INSTRUMENT TEST REPORT 674

Investigation of the Uncertainty of Sunshine Duration in the Solar and Terrestrial Radiation Network

Paul Dyson

Regional Instrument Centre
24 September 2003

Authorisation

Bruce W Forgan
Supervisor
Regional Instrument Centre, OEB

Distribution (email) All RDs, ROMs, RESMs, HYDs, CCSs
STAW, STNM, STIE, STES, SRLR, SROG, SROO, SRPP
SRSI, SRUM, BAPS, R. Hibbins,
STCC, SRDS, SRCA, R Hutchinson
STHY, SRWR, SRFW, STTR, CSR, T. Keenan, LIB

19 Pages

1. PURPOSE

The WMO requires that the uncertainty of measurements of *sunshine duration* be +/- 0.1 hours [1], and we investigate whether the Solar and Terrestrial Radiation Network of the Commonwealth Bureau of Meteorology (the Bureau) [2] fulfils this requirement. This report examines several models to estimate the sunshine duration and its uncertainty from the statistical minute data.

2. CURRENT DATA COLLECTION

The WMO states that the "*sunshine duration* during a given period is defined as the sum of that subperiod for which the direct solar irradiance exceeds 120 W m^{-2}". The WMO also states that the uncertainty of the *sunshine duration* should be +/- 0.1 hours [1]. We will refer to this condition as the *uncertainty requirement*, and to 120 W m^{-2} as the *threshold*. In two of the four recommended calibration methods the WMO indicates that the pyrheliometer should be adjusted if calibration reveals it to be more than 20% from 120 W m^{-2}. As the required *sunshine duration* uncertainty of 0.1 hours is less than 1% of daylight hours for a typical day, the allowance for the *threshold* to vary by up to 20% is surprising.

In this discussion the term "uncertainty" refers to the *expanded uncertainty* as defined by the ISO to "encompass a large fraction of the distribution of values" [3]. Here we choose a "large fraction" to be 95% and denote the uncertainty as U_{95}.

In the Solar and Terrestrial Radiation Network [2] the direct solar irradiance is measured once per second by pyrheliometers. For the sixty values measured each minute, the minimum (I_{min}), the maximum (I_{max}), the mean (μ) and the standard deviation (σ) are recorded. The individual measurements for each second are not stored.

Currently, as a simple approximation, the number of minutes with means greater than the threshold is used as a measure of sunshine duration for each day. As this method makes the unjustified assumption that the direct irradiance is constant and equal to the mean for the entire minute, this estimate deserves some evaluation. Also, the current method does not determine an uncertainty for the sunshine duration. This makes it difficult without further analysis to determine whether the WMO uncertainty requirement is met.

Since the minimum, maximum and standard deviation are recorded for each minute, a more refined calculation of the sunshine duration are possible, as are estimates of the uncertainty. This report investigates various methods of calculating of the sunshine duration and its uncertainty in order to assess the degree to which the uncertainty requirement is being attained.

3. METHODS AND RESULTS

This analysis used the data set for the year 2002 from the 16 stations listed in Table 1. The stations are those existing in 2002 in the Solar and Terrestrial Radiation Network and a National Institute for Water and Atmosphere (NIWA) station configured to almost the same specification at Lauder, South Island NZ. These stations are spread across Australasia with a wide range of latitudes, longitudes, terrain, and radiation climate.

The data set consists of direct irradiances for sun elevations of at least 1.0 degree above the horizon. The data set for a station does not include days for which no data are available.

Table 1. Solar and Terrestrial Network Sites.

Station Number	Location	Latitude	Longitude
201	Wagga Wagga	35° 09.59'	147° 27.37'
301	Melbourne	37° 40.50'	144° 50.53'
351	Mildura	34° 14.23'	142° 05.13'
401	Rockhampton	23° 22.62'	150° 28.59'
402	Cairns	16° 52.52'	145° 44.64'
501	Mt. Gambier	37° 44.92'	140° 36.44'
601	Cape Grim	40° 40.90'	144° 41.35'
701	Lauder(NZ)	45° 2.68'	169° 41.33'
801	Alice Springs	23° 47.86'	133° 52.28'
803	Darwin	12° 25.45'	130° 53.55'
851	Tennant Creek	19° 38.61'	134° 10.91'
921	Broome	17° 56.93'	122° 14.04'
922	Learmonth	22° 14.51'	114° 05.74'
923	Geraldton	28° 47.49'	114° 41.75'
924	Kalgoorlie	30° 47.45'	121° 27.88'
998	Glenlitta	37° 41.48'	144° 56.82'

