ON THE GENERATION OF AN OPTIMIZED FRACTIONAL CLOUDINESS TIME SERIES USING A MULTI-SENSOR APPROACH*

Wiel Wauben[1], Marijn de Haij[1], Reinout Boers[2], Henk Klein Baltink[3], Bert van Ulft[3] and Mark Savenije[4]

[1] R&D Information and Observation Technology, [2] Climate Observations Department,

[3] Regional Climate Department, [4] Weather Research and Development Department

Royal Netherlands Meteorological Institute (KNMI)

P.O. Box 201, 3730 AE De Bilt, The Netherlands

Tel. +31-30-2206 482, Fax +31-30-2210 407, e-mail: Wiel.Wauben@knmi.nl

ABSTRACT

Five active and passive remote sensing techniques that observe clouds are operated nearly continuously at the Cabauw Experimental Site for Atmospheric Research [CESAR]. The instruments involved are: 2 types of ceilometers, including the sensor that is used operationally by KNMI for generating the automated cloud observations; a cloud radar; an infrared pyrgeometer; a scanning pyrometer and a total sky imager. The total cloud cover has been determined for each of these techniques at 10-minute intervals and their characteristics have been investigated. Next an algorithm was designed to derive a continuous and optimized record of fractional cloudiness that takes into account the strength and weaknesses of the individual observational techniques. The algorithm accounts for breaks in instrumental availability and assigns an error to the data. The fractional cloudiness characteristics of the individual instruments and the synthetic algorithm have also been evaluated against a climatology of human cloud observations.

The paper presents details and the challenges encountered in the evaluation of the cloudiness obtained with the individual instruments and the construction of the optimized fractional cloudiness series. Particularly the lack of a reference is a serious issue. The measure of uncertainty of the resulting series and the relative contribution of each instrument will be discussed.

1. INTRODUCTION

Cloud fraction or coverage has been measured by human observers for well over 100 years. For that purpose an observer looks at the sky and estimates cloud fraction in okta and the height of the cloud base. The widespread global use of this observation technique together with the presence of many very long time series is evidence of its usefulness to meteorology and its potential applicability for climate research and monitoring. Cloudiness is regarded as a priority 2 Essential Climate Variable [WMO, 2007]. Presently the human observer is under threat. Cost reductions have already resulted in automation of the cloud observations at the expense of human observers. This transition from human observer to instrument for the purpose of measuring fractional cloudiness has not always been made with proper consideration of the manner in which the introduction of new instruments would impact the continuity / accuracy of the time series. The transition has resulted in unfortunate discontinuities in the time series of cloud observations that can not be rectified a posteriori. It is therefore paramount to investigate the value of these observational techniques that can or already have replaced the observer. The differences in cloud characteristics introduced by the automation at KNMI using cloud ceilometers in 2002 are described in Wauben [2002] and Wauben et al. [2006].

* This work is based on the paper entitled “Optimized Fractional Cloudiness Determination from Five Ground - based Remote Sensing Techniques” by R. Boers, M.J. de Haij, W.M.F. Wauben, H. Klein Baltink, L.H. van Ulft, M. Savenije and C.N. Long that has been submitted to J. Geophys. Res. - A.

The purpose of this paper is to provide a complete and continuous time series of fractional cloudiness at a single site based on combinations of several different instruments that can be substituted in case of partial failures of individual instruments. The observations were carried out at the Cabauw Experimental Site for Atmospheric Research (CESAR, 51°58'N, 4°55'E) in the central western flat rural region of the Netherlands. CESAR is one of the candidate sites for the GCOS Global Reference Upper Atmosphere Network [GRUAN] so that the study benefits from the many synergistic observational techniques that are employed there but also can provide a guideline for continuous cloud observations at a reference site. In this paper cloudiness or total cloud amount is considered, i.e. the fraction of the sky covered with cloud irrespective of its height, although the height is used in the analysis since the detection threshold of cloud is generally height dependent. A time period of one year duration was chosen for this study [15 May 2008 – 14 May 2009]. The study was designed in such a manner that continuity of observations was given the highest priority. Also, the algorithm for deriving fractional cloudiness was constructed such that it chooses the optimum combination of instruments at all times of observation. Thus, it withstands the unavoidable instrumental failures by changing to different combinations of instruments at the moment that such failures occur, and changing back when they are available again.

2. INSTRUMENTATION

At CESAR two ceilometers are in continuous operation, namely a Vaisala LD40 and a Vaisala CT75K. A ceilometer is a low-powered LIDAR [LIght Detection And Ranging]. The ceilometers report the cloud base with a vertical resolution of 25 ft every 15 seconds (LD40) and 15 meters every 30 seconds (CT75K) respectively, although for higher clouds it is common to use a longer integration period e.g. up to 10 minutes for the LD40. The LD40 is the ceilometer that is used at about 40 locations in the meteorological network of KNMI and is capable of detecting clouds up to about 13.6 km although generally the sensitivity is reduced above 7 km due to atmospheric turbidity. The CT75K normally can detect clouds up to the stratosphere with a maximum range 11.2 km, but in the time period used here, the sensor was at the end of its life time and the maximum altitude for cloud detection was about 4 km. The ceilometers have a narrow field of view of about 1 mrad and sample only a small part of the sky directly overhead. A time series of cloud base measurement is used to derive fractional cloudiness. In this study, cloudiness is determined by the fraction of the ceilometer data per 10-minute interval with clouds. This is identical to the total cloud cover reported for aeronautical purposes by KNMI [Wauben et al., 2006].

At CESAR a Degreane 35 GHz radar is operational and capable of detecting clouds up to an altitude of 12 km. The beam width of the cloud radar in operation at CESAR is 0.3 degrees. Cloud base detection for low water clouds is often problematic for 35 GHz cloud radar systems because cloud droplets are very small near the cloud base. This means that the lowest height of water cloud detection by radar does not necessarily correspond to actual cloud base. As a result the radar / LIDAR combination is often used to detect the complete spectrum of clouds throughout the troposphere. One such procedure, the Cloudnet procedure is employed in this study. The Cloudnet algorithm takes the radar data spacing in time (approximately 15 s) and height (89 m) as the grid to which the other observations like the CT75K ceilometer backscatter are mapped. For each grid a set of rules determines the class of the data in the grid cell, e.g. rain, ice, mixed phase, etc. Details of this Cloudnet study and procedures to calculate cloudiness are provided by Illingworth et al. [2007]. In this study cloudiness is calculated from the number of profiles within the 10 minute-period that contain a cloud hit.

At Cabauw a Baseline Surface Radiation Network [BSRN] site has been operational since 2005. One minute averaged longwave downward radiation data from the infrared pyrgeometer (Kipp & Zonen, CG4) was used as input for the Automated Partial Cloud Amount Detection Algorithm APCADA) [Dürr and Philipona, 2004]. Here, as a first step the clear sky radiation must be estimated, which is done by a parameterization of the standard atmosphere using measured temperature and humidity at screen level height. Next, the algorithm relates the difference between the actually observed radiation and the calculated clear sky radiation to a cloud index. At the last step the algorithm takes into account the 60-minute variability of the longwave downward radiation to be able to distinguish between cloud fraction types: broken clouds strongly influence the variability signal, while overcast and cloud free lead to a low variability. By use of a lookup table an okta value for every 10-minute interval is derived given the cloud index and the variability. The APCADA algorithm has been tuned for seasonal and diurnal variations in the atmosphere at Cabauw using 3 full years of BSRN observations.

