CHAPTER 4

QUALITY CONTROL

"Quality control must take place during the preparations towards data collection, during the measurement process and during transmission of the numerical values. Once data are recorded, only quality assessment (quality assurance) can be performed". (NREL, 1993)

"If we think of taking measurements as a test of our quality control process, quality assurance is the marking of the test. Unfortunately, as markers, we do not know the correct answers and we only know a fraction of the (obviously) incorrect ones. This means that even after all our efforts, we can never be 100% certain that the data is correct. Quality assurance procedures provide us with guides to flag data we believe to be suspect. In some cases, aided by operator notes, we can assure data is bad. In other cases, the best we can do is educated guesses." (Mc Arthur, 1995)

Since data accuracy and therefore data quality is the cornerstone of all BSRN measurements, this Chapter is devoted to quality control in the following manners :

- Different terminology and how quality control is applied in the BRSN by the WRMC.
- Elucidation of quality assurance procedures as applied by the WRMC database.
- Application and analysis of the WRMC procedures using real De Aar data.

To balance the high ideals sketched for a prefect measurement environment with practical reality, the two quotes at the beginning of this Chapter describe in essence what the process of quality control with data is all about in general, and BSRN in particular – it can never be 100% correct. Although the ideal of any measurement endeavour is to deliver the best and most accurate data at all times, practical experience shows, that there is always a quantity of data that is questionable to some degree. The true site scientist will develop ways in which to handle this data in a scientifically responsible way and keep the uncertainty, regarding the quality of such data, to a minimum.

Although scrutinizing data in itself is not pure science, it leads to improved data procedures, and a database with high integrity, which forms the basis from which good science can follow. This is actually the context in which the discussion of this Chapter is conducted.

4.1 CONTROL OVER THE QUALITY OF MEASUREMENTS

In the realm of concepts concerning control over the quality of measured data, quality control typically involves the following elements (NREL, 1993):

- Control in preparations for data collection (selection, calibration and installation of instruments);
- Control during the measurement process (inspection, calibration and maintenance of instruments);
- Control during the transmission and recording of numerical values (data acquisition systems, data archival and subsequent management);
- Controlling quality improvements by limited retrospective enhancement of measurements in cases of obvious and rectifiable mistakes.

If quality assessment is performed in real time or soon after the measurement process is completed, it can also provide valuable input to control the quality of future measurements, such as a real time data display on site (Long, 1996). A few good measurements widely spaced apart during the course of a day, have less value than a systematic set of measurements of comparable accuracy for the same day. Therefore, in essence, quality assurance has a definite dynamic character, something that site scientists must be aware of.

4.1.1 Terminology

The following definitions contextualize terms as they are used in this Chapter:

- Quality control: It is the larger holistic (Marion, 1993) and dynamic process involving the entire measurement effort starting with selection of instruments, continuing through the entire data acquisition process and maintenance routines, ending when the last backup of the data is being made. This encompasses all other processes striving to enhance data quality, whether it be past, present or future data.

- Quality assessment: It is the process of determining the value or rating the measurement by means of insertion of quality flags. This can be done using external ways of gauging, such as models, or internal ways ("comparing the data with itself"). The WRMC applies the latter in evaluating BSRN data.

- Quality assurance: After the measurement process, this is the "rubber stamp" of approval, which in itself has no control whatsoever over the measurement process itself. After assessment is made, only the assurance (or lack thereof) can be endorsed upon the recorded data.

- Quality enhancement: Retrospective or other means of attempting to enhance the quality of already recorded data. This is the only process whereby actual data is being changed. The utmost care and precision must be taken, so that the effort actually materializes in data quality *enhancement* and not *deterioration*. The BSRN site scientist of a particular site is the only person authorized to do quality enhancement on BSRN data for that specific site. This is because the site scientist is the person with access to all the relevant information leading to a questionable measurement, but he/she also has the best sense of judgement to offer solutions for data quality enhancement.

4.1.2 Measurement redundancy

Redundancy of measurements by means of a pair of operating instruments on as many parameters as possible, is a sound consideration for improvement of the quality of radiation measurements, as suggested by Mc Arthur (1998). Whilst two identical instruments in operation at a site can serve as "watchdogs" for one another, interruption of the measurement programme is less likely as a result of instrument failure if a backup is operating. The cost-factor of an instrument pair per measurement, makes this *modus operandum* impractical at De Aar and with exception of the double pyrheliometer as described in Sections 3.1.4 and 3.1.5, all the other measurements are carried out using only one instrument of each type.

The ideal way of exercising quality control, not only on the data, but actually on measurement processes, is by doing it on a continuous basis while the data is still as "fresh" as possible, as proposed by Long and Ackerman (1994). In this way, instrument or other hardware errors causing erroneous data, are identified and fixed as soon as possible. This also ensures to drastically limit the generation of erroneous data as a result of any specific error.

4.1.3 Keeping data fresh

The application of fresh instrument calibration constants when calculating irradiance values is one way of adding quality to data by applying retrospective methods.

The black surface of a radiometer thermopile becomes more reflective with prolonged solar exposure, hence less radiation is absorbed by the sensing element, and the instrument's calibration factor gradually becomes smaller. Regular calibration (the recommended frequency is once every six months) and subsequent updating of the radiometer's calibration constant is therefore imperative if the recorded irradiance values are expected to be a true reflection of the irradiance values.

If the irradiance values are calculated while thermopile measurements are sampled, care must be exercised that the latest (freshest) calibration constant is applied featured in the system for real–time quality control on values as close as possible to the actual values. If this is not done, necessary adjustments must be made on recorded data, so that archived irradiance values reflect the latest thermopile sensitivity. A simple interpolation of the calibration factors between calibration episodes helps keeping the values as close as possible to the actual sensitivity during a specific month, as applied at the De Aar BSRN station.

4.1.4 Handling seemingly erroneous data

It must be emphasized, that if data is flagged, it does not mean that it is invalid. It is merely an indication that data lies outside the expected ranges (NREL, 1992). The procedure of WRMC to flag data and to report it to the station scientist, is aimed towards enabling the station scientist to rectify or delete erroneous data.

However, care must be taken not to simply delete data without a good reason, since data inherently also contains a record of instrument performance, which can contain valuable information to aid in any form of future remedial processes. The key element is, that retrospective remedial actions on data must be based upon a good reasoning and a logical explanation as to why it is done.

The quote of Mc Arthur (1995) at the beginning of this Chapter emphasizes that quality assurance, amidst all the exact scientific endeavours, is not an exact science in itself . The judgement of data quality and enhancement has a subjective component, which is a function of pure experience.

4.2 WRMC VALIDATION CHECKS

The WRMC uniformly and indiscriminately applies specific quality assurance procedures on BSRN radiation data submitted to the archive. Although the WRMC is not allowed to change any radiation data (that privilege belongs to the site scientist only), all data are flagged using a specific set of quality control procedures, as described in Sections 4.2.1 to 4.2.3. The data flags are reported to the site scientist, who re-evaluates the data and in the event of identification of errors, rectifies and re-submits the newer version of the dataset. The re-submitted data is re-evaluated at the database, re-flagged and returned. This process is repeated until the data can be regarded as reliable (Gilgen *et al*., 1991).

In the early BSRN years, the following main quality control procedures, were agreed upon (Gilgen *et al.*, 1993):

- Procedure 1: Physically possible
- Procedure 2: Extremely rare
- Procedure 3: Across quantities
- Procedure 4: Comparison with a model
- Procedure 5: Eye check of time series plots

Each of these procedures consists of sub-procedures, each with specifications involving certain boundaries of various radiation quantities. In a later revision of the data quality control techniques (Hegner *et al.*, 1998), it was mentioned that Procedures 4 and 5 were still under development at that time and were therefore not implemented. For the remaining three procedures, boundary values and descriptions were revised. These three main quality assurance procedures, as applied to BSRN data, are now detailed in the next three Sections (4.2.1 to 4.2.3). An evaluation with real De Aar data collected in the three years between June 2000 and May 2003, in Section 4.3.

4.2.1 Procedure 1 ("Physically possible")

This is the first, most basic, and roughest of the procedures. The intention is only to identify large and random errors that might be introduced in the initial data processing (Hegner *et al.*, 1998, Ohmura *et al.*, 1998). Data which are reasonably well controlled, should (at least in principle) be able to pass this test. Procedure 1 is applied to each radiation quantity individually, independent from other radiation fluxes or non-radiation data.

The sub-procedures for procedure 1 are listed in Table 4.1, together with the latest WRMC definition of lower and upper bounds per specified radiation quantity (Hegner *et al.,* 1998). All these quantities are based upon absolute maximum or minimum radiation quantities conceivable under general present climatic conditions of the Earth (Ohmura *et al.*, 1998).

Table 4.1 *Sub-procedures of WRMC procedure 1: Physically possible quantities*

Sub-procedure	Lower bound	Radiative quantity	Upper bound
1.1	0 W.m^{-2}	Direct solar irradiance (DSDIR)	1368 W.m^{-2} (annual mean solar constant)
1.2	0 W.m^{-2}	Diffuse sky irradiance (DSDFS)	TOA irradiance + 10 W.m^{-2}
1.3	0 W.m^{-2}	Global irradiance 2: measured with pyranometer (DSGL2)	1368 W.m^{-2} (annual mean solar constant)
1.4	0 W.m^{-2}	SW reflected irradiance (USR)	TOA irradiance + 10 W.m^{-2}
1.5	50 W.m^{-2}	LW Downwelling radiation (DL)	700 W.m^{-2}
1.6	50 W.m^{-2}	LW Upwelling radiation (UL)	700 W.m^{-2}

The following are relevant to Table 4.1:

- In this procedure, as well as other procedures, where applicable, TOA radiation (in W.m^{-2}) over a one-minute is evaluated by using the following equation (Iqbal, 1983):

$$TOA = \frac{S}{A_0}(\sin\delta\sin\varphi + \cos\delta\cos\varphi\cos(2\pi\xi + \varphi - \pi)) \qquad (4.1)$$

where

S = annual average of solar constant (1368 W.m^{-2})

A_0 = radius vector of solar distance
δ = solar declination in radians
φ = site latitude in radians
ξ = equation of time in radians

- The maximum theoretically possible TOA for De Aar is 1402 W.m^{-2}, using Equation 4.1 at the annual solar solstice during 22-23 December, when the solar zenith angle is just 7.2° (De Aar latitude (-30.7°) minus solar latitude (-23.5°) on that day).

- Using an atmospheric transmission coefficient of 0.95 (highest ever measured at De Aar) and the absolute maximum TOA of 1402 W.m^{-2}, the associated global radiation at the summer solstice is expected to be 1332 W.m^{-2}, This is significantly less than the boundary of 1368 W.m^{-2} set in Sub-procedure 1.3, therefore all values of DSGL2 are expected to be less than 1368 W.m^{-2}.

- In Sub-procedures 1.5 and 1.6, the extremes of 50 W.m^{-2} and 700 W.m^{-2} for DL is based on the highly unlikely blackbody effective sky temperatures of -100°C and 60°C, respectively.

4.2.2 Procedure 2 (“Extremely rare”)

This procedure is applied to data immediately after having passed Procedure 1. The specifications for interval limits are narrower than Procedure 1, in order to identify erroneous data that have escaped Procedure 1. The sub-procedures depicted in Table 4.2, are defined only in terms of the upper bound of radiative quantities which are physically possible. Under certain conditions, quantities can momentarily overshoot the limits, as detailed in Section 1.3. It was suggested (Dutton, 2002), that the sub-procedure limits are characterized per station to best fit extremely rare results specifically applicable to a certain location. However, the elucidation and evaluation of Procedure 2 is applied to stations in general, not on a one by one basis.

Kasten optical air mass value m (Kasten *et al.*, 1989) is defined by

$$m = (\cos Z + 0.50572(96.07995 - Z)^{-1.6364})^{-1} \qquad (4.2)$$

where

Z = Solar zenith angle measured in degrees, not radians.

Table 4.2 *Sub-procedures of WRMC procedure 2: Extremely rare quantities*

Sub-procedure	Radiative quantity	Upper bound
2.1	Global irradiance 2 : measured with pyranometer (DSGL2)	TOA irradiance for $Z < 80°$ TOA + $0.56(Z-93.9)^2$ for $Z \geq 80°$
2.2	Upward reflected SW irradiance (USR)	0.95 * DSGL2
2.3	Diffuse sky irradiance (DSDFS)	700 $W.m^{-2}$
2.4	Direct solar irradiance (DSDIR)	TOA irradiance * 0.9^m (where 'm' is optical air mass defined in Equation 4.2)
2.5	Downwelling LW irradiance (DL)	Upwelling LW irradiance + 30 $W.m^{-2}$
2.6	Upwelling LW irradiance (UL)	Downwelling LW irradiance - 30 $W.m^{-2}$

The following is applicable to table 4.2:

- Solar zenith angle Z (as applicable to sub-procedures 2.1 and 2.4) is measured in degrees, not radians.

- Very high diffuse radiation is likely to be achieved in overcast conditions, with highly transmitting cloud and high ground albedo. The current globally measured, audited and controlled record is 700 $W.m^{-2}$ for DSDFS[1].

- Global radiation values larger than TOA irradiance can be measured in areas of low latitude when scattered clouds pass in front of the sun and add reflected radiation from the cloud bottom to the already high level of global radiation.

- Since the Rayleigh extinction of normal beam radiation is about 9% at mean sea level[2], direct normal transmission coefficients are rarely more than 0.9 – although the altitude (1287m above sea level) of De Aar is expected to play a prominent role by increasing the effective transmission coefficient. The behaviour of transmission coefficients evaluating real De Aar data, are discussed in detail in Section 1.3.2.4.

---

[1] http://bsrn.ethz.ch/quality/procedure2

[2] http:// bsrn.ethz.ch/quality/procedure2

4.2.3 Procedure 3 ("Across quantities")

This procedure is applied immediately after Procedure 2, and the interval limits per sub-procedure are again narrower than those in Procedure 2, in order to identify errors that could have escaped Procedures 1 or 2. This procedure is unique in the sense, that it involves more than one quantity in any specific comparison process, hence the name "Across quantities".

The sub-procedures are based upon general, globally obtained, empirical relations of the different quantities measured (Hegner *et al.,* 1998) and therefore the best application for customization with respect to specific sites using local conditions, exists here. Details of the sub-procedures of Procedure 3 are listed in Table 4.3.

Table 4.3 *Sub-procedures of WRMC procedure 3: Across quantities*

Sub-procedure	Lower bound	Radiative quantity	Upper bound
3.1	$0.7 * \sigma T^4$	Downwelling LW irradiance (DL)	$1.0 * \sigma T^4$
3.2	$\Sigma(T - 10)^4$	Upwelling LW irradiance (UL)	$\sigma(T + 10)^4$
3.3	Global 2 (DSGL2) – Diffuse sky (DSDFS) - 50 $W.m^{-2}$	Direct solar irradiance horizontal component (DSDIR*cosZ)	Global 2 (DSGL2) – Diffuse sky (DSDFS) + 50 $W.m^{-2}$
3.4	Direct solar irradiance horizontal component (DSDIR*cosZ)-50 $W.m^{-2}$	Global 2 (DSGL2) – Diffuse sky (DSDFS)	Direct solar irradiance horizontal component (DSDIR*cosZ) + 50 $W.m^{-2}$

<u>The following are applicable to Table 4.3:</u>

- LWD is the equivalent of DL, the latter being used in the table for consistency towards the terms used in Hegner *et al.* (1998).

- T = (Stevenson) screen temperature in Kelvin measured at LW instrument height.

4.2.3.1 LW radiation

Sub-procedure 3.1 is based upon Stefan–Bolzmann's Equation, assuming the atmosphere as a blackbody radiation emitter:

$$E_{LW} = \varepsilon\sigma T^4 \tag{4.3}$$

where E_{LW} = Downwelling longwave irradiance in W.m^{-2}
T = (Stevenson) screen temperature in Kelvin
ε = Atmopsheric emmissivity
σ = Stefan-Bolzmann constant

Rearranging Equation 4.3 yields

$$\varepsilon = \frac{E_{LW}}{\sigma T^4} \tag{4.4}$$

Therefore, all values of E_{LW} passing sub-procedure 3.1 are those E_{LW} for which $0.7 < \varepsilon < 1.0$.

There are a number of parameterization schemes in the literature, that were used through the years to calculate ε in terms of other surface parameters, such as surface (screen) temperature and in some cases, surface humidity expressed as surface vapour pressure (e).

Bearing in mind that E_{LW} is the result of LWD radiation in an air column above a measuring point, including the entire atmosphere up to the TOA, the only sure way of determining LWD is to evaluate the radiative properties of the entire column in small increments using upper-air soundings. On the other hand, one point measurement on the surface cannot truly represent an entire air column.

However, one point measurement on the surface is sufficient under clear skies and the associated stable atmosphere as the result of settled weather. A stable upper-air profile follows the Standard Atmosphere to such an extent, that only the surface temperature is needed to describe it accurately enough for accurate LWD parameterizations.