3.1 INITIAL CLASSIFICATION

Using the maximum and minimum irradiances, each minute in the data set was classified as one of four types: *sunshine minute* (M_s); *no-sunshine minute* (M_n); *questionable sunshine minute* (M_q) (see figure 1); minutes for which there were no data available were classified as M_m.

Minutes with I_{min} greater than the threshold represent a full minute of sunshine duration and so were classified as M_s. Minutes with I_{max} below or equal to the threshold represent a full minute with zero sunshine duration and so were classified as M_n. All other minutes for which data was available (I_{max} above the threshold and I_{min} less than or equal to the threshold) were classified as M_q. The results for the 2002 data set are shown in table 2.

Figure 1. Classification of minutes using the threshold.

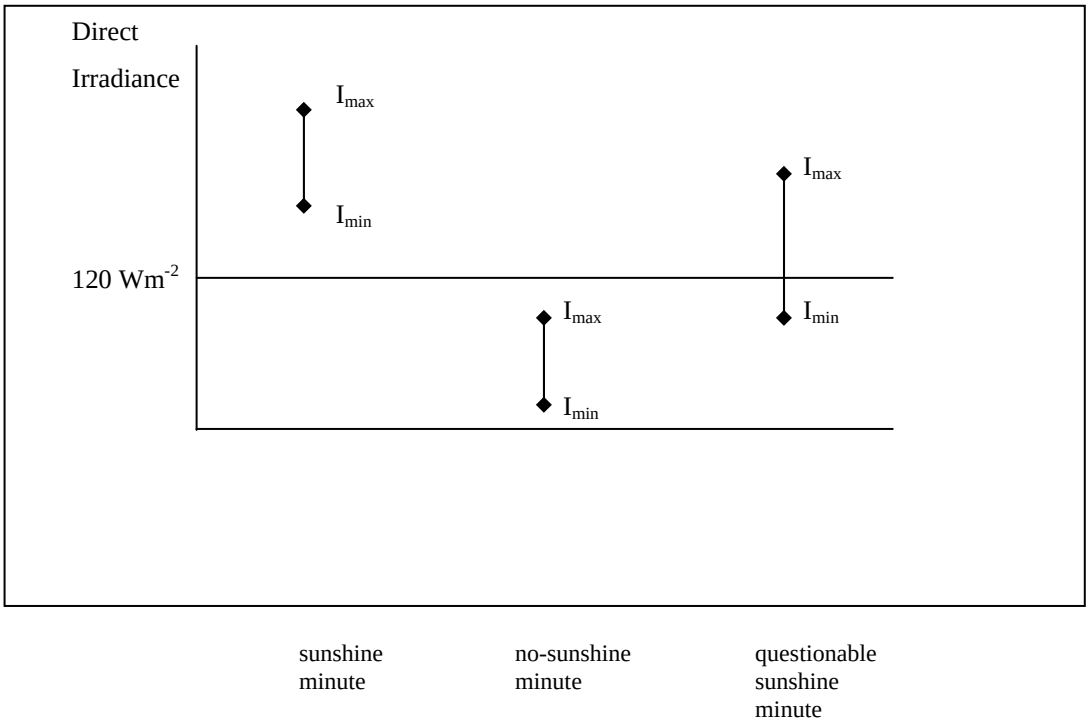

Table 2. Classification of minute data. Total minutes and missing minutes do not include days for which no direct irradiance data are available. M_m (%) indicates M_m as a percentage of the total minutes. M_s (%), M_n (%) and M_q (%) indicate M_s, M_n and M_q respectively as a percentage of the total minutes when the missing minutes are excluded.