The NubiScope system, manufactured by IMK/Sattler-SES, is a passive remote sensing instrument which consists of a pyrometer mounted on a pan-and-tilt unit. The pyrometer measures in the spectral range of 8-14 μm, i.e. in the atmospheric thermal infrared window, and has a field of view of 3°. Every 10 minutes a full scan of the sky is performed which takes about 6½ minutes. The NubiScope detects cloudiness when the atmospheric brightness temperature is above the clear sky background. In the 8-14 μm window there is a

contribution of water vapor to the measured brightness temperature. Hence the clear sky brightness temperature increases with zenith angles and varies over time. For that purpose the NubiScope adapts the clear sky reference dynamically at each scan when sufficient cloud free scenes at various elevations are available [Collet et al. 2009]. The NubiScope scan at 30 zenith angles (1.5° to 88.5° in steps of 3°) and 36 azimuth angles (5° to 355° in steps of 10°) covers the whole sky from zenith to horizon, but the cloud determination is only performed for zenith angles <70°. The brightness temperatures near the horizon are used to estimate the ambient temperature and furthermore the measurements at low elevations are used to discriminate between cloudiness and fog. Details of the NubiScope are given in Wauben [2006] and Wauben et al. [2010].

The Total Sky Imager (YesInc, TSI-440) consists of a digital camera pointed downward at a hemispheric mirror on which the reflection of the sun is blocked by a dark strip [shadow band]. The digital information is stored in color images (JPEG format), so that information on the relative red, green, and blue content is preserved. The TSI is set-up to take one sky image every minute when the sun is more than 5 degree above the horizon. Cloudiness analysis is performed when the solar elevation is more than 10 degrees above horizon. The sky FOV analyzed is 160 degrees. The TSI software determines the fraction of thin and opaque clouds in the FOV based on two thresholds for the red/blue ratio of each pixel. From visual assessment of the sky images recorded and the related output of the TSI we determined that most suitable values for the clear/thin and thin/opaque threshold are 65 and 90 respectively. This compares well with values found from an analysis of TSI images at Desert Rock SURFRAD site, USA [Hodges, 2003]. All images were reprocessed using the above settings and small corrections for misalignment of the shadow band were applied as well. A second post-processing step was performed to correct for cloud classification errors due to forward scatter near the sun and the horizon at low solar elevation [Long et al., 2006, Long, 2010].

Note that all techniques except APCADA report cloudiness in percentage which is converted into okta when required. According to the guidelines of the World Meteorological Organization [WMO, 2008] the presence of only a single cloud on a clear sky should be reported as 1 okta and similarly a tiny gap in an otherwise closed cloud deck is reported as 7 okta, since 0 and 8 okta correspond to completely clear (0 %) and completely cloudy situations (100 %). This means that 1 okta and 7 okta are associated with a larger range of fractional cloudiness conditions (18.75 %) than the intermediate (12.5 %) where the fraction corresponds to the nearest okta value. This is an issue that cannot be avoided in the analysis.

3. INTERCOMPARISONS OF TECHNIQUES

The techniques can be divided in to two main groups: column and hemispheric methods. The ceilometer and Cloudnet are column methods that sample the sky only in zenith direction, while NubiScope, TSI and APCADA are hemispheric methods that sample the sky over a large FOV from zenith (almost) down to the horizon. All techniques were intercompared to assess relative differences in output. In this section this is illustrated by a comparison of all techniques to the NubiScope in the form of contingency matrices. The matrix for the LD40 and the NubiScope is shown in Table 1. A total of 42197 data points, when the results of both techniques were available, are considered in this intercomparison. The okta indicated in the vertical direction represent the data from the LD40, the okta in the horizontal direction are those from the NubiScope. The meaning of each value in the colour codes table can be understood as follows: The 9.0% in the top left cell means that for 9.0% of all cases, the NubiScope and the LD40 technique yield identical values of cloudiness, namely 0 okta. In the bottom left, similarly for 26.8% of all cases both techniques yield 8 okta. The number 0.0 in the lower left indicates that for 0.0% of the values the LD40 cloudiness is 8 okta, while at the same time the NubiScope yields 0 okta. Or conversely, for the top right number of 0.4 means that according to the LD40 it is completely clear at 0 okta for 0.4% of the time, while according to the NubiScope instrument it is completely overcast at 8 okta.

The column and row denoted by "%" give the percentage of observations in an okta bin for LD40 and NubiScope, respectively. The okta distribution will be discussed in detail below. The colour coding expresses the degree of agreement between the outputs of the two procedures. If the cell is green, both procedures yield the same cloudiness in okta. According to the numbers at the bottom this happens 45.9% of the time. If it is yellow there is 1 okta difference. So, in 81.0% of the data points, there is 1 okta or less difference between the two procedures. For red and cyan cells the difference in okta between procedures is larger than 2 okta. In fact for 6.7% of cases the LD40 data yields fractional cloudiness values that are at least 3 okta smaller than the NubiScope, while in 4.5% of cases they are at least 3 okta higher than the NubiScope. So, in other words in 11.2 % of the cases the NubiScope and the LD40 differ from each other in fractional cloudiness determination by more than 2 okta, a significant fraction of the time. The last line indicates that the averaged difference in fractional cloudiness ($\Delta = \langle n_{LD40} - n_{NubiScope} \rangle$) is slightly negative ($> -0.05$) whereas the averaged absolute deviation ($|\Delta| = \langle |n_{LD40} - n_{NubiScope}| \rangle$) is 1.0 okta. The slight underestimation of cloudiness

by the LD40 is probably the result of the lesser sensitivity of the LD40 to high clouds as compared to the NubiScope [Wauben et al., 2010].

Table 1: Contingency matrix of LD40 (vertical) versus NubiScope (horizontal) cloud cover fraction for the period 15 May 2008 – 14 May 2009.

	0	1	2	3	4	5	6	7	8	%	count
0	9.0	8.0	1.0	0.4	0.3	0.2	0.2	0.7	0.4	20.2	8504
1	3.0	3.5	1.0	0.5	0.3	0.2	0.2	0.7	0.4	9.7	4096
2	0.6	0.9	0.5	0.4	0.2	0.2	0.1	0.4	0.2	3.4	1449
3	0.3	0.5	0.5	0.4	0.3	0.1	0.1	0.5	0.2	2.9	1216
4	0.2	0.4	0.4	0.5	0.4	0.2	0.2	0.6	0.2	3.3	1409
5	0.1	0.3	0.3	0.4	0.4	0.4	0.3	0.7	0.3	3.1	1298
6	0.1	0.1	0.2	0.3	0.4	0.4	0.3	1.0	0.4	3.3	1375
7	0.0	0.2	0.1	0.3	0.6	1.0	1.3	4.5	2.7	10.7	4512
8	0.0	0.2	0.1	0.1	0.2	0.4	1.2	14.5	26.8	43.5	18338
%	13.4	14.1	4.1	3.3	3.0	3.1	3.7	23.7	31.6	100.0	
count	5639	5944	1733	1389	1281	1329	1562	9981	13339		42197

Band 0: 45.9% **Band 1: 81.0%** **Band 2: 88.8%** **Over: 4.5%** **Under: 6.7%** **Δ: -0.0** **|Δ|: 1.0**

Table 2: Contingency matrix of Cloudnet (vertical) versus NubiScope (horizontal) cloud cover fraction for the period 15 May 2008 – 14 May 2009.

