Analysis Algorithm for Sky Type and Ice Halo Recognition in All-Sky Images

Sylke Boyd, Stephen Sorenson, Shelby Richard, Michelle King, Morton Greenslit
Division for Science and Mathematics, University of Minnesota-Morris, 500 E 4th Street, Morris, MN

5   *Correspondence to*: Sylke Boyd (sboyd@morris.umn.edu)

Abstract. Ice halos, in particular the 22° halo, have been captured in long-time series of images obtained from Total Sky Imagers (TSI) at various Atmospheric Radiation Measurement (ARM) sites. Ice halos form if smooth-faced hexagonal ice crystals are present in the optical path. We describe an image analysis algorithm for long-time series of TSI images which identifies images with 22° halos. Each image is assigned an ice halo score (IHS) for 22° halos, as well as a sky type score

10   (STS), which differentiates cirrostratus (CS), partially cloudy (PCL), cloudy (CLD), or clear (CLR) within a near-solar analysis area. The colour-resolved radial brightness behaviour of the near-solar region is used to define the characteristic property spaces used for STS and IHS. The scoring is based on distance from a region in that property space, using tools of multivariate Gaussian analysis. An external expandable master table of characteristic properties allows continued training of the algorithm. Scores are assigned to the standardized sun-centred image produced from the raw TSI image after a series of calibrations,

15   rotation, and coordinate transformation. We present test results on halo observations and sky type for the first four months of the year 2018, for TSI images obtained at the Southern Great Plains (SGP) ARM site. A detailed comparison of visual and algorithm scores for the month of March 2018 shows that the algorithm is about 90% reliable in discriminating the four sky types, and identifies 86% of all visual halos correctly. Numerous instances of halo appearances were identified for the period January through April 2018, with persistence times between 5 and 220 minutes. Varying by month, we found that between 9%

20   and 22% of cirrostratus skies exhibited a full or partial 22° halo.

Introduction

Modelling and predicting the Earth's climate is a challenge for physical science, even more so in light of the already observable changes in Earth's climate system (Fasullo and Balmaseda, 2014; Fasullo et al., 2016; IPCC, 2013, 2014). Global circulation

25   models (GCMs) describe the atmosphere in terms of a radiative dynamic equilibrium. The Earth receives solar shortwave (SW) radiation and discards energy back into space in form of terrestrial long-wave (LW) radiation. The radiation balance of the earth has been subject to much study and discussion (Fasullo and Balmaseda, 2014; Fasullo and Kiehl, 2009; Kandel and Viollier, 2010; Trenberth et al., 2015). Global Circulation Models (GCMs) describe the influence of various parts of the earth system in terms of radiative forcing factors (Kandel and Viollier, 2010; Kollias et al., 2007). Clouds may restrict the SW flux

Atmospheric
Measurement
Techniques
Discussions

reaching the surface, but they also influence the LW emissions back into space. While low stratus and cumulus clouds exhibit a net negative radiative forcing, high cirroform clouds are more varied in their radiative response, varying between negative and positive forcing depending on time of day, season, and geographical location (Campbell et al., 2016). The Fifth Assessment Report from the IPCC in 2013 (IPCC, 2013) identified ice and mixed clouds as major contributors to the low confidence level

5 into the aerosol/cloud radiative forcing. The uncertainty in the aerosol/cloud forcing has implications for the confidence in and for the variance of the predictions of global circulation models (Fu et al., 2002; Trenberth et al., 2015). Closing the radiation budget of the Earth hinges on reliable cloud data (Hammer et al., 2017; Knobelspiesse et al., 2015; Schwartz et al., 2014). Traditionally, cloud radiative forcing is modelled using a cloud fraction based on sky images (Kennedy et al., 2016; Kollias et al., 2007; Schwartz et al., 2014). Cirrostratus clouds, lacking sharp outlines, pose a challenge to this approach (Schwartz et

10 al., 2014). The uncertainty about the role of cirrus in the global energy balance has been attributed to limited observational data concerning their composition, and temporal and spatial distribution (Knobelspiesse et al., 2015). Cirroform clouds, at altitudes between 5000-12,000 m, are effective LW absorbers. Cloud particle sizes can range from 0.1 microns to a few millimetres (Cziczo and Froyd, 2014). Methods to probe cirrus cloud particles directly involve aircraft sampling (Delene, 2011) and mountainside observations (Hammer et al., 2015). Ground- and satellite-based indirect radar and LIDAR

15 measurements (Hammer et al., 2015; Tian et al., 2010) give reliable data on altitudes, optical depths, and particle phase. All of these methods are restricted to a particular time, or the path of an aircraft, or in terms of the nature of information that can be gleaned from the data. It is clear, that no single method has all the composition information.

Optical scattering behaviour is influenced by the types of ice particles, which may be present in very many forms, including crystalline hexagonal habits in form of plates, pencils and prisms, hollow columns, bullets and bullet rosettes, and amorphous

20 ice pellets, fragments, rimed crystals and others (Delene, 2011; Ewald et al., 2013; Leisner et al., 2013; Tian et al., 2010; van Diedenhoven, 2014; Weihs et al., 2015). Only ice particles with a simple crystal habit can lead to observable symmetric scattering patterns. Usually, this will be the hexagonal prism habit, which we can find in plates, columns, bullet rosettes, pencil crystals, etc. If no preferred orientation exists, a clear tell-tale sign for their presence is the 22° halo around a light source in the sky, usually sun or moon. More symmetry in the particle orientations will add additional ice halo features. If many of the

25 ice particles are amorphous in nature, or did not form under conditions of crystal growth- for example by freezing from super-cooled droplets, or by riming – the scattering pattern will be similar to what we see for liquid droplets: a white scattering disk surrounding the sun, but no halo. If the particles in the cirroform cloud are very small, e.g. below 1 μm, diffraction will lead to a corona. Hence, we believe that a systematic observation of the optical scattering properties can add information to our data on cirrus composition. The authors observed the sky at the University of Minnesota Morris, using an all sky camera,

30 through a five-month period in 2015, and found an abundance of halo features. The data and methods of image analysis were presented at the Gordon Research Conference on Radiation and Climate in 2015 and confirmed that ice halos do form often and throughout the year (Sassen et al., 2003; Seefeldner et al., 2017; van Diedenhoven, 2014). There are a few studies pursuing a similar line of inquiry (Forster et al., 2017; Sassen et al., 2003).

The study by Sassen et al (Sassen et al., 2003) showed a prevalence of the 22° halo, full in 6% and partial in 37.3% of cirrus periods, based on a ten-year photographic and LIDAR record of mid-latitude cirrus clouds, also providing data on parhelia, upper tangent arcs, and other ice halo features, as well as coronas. The photographic record was based on 20-minute observation intervals; cirrus identification was supported by LIDAR. The authors found an interesting geographical variability

5   in halo displays, related to air mass origin, and suggest that optical displays may serve as tracers of the cloud microphysics involved. Forster *et al.* (Forster et al., 2017) used a sun-tracking camera system to observe ice halo details over the course of several months in Munich, Germany, and a multi-week campaign in the Netherlands in November 2014. A carefully calibrated camera system provided high-resolution images, for which a halo detection algorithm was presented, based on a decision tree and random forest classifiers. Ceilometer data and cloud temperature measurements from radio probes were used to identify

10   cirrus clouds. The authors report 25% of all cirrus clouds also produced ice halos, in particular in the sky segments located above the sun. The fraction of smooth crystals necessary for ice halo appearance is 10% for columns, and 40% for plates (van Diedenhoven, 2014). The consideration of the percentage of cirrus clouds that display optical halo features allows a direct conclusion with respect to the fraction of crystalline habit in the cloud, and, upon further study, about the microphysical conditions in the cloud. This raises interest in examining existing long-term records of sky images.

15   Long-term records of sky images have been accumulated in multiple global sites. The Office of Science in the US Department of Energy has maintained Atmospheric Radiation Measurement (ARM) sites. These sites, among other instruments, contain a Total Sky Imager (TSI), and have produced multi-year records of sky images. In this paper, we introduce a computational method to analyse these long-term records for the presence of ice halos in the images. We are introducing an algorithm that will read, standardize, and analyse long sequences of TSI data, and produce a time record of near-solar sky type, differentiated

20   as cirrostratus (CS), partly cloudy (PCL), cloudy (CLD), and clear (CLR) sky types, as well as assign an ice halo score (IHS). The resolution and distortion of the TSI images restricts the halo search to the common 22° halo. Other halo features, such as parhelia, can occasionally be seen in a TSI image, but often are too weak or too small to reliably discriminate them from clouds and other features. Coronas are obscured by the shadow strip, and often also by over-exposure in the near-solar area of the image. The algorithm offers an efficient method of finding ice halo incidences, full or partial. Since ARM sites also have

25   collected records of LIDAR and photometric data, the halo algorithm is intended to be combined and correlated to other instrumental records from the same locations and times. This will be addressed in future work.

Section 1 describes the TSI data used in this work. Section 2 presents the details of the image analysis algorithm, including a refinement of the algorithm goals, the image preparation, and sky type and halo scoring. Section 3 applies the algorithm to the TSI data record of the first four months of 2018, and examines effectiveness and types of data available for this interval.

30   Summary and outlook are given in Section 4.

Atmospheric
Measurement
Techniques
Discussions

1 TSI images

Images used in this paper where obtained from Atmospheric Research Measurement (ARM) Climate Research Facilities in three different locations: Eastern North Atlantic (ENA) Graciosa Island, Azores, Portugal; North Slope Alaska (NSA) Central Facility, Barrow AK; and Southern Great Plains (SGP) Central Facility, Lamont, OK (ENA, 2018; OLI, 2018; SGP, 2018). The

5   ranges and dates vary by location, as listed in Table 1. The images were taken with Total Sky Imagers, which consist of a camera directed downward toward a convex mirror to view the whole sky to the horizon. A sun-tracking shadow strip obscures a strip of sky from zenith to horizon, containing the sun itself. Images were recorded every 30 seconds. The longest series was taken at the Southern Great Plains (SGP) location, reaching back to July 2000. The images, in jpeg format, have been taken continuously during day time. Aside from night time and polar night, there are some additional gaps in the data, perhaps due to instrument failure or other causes.

10   Camera quality, exposure, image resolution, and image orientation varies over time as well as by location. The angular resolution varies with zenith angle but can rise above 0.7 ° for the smaller images (0.4° for the larger size), in particular for sky sections close to the horizon. The images are sorted into a folder tree by day and month, labelled with markers for location, date, and time in accordance with the naming conventions at the ARM sites. We used eighty seed images taken from across the TSI record and across all available years to start the master table (ENA, 2018; NSA, 2018; SGP, 2018). This included images visually identified

15   as CS, PCL, CLD, CLR, and halo-bearing. The seed samples were used to develop the algorithm and define a suitable set of characteristic properties for STS and IHS.

