82

Halo Identification Algorithm for Sky Images Produced in TSI Series

Sylke Boyd*, Stephan Sorenson, Shelby Richard, Michelle King, and Morton Greenslit
University of Minnesota-Morris, Morris, Minnesota

1. Motivation and Background

The radiative effects of cirroform clouds are challenging and varied. The goal of this project is to extract information about cloud composition, as well as about the spatial and temporal distribution of these clouds based on photographic ice halo observations. Cirrus optical scattering behavior is heavily influenced by the types of ice particles. These may be present in many forms, including crystalline hexagonal habits in form of plates, pencils and prisms, hollow columns, bullets and bullet rosettes, and also as amorphous ice pellets, fragments, rimed crystals and others. If a significant fraction of pristine crystal habits is present, the optical scattering behavior of the cirrus cloud gives information about the cloud particle types in form of ice halos, most frequently appearing as a bright ring of 22° radius around sun or moon. Vice versa, the absence of the ice halo would suggest that the ice particles are of a shape and structure different from hexagonal plates, pencils, bullets or prisms (van Diedenhoven, 2014).

The question becomes how to use ice halos to improve our knowledge about the composition of cirroform clouds and the conditions in the upper troposphere. It is essential to establish frequency of halo appearances across seasons and years, as well as geographically. A couple of such data collections can be found in the literature. 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. 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. Sky images have been collected for decades at several research facilities. We are using series of images produced by Total Sky Imagers (TSI) at Atmospheric Radiation Measurement (ARM) Climate Research Facilities in order to assess the presence of ice halos on these sites. We have obtained long-term series of images from three locations: (1) Southern Great Plains Central Facility (SGP, 2018), (2) Eastern North Atlantic Graciosa Island, Azores (ENA, 2018), and (3) the North Slope Alaska Facility (NSA, 2018). These images were taken every 30 seconds over many years, producing records of millions of images capturing the appearance of the sky throughout daylight hours. The TSI views the whole sky in a convex upward mirror, and captures it with a centered downward camera. These captures include appearances of the 22°halo. Other halo features cannot reliably differentiated in these images, although occasionally one might spot a parhelion or a circumzenithal arc.

We present an image-processing algorithm to automatically identify 22° ice halos in TSI images. The TSI data present as jpg images of different sizes (over the years) and different quality. While images from all three data sets were used to build and train the algorithm, our current results focus on the SGP location and the time interval between January 2013 and April 2018.

2. Algorithm

2.1. Basic Idea

The algorithm assigns a score to each image, evaluating how well it fits the description of the scored target listed in Table 1. The software itself can be found in github. (Boyd, 2018a). An extended description of the algorithm is under review at Atmospheric Measurement Techniques, manuscript number amt-2018-401 (Boyd *et al.*, 2018b).

The evaluation of the image is based on a Gaussian multivariate analysis of selected characteristic properties. For each scored target, a vector of N_p characteristic properties of the image

* *Corresponding author address:* Sylke Boyd, Division of Science and Mathematics, University of Minnesota-Morris, 600 E 4th St, Morris, MN 56267; e-mail: sboyd@morris.umn.edu

$$X = \{x_i\}_i^{N_p} \quad (1)$$

is defined. The image properties are then related to the expected mean vector

$$M = \{\mu_i\}_{i=1}^{N_P} \text{ with } \mu_i = \frac{1}{N_{master}} \sum_{k=1}^{N_{master}} x_{ik} \quad (2)$$

and the covariance matrix

$$\Sigma = \overline{(X-M)(X-M)^T} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots \\ \sigma_{21} & \sigma_{22} & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix} \quad (3)$$

To define a raw score to the image

$$F = C_0 exp\left(-\frac{1}{2}\left(X_{image} - M\right)^T \Sigma^{-1} \left(X_{image} - M\right)\right) \quad (4)$$

That means, in order to score an image with respect to a target, one must provide M and Σ^{-1}.

Table 1. Scored target types.

Sky Type Scores (STS)
CS: Cirrostratus. *Muted blue, no sharp cloud outlines; solar position clearly visible, bright scattering disk or halo may be present; changes are gradual and slow*
PCL: Partly cloudy. *Variable sky with sharply outlined stratocumulus or altocumulus; variations between sky quadrants; sun may be obscured; changes are abrupt and fast.*
CLD: Cloudy. *Sun is obscured; low brightness; low blue intensity values; stratus, nimbostratus, altostratus, or cumulonimbus; changes occur slowly.*
CLR: Clear. *Blue, cloud-free sky; sun clearly visible and no bright scattering disk around it; changes are slow*
N/A no data. *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.*
Ice Halo Score (IHS) *An ice halo is present in one or more quadrants at expected positions around the sun.*

These are computed in an external master table from visually scored images. The master table is continually trained and expanded. We are using a spreadsheet for this purpose, allowing to compute means and inverse covariance matrices in an efficient fashion. Figure 1 shows a basic outline for the algorithm.

