356	0.7115850	0.12587295	0.06632387	42.082294	3208
Montana	0.7588199	0.7453723	0.11908180	0.05533643	38.272120	2396
Alice Springs	0.8866533	0.8285907	0.08532740	0.02770627	16.726160	3922
Barrow	0.7027471	0.5089345	0.15522106	0.12485502	38.645418	3514
Bermuda	0.8039380	0.7200999	0.10393480	0.07300610	34.755625	3867
Cabauw	0.8850711	0.8688919	0.07143916	0.04867081	55.313634	3858
Carpentras	0.8995191	0.8574744	0.08807254	0.03899609	35.828450	3754
Chesapeake Light	0.9179134	0.9018380	0.06786469	0.04223711	44.218415	3736
Coco Island	0.8349969	0.7906055	0.08207378	0.05718751	39.323698	3667
Dome C	0.5653018	-2.6281581	0.11898770	0.09185393	21.090132	2807
Desert Rock	0.8169632	0.7263950	0.08405736	0.04147044	15.814777	3952
Goodwin Creek	0.8853985	0.8316392	0.09160615	0.04827821	33.817804	3853
Georg von Neumayer	0.7485279	-0.1731673	0.23317960	0.20953988	15.601266	3160
Momote	0.8287951	0.7973876	0.09462916	0.04591585	42.386831	4131
Ny-Alesund	0.8006920	0.6641574	0.13097578	0.09395672	42.437086	3775
Tamanrasset	0.8032452	0.6114248	0.10761813	0.08151543	20.320041	3937
Tateno	0.9146441	0.8998668	0.06412966	0.04608334	54.927501	4069
Toravere	0.9047302	0.8776950	0.07155062	0.02471932	56.541678	3707
Payerne	0.8778039	0.8365510	0.08790467	0.05782691	49.459288	3144
Brasilia	0.8688476	0.8294349	0.09461297	0.03567818	31.982889	3974
De Aar	0.9253562	0.8831964	0.08424084	0.02867434	26.710684	3332
Darwin	0.8409219	0.7624125	0.10396678	0.03799748	29.564795	4113
Xianghe	0.8916573	0.7121721	0.11349129	0.08711956	31.774761	3978
Solar Village	0.8407930	0.7760512	0.07412578	0.06001529	28.543355	4071
Ilorin	0.5879029	0.2944510	0.11496362	0.06699045	38.439761	3179
Izana	0.8314110	0.7230235	0.08531861	0.04424831	14.968999	3387
Lauder	0.8735181	0.8416139	0.09835210	0.04892695	35.365226	3888
Summit Station	0.6861280	-0.6820207	0.24855270	0.20784162	7.928268	2119

Table E.1: Statistical results for the Reindl-Helbig model.

Station	R^2	E	MAE	MBE	P_d	N_d
Robbia	0.7479716	0.6054957	0.14062677	0.03293994	29.426861	2338
Moleson	0.7112182	0.6362427	0.13486112	0.05750162	31.605003	3838
Montana	0.8527031	0.8120257	0.09021507	0.06423317	37.761852	3628
Chasseral	0.7942031	0.7248408	0.11155334	0.07684805	38.151310	3397
Pilatus	0.7159098	0.6322808	0.12480369	0.05587849	43.128525	3369
Altdorf	0.7773555	0.6988003	0.11801262	0.03720137	43.661972	2414
Geneva	0.8882122	0.8872594	0.07649497	0.04422209	51.083510	3784
Visp	0.7410529	0.6409384	0.12791166	0.04899446	34.208919	3274
Buchs	0.7099893	0.6280870	0.12665781	0.04968360	43.070730	3478
Goesgen	0.7127231	0.6414368	0.12340326	0.05102440	41.283041	3367
Comprovasco	0.7781875	0.6413535	0.12891860	0.03167270	40.853432	2695
Hoernli	0.7488947	0.7119642	0.11021061	0.05339457	45.698254	3208
Montana	0.7711007	0.6929969	0.12110109	0.04273816	37.186978	2396
Alice Springs	0.8641489	0.7946755	0.08780989	0.02291353	19.811321	3922
Barrow	0.6909929	0.5391603	0.13814826	0.11615954	44.052362	3514
Bermuda	0.8496959	0.7908202	0.09131790	0.06508211	38.401862	3867
Cabauw	0.9006764	0.8919363	0.06400712	0.04246298	57.024365	3858
Carpentras	0.9155086	0.8914894	0.07307125	0.03521344	40.037294	3754
Chesapeake Light	0.9202600	0.9101787	0.06337352	0.03773306	46.654176	3736
Coco Island	0.8592840	0.8282531	0.07753942	0.05185099	39.432779	3667
Dome C	0.2176861	-2.4588793	0.15980288	0.06979002	4.061275	2807
Desert Rock	0.8294056	0.7840923	0.07345677	0.03611684	26.644737	3952
Goodwin Creek	0.8996763	0.8649314	0.08208318	0.04388682	39.242149	3853
Georg von Neumayer	0.6681572	-0.3970748	0.21723094	0.19414642	14.683544	3160
Momote	0.8624036	0.8336970	0.08417065	0.03996155	45.073832	4131
Ny-Alesund	0.8111389	0.7120026	0.10539113	0.08406667	49.801325	3775
Tamanrasset	0.7897895	0.6427919	0.11221113	0.07612661	16.916434	3937
Tateno	0.9126942	0.9032925	0.06268706	0.04173550	55.763087	4069
Toravere	0.9131026	0.8815805	0.07229306	0.02189618	57.243054	3707
Payerne	0.8971442	0.8710585	0.07597807	0.05033665	52.385496	3144
Brasilia	0.8752521	0.8220122	0.09205560	0.03025625	34.272773	3974
De Aar	0.9050819	0.8445501	0.09352727	0.02563715	26.890756	3332
Darwin	0.8866132	0.8366397	0.08561959	0.03191094	33.041575	4113
Xianghe	0.9016999	0.7659253	0.10905396	0.08153797	36.902966	3978
Solar Village	0.8457477	0.7597211	0.09025462	0.06366843	19.405551	4071
Ilorin	0.5981215	0.3077692	0.11387525	0.05542284	40.578798	3179
Izana	0.7315111	0.6044272	0.11061052	0.03444699	12.016534	3387
Lauder	0.8773286	0.8159544	0.09866153	0.03864289	34.696502	3888
Summit Station	0.5187612	-1.2902064	0.23043027	0.20143590	9.910335	2119

Table E.2: Statistical results for the Skartveit and Olseth model.

Station	R^2	E	MAE	MBE	P_d	N_d
Robbia	0.5339677	0.7430305	0.09839203	0.03521911	21.086399	2338
Moleson	0.6812201	0.6890893	0.12663199	0.05630590	31.005732	3838
Montana	0.7659391	0.7552519	0.10380693	0.06911121	28.335171	3628
Chasseral	0.7096287	0.6904590	0.11543628	0.07325546	31.439505	3397
Pilatus	0.6802588	0.7295095	0.11544949	0.05019274	40.397744	3369
Altdorf	0.6485310	0.7452673	0.10813795	0.03812408	35.666943	2414
Geneva	0.8269857	0.8602988	0.07582718	0.03819199	47.912262	3784
Visp	0.5643906	0.6497179	0.13780290	0.05633696	22.266341	3274
Buchs	0.6034045	0.6467216	0.13253586	0.05180542	34.215066	3478
Goesgen	0.5927138	0.6308204	0.13851719	0.05417101	30.501931	3367
Comprovasco	0.6707144	0.7620895	0.10455593	0.03317614	34.805195	2695
Hoernli	0.6599088	0.7045343	0.12067620	0.05318091	36.471322	3208
Montana	0.6327508	0.7525408	0.10735937	0.04614240	29.799666	2396
Alice Springs	0.8810923	0.8566982	0.07197961	0.02994504	21.111678	3922
Barrow	0.6601882	0.4441032	0.16863934	0.13180445	38.616961	3514
Bermuda	0.8077770	0.7617100	0.09618921	0.05984183	35.479700	3867
Cabauw	0.8676571	0.8642308	0.07042890	0.03326044	59.020218	3858
Carpentras	0.8991856	0.8735414	0.08242041	0.03327231	37.426745	3754
Chesapeake Light	0.9038189	0.8985416	0.06958235	0.03811661	45.074946	3736
Coco Island	0.8470771	0.8194021	0.07813074	0.05515308	41.150804	3667
Dome C	0.5235350	-2.7406773	0.18478684	0.10095282	2.137513	2807
Desert Rock	0.8094938	0.7010376	0.08013175	0.05030165	21.761134	3952
Goodwin Creek	0.8167309	0.8131966	0.09351742	0.04542716	29.120166	3853
Georg von Neumayer	0.7543687	0.1033020	0.23719585	0.19002296	6.645570	3160
Momote	0.8537867	0.8336868	0.08601463	0.03785070	45.775841	4131
Ny-Alesund	0.8061093	0.7292549	0.11712974	0.06822810	46.649007	3775
Tamanrasset	0.8105975	0.5635298	0.11879180	0.08812503	17.145034	3937
Tateno	0.9168311	0.9089322	0.06191567	0.03538576	58.294421	4069
Toravere	0.8665628	0.7961298	0.08436168	0.01633279	54.599407	3707
Payerne	0.8702683	0.8481935	0.08478929	0.04433897	50.667939	3144
Brasilia	0.8815018	0.8597843	0.08413417	0.03157646	35.656769	3974
De Aar	0.9302499	0.9130804	0.07188879	0.02499516	29.921969	3332
Darwin	0.8754014	0.8358246	0.08755820	0.03309197	33.843910	4113
Xianghe	0.9067003	0.8011198	0.09894410	0.06651191	33.634992	3978
Solar Village	0.8393147	0.7325519	0.08763783	0.06665254	22.058462	4071
Ilorin	0.5219174	0.1434545	0.11084028	0.04923917	36.017616	3179
Izana	0.8446238	0.7256009	0.07762583	0.04791371	23.176853	3387
Lauder	0.8671643	0.8401809	0.09910976	0.04309259	35.468107	3888
Summit Station	0.6576514	-1.1789869	0.29101653	0.22392907	1.226415	2120