Stn	Days of Data	Total Minutes	M_m	M_m (%)	M_s	M_n	M_q	M_s (%)	M_n (%)	M_q (%)
201	359	253833	1105	0.4	174434	63086	15208	69.0	25.0	6.0
301	365	257895	476	0.2	112951	121169	23299	43.9	47.1	9.1
351	365	258219	1788	0.7	183813	57043	15575	71.7	22.2	6.1
401	42	27304	0	0.0	20923	4805	1576	76.6	17.6	5.8
402	24	18376	43	0.2	10762	6316	1255	58.7	34.5	6.8
501	358	253608	9446	3.7	100376	117468	26318	41.1	48.1	10.8
601	318	218422	1742	0.8	86295	102911	27474	39.8	47.5	12.7
701	359	251593	2865	1.1	111673	121141	15914	44.9	48.7	6.4
801	358	254166	2768	1.1	205028	36913	9457	81.6	14.7	3.8
803	222	156582	4112	2.6	115911	26893	9666	76.0	17.6	6.3
851	348	246030	742	0.3	209761	29208	6319	85.5	11.9	2.6
921	185	138577	1219	0.9	102034	27229	8095	74.3	19.8	5.9
922	347	246641	2590	1.1	212009	25286	6756	86.9	10.4	2.8
923	363	257061	3068	1.2	175360	58799	19834	69.0	23.1	7.8
924	315	226063	6260	2.8	158272	48226	13305	72.0	21.9	6.1
998	348	244304	4268	1.7	99926	119127	20983	41.6	49.6	8.7
Total	4676	3308674	42492	1.3	2079528	965620	221034	63.7	29.6	6.8

It is clear from the results in table 2 that there were a sufficient number of *questionable sunshine minutes* to require their further analysis. Being 6.8% of the usable data (2.6% at the best station), they represent significantly more than 0.1 hours per day. (By way of comparison, 0.1 hours is 0.83% of 12 hours, the mean time the sun is above the horizon during a day.)

In order to provide a benchmark to assess these various methods, the data were examined to determine N_0, the number of days for which the number of *questionable minutes* is greater than 6. The results are in table 3.

Table 3. Sunshine days when more than six minutes could not be classified as sunshine or not sunshine.

Station Number	Number of Days with Valid Data	N_0	N_0 (%)
201	359	245	68.2
301	365	315	86.3
351	365	246	67.4
401	42	26	61.9
402	24	23	95.8
501	358	324	90.5
601	318	298	93.7
701	359	299	83.3
801	358	164	45.8
803	222	156	70.3
851	348	127	36.5
921	185	141	76.2
922	347	134	38.6
923	363	260	71.6
924	315	211	67.0
998	348	313	89.9
Total	4676	3282	70.2

The worst case is to treat all of the *questionable sunshine minutes* as 100% uncertain, resulting in N_0 days that do not meet the *uncertainty requirement*.

To improve upon this result the *questionable sunshine minutes* need to be further examined for the 70.2% of days where more than 6 valid minutes are not classified as either *sunshine* or *no-sunshine minutes*.

3.2 THRESHOLD TOLERANCE

The data were also examined to take into account the 20% tolerance allowed in the *threshold* (see figure 2). Thus we classified minutes with I_{min} greater than the *threshold* - 20% (96 Wm^{-2}) as *sunshine minutes*; the remaining minutes with I_{max} below or equal to the *threshold* + 20% (144 Wm^{-2}) as *no-sunshine minutes*; and all other valid data minutes as *questionable sunshine minutes*. As before, minutes for

which there were no data available were classified as *missing minutes*. This gives the results in table 4.

Figure 2. Classification of minutes using threshold +/- 20%.