2.2. Average Radial Intensity I(s)

Image calibrations for orientation, perspective distortion, masks, and possible discoloration are performed. The calibrations are consistent over periods of weeks and months, and do not need to be adjusted for each individual image. The solar position is marked using geographical location and time. Then the Local Sky Map (LSM) is produced by rotating and cropping the image to a sun-centered square image, see Figure 2. The LSM is divided into four quadrants, each analyzed separately.

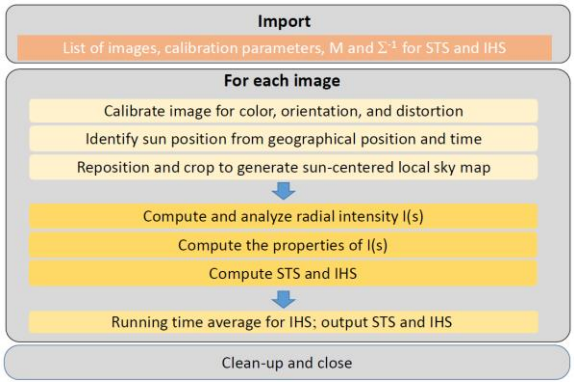

Figure 1. The algorithm is implemented in a program named haloloop. The principal flow is outlined here.

For each color channel, the radial intensity as a function of radial distance from the sun, I(s), is computed by averaging over a 4-pixel wide ribbon at distance s. In Figure 3, the average radial intensity for each of the three color channels is shown versus radial distance. On the left, graphs are for the halo images from Figure 2, on the right for the partially-cloudy image. The bottom row of panels in Figure 3 show the deviation of I(s) from its own running average (taken over 6 LSM units). One LSM unit corresponds roughly to an angle of 1 degree, however, due to some approximations made in the removal of the mirror distortion, this is only really true for zenith angles close to 45 degrees.

The examples in Figure 3 show differences between halo and non-halo images. Halo images exhibit a maximum at 21 LSM in every color channel. The relative contributions of B and G are reversed compared to a clear sky or partially cloudy sky. The radial intensity gradient, measured by the slope of a fit line in the analysis area, is a useful indicator for sky type assessments. A clear sky, for example, will exhibit a very steep initial gradient, but then decline slower, dominated by the B channel. A cloudy sky exhibits a different color ratio, and sometimes even a positive radial intensity gradient. We selected 10 properties of I(s) to assess sky type, and 31 properties of I(s) to assess the presence of a 22-degree halo.

In order to assess STS, we build the vector of characteristic properties in Eqn. 1 using slope and intercept in each color channel, the variance of the intensity across the analysis area in each color channel, and the average color ratio

$$ACR = \frac{B^2}{GR} \qquad (5)$$

as taken across the analysis area.

For the IHS, we add parameters that assess the presence of the halo intensity bump in the deviation of I(s) from running average, which adds 21 additional parameters. In addition, the resulting raw halo scores from Eqn. 4 are Gaussian-broadened in time to accommodate effects of halo persistence, and minimize false scores from incidental cloud arrangements that occasional may present a halo signal. The half width of the broadening is usually chosen to be 3 to 4 minutes. This means, halo appearances shorter than 3 minutes cannot be resolved with this approach.

The master table for each scored image quality, together with mean vector (Eqn. 2) and inverse covariance matrix (Eqn 3) was seeded using an assortment of 20 images of each target type, and has since been continually complemented with further sets of properties, yielding more than 100 reference sets for each scored quality.

1.1. Algorithm Output

The algorithm produces a time-resolved data set of sky type and halo scores for each quadrant, and for the whole LSM. The four STS scores are normalized to 100%, respectively, and can be interpreted as a probability. The IHS as a function of time is also produced. While the absolute value of IHS is arbitrary and dependent on the choice of pre-factor in Eqn.4 as well as the success of the image calibration, the discrimination between halo and no-halo images is quite clear.