The first attempts to estimate LWD, using surface parameters, were done by Angström in 1918, followed by Brunt in 1932 (Jiménes *et al.*, 1987, Iziomon *et al.*, 2003). The majority of these attempts were empirical in nature. One exception is the pure theoretical approach as performed by Brutsaert (1975), using a clear-sky equation and making use of the standard atmosphere.

An intercomparison of selected parameterization schemes is now being offered. For the three-year data period at De Aar (June 2000 to May 2003), LWD data and simultaneous five-minute measurements of temperature and humidity (converted to vapour pressure) were used, together with SYNOP identification of clear skies (all in one observed dataset), to calculate percentage differences between measured LWD and calculated LWD from a set of parameterization schemes collected by Pirazzini (1998) and Olivieri (pers comms, 2003).

In total, there were 101 074 data points in the three-year period. The percentage differences (calculated LWD minus measured LWD), are listed in Table 4.4 as Mean (averages of the errors) and Root Mean Square (RMS) error (standard deviation of errors). The parameterization schemes are ordered in ascending RMS error order in Table 4.4.

Table 4.4 *Intercomparison of a number of LWD parameterization schemes using De Aar measured LWD and surface meteorological data*

Author(s)	**Year**	**Equation**	**Mean error**	**RMS error**
Centeno *	1982	$E_{l\,w} = [(5.77+0.996*0.601^{H})T^{1.1893}(U^{0.0665}/10^{4})]\ \sigma T^{4}$	-8.85	3.49
Satterlund	1979	$E_{l\,w} = 1.08\ [1-\exp(-e^{T/2016})]\ \sigma T^{4}$	7.13	4.16
Konzelmann *et al.*	1994	$E_{l\,w} = [\ 0.23 + 0.484\ (\ e/T\ )^{1/8}\ ]\ \sigma T^{4}$	-35.04	4.77
Andreas & Ackley	1982	$E_{l\,w} = [\ 0.601 + 5.95 \times 10^{-5}\ e\ \exp(1500/T)\ ]\ \sigma T^{4}$	-7.98	4.85
Marshunova	1966	$E_{l\,w} = (0.67+0.05e^{0.5})\ \sigma T^{4}$	8.69	4.88
Efimova	1961	$E_{l\,w} = (0.746+0.0066^{e})\ \sigma T^{4}$	3.82	4.95
Prata	1996	$E_{l\,w} = \{1-(1+46.5(e/T)\exp[-(1.2+3(46.5(e/T))^{1/2}]\}\ \sigma T^{4}$	29.61	5.05
K- Langlo *et al.*	1994	$E_{l\,w} = 0.765\ \sigma T^{4}$	4.00	5.19
Brutsaert	1975	$E_{l\,w} = 1.24\ (\ e/T\ )^{1/7}\ \sigma T^{4}$	-1.66	5.27
Brunt	1932	$E_{l\,w} = (0.51+0.066e^{0.5})\ \sigma T^{4}$	-7.54	5.53
Ohmura	1981	$E_{l\,w} = (8.733 \times 10^{-3}T^{0.788})\ \sigma T^{4}$	3.74	6.00
Guest	1998	$E_{l\,w} = -85.6 + \sigma T^{4}$	6.82	6.43
Idso and Jackson	1969	$E_{l\,w} = [1-0.26\exp(-7.77\times10^{-4}(273-T)^{2})]\ \sigma T^{4}$	8.59	7.86
Swinbank	1963	$E_{l\,w} = (9.365 \times 10^{-6}T^{2})\ \sigma T^{4}$	7.40	8.23

* *NOTE:- The Centeno Equation was obtained from personal communication with Olivieri (2003) quoting* International Energy Agency *Task 17: Measuring and Modelling Spectral Radiation Affecting Solar Systems and Buildings: Review and Test of Parameterizations of Atmospheric Radiation* Report No. IEA-SHCP-17F-2 (December 1994). In this Equation, H represents the station height in km, and U the surface relative humidity in %.

The following is concluded from Table 4.4:

- All the listed parameterization schemes have RMS errors of between 3% and 10%.

- With the exception of the parameterization schemes of Konzelmann *et al*. (-35.04%) and Prata (29.61%), all featured parameterizations have mean errors of between -9% and 9%.

- The widely used equation of Swinbank (1963) has the highest RMS error of the quoted parameterization schemes for the De Aar data.

- By virtue of the lowest RMS error and a modest mean error, the equation of Centeno is found to be the best. It is fine-tuned for usage at De Aar by reverse-applying the systematic error of -8.852 %:

$$E_{LW} = [(6.28 + 1.084 * 0.601^{H})T^{1.1893}\frac{U^{0.0665}}{10^{4}}]\sigma T^{4} \qquad (4.5)$$

Equation 4.5 can be used for estimating LWD when the operational pyrgeometer is not measuring LWD for any one of many reasons.

The usefulness of parameterization schemes is that, apart from controlling the quality of E_{LW} data, it can also be used to retrospectively estimate E_{LW} for periods where T and U were available, in order to fill data gaps, such as the completion of hourly, daily and monthly averages, in a long-term time series of those quantities.

For a stable, cloud-free atmosphere, 0.7 is regarded as a generally accepted value for emmissivity[3]. Any increase in this number is brought about by the presence of LW radiators, such as water vapour and clouds. A water film on the silicon pyrgeometer dome can also artificially register a higher value for LWD and therefore of the emmissivity, due to the higher hygroscopic properties of silicon. Other reasons for registering a higher value for the emmissivity may be mis-shading of the dome under sunny conditions, or wrong compensation for dome temperature in cases of no shading.

---

[3] http://bsrn.ethz.ch/quality/procedure3

4.2.3.2 Shortwave radiation

The last two sub-procedures in Table 4.3 are comparisons between global, diffuse and direct irradiances using solar zenith angle. The relationship

$$global(t) = diffuse(t) + direct(t)\cos Z \qquad (4.6)$$

should be satisfied in principle for all values of time (t) when $Z < 90.83°$ (i.e., when the sun is visible). However, often this is not the case due to:

- Misalignment of the tracker device leading to erroneous measurement of either one of the quantities or a combination thereof.
- Incorrect calibration of one or a combination of the involved thermopile instruments.
- Differences in $1/e$ reaction times between the sensors in rapidly changing circumstances, resulting in Equation 4.6. not being satisfied at specific points in time.
- Skewness of the pyranometers brought about by incorrect levelling, resulting in an over-registering when titled slightly more towards the sun, or an under-registration when tilted away from the sun. The diffuse pyranometer (on the tracker) is particularly prone to such errors due to the rotating tracker table being not exactly level at all times.

In Table 4.3, the two sub-procedures in Points 3.3 and 3.4 are exactly the same, i.e., the defining inequalities are describing the same conditions. This statement can be substantiated as follows:

If $X = DSDIR*\cos Z$ and $Y = DSGL2 - DSDFS$

Then sub-procedure 3.3 can be rewritten: The test is passed if $(Y - 50\ \mathrm{W.m^{-2}}) < X < (Y + 50\ \mathrm{W.m^{-2}})$

(subtract Y consistently) $$\mathbf{-50\ W.m^{-2} < X - Y < 50\ W.m^{-2}} \qquad (4.7)$$

Sub-procedure 3.4. can be rewritten as: The test is passed if $(X - 50\ \mathrm{W.m^{-2}}) < Y < (X + 50\ \mathrm{W.m^{-2}})$

(subtract X consistently) $-50\ \mathrm{W.m^{-2}} < Y - X < 50\ \mathrm{W.m^{-2}}$

(multiply by -1) $50\ \mathrm{W.m^{-2}} > X - Y > -50\ \mathrm{W.m^{-2}}$

(read from the right) $$\mathbf{-50\ W.m^{-2} < X - Y < 50\ W.m^{-2}} \qquad (4.8.)$$

Equation 4.8 and therefore the latter of the two sub-procedures, is hence exactly the same sub-procedure as the first in Equation 4.7. Although 50 $W.m^{-2}$ is considered to be a large error in any one of the individual LW quantities, the margin of 50 $W.m^{-2}$ is allowed to give room to the addition of uncertainties in all three of the individual sensors (direct, global and diffuse). This is justified by the typical occurrence of ($DSGL2 - DSDFS - DSDIR*\cos Z$) outside a boundary of 50 $W.m^{-2}$.

4.3 APPLICATION OF THE WRMC PROCEDURES

In this Section, each of the WRMC procedures discussed in Section 1.2 are applied to real De Aar data consisting of 36 monthly files for three complete years, viz., June 2000 until May 2003 inclusive. The data is presented here with the minimum procedures applied, but taking thermopile deterioration into account, as discussed in Section 4.1.3. The term “datapoint” refers to the one-minute average value, as recorded in the BSRN one-minute statistics.

For every sub-procedure in Section 4.2, the data is analyzed using a table, listing on a monthly basis how the one-minute data is distributed with respect to the upper and/or lower bounds of the sub-procedure. All the Tables (Tables 4.5 to 4.13) are presented as frequency tables. In two cases (Table 4.6 and Table 4.9) the table is only one column containing the frequency of boundary violations of the sub-procedure, plus the number of missing datapoints and/or the possible datapoints (the number of minutes per specific month). In all the other cases, a frequency distribution using bins related to fractions of the distance between the boundaries, is employed. In this way, the relative seriousness of boundary violations can be assessed. Where applicable, columns containing data, which violate boundaries of a specific sub-procedure are shaded.

4.3.1 Procedure 1 – “Physically possible”

4.3.1.1 Sub-procedure 1.1 *(The test for $0\ W.m^{-2} < DSDIR < 1368\ W.m^{-2}$)*

The lower and upper boundaries for this sub-procedure are 0 $W.m^{-2}$ and 1368 $W.m^{-2}$, respectively. The range (between lower and upper boundary) is therefore 1368 $W.m^{-2}$, which equals 12 x 114 $W.m^{-2}$. Therefore the data in Table 4.5 was placed in 12 equally spaced frequency bins, each sized 114 $W.m^{-2}$. This is intended to show how the data is distributed in the range between lower and upper boundary, as well as the relative seriousness of boundary violations.

Added columns, are: the number of points equal to 0 W.m^{-2} , the number of missing points and the possible number of points. For each frequency bin, the total number expressed as a percentage, is listed. This is also expressed as the total of seasonal averages, i.e. December-January-February (DJF) for summer, March-April-May (MAM) for autumn, June-July-August (JJA) for winter and September-October-November (SON) for spring.

Table 4.5 *Frequency distribution of DSDIR: June 2000 to May 2003, ref.to sub-procedure 1.1*

		Data bins in W.m^{-2}												
Month	= 0	1 to 114	115 to 228	229 to 342	343 to 456	457 to 570	571 to 684	685 to 798	799 to 912	913 to 1026	1027 to 1140	1141 to 1254	1265 to 1368	Mis-sing	Possible
Jun 2000	26120	1322	621	672	772	936	1416	2748	5570	2978	0	0	0	45	43200
Jul 2000	26808	1297	605	685	819	1074	1580	3100	6158	2245	0	0	0	269	44640
Aug 2000	24757	1356	848	875	1022	1252	1777	3193	5165	4367	0	0	0	28	44640
Sep 2000	25298	2242	569	643	675	870	1104	1943	3693	5255	880	0	0	28	43200
Oct 2000	22033	2926	1058	1006	1035	1080	1446	2277	3414	4710	2011	0	0	1644	44640
Nov 2000	19852	2435	784	826	1144	1181	1416	1553	2217	6018	2880	0	0	2894	43200
Dec 2000	19653	3381	1251	1727	3761	4201	3058	767	1306	3744	1759	0	0	32	44640
Jan 2001	20340	2821	745	700	810	898	1169	1772	2806	6822	5725	0	0	32	44640
Feb 2001	19747	3023	633	612	639	765	1042	1556	2802	6601	2867	0	0	33	40320
Mar 2001	23121	3885	746	674	736	840	1189	1640	3226	6436	692	0	0	1455	44640
Apr 2001	25965	4954	829	763	713	805	945	1431	2785	3951	0	0	0	59	43200
May 2001	26005	2694	749	732	836	1055	1445	2310	4962	3702	0	0	0	150	44640
Jun 2001	25255	2278	583	665	790	1018	1619	2900	5120	1889	0	0	0	1083	43200
Jul 2001	27401	2950	749	718	811	1110	1623	2579	3902	2766	0	0	0	31	44640
Aug 2001	24728	1639	735	798	848	1038	1310	2003	4675	6843	0	0	0	23	44640
Sep 2001	24701	2491	718	695	809	1096	1447	2287	3233	5158	10	0	0	555	43200
Oct 2001	23556	2183	716	774	938	1197	1640	2834	4590	5197	985	0	0	30	44640
Nov 2001	22950	3885	821	712	761	847	1115	1544	2547	5105	2881	0	0	32	43200
Dec 2001	20301	2153	686	697	804	969	1327	1941	3018	7546	5152	0	0	46	44640
Jan 2002	20949	2327	726	684	775	865	1178	1693	2786	6734	5884	0	0	39	44640
Feb 2002	19847	1904	601	615	728	905	1220	1628	2926	6488	3421	0	0	37	40320
Mar 2002	24042	2604	785	720	813	1007	1498	2587	5310	4776	197	0	0	301	44640
Apr 2002	25161	3661	934	861	781	809	1123	1571	3067	4719	480	0	0	33	43200
May 2002	27717	2184	739	796	946	1189	1555	2985	5924	565	0	0	0	40	44640
Jun 2002	26352	1635	607	713	843	1061	1606	3154	6790	115	0	0	0	324	43200
Jul 2002	26274	1801	785	853	1039	1405	2071	4399	5940	27	0	0	0	38	44640
Aug 2002	26795	2245	758	712	799	1019	1460	2406	4130	4285	0	0	0	31	44640
Sep 2002	23568	2627	829	829	1010	1132	1471	1874	3891	5801	141	0	0	27	43200
Oct 2002	22823	2498	953	928	1051	1120	1454	2196	3629	5985	1980	0	0	23	44640
Nov 2002	20032	2024	730	723	817	903	1266	1660	2751	6482	5787	0	0	25	43200
Dec 2002	21673	2557	685	724	841	991	1304	1892	3007	6824	4113	0	0	29	44640
Jan 2003	22026	3395	699	675	702	915	1225	1834	3042	7726	2357	0	0	44	44640
Feb 2003	22524	3190	729	829	670	755	948	1505	2638	6135	366	0	0	31	40320
Mar 2003	25579	2760	523	543	629	746	1026	1647	3099	7780	285	0	0	23	44640
Apr 2003	24950	2412	866	829	1026	1126	1694	2802	4100	3367	0	0	0	28	43200
May 2003	26599	2143	726	716	816	1036	1537	2796	5523	2725	0	0	0	25	44640
Total %	54.26	5.83	1.72	1.76	2.09	2.49	3.26	5.02	8.88	10.90	3.23	0.00	0.00	0.58	100
DJF %	48.13	6.37	1.74	1.87	2.50	2.90	3.22	3.76	6.28	15.09	8.14	0.00	0.00	0.03	100
MAM %	57.94	6.90	1.74	1.68	1.84	2.18	3.05	5.00	9.62	9.62	0.42	0.00	0.00	0.50	100
JJA %	59.27	4.18	1.59	1.69	1.96	2.51	3.66	6.70	12.01	6.45	0.00	0.00	0.00	0.45	100
SON %	52.79	6.01	1.85	1.84	2.12	2.44	3.19	4.69	7.73	12.81	4.53	0.00	0.00	1.33	100

From Table 4.5, the following inferences can be made:

- Missing DSDIR values are the result of window cleaning, calibration, tracker stoppages or other infrequent occurrences deeming deletion of data. For this dataset, missing values are only 0.58% of all possible values – a very low number, indicating an excellent performance of the twin pyrheliometers and solar tracker.

- The lower bound (0 $W.m^{-2}$) is satisfied for all values of DSDIR.

- The upper bound (1368 $W.m^{-2}$) is also satisfied – all DSDIR fall well within the required range. In fact, for 14 of the 36 months, DSDIR does not exceed 9/12 of the upper limit of 1368 $W.m^{-2}$.

4.3.1.2 **Sub-procedure 1.2** *(The test for $0\ W.m^{-2} < DSDFS < TOA + 10\ W.m^{-2}$)*

Since TOA radiation is a varying parameter, exact division in frequency classes like in Table 4.5, was not practically possible. Instead, TOA radiation for every datapoint was calculated, and the frequency of DSDFS > TOA + 10 $W.m^{-2}$ listed in Table 4.6.

From Table 4.6, the following deductions are made:

- The majority of months have zero violations. Only 10 of the 36 months have more than 10 datapoint violations per month.

- The month with the most violations (202) is September 2001, followed by May 2001 (53 violations). This is illustrated in Figure 4.1.

- There is apparently no specific preferential season for the occurrence of violations – if the 202 violations of September 2001 are ignored, an even distribution between the seasons exists, DJF having slightly less violations than the other three seasons.

- One must bear in mind that the violations of procedure 1.2 only account for 0.033% of the entire dataset. Therefore, it involves a relatively small and in the most cases, insignificant number of values. Deleting all the datapoints violating this criterium will not add significantly to the amount of already missing data. However, the reason for this violation is investigated in Section 4.3.1.2.1.