	0	1	2	3	4	5	6	7	8	%	count
0	11.3	7.6	0.7	0.2	0.1	0.0	0.0	0.1	0.1	20.0	7551
1	1.3	2.6	0.8	0.3	0.1	0.1	0.0	0.1	0.0	5.2	1970
2	0.0	0.7	0.4	0.3	0.2	0.1	0.0	0.0	0.0	1.7	643
3	0.0	0.5	0.4	0.4	0.2	0.1	0.1	0.1	0.0	1.8	692
4	0.0	0.3	0.3	0.3	0.3	0.2	0.0	0.1	0.0	1.5	581
5	0.0	0.3	0.3	0.3	0.4	0.3	0.1	0.1	0.0	1.9	734
6	0.0	0.2	0.2	0.3	0.4	0.4	0.3	0.3	0.0	2.1	795
7	0.0	0.4	0.3	0.3	0.5	0.7	0.9	1.2	0.2	4.5	1708
8	0.1	1.3	0.7	0.7	0.8	1.2	2.1	21.8	32.4	61.1	23018
%	12.9	13.8	4.0	3.2	2.9	3.1	3.6	23.7	32.8	100.0	
count	4867	5197	1509	1203	1106	1163	1375	8923	12349		37692

Band 0: 49.2% **Band 1: 85.1%** **Band 2: 90.9%** **Over: 8.0%** **Under: 1.1%** **Δ: +0.5** **|Δ|: 0.9**

Table 3: Contingency matrix of APCADA (vertical) versus NubiScope (horizontal) cloud cover fraction for the period 15 May 2008 – 14 May 2009.

	0	1	2	3	4	5	6	7	8	%	count
0	8.6	3.5	0.1	0.0	0.0	0.0	0.0	0.0	0.0	12.4	5195
1	3.7	5.6	0.9	0.4	0.2	0.2	0.1	0.4	0.2	12.0	5053
2	1.1	4.1	1.8	1.1	0.7	0.4	0.3	0.7	0.2	10.4	4362
3	0.1	0.6	0.9	1.0	0.8	0.6	0.4	0.8	0.1	5.3	2212
4	0.0	0.1	0.2	0.4	0.6	0.6	0.5	1.1	0.2	3.8	1585
5	0.0	0.1	0.0	0.0	0.1	0.1	0.1	1.2	0.6	2.2	913
6	0.0	0.0	0.1	0.2	0.6	1.1	1.8	6.1	0.8	10.8	4537
7	0.0	0.0	0.0	0.0	0.0	0.1	0.4	10.4	6.3	17.2	7240
8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.4	23.7	26.1	10965
%	13.5	14.1	4.1	3.3	3.0	3.1	3.7	23.2	31.9	100.0	
count	5695	5930	1725	1392	1272	1314	1545	9760	13429		42062

Band 0: 53.7% **Band 1: 86.2%** **Band 2: 93.1%** **Over: 0.8%** **Under: 6.1%** **Δ: -0.3** **|Δ|: 0.7**

The Cloudnet results (cf. Table 2) shows better agreement with the NubiScope than the LD40. Cloudnet reports, however, significantly higher cloudiness values (Δ=+0.5) due to its higher sensitivity to cirrus cloud. APCADA (Table 3) shows even better overall agreement with the NubiScope than the previous 2 two column techniques. APCADA underestimated cloudiness (Δ=–0.3) which might be the result of sensitivity and the fact that it uses integrated hemispheric information. TSI and NubiScope (cf. Table 4) show the best overall agreement, which could be expected since both are true hemispheric techniques. Concerning the differences of more than 2 okta TSI versus NubiScope shows the best result (only 6.6% of cases where the okta values differ by 3 okta or more), and the comparison between the TSI and the LD40 show the worst result (17.7%). Furthermore, all intercomparisons between the three hemispheric techniques for okta differences in excess of 2 yield percentage values below 10%, while all other intercomparisons yield values in excess of 10 %, with the exception of the NubiScope versus Cloudnet (9.1%). In other words, the hemispheric methods tend to agree more with each other, than with the column methods. Also, the two column techniques (Cloudnet versus LD40, 13.6%) do not agree so much among themselves as the hemispheric methods tend to agree among themselves.

The columns of Tables 1 to 3 show generally the same numbers, as could be expected since nearly the same set of NubiScope data is always used. In Table 4, however, only day time data is considered for the comparison with TSI data. The column totals shows that the fraction of overcast (8 okta) and particularly clear sky (0 okta) cases is reduced during day time while the fraction of cases in 1 to 7 okta bins increase slightly. This difference between day and night time cloud distributions can be observed for all measuring techniques (except APCADA at 5, 7 and 8 okta) and can also be observed in the climatology of human observations (except at 1 and 2 okta) as is shown in Boers et al. [2010].

Table 4: Contingency matrix of TSI (vertical) versus NubiScope (horizontal) cloud cover fraction for the period 15 May 2008 – 14 May 2009.

	0	1	2	3	4	5	6	7	8	%	count
0	2.8	1.8	0.0	0.0	0.0	0.0	0.0	0.1	0.1	4.8	829
1	4.0	9.5	1.6	0.5	0.2	0.1	0.1	0.2	0.1	16.1	2777
2	0.2	1.3	1.7	0.9	0.3	0.2	0.2	0.2	0.1	5.1	873
3	0.4	1.2	1.1	1.4	0.8	0.4	0.3	0.4	0.1	6.1	1046
4	0.2	0.6	0.5	1.0	1.1	0.8	0.4	0.6	0.1	5.4	937
5	0.1	0.3	0.3	0.4	0.8	1.0	0.8	0.9	0.1	4.7	805
6	0.0	0.2	0.2	0.2	0.5	0.9	1.1	1.8	0.3	5.3	921
7	0.0	0.1	0.1	0.2	0.4	0.7	1.6	13.7	8.0	24.8	4273
8	0.0	0.0	0.0	0.0	0.0	0.0	0.1	7.7	19.9	27.8	4795
%	7.7	15.0	5.5	4.5	4.1	4.3	4.5	25.7	28.7	100.0	
count	1334	2585	947	784	713	734	778	4430	4951		17256
Band 0: 52.1% **Band 1: 87.0%** **Band 2: 93.4%** **Over: 3.5%** **Under: 3.1%** **Δ: +0.0** **|Δ|: 0.7**											

No Observer data was available in the period of interest [15 May 2008 – 14 May 2009] to verify the cloudiness data taken by the instruments. Therefore it is useful to compare the one-year data records against the statistical Observer cloudiness information that has been collected in the period 1971 – 2000, the last 30 years for which a climatological average based on human observations can be constructed at KNMI. Figure 1 shows a histogram of all instrumental data plotted together with the histogram based on the Observer records. The latter is an average calculated for the human cloud observations of Rotterdam and of De Bilt. Both stations are almost equidistant from CESAR (22 km), but Rotterdam is located near the coast, while De Bilt is located some 45 km away from the coast. Superimposed on the Observer record are the absolute maximum and minimum values of the annual frequency of occurrence recorded in the 30 years.