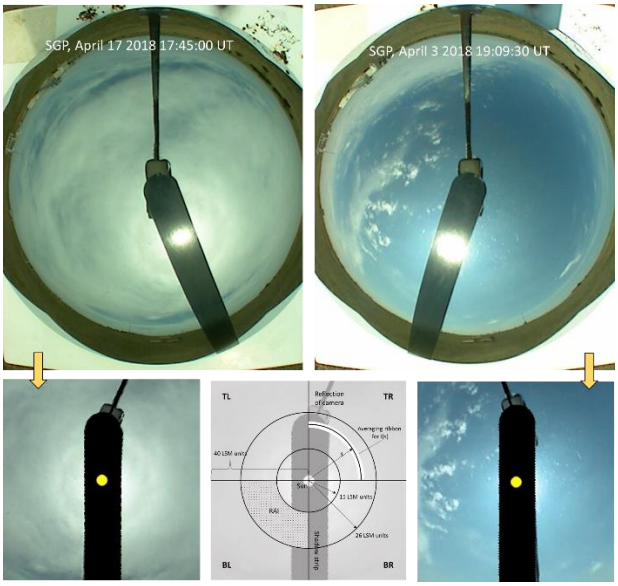

Figure 2. Two Examples of TSI images. Halo on left, partially cloudy on right. The transition from raw image to local sky map (LSM) is illustrated. The center of the bottom row illustrates how the LSM is divided for analysis.

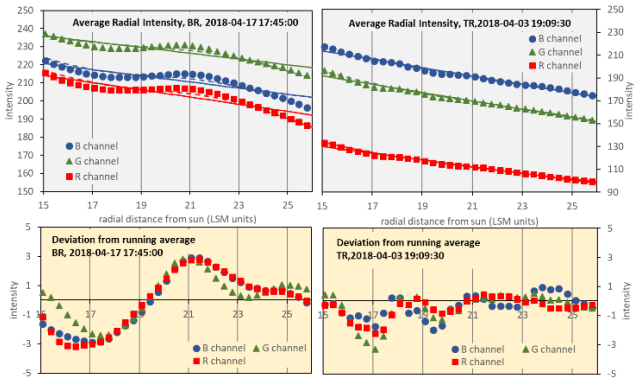

Figure 3. Average radial intensity I(s) for all three color channels. Left and right column belong to the left and right images above in Figure 2. Bottom row shows the deviation from running average.

Figure 4 shows an example of a single-day analysis. Four TSI snapshots at different times are included to illustrate such a time line. This particular day was chose for its variability. The halo is clearly detected, as well as its strength in terms of how closely it resembles the reference set in the master table at different times. Sky types have been correctly assessed. It is interesting that one can have multiple sky types in close resemblance, and observe their varying contributions along the time line.

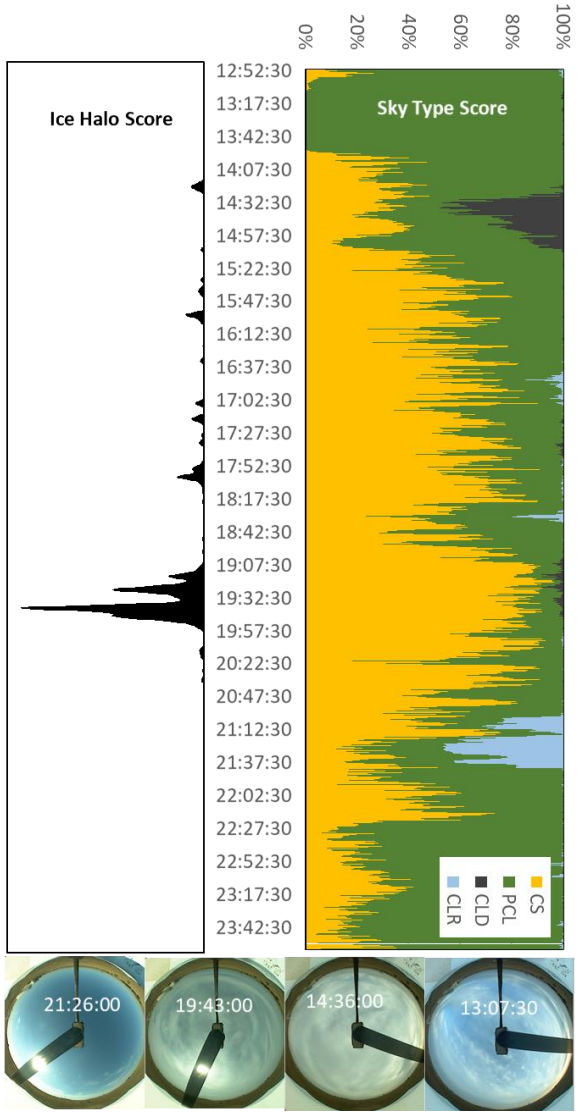

Figure 4. Example of a single-day analysis. This is for SGP March 10 2018.