Table 4.6 *Frequencies of DSDFS>TOA + 10 W.m^{-2}: June 2000 to May 2003, ref. to sub-procedure 1.2*

Month	Number of violations	Possible data points	Percentage violations
Jun 2000	0	43200	0.0000
Jul 2000	0	44640	0.0000
Aug 2000	0	44640	0.0000
Sep 2000	0	43200	0.0000
Oct 2000	4	44640	0.0090
Nov 2000	4	43200	0.0093
Dec 2000	17	44640	0.0381
Jan 2001	0	44640	0.0000
Feb 2001	2	40320	0.0050
Mar 2001	12	44640	0.0269
Apr 2001	9	43200	0.0208
May 2001	53	44640	0.1187
Jun 2001	2	43200	0.0046
Jul 2001	4	44640	0.0090
Aug 2001	26	44640	0.0582
Sep 2001	202	43200	0.4676
Oct 2001	0	44640	0.0000
Nov 2001	21	43200	0.0486
Dec 2001	8	44640	0.0179
Jan 2002	0	44640	0.0000
Feb 2002	2	40320	0.0050
Mar 2002	0	44640	0.0000
Apr 2002	9	43200	0.0208
May 2002	0	44640	0.0000
Jun 2002	8	43200	0.0185
Jul 2002	19	44640	0.0426
Aug 2002	33	44640	0.0739
Sep 2002	8	43200	0.0185
Oct 2002	35	44640	0.0784
Nov 2002	5	43200	0.0116
Dec 2002	0	44640	0.0000
Jan 2003	6	44640	0.0134
Feb 2003	28	40320	0.0694
Mar 2003	0	44640	0.0000
Apr 2003	0	43200	0.0000
May 2003	1	44640	0.0022
Average	14.39	43800	0.0329
DJF %	0.0162	MAM %	0.0211
JJA %	0.0231	SON %	0.0710

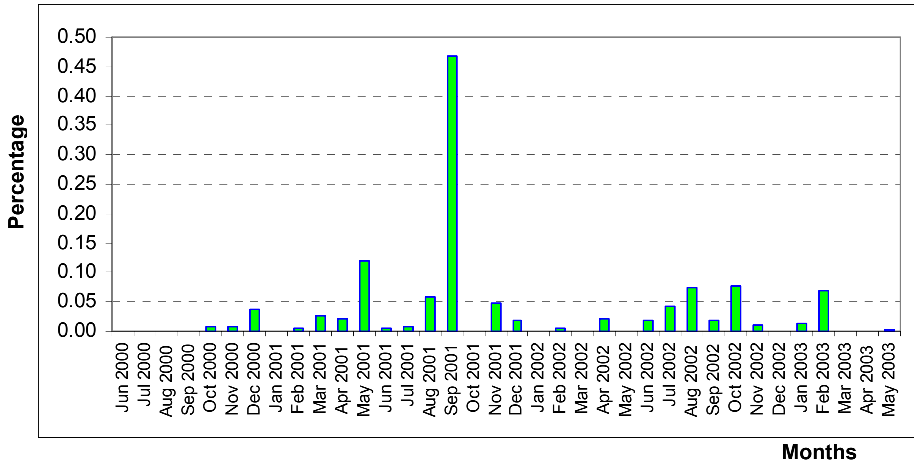

Figure 4.1 *From Table 4.6: Violations of sub-procedure 1.2, expressed as a percentage for all datapoints*

The last column of Table 4.6, viz., the violations of sub-procedure 1.2, are illustrated in Figure 4.1 as a percentage of all datapoints. The anomalous occurrence of violations of sub-procedure 1.2 during the month of September 2001 within a 36 month-period of observations, prompts the investigation in Section 4.3.1.2.1.

4.3.1.2.1 Closer inspection of violations of sub-procedure 1.2

As a result of diffuse radiation being present in the atmosphere a few minutes before sunrise and a few minutes after sunset, while TOA radiation equals zero, DSDFS and TOA radiation are bound to have a difference at the start and end of every day. This, however, is less than 10 $W.m^{-2}$ as a general rule, and does not explain the violations of sub-procedure 1.2.

Closer inspection reveals, that a difference of more than 10 $W.m^{-2}$ between DSDFS and TOA radiation does, however, appear near the end or the beginning of a specific day. The quantities DSDFS and TOA radiation are small at the beginning and end of a day, and a difference of 10 $W.m^{-2}$ between them is significant. If the time of observation is slightly offset between DSDFS and TOA, the difference can easily become larger than 10 $W.m^{-2}$.

A closer look at September 2001 reveals that there were small discrepancies between DSDFS and TOA radiation for only the first three days of the month, and the system recovered during the night of 3-4 September 2001. No anomalous system behaviour was reported or observed for these specific days. To illustrate this phenomenon, consider Figure 4.2 - a graph of TOA + 10 $W.m^{-2}$ and DSDFS for 2 September 2001.

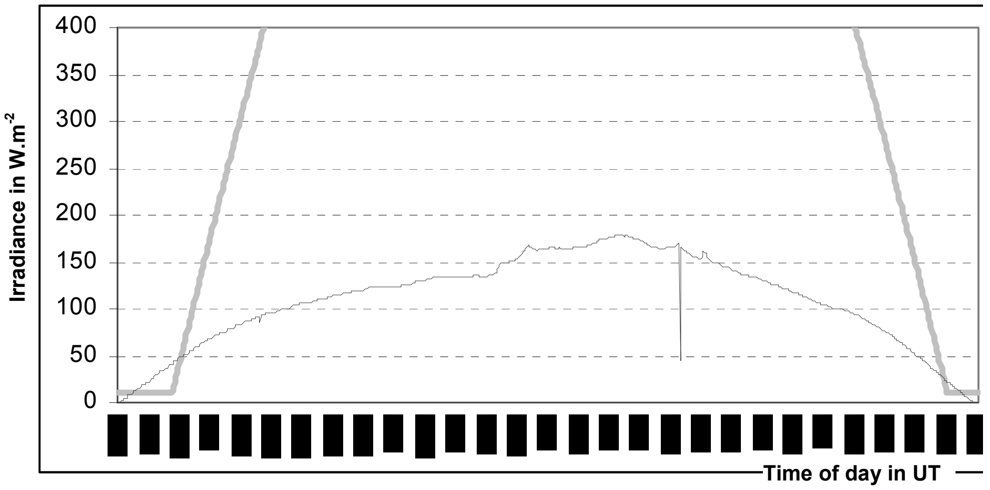

Figure 4.2 *TOA radiation + 10 $W.m^{-2}$ (thick grey line) and diffuse radiation (thin black line) for 2 September 2001*

A strange double offset w.r.t. the time at the beginning and the end of the day is observed, with DSDFS overshooting (TOA+10 $W.m^{-2}$) by a substantial amount between 04:04 UT and

04:39 UT at the beginning of the day, as well as 16:00 UT and 16:09 UT at the end of the day. (Sunsrise and sunset times at De Aar for 2 September 2001 are 04:39 UT and 16:09 UT, respectively). One last remark is, that this data possibly could have been incorrectly date-stamped by the logger (i.e. the DSDFS data attributed to 2 September 2001 actually belongs to a another, longer day). If the sunrise and sunset times were further spaced apart, more space would have allowed the DSDFS profile from sunrise to sunset, to fit in the space below TOA radiation + 10 $W.m^{-2}$. An accurate on-site GPS receiver updates the system time and transfers the time to the logger at least three times a day, so it is highly unlikely that this could have happened.

4.3.1.3 Sub-procedure 1.3 *(The test for $0\ W.m^{-2} < DSGL2 < 1368\ W.m^{-2}$)*

In Table 4.7, frequencies of DSGL2 datapoints with respect to twelfths of 1368 $W.m^{-2}$ – like in Table 4.5, are presented. The only exception is, that a column making provision for the number of datapoints larger than 1368 $W.m^{-2}$, is added. This column is shaded to indicate that it represents the number of datapoint violations of sub-procedure 1.3.

From Table 4.7 it is noted that:

- Out of all DSGL2 datapoints, the small percentage of 0.35 % are missing. The reason for missing values are generally the same for DSDIR (calibration, maintenance, cleaning), except tracker failures, since DSGL2 does not require solar pointing or tracking. This emphasizes that the global pyranometer was in excellent working condition.

- A number of one-minute global radiation values (89) are already violating this crude criterium. (Shaded column in table 4.7.) Even though this number only accounts for 0.0056 % of all the datapoints, it is still significant since violation of this crude criteria is not expected in BSRN. A closer investigation is performed in Section 4.3.1.3.1.

- The violations are restricted to summer months, which is logical since DSGL2 is expected to reach its highest value, approach and/or temporarily overtake the boundary of 1368 $W.m^{-2}$. The month of January 2001 (investigated in Section 4.3.1.3.1) has the most violations (26), followed by November 2001 (16).

Table 4.7 *Frequency distribution of DSGL2: June 2000 to May 2003 ref.to sub-procedure 1.3. Shaded area represents violations.*

	Data bins in W.m^{-2}														
Month	<= 0	0 to 114	115 to 228	229 to 342	343 to 456	457 to 570	571 to 684	685 to 798	799 to 912	913 to 1026	1027 to 1140	1141 to 1254	1255 to 1368	> 1368 W.m^{-2}	Missing	Possible
Jun 2000	24487	4187	2632	2391	2470	3785	3126	74	3	0	0	0	0	0	45	43200
Jul 2000	24989	4173	2713	2410	2562	3679	3560	160	44	1	0	0	0	0	349	44640
Aug 2000	23829	3925	2727	2603	2821	3679	4445	579	4	0	0	0	0	0	28	44640
Sep 2000	21625	5412	2687	2254	2055	2082	2686	3230	932	49	0	0	0	0	188	43200
Oct 2000	20490	3697	2427	2204	2220	1994	1888	1894	2390	3729	1031	0	0	0	676	44640
Nov 2000	18377	4407	2422	2088	1784	1614	1595	1645	1905	2155	3721	232	28	0	1227	43200
Dec 2000	18139	4352	2403	2130	1800	1758	1768	1849	2109	2600	4571	975	141	13	32	44640
Jan 2001	18720	3967	2388	2104	1949	1814	1717	1772	2060	2332	4143	1464	164	26	20	44640
Feb 2001	18114	3762	2267	2031	1658	1589	1549	1540	1864	2481	2938	441	61	2	23	40320
Mar 2001	21693	4354	2484	2378	1945	1848	1710	1819	2529	3014	654	61	11	0	140	44640
Apr 2001	22612	5525	3497	2426	1784	1718	1686	2312	1163	378	39	1	0	0	59	43200
May 2001	24515	4121	2940	2599	2494	2817	4105	865	90	7	1	0	0	0	86	44640
Jun 2001	24571	4406	2818	2531	2460	3700	2637	44	0	0	0	0	0	0	33	43200
Jul 2001	25293	5097	2885	2409	2373	2929	3288	250	78	6	0	0	0	0	32	44640
Aug 2001	23851	3622	2199	2088	2451	2680	3651	3614	427	32	2	0	0	0	23	44640
Sep 2001	21456	4880	2602	2122	1924	1758	2032	2826	2585	387	25	0	0	0	603	43200
Oct 2001	20905	4233	2771	2231	2054	1904	2209	2848	2200	2497	589	139	11	1	48	44640
Nov 2001	18379	5229	2705	2683	2239	1773	1545	1473	1660	1780	2911	552	223	16	32	43200
Dec 2001	18032	4522	2474	2261	1833	1708	1707	1760	2044	2459	4591	994	219	10	26	44640
Jan 2002	18630	4260	2441	2109	1903	1672	1709	1852	2036	2404	4404	1029	150	12	29	44640
Feb 2002	18023	3481	2085	1896	1863	1717	1633	1733	2037	2985	2678	147	14	1	27	40320
Mar 2002	21661	3989	2734	2228	2066	1954	1861	2164	3127	2013	680	80	10	0	73	44640
Apr 2002	22527	4404	2860	2400	2049	2078	2144	3105	1514	68	18	0	0	0	33	43200
May 2002	24836	4779	2963	2312	2312	2719	3837	813	38	1	0	0	0	0	30	44640
Jun 2002	24506	4030	2876	2436	2494	3555	3222	41	0	0	0	0	0	0	40	43200
Jul 2002	24941	4083	2583	2532	2665	3457	4093	210	41	7	0	0	0	0	28	44640
Aug 2002	23841	4834	2504	2552	2247	2147	2936	2953	524	68	8	1	0	0	25	44640
Sep 2002	21586	3829	2450	2297	2051	1916	1986	2591	3559	835	60	9	1	0	30	43200
Oct 2002	20516	4412	2115	1878	1854	2000	1954	2117	2471	4026	1145	105	17	0	30	44640
Nov 2002	18334	3821	2204	1986	1736	1856	1757	1919	2217	2593	4223	484	34	1	35	43200
Dec 2002	18571	4662	2835	2291	1952	1784	1634	1655	1991	2292	4060	772	110	7	24	44640
Jan 2003	19368	4581	2509	2305	1859	1693	1656	1712	2105	3062	2986	79	1	0	724	44640
Feb 2003	18113	5148	2566	2096	1602	1435	1413	1472	1621	2406	1587	0	0	0	861	40320
Mar 2003	21601	5530	2800	1972	1683	1557	1695	1975	3203	2339	100	0	0	0	185	44640
Apr 2003	22667	4007	2629	2403	2109	2044	2401	3881	935	74	18	0	0	0	28	43200
May 2003	24763	4397	2756	2446	2392	3073	4241	544	5	0	0	0	0	0	23	44640
Total %	49.1	10	5.9	5.2	4.79	5.16	5.5	3.8	3.3	3.1	3.1	0.6	0.1	0.0056	0.3499	100
DJF %	42.80	10.00	5.70	5.00	4.24	3.92	3.80	4.00	4.60	5.95	8.30	1.50	0.20	0.00	0.00	100
MAM %	52.10	10.40	6.50	5.30	4.75	4.99	6.00	4.40	3.20	1.99	0.40	0.00	0.00	0.00	0.00	100
JJA %	55.50	9.68	6.00	5.50	5.68	7.46	7.80	2.00	0.30	0.03	0.00	0.00	0.00	0.00	0.00	100
SON %	46.50	10.10	5.70	5.00	4.53	4.27	4.40	5.00	5.10	4.57	3.80	0.80	0.10	0.00	0.00	100

4.3.1.3.1 Inspection of January 2001

Figure 4.3 depicts DSGL2 datapoints for the month of January 2001, with all the night-time (zero) values removed for brevity. The Y-axis is divided in the same increments as Table 4.7.

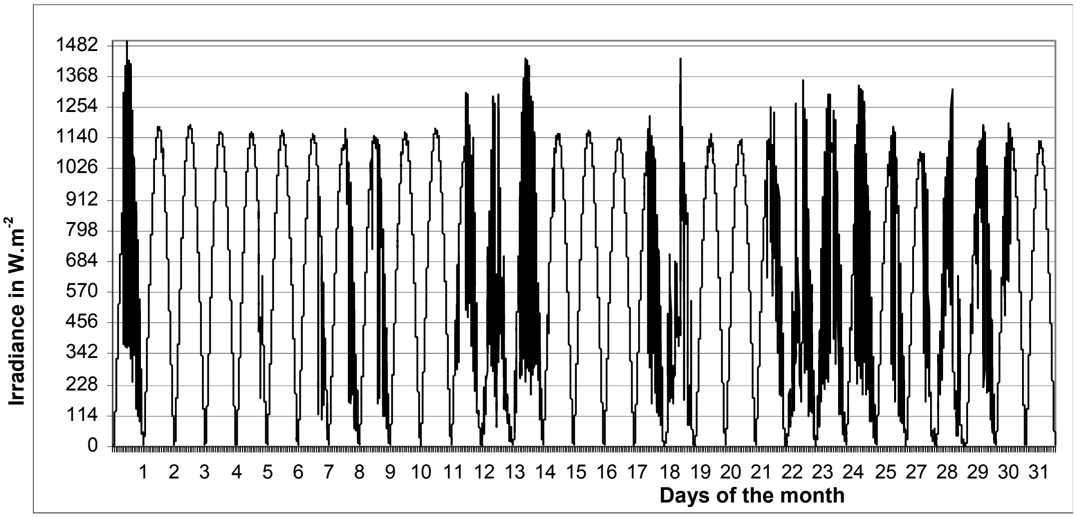

Figure 4.3 *DSGL2 datapoints for January 2001*

In Figure 4.3, note that clear-sky days or days with little or no clouds in the sky, are identified as days with smooth curves. Examples are: the 2nd to 10th as well as 15th to 17th, 20th and 21st and 31st. Days with clouds in the sky (partly cloudy to overcast) renders a distinctive unsmooth pattern, such as the 1st, 12th to 14th,18th, 19th and 22nd to 30th. It follows that the daily maximum value of DSGL2 for clear sky days, is remarkably consistent (ca. 1140 $W.m^{-2}$), whilst the 1368 $W.m^{-2}$ boundary is violated or almost violated on days with cloud interference. The best example is the 1st. (Daily maximum of DSGL2 > 1482 $W.m^{-2}$).