2 Algorithm

2.1 Goal and Strategy

The algorithm aims to process very large numbers of images, and return information about the presence of ice halos, as well

20   as the general sky conditions. The program is written in C++ and uses the opencv library for image processing. If given a list of image directories, the algorithm proceeds to sequentially import, process, and score TSI images resulting in a sky type score (STS) and an ice halo score (IHS). In order to discriminate the sky types listed in Table 2, for example, or to single out the relatively weak halo signature from an image we use a multivariate Gaussian analysis. This begins with the definition of a set of N_P properties of the image, selected to be characteristic for a sky type or a halo. Let this set of properties be a vector

25
$$X = \{x_i\}_i^{N_p}$$

A master table is created from N_{master} images that visually exhibit the target feature, i.e. a halo or a clear sky. This set defines an ellipsoidal region in the property space of X. The region is centred at the vector of mean values

$$M = \{\mu_i\}_{i=1}^{N_P} \tag{1}$$

where the vector elements are the mean values of the master set

30
$$\mu_i = \frac{1}{N_{master}} \sum_{k=1}^{N_{master}} x_{ik}$$

Atmospheric
Measurement
Techniques
Discussions

Open Access

EGU

The stochastic ellipsoid is described by the $N_P \times N_P$ covariance matrix

$$\Sigma = \overline{(X-M)(X-M)^T} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \dots \\ \sigma_{21} & \sigma_{22} & \dots \\ \dots & \dots & \dots \end{pmatrix}, \tag{2}$$

evaluated for the sets in the master table. The elements of the covariance matrix are computed as

$$\sigma_{ij} = \frac{1}{N_{master}} \sum_{k=1}^{N_{master}} x_{ik} x_{jk} - \mu_i \mu_j$$

5   The property vector of any further image X_{image} will then be referenced with M and Σ in form of a multivariate normal distribution

$$F_{image} = C_0 exp\left(-\frac{1}{2}(X_{image} - M)^T \Sigma^{-1}(X_{image} - M)\right) \tag{3}$$

in which the exponent is known as the square of the Mahalanobis distance in property space. The closer an image places to the region of interest, the higher its score will be. In order to score a time series of property vectors X_{image}, one only needs to import

10   M and Σ^{-1} once at the start of the analysis run. Both, M and Σ^{-1}, are computed a priori in a master table via Equations (1) and (2). We are using a flexible spreadsheet for this purpose, allowing the addition of reference property vectors as more images are analysed. This allows to continually train the algorithm toward improvement of scoring. The pre-factor C_0 in Eqn.(3) is chosen to be 10^6, placing the typical value for F into an order of magnitude larger than 1 for a halo image. This basic algorithm structure is used on a standardised local sky map, described in 2.2. The details of STS and IHS will be treated separately below

15   as well. The code and accessories can be accessed at a GitHub repository (Boyd et al., 2018).

2.2 Image preparations and local sky map (LSM)

The goal of the image preparation is to create a local sky map centred at the sun, in easy-to-use coordinates, after a minimal colour calibration, and after extraneous image parts have been masked. Some sample steps in the image preparation are

20   illustrated in Figure 1. The images were taken at the Southern Great Plains ARM site in March and April of 2018 (SGP, 2018). One of the images contains a solar 22° halo, the other is a partly cloudy sky without any halo indications. Both original images have a slightly green tinge, which is typical for images from the TSI at this location, in particular after an instrument update in 2010. This is noticeable in particular if images are compared to earlier TSI data from the same location, and can become a problem for the planned analysis, in particular for the use of relative colour values.

25   The very first step is therefore a colour correction. Colour drift is slow, and occurs on a time scale of years, rather than days. The calibration is based on sampling of clear-sky colour channels (BGR values) to define weighed scaling factors for a whole series of images. In a discoloured series, measurements of BGR were taken in clear-sky images (indexed CLR), and a scaling factor and weight for each colour channel defined based on this information:

Atmospheric
Measurement
Techniques
Discussions

$$\left.\begin{array}{l} \beta_B = 1.00 \\ \beta_G = \dfrac{G_{ref}}{G_{CLR}} \times \dfrac{B_{CLR}}{B_{ref}} \\ \beta_R = \dfrac{R_{ref}}{R_{CLR}} \times \dfrac{B_{CLR}}{B_{ref}} \end{array}\right\} \text{ with } \left(B_{ref}, G_{ref}, R_{ref}\right) = (180, 120, 85)$$

The reference values are based on colour values for average clear sky images from across all TSI records available to the authors. Once these colour-scaling factors are determined for a series, every image was then tinted by generating an average colour $\bar{B}$ for a small near-zenith sky-sample and applying

$$B = \left[B + \alpha\left(\beta_B \bar{B} - B\right)\right]$$

to each colour channel and pixel, respectively, followed by a simple scaling to preserve the total brightness of the pixel. For the series SGP 2018, these factors were $\beta = (0.9, 0.78, 1)$ and $\alpha=0.4$ (regulates the strength of the tinting; $\alpha=0$ no tint, $\alpha=1$ monochrome image). This tinting is minimal, and linear colour behaviour is a reasonable assumption. The second step identifies the horizon circle, stretches the visible horizon ellipse to circular shape, if necessary, and centres the circle as close to the zenith as possible. A north-south alignment correction may also have to be applied. These calibrations become necessary if the TSI was not perfectly aligned in the field. In addition, the horizon is chosen at a zenith angle smaller than 90°, often between 85° and 79°, to eliminate the strong view distortion close to the horizon, and in some cases objects present in the view. The current process requires to find these calibrations for a handful of images in a series, and then apply them to all images in the series. The projection of the sky onto the plain of an image introduces a perspective distortion, as described in Long et al (Long et al., 2006). Once the horizon circle is centred and aligned, a coordinate transformation is performed to represent the sky in terms of azimuth and zenith angles. The azimuth is the same in both projections. Zenith angle θ relates to the radial distance from the centre of the horizon circle as $R \sin \theta$. The coordinate transformation represents the sky circle in a way in which radial distance from zenith s_z scales with zenith angular coordinate θ. Image horizon radius and maximum zenith angle provide one known point to allow for proportional scaling. We tested the influence of the spherical mirror reflection on the distortion. For camera positions at height h above a convex mirror of radius A where $h/A \approx 1$, the assumption $\theta \sim s_z$ is reasonable. One of the visible effects of this transformation concerns 22° halos: in the original TSI image, a halo appears as a horizontal ellipse; after the transformation it will have a shape closer to a circle. The position of the sun is marked based on the geographical position and the Universal Time (UTC) of the image. Extraneous details are being masked. The centre panel of Figure 1 shows the image produced by all these adjustments. Since often the position of the sun is detectable in the image, the marked sun position serves to refine the calibrations described above. The local sky map (LSM) is created by rotating and cropping the image to approximately within 40 sky degrees of the sun, with the sun at its centre. Units of measurement in the LSM are closely related to angular degrees, but do not match perfectly due to a zenith-angle dependence of the azimuth arc length. At θ=45°, the arc length of azimuth angle φ is equivalent to the arc length of a θ of same size; however, if θ>45° the azimuth arc is stretched, requiring an additional horizontal compression. The algorithm is robust enough

Atmospheric
Measurement
Techniques
Discussions

to allow this scaling by solar position alone, without loss of efficacy. We will measure all distances across the LSM in LSM units. A sketch of the layout of the LSM is provided in Figure 2.

2.3 Computing Sky Type and Halo Properties

2.3.1 Average radial intensity (ARI)

5 Halos, as sun-centred circles, are creating a brightness signal at a radius of 22°. We found it useful to analyse the radial intensity behaviour $I(s)$ with s being the radial distance from the sun in the image plane. There is a physical reason for using $I(s)$ in cloud assessment. The presence of scattering centres in the atmosphere, near to the line-of-sight to the sun, influences the properties of sky brightness in the near-sun sky section. A very clear atmosphere, for example, exhibits a very fast initial radial decline of brightness, followed by a relatively low gradient at low brightness, but with relatively high intensity values in the

10 blue channel. The presence of cirrostratus will soften the initial radial decline and lead to a longer radial range of brightness, with more evenly distributed intensities in the red, green, and blue channels. In a partially cloudy sky, we would find sharp variations in $I(s)$, varying with colour channel. A cloudy sky, on the other hand, may exhibit no decline in radial brightness, and will generally have low intensity values across all colour channels. A sketch of the LSM is given in Figure 2. The radial intensity $I(s)$ is computed using the colour intensity values of the image (0 to 255), separated by colour channel. The LSM is

15 divided into four quadrants: TR = top right, BR = bottom right, BL = bottom left, TL = top left, analysed separately, and then recombined for the image scores. The division into quarters allows to accommodate partial halos, low solar positions, and the influence of low clouds in partially obstructing the view to cirrostratus. The algorithm uses various properties of $I(s)$ to assign STS and IHS, as detailed below. The average radial intensity $I(s)$ is computed as an average over pixels at constant radial distance s from the sun. Due to the low resolution of the images, and due to some noise in the data, we average $I(s)$ over a

20 circular ribbon with a width of 4 pixels, centred at s. Figure 3 shows the radial intensity of the red channel (R) in the bottom right quadrants of the LSMs featured in Figure 1. Panel A includes $I(s)$, a linear fit, as well as the running average $\overline{I_6}$, plotted versus radial distance s. The running average is taken as the average of $I(s)$ over a width of 6 LSM units centred at s:

$$\overline{I_6}(s) = \frac{1}{N_a} \sum_{a=s-3LSMunits}^{a=s+3LSMunits} I(a)$$

The clear-sky image exhibits a lower red intensity overall than the halo image. The halo presents as a brightness fluctuation at

25 about 21 LSM units. The resolution of the TSI images only allows to resolve 0.4° to 0.7° with certainty, and variations in calibration and zenith angle do influence deviations from the expected 22° position. The analysis of $I(s)$ is undertaken in an interval between 15 and 26 LSM units, called the radial analysis interval (RAI). The RAI is marked in Figure 2. A linear fit yields a slope and intercept value used for the STS. We define the radial intensity deviation as

$$\eta(s) = I(s) - \overline{I_6}(s) \qquad\qquad (4)$$

30 Panels B in Figure 3 show $\eta(s)$ for both situations. The details of the halo signal in $\eta(s)$ contribute in particular to the computation of the IHS.