2. Performance Testing

The algorithm was trained and tested for accuracy in both, STS and IHS, based on the entire SGP TSI record for March 2018; the month contains 44,026 images. We proceeded iteratively. Each iteration used the algorithm assignments to visually inspect selected segments of the image timeline and refine the master table. In the result, a complete visually assigned record for March was produced, a working master table had been created, and a good understanding of the performance and limits of the algorithm had been gained. In Table 2, we show the percentage of visual types as they are classified by the algorithm. For example, 97 % of all CLD images are correctly recognized by the algorithm, while 8 % of visual CS is classified as CLR.

Table 2. STS Performance: Percentage of each visual type as classified by the algorithm.

STS	Visually assigned				
Algorithm assigned	%	CS	PCL	CLD	CLR
	CS	**88**	11	1	4
	PCL	2	**87**	3	2
	CLD	1	1	**97**	0
	CLR	8	2	0	**95**
	N/a	40% of all images			

Table 3. STS Performance: Percentage of each algorithm type as it corresponds to visual types.

STS	Visually assigned				
Algorithm assigned	%	CS	PCL	CLD	CLR
	CS	**86**	9	0	5
	PCL	3	**91**	3	3
	CLD	1	1	**98**	0
	CLR	6	1	0	**93**
	N/a	40% of all images			

Similar, in Table 3 we examine the type assessments by the algorithm and relate them to the visual inspections of the images. The algorithm performs extremely well for CLD and CLR. For CS and PCL, we find that 9% of the CS assignments would visually present as PCL. It is, however, necessary to note that visual assignments from TSI images are carrying an uncertainty of their own. The distinction between CLR and a thin CS visually can often only be made by

inspecting an animation. In addition, a CS sky can present as inhomogeneous, triggering an affinity to PCL. Therefore, we will engage with the ceilometer and pulsed LIDAR records, if available, in our future work, to alleviate the remaining uncertainty as to sky type assignments. Table 4 shows a similar assessment for the halo recognition performance of the algorithm. Twelve percent of the algorithm assignments refer to images that do not visually show an ice halo. Often these are altocumulus skies, which are very challenging to sort out from a halo signal. The halo score is of low value for these images, but similarly low for some actual halo images. Thus, we cannot exclude these from the record using the IHS alone. The 15% of visual halo assignments that are missed by the algorithm often refer to images with low clouds, triggering disturbances in I(s) that prevent a correct interpretation of the halo signal.

Table 4. IHS Performance. Percentage of visual type recognized by the algorithm (%vis), and percentage of algorithm types corresponding to visual types (%alg).

IHS		Visually assigned			
Algorithm assigned	%	Halo present		No halo	
		%vis	%alg	%vis	%alg
	halo	**85**	**88**	1	12
	No halo	15	1	**99**	**99**

3. Findings for SGP 2011-2018

We applied the algorithm to the TSI data from the Southern Great Plains ARM site, dated January 2011 to April 2018. During this time period, we found on average 50 incidences of ice halos per month. These were not evenly distributed, but rather occurred in groups related to the changes in weather pattern. The average halo duration observed was 21 minutes, however, on average the maximum monthly persistence time was 143 min. The longest persistence of a 22-degree halo in this 5-year period was found with 428 minutes on March 7 2012, followed by 412 minutes on September 12 2017.

Since sky quadrants are analyzed separately, we were able to estimate fractional halos versus full halos as well. We found full halos in 34% of all cases. Most commonly with 35% were three-quarter halos, followed by halos only in the upper two quadrants with 26%. The latter is particularly common during the winter months when the low solar position does not allow to reliably resolve the horizon-near portions of the LSM.

We examined seasonal dependencies of the scored image properties. In Figure 5, the annual variation in dominant sky type has been summarized. Figure 6 shows the percentage of classifiable images that also show a halo signal. Overall, there is a 10-15% chance to find a halo in the sky, maximizing in March and April. That surprisingly coincides with a local minimum in the chance for CS.

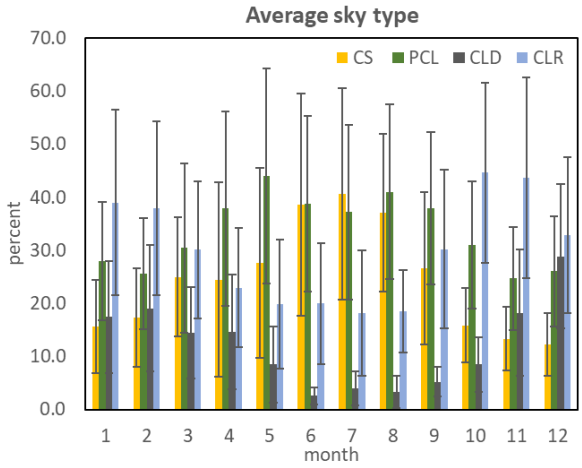

Figure 5. Average Sky Type by month. The data include January 2011 to April 2018.