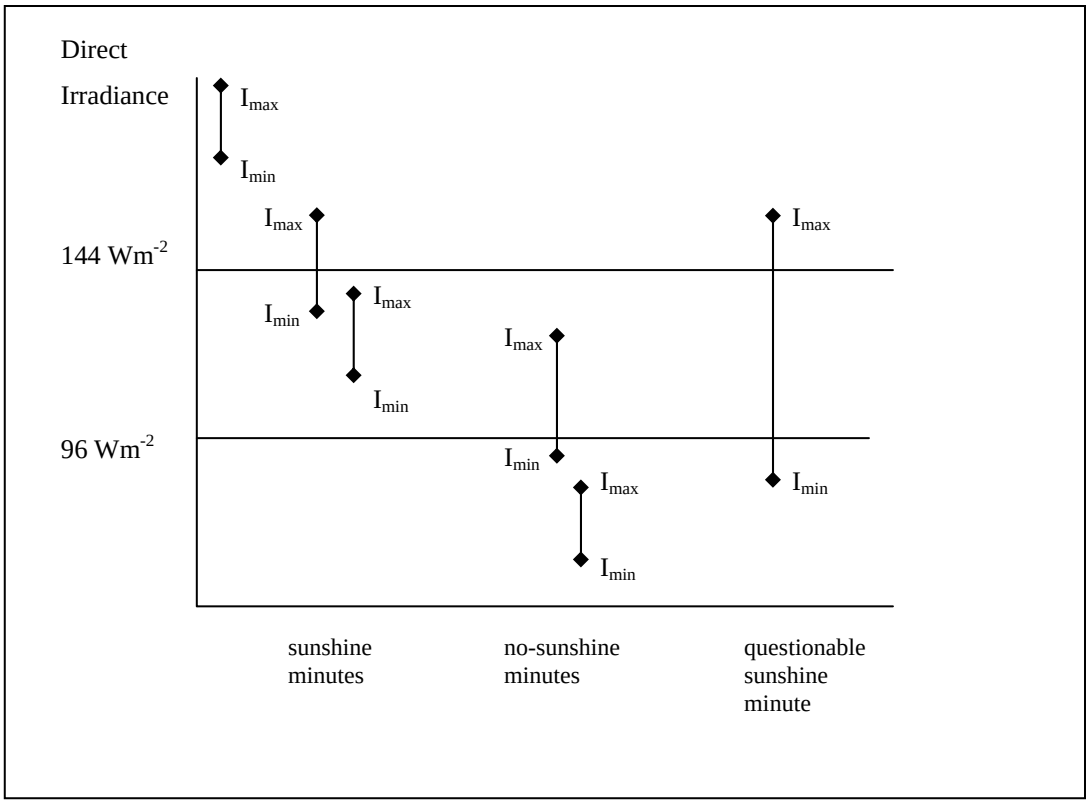

Table 4. Classification of minute data using the *threshold* with a tolerance of +/- 20%. M_s (%), M_n (%) and M_q (%) indicate M_s, M_n and M_q respectively as a percentage of the total minutes when the missing minutes are excluded.

Station Number	M_s	M_n	M_q	M_s (%)	M_n (%)	M_q (%)
201	176660	63790	12278	69.9	25.2	4.9
301	115574	122526	19319	44.9	47.6	7.5
351	186110	57903	12418	72.6	22.6	4.8
401	21244	4895	1165	77.8	17.9	4.3
402	10921	6398	1014	59.6	34.9	5.5
501	102967	119026	22169	42.2	48.7	9.1
601	89221	104633	22826	41.2	48.3	10.5
701	114320	122428	11980	46.0	49.2	4.8
801	206830	37420	7148	82.3	14.9	2.8
803	117752	26771	7947	77.2	17.6	5.2
851	211447	29451	4390	86.2	12.0	1.8
921	103695	27338	6325	75.5	19.9	4.6
922	213820	25539	4692	87.6	10.5	1.9
923	177866	59962	16165	70.0	23.6	6.4
924	160316	49049	10438	72.9	22.3	4.7
998	102191	120480	17365	42.6	50.2	7.2
Total	2110934	977609	177639	64.6	29.9	5.4

Compared with the earlier classification based upon the *threshold*, the percentage of *questionable sunshine minutes* in this case is between 0.8% and 2.2% less for each station and 1.4% less for the data set as a whole.

The number of days for which the number of *possible minutes* is greater than 6, N_{20}, are shown table 5.

Table 5. Number and % of days where more than 6 minutes in the day could not be certain of being either sunshine or not sunshine when using the +/- 20% tolerance to the limit.