2.3.2   Sky type score (STS)

Properties of *I(s)* were computed for the set of seed images mentioned in section 1. Twenty images for each sky type were divided further by sky quadrants, yielding between 60 and 80 property sets to seed the master table. Some quadrants were eliminated by horizon-near solar positions. These image quadrants were used to apprise the utility of properties in making sky

5   type assignments, with focus on the radial analysis interval (RAI) between 15 and 26 LSM units. The ten image properties used to compute the STS are listed in Table 3, together with their average values for the master table. They include the slope and intercept of the line fit to *I(s)* for each colour channel, where the slope characterizes a general brightness gradient, and the intercept gives access the overall brightness in the RAI. The line fit alone will not allow to differentiate partially cloudy skies from other sky types. However, the presence of sharply outlined clouds leads to a larger variation in intensity values, even for

10   the same radial distance from the sun. The areal standard deviation (ASD) is an average of the standard deviation of *I(s)* for each radial distance s, averaged over all radii separated by colour channel. To set apart CLR skies, the average colour ratio (ACR) in the analysis area is computed as

$$ACR = \frac{B^2}{GR}$$

In Figure 4, the STS property set is represented graphically, including means, standard deviations, and extreme values as

15   observed in the master table. Clearly, no single property alone will suffice to assign sky type reliably. There is overlap in the extreme ranges. Relations between the colour channels are influential, as well. We are using the mechanism described in section 2.1 and the continually refined master table to define a mean value vector *M* and inverse covariance matrix Σ^{-1} for each sky type. As an image is processed, its STS properties are computed for each sky quadrant. Subsequently, a score is computed for each sky type using equation (3). The raw values of F (equation 3) vary greatly even between similar looking images, hence

20   STS is computed as a relative contribution between 0 and 100% for each sky type and each quadrant. For the CS score this would mean:

$$STS(CS) = \frac{F_{CS}}{F_{CS} + F_{PCL} + F_{CLD} + F_{CLR}} \times 100\%$$

This means, a single image quadrant can carry scores of 45% for CS, 35% for PCL, and 20% for CLD. The dominant sky type

25   then is CS for this quadrant. The STS for the image is assigned as the average over all quadrants. If the raw scores F for all sky types were smaller than 10^{-8} the images is classified as N/A. It simply means that its properties are not close to any of the sky type categories. A one-day sample of sky type data is shown in Figure 5, for 10 March 2018. The day was chosen for its variability, including periods of each of the sky types, as well as clearly visible halo periods. The central panel tracks STS for all sky types through the day, taken for the combined sky. It is important to note that the sky type only can be representative

30   of the section of sky near to the sun. The white areas of 25 or 50 % are introduced when the solar position nears the horizon, eliminating the two bottom quadrants of the LSM from analysis. Some of the late images contain quadrants that were eliminated due to overexposure. The white scattering disk around the sun near the horizon does not allow for analysis,

exemplified in the sample image at 22:53:00 included in Figure 5. For large portions of the day, the dominant sky types have been classified as CS and PCL, and the images corroborate this. The 14:36:00 image shows a thicker cloud cover, and the algorithm correctly responds by increasing the CLD score. At 21:00:00, the algorithm indicates an increased CLR score, consistent with the visual inspection of the TSI image at the time. Given the simplicity of this sky type assessment, we believe

5 that this radial scattering analysis around the sun has the potential to address some of the challenges that have been encountered using a simple cloud fraction in radiation modelling (Calbó and Sabburg, 2008; Ghonima et al., 2012; Kollias et al., 2007). That will be a direction to discuss and explore in the future.

2.3.3 Ice halo score (IHS)

The 22° halo is a signal in the image that overlays many other image features, including low clouds, partial clearings,

10 inhomogeneous cirrostratus, regions of over-exposure, and near-horizon distortions. While most halos are formed in CS skies, halo signals also can be found in thinner CLD skies, and even in skies that classified as CLR (even though a very thin pristine ice crystal layer must have been present). The challenge to extract the halo from such a wide variety of sky conditions is formidable. While the statistical approach described in 2.1 will again form the core of the approach, the challenge shifts to defining a set of suitable properties of the image. In addition to the properties used in sky type assignment, the halo scoring

15 must seek features in $\eta(s)$ (Eqn.4) that are unique in halo images, such as a minimum followed by a maximum at halo distance from the sun. The absolute values of $\eta(s)$ are dependent on various image conditions. Due to the variety of sky conditions, and variations in calibration and image quality, the values of maximum and minimum alone are not sufficient to reliably conclude the presence of a halo. We have found instances in which $\eta(s)$ does exhibit the halo maximum, but does not dip to negative values first. However, the upslope-crest-downslope sequence is consistently present in all cases of ice halo. The halo search

20 should be undertaken for a sequence of upslope – crest – downslope in terms of radial positions and range of slopes. All three characteristics present clearly in the derivative of the $\eta(s)$, the radial intensity deviation derivative $\eta'(s)$. This derivative of the discrete series $\eta(s)$ is approximated numerically by a secant methods as

$$\eta'_i \approx \frac{\eta_{i+1} - \eta_{i-1}}{s_{i+1} - s_{i-1}}$$

In Figure 6, both $\eta(s)$ and $\eta'(s)$ are shown for the bottom-right quadrant of the green channel of the halo image in Figure 1.

25 The sequence of radial halo markers is illustrated in Figure 6. The algorithm computes $\eta'(s)$ and seeks the positive maximum and the subsequent negative minimum, plus the radial position of the sign-change between them. This produces a sequence of radial locations s_{up}, s_{max}, and s_{down} which basically outline the halo bump in width and location. There are often multiple maxima of $\eta'(s)$ contained in the RAI. A halo image typically has fewer maxima than a non-halo image, but of larger amplitude. Therefore, the number of maxima as well as the upslope value η'_{up} and down-slope derivative η'_{down} join the set of halo

30 indicators. If multiple maxima are found, the dominant range is used. Lastly, if a radial sequence is found in one colour channel, it should be found in the same locations in all colour channels, which is why a standard deviation of all three radial positions

across the three colour channels was added to the halo scoring set of properties. We arrive at a set of 31 properties for the computation of the IHS, listed in Table 4. The means and standard deviations have been computed for the master table. The mean value vector M and the inverse covariance matrix Σ^{-1} are computed in the master table and then imported by the halo searching algorithm for use in equation (3). The raw halo score F is computed for the four quadrants of an individual image,

5   their sum is used to assign the raw score for the whole image. While the individual scoring works very well for halo images, it does trigger the occasional halo score for images that do not exhibit a halo. This may occur due to inhomogeneities in a broken cloud cover, or other isolated circumstances. These false halo scores often occur on isolated images. We utilize the factor of residence time of a halo to address this. In a 30-s binned series of TSI images, the halo will appear usually in a sequence of subsequent images, often in the order of minutes or even hours. We added a Gaussian broadening to the time

10   series of halo scores F_i, taken at times t_i:

$$IHS(t) = \sum_{t_i=t-3w}^{t_i=t+3w} F(t_i) \exp\left[-\frac{(t_i-t)^2}{2w^2}\right] \qquad (5)$$

This de-emphasizes isolated instances, and enforces sequences of halo scores, even if they exhibit weak signals or gaps. This procedure reduced the false halo identifications significantly. In Figure 5, the time series of the IHS is shown in the bottom panel. Just as for the sky type, the Master table for sky type scores and halo scores is being complemented as more images are

15   analysed. The broadening w in Eqn (5) was chosen as 4 images for this example, which means the Gaussian half width corresponds to 2 minutes. The clear 22° halo between 19:00 and 20:00 UTC produces a strong IHS. There are a few weaker signals, and upon inspection of the images we find that these correspond to partial halos (like at 17:07:00), or halos in a more variable sky.

3   Results for January through April 2018

20   We chose the record of the month of March of 2018 at the SGP location for a thorough comparison of algorithm results to visual image inspection. The complete month TSI record, starting at 3/1/2018 0:00:00 UTC and ending at 3/31/2018 23:59:30 UTC, contains 44,026 images. An image IHS and STS are assigned as the average over all scoring quadrants. Visual image classification for so many images poses a considerable challenge, which we approached in form of an iteration. The classifications produced by an initial run of the algorithm were inspected visually as isolated images as well as animations.

25   The visual assessments then were refined, while the master table was complemented by property sets from images that were not classified correctly. The process was repeated several times until no more gains in accuracy were observed.

The resulting time lines for STS and IHS for the month of March are plotted in Figure 7. Many of the images exhibit strong indicators for multiple sky types. The largest STS is used to assign a sky type to an image. The IHS was computed using a pre-factor C of 10^6 (Eqn.3) and a half width broadening in time of w=3.5 minutes (Eqn.2). It is interesting to observe that the high

30   halo scores coincide with strong CS signals. Noteworthy is also, that there are a number of days in which CS does not carry an ice halo. Upon inspection of the numerical values for IHS, it becomes clear that `a cut-off is needed to assign an image with

a label of halo/no halo. This cut-off value is arbitrary and dependent on factors such as w and C_0, as well as the quality of the calibration. Our testing places it at around 3500 for the month of March.

In Table 5, visual and algorithm results of the sky type assignments are cross-listed. Cloudy skies are reliably identified by the algorithm. A small percentage of visual CLD skies trigger a PCL signal, mostly due to inhomogeneities in cloud cover. The

5 algorithm is very successful in pinning down the CLR skies as well. CS and PCL are very successful, but exhibit some difficulties. For example, 8.5% if visual CS skies scored a CLR signal, and 10% of images classified as CS were visually assigned a PCL sky type. In these cases we often found that the algorithm assignment might be more persuasive than the visual assignment – a visual assignment is a subjective call, and open to interpretation of the observer. Combined with image distortion and resolution limits, it is quite possible that the visual assignments carry a considerable uncertainty. Some of the

10 visual CS skies, for example, present to the eye as CLR, but reveal the movement of a cirrostratus layer if viewed in context of time-development (animation). Similarly, cirrostratus may present as an inhomogeneous layer in transition skies, triggering a partly-cloudy assessment in the algorithm. Low solar positions are prone to larger image distortion, which may lead to misinterpretation. It is also worth noting that every image receives an STS for all sky types, and that the total image assessment merely selects for the STS with the highest contribution. In cases of mismatch, we often find that the two sky types at conflict

15 both contribute significantly to the STS of the image. If the solar zenith angle is above 68°, no sky type assignments were made. Most of the 397 CLR images that presented as CS to the algorithm were taken at very low sun, with a significant over-exposure disk in near-solar position. Table 5 also lists a comparison of visual halo identifications with the algorithm scores. According to this assessment, the algorithm correctly calls 85 % of visual halo images, while not diagnosing 15 % of them. On the other hand, 12 % of all halo signals do not correspond to a halo in the image. One can improve the correct identification

20 rate by lowering the cut-off score, on the cost of an increase the signal from false identifications. Balancing the false positive and false negatives yields a reliability of about 12 to 14 %. Some of the false negatives arise from altocumulus skies, in which the outlines of cloudlets may trigger halo signals by their distribution and size. These are very difficult to discriminate from visual halo images. Some caution is advised in relying heavily on visual classifications of TSI images alone. The visual sky type and halo assignments themselves have an uncertainty due to subjectivity. While it is easy to distinguish a partially cloudy

25 sky from a clear sky, this may become difficult for the difference between thick cirrostratus and stratus. Their visual appearances may be quite similar. Sometimes, an assignment can be made in context of temporal changes. Some clear-appearing skies reveal a thin cirrostratus presence if viewed in a time series instead of in an individual image. We also have found visual halos in images which this algorithm flagged, and the presence of a weak halo revealed itself only after secondary and tertiary inspection of the image. It is therefore a future necessity to combine the visual assignments of sky types with

30 LIDAR data for altitude, optical thickness, and depolarization measurements to make an accurate assessment of the efficacy of the halo detection.