Table E.3: Statistical results for the BRL model.

Station	R^2	E	MAE	MBE	P_d	N_d
Robbia	0.67783881	0.61721872	0.14916194	0.04459105	29.170231	2338
Moleson	0.69988614	0.67305713	0.13580210	0.07223995	31.344450	3838
Montana	0.79431383	0.71008850	0.11869370	0.08720270	32.414553	3628
Chasseral	0.74954430	0.64211148	0.13726130	0.10128709	34.088902	3397
Pilatus	0.60246518	0.54721152	0.14431000	0.07448434	43.158207	3369
Altdorf	0.72258329	0.69386699	0.12790283	0.04717073	42.750621	2414
Geneva	0.84292055	0.83118235	0.09795993	0.06158284	44.476744	3784
Visp	0.63104129	0.56694885	0.14934858	0.06640408	33.414783	3274
Buchs	0.61794373	0.57058849	0.14621858	0.06641948	41.058079	3478
Goesgen	0.61667645	0.57657061	0.14680353	0.06643992	38.996139	3367
Comprovasco	0.74173242	0.69591287	0.12690119	0.04276991	41.855288	2695
Hoernli	0.65666698	0.63272226	0.13214807	0.06678827	43.827930	3208
Montana	0.71445828	0.68487922	0.13039181	0.05325184	37.437396	2396
Alice Springs	0.88314297	0.85986882	0.06957296	0.02994383	24.426313	3922
Barrow	0.65344772	0.42630954	0.16893211	0.14460477	38.133182	3514
Bermuda	0.77906687	0.69442673	0.11409002	0.08406730	33.514352	3867
Cabauw	0.88236823	0.85265203	0.08156009	0.05950569	53.032659	3858
Carpentras	0.90207695	0.88530927	0.07619352	0.04842767	40.197123	3754
Chesapeake Light	0.91010602	0.89842543	0.06980593	0.04957289	43.629550	3736
Coco Island	0.78632926	0.74832378	0.09470237	0.06365921	34.687756	3667
Dome C	0.02475112	-4.96448138	0.19059465	0.07903035	4.488778	2807
Desert Rock	0.83598963	0.73775714	0.07248430	0.05030667	27.277328	3952
Goodwin Creek	0.88167702	0.84260997	0.08365465	0.05786031	43.498573	3853
Georg von Neumayer	0.59352878	-1.06626665	0.25488811	0.23613864	8.037975	3160
Momote	0.79779656	0.77309134	0.10019960	0.05009545	42.265795	4131
Ny-Alesund	0.77384047	0.61889391	0.13477258	0.11117019	42.596026	3775
Tamanrasset	0.76652234	0.50511719	0.12600225	0.09279889	16.535433	3937
Tateno	0.90927810	0.89154005	0.06812022	0.05196420	55.394446	4069
Toravere	0.90078986	0.90007445	0.07082736	0.03344563	57.971405	3707
Payerne	0.87105815	0.82376702	0.09403032	0.06825904	47.805344	3144
Brasilia	0.86351178	0.85020051	0.08480482	0.03810019	37.317564	3974
De Aar	0.90145686	0.87390448	0.07771343	0.03512430	32.202881	3332
Darwin	0.82372413	0.78733528	0.09806266	0.04298506	31.096523	4113
Xianghe	0.88796752	0.71355208	0.12328461	0.09505587	30.744093	3978
Solar Village	0.82464799	0.64969631	0.09504822	0.07550460	21.272415	4071
Ilorin	0.49141223	0.06707967	0.12717591	0.06037992	37.055678	3179
Izana	0.69047385	0.55343497	0.11231744	0.04726565	11.662238	3387
Lauder	0.86939713	0.84458473	0.09126971	0.05406260	38.143004	3888
Summit Station	0.37641493	-2.49301482	0.27466443	0.24026858	7.456347	2119

Table E.4: Statistical results for the Maxwell model.

Appendix F

Additional graphics and tables

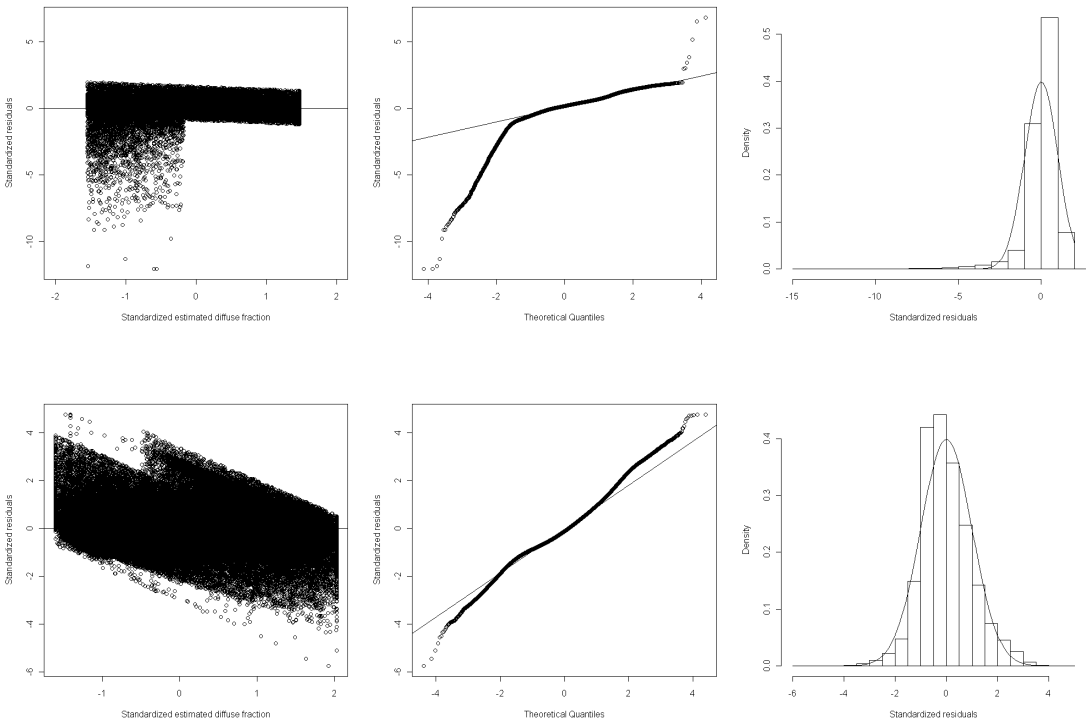


Figure F.1: Residual analysis for $k_t \leq 0.3$(top) and $0.3 < k_t \leq 0.78$ (bottom): standardized residuals plotted against the standardized estimated diffuse fraction (left) and against the theoretical quantiles (right).