Station Number	N_{20}	N_{20} (%)
201	216	60.2
301	304	83.3
351	216	59.2
401	23	54.8
402	18	75.0
501	305	85.2
601	286	89.9
701	260	72.4
801	137	38.3
803	149	67.1
851	109	31.3
921	125	67.6
922	112	32.3
923	242	66.7
924	196	62.2
998	287	82.5
Total	2985	63.8

3.3 THE RANDOM MODEL

As a first approximation, we assumed that each *questionable sunshine minute* is equally likely to be a complete minute of sunshine, or a complete minute with no sunshine. Thus a sunshine duration of 0.5 +/- 0.5 minutes was assigned to each *questionable sunshine minute*. In this way an uncertainty for the *sunshine duration* can be calculated for each day at each station.

We can calculate the number of *questionable sunshine minutes* in a day (n) before the required uncertainty (0.1 hours = 6 minutes) is exceeded:

$U_{95} = \sqrt{(n)} \times 0.5 \times 1.96 <= 6$

$n <= 37.4$

Thus the uncertainty will be less than or equal to 6 minutes when there are 37 or fewer *questionable sunshine minutes* in a day. The number and percentage of days for which this does not occur for each station, N_{random}, along with the number of days with valid data, is shown in table 6.

Table 6. Number and % of days where more than 6 minutes in the day could not be certain of being either sunshine or not sunshine when using the random model.

Station Number	N_{random}	N_{random} (%)
201	148	41.2
301	219	60.0
351	141	38.6
401	16	38.1
402	13	54.2
501	237	66.2
601	238	74.8
701	169	47.1
801	96	26.8
803	115	51.8
851	70	20.1
921	81	43.8
922	58	16.7
923	188	51.8
924	130	41.3
998	206	59.2
Total	2125	45.4

More than 5% of the days have an uncertainty greater than the *threshold* for all of the 16 stations. So U_{95} is greater than the 0.1 hours at all stations, and hence the *uncertainty requirement* is not fulfilled.

3.4 THE MEAN MODEL

The mean irradiance was used to develop a simple model of the *sunshine duration* of the *questionable sunshine minutes*. In this model we assume that the *sunshine duration* is certain for half of the interval (0.5 minutes) and assign it either 0.5 minutes of sunshine or 0 minutes of sunshine, depending upon the mean. The sunshine for the other half of the interval is unknown – equally likely to be 0.5 minutes of sunshine or 0 minutes of sunshine, and so is assigned a value of 0.25 +/- 0.25 minutes.

When the mean is greater than the *threshold*, we have a *sunshine duration (D)* of:

D = 0.5 + (0.25 +/- 0.25) minutes

= 0.75 +/- 0.25 minutes.

When the mean is less than or equal to the *threshold*, we have:

D = 0.0 + (0.25 +/- 0.25) minutes

= 0.25 +/- 0.25 minutes.

With this model, the uncertainty for each *questionable sunshine minute* is half of that for the random model. The mean model relies upon the additional assumption that the mean is an appropriate approximation of the median.

The number of *questionable sunshine minutes* in a day (n) before the *uncertainty requirement* is exceeded is:

$U_{95} = \sqrt{(n)} \times 0.25 \times 1.96 <= 6$

$n <= 149.9$

Thus the uncertainty will be less or equal to 6 minutes when there are 149 or fewer *questionable sunshine minutes* in a day. The number of days for which this does not occur, N_{mean}, is shown for each station and the overall total in table 7.

Table 7. Number and % of days where more than 6 minutes in the day could not be certain of being either sunshine or not sunshine when using the mean model.

Station Number	N_{mean}	N_{mean} (%)
201	14	3.9
301	28	7.7
351	17	4.7
401	0	0.0
402	0	0.0
501	33	9.2
601	58	18.2
701	10	2.8
801	10	2.8
803	1	0.5
851	1	0.3
921	8	4.3
922	8	2.3
923	24	6.6
924	13	4.1
998	22	6.3
total	247	5.3

For five of the 16 stations more than 5% of the days do not meet the *uncertainty requirement*. So although this model indicates satisfactory performance for most stations, it still shows that the *uncertainty requirement* cannot be met for all stations.

These results indicate that accurate calculation of *sunshine duration* is harder at the mid-latitude locations of Cape Grim (601), Mt. Gambier (501), and Melbourne (301) and Geraldton (923). Small uncertainties are attained at the more northerly stations of Rockhampton (401), Cairns (402), Darwin (803) and Tennant Creek (851). Other sites, Broome (921), Learmonth (922), Lauder (NZ, 701), Alice Springs (801) and even Wagga Wagga (201) return very low uncertainties.