The column methods LD40 and Cloudnet show a tendency for cloudiness observations at 0 and 8 okta that fall well outside the range that could be expected based on 30 years of climatological data. This is particularly so for Cloudnet, which measures 8 okta no less than 61% of the time. Because of their excessive emphasis on 0 and 8 okta Cloudnet and LD40 are reduced in the intermediate values of cloudiness (2 – 6) when compared to the Observer. Weakening of the 0 % and 100 % thresholds for reporting 0 and 8 okta would increase the cloudiness observations at 0 and 8 okta even further. The enhanced cloudiness observations at 0 and 8 okta of LD40 and Cloudnet are the result of their lack of spatial representativeness. This is illustrated in Figure 2, which shows the frequency distribution per okta interval of the cloudiness at CESAR obtained with the LD40 and with the NubiScope by evaluating different portions of the sky. A NubiScope curve labelled by Z < x° denotes that the cloudiness is derived for the NubiScope for values of the zenith angles smaller than x°. Ideally, if the NubiScope and LD40 are equivalent instruments, their distributions should be the same. However, full correspondence between the degraded NubiScope output (Z < 3°) and the LD40 is not achieved due to differences in FOV, time window and sensitivity to cloudiness. As the NubiScope angle is increased to Z < 30° and Z < 69° the number of 0 and 8 okta cases reduces significantly whereas the number in the other okta bins and particularly 1 and 7 okta increases. Hence by processing a larger fraction of the sky the NubiScope cloudiness more and more starts to resemble that of the Observer and more often detects isolated clouds in clear sky situations or gaps in overcast situations that increase the 1 and 7 okta cases at the expense of 0 and 8 okta, respectively. Again, full correspondence between the NubiScope and Observer is not achieved since the observer views the entire sky and also an observer evaluates the cloudiness as if he were looking vertically upward at each point of the sky. In practice this means that an observer determines the cloudiness by evaluation of the observed cloud bases as far as possible in order to avoid the so-called “packing” effect, i.e. the screening of gaps in the cloud deck when

looking at slant angles at clouds in the foreground. In Figure 1, the TSI, which only yields daytime observations, is the highest in the intermediate range 2 – 6 okta. The increased number of 1 – 7 okta cases during day time was also observed for the other observing techniques (cf. e.g. Table 4). The APCADA values at 2 and 6 okta are raised with respect to the others, and are in fact even well above the Observer values. Apparently this behaviour is a manifestation of the peculiarities of determining cloudiness based on the two parameter discrimination method inherent to the APCADA algorithm. The orange columns are the values based on the reference algorithm to be discussed below.

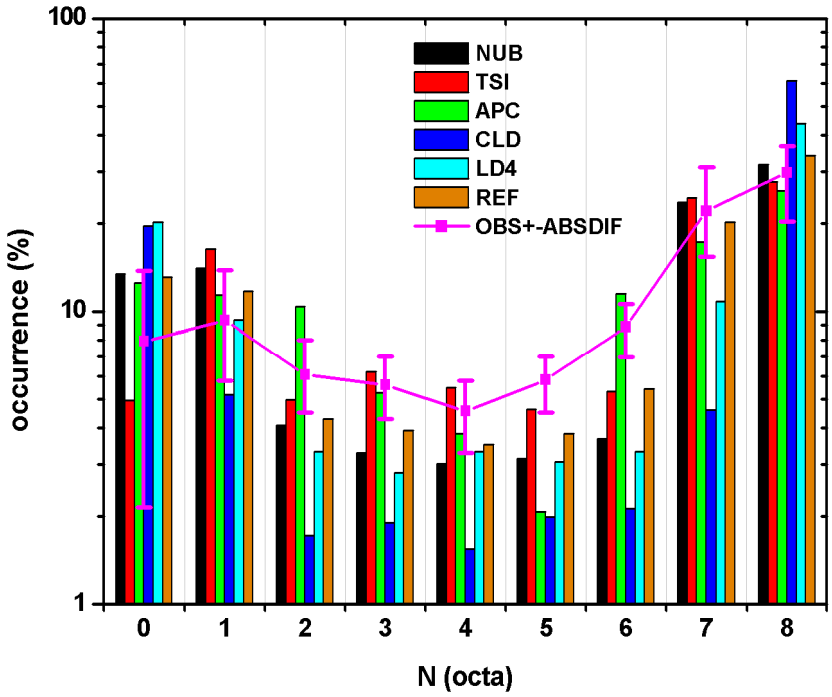

Figure 1: Cloud cover fraction histogram for all instruments and the reference algorithm (cf. Section 4). The Observer data combined from Rotterdam and De Bilt are showed in the purple line with error bar. The error bar denotes the absolute maximum and absolute minimum values for the last 30 year climate record [1971 – 2000].

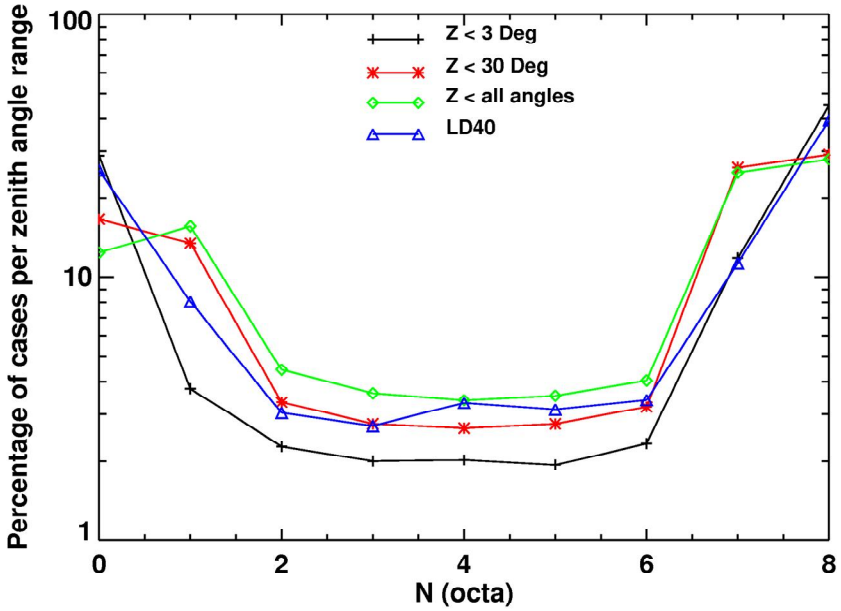

Figure 2: Percentage of observations in a zenith angle range as a function of okta value, for the NubiScope. Z < 3, 30, and all angles indicate the zenith angle range that was used to process the NubiScope data. The LD40 data is shown for comparison.

Generally it can be stated that the higher the cloud, the more difficult will it be to detect it for each cloud observing technique. Not all techniques give information on the cloud base height (APCADA and TSI give no height information at all and the height information from the NubiScope – which is inferred from the observed sky brightness temperature assuming a standard temperature profile – is rather poor). In order to investigate

the effect of cloud base height the fractional cloudiness for each cloud observing technique is shown as a function of measured Cloudnet 10-minute lowest cloud base in Figure 3. It should be stressed that this plot is not the same as fractional cloudiness as a function of height. An 11 year [1990 – 2000] climatological record of fractional cloudiness versus cloud base height for De Bilt and Rotterdam combined is shown for comparison. During these years the observers at De Bilt and Rotterdam had access to the data of a local ceilometer in order to be able to make an accurate assessment of the cloud base height. Interestingly the results from the Observer show a distinctly different height profile, with a large portion of the troposphere below 6 km altitude with little variation in cloudiness [50 – 60%]. Based on Figure 3 we can conclude that TSI and the NubiScope largely yield similar cloudiness information even though the techniques are entirely different and TSI only day time cloudiness. Both show little variation in cloudiness [70 – 80%] between 1 and 4 km and between 6 and 8 km [40 – 50%] and generally report higher cloudiness than the Observer. LD40 and APCADA give similar values between 2 and 3 km, but for higher cloud base values both give significant lower cloudiness than the Observer. Cloudnet by far report the highest cloudiness [80 – 90%] with little variation between 1 and 7 km. The Cloudnet technique combines output from radar and LIDAR. We believe that the ability of Cloudnet to detect very thin and high clouds together with the fact that it combined two column techniques may be a reason why Cloudnet produces high cloudiness values and skews the histogram values of cloudiness as presented in Figure 1 so highly towards 8 okta. The correctness of the Cloudnet detection of high clouds was shown on several occasions by using data of a collocated Raman LIDAR system at CESAR, which in fact often reported even more high clouds.