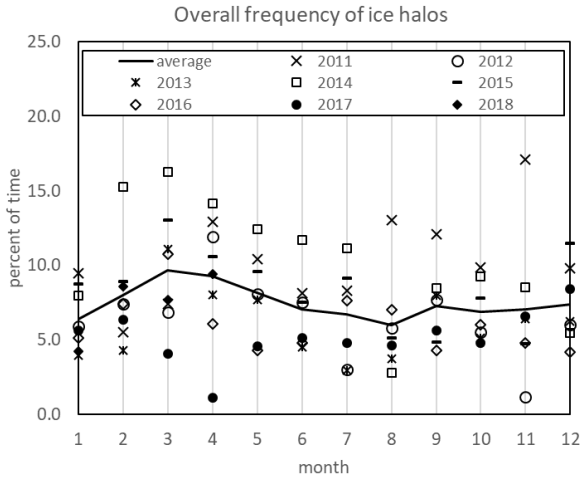

Figure 6. Percentage of classifiable images that also show an ice halo versus month of the year.

Figure 7 examines the relation between CS and halos. March and April also are the months in which CS is most likely to produce ice halos, in some years exceeding 20 % of all CS skies. That fraction has a minimum at the end of summer, in August. This would indicate that cirrostratus in March and April has a higher crystalline component than in other months of the year. CLD and CLR do not show any significant halo component.

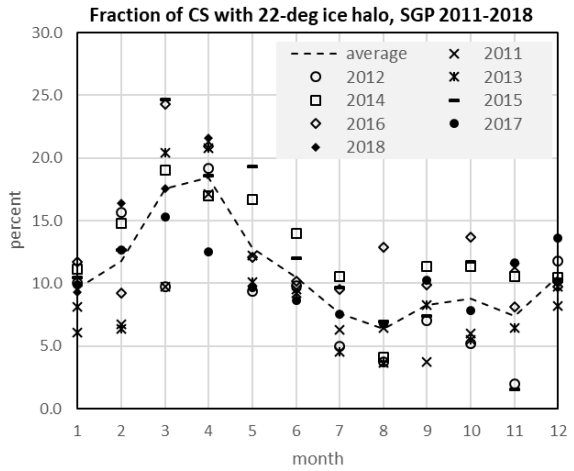

Figure 7.Percentage of all CS image quadrants that also show a halo signal.

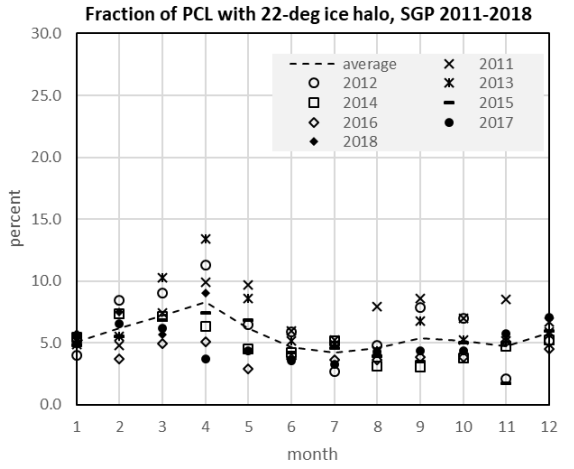

Figure 8. Percentage of Partly cloudy skies (PCL) that also show a halo signal.

We find that 5 to 10 % of PCL images do show a full or partial halo, as indicated in Figure 8. From our performance testing we know that some CS is qualified as PCL, either due to inhomogeneities in the cirroform layer, or due to the presence of low-layer stratocumulus. One should not interpret this as low-layer clouds generating halos. The PCL – halo relation shows the same annual behavior as the CS-halo relation.

4. Summary

An algorithm for the detection of ice halos in TSI images has been developed and applied to the recent several years of the TSI record collected at the SGP ARM site. This algorithm was tested and trained on a complete month's record, taken at the SGP site in March 2018.

The algorithm is flexible and trainable, and can be expanded for other target features. Sky type and ice halo scores are assigned based on the behavior of the radial brightness gradient in the near-solar region of an image