We applied the algorithm to the TSI record for the first four months of 2018 for the SGP ARM site. It is worth noting that this is not intended to be a complete exploration of the ARM record concerning ice halos, but rather a demonstration of capacity of the algorithm presented here. Table 6 summarizes our findings. It lists the percentages for the four sky types by month. A

portion of the images has not been assigned with an STS. The conditions under which this occurs have been alluded to earlier, and include horizon-near solar positions, images with over exposure in the RAI, and images for which the raw STS for each sky type was numerically too low to be considered a reliable assessment. Therefore, sky type percentages refer only to all identified images. January and March exhibited a large fraction of clear skies. February was dominated by cloudy skies, while

5   April registered a high percentage of CS - however, only a partial month of images was available for April. Sky type depends strongly on the synoptic situation. That means that no further conclusions should be made from these data. The ice halo statistics in Table 6 gleaned the data for various dimensions of the record. The number of separate halo incidences counts sequences of images for which the IHS did not fall below the cut-off value of 3500. While it is worth noting that the number of incidences lies in the order of magnitude of the number of days in a month, it is certain that the halo instance are not evenly

10   distributed. Figure 7 does demonstrate this behaviour. However, even on a day of persistent cirrostratus with ice halo, interruptions of its visibility can occur. Sometimes low stratocumulus may obscure the view of the halo, sometimes the cirrus layer is not homogeneous. This may lead to a large number of separate halo incidences in a short time, while none are counted at other times. The mean duration of a halo incident lies between 16 and 34 minutes, depending on month. We listed the maximum duration found in each month as well. The longest halo display in the time period occurred in April, with nearly 3.5

15   hours. Mean values are easily skewed by a few long display. Figure 8 shows the distribution of ice halo durations for the four months. The most common duration of a 22° halo lies between 5 and 10 minutes, followed by 10 to 15 minutes. Due to the time-broadening applied via Eqn (5), the display time cannot be resolved below 3 minutes. We consider the fraction of images in which a halo was registering. That marker varied between 3.9% for January and 9.4% for April. In accord with findings in (Sassen et al., 2003), we find a low amount of ice halo activity in January. However, this may be influenced by the large zenith

20   angles for the sun in January. The closer the sun to the horizon, more TSI images have been excluded from the analysis, and the stronger the influence of distortion – both effects that would influence both studies.

Occasionally, only partial halos will be seen, depending on positioning of the cirroform clouds and on obstruction by low clouds. The division of the LSM into quadrants allows to assess the possibility of fractional halos, as indicated in Table 6. The overwhelming portion of halo incidences shows full or 75% halo. This means that in four or three of the quadrants, the IHS

25   has exceeded its minimum cutoff. Quarter halos have only rarely registered in the algorithm. Many of the half halos can be found for images taken close to sunrise or sunset. That explains their relative frequency in January and February.

We started the project with the goal to find information on cirrostratus composition. The bottom part of Table 6 investigates the relation between sky type and ice halo incidences. The first set of data in the Relations section of Table 6 gives the fraction of each sky type, as it produced an ice halo incident. For example, in January we found that 9 % of all cirrostratus skies were

30   accompanied by a 22° halo. In the data for April, this fraction increased to 22% of all cirrostratus skies. We also have registered halos for a portion of partly cloudy skies, and for cloudy skies. No halos have been registered in any of the clear skies. However, we must consider reasons for the PCL and CLD halo incidences. Upon random sampling of these combinations we find the following: The PCL indicator has been assigned to images that have a highly varied cirroform sky, including halo appearances. In a few instances, low clouds triggered the PCL indicator, however, a CS layer at higher altitude still contributed a halo score

above the threshold. Many of the halo scores in CLD skies belong to images with a cloudy appearance, however, most likely belong to a thickening and lowering altostratus as often found in warm front approaches. So, these are not false signals, but certainly conditioned by the limitations of the sky type classification. The second set of numbers shows the fraction of all halos associated with the various sky types, which may be a little easier to interpret. In January, 49% of all halo incidences occurred

5   in CS skies, while in March this number was 87%. One of the conclusion to be made from the relation between STS and IHS concerns the confidence in the presence of smooth crystalline habits among the cloud particles, as shown only in a one-fifth fraction of all cirrostratus. If the halo algorithm is used on TSI images to identify the appearance of ice halos, the next useful and logical step will be to relate these data to other instrument records: LIDAR for altitude, particle density, and particle phase (solid or liquid), photometric measurements to glean information on radiative flux. ARM sites have accumulated such

10   instrumental data. The algorithm proposed here will make such data investigation possible, and delivers support for crystal identification.

4   Summary

ARM sites have produced long-term records of sky images. With the goal of using these long-term image records to provide supporting information on cirrus composition, we developed an algorithm that assigns sky type and halo scores to long-term

15   series of TSI images. We described this algorithm in this paper, including the image preparation to generate a standardized image section centred at the sun, called the Local Sky Map (LSM). A multivariate analysis of selected LSM properties, as supported by a continually developed master table, allows the assignment of scores with respect to sky type and ice halo presence in the solar-near section of the sky. In particular, we focus on the properties associated with the radial brightness behaviour around the sun. Physically, the number and type of scattering centres in the atmosphere does influence the brightness

20   gradient, thus giving us access to an assessment of cloud type and cloud cover. The brightness fluctuation associated with the 22° halo provides a further set of properties specific to the presence of a 22° halo. We analyse all four quadrants adjacent to the sun separately, then combine the scores into a raw image score. For the ice halo score, we also apply a Gaussian broadening across the time series. The algorithm has been found to be about 90% in agreement with the visually assigned sky type, and 86% in agreement with the visually identified ice halo score. The application to the first four months of the TSI records from

25   SGP ARM site indicates periods of ice halo presences, with a most common duration of about 5 to 10 minutes, but lasting up to 3 hours. It allowed to identify the fraction of cirrostratus skies that do produce ice halos, as well as find such data for other sky types as well. The algorithm now will be applied to deliver ice halo data for the long-term TSI records accumulated in various geographical locations of ARM sites, and allows further investigation into correlations with other instrumental records from those sites. In particular, LIDAR data for altitude and optical thickness measurements, as well as depolarization analysis

30   will be a useful combination with this photographic ice halo record. It is reasonable to expect that the reference set for sky type determination will improve with the support of LIDAR data. The method described here may be suitable to expand to other types of sky analysis on TSI images.

Atmospheric
Measurement
Techniques
Discussions

Open Access

EGU

Author contribution

Sylke Boyd is the main author of this paper and the code. The four co-authors worked on the algorithm as undergraduate researchers. Stephen Sorenson decided on the use of C++ and opencv3.2 for image manipulation, and initiated the program code. Shelby Richard worked out the details of the radial intensity computation and properties. Michelle King and Morton

5   Greenslit contributed algorithm parts to eliminate optical distortions and low-cloud obstruction, and input management. SR, MK, and MG all contributed to data collection and analysis.

Competing Interests

15   The authors declare that they have no conflict of interest.

Code availability

Code and accessory files are made available at GitHub under DOI 10.5281/zenodo.8475 (Boyd et al., 2018).

References

20   Boyd, S., Sorenson, S., Richard, S., King, M., and Greenslit, M.: Haloloop software. https://zenodo.org/badge/latestdoi/157256231, 2018.

Calbó, J. and Sabburg, J.: Feature Extraction from Whole-Sky Ground-Based Images for Cloud-Type Recognition, Journal of Atmospheric & Oceanic Technology, 25, 3-14, 2008.

25   Campbell, J. R., Lolli, S., Lewis, J. R., Gu, Y., and Welton, E. J.: Daytime Cirrus Cloud Top-of-the-Atmosphere Radiative Forcing Properties at a Midlatitude Site and Their Global Consequences, Journal of Applied Meteorology & Climatology, 55, 1667-1679, 2016.

Cziczo, D. J. and Froyd, K. D.: Sampling the composition of cirrus ice residuals, Atmospheric Research, 142, 15-31, 2014.

30   Delene, D. J.: Airborne data processing and analysis software package, Earth Science Informatics, 4, 29-44, 2011.

ENA, A.-T. a.: 2013, updated hourly. Total Sky Imager (TSISKYIMAGE). 2013-10-01 to 2018-05-28, Eastern North Atlantic (ENA) Graciosa Island, Azores, Portugal (C1). Compiled by V. Morris. Atmospheric Radiation Measurement (ARM) Climate Research Facility Data Archive: Oak Ridge, Tennessee, USA. Data set accessed 2018-06-07 at http://dx.doi.org/10.5439/1025309., 2018.

5   Ewald, F., Zinner, T., and Mayer, B.: Remote sensing of particle size profiles from cloud sides: Observables and retrievals in a 3D environment, AIP Conference Proceedings, 1531, 83-86, 2013.

Fasullo, J. T. and Balmaseda, M. A.: Earth's Energy Imbalance, Journal of Climate, 27, 16p, 2014.

10   Fasullo, J. T. and Kiehl, J.: Earth's Global Energy Budget, Bulletin of the American Meteorological Society, 90, 13p, 2009.

Fasullo, J. T., von Schuckmann, K., and Cheng, L.: Insights into Earth's Energy Imbalance from Multiple Sources, Journal of Climate, 29, 11p, 2016.

15   Forster, L., Seefeldner, M., Wiegner, M., and Mayer, B.: Ice crystal characterization in cirrus clouds: a sun-tracking camera system and automated detection algorithm for halo displays, Atmos. Meas. Tech., 10, 2499-2516, 2017.

Fu, Q., Lohmann, U., Mace, G. G., Sassen, K., and Comstock, J. M.: High-Cloud Horizontal Inhomogeneity and Solar Albedo Bias, Journal of Climate, 15, 19p, 2002.

20   Ghonima, M. S., Urquhart, B., Chow, C. W., Shields, J. E., Cazorla, A., and Kleissl, J.: A method for cloud detection and opacity classification based on ground based sky imagery, Atmospheric Measurement Techniques, 5, 2881-2892, 2012.

Hammer, A., Monahan, A. H., Schmidt, T., and Heinemann, D.: Simulating clear-sky index increment correlations under mixed sky
25   conditions using a fractal cloud model, Solar Energy, 150, 10p, 2017.