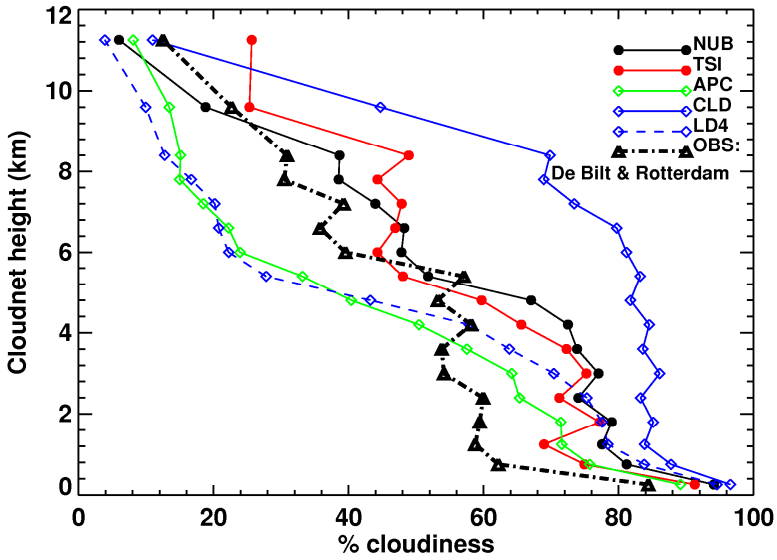

Figure 3: Fractional cloudiness as a function of cloud base height as determined by the Cloudnet procedure. Also shown is the climatology of observed values from the De Bilt and Rotterdam [1990 – 2000].

4. REFERENCE ALGORITHM

The first attempts to construct a reference algorithm used the height and cloudiness information of all observing techniques and applied empirical rules on how to combine the results in various situations. The rules were obtained from specific cases which illustrated strengths and shortcomings of each technique. However, this approach was rather subjective since no true reference for cloudiness was available and often the rules could not be applied universally. In order to obtain a robust and general reference algorithm that provides a continuous set of optimized cloudiness values a more basic approach was finally adopted based on the results shown in the previous section. If cloud base height is low, then there seems to be not much difference between methods, and any of the methods (if available) could have been used to compute the reference cloudiness values. If the cloud base height increases then it appears that the NubiScope and the TSI yield comparable results, but the LD40 and APCADA yield much smaller values than the NubiScope / TSI, while the Cloudnet procedure yields much larger values. We decided to base our reference algorithm on a weighting of the individual instruments taking into account the cloud base height as a discriminating factor. Generally, the reference cloudiness is determined in the following manner:

$$R_j = \frac{\sum_i H_{i,j} \bullet W_{i,j} \bullet C_{i,j}}{\sum_i H_{i,j} \bullet W_{i,j}} \qquad (1)$$

where R_j is the reference cloudiness (in percentage) at time j, $W_{i,j}$ is the weighting value at time j for the i-th instrument where i = 0 for NubiScope, 1 for TSI, 2 for APCADA, 3 for Cloudnet and 4 for LD40, $H_{i,j}$ = 1 when

the i-th instrument has a valid output at time j, and $H_{i,j} = 0$ when there is no valid output, and $C_{i,j}$ is the cloudiness (in percentage) measured by the i-th instrument at time j. Of course, of decisive importance in this analysis is the choice of $W_{i,j}$.

The first decision involves the NubiScope and the TSI instrument. Even though on a point-by-point basis, the okta differences of 1 or 2 are common, the analysis in the previous section together with inspection of the individual time series strongly suggests that their output should be considered equivalent. Thus, we decided to put $W_{0,j} = W_{1,j} = 1$.

The second decision involves the three other methods (APCADA, Cloudnet and LD40), for which this equivalence does not hold. The analysis of the previous section clearly suggests that the departure of the output from the remaining methods from the NubiScope and TSI is a strong function of cloud base altitude. Therefore a plausible weighting function would decrease the emphasis of the algorithm on the remaining instruments with increasing cloud base altitude (see Figure 3), We experimented with several weighting functions but in the end decided on a simple model whereby the value of the weighting was exponentially dependent upon cloud base altitude by the Cloudnet procedure within the 10-minute time interval:

$$W_{i,j} = \exp(-D_{3,j} / D_{scale}) \tag{2}$$

where $D_{3,j}$ is the observed minimum Cloudnet cloud base height in the 10-minute period at time j, D_{scale} is set at 1.3 km, which is based on rapidly increasing difference between the NubiScope / TSI output with respect tot the others.

Although they form the minority of data there are some periods where there is no altitude information. Of course during those periods the other available data should be used. Here we assume that the cloud base height distribution for those periods that the Cloudnet could did not produce output was the same as the annual mean observed cloud base height distribution. The mean weight was determined as:

$$\overline{W} = \frac{\sum_j H_{3,j} \bullet \exp(-D_{3,j} / D_{scale})}{\sum_j H_{3,j}} = 0.83 \tag{3}$$

where $H_{3,j} = 1$ when Cloudnet has a valid cloud base at time j and otherwise $H_{3,j} = 0$. This mean weight corresponds to a cloud altitude of 1.58 km.

We define two uncertainties E_{i+} and E_{i-} related to the reference cloud algorithm according to:

$$E_{i+} = \langle \frac{1}{N_+} \sum_{j,I+} (R_j - C_{i,j})^2 \rangle^{1/2} \text{ for all } R_j - C_{i,j} > 0 \tag{4a}$$

and

$$E_{i-} = \langle \frac{1}{N_-} \sum_{j,I-} (R_j - C_{i,j})^2 \rangle^{1/2} \text{ for all } R_j - C_{i,j} < 0 \tag{4b}$$

where I+ and I- are summations for which the difference between the reference cloudiness value and the cloudiness from the i-th instrument is positive and negative, respectively. The range (E_{i-}, E_{i+}) represents the uncertainty of the reference cloudiness value derived from the individual instruments. Note that the uncertainties E_{i+} and E_{i-} are averaged parameters over all time intervals in the year of analysis, but if desired a daily or even an instantaneous uncertainty of the reference cloudiness can be deduced by adjusting the summation over time j accordingly.

5. RESULTS AND DISCUSSION

Useful information is now provided by comparing the okta output from the individual instruments with the one obtained from the reference algorithm. If the reference algorithm provided an absolute reference, then a comparison would yield objective accuracy / precision information for each technique. Since the reference is not a true absolute reference but based on a subjective choice by the investigators, a comparison only provides a loose guideline for the performance of one technique versus the other. Also it should be remembered that the reference algorithm is based on a combination of instruments, so that the variations in the output from the reference algorithm are also partly based on those from the individual instruments, and on the times when selected instruments are available. Nevertheless as we shall see there may be some general conclusions to be drawn.