A closer inspection is now given to one individual day having the highest daily DSGL2 maximum for January 2001 (Figure 4.4). Note that the same increment of 114 $W.m^{-2}$ is also used in this graph, but the Y-axis extends two increments further than Figure 4.3.

From Figure 4.4, it can be deduced that cloudless skies, indicated by smooth profiles of DSGL2 (global), DSDIR (direct) and DSDFS (diffuse) existed since sunrise (03:25 UT) up to about 06:25 UT, after which global (thin black line) overshoots the boundary of 1368 $W.m^{-2}$ on a few occasions during the middle of the day. It reached a momentary peak of 1590 $W.m^{-2}$ roughly at 10:00 UT and occasionally also overshot the TOA value, mostly in the presence of low direct radiation (thick grey line). This pattern is maintained sunset at 17:30 UT. Thick and heavy clouds were reported during the afternoon, but no precipitation was recorded by the Automatic Weather Station (AWS).

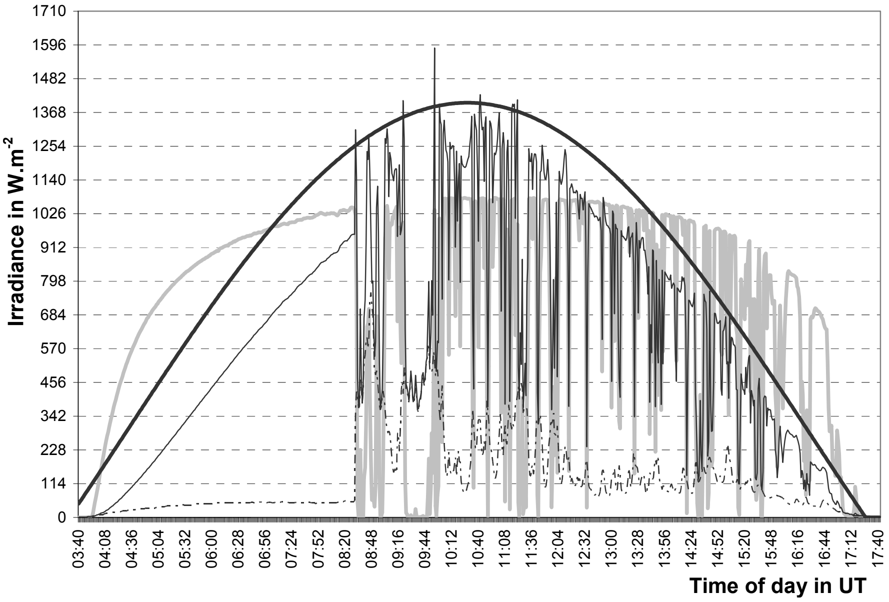

Figure 4.4 *Diurnal variation of TOA radiation (thick smooth black line), DSDIR (thick grey line), DSGL2 (thin black line) and DSDFS (dashed line) on 1 January 2001*

Consider another example on 20 November 2001 (Figure 4.5), The same scale and definitions as in Figure 4.4 are used.

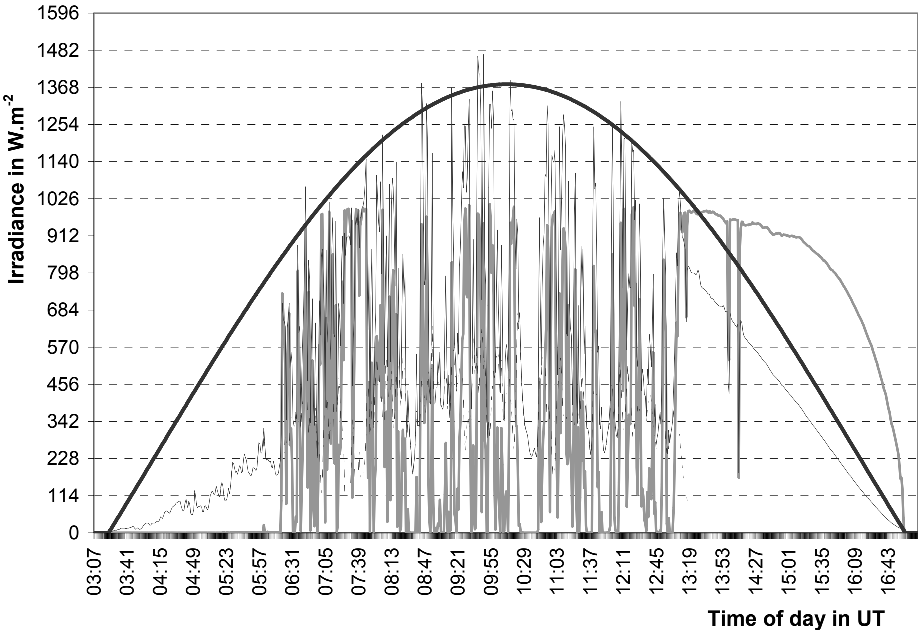

Figure 4.5 *Diurnal variation of TOA radiation (thick smooth black line), DSDIR (thick grey line), DSGL2 (thin black line) and DSDFS (dashed line) on 20 November 2001*

On this day, overcast conditions prevailed since sunrise (03:16 UT) until roughly 06:00 UT (global and diffuse graphs run concurrently with direct being equal to zero). Thereafter the same pattern, as shown in Figure 4.4 was followed: moderate diffuse values prevailed, TOA radiation was overshot by global in numerous occasions. The situation restored at about 13:00 UT to smooth lines for global and direct, with diffuse having relatively low values (cloudless sky) until sunset at 17:03 UT.

The presence of clouds in hours around noon seems to be the responsible factor in these two cases. Clouds have both a diffusing and blocking effect on radiation. When broken cloud passes the pyranometer, the full direct beam is measured in intervals with diffuse radiation, plus reflected radiation from clouds, added in the same interval. The fact, that broken cloud creates streaks of intense sunlight and shadow on the pyranometer, forcing it to respond quickly to rapidly changing signals, seems to aggravate the situation. Note that the anomalously high peak-values are short bursts, i.e. seldom lasting for more than one minute at a time.

Since only 0.35 % of all global irradiance values are violating sub-procedure 1.3, simple deletion of these values is not deemed to make a significant impact on the dataset.

4.3.1.4 Sub-procedure 1.4 *(The test for $50 \, W.m^{-2} < LWD < 700 \, W.m^{-2}$)*

The quoted limits for LWD in this sub-procedure corresponds to sky blackbody temperatures of respectively -100°C and 60°C, respectively. In Table 4.8, the distribution per month is listed as frequencies as in Tables 4.5 and 4.7, but with 50 $W.m^{-2}$ bins.

From Table 4.8 it is deduced that:

- Missing values (a minimum of 23 per month) are due to dome cleaning, maintenance, as well as tracker stoppages (the shading device is necessary to render good shading for correct application of the pyrgeometer Equation), and download failures as a result of lightning strikes. One such incident (the highest) occurred in December 2002.

- The total number of missing LWD datapoints is a relatively high percentage (1.132%), but this is biased by the high number of 3718 missing datapoints in December 2002. The average missing value percentage for the other 35 months is only 0.389 % - which is in line with DSGL2, DSDIR and DSDFS.

- No boundary violations - all the LWD datapoints fall within 50 $W.m^{-2}$ to 700 $W.m^{-2}$.

- In line with expected seasonal trends, datapoints are clustered in a lower range, 201 W.m^{-2} to 400 W.m^{-2}, during the winter months. During the summer months, the highest concentration of datapoints is between 251 W.m^{-2} to 500 W.m^{-2}.

- The three-year LWD record minimum is 218 W.m^{-2} and the record maximum is 524 W.m^{-2} – which are respectively at cushion margins of 168 W.m^{-2} and 176 W.m^{-2}, respectively, within the boundaries of the rough Procedure 1.4.

Table 4.8 *Frequency distribution of LWD: June 2000 to May 2003 ref. to sub-procedure 1.4*

	Data bins in W.m^{-2}											
Month	101-150	151-200	201-250	251-300	301-350	351-400	401-450	451-500	501-550	551-600	Missing	Total
Jun 2000	0	0	4908	28435	8054	1769	0	0	0	0	34	43200
Jul 2000	0	0	7548	29246	4476	387	0	0	0	0	2983	44640
Aug 2000	0	0	3970	28177	11358	1109	1	0	0	0	25	44640
Sep 2000	0	0	3745	20025	15201	4202	0	0	0	0	29	43200
Oct 2000	0	0	1083	16193	18751	7919	670	0	0	0	24	44640
Nov 2000	0	0	8	8611	24413	9157	961	26	1	0	23	43200
Dec 2000	0	0	0	3285	25244	14082	1971	30	0	0	28	44640
Jan 2001	0	0	0	4946	21616	14927	3102	19	0	0	30	44640
Feb 2001	0	0	0	2230	18201	16930	2898	19	1	0	41	40320
Mar 2001	0	0	0	1729	18876	21400	2589	19	0	0	27	44640
Apr 2001	0	0	0	8955	20115	13244	860	0	0	0	26	43200
May 2001	0	0	0	2403	22588	16014	3607	6	0	0	28	44640
Jun 2001	0	0	2058	29838	10041	1239	0	0	0	0	24	43200
Jul 2001	0	0	10165	25412	8294	740	0	0	0	0	29	44640
Aug 2001	0	0	7321	28312	8427	550	0	0	0	0	30	44640
Sep 2001	0	0	2322	22616	12454	5773	7	0	0	0	28	43200
Oct 2001	0	0	272	11149	21450	10257	604	3	0	0	905	44640
Nov 2001	0	0	0	6101	19214	15757	2098	0	0	0	30	43200
Dec 2001	0	0	0	3679	26615	12115	2176	12	0	0	43	44640
Jan 2002	0	0	0	6345	21506	13114	3563	57	5	0	50	44640
Feb 2002	0	0	0	2787	20753	13645	3089	19	0	0	27	40320
Mar 2002	0	0	310	3883	19726	18787	1887	11	0	0	36	44640
Apr 2002	0	0	591	9348	19914	12618	701	5	0	0	23	43200
May 2002	0	0	4426	25689	10726	3759	10	0	0	0	30	44640
Jun 2002	0	0	9620	27080	5812	663	0	0	0	0	25	43200
Jul 2002	0	0	7992	30121	6253	246	0	0	0	0	28	44640
Aug 2002	0	0	7165	22689	10420	4340	0	0	0	0	26	44640
Sep 2002	0	0	889	21239	18454	2572	21	2	0	0	23	43200
Oct 2002	0	0	1475	15364	16822	8550	980	9	0	0	1440	44640
Nov 2002	0	0	1503	16177	19298	5316	860	19	1	0	26	43200
Dec 2002	0	0	1045	16906	12732	2233	6	0	0	0	3718	44640
Jan 2003	0	0	0	1440	21165	19079	2901	22	4	0	29	44640
Feb 2003	0	0	0	167	9797	22021	7468	97	5	0	765	40320
Mar 2003	0	0	0	2581	22872	16965	2198	0	0	0	24	44640
Apr 2003	0	0	0	9285	19544	13822	522	0	0	0	27	43200
May 2003	0	0	1123	24303	17509	1680	2	0	0	0	23	44640
Total %	0	0	5.048	32.78	37.364	20.75	2.907	0.02	0.001	0	1.132	100
DJF %	0.00	0.00	0.28	11.11	47.23	34.08	7.23	0.07	0.00	0.00	3.32	100
MAM %	0.00	0.00	1.63	22.19	43.28	29.77	3.13	0.01	0.00	0.00	0.01	100
JJA %	0.00	0.00	2.89	35.19	42.52	17.80	1.59	0.02	0.00	0.00	0.60	100
SON %	0.00	0.00	15.41	63.22	18.58	2.80	0.00	0.00	0.00	0.00	0.76	100

4.3.1.5 Sub-procedure 1.5 *(The test for* $50\ W.m^{-2} < UL < 700\ W.m^{-2}$*)*

At the time of preparation of this document, no UL data was measured, hence no data is available for verification. The focus is now turned to Procedure 2.

4.3.2 Procedure 2 – "Extremely rare"

The criteria is stricter under the "extremely rare" tests compared to Procedure 1, hence some datapoints that passed the test in Procedure 1, are bound to be identified as errors in Procedure 2.

4.3.2.1 Sub-procedure 2.1 *(The test for DSGL2 < TOA if Z < 80°, and DSGL2 < TOA + 0.56(Z-93.3°)² for Z ≥ 80°)*

This procedure involves calculating TOA radiation as well as Z for every DSGL2 datapoint. For brevity and simplicity, a new variable is defined: ***DSTM***. The definition is as follows:

If $Z < 80°$ then DSTM = DSGL2 - TOA
if $Z \geq 80°$ then DSTM = DSGL2 – (TOA + 0.56 (Z - 93.3°)²)
Violations of the sub-procedure are therefore cases where DSTM < 0.

Consider Table 4.9, showing frequencies of violations in the same pattern as Table 4.6. Note, that the complex definition of DSTM, plus the absence of fixed borders, does not allow for a frequency distribution table like (for example) Table 4.8.

From Table 4.9, it is deduced that:

- During 33 out of 36 months, datapoints with DSTM < 0, occurred with monthly frequencies between 3% and 6%. They were evenly distributed between seasons.

- June 2002 has the highest number (5.77%) and June 2001 (0.07%) the lowest.

- Overall average violation of sub-procedure 2.1 is 4.10%, which is a significant figure. Therefore, these datapoints violating the criterium, cannot all be summarily deleted. Reasons for the violations are seeked in Sections 4.3.2.1.1 and 4.3.2.1.2.

Table 4.9 *Frequencies of DSTM < 0: June 2000 to May 2003, ref.to sub-procedure 2.1*

Month	Number of violations	Possible data points	Percentage violations
Jun 2000	2357	43200	5.46
Jul 2000	2311	44640	5.18
Aug 2000	2	44640	0.00
Sep 2000	1770	43200	4.10
Oct 2000	1881	44640	4.21
Nov 2000	2108	43200	4.88
Dec 2000	2007	44640	4.50
Jan 2001	2040	44640	4.57
Feb 2001	1700	40320	4.22
Mar 2001	1671	44640	3.74
Apr 2001	1560	43200	3.61
May 2001	251	44640	0.56
Jun 2001	31	43200	0.07
Jul 2001	1669	44640	3.74
Aug 2001	2268	44640	5.08
Sep 2001	1828	43200	4.23
Oct 2001	2153	44640	4.82
Nov 2001	1617	43200	3.74
Dec 2001	2417	44640	5.41
Jan 2002	2060	44640	4.61
Feb 2002	1833	40320	4.55
Mar 2002	1720	44640	3.85
Apr 2002	1549	43200	3.59
May 2002	1726	44640	3.87
Jun 2002	2494	43200	5.77
Jul 2002	2497	44640	5.59
Aug 2002	1978	44640	4.43
Sep 2002	1792	43200	4.15
Oct 2002	1852	44640	4.15
Nov 2002	2317	43200	5.36
Dec 2002	1796	44640	4.02
Jan 2003	1119	44640	2.51
Feb 2003	1200	40320	2.98
Mar 2003	1634	44640	3.66
Apr 2003	1604	43200	3.71
May 2003	2172	44640	4.87
Average	1750	43800	4.104
DJF %	4.159	JJA %	4.362
MAM %	3.494	SON %	4.405

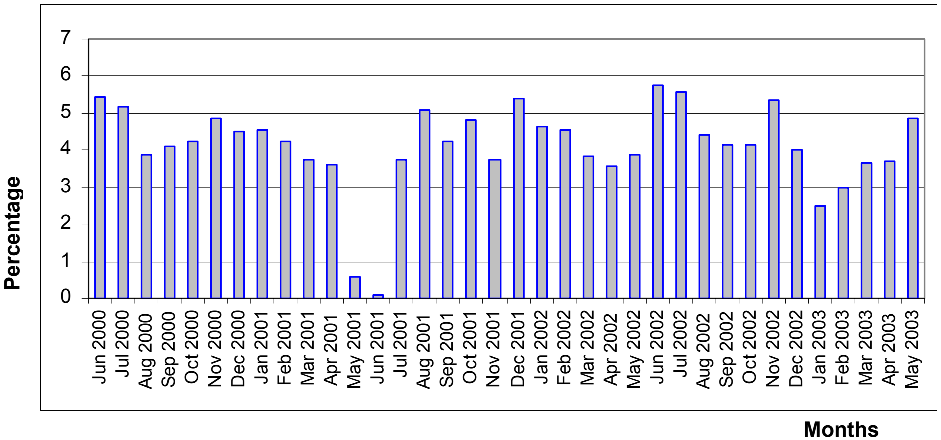

Figure 4.6 *From Table 4.9: violations of sub-procedure 2.1, expressed as a percentage for all datapoints*

The percentage violations per month of sub-procedure 2.1 is illustrated in Figure 4.6. Two extreme months, May 2001 (second-lowest) and June 2002 (highest), are now investigated.