Hammer, E., Bukowiecki, N., Luo, B. P., Lohmann, U., Marcolli, C., Weingartner, E., Baltensperger, U., and Hoyle, C. R.: Sensitivity estimations for cloud droplet formation in the vicinity of the high-alpine research station Jungfraujoch (3580 m a.s.l.), Atmos. Chem. Phys., 15, 10309-10323, 2015.
30

IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge, United Kingdom and New York, USA, 1535 pp., 2013.

IPCC: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the
35   Intergovernmental Panel on Climate Change Geneva. Switzerland, 151 pp., 2014.

Kandel, R. and Viollier, M.: Observation of the Earth's radiation budget from space, Observation du bilan radiatif de la Terre depuis l'espace, 342, 286-300, 2010.

40   Kennedy, A., Dong, X., and Xi, B.: Cloud fraction at the ARM SGP site: reducing uncertainty with self-organizing maps, Theoretical & Applied Climatology, 124, 43-54, 2016.

Knobelspiesse, K., van Diedenhoven, B., Marshak, A., Dunagan, S., Holben, B., and Slutsker, I.: Cloud thermodynamic phase detection with polarimetrically sensitive passive sky radiometers, Atmos. Meas. Tech., 8, 1537-1554, 2015.
45

Kollias, P., Tselioudis, G., and Albrecht, B. A.: Cloud climatology at the Southern Great Plains and the layer structure, drizzle, and atmospheric modes of continental stratus, Journal of Geophysical Research: Atmospheres, 112, 2007.

Leisner, T., Duft, D., Möhler, O., Saathoff, H., Schnaiter, M., Henin, S., Stelmaszczyk, K., Petrarca, M., Delagrange, R., Hao, Z., Lüder, J.,
50   Petit, Y., Rohwetter, P., Kasparian, J., Wolf, J.-P., and Wöste, L.: Laser-induced plasma cloud interaction and ice multiplication under cirrus cloud conditions, Proceedings of the National Academy of Sciences, 110, 10106-10110, 2013.

Long, C. N., Sabburg, J. M., Calbó, J., and Pagès, D.: Retrieving Cloud Characteristics from Ground-Based Daytime Color All-Sky Images, Journal of Atmospheric and Oceanic Technology, 23, 633-652, 2006.
55

NSA, A.-T. a.: Atmospheric Radiation Measurement (ARM) Climate Research Facility. 2006, updated hourly. Total Sky Imager (TSISKYIMAGE). 2006-04-25 to 2018-04-11, North Slope Alaska (NSA) Central Facility, Barrow AK (C1). Compiled by V. Morris. Atmospheric Radiation Measurement (ARM) Climate Research Facility Data Archive: Oak Ridge, Tennessee, USA. Data set accessed 2018-06-07 at http://dx.doi.org/10.5439/1025309., 2018.

OLI, A.-T. a.: Atmospheric Radiation Measurement (ARM) Climate Research Facility. 2013, updated hourly. Total Sky Imager (TSISKYIMAGE). 2013-08-30 to 2018-05-24, ARM Mobile Facility (OLI) Olikiok Point, Alaska; AMF3 (M1). Compiled by V. Morris. Atmospheric Radiation Measurement (ARM) Climate Research Facility Data Archive: Oak Ridge, Tennessee, USA. Data set accessed 2018-06-07 at http://dx.doi.org/10.5439/1025309., 2018.

Sassen, K., Zhu, J., and Benson, S.: Midlatitude cirrus cloud climatology from the Facility for Atmospheric Remote Sensing. IV. Optical displays, Appl. Opt., 42, 332-341, 2003.

Schwartz, S. E., Charlson, R. J., Kahn, R., and Rodhe, H.: Earth's Climate Sensitivity: Apparent Inconsistencies in Recent Assessments, Earth's Future, 2, 601-605, 2014.

Seefeldner, M., Wiegner, M., and Mayer, B.: Ice Crystal Characterization in Cirrus Clouds: A Sun-tracking Camera System and Automated Detection Algorithm for Halo Displays, Atmospheric Measurement Techniques Discussions, doi: 10.5194/amt-2017-17, 2017. 24p, 2017.

SGP, A.-T. a.: Atmospheric Radiation Measurement (ARM) Climate Research Facility. 2000, updated hourly. Total Sky Imager (TSISKYIMAGE). 2000-07-02 to 2018-04-19, Southern Great Plains (SGP) Central Facility, Lamont, OK (C1). Compiled by V. Morris. Atmospheric Radiation Measurement (ARM) Climate Research Facility Data Archive: Oak Ridge, Tennessee, USA. Data set accessed 2018-06-07 at http://dx.doi.org/10.5439/1025309., 2018.

Tian, L., Heymsfield, G. M., Li, L., Heymsfield, A. J., Bansemer, A., Twohy, C. H., and Srivastava, R. C.: A Study of Cirrus Ice Particle Size Distribution Using TC4 Observations, Journal of the Atmospheric Sciences, 67, 195-216, 2010.

Trenberth, K. E., Zhang, Y., and Fasullo, J. T.: Relationships among top-of-atmosphere radiation and atmospheric state variables in observations and CESM, Journal of Geophysical Research: Atmospheres, 120, 10,074-010,090, 2015.

van Diedenhoven, B.: The prevalence of the 22 deg halo in cirrus clouds, Journal of Quantitative Spectroscopy & Radiative Transfer, 146, 5p, 2014.

Weihs, P., Rennhofer, M., Baumgartner, D. J., Gadermaier, J., Wagner, J. E., Gehring, J. E., and Laube, W.: Potential impact of contrails on solar energy gain, Atmos. Meas. Tech., 8, 1089-1096, 2015.

Atmospheric
Measurement
Techniques
Open Access
EGU
Discussions

Table 1. TSI data set properties

Location	Dates and times (UTC)		Image interval	Resolution (pixels)
Southern Great Plains (SGP, 2018)	2 Jul 2000 0:35:00	15 Aug 2011 01:17:30	30 s	288×352
36° 36′ 18″ N, 97° 29′ 6″ W	15 Aug 2011 22:17:30	19 Apr 2018 01:02:00	30 s	480×640
North Slope of Alaska (NSA, 2018)	25 Apr 2006 21:44:00	2 Nov 2010 21:31:00	30 s	288×352
71° 19′ 22.8″ N, 156° 36′ 32.4″ W	9 Mar 2011 01:08:30	11 Apr 2018 18:59:30	30 s	480×640
Eastern North Atlantic (ENA, 2018)	1 Oct 2013 08:13:00	28 May 2018 21:04:00	30 s	480×640
39° 5′ 29.76″ N, 28° 1′ 32.52″ W				

Atmospheric
Measurement
Techniques
Open Access
Discussions

Table 2 Sky Type descriptions

Sky type		Visual description
Cirrostratus	CS	Muted blue, no sharp cloud outlines; solar position clearly visible, bright scattering disk or halo may be present; changes are gradual and slow (several minutes)
Partly cloudy	PCL	Variable sky with sharply outlined stratocumulus or altocumulus; variations between sky quadrants; sun may be obscured; changes are abrupt and fast (less than two minutes)
Cloudy	CLD	Sun is obscured; low brightness; low blue intensity values; stratus, nimbostratus, altostratus, or cumulonimbus; changes occur slowly (order of hours)
Clear	CLR	Blue, cloud-free sky; sun clearly visible and no bright scattering disk around it; changes are slow (order of hours)
No data	N/A	This may occur at low sun positions for the bottom quadrants of the LSM, or due to overexposure in the near-solar region of the image; it's the default at night.

Atmospheric
Measurement
Techniques
Discussions

Table 3 STS properties, their averages, and standard deviations for the each sky type in the Master table. All units based on colour intensity values and LSM units. The sky type assignment is based on visual assessment the images. Number of records for each sky type is indicated in parentheses.

STs property	CS (155)	PCL (99)	CLD (93)	CLR (96)
Slope a	B -3.0 ±1.5	B -1.6 ±2.2	B -0.7 ±1.7	B -2.3 ±1.6
	G -3.2 ±1.7	G -1.6 ±2.2	G -0.7 ±1.7	G -2.8 ±1.6
	R -3.6 ±1.9	R -1.9 ±2.6	R -0.8 ±1.8	R -2.8 ±1.7
Intercept b	B 276 ±34	B 248 ±46	B 193 ±40	B 248 ±43
	G 271 ±33	G 240 ±53	G 195 ±44	G 233 ±47
	R 255 ±48	R 228 ±65	R 179 ±47	R 184 ±47
ASD[1]	B 13.1 ±5.3	B 20.5 ±7.0	B 14.2 ±5.0	B 15.4 ±5.2
	G 15.0 ±6.0	G 22.9 ±7.7	G 15.0 ±5.1	G 16.3 ±5.3
	R 16.6 ±6.6	R 25.5 ±8.1	R 15.8 ±5.6	R 14.8 ±5.7
ACR[2]	1.33 ±0.36	1.24 ±0.32	1.08 ±0.12	2.07 ±0.11

[1] Areal Standard Deviation; [2]Average Colour Ratio

5

Table 4 Halo scoring properties. These 31 properties define the space in which an image is scored for a halo. The averages given are from the master file and include 188 records from halo-containing sky quadrants, visually assessed.

IHS property	B	G	R
Slope a	-3.3 ±1.5	-3.3 ±1.6	-3.8 ±1.8
Intercept b	279 ±35	278 ±37	268 ±45
ASD	12.6 ±4.7	14.8 ±6.0	16.2 ±6.4
Maximum upslope η'_{up}	2.1 ±1.3	2.1 ±1.4	2.5 ±1.6
Maximum downslope η'_{down}	-1.6 ±1.0	-1.6 ±1.0	-1.8 ±1.1
Upslope location s_{up}	17.5 ±1.9	17.8 ±2.3	17.5 ±2.1
Maximum location s_{max}	18.9 ±1.9	19.1 ±2.3	18.8 ±2.1
Downslope location s_{down}	20.0 ±2.1	20.2 ±2.4	19.9 ±2.2
Number of maxima n_{max}	2.4	2.6	2.5
BGR consistency	$\sigma_{BGR}(s_{up}) = 0.8$	$\sigma_{BGR}(s_{max}) = 0.8$	$\sigma_{BGR}(s_{down}) = 0.9$
ACR		1.2±0.3	

5

Table 5. STS and IHS test results for SGP March 2018. Visual assignments were made iteratively in step with the algorithm results. Given are the percentages of images of visual type that have been assigned an algorithm type (%vis), and the percentages of the algorithm type that correspond to a visual type (%alg). Agreement combinations in bold. IHS > 4000 to count an algorithm halo.