The contingency matrices for the cloudiness of each of the 5 observing techniques versus the reference are given in Tables 5 to 9 below. As could be expected the NubiScope cloudiness is closest to the reference followed by TSI, then APCADA (which shows an underestimation) and finally Cloudnet (which shows an overestimation) and LD40. Getting back to the okta histograms that were discussed earlier, the orange bars in Figure 1 represent the okta histogram for the reference. Except for okta values 2 - 3 and 5 - 6 all cloudiness frequencies based on the reference algorithm fall within the absolute maximum and minimum for the 30 year period of human observations (1971 – 2000). The references values are on the low side in the 2-6 okta range, but they are better than the NubiScope, Cloudnet and LD40 results individually since the references utilizes the higher values reported by TSI (although this is partly a day-time only effect) and APCADA. Although this should not be used as an absolute discriminator for the validity of the reference cloudiness output, it demonstrates an elegant compromise whereby the NubiScope and TSI are considered to be a higher quality product that the others, but whereby the output from the others is not considered to be valueless. Thus we find that for the okta values in the mid range 3 - 6 the reference cloudiness is raised above the excessively low values from either Cloudnet or LD40, about midway [but not exactly midway] between the NubiScope and the TSI.

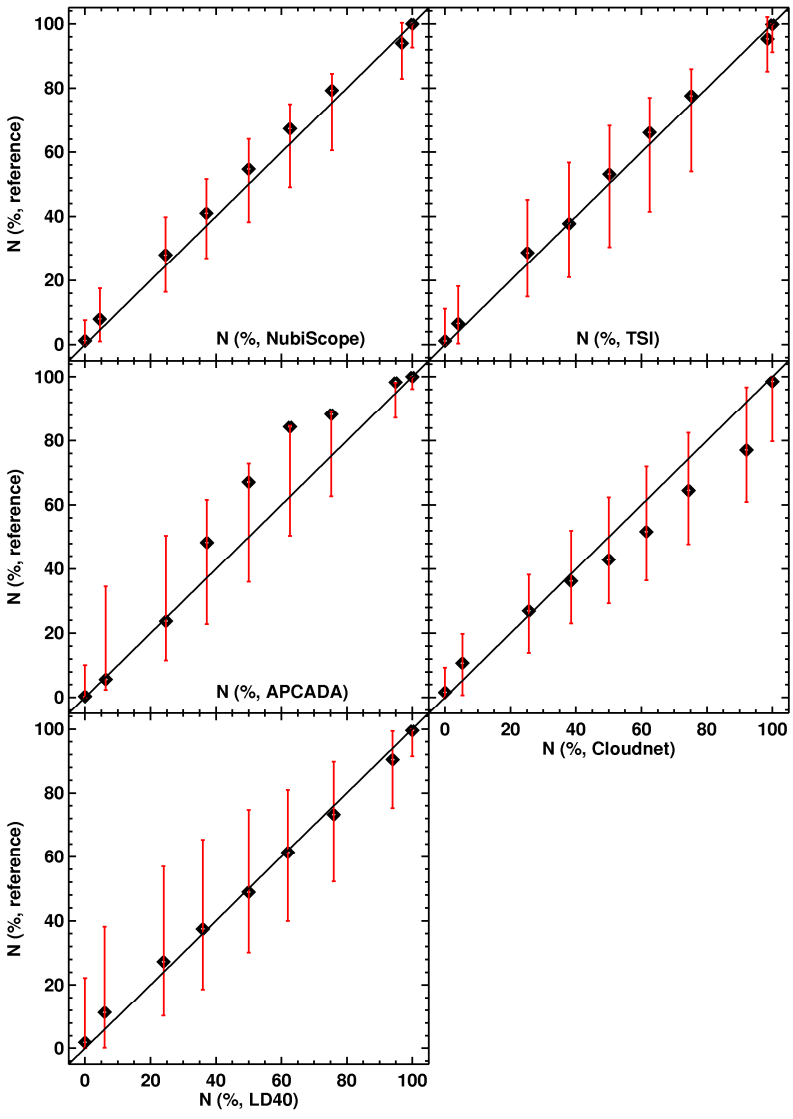

Figure 4: Comparison of cloudiness reported by individual instruments versus the reference algorithm.

Figure 4 shows a set of panels comparing the cloudiness of each instrument against the reference algorithm for the nine cloudiness ranges that correspond to the 0-8 okta observations. The errors as determined with Equation 4 were plotted as red error bars on the graphs. The individual points represent the mean value for R and C. These figures may be best understood by focusing on a couple of examples. In the APCADA panel we note that the reference output values usually exceed those from the APCADA algorithm, while for Cloudnet the opposite occurs. This means that if APCADA / Cloudnet would be the only instruments to determine the fractional cloudiness reported by the reference algorithm, then its values would be

systematically too low / high. The asymmetric error bars inform us on the two-sided uncertainty in the observations. For example for the APCADA N = 6 okta value the uncertainty bar for APCADA > Reference Value appears to be almost non-existent. This means that the observations of APCADA are so systematically highly skewed towards the lower side that most of the time the true cloudiness is expected to be higher than observed by APCADA, and according to the size of the error bar on the lower side of this point it is expected to be 30% higher. From the previous section we know that this is caused by the fact that as cloud height increases, APCADA will systematically underestimate cloudiness. The opposite is true for Cloudnet. We know that 61% of Cloudnet observations are in the 8 okta bin. This means that it probably overestimates cloudiness most of the time although the situation is not as imbalanced as APCADA, judging from the error bars. Based on this analysis it can be concluded that in the intermediate range of N = 2 – 6 it is, as a rule, safe to impose an error bar of at least 15% on either side of the observations. As expected based on our choices, the error bars on the NubiScope and TSI observations are the smallest. For the LD40 larger error bars can be imposed (20%), while APCADA and Cloudnet often have large offsets from the reference values. Again it should be stressed that these numbers are only to be used as a general guideline; they are not precise measures due to the interdependence of the reference data on the individual instruments.

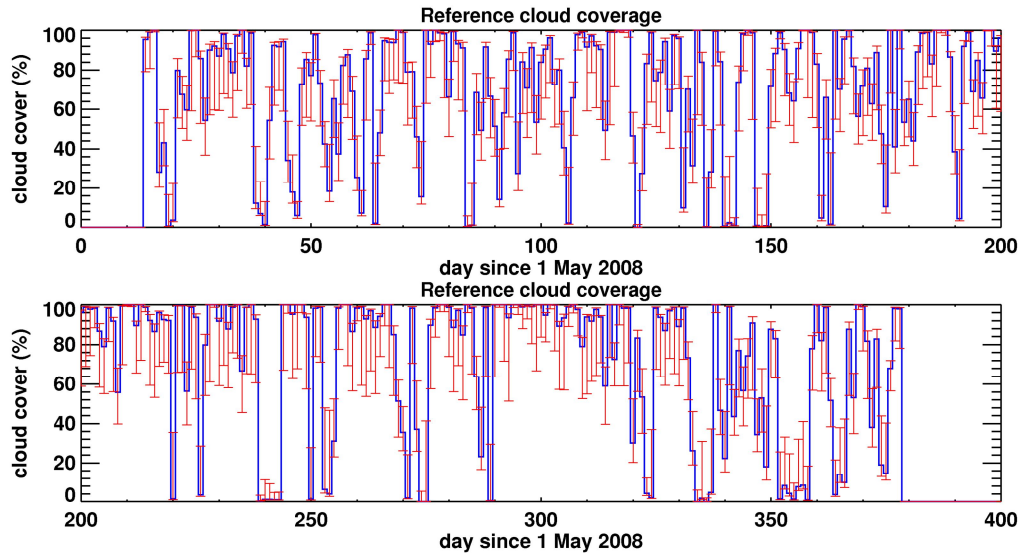

Figure 5: Visualization of the year-long record of daily cloudiness at CESAR based on the reference algorithm including error bars.