4.3.2.1.1 **May 2001** *(Second-lowest number of boundary violations: 0.56%).*

Although June 2001 had the absolute minimum number of violations, viz. 0.07%, it was decided to use the second-lowest month, May 2001, for illustrating violations of sub-procedure 2.1. because the very few violations occurring during June 2001 could not satisfactorily illustrate reasons for violations.

Figure 4.7 depicts DSTM in a time-series for the complete month of May 2001, but the night-time datapoints (zero), are removed for clarity. Note that smooth curves are associated with clear, cloud-free days, while the presence of clouds disturbs the pattern.

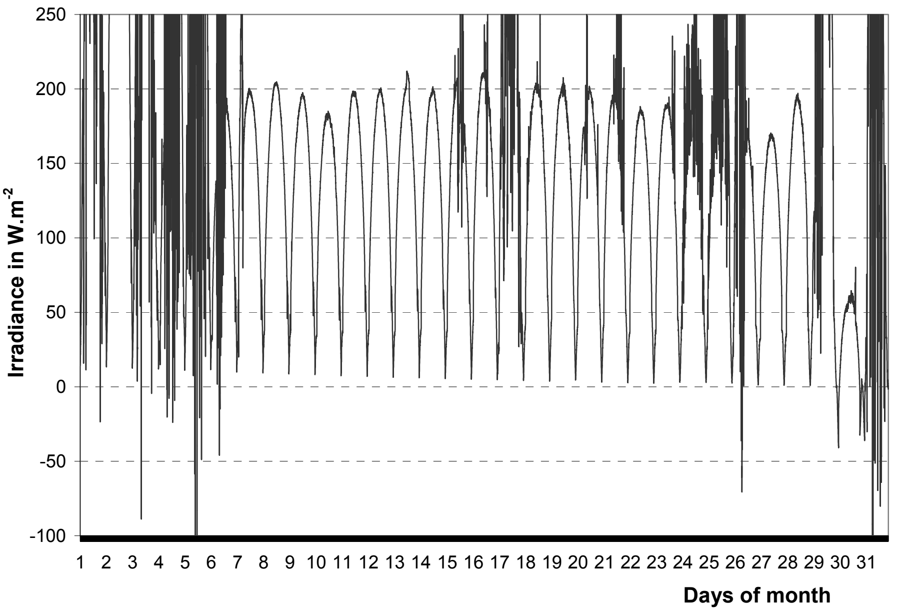

Figure 4.7 *Time-series for DSTM: May 2001*

From Figure 4.7 it can be deduced that datapoints for which DSTM<0 are restricted to certain days. The lower and upper boundaries for DSTM seem to follow 0 W.m^{-2} and 200 W.m^{-2}, respectively, on clear days. On days with broken clouds, the lower and upper boundaries are both violated on the same day, like the 5th and 31st. A total of 251 datapoints do not satisfy the criterium DSTM<0, being 0.56% of all the values for that specific month.

Continuing with May 2001, the dependence of DSTM with respect to the solar angle is investigated in Figure 4.8, which is a scatter diagram of DSTM versus cos$Z$. The majority of night-time values were removed for simpilcity. The numbered notations on the graph are:

1. Point of sunrise: Z = 90.83°, therefore cos$Z$ = -0.014

2. Point of maximum solar elevation angle (Z and therefore, cos$Z$ is a minimum)

3. Inflection line, where the Z=80° (and cos$Z$ = 0.17) criterium separates DSTM values for low solar angles from DSTM values for high solar angles

4. For high solar elevation, there are also a few loose values of DSTM<0

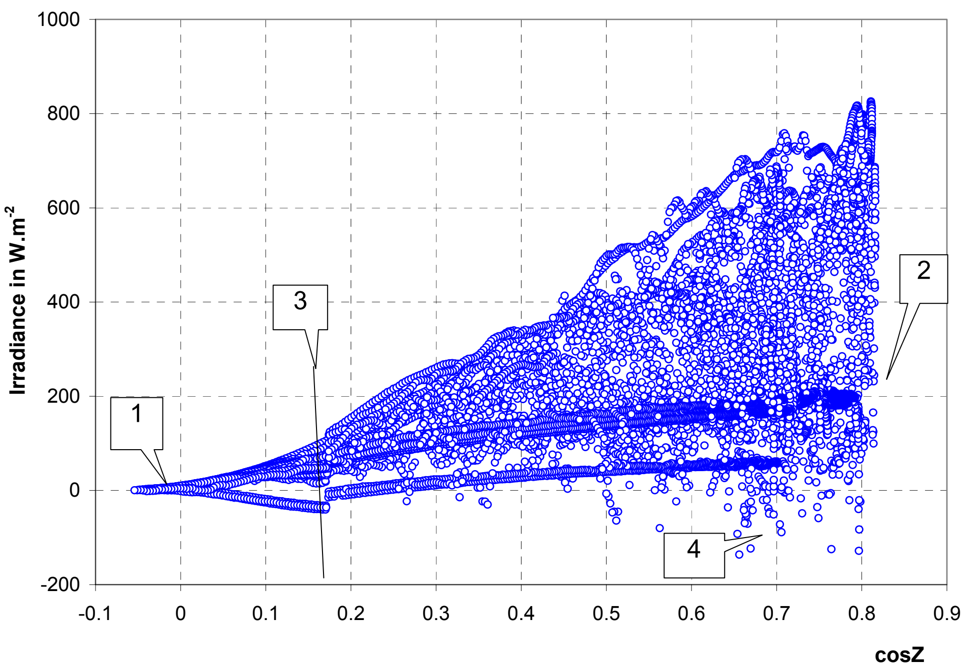

Figure 4.8 *DSTM versus cos$Z$: May 2001*

It follows from Figure 4.8, that the majority of DSTM datapoints < 0 are lying to the left of the inflection line (3), starting at zero for cos$Z$ = 0° and becoming significantly less than zero as cos$Z$ increases up to about 0.17. There are, however, also single values of DSTM less than zero for relatively high values of cos$Z$, in other words, high solar angles.

It should be kept in mind that this is the “second-best” month, hence the remaining 34 months in the 36 month- record have more violations of sub-procedure 2.1.

4.3.2.1.2 June 2002 *(Highest number of boundary violations: 5.77%)*

In Figure 4.9, similar to Figure 4.7, depicts a time-series of DSTM for June 2002.

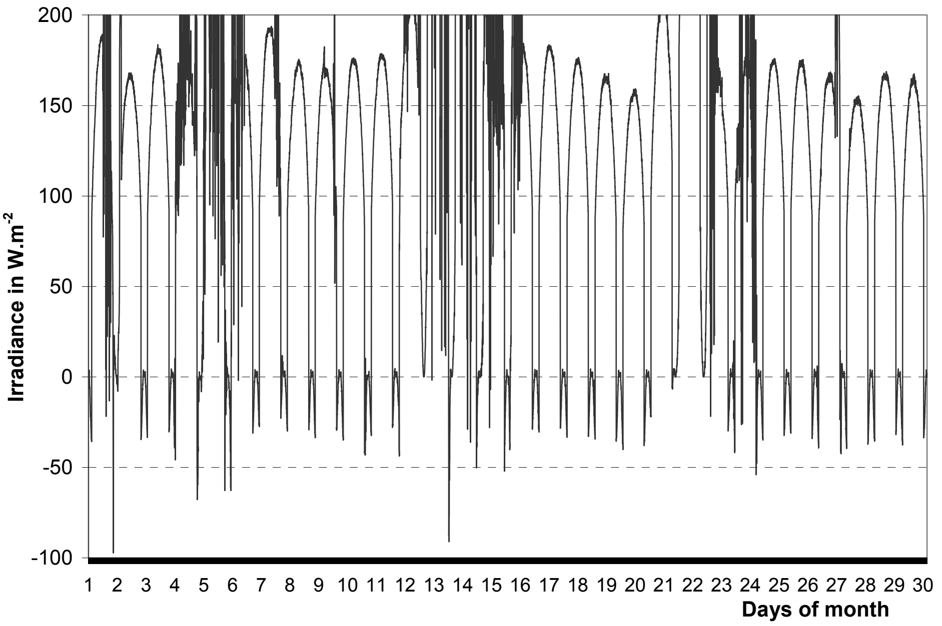

Figure 4.9 *Time-series for DSTM: June 2002*

For clear days in Figure 4.9, similar to Figure 4.7, a virtual maximum of 150 W.m^{-2} to 190 W.m^{-2} is maintainted, but with the exception that DSTM<0 violating days are now not only restricted to days where the virtual upper boundary is violated, but also portions of almost every day on June 2002.

In Figure 4.10, the scatter diagram of DSTM versus cos$Z$ is featured in a similar way as in Figure 4.8. The numbered notations for Figure 4.10 are:

1. Point of sunrise: Z = 90.83°, therefore cos$Z$ = -0.014
2. Point of maximum solar elevation angle (Z and therefore, cos$Z$, is a minimum)
3. Inflection line, where the Z=80° (and hence cos$Z$=0.17) criterium separates DSTM values for low solar angles from DSTM values for high solar angles
4. For high solar elevation, there are also a few loose values of DSTM<0

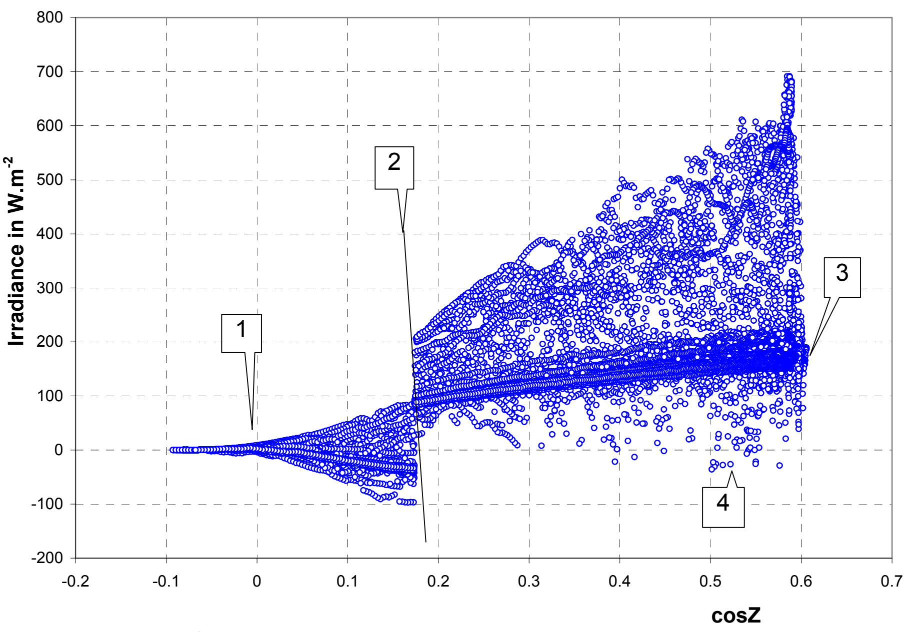

Figure 4.10 *DSTM versus cosZ: June 2002*

From Figure 4.10 it follows that DSTM datapoints less than zero are concentrated to the left of the inflection line (3), starting at zero for cos$Z$ =0, and becoming significantly lower than zero as cos$Z$ increases up to the inflection point, cos$Z$ =0.17. The relatively few values of DSTM less than zero for relatively high values of cos$Z$ are less in Figure 4.10, than Figure 4.8.

The values for DSTM < 0 for low values of cos$Z$ are attributed to a global pyranometer that reads values which are too high compared to an expected TOA value for these angles. This could either be the result of a skew pyranometer, tilted towards one end, or one that is incorrectly calibrated. The checking of pyranometer balance as part of the daily routine, should therefore be given a high priority in an effort to reduce these errors.

4.3.2.2 Sub-procedure 2.2 *(The test for USR < 0.95 * DSGL2)*

Since USR is not a basic BSRN measurement quantity that is measured at the De Aar BSRN station, no real data is available for the evaluation of this sub-procedure.

4.3.2.3 Sub-procedure 2.3 *(The test for DSDFS < 700W.m^{-2})*

Similar to Table 4.8, the DSDFS datapoints in Table 4.10 are placed in bins of 50 W.m^{-2} intervals. The violations of sub-procedure 2.3 are shaded.

Table 4.10 *Frequency distribution of DSDFS: June 2000 to May 2003, ref.to sub-procedure 2.3. Shaded area represents violations*

Months	Data bins in W.m^{-2}																Missing	Possible
	0	1-50	51-100	101-150	151-200	201-250	251-300	301-350	351-400	401-450	451-500	501-550	551-600	601-650	651-700	> 700		
Jun 2000	24512	11084	5216	1108	635	304	210	70	13	3	0	0	0	0	0	0	45	43200
Jul 2000	25013	8618	7863	1336	595	391	307	134	33	3	0	0	0	0	0	0	347	44640
Aug 2000	23069	6111	8239	2837	888	651	468	457	289	127	40	4	2	0	0	0	1458	44640
Sep 2000	21635	5570	8455	2487	1360	983	754	603	465	313	249	155	105	28	9	1	28	43200
Oct 2000	20504	5991	7801	2318	1816	1225	960	705	752	530	261	84	38	9	2	0	1644	44640
Nov 2000	17579	7135	8298	1694	1400	1114	991	629	583	336	235	141	111	42	15	3	2894	43200
Dec 2000	18206	5480	10403	2995	2143	2248	1154	604	495	410	262	96	83	47	11	1	32	44640
Jan 2001	18853	10633	8459	1938	1301	923	619	583	417	341	222	117	74	80	15	3	32	44640
Feb 2001	18265	7750	7553	1967	1458	1208	770	422	316	250	192	111	35	0	0	0	23	40320
Mar 2001	21083	8035	6078	2213	1611	1235	946	650	522	315	210	170	59	18	0	1	1494	44640
Apr 2001	22564	7907	4357	2640	2109	1197	970	703	402	182	72	30	6	1	1	0	59	43200
May 2001	24449	11640	4017	1298	1156	784	472	328	210	86	33	16	1	0	0	0	150	44640
Jun 2001	24010	9780	5582	1204	739	420	294	48	32	7	0	0	0	0	0	0	1084	43200
Jul 2001	25232	9473	5102	1852	1308	834	474	225	75	28	6	0	0	0	0	0	31	44640
Aug 2001	23806	12302	4806	1158	877	596	442	243	170	161	49	7	0	0	0	0	23	44640
Sep 2001	21456	6940	6810	2530	1753	1189	759	466	327	189	74	51	36	10	2	1	607	43200
Oct 2001	20959	8027	7888	3441	1763	1148	656	345	136	117	68	21	3	0	0	0	68	44640
Nov 2001	18483	8366	6630	1862	1481	1565	1400	1164	837	551	398	238	125	53	15	0	32	43200
Dec 2001	18131	9828	10405	2014	1116	1064	831	562	332	203	89	25	11	2	1	0	26	44640
Jan 2002	18708	9426	9761	1957	1308	1062	834	536	409	252	185	98	57	17	2	0	28	44640
Feb 2002	18098	7955	8642	1473	1142	920	525	462	443	324	160	92	38	9	0	0	37	40320
Mar 2002	21717	8905	5957	2485	1950	1295	684	481	350	231	143	58	23	17	1	0	343	44640
Apr 2002	22581	8687	4787	2109	1338	1008	1101	672	395	220	185	84	0	0	0	0	33	43200
May 2002	24874	11204	3470	2124	1140	910	413	301	157	15	2	0	0	0	0	0	30	44640
Jun 2002	24546	12611	2820	1065	750	568	301	138	64	13	0	0	0	0	0	0	324	43200
Jul 2002	24983	9943	6101	1705	753	471	289	270	76	19	1	1	0	0	0	0	28	44640
Aug 2002	23883	10853	5619	1403	1251	784	471	245	93	14	0	0	0	0	0	0	24	44640
Sep 2002	21639	5978	9013	2043	1356	948	784	597	440	261	81	28	2	0	0	0	30	43200
Oct 2002	20574	7233	8717	2098	1677	1276	847	642	492	353	360	253	78	0	0	0	40	44640
Nov 2002	18398	8370	10854	1705	870	688	624	433	337	281	296	217	84	4	4	0	35	43200
Dec 2002	18700	10427	7904	2464	1619	1018	694	579	442	291	206	103	52	70	28	4	39	44640
Jan 2003	19449	7862	9275	2289	1610	1461	951	568	333	271	215	88	94	26	17	13	118	44640
Feb 2003	18164	4984	7943	2627	1946	1728	1063	675	397	358	240	72	74	9	0	0	40	40320
Mar 2003	21660	9247	7840	2020	1429	839	566	393	266	172	129	43	6	0	0	0	30	44640
Apr 2003	22724	7846	4711	2936	1601	1279	873	630	377	179	15	0	0	0	0	0	29	43200
May 2003	24815	9613	6327	1803	966	465	325	173	106	24	0	0	0	0	0	0	23	44640
%	39.5	16.605	12.53	3.71	2.5062	1.837	1.27	0.86	0.568	0.362	0.23	0.122	0.054	0.02	0.005	0.0014		
DJF %	42.87	19.13	20.68	5.09	3.54	3.00	1.91	1.28	0.92	0.69	0.46	0.21	0.13	0.07	0.02	0.01		
MAM %	52.22	21.01	12.03	4.97	3.37	2.28	1.61	1.10	0.70	0.36	0.20	0.10	0.02	0.01	0.00	0.00		
JJA %	55.58	23.03	13.04	3.47	1.98	1.27	0.83	0.46	0.21	0.10	0.02	0.00	0.00	0.00	0.00	0.00		
SON %	46.73	16.40	19.21	5.21	3.48	2.63	2.00	1.44	1.13	0.76	0.52	0.31	0.15	0.04	0.01	0.00		

From Table 4.10, it can be deduced that:

- The occurrence of DSDFS > 700 W.m^{-2} (violations of sub-procedure 2.3) is very rare. In total, there were only 27 datapoints (0.0014% of the entire record). January 2003 is the largest single contributor with 13 datapoints.