	STS	CS			PCL			CLD			CLR		
		N	%vis	%alg	N	%vis	%alg	N	%vis	%alg	N	%vis	%alg
	CS	**6675**	**88**	**86**	683	11	9	38	1	0	397	4	5
	PCL	182	2	3	**5513**	**87**	**91**	176	3	3	191	2	3
	CLD	61	1	1	47	1	1	**6129**	**97**	**98**	0	0	0
	CLR	641	8	6	136	2	1	0	0	0	**10529**	**95**	**93**
	N/A	12597 (40% of all images)											
	IHS	22° halo						No 22° halo					
		N	%vis	%alg				N	%vis	%alg			
	22° halo	**1996**	**85**	**88**				272	1	12			
	No 22° halo	349	15	1				**41409**	**99**	**99**			

Algorithm assignment

Visual assignment

21

Table 6. Sky type assignments and ice halo formations during the months of January through April 2018, SGP. Percentages are with respect to all classifiable images. Times are TUC.

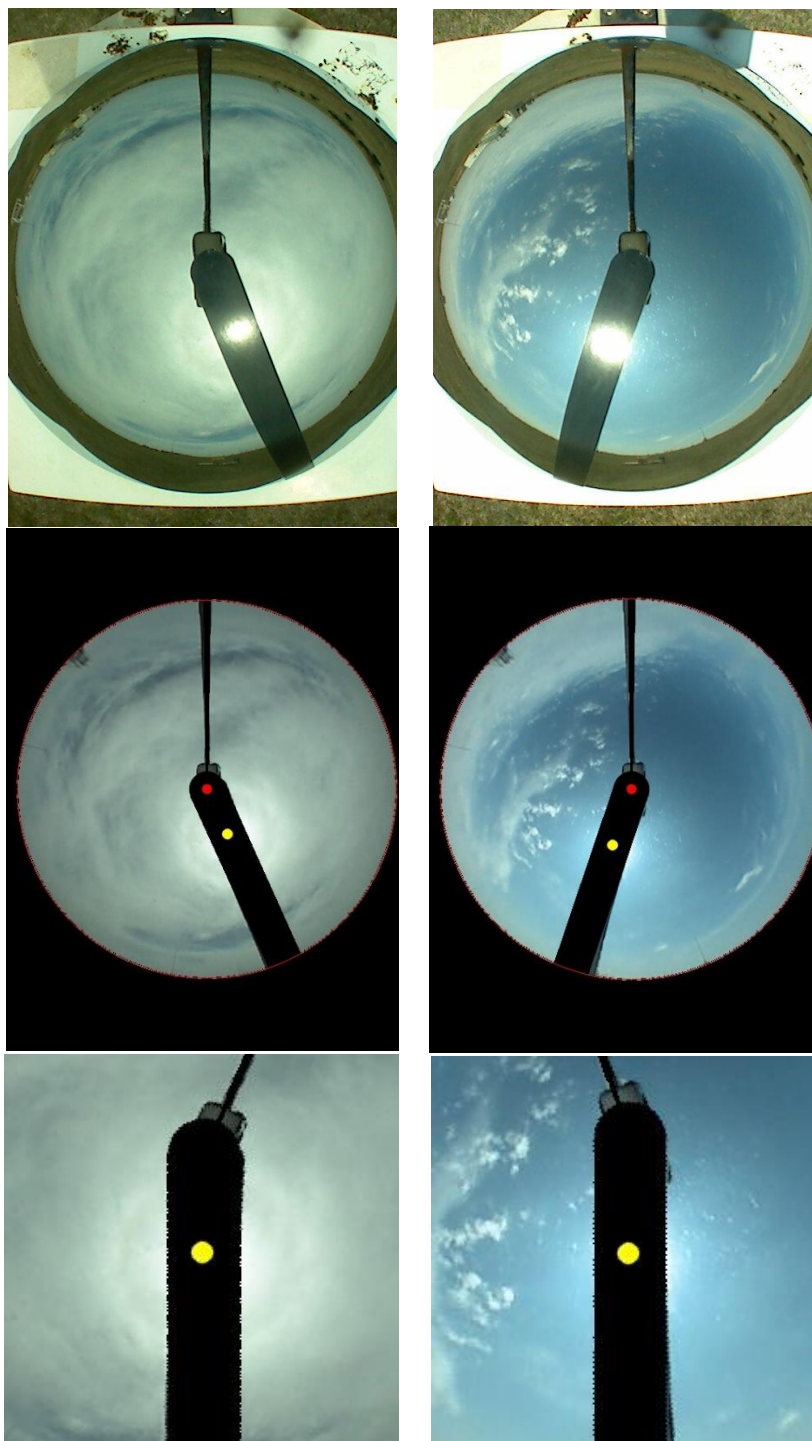|			Jan 2018	Feb 2018	Mar 2018	Apr 2018[1]
		total number of images	36632	36011	44057	27741
		Number with valid STS	21238	23604	31398	20436
		begin date of record	1 Jan 2018 13:47:00	1 Feb 2018 13:36:00	1 Mar 2018 0:00:00	1 Apr 2018 0:00:00
		end date of record	31 Jan 2018 23:50:00	28 Feb 2018 23:59:30	31 Mar 2018 23:59:30	19 Apr 2018 1:02:00
Sky types		CS	20 %	18 %	25 %	34 %
		PCL	24 %	24 %	19 %	19 %
		CLD	11 %	33 %	20 %	25 %
		CLR	45 %	25 %	36 %	22 %
Ice Halos		Number of separate halo incidents	26	45	34	46
		Mean duration	16 min	22 min	34 min	21 min
		Maximum duration	62 min	136 min	171 min	208 min
		Total halo time	411 min	998 min	1160 min	963 min
		% halo instances with				
		Full 22° halo	29 %	42 %	77 %	42 %
		¾ 22° halo	38 %	31 %	13 %	40 %
		½ 22° halo	32 %	25 %	10 %	18 %
		¼ 22° halo	1 %	1 %	0 %	0 %
Relations		% halo instances of all sky type instances				
		CS	9 %	16 %	18 %	22 %
		PCL	6 %	7 %	6 %	9 %
		CLD	4 %	5 %	10 %	12 %
		CLR	0 %	0 %	0 %	0 %
		All STS	3.9 %	8.5 %	7.4 %	9.4 %
		% sky type of all halo instances				
		CS	49 %	60 %	87 %	78 %
		PCL	42%	33 %	9 %	14 %
		CLD	2 %	5 %	3 %	5%
		CLR	0 %	0 %	0 %	0 %
		N/A	7 %	2 %	1 %	3 %