Finally Figure 5 shows a time series of the daily median fractional cloudiness values obtained at CESAR for the one year time period 15 May 2008 – 14 May 2009 with the reference algorithm together with error bars. The error bars represents the mean positive and mean negative deviations of the daily time series with respect to the daily median value. Clearly visible are the synoptic scale variations with periods of low cloudiness (around day 350) or high cloudiness (around day 300).

6. CONCLUSIONS

The intercomparison of a 1-year data set of fractional cloudiness from 5 different ground-based remote sensing cloud observing techniques at CESAR and a climatology of human cloud observations from 2 nearby locations showed:

- All instruments reported fractional cloudiness values that deviate 83 – 94% of the time 2 okta or less from one another.
- The column techniques (Cloudnet and LD40) showed enhanced number of cloudiness observations with 0 and 8 okta as compared to the hemispheric techniques (NubiScope, TSI and APCADA). The hemispheric techniques are more capable to detect the presence of only a single cloud on a clear sky or a tiny gap in an otherwise closed cloud deck which is reported as 1 and 7 okta, respectively.
- Cloudnet recorded by far the highest cloudiness values with more than 60% of all values in the 8 okta bin. Our analysis suggests that these high values are due to the sensitivity of this technique to detect even optically thin cirrus clouds at altitudes above 6 km, and the fact that the technique combines two remote sensing instruments, namely a LIDAR and a cloud radar.
- All instruments and the Observer detect more sky clear periods and more completely overcast periods at night than during day.

- Between 1 and 7 okta almost all instruments report more clouds during the day than at night. The Observer sees progressively more clouds during the daytime as cloudiness increases towards higher okta values up to 7 okta.
- Compared to the Observer, all techniques except the TSI yield cloudiness values at the absolute lowest range of values from the 30 year record of Observer cloudiness in the 2 – 6 okta range indicating that the Observer data are systematically different than those from the instruments.
- The altitude of the cloud base can be used as an important discriminator between methods. We constructed a reference algorithm as a linear combination of the output from the five instruments with a weight factor derived from the Cloudnet lowest cloud base value in each 10-minute time period.
- The two hemispheric techniques that constructed the cloudiness values from digital sky imagery (NubiScope and TSI) yield comparable cloudiness values for all cloud base levels. Thus they were taken to be equivalent for all times at which they yielded credible output values. The other three methods were weighted exponentially with cloud base altitude taken from Cloudnet.
- The uncertainty range of the reference cloudiness is determined from the negative and positive differences between the reference and the cloudiness reported by each technique over the time period under consideration.
- Although the constructed reference is not a true reference, it is a general and robust approach that produces useful results. It is an elegant compromise whereby the NubiScope and TSI are considered to be a higher quality product than the others, but whereby the output from the others is not considered to be valueless.

The most important conclusion that can be drawn from this paper is that for fundamental climate data records and their associated essential climate variables, such as fractional cloudiness it is of the highest importance to properly manage the change from one observing technique to the next. At many institutes the Observer has been retired, and replaced by a column method such as a ceilometer. It was shown here that such a column method shifts cloudiness values in the 0 and 8 okta bin for short averaging intervals, so that the characteristic of the cloudiness distribution functions was subjected to a large change. Our results strongly suggest that the candidate replacement for the Observer is not a column method (such as the LD40). We suggest here that it should be a combination of a hemispheric method such as the NubiScope with a column method such as the LD40 (or comparable) mounted on one single pan and tilt unit. It should be noted that it is unrealistic to expect complete similarity between the Observer and any instrumental technique. There are many situations (at night or very high thin clouds) where the Observer is not able to detect clouds, even though they may have been present. Indeed, the results here demonstrate that a universal definition of cloud, which depends on some subjective form of threshold detection such as LIDAR-based optical depth, may be out of reach by the more commonly deployed instruments. Note that the definition of a cloud may depend on the application. However, a physical definition is required in order to be able to achieve a true reference and to determine the performance of each observing technique.

7. ACKNOWLEDGEMENTS

We are indebted to Dr. C.N. Long from Pacific Northwest National Laboratory, USA for providing the software for reprocessing of the TSI data.

8. REFERENCES

Boers, R., M.J. de Haij, W.M.F. Wauben, H. Klein Baltink, L.H. van Ulft, M. Savenije and C.N. Long (2010), Optimized Fractional Cloudiness Determination from Five Ground - based Remote Sensing Techniques, submitted to J. Geophys. Res.

Dürr, B., and R. Philipona (2004), Automatic cloud amount detection by surface longwave downward radiation measurements, J. Geophys. Res., doi:10.1029/2003JD004182.

Hodges, G. (2003) A comparison of fractional cloud cover determinations by a sky imager with visual estimations by trained national weather service observers, 12th Symposium on Meteorological Observations and Instrumentation, 10-13 february, Long Beach, California, USA.

Illingworth, A. J., and others (2007), CloudNet: Continuous evaluation of cloud profiles in seven operational models using ground-based observations, Bull. Amer. Meteor. Soc., 883 – 989.

Long, C.N., J.M. Sabburg, J. Calbo and D. Pages (2006), Retrieving cloud characteristics from ground-based daytime color all-sky images, J. Atmos. Oceanic. Technol., 23, 633-652.

Long, C. N. (2010): Correcting for Circumsolar and Near-Horizon Errors in Sky Cover Retrievals from Sky Images, TOASJ, 4, 45-52, doi: 10.2174/1874282301004010045.

Wauben, W.M.F. (2002), Automation of visual observations at KNMI: (ii) Comparison of automated cloud reports with routine visual observations, AMS Annual Meeting, Orlando, Florida, 2002.

Wauben, W. (2006), Evaluation of the NubiScope, Technical Report No. 291, KNMI, De Bilt.

Wauben, W., H. Klein Baltink, M. de Haij, N. Maat and H. The (2006), Status, Evaluation and new Developments of the Automated Cloud Observations in the Netherlands, IOM 94 (TD 1354) WMO Technical Conference, Geneva, Switzerland.

Wauben, W., F. Bosveld, H. Klein Baltink (2010), NubiScope Laboratory Tests and Field Evaluation, WMO-CIMO Technical Conference, paper 1(8), Helsinki.

WMO (2007), GCOS Reference Upper Air Network: Justification, Requirements, Siting and Instrumentation options, GCOS 112, WMO/TD 1379.

WMO (2008), WMO Guide to Meteorological Instruments and Methods of Observation, WMO-No. 8, Seventh edition, Geneva, Switzerland.

Table 5: Contingency matrix of NubiScope (vertical) versus Reference (horizontal) cloud cover fraction for the period 15 May 2008 – 14 May 2009.