- Adjudicating susceptibility for violations, the summer months (DJF) have the most, whilst the winter months (JJA) have 4 bins clear to the left of the sub-procedure boundary.

Although the occurrence of violations of sub-procedure 2.3 is extremely low, January 2003 as the largest single contributor towards violations, (almost half of all), is now inspected.

4.3.2.3.1 January 2003 *(Largest contributor towards violations of sub-procedure 2.3.)*

Further inspection of DSDFS reveals that the 13 violating datapoints are all concentrated in one day, viz. 28 January 2003. This is illustrated in Figure 4.7.

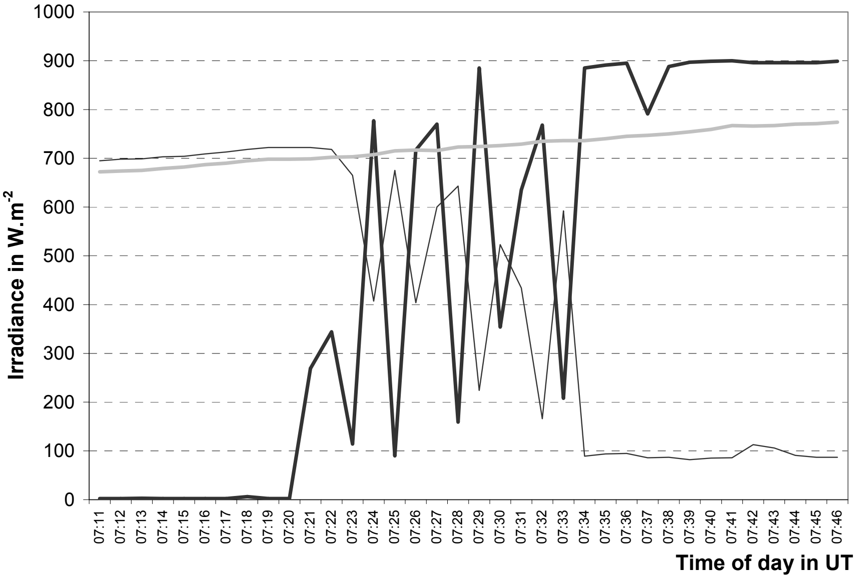

Figure 4.11 *A close-up view of 28 January 2003: (07:11UT-07:46UT): DSDIR (thick black line), DSDFS (thin black line) and DSGL2 (thick grey line)*

In Figure 4.11, the overcast morning (up to 07:20 UT) and a clear section (from 07:34 UT) sandwich an episode of intense bursts of sunshine, during which the boundary of 700 W.m^{-2} had not even been reached, alhough it can be expected due to the differential instrumental reaction times in situations such as this one, as mentioned in Section 3.1.4.2. The actual violations occur in the time period between 07:12 UT to 07:22 UT, where the diffuse (thin black line) was overtaking the global by a systematic difference, which is just enough to push DSDFS over 700 W.m^{-2}. This could be the result of a slight misalignment in the balancing of either of the two pyranometers. However, this is a sporadic occurrence of temporary over-estimation of the diffuse irradiance and hence no cause for concern. The margin of over-estimation, (i.e., just over 700 W.m^{-2}) is also negligible.

4.3.2.4 Sub-procedure 2.4 *(DSDIR < TOA * 0.9^{m})*

The definition of DSDIR in terms of the transmission coefficient τ and optical air mass m in sub-procedure 2.4, is as follows (Hegner *et al.*, 1998):

$$E_{dir} = E_{TOA}\tau^{m} \tag{4.9}$$

where

E_{dir} = Datapoint of DSDIR

E_{TOA} = Datapoint of TOA radiation

τ = transmission coefficient

m = Kasten optical air mass

Equation 4.9. may be re-arranged to:

$$\tau = (E_{dir} / E_{TOA})^{\frac{1}{m}} \tag{4.10}$$

Equation 4.10 means that violations of sub-procedure 2.4 equate to all datapoints for which $\tau > 0.9$. This is illustrated in Figure 4.12 by means of the last column of Table 4.11.

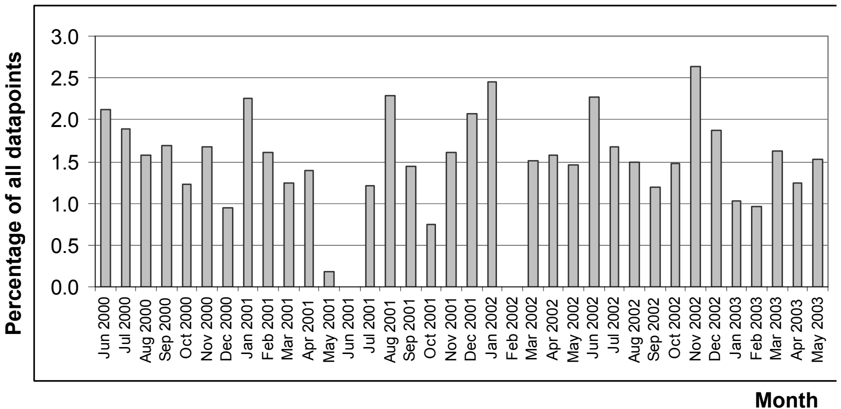

Figure 4.12 *From Table 4.11: violations of sub-procedure 2.4, expressed as a percentage of all datapoints per month.*

In Table 4.11, the frequency distribution of τ between 0.35 and 1.00, with increments of τ = 0.05, is presented for all datapoints per month of the data evaluation period. Violations of sub-procedure 2.4. are expected for $0.9 \leq \tau < 1.0$, and the two columns containing the number of datapoints for which this is true, are shaded. An extra column to allow for $\tau > 1.0$ was created, although it is physically possible that $\tau > 1.0$, due to DSDIR < TOA as a general rule (Equation 4.10). However, since DSDIR and TOA are independently acquired, allowance was made for $\tau > 1.0$ so that these cases could be contextualized and added to the true number of violations of sub-procedure 2.4. This could also serve as a test for the soundness of the complex calculations concerning TOA, DSDIR and τ. Lastly, in Table 4.11, the "missing data"-column was not included, since the table does not reflect the expected presence total of measured datapoints. It evaluates properties of datapoints already present.

Table 4.11 *Frequency distribution of τ: June 2000 to May 2003, ref.to. sub-procedure 2.4. Shaded area represents violations*

Month	Data bins in fractions of m														
	Up to 0.35	0.35-0.40	0.40-0.45	0.45-0.50	0.50-0.55	0.55-0.60	0.60-0.65	0.65-0.70	0.70-0.75	0.75-0.80	0.80-0.85	0.85-0.90	0.90-0.95	0.95-1.00	> 1.00	% > 0.9
Jun 2000	27006	68	70	88	101	134	179	333	1461	5929	4585	2323	915	0	0	2.12
Jul 2000	27960	90	93	109	107	106	165	221	1280	5522	5603	2536	840	0	0	1.88
Aug 2000	25939	163	167	208	228	281	359	1646	3621	5264	3975	2078	676	27	0	1.57
Sep 2000	27400	140	165	191	217	289	553	1054	1953	5049	3671	1776	732	0	0	1.69
Oct 2000	26667	320	321	368	442	593	1366	1405	2879	5367	2777	1583	546	0	0	1.22
Nov 2000	25333	467	376	377	451	393	425	475	813	5692	5264	2405	721	0	0	1.67
Dec 2000	22908	1824	1455	2392	1567	1148	1052	1002	1099	3462	4529	1776	419	0	0	0.94
Jan 2001	23177	214	237	208	209	222	285	305	941	4169	9673	3984	1008	0	0	2.26
Feb 2001	22666	163	162	161	188	179	259	447	1104	5135	6524	2679	645	0	0	1.60
Mar 2001	28286	262	235	257	298	263	324	467	2184	5418	4298	1784	556	0	0	1.25
Apr 2001	30834	269	272	276	234	309	344	418	1123	4654	2540	1315	602	0	0	1.39
May 2001	28706	230	209	223	238	281	332	742	4459	6673	2306	154	79	0	0	0.18
Jun 2001	28322	117	143	170	274	458	664	2873	5831	3894	418	28	0	0	0	0.00
Jul 2001	29878	183	233	215	274	370	540	688	1726	4281	4021	1682	542	0	0	1.21
Aug 2001	26128	188	263	257	279	312	327	416	1319	7180	4603	2340	1005	15	0	2.28
Sep 2001	27706	177	205	234	322	353	760	1364	1961	5122	2925	1441	592	30	0	1.44
Oct 2001	25526	210	242	248	313	454	892	2166	4149	5553	3174	1371	333	1	0	0.75
Nov 2001	27120	249	289	326	339	345	361	414	989	4101	5703	2258	696	0	0	1.61
Dec 2001	22573	193	204	221	177	219	263	334	869	5226	9796	3634	912	12	0	2.07
Jan 2002	23323	200	184	220	234	253	289	356	606	4010	10084	3778	1094	0	0	2.45
Feb 2002	33924	1394	1253	804	554	472	446	395	357	328	319	64	0	0	0	0.00
Mar 2002	26887	211	194	233	273	315	399	1016	5080	4737	2949	1667	670	0	0	1.50
Apr 2002	28544	293	349	316	301	381	361	517	1436	4419	3869	1722	682	0	0	1.58
May 2002	29598	196	208	240	274	352	462	961	2537	5227	2545	1382	649	0	0	1.45
Jun 2002	27951	126	138	165	155	172	227	584	1479	6295	2996	1921	981	0	0	2.27
Jul 2002	27531	154	157	202	217	274	372	1179	4429	4981	2698	1691	745	0	0	1.67
Aug 2002	28739	159	188	197	239	235	463	610	2806	5052	3540	1738	638	26	0	1.49
Sep 2002	25928	294	278	296	332	381	563	884	3268	5898	3053	1504	513	0	0	1.19
Oct 2002	25581	341	327	349	400	492	691	1143	2869	4965	4859	1956	645	12	0	1.47
Nov 2002	22249	183	186	151	200	236	281	391	757	3945	9886	3586	1139	0	0	2.64
Dec 2002	24178	234	207	266	277	304	352	437	863	5058	8551	3069	836	0	0	1.87
Jan 2003	25281	243	238	274	255	314	390	499	2320	7026	5557	1773	460	0	0	1.03
Feb 2003	25881	284	271	271	324	295	354	443	1400	6291	2895	1215	386	0	0	0.96
Mar 2003	28269	170	165	124	169	158	269	506	1273	7011	3921	1867	728	0	0	1.63
Apr 2003	27091	291	303	339	300	506	824	1732	2621	4154	3047	1445	539	0	0	1.25
May 2003	28051	149	168	213	210	288	349	531	2762	5885	3434	1910	680	0	0	1.52
Total %	61.08	0.660	0.640	0.710	0.700	0.771	1.050	1.836	4.859	11.603	10.180	4.404	1.470	0.0078	0	1.4778
DJF %	57.59	1.22	1.08	1.24	0.97	0.87	0.95	1.08	2.46	10.47	14.90	5.65	1.48	0.00	0	1.46
MAM %	64.48	0.52	0.53	0.56	0.58	0.71	0.92	1.734	5.91	12.12	7.27	3.33	1.30	0.00	0	1.31
JJA %	59.40	0.61	0.61	0.65	0.77	0.89	1.50	2.36	4.99	11.62	10.51	4.55	1.51	0.01	0	1.52
SON %	66.79	0.37	0.41	0.43	0.48	0.58	0.80	1.79	5.17	10.91	7.21	3.60	1.43	0.01	0	1.44

The following conclusions can be drawn from Table 4.11 and Figure 4.12:

- Two months have no sub-procedure violations, i.e. June 2001 and February 2002.
- The month with the highest number of sub-procedure violations, is November 2002 (2.64%), and the lowest non-zero number occurs in May 2001 (0.18%).
- All months, except the four months mentioned, have an incidence of violations of between 1.0% and 2.5%.
- The total number of violations is 1.48% of all datapoints. If one puts this into context with the number of *missing* datapoints (0.58% from table 4.5.) for DSDIR from which τ is derived, then this is more than two and a half times the number of missing DSDIR datapoints, and therefore significant in its own right.

Inspection of τ within a number of randomly selected diurnal cycles, revealed, that highest τ-values consistently occurs when the sun is near the horizon, i.e., when solar zenith angle (Z) is about 90°. This suggests a possible correlation between τ and Z. April 2001 (1.39%) was chosen (1.39%) for a first investigation as a representative month with sub-procedure violations close to the 36-month average (1.48%), and a scatter diagram of τ versus Z was drawn (Figure 4.13)

4.3.2.4.1 April 2001 *(Sub-procedure violations representative of the 36-month average)*

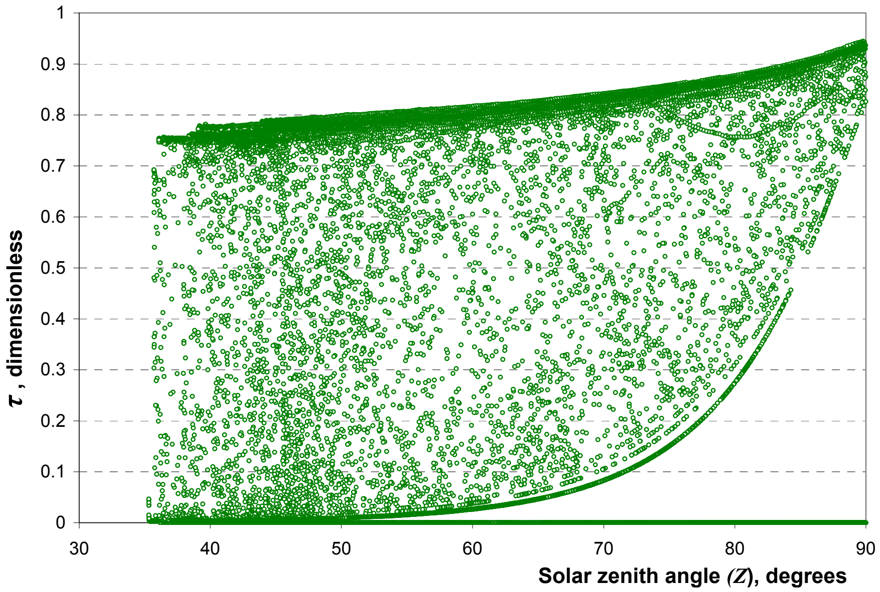

Figure 4.13 *Scatter diagram of τ versus Z: April 2001*

In Figure 4.13, all datapoints are encapsulated in an area defined by three distinct curves. The bottom curve represents the theoretical minimum TOA radiation value for any given Z. The quasi-vertical left boundary represents the highest solar elevation angle (lowest Z) possible for that month, and the curve on the top are datapoints representing cloudless sky. The domination of clear sky over the De Aar BSRN site in winter and even in April, as in this case, gravitates the majority of datapoints towards the top boundary. This is also the curve where the violating datapoints cluster. ($\tau > 0.9$ for Z ca. 90°). In this specific case, the lowest Z for which $\tau > 0.9$, is $Z = 83.71°$. This is still more than (a disturbing!) 7° separation from Z at sunset /sunrise (90.83°).

Looking at Equation 4.7 and realizing, that m, τ and TOA are all theoretically established values, the only explanation for an unusually high τ in this sub-procedure is that the actual value of DSDIR is too high. This leads to the conclusion that $\tau = 0.9$ for the violation boundary possibly needs some empirical justification to allow more flexibility, taking take local conditions into account. At the moment, $\tau = 0.9$ is applied by the WRMC as a universal measure for all BSRN stations, which clearly does not work for De Aar.

4.3.2.4.2 June 2001 *(Representative month with no sub-procedure violations)*

A month (June 2001) where no violations occurred, is represented in Figure 4.14.