[1]incomplete month

Figure 1 Two examples for image preparation. The left column develops an image from SGP 17 April 2018 17:45:00 UTC, the right image was taken on SGP 3 April 2018 19:09:30 UTC. Top row: original image; centre row: image after colour correction, distortion removal, masking of horizon and equipment, and sun mark were applied; bottom row: final local sky map with sun at centre and a width of about 80 LSM units.
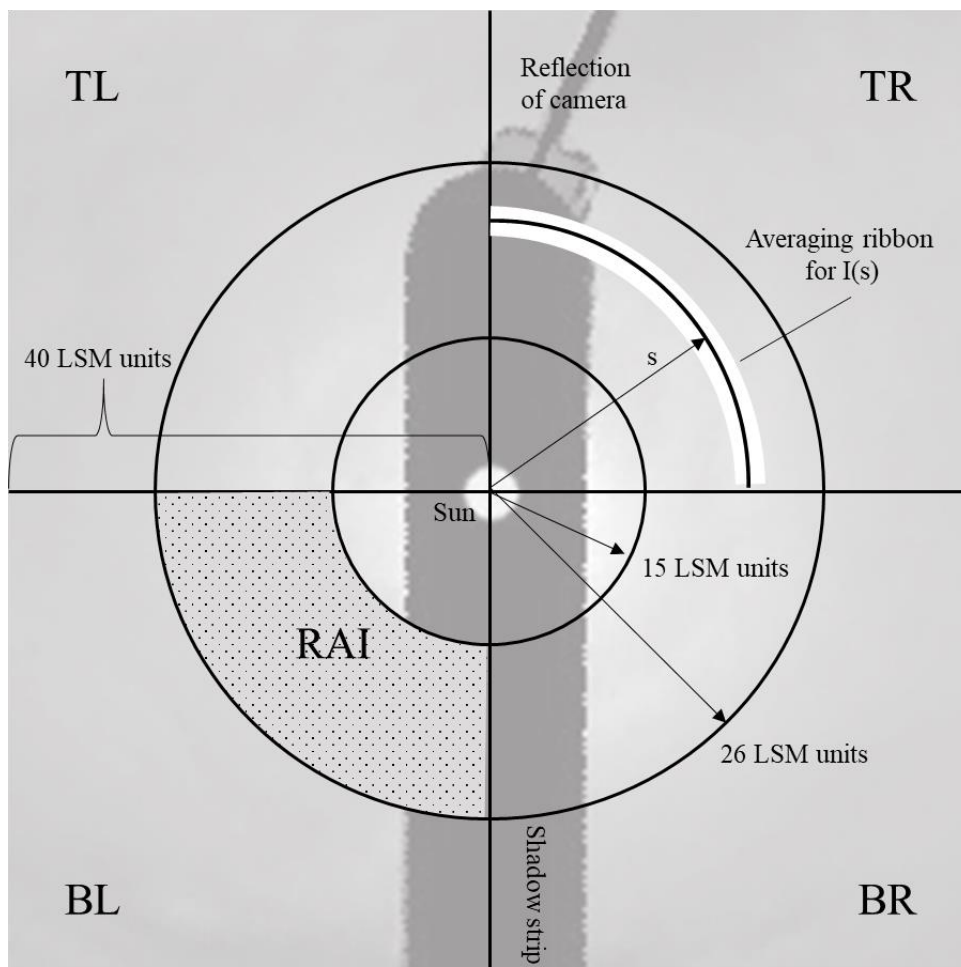
Figure 2. Layout of the local sky map (LSM). The LSM is divided into four quadrants. Shadow strip and camera are excluded from analysis.
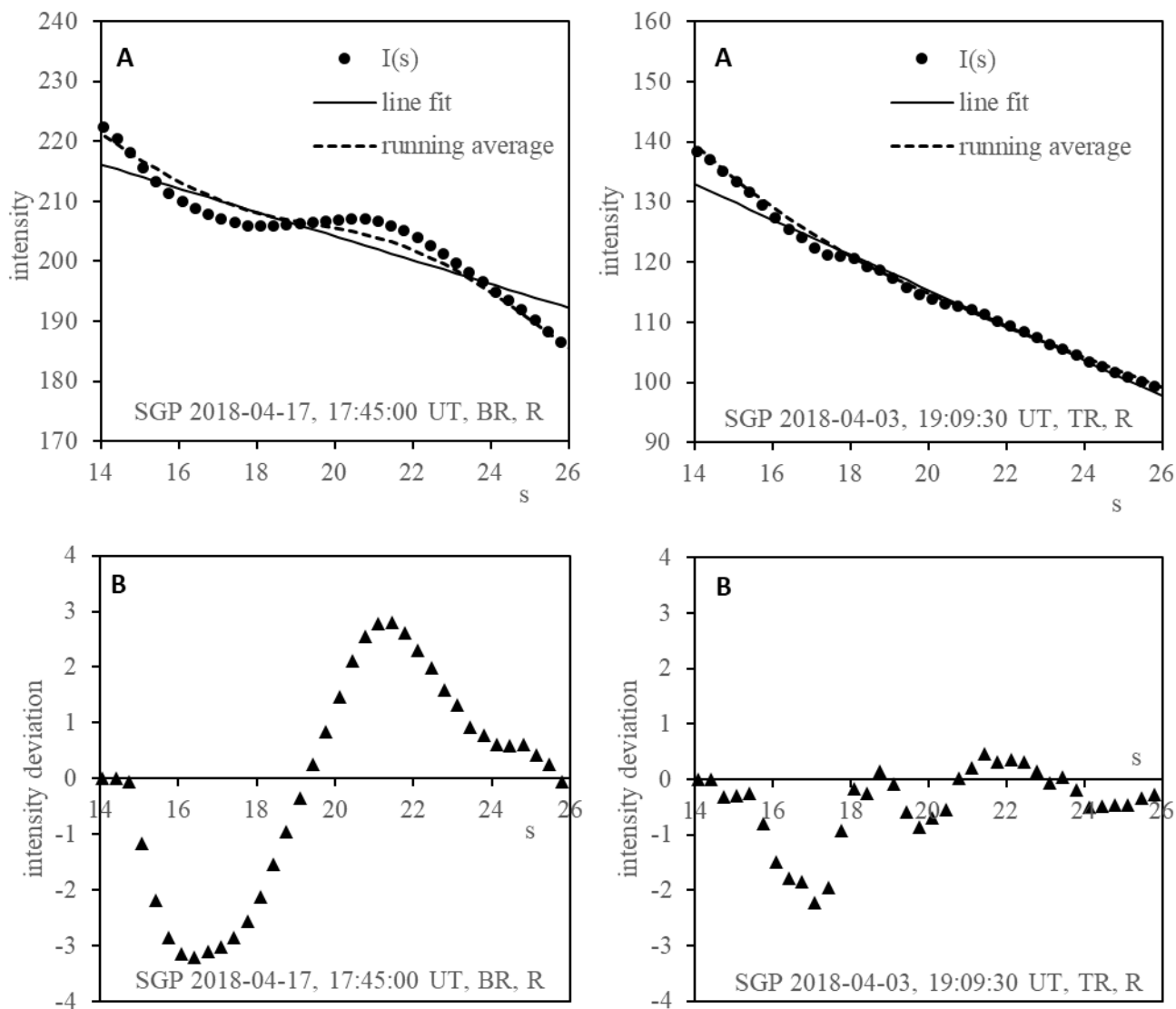
Figure 3 Average radial intensity of the red channel is shown versus radial distance s, measured in LSM units, for the two images of Figure 1, halo at left. Panel (A) includes the average intensity $I(s)$, a linear fit, and the running average $\bar{I}_6(s)$ as averaged over a width of 6 LSM units. (B) shows the radial intensity deviation $\eta(s)$.
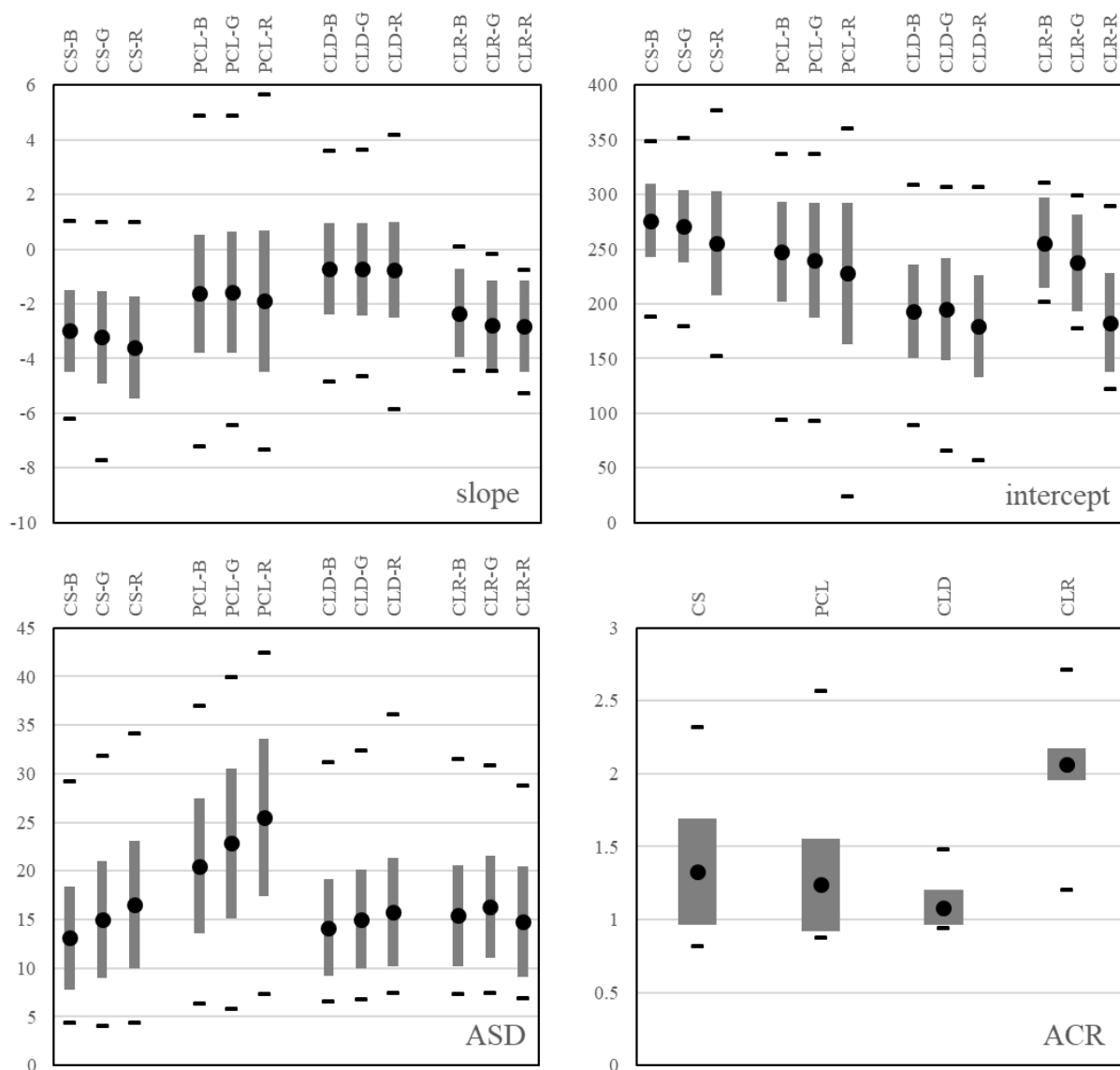
5