3.5 SPLIT INTO THREE

The mean model can be refined, using the mean and standard deviation, to divide each minute into three parts, with each part representing one third of the distribution. For two of these parts the *sunshine duration* will be assumed known, and for the remaining part, the *sunshine duration* is equally likely to be either 0.333 minutes or zero minutes.

In order to divide the minute into three equal parts a distribution must be assumed. In this investigation we compared the results for six different distributions, in order to cover the range of distributions. We used four distributions that are symmetric about the mean: flat, normal, triangular and inverted triangular; and two asymmetric distributions: ramp up and ramp down.

The interval was divided into three parts, each having one third of the distribution. Table 8 shows the interval division points (p_1 and p_2 for the different distributions.

Table 8. Division of interval into three parts for the various distributions.

Distribution	p_1	p_2
Flat	$\mu - 0.577\sigma$	$\mu + 0.577\sigma$
Normal	$\mu - 0.431\sigma$	$\mu + 0.431\sigma$
Triangular	$\mu - 0.449\sigma$	$\mu + 0.449\sigma$
Ramp Up	$\mu - 0.379\sigma$	$\mu + 0.636\sigma$
Ramp Down	$\mu - 0.636\sigma$	$\mu + 0.379\sigma$
Inverted Triangular	$\mu - 0.816\sigma$	$\mu + 0.816\sigma$

Sunshine duration and *sunshine duration uncertainties* (U_D) were assigned to each minute using table 9 below, with the value depending upon where p_1 and p_2 lie with respect to the *threshold*.

Table 9. Assignment of *sunshine duration* and *sunshine duration uncertainties*.

	D (*minutes*)	U_D (*minutes*)
$p_2 < threshold$	0 + 0 + 0.167 = 0.167	0.167
$p_1 < threshold <= p_2$	0 + 0.167 + 0.333 = 0.5	0.167
$threshold <= p_1$	0.167 + 0.333 +0.333 = 0.833	0.167

The *total uncertainty*(U_{95}) was evaluated by calculating six *sunshine duration* values, one for each distribution, and comparing the minimum (D_{min}) and maximum (D_{max}) of these six values.

For each minute:

$U_{95} = (D_{max} - D_{min})/2 + U_D/ \sqrt{(2)}$

$= (D_{max} - D_{min})/2 + 0.1667 / \sqrt{(2)}$

The number of days with U_{95} greater than 0.1 hours, N_{split}, is shown in table 10.

Table 10. Number and % of days where more than 6 minutes in the day could not be certain of being either sunshine or not sunshine when using the "split" model.

Station Number	N_{split}	N_{split} (%)
201	105	29.3
301	172	47.1
351	107	29.3
401	13	31.0
402	9	37.5
501	201	56.2
601	194	61.0
701	109	30.4
801	61	17.0
803	78	35.1
851	45	12.9
921	51	27.6
922	41	11.8
923	137	37.7
924	94	29.8
998	156	44.8
Total	1573	33.6

Under this analysis none of the stations met the *uncertainty requirement*.

3.6 SPLIT INTO THREE – SELECT DISTRIBUTIONS

The method of 3.5 was refined by excluding the ramp up and ramp down distributions from the calculations for minutes that were not *sufficiently symmetric*. *Sufficiently symmetric* being when the mean is within the middle third of the interval between the minimum and maximum values:

$$| (I_{max} + I_{min})/2 - \mu | < (I_{max} - I_{min})/6$$

This refinement produces the results of table 11.

Table 11. Number and % of days where more than 6 minutes in the day could not be certain of being either sunshine or not sunshine when using the "select" model. N_{select} is the number of days with U_{95} > 0.1 hours.

Station Number	N_{select}	N_{select} (%)
201	101	28.1
301	169	46.3
351	105	28.8
401	13	31.0
402	9	37.5
501	197	55.0
601	190	59.8
701	107	29.8
801	59	16.5
803	75	33.8
851	44	12.6
921	50	27.0
922	39	11.2
923	134	36.9
924	92	29.2
998	150	43.1
Total	1534	32.8

Comparison with table 10 shows that there is little improvement.