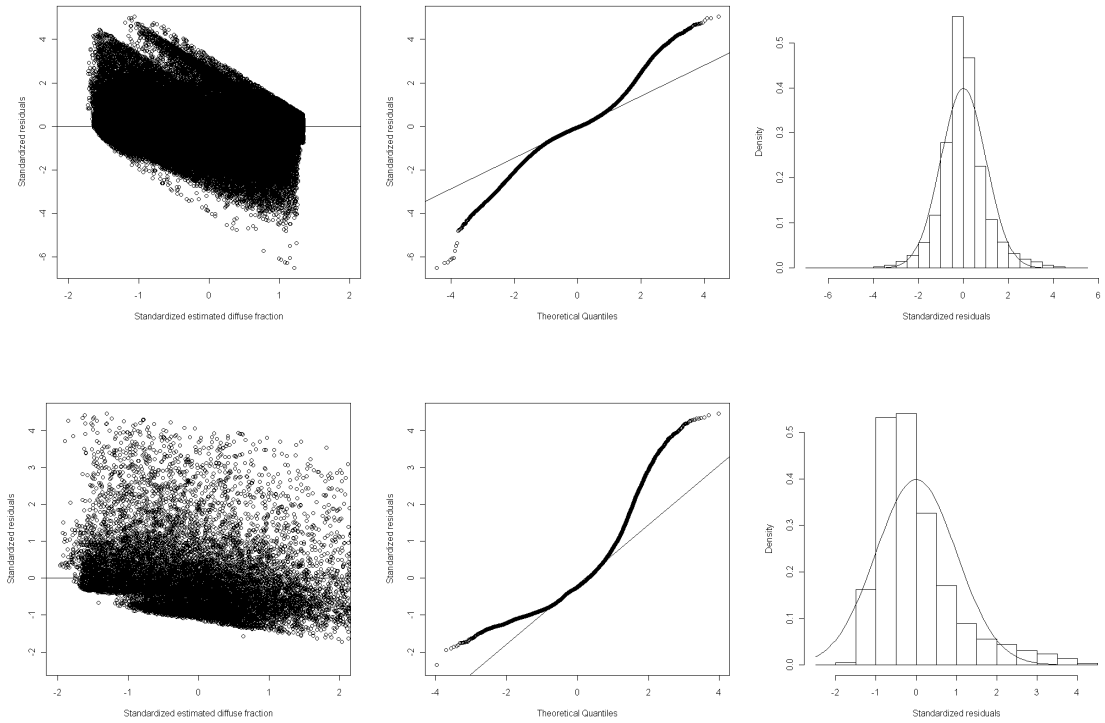


Figure F.2: Residual analysis for $k_t \leq 0.8$ (top) and $k_t > 0.8$ (bottom): standardized residuals plotted against the standardized estimated diffuse fraction from BRL (left) and against the theoretical quantiles (right).

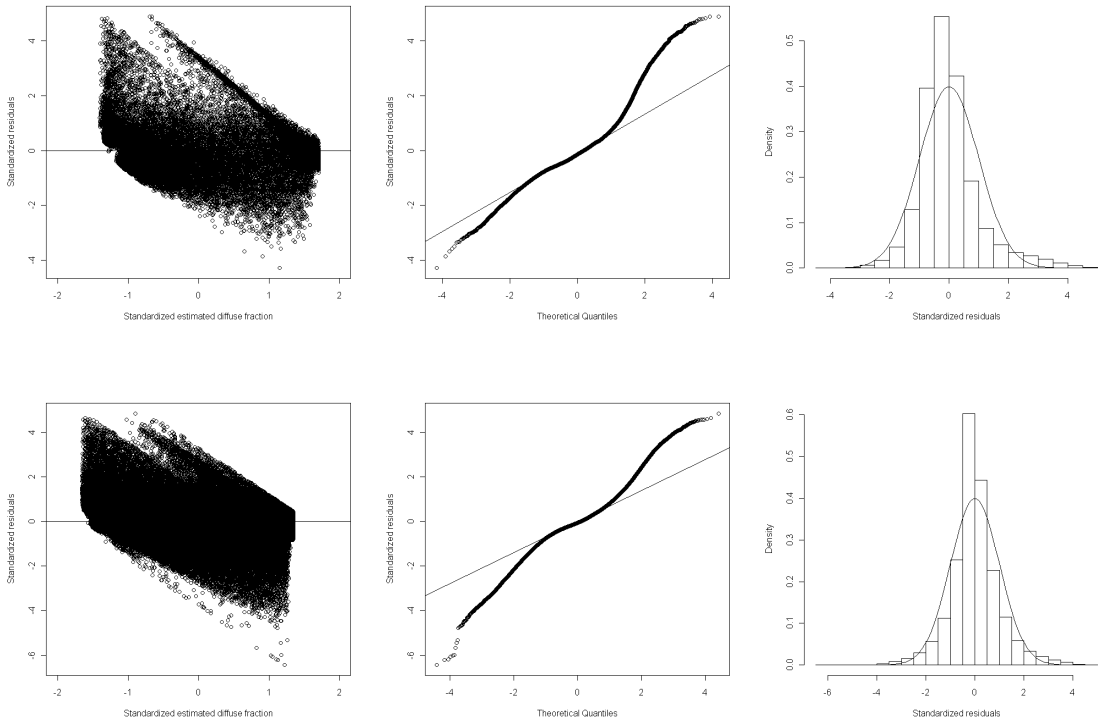


Figure F.3: Residual analysis for Southern (up) and Northern (down) Hemisphere station's data: standardized residuals plotted against the standardized estimated diffuse fraction from BRL (left) and against the theoretical quantiles (right).

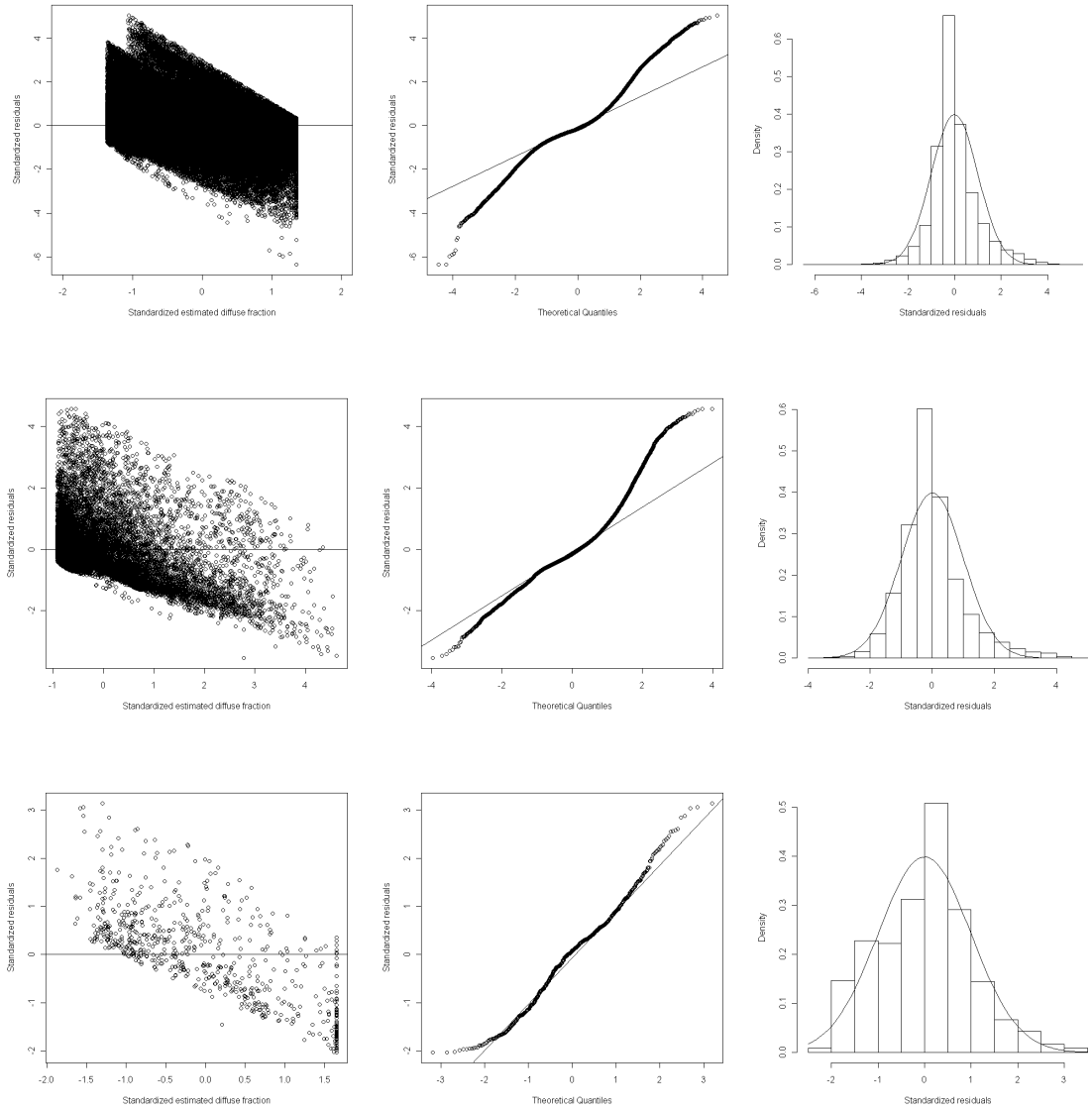


Figure F.4: Residual analysis for three clearness index intervals ($k_t \leq 0.80$, $0.80 < k_t \leq 1.00$ and $k_t > 1.00$): standardized residuals plotted against the standardized estimated diffuse fraction (left) and against the theoretical quantiles (right).