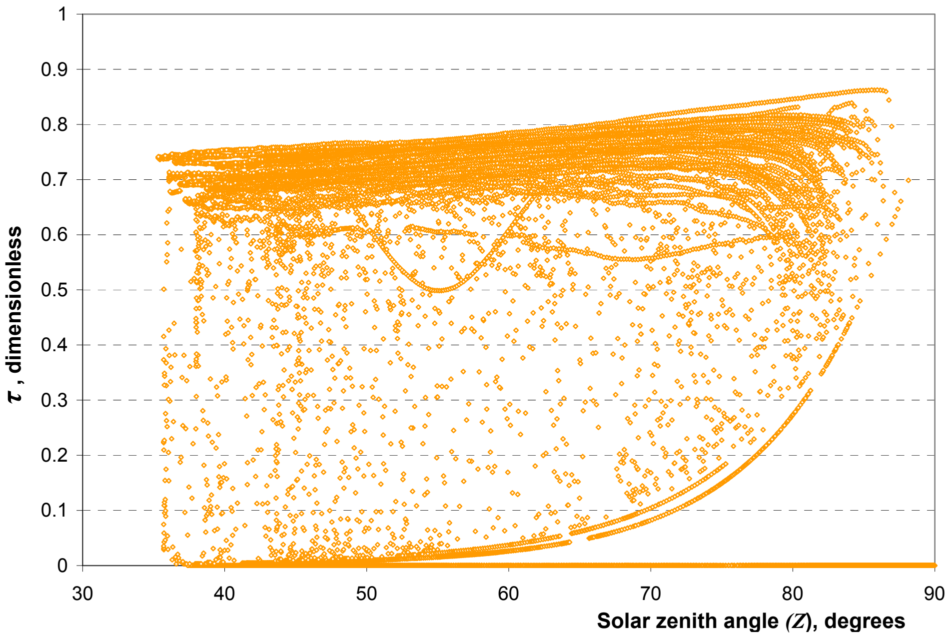

Figure 4.14 *Scatter diagram* τ *versus* Z*: June 2001*

The description of Figure 4.13 is also applicable to Figure 4.14. One exception is the area for 85° < Z < 90°, where significantly fewer points are found in Figure 4.13, possibly due to a larger number of datapoints having gravitated towards the clear-sky (top) boundary in Figure 4.14. This confirms, that in winter, the majority of datapoints over the De Aar BSRN station represents clear (cloudless) sky.

To summarize, violations of sub-procedure 2.4 occurs when the sun is near the horizon. Deeper investigation into the applicability of station - specific boundaries for τ instead of the globally applied number of 0.9, is a suggested way of addressing this problem.

4.3.2.5 Sub-procedure 2.5 ($LWD < UL + 30\ W\ m^{-2}$) and

4.3.2.6 Sub-procedure 2.6 ($UL > LWD - 30\ W\ m^{-2}$)

These two sub-procedures involve UL, which is not a basic quantity measured at De Aar during this period. In fact, the defining inequalities are describing exactly the same conditions and it is therefore the same sub-procedure, in the same fashion that the two sub-procedures in Section 4.2.3.2 are the same.

4.3.3 Procedure 3 – "Across quantities"

This procedure, the third group of sub-procedures, now implements a finer sifting process than Procedure 2, therefore values that have escaped the previous procedures, are expected to be identified by this procedure to be suspect or erroneous.

Bear in mind that up to now, the radiation data were only compared against themselves or similar parameters, such as TOA radiation. No completely independent outside source was involved. A real test of data quality happens, when a comparison with independently measured quantities is done. In procedure 3, exactly this is implemented, but also involving more than just one measurement quantity. This in itself weakens the comparison accuracy, by introducing the inherent (and by the nature of the procedure) independent uncertainties in both quantities, creating a combined relatively larger uncertainty.

4.3.3.1 Sub–procedure 3.1 *(The test for DL between $0.7{*}\sigma T^4$ and σT^4)*

The first of the finer sub-procedures involves the comparison of surface (Stevenson screen) air temperature T from the 5-minute meteorological datasets, to LWD radiation. In Section 4.2.3.1, Equation 4.4 reflects the conditions describing this sub-procedure, i.e., $0.7 < \varepsilon < 1.0$. Violations are therefore those datapoints for which $\varepsilon < 0.7$ or $\varepsilon > 1.0$. The distribution of ε in bins versus months for the 36-month period, shown in Table 4.12, hence directly indicates LWD with respect to T. The chosen bin ranges for ε are 0.60 to 1.0, with increments of 0.05. Note that the number of possible datapoints for this sub-procedure are one fifth of the corresponding number for BSRN data since five-minute meteorological data was reconciled with the one-minute BSRN data. Shaded columns again represent sub-procedure violations: Lighter shading represents the number of violating one-minute values per specific month, darker shading the corresponding percentages.

Table 4.12 *Frequency distribution of ε: June 2000 to May 2003, ref.to sub-procedure 3.1. Shaded areas represent violations*

Month	% less than 0.70	Less than 0.60	0.60 to 0.65	0.65 to 0.70	0.70 to 0.75	0.75 to 0.80	0.80 to 0.85	0.85 to 0.90	0.90 to 0.95	0.95 to 1.00	More than 1.00	% more than 1.00	Possible data points
Jun 2000	3.88	0	0	335	3709	2900	563	356	267	162	58	0.67	8640
Jul 2000	13.43	144	49	1006	4260	2095	336	376	330	105	24	0.27	8928
Aug 2000	18.44	0	319	1327	3387	2630	572	230	152	24	1	0.01	8928
Sep 2000	15.60	0	151	1194	2812	1949	576	423	504	742	134	1.58	8640
Oct 2000	14.35	0	2	1238	3438	2329	990	374	168	32	157	1.80	8928
Nov 2000	3.95	0	4	339	2765	2965	780	504	540	392	60	0.72	8640
Dec 2000	6.78	0	9	596	3389	2870	936	396	253	135	23	0.27	8928
Jan 2001	8.08	0	9	712	3245	2215	1374	664	275	81	13	0.15	8928
Feb 2001	1.48	0	0	119	2013	2922	1491	693	354	112	3	0.04	8064
Mar 2001	0.37	0	0	33	1656	3229	1651	982	828	175	18	0.21	8928
Apr 2001	0.00	0	0	0	993	2895	1772	938	850	823	116	1.38	8640
May 2001	3.52	0	18	296	3333	3154	741	354	399	279	30	0.35	8928
Jun 2001	7.12	0	0	615	3530	2895	543	274	212	228	79	0.94	8640
Jul 2001	16.90	0	215	1294	3947	1534	525	368	363	332	201	2.29	8928
Aug 2001	22.47	0	25	1981	4085	1731	433	167	124	40	0	0.00	8928
Sep 2001	16.48	2	0	1422	3498	1263	616	259	460	772	81	0.97	8640
Oct 2001	6.81	0	50	558	3200	2396	930	754	619	69	7	0.08	8928
Nov 2001	3.04	0	0	263	2377	1964	1339	808	989	554	102	1.22	8640
Dec 2001	2.93	0	0	262	3221	3142	1038	393	282	241	25	0.29	8928
Jan 2002	11.55	0	32	999	2969	2170	1226	579	445	168	15	0.17	8928
Feb 2002	4.99	0	0	402	2994	2792	696	429	219	44	0	0.00	8064
Mar 2002	7.28	0	0	650	1818	2601	2102	807	536	72	11	0.13	8928
Apr 2002	8.95	0	39	734	2300	2404	1321	937	428	135	3	0.04	8640
May 2002	24.25	173	394	1598	2906	1735	451	419	509	558	101	1.14	8928
Jun 2002	15.49	0	2	1336	3925	1819	289	253	350	316	41	0.49	8640
Jul 2002	8.60	0	56	712	4517	1959	570	395	179	194	24	0.28	8928
Aug 2002	19.21	0	128	1587	3006	1782	630	390	509	552	80	0.92	8928
Sep 2002	10.66	0	0	921	3769	1962	632	553	254	206	25	0.30	8640
Oct 2002	26.61	0	138	2238	2959	1764	837	372	159	119	10	0.12	8928
Nov 2002	23.58	0	133	1904	3748	1557	522	274	214	60	16	0.19	8640
Dec 2002	5.49	0	0	490	2711	2870	1278	582	322	214	29	0.34	8928
Jan 2003	6.21	0	0	554	2386	3471	1644	551	285	31	6	0.07	8928
Feb 2003	0.35	0	0	28	781	2242	2318	1197	973	474	51	0.63	8064
Mar 2003	1.28	0	0	114	2006	3653	1045	679	656	567	208	2.33	8928
Apr 2003	2.60	0	0	225	1645	3629	1411	713	629	369	19	0.22	8640
May 2003	2.92	0	0	261	2444	3994	1251	468	263	206	41	0.46	8928
PERCENT	9.08	0.1	0.58	4.57	17.9	18.4	8.95	5.24	4.36	2.93	0.59	0.59	
DJF %	5.41	0.00	0.06	5.35	30.49	31.76	15.43	7.05	4.38	1.93	0.22	0.22	
MAM %	5.71	0.22	0.57	4.92	24.03	34.34	14.78	7.92	6.41	4.01	0.70	0.70	
JJA %	14.00	0.18	1.00	12.82	43.24	24.34	5.61	3.53	3.13	2.46	0.66	0.66	
SON %	13.43	0.00	0.61	12.82	36.33	23.08	9.19	5.49	4.97	3.75	0.77	0.77	

From Table 4.12, the following conclusions are drawn:

- Lower boundary violations (datapoints for which $\varepsilon < 0.7$) are on average substantially more (9.08%) than upper boundary ($\varepsilon > 1.0$) violations (0.59%). Bearing in mind that the lower boundary is tested under clear-sky conditions and the upper boundary under overcast conditions, the abundance of clear sky at BSRN De Aar for most of the year, explains the reasons for this phenomenon.

- Lower boundary violations are biased towards a maximum in the latter half of the year (JJA and SON), while upper boundary violations shows a minimum in summer (DJF).

- Only one month (April 2001) has no lower boundary violations, while two months (August 2001 and February 2002) have no upper boundary violations.

- Lower boundary violations occur at more than 20% of the time in the months of October 2002 (26.61%), May 2002 (24.25%), November 2002 (23.58%) and August 2001 (22.47%), while upper boundary violations exceed 2% of the time in the months of March 2003 (2.33%) and July 2001 (2.29%).

To illustrate the boundary violations towards both sides, Figure 4.15 represents a typical month, July 2001, having a 16.9% lower-boundary violation and a 2.29% upper-boundary violation rate. For clarity, individual time-series of LWD, σT^4 (T = surface temperature) as well as ε (σT^4/LWD) are presented.

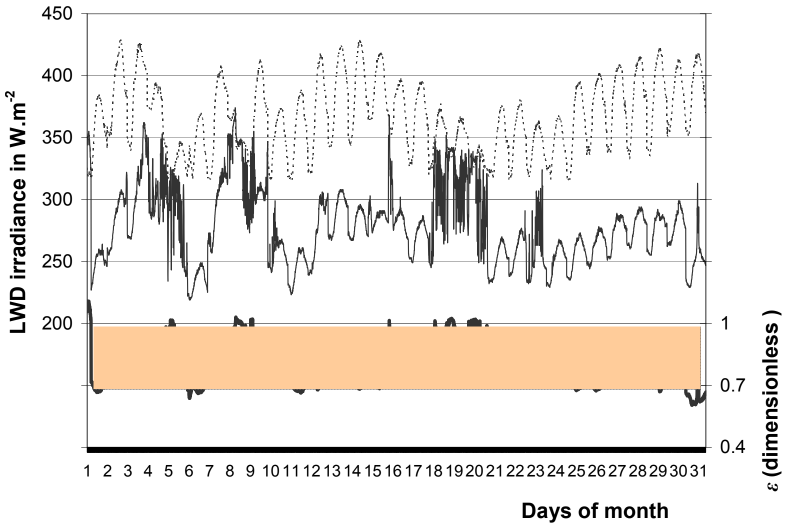

Figure 4.15 *Time-series for July 2001: Measured LWD (thin black line), calculated σT^4 (dashed line) and ε (thick black line). Area between boundaries is shaded.*

In general, LWD and σT^4 show opposing, as well as concurrent trends to a large degree. Boundary violations (ε outside the shaded area in Figure 4.15) occur randomly and scattered throughout the month – there is no preferred time, although the lower boundary in general, is more frequently and severely tested.

For July 2001, the highest ε occurs on the 1st, 4th, 5th, 8th, 9th, 16th and 18th to 20th, representing relatively frequent but only slight passing through the 1.0 boundary. The violations of the lower boundary (cf Figure 4.15) only appear significant on the 31st. Lowest values of ε appear on the 10th to 17th and 23rd to 31st, suggesting that the sky was indeed clear during those times. If the threshold value for ε was only slightly lower than 0.7, a lot less violations would have occurred. This is again an example of how an indiscriminate application of one sub-procedure parameter to all stations in the world leads to a large number of violations. More research into acquiring site-specific parameters, would be a step towards addressing this problem.

One last comparison to elucidate the behaviour of ε, is an overall look at the relationship between ε and U. (U = relative humidity, expressed as a percentage). Figure 4.16 is a scatter diagram of ε versus U expressed in 18915 hourly values for June 2000 to May 2003.

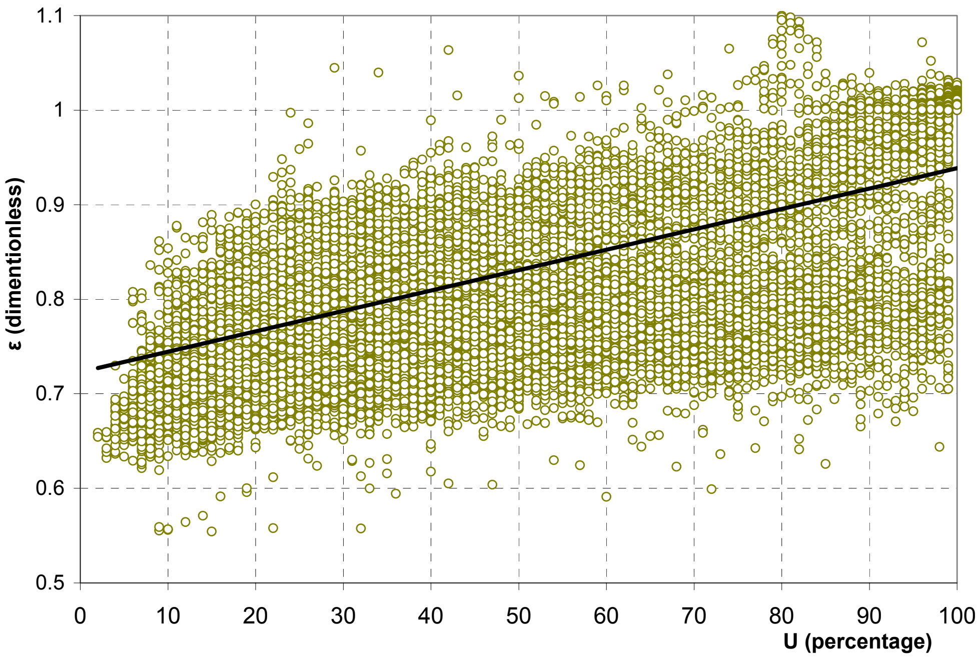

Figure 4.16 *Scatter diagram U versus ε: June 2000 to May 2003 (hourly values) with straight regression line*

From Figure 4.16, the impression (also indicated by the regression line) is that there is a trend for lower ε to be gravitated towards lower U and vice versa, although there is a large tolerance both sides of the regression line, which is calculated as **$\varepsilon = 0.725 + 0.0013U$.** Correlance coefficient of R^2 is calculated as 0.2704, which is relatively low, but still has significance given the long record. Generally, upper boundary violations ($\varepsilon > 1.0$) occur towards higher U and lower boundary violations ($\varepsilon < 0.7$) towards lower U, although there are a number of outliers in both cases.

4.3.3.2 Sub–procedure 3.2 *(The test for UL between $\sigma(T-10)^4$ and $\sigma(T+10)^4$)*

Since UL is not a basic parameter measured at De Aar, there is no data and no analysis.

4.3.3.3 Sub–procedure 3.3 *(The test for DSDIR*cosZ between DSGL2 – DSDFS – 50 W.m^{-2} and DSGL2 – DSDFS + 50 W.m^{-2})*

and

4.3.3.4 Sub–procedure 3.4 *(The test for DSGL2 – DSDFS between DSDIR*cosZ - 50 W.m^{-2} and DSDIR*cosZ + 50 W.m^{-2})*

In Section 4.2.3.2 it was shown that procedures 3.3 and 3.4 refer to the same conditions, therefore these two procedures are treated as one in this Section. Equation 4.8 implies that evaluation involves calculating MM = DSGL2–DSDFS–DSDIR*cos$Z$ for all datapoints. Violations of sub-procedure therefore are cases where MM < -50 W.m^{-2} or MM > 50 W.m^{-2}.
It follows from Table 4.13, that:

- In general, datapoints are gravitated towards the upper boundary, i.e., DSGL2 > DSDFS + DSDIR * cos$Z$ to a larger extent than DSGL2 < DSDFS + DSDIR * cos$Z$.
- For every month, the majority of all datapoints lies between the borders, but there are certain months where substantial, and possibly serious, boundary violations occur.
- Most serious lower boundary violations (DSGL2–DSDFS–DSDIR*cos$Z$ < -50 W.m^{-2}) occurred in December 2000 (19.40%) and October 2001 (13.09%). The worst upper boundary violations (DSGL2–DSDFS–DSDIR*cos$Z$ > 50 W.m^{-2}) happened in December 2000 (17.86%), May 2001 (15.86%) and March 2002 (14.18%).