The data of this latest model can also be examined to determine the value U_{95}, that is, the uncertainty of the daily sunshine value in minutes.

Table 12. Likely uncertainties in minutes for sunshine hours.

Station Number	U_{95} (*minutes*)
201	12.0
301	13.9
351	12.4
401	10.1
402	10.9
501	15.0
601	14.7
701	11.1
801	10.4
803	10.0
851	7.7
921	11.0
922	8.7
923	13.2
924	12.3
998	13.0

So under this model the uncertainty in the *sunshine duration* for the whole network is about 15 minutes. However, the results differ for station to station, going as low as 7.7 minutes for Rockhampton.

3.7 SUMMARY OF RESULTS

The results for the different models is shown in table 13. The mean model appears to give the best results, however, unlike N_{split} and N_{select} this model assumes that the mean approximates the median.

Table 13. Summary of results from all methods.

Station Number	N_0 (%)	N_{20} (%) sec 3.2	N_{random} (%) sec 3.3	N_{mean} (%) sec 3.4	N_{split} (%) sec 3.5	N_{select} (%) sec 3.6
201	68.2	60.2	41.2	3.9	29.3	28.1
301	86.3	83.3	60.0	7.7	47.1	46.3
351	67.4	59.2	38.6	4.7	29.3	28.8
401	61.9	54.8	38.1	0.0	31.0	31.0
402	95.8	75.0	54.2	0.0	37.5	37.5
501	90.5	85.2	66.2	9.2	56.2	55.0
601	93.7	89.9	74.8	18.2	61.0	59.8
701	83.3	72.4	47.1	2.8	30.4	29.8
801	45.8	38.3	26.8	2.8	17.0	16.5
803	70.3	67.1	51.8	0.5	35.1	33.8
851	36.5	31.3	20.1	0.3	12.9	12.6
921	76.2	67.6	43.8	4.3	27.6	27.0
922	38.6	32.3	16.7	2.3	11.8	11.2
923	71.6	66.7	51.8	6.6	37.7	36.9
924	67.0	62.2	41.3	4.1	29.8	29.2
998	89.9	82.5	59.2	6.3	44.8	43.1
Total	70.2	63.8	45.4	5.3	33.6	32.8

4. CONCLUSION

Given the assumption that the mean is a good approximation of the median, 11 of the stations in the Solar and Terrestrial Radiation Network satisfy that *uncertainty requirement* that the uncertainty of the *sunshine duration* be no more than 0.1 hours (6 minutes). However, as that assumption is likely to be violated, the current recorded quantities do not allow the *uncertainty requirement* to be attained.

The bulk of this analysis examined the uncertainties that arise from incomplete information and did not take into account uncertainties in the measurement of direct solar irradiance. Preliminary investigation indicated that introducing a 20% tolerance in the threshold irradiance would make the *uncertainty requirement* slightly easier to attain.

Using one of the models ("select") the likely uncertainty of the current sunshine data for any day and station is 15 minutes, or 3% of a mean solar day.

5. RECOMMENDATION

It is recommended that to achieve the WMO specification the sunshine status of each second of data should be determined and recording the number of seconds of sunshine for each minute. Suitable statistics can be achieved by recording the number of seconds for which the pyrheliometer irradiance exceeds three different values: the *threshold*, the *threshold* plus 20%, and the *threshold* minus 20%. By recording three values an uncertainty assessment can be performed on the resultant data. Initial indications are that the current data system used in the Solar and Terrestrial Radiation Network can accommodate this modification.

6. REFERENCES

[1] *Guide to Meteorological Instruments and Methods of Observation, sixth edition*, 1996, Secretariat of the World Meteorological Organization, Geneva, Switzerland, Part I, 8.1.1 for definition of *sunshine duration*; Part I, 8.1.3 for uncertainty requirement.

[2] *Project Development Plan - Radiation Network*, 1994, Bureau of Meteorology, Commonwealth of Australia.

[3] *Guide to the Expression of Uncertainty in Measurement*, Section 3.3 "Uncertainty", International Organization for Standardization, 1995.