Solar elevation	R^2	E	MAE	MBE	N_d	P_d
$\phi \leq 10\,^\circ$	0.5118865	0.4861008	0.15582004	0.07617900	11415	35.74244
$10\,^\circ < \phi \leq 15\,^\circ$	0.6923237	0.6858108	0.12683692	0.06182501	12102	39.78681
$15\,^\circ < \phi \leq 20\,^\circ$	0.7264461	0.7257030	0.12435764	0.06136565	13577	38.48420
$20\,^\circ < \phi \leq 25\,^\circ$	0.7749891	0.7748224	0.11448052	0.05614469	13917	38.74398
$25\,^\circ < \phi \leq 30\,^\circ$	0.7989778	0.7988985	0.10776692	0.05289655	13660	37.38653
$30\,^\circ < \phi \leq 35\,^\circ$	0.8164619	0.8158501	0.10171927	0.04819227	12186	36.70606
$35\,^\circ < \phi \leq 40\,^\circ$	0.8334284	0.8325538	0.09519588	0.04737749	11616	35.95041
$40\,^\circ < \phi \leq 45\,^\circ$	0.8649203	0.8643109	0.08523975	0.04315256	9422	34.48312
$45\,^\circ < \phi \leq 50\,^\circ$	0.8780470	0.8752827	0.08183435	0.04567789	8889	36.50579
$50\,^\circ < \phi \leq 55\,^\circ$	0.8946735	0.8929240	0.07676437	0.04150198	7516	35.24481
$55\,^\circ < \phi \leq 60\,^\circ$	0.8934487	0.8909714	0.07712192	0.04149052	7047	37.88846
$60\,^\circ < \phi \leq 65\,^\circ$	0.8963855	0.8924510	0.07641858	0.04147725	6051	37.23352
$65\,^\circ < \phi \leq 70\,^\circ$	0.8989533	0.8948823	0.07456672	0.04147725	3648	34.12829
$70\,^\circ < \phi \leq 75\,^\circ$	0.8896879	0.8834576	0.07769438	0.04147725	2371	33.99410
$75\,^\circ < \phi \leq 80\,^\circ$	0.8922374	0.8866747	0.07445195	0.04147725	1481	32.81567
$80\,^\circ < \phi \leq 85\,^\circ$	0.8838311	0.8741087	0.08033286	0.04147725	945	29.41799
$85\,^\circ < \phi \leq 90\,^\circ$	0.8871041	0.8793356	0.07761098	0.04147725	2664	31.23123

Table F.1: Summary of statistic of Skartveit and Olseth depending on the solar elevation.

Appendix G

Comparison graphics

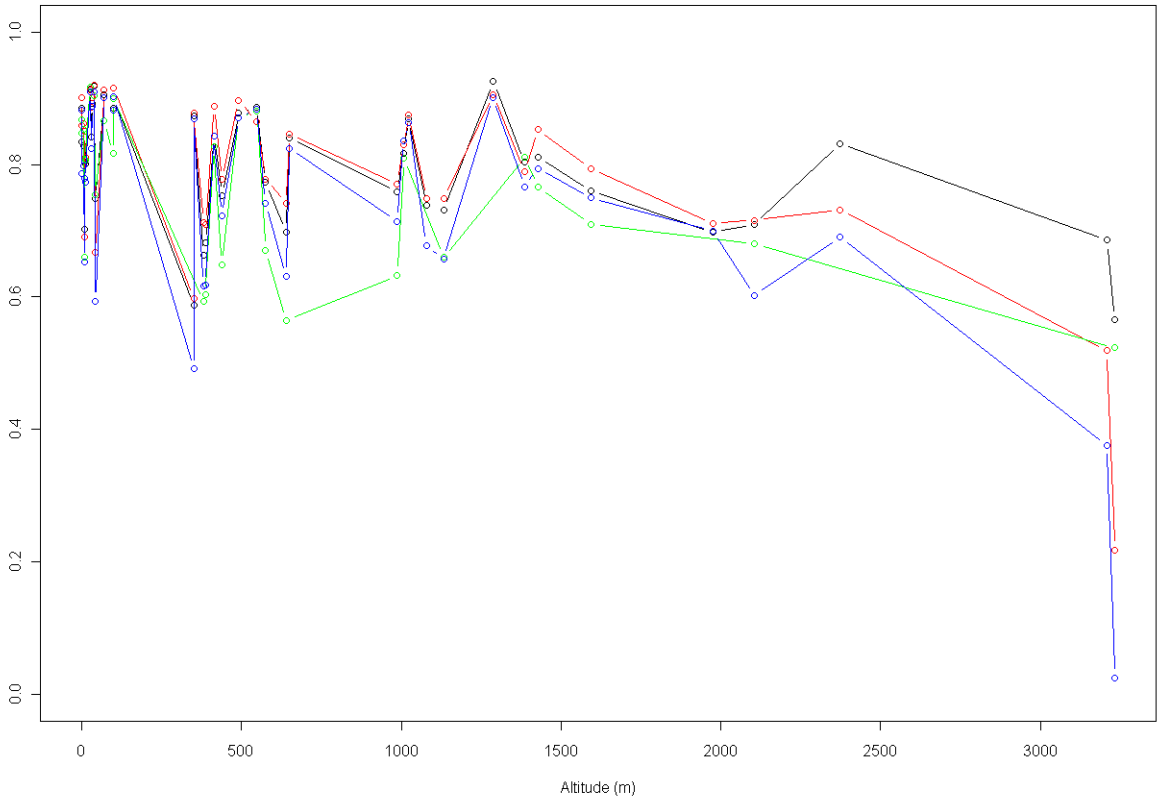


Figure G.1: The coefficient of determination R^2 for Reindl-Helbig (in black), Skartveit and Olseth (in red), BLR (in green) and Maxwell (in blue), in function of the stations **altitude**.

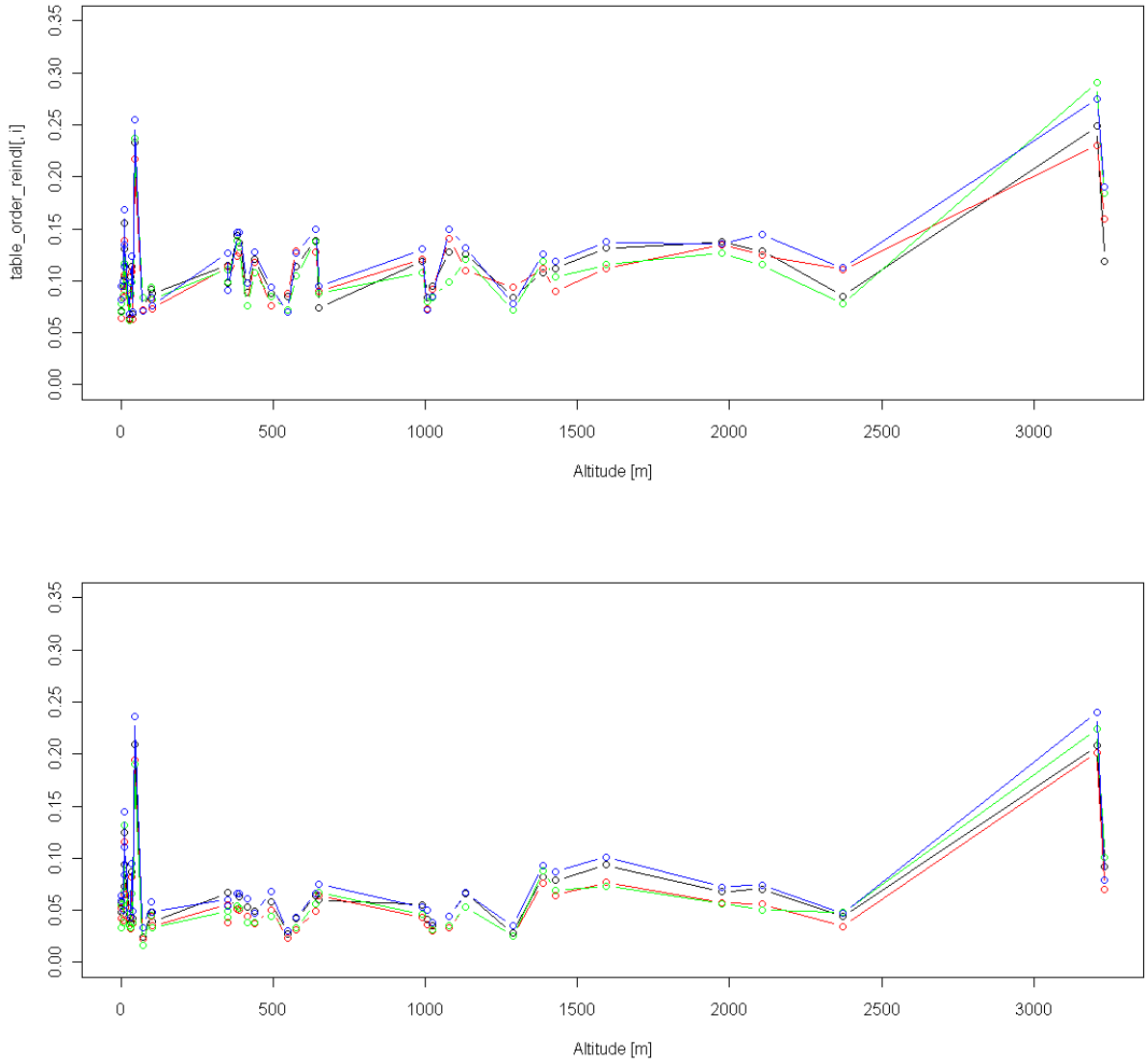


Figure G.2: The Mean Absolute Error, **MAE**, (up) and the Mean Bias Error, **MBE**, (down) for Reindl-Helbig (in black), Skartveit and Olseth (in red), BLR (in green) and Maxwell (in blue), in function of the stations **altitude**.

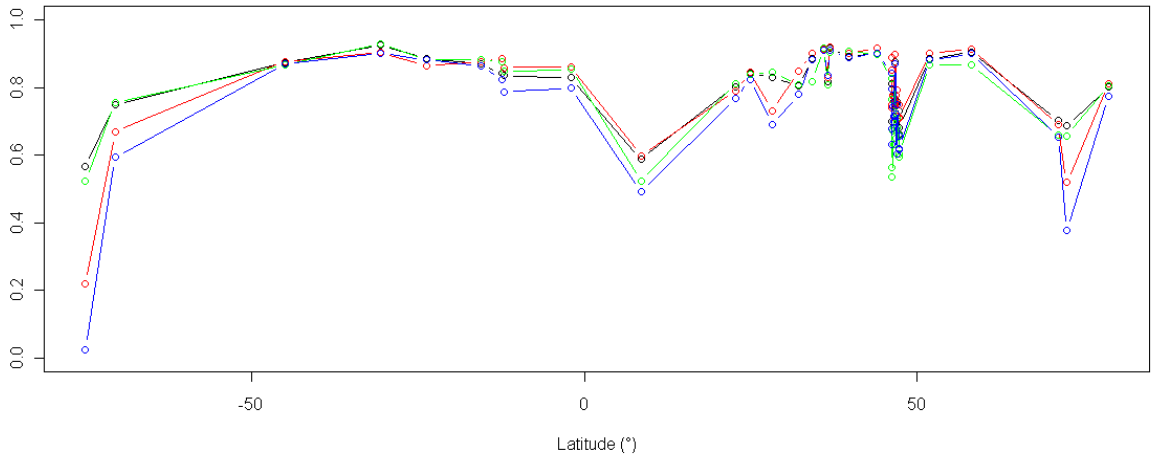


Figure G.3: The coefficient of determination R^2 for Reindl-Helbig (in black), Skartveit and Olseth (in red), BLR (in green) and Maxwell (in blue), in function of the stations **latitude**.

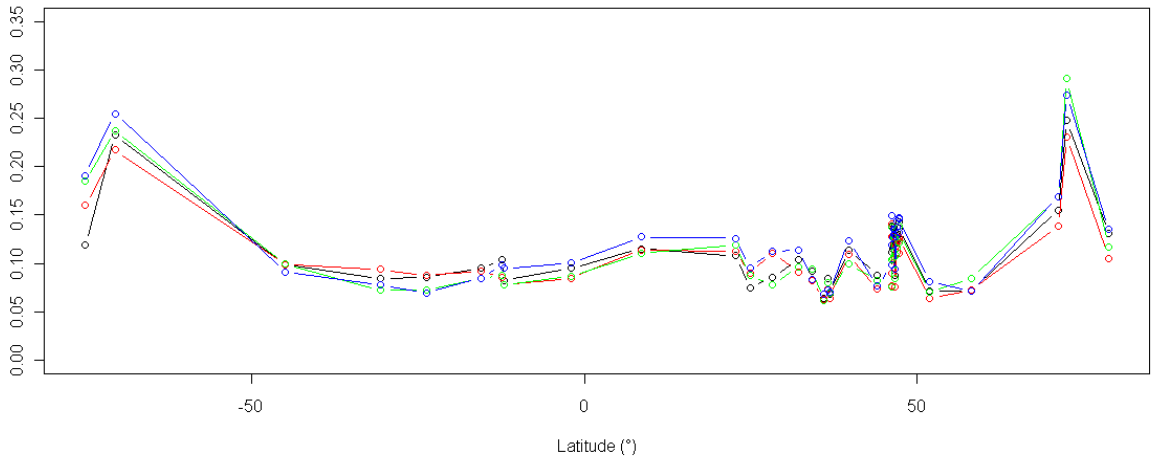


Figure G.4: The Mean Absolute Error, **MAE**, for Reindl-Helbig (in black), Skartveit and Olseth (in red), BLR (in green) and Maxwell (in blue), in function of the stations **latitude**.

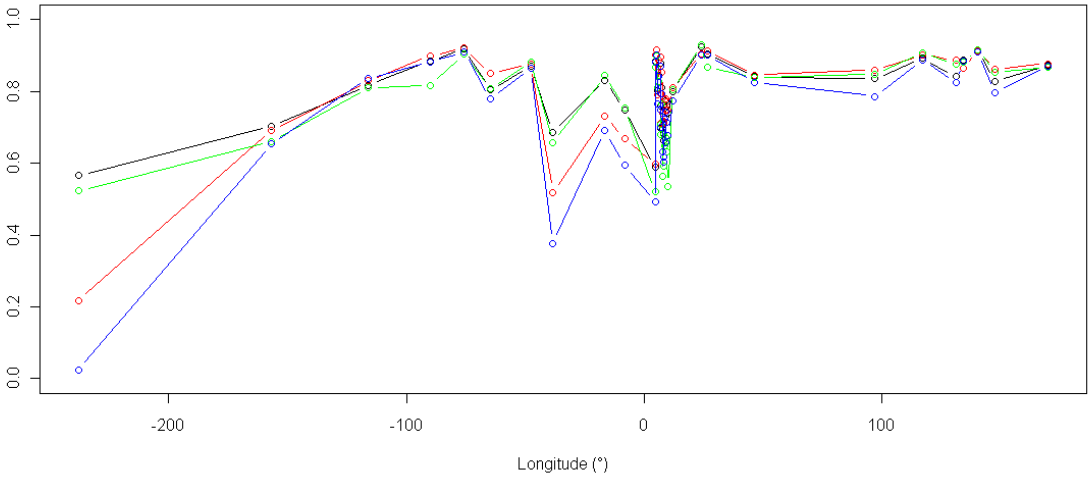


Figure G.5: The coefficient of determination R^2 for Reindl-Helbig (in black), Skartveit and Olseth (in red), BLR (in green) and Maxwell (in blue), in function of the stations **longitude**.

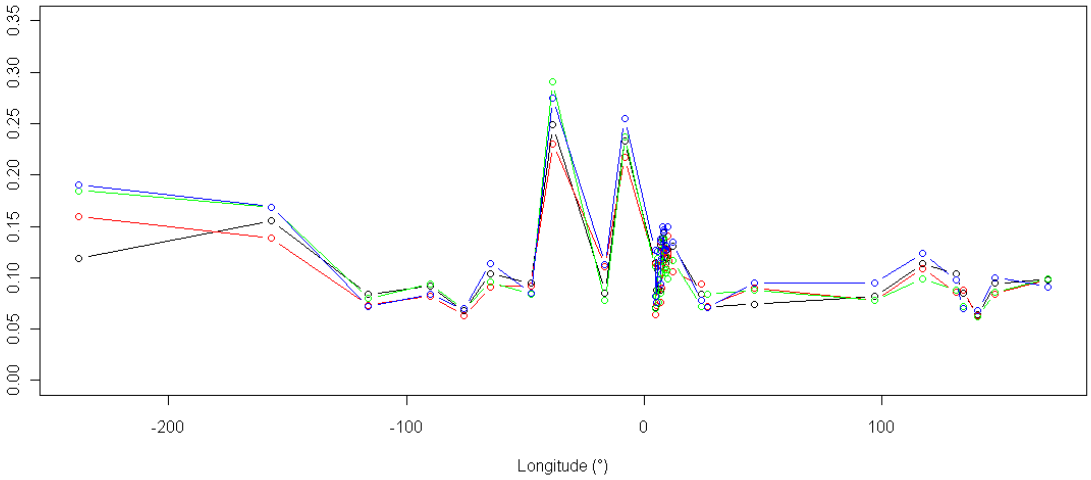


Figure G.6: The Mean Absolute Error, **MAE**, for Reindl-Helbig (in black), Skartveit and Olseth (in red), BLR (in green) and Maxwell (in blue), in function of the stations **longitude**.

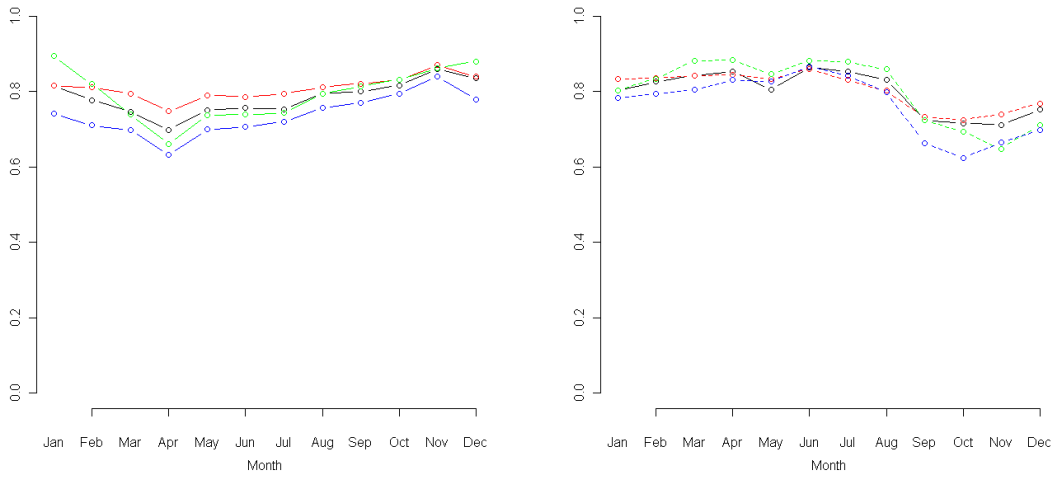


Figure G.7: The coefficient of efficiency **E** for Reindl-Helbig (in black), Skartveit and Olseth (in red), BRL (in green) and Maxwell (in blue), in function of the **months**, for stations in the Northern Hemisphere (left) and Southern Hemisphere (right).

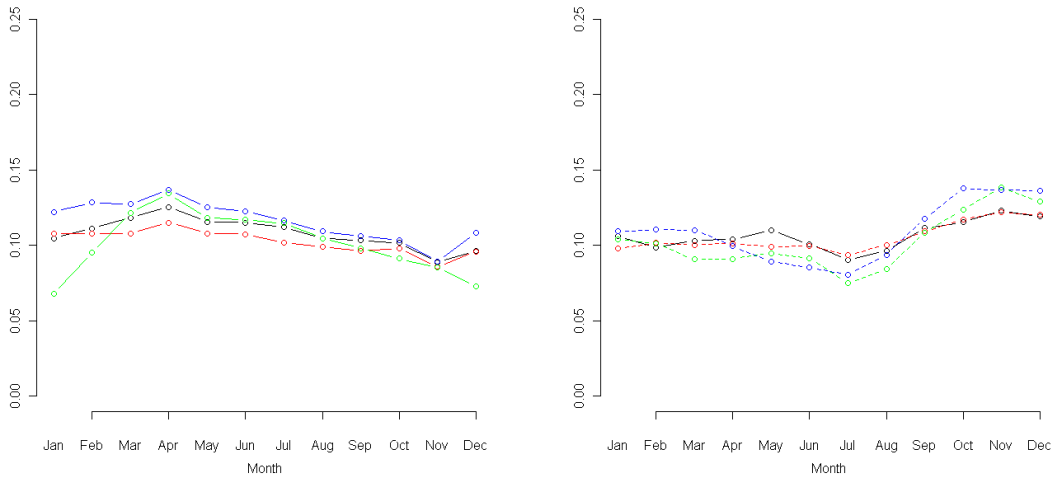


Figure G.8: The Mean Absolute Error, **MAE**, for Reindl-Helbig (in black), Skartveit and Olseth (in red), BRL (in green) and Maxwell (in blue), in function of the **months**, for stations in the Northern Hemisphere (left) and Southern Hemisphere (right).

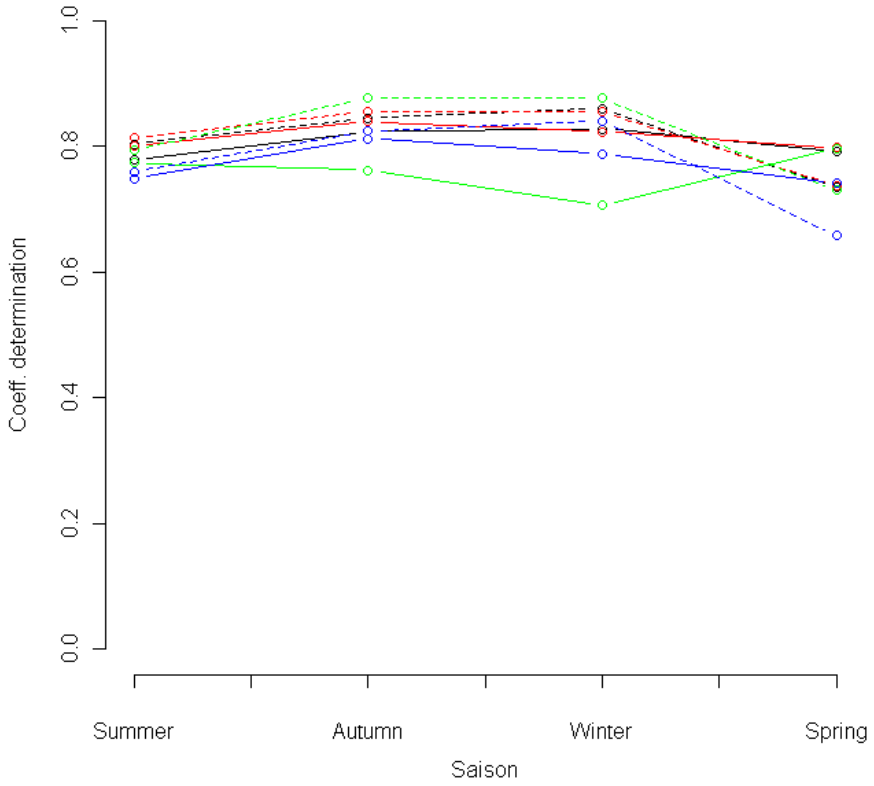


Figure G.9: The coefficient of determination R^2 for Reindl-Helbig (in black), Skartveit and Olseth (in red), BRL (in green) and Maxwell (in blue), in function of the **seasons**, for stations in the Northern Hemisphere (line) and Southern Hemisphere (dashed).

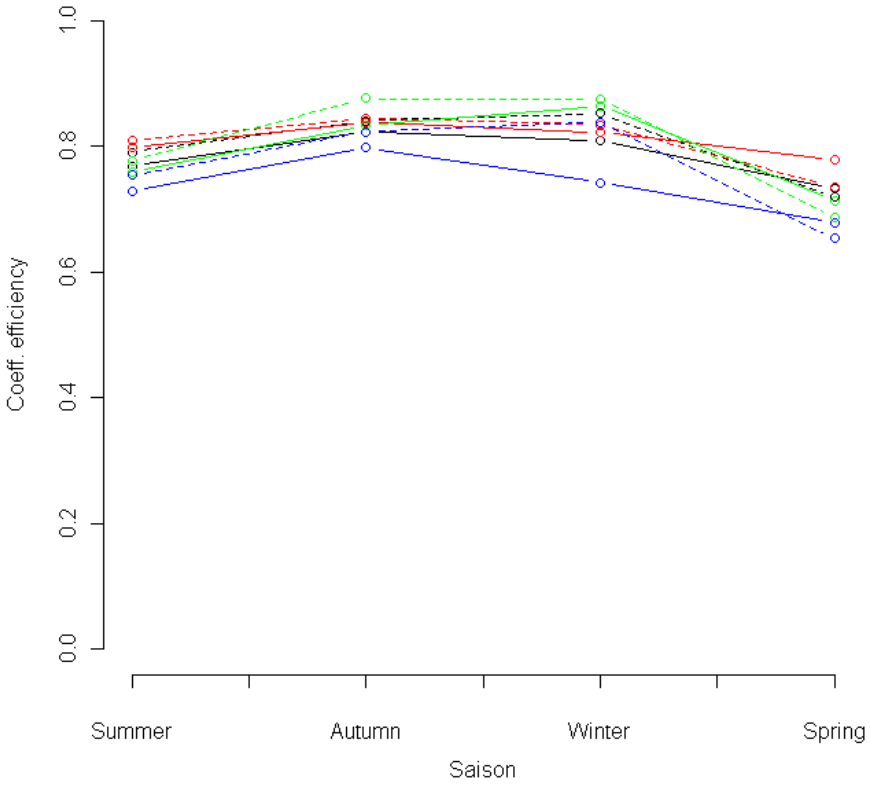


Figure G.10: Representation of the coefficient of efficiency **E** for Reindl-Helbig (in black), Skartveit and Olseth (in red), BRL (in green) and Maxwell (in blue), in function of the **seasons**, for stations in the Northern Hemisphere (line) and Southern Hemisphere (dashed)

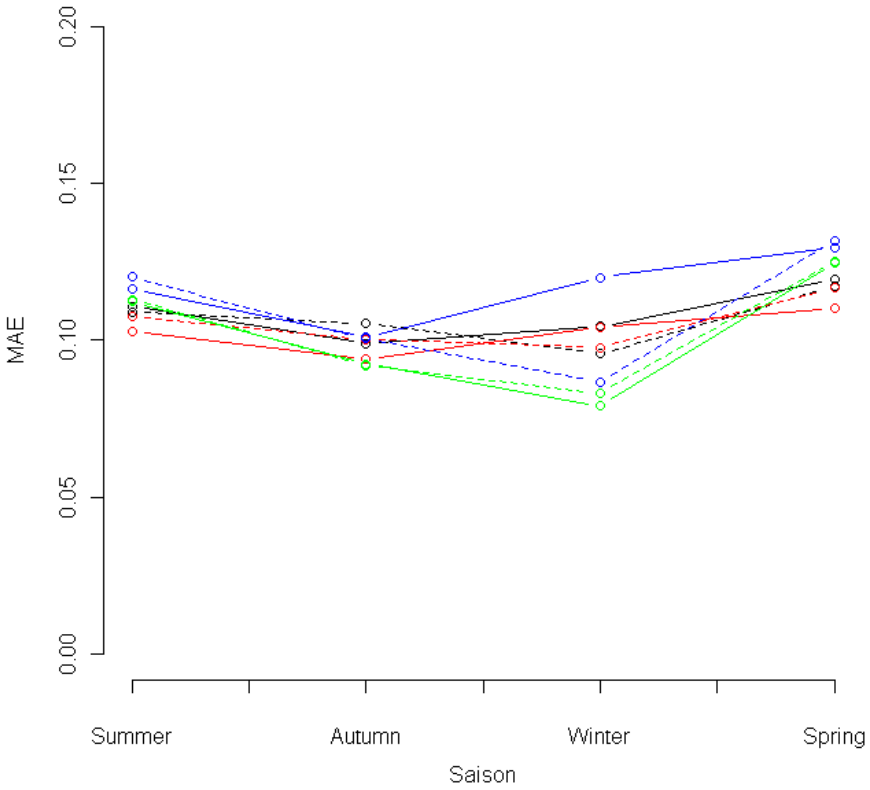


Figure G.11: Representation of the Mean Absolute Error (**MAE**) for Reindl-Helbig (in black), Skartveit and Olseth (in red), BRL (in green) and Maxwell (in blue), in function of the **seasons**, for stations in the Northern Hemisphere (line) and Southern Hemisphere (dashed)

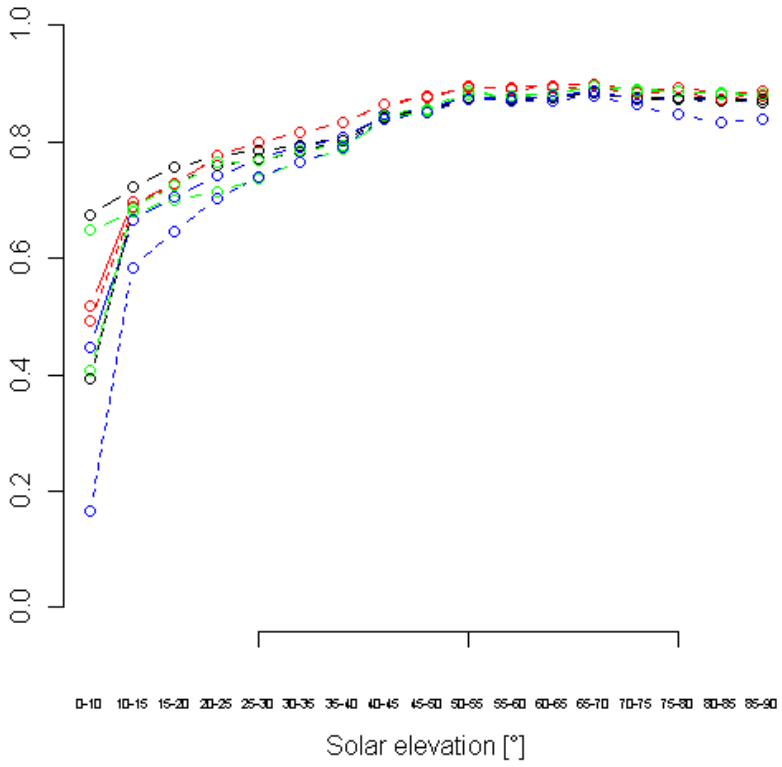


Figure G.12: The coefficient of determination R^2 (lined) and the coefficient of efficiency E (dashed) for Reindl-Helbig (in black), Skartveit and Olseth (in red), BRL (in green) and Maxwell (in blue), in function of the **solar elevations**.

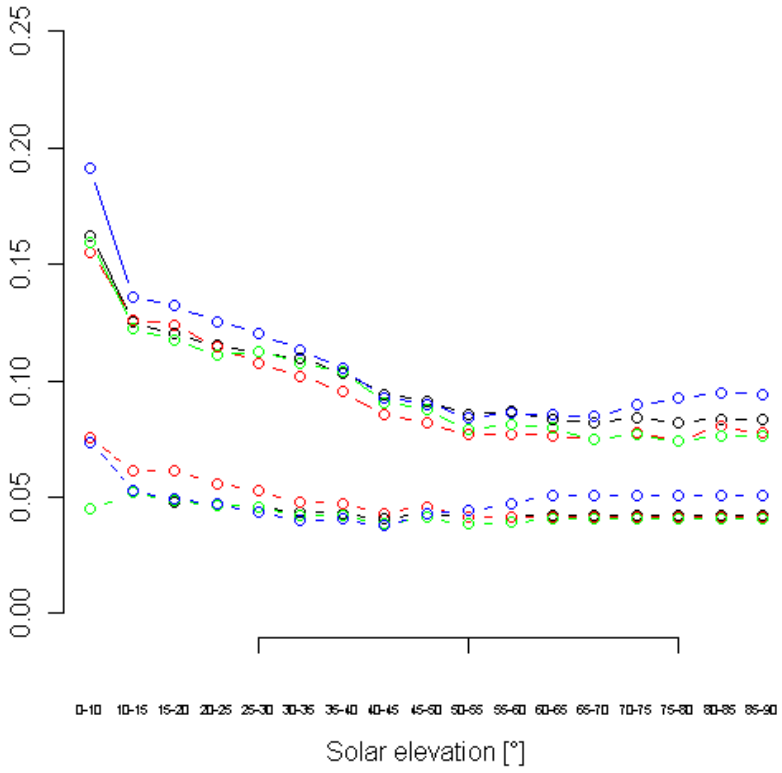
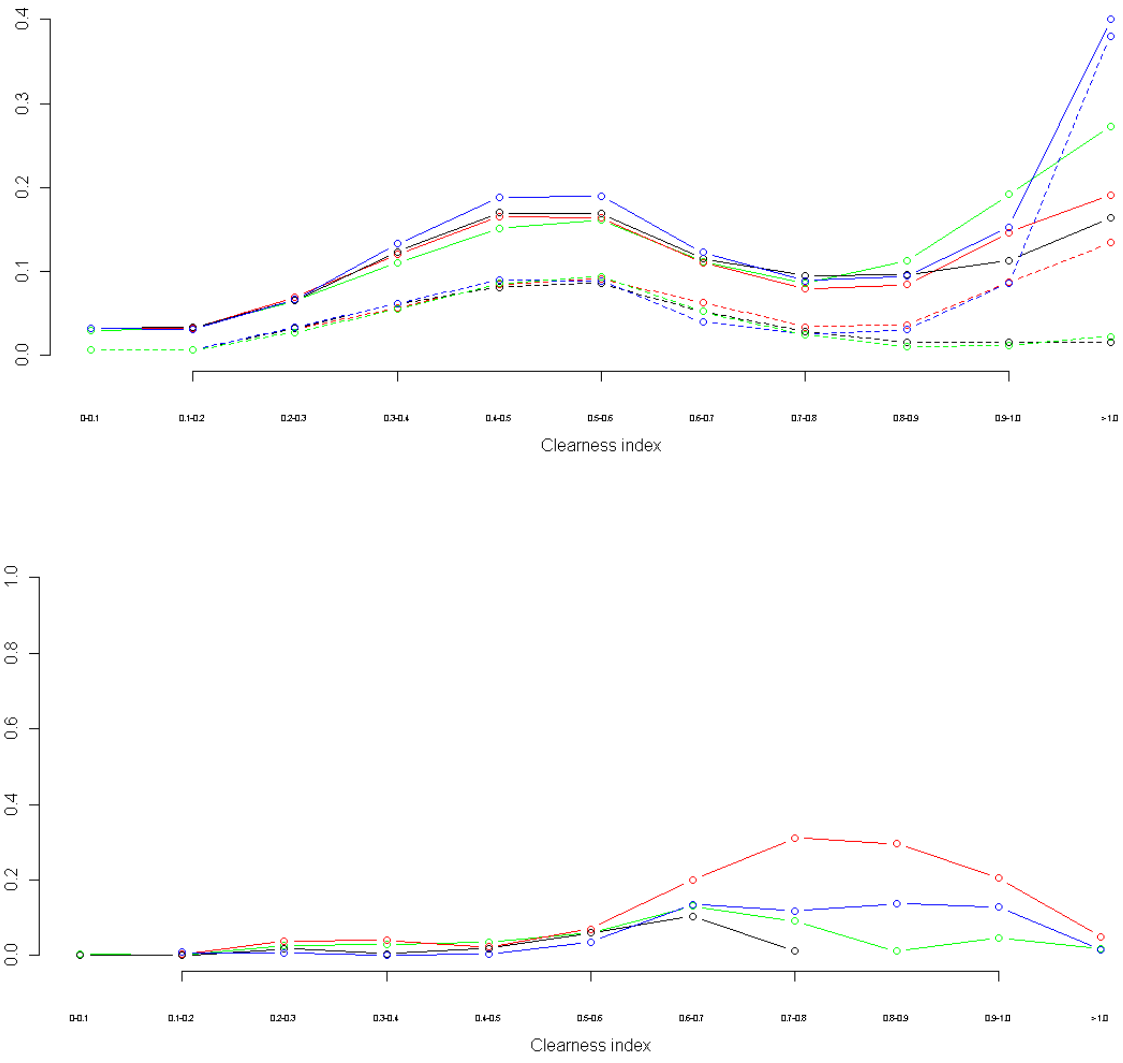


Figure G.13: Representation of the Mean Absolute Error, **MAE**, (lined) and the Mean Bias Error, **MBE**, (dashed) for Reindl-Helbig (in black), Skartveit and Olseth (in red), BRL (in green) and Maxwell (in blue), in function of the **solar elevations**.



captionThe Mean Absolute Error, **MAE**, (up) and the coefficient of determination R^2 (down) for Reindl-Helbig (in black), Skartveit and Olseth (in red), BRL (in green) and Maxwell (in blue), in function of the **clearness index**.

Appendix H

Nomenclature

I	Global radiation (in W m^{-2}
I_0	Extraterrestrial radiation (in W m^{-2}
I_b	Direct radiation (in W m^{-2}
I_d	Diffuse radiation (in W m^{-2}
r_0	Mean distance between the Earth and the Sun (in m)
λ	Solar wavelength (in m)
E_0	Eccentricity correction factor (dimensionless)
E_t	Equation of time (in minutes)
L_{loc}	Local longitude (in degree)
L_{st}	Standard longitude (in degree)
ω	Hour angle (in radian)
δ	Solar declination angle (in radian)
Γ	Day angle (in radian)
d_n	Day number (dimensionless)
Γ_{ne}	Day angle (in radian)
d_{ne}	Day equinox number (dimensionless)
n_0	Spring equinox time (dimensionless)
N_d	Number of data (dimensionless)
P_d	Percentage of data explained by 10%
Φ	Solar elevation (in degree)
Φ_{horiz}	Horizon angle (in degree)
h_0	Station altitude (in m)
h_{max}	Highest mountain altitude (in m)
d	Distance between the station and the highest mountain (in m)
k_t	Clearness index (dimensionless)
k_d	Diffuse fraction (dimensionless)
θ_z	Zenith angle (in degree)
ω_3	Hourly variability index (dimensionless)
$_{t,i}$	Cloudless clearness index, $i = 1, 2$ (dimensionless)
θ	Clear sky index (dimensionless)
α	Extreme beam transmittance (dimensionless)
A_s	Apparent Solar Time (in hour)
Ψ	Persistence factor (dimensionless)
K_t	Daily clearness index (dimensionless)
T_d	Dew temperature (in Celsius)
T_{air}	Surface temperature (in Celsius)

Continued on Next Page...

R_h	Relative humidity (in percent)
$m_a ir$	Air mass
K_n	Direct radiation transmittance (dimensionless)
$I_{b,DISC}$	Normal beam radiation estimated (W m^{-2})
χ	Correction matrix
k_t'	Normalized clearness index (dimensionless)
W	Precipitable water
$\Delta k_t'$	stability index (dimensionless)
q	Transmissivity (dimensionless)
q_{avg}	Transmissivity average (dimentionless)
q_{lim}	Transmissivity limit (dimentionless)
q_{diff}	Lower limit of q for cloud-free conditions (dimentionless)
q_{std}	Transmissivity standard deviation (dimentionless)
q_r	Transmissivity range (dimensionless)
Δq	Transmissivity offset (dimensionless)
Δa	Albedo offset (dimensionless)
t_0	Sunrise time (hour)
t_n	Sunset time (hour)
$\bar{t}$	Time at solar noon (hour)
MBE	Mean bias error (in the unit of the variables)
MAE	Mean absolute error (in the unit of the variables)
r	Correlation coefficient (dimensionless)
R^2	Coefficient of determination (dimensionless)
e	Residuals (in the unit of the variables)

ASRB	Alpine Surface Radiation Budget
BSRN	Baseline Surface Radiation Network
WCRP	World Climate Research Programme
WMO	World Meteorological Organization
ICSU	Intenational Council for Science

Continued on Next Page...

IOC Intergovernamental Oceonographic Commission