	0	1	2	3	4	5	6	7	8	%	count
0	10.4	2.8	0.2	0.0	0.0	0.0	0.0	0.0	0.0	13.4	5771
1	3.4	8.2	1.7	0.5	0.2	0.0	0.0	0.0	0.0	14.0	6050
2	0.0	0.7	1.8	1.1	0.4	0.1	0.0	0.0	0.0	4.1	1758
3	0.0	0.0	0.6	1.3	0.9	0.4	0.0	0.0	0.0	3.3	1416
4	0.0	0.0	0.1	0.5	1.0	1.0	0.4	0.0	0.0	3.0	1297
5	0.0	0.0	0.0	0.1	0.5	1.0	1.2	0.3	0.0	3.1	1355
6	0.0	0.0	0.0	0.0	0.2	0.6	1.4	1.6	0.0	3.7	1582
7	0.0	0.0	0.0	0.1	0.1	0.6	1.9	14.4	6.5	23.6	10172
8	0.0	0.0	0.0	0.0	0.0	0.1	0.4	3.9	27.3	31.7	13668
%	13.8	11.7	4.4	3.7	3.4	3.8	5.3	20.2	33.8	100.0	
count	5929	5044	1907	1573	1461	1639	2278	8681	14557		43069

Band 0: 66.7% **Band 1: 95.6%** **Band 2: 99.1%** **Over: 0.4%** **Under: 0.5%** **Δ: -0.1** **|Δ|: 0.4**

Table 6: Contingency matrix of TSI (vertical) versus Reference (horizontal) cloud cover fraction for the period 15 May 2008 – 14 May 2009.

	0	1	2	3	4	5	6	7	8	%	count
0	3.8	0.9	0.1	0.1	0.0	0.0	0.0	0.0	0.0	5.0	1035
1	4.4	9.3	1.7	0.5	0.2	0.1	0.1	0.0	0.0	16.4	3417
2	0.0	1.1	1.8	1.2	0.5	0.2	0.1	0.1	0.0	5.0	1040
3	0.0	0.8	1.3	1.8	1.3	0.7	0.3	0.1	0.0	6.2	1297
4	0.0	0.2	0.5	0.9	1.3	1.4	0.7	0.3	0.0	5.5	1142
5	0.0	0.0	0.3	0.3	0.7	1.3	1.4	0.7	0.0	4.6	965
6	0.0	0.0	0.1	0.2	0.4	0.8	1.8	2.0	0.0	5.3	1113
7	0.0	0.0	0.0	0.1	0.2	0.7	1.9	12.5	8.9	24.4	5095
8	0.0	0.0	0.0	0.0	0.0	0.1	0.6	6.5	20.5	27.7	5777
%	8.2	12.4	5.8	5.2	4.6	5.3	6.9	22.2	29.4	100.0	
count	1714	2583	1204	1082	970	1105	1450	4633	6140		20881

Band 0: 54.2% **Band 1: 90.7%** **Band 2: 97.2%** **Over: 1.3%** **Under: 1.6%** **Δ: -0.0** **|Δ|: 0.6**

Table 7: Contingency matrix of APCADA (vertical) versus Reference (horizontal) cloud cover fraction for the period 15 May 2008 – 14 May 2009.

	0	1	2	3	4	5	6	7	8	%	count
0	9.5	2.4	0.3	0.2	0.0	0.0	0.0	0.0	0.0	12.4	6361
1	3.4	5.0	1.2	0.7	0.4	0.2	0.2	0.2	0.1	11.3	5805
2	0.2	4.0	2.1	1.6	1.1	0.7	0.4	0.2	0.1	10.4	5316
3	0.0	0.3	0.7	1.1	1.2	1.0	0.7	0.2	0.0	5.3	2703
4	0.0	0.0	0.1	0.3	0.6	1.0	1.1	0.7	0.0	3.8	1964
5	0.0	0.0	0.0	0.0	0.1	0.1	0.5	1.3	0.0	2.1	1065
6	0.0	0.0	0.0	0.0	0.2	0.7	2.1	8.4	0.0	11.5	5909
7	0.0	0.0	0.0	0.0	0.0	0.0	0.2	7.2	9.9	17.3	8899
8	0.0	0.0	0.0	0.0	0.0	0.0	0.1	1.6	24.2	25.9	13301
%	13.1	11.8	4.3	3.9	3.5	3.8	5.4	19.9	34.2	100.0	
count	6721	6065	2222	2016	1798	1955	2752	10227	17567		51323

Band 0: 52.0% **Band 1: 89.1%** **Band 2: 95.6%** **Over: 0.1%** **Under: 4.3%** **Δ: -0.4** **|Δ|: 0.7**

Table 8: Contingency matrix of Cloudnet (vertical) versus Reference (horizontal) cloud cover fraction for the period 15 May 2008 – 14 May 2009.

	0	1	2	3	4	5	6	7	8	%	count
0	11.4	6.9	0.8	0.3	0.1	0.0	0.0	0.0	0.0	19.5	9080
1	1.2	2.5	1.1	0.3	0.1	0.0	0.0	0.0	0.0	5.1	2372
2	0.0	0.4	0.7	0.5	0.1	0.0	0.0	0.0	0.0	1.7	793
3	0.0	0.3	0.3	0.7	0.3	0.1	0.0	0.0	0.0	1.9	869
4	0.0	0.2	0.2	0.4	0.5	0.2	0.0	0.0	0.0	1.6	730
5	0.0	0.2	0.2	0.3	0.5	0.6	0.2	0.0	0.0	2.0	912
6	0.0	0.1	0.1	0.1	0.3	0.6	0.6	0.2	0.0	2.1	989
7	0.0	0.2	0.2	0.2	0.4	0.5	1.1	1.7	0.1	4.6	2133
8	0.1	0.5	0.6	1.1	1.2	1.7	3.3	18.0	35.2	61.6	28654
%	12.7	11.3	4.2	3.9	3.4	3.8	5.3	20.1	35.3	100.0	
count	5909	5236	1971	1806	1602	1758	2473	9333	16444		46532

Band 0: 53.8% **Band 1: 86.0%** **Band 2: 92.2%** **Over: 7.2%** **Under: 0.6%** **Δ: +0.5** **|Δ|: 0.8**

Table 9: Contingency matrix of LD40 (vertical) versus Reference (horizontal) cloud cover fraction for the period 15 May 2008 – 14 May 2009.

	0	1	2	3	4	5	6	7	8	%	count
0	10.5	6.1	1.2	0.8	0.4	0.3	0.4	0.2	0.1	20.1	10309
1	2.1	3.8	1.1	0.7	0.4	0.3	0.4	0.3	0.1	9.3	4769
2	0.3	0.8	0.8	0.5	0.3	0.2	0.2	0.3	0.0	3.3	1700
3	0.1	0.4	0.6	0.6	0.3	0.2	0.2	0.3	0.0	2.8	1445
4	0.1	0.2	0.3	0.8	0.6	0.4	0.4	0.5	0.0	3.3	1706
5	0.0	0.1	0.1	0.3	0.7	0.6	0.5	0.7	0.0	3.1	1574
6	0.0	0.1	0.1	0.1	0.4	0.7	0.7	1.2	0.0	3.3	1712
7	0.0	0.0	0.1	0.1	0.3	0.8	1.8	6.0	1.8	10.9	5575
8	0.0	0.0	0.0	0.1	0.1	0.3	0.9	10.6	32.0	44.0	22592
%	13.2	11.6	4.3	3.9	3.5	3.8	5.4	20.2	34.0	100.0	
count	6777	5981	2201	2024	1812	1972	2790	10377	17448		51382

Band 0: 55.6% **Band 1: 85.6%** **Band 2: 92.4%** **Over: 2.0%** **Under: 5.5%** **Δ: -0.1** **|Δ|: 0.7**