Although violations are numerous in specific months during which they occur, overall they only amount to 0.0038% for the lower boundary and 0.0044% for the upper boundary. The higher number for the latter confirms the datapoint gravitation towards the upper boundary.

Table 4.13 *Frequency distribution of DSGL2-DSDFS-DSDIR*cosZ: June 2000 to May 2003, ref.to sub-procedures 3.3 and 3.4. Shaded areas represent violations*

Month	% <-50	Data bins in W.m^{-2}															% >+50
		< -70	-70 to -60	-60 to -50	-50 to -40	-40 to -30	-30 to -20	-20 to -10	-10 to 0	0 to 10	10 to 20	20 to 30	30 to 40	40 to 50	50 to 60	60 to 70	> 70	
Jun 2000	0.00	0	0	0	0	0	0	3	26479	16106	556	24	22	2	0	0	0	0.00
Jul 2000	2.93	1072	117	114	111	106	94	110	26987	15125	647	8	16	4	5	1	116	0.27
Aug 2000	0.00	0	0	0	0	1	3	19	37574	7018	15	0	0	2	0	0	0	0.00
Sep 2000	0.00	0	0	0	0	0	0	1	20586	8865	4105	5874	3290	456	11	0	0	0.03
Oct 2000	0.00	0	0	0	0	2	2	2	24364	10520	8384	1353	5	1	1	0	0	0.00
Nov 2000	0.02	9	0	0	0	0	1	1	20990	7230	3956	4043	3283	1468	290	164	1757	5.34
Dec 2000	19.40	7171	231	234	226	255	293	361	17561	7612	891	904	876	990	880	869	5279	17.86
Jan 2001	0.29	0	47	82	48	49	38	5	15236	11639	4332	6058	4058	1869	495	365	311	2.64
Feb 2001	0.00	1	0	0	0	0	0	1	12176	14206	4518	7211	1897	273	26	0	3	0.07
Mar 2001	0.00	0	0	0	0	0	1	0	16774	15223	5023	6058	740	371	155	155	132	0.99
Apr 2001	0.03	0	9	4	6	6	3	7	21110	14791	5342	1203	581	42	20	24	42	0.20
May 2001	0.00	0	0	1	221	519	940	1821	11661	2640	1724	1892	2669	1682	761	646	2902	15.86
Jun 2001	9.72	1514	716	1899	2232	2325	1644	1113	22695	4823	1077	779	596	440	335	279	733	3.17
Jul 2001	0.58	0	0	261	737	709	830	535	24001	14872	2534	109	20	21	4	0	0	0.01
Aug 2001	0.00	0	0	0	0	0	0	0	26434	17873	120	7	6	6	11	10	165	0.42
Sep 2001	7.06	2468	253	328	311	476	362	338	15332	16219	5347	1429	107	102	44	76	0	0.28
Oct 2001	13.09	5282	275	278	251	246	218	192	9493	19943	5790	2043	258	85	75	149	54	0.62
Nov 2001	0.00	0	0	0	0	1	0	2	14870	15666	5754	5749	961	152	30	2	3	0.08
Dec 2001	0.04	0	7	9	11	14	0	71	16465	12028	4487	5869	3938	1162	221	249	102	1.28
Jan 2002	0.34	0	94	59	46	49	38	48	9290	20056	7672	5713	1427	133	5	0	1	0.01
Feb 2002	0.01	3	0	0	1	0	0	1	9520	25236	4648	720	147	33	1	0	0	0.00
Mar 2002	1.38	314	129	142	80	27	227	90	13269	15621	3232	2567	1520	1418	1568	2069	2358	14.18
Apr 2002	0.00	0	0	0	0	0	0	0	10511	27200	4467	896	80	25	11	0	0	0.03
May 2002	0.00	0	0	0	134	1109	783	295	11107	20893	2089	1991	2447	3397	382	4	0	0.86
Jun 2002	0.00	0	0	0	0	0	0	5	10590	20440	2549	2655	3982	2911	46	1	11	0.13
Jul 2002	0.00	0	0	0	0	0	1	8	11518	19398	2852	2954	4242	3223	85	61	288	0.98
Aug 2002	0.00	0	0	0	0	0	0	4	11290	30973	1479	132	44	38	48	42	580	1.52
Sep 2002	0.00	0	0	0	1	1	1	1	7694	29694	5543	250	7	0	0	0	0	0.00
Oct 2002	0.00	0	0	0	0	0	0	0	9829	23256	9623	1407	149	52	78	44	192	0.71
Nov 2002	0.00	0	0	0	0	0	1	0	9977	20708	10016	2307	149	11	4	4	13	0.05
Dec 2002	0.00	1	0	0	1	2	15	16	9354	22700	8557	3462	428	76	13	0	7	0.04
Jan 2003	3.86	60	467	1192	849	903	1730	1060	17312	17700	3039	255	18	1	1	1	42	0.10
Feb 2003	0.40	0	0	159	27	18	18	10	18943	11802	8296	641	36	12	8	5	335	0.87
Mar 2003	0.02	0	5	5	7	11	17	17	25940	18143	445	25	5	1	3	1	5	0.02
Apr 2003	0.17	31	23	20	13	9	10	7	23906	18588	555	23	7	0	0	0	0	0.00
May 2003	0.75	167	73	93	37	24	37	50	25919	17547	623	27	18	15	0	0	0	0.00
%	0.004	1.147	0.155	0.309	0.339	0.435	0.463	0.393	39.11	37.57	8.897	4.860	2.412	1.298	0.356	0.331	0.979	0.004
DJF %	0.01	1.86	0.22	0.45	0.31	0.33	0.55	0.40	32.37	36.77	11.94	7.93	3.30	1.17	0.42	0.38	1.56	0.01
MAM %	0.00	0.13	0.06	0.07	0.13	0.43	0.51	0.58	40.31	37.90	5.91	3.69	2.03	1.75	0.73	0.73	1.37	0.01
JJA %	0.01	1.97	0.13	0.15	0.14	0.18	0.15	0.14	33.87	38.69	14.89	6.22	2.09	0.59	0.14	0.11	0.51	0.00
SON %	0.00	0.65	0.21	0.57	0.77	0.79	0.65	0.45	49.71	36.89	2.98	1.68	2.25	1.67	0.13	0.10	0.48	0.00

- Since three radiometers (global and diffuse pyranometers) and a pyrheliometer are involved in the evaluation of sub-procedues 3.3 and 3.4, using Equation 4.8, the margin of error is set to the relatively large value of 50 W.m^{-2}, as mentioned in Section 4.2.3.2. The distribution of MM = DSGL2–DSDFS–DSDIR*cos*Z* have been tabulated using bin sizes of 10 W.m^{-2}, distributed between -70 W.m^{-2} and 70 W.m^{-2} (Table 4.13). Columns either side of the boundaries are shaded as in Table 4.12: datapoint numbers in a light shade, and percentages in a darker shade.

4.3.3.4.1 May 2001 *(Typical boundary violations of sub-procedure 3.3 and 3.4)*

A typical month exhibiting boundary violations (15.86% of the upper boundary), is now investigated. Figure 4.17 depicts a time-series of (DSGL2–DSDFS–DSDIR*cos$Z$) for May 2001.

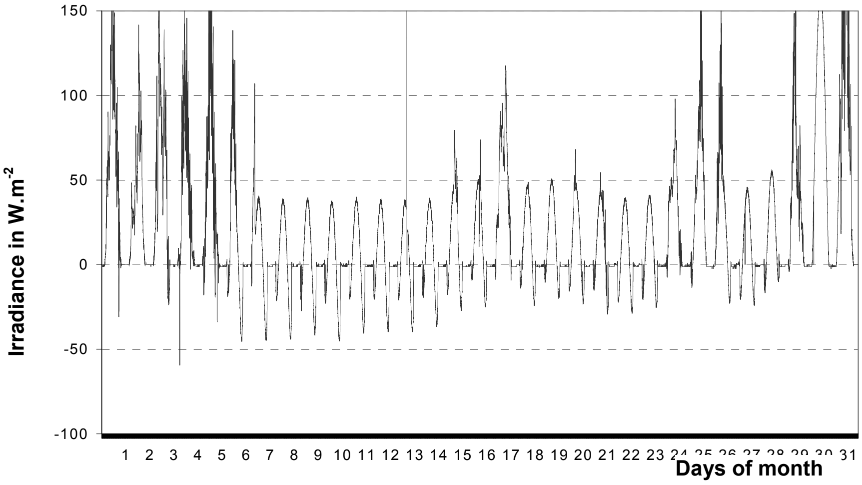

Figure 4.17 *Time-series of (DSGL2–DSDFS–DSDIR*cos$Z$): May 2001*

Violations of the 50 W.m^{-2} upper boundary are apparently distributed at random days throughout the month (Figure 4.17). A single spike violation of the upper boundary also exists.

Note that the first 6 days of the month all have large occasional overshootings of the upper boundary, as well as the 17th, 25th, 26th, and from the 19th to the 31st. Two specific days, typical of this behaviour, are now investigated. They are the 17th and 31st.

Firstly, consider 17 May 2001. It was a sunny day with frequent broken cloud passing the sun. Time-series of datapoints for the respective radiometers are depicted in Figure 4.18.

It follows from Figure 4.18 that the boundary violations (open circles above the 50 W.m^{-2} line against the right-hand axis) are restricted to the middle of the day, whilst DSGL2 does not overshoot TOA radiation. In fact, it maintains a healthy cushion throughout the day. Note that the time-series of DSDFS (dashed line) followed a similar trend as that of DSGL2-DSDFS-DSDIR*cos$Z$ (open circles) and DSDFS (dashed line) (Figure 4.18). This indicates that the features of DSDFS are echoed during violations. Thus, the violations are most likely the cause of DSDFS (dashed line) being too high.

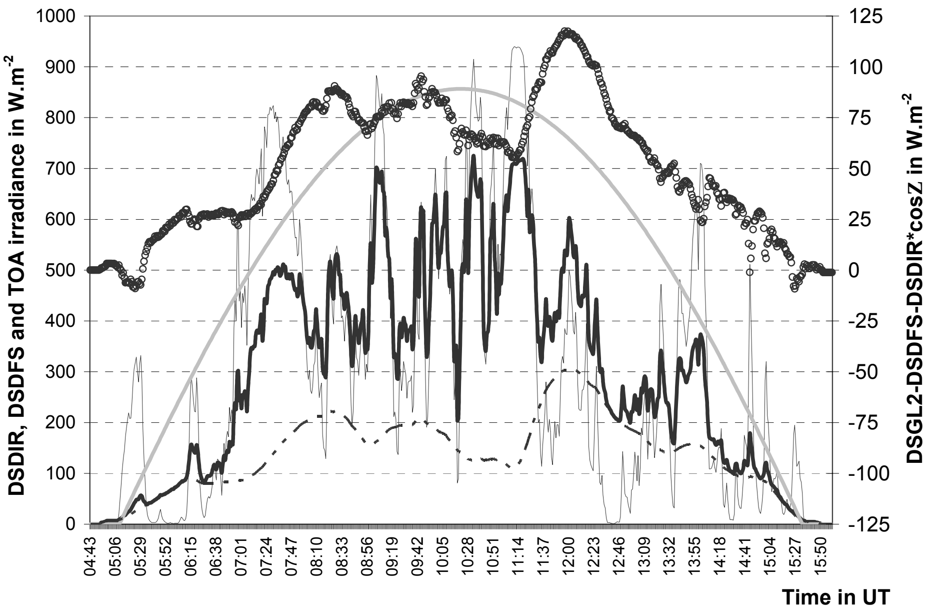

Figure 4.18 *Time-series on 17 May 2001 for DSDIR (thin black line), DSDFS (dashed line),DSGL2 (thick black line), TOA radiation (thick smooth grey line) and DSGL2-DSDFS-DSDIR*cos$Z$ (open circles).*

An anomalously high DSDFS can be attributed to pyranometer skewness, or more likely, since the tracking is not active, a case of the shading device drifting off track and not fully shading the diffuse pyranometer, resulting in anomalously high values.

One of the most important recommendations is therefore, that the shading must be controlled regularly on site at different times of the day, and as soon as possible, the tracker be upgraded to an active tracker, which will be a major step towards eliminating this error.

Secondly, consider 31 May 2001, which was also a sunny day with the occasional passage of broken clouds. However, the clouds were thinner than those observed on the 17th. In Figure 4.19, a graph similar to Figure 4.18 is depicted.

From Figure 4.19, behaviour of DSGL2-DSDFS-DSDIR*cos$Z$ similar to Figure 4.18 is observed: a large number of upper-boundary violations (open circles above 50 W.m^{-2}). One exception is, that the relationship between DSGL2-DSDFS-DSDIR*cos$Z$ (open circles) and DSDFS (dashed line) is not as clearly exhibited as in Figure 4.18. The violations are mainly the result of DSGL2 (thick black line) overshooting TOA radiation (thick smooth grey line), leading to anomalously high values for DSGL2-DSDFS-DSDIR*cos$Z$.

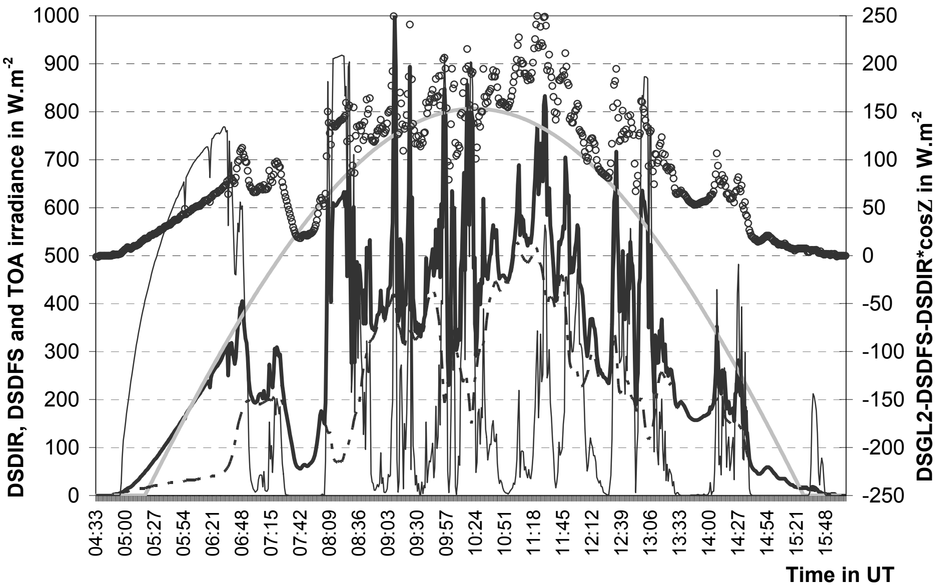

Figure 4.19 *Time-series on 31 May 2001 for DSDIR (thin black line), DSDFS (dashed line), DSGL2 (thick black line), TOA radiation (thick smooth grey line) and DSGL2-DSDFS-DSDIR*cosZ (open circles).*

Abnormally high DSGL2 can also be the result of temporary pyranometer skewness, since the other possible causes (misapplication of calibration or a mistake in the logger program) are highly unlikely.

This sub-procedure, the last of the finer sub-procedures, also exhibits by far the largest number of boundary violations in comparison to all the preceding sub-procedures.

4.3.4 Procedures 4 and 5

The originally proposed procedures 4 and 5 were not yet implemented at WRMC at the time of writing this document, therefore no further discussion is included here.

4.4 CONCLUSIONS

Quality assurance, as defined at the beginning of this Chapter, is in essence a dynamic process. In the BSRN context, it takes the form of continuous evaluation and effective feedback, in order to sustain the best possible measurements, as well as a retrospective character to apply corrections truly believed to enhance the quality of collected data. The WRMC employs a rigorous set of quality control procedures, which had all been investigated in detail throughout this Chapter. They identify a wide variety of errors in the standard BSRN one-minute data for a wide variety of diverse reasons.

When the WRMC procedures are applied to the experimental three-year De Aar dataset between June 2000 and May 2003, errors are identified. Some of these sub-procedure violations identify erroneous data which make up a small percentage of the overall dataset, and can thus be safely deleted. There are, however, single cases of substantial differences which would require deeper investigation, but such details are beyond the scope of this research.

Like all procedures of this kind, the WRMC quality control measures are not perfect, and the author acknowledges the fact, that WRMC is also in the process of undergoing stages of development. Two pairs of sub-procedures (2.5. and 2.6 as well as 3.3 and 3.4), listed in the database design notes as separate sub-procedures, had been demonstrated to refer to the same circumstances and are therefore identical. Another flaw in the procedures, which the author wishes to point out, is the inflexible definition of certain sub-procedure boundaries globally applied to all BSRN stations in the network.

This inflexible situation, leading to a large number of possibly falsely identified sub-procedure violations, could be improved with some refinement measures to the procedures, making allowance for local and site-specific circumstances, which can be determined by a combination of empirical and analytical methods. However, in the opinion of the author, this falls beyond the scope of this document.