Figure 4 Sky type properties for the four sky types in the master table. Slope and intercept (top row) for the radial fit; areal standard deviation (AST) of brightness (bottom left); average colour ratio (ACR) (bottom right). Sky types were assigned visually. Black circles indicate the mean, gray boxes the range of the first standard deviation, black bars limit the extreme values found in the master table.
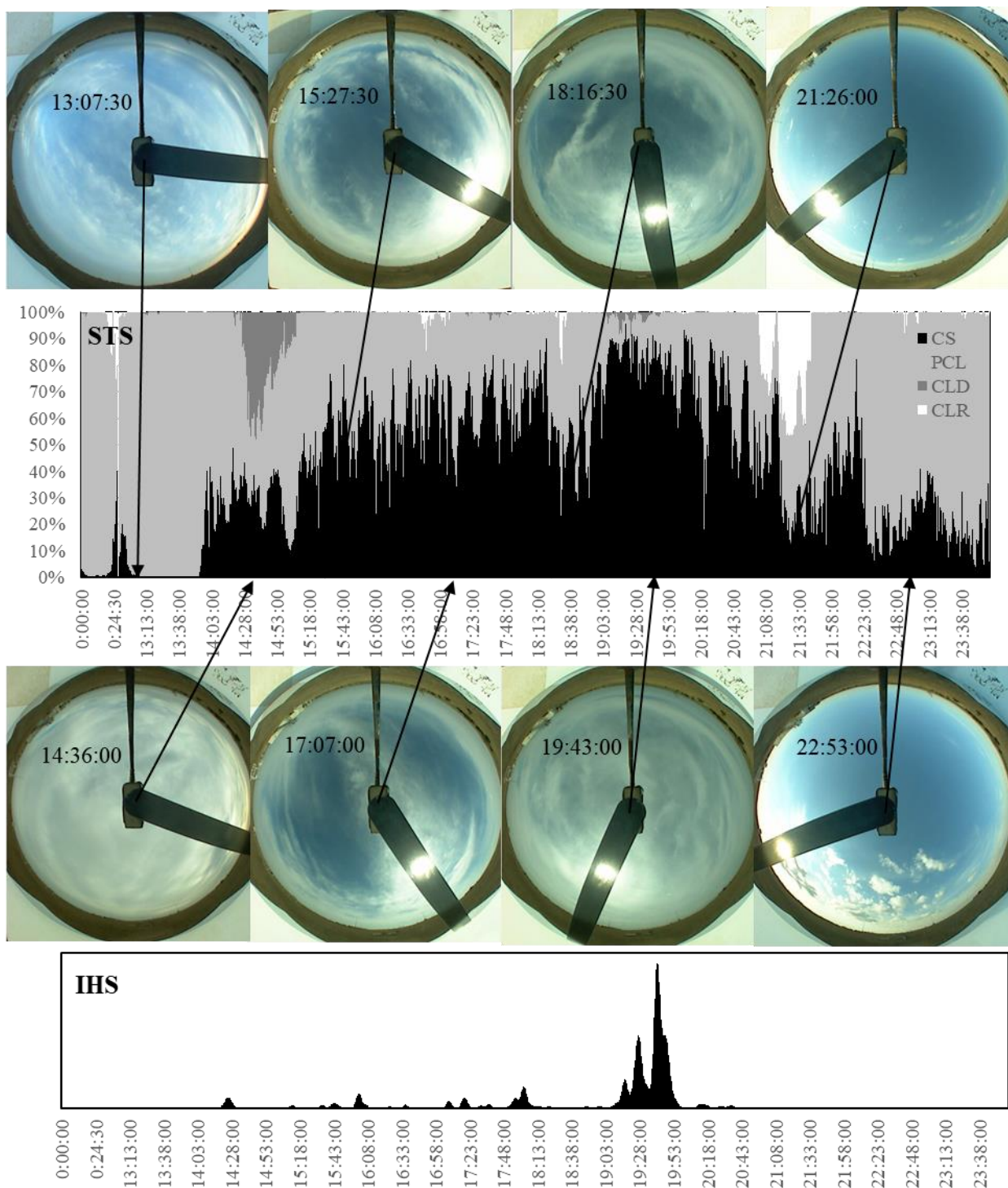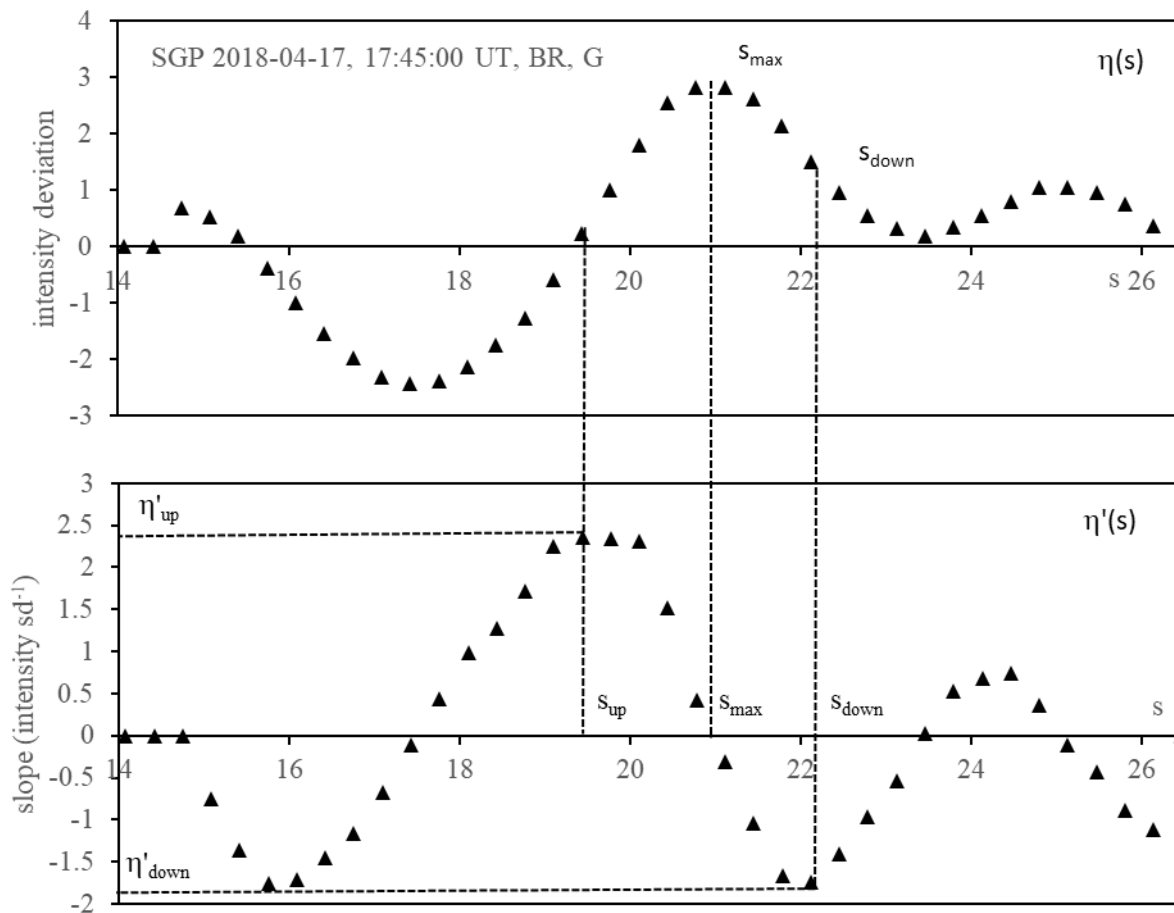
Figure 5 One-day example for STS and IHS (SGP March 10, 2018). Sample TSI images are included. The middle panel shows STS versus time of day (N/A excluded). Bottom panel is IHS versus time. All times in TUC.

Figure 6. Radial markers used in halo scoring. The data belong to the green channel of the TSI image from SGP, April 17, 2018, see Figure 1. The top panel shows the radial intensity deviation $\eta(s)$; the bottom panel shows its derivative $\eta'(s)$. The sequence of radial locations used in halo scoring is indicated, as well as the interpretation of the up- and down-slope markers.

Atmospheric
Measurement
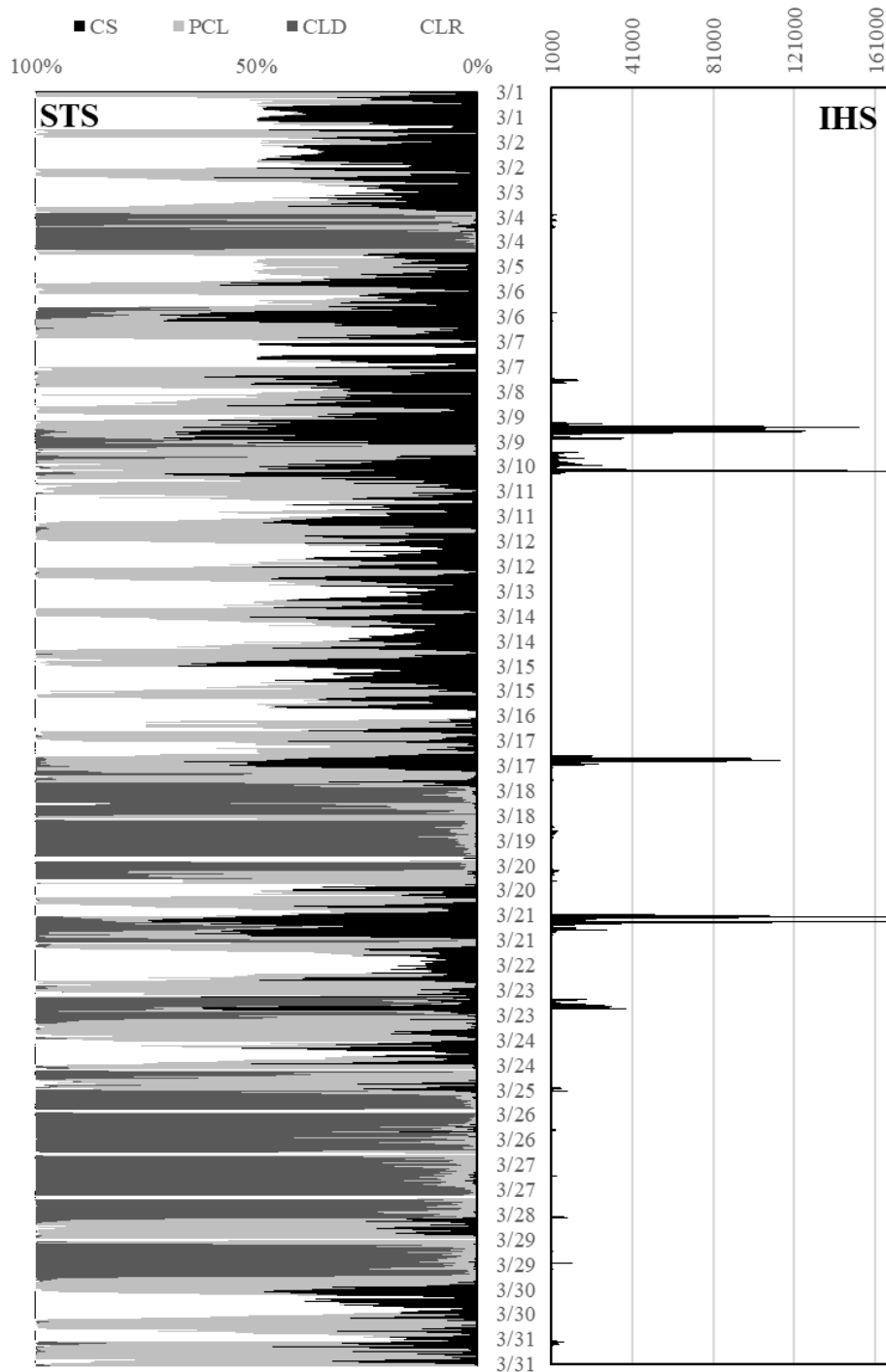Techniques
Discussions

Figure 7. Time line of Sky type scores (STS) and ice halo scores (IHS) for TSI images from SGP March 2018. Left panel shows the STS: CS – black, PCL – light grey, CLD – dark grey, CLR – white. Right panel: IHS Pre-factor C=10^6, broadening w=3.5 minutes.

Atmospheric
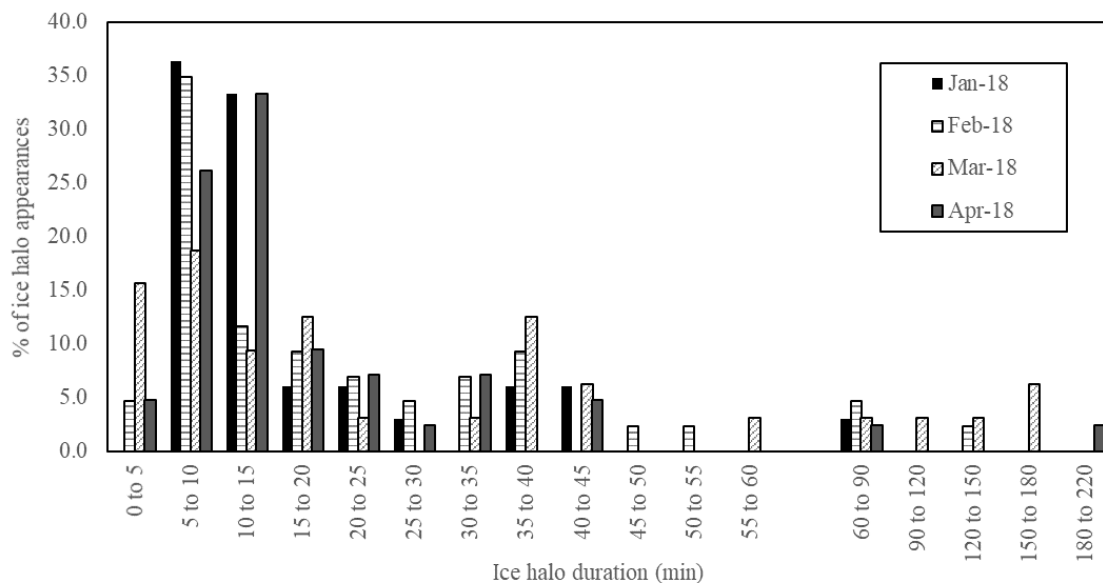Measurement
Techniques
Discussions

Figure 8. Distribution of observed ice halo durations for the first four months of 2018 at SGP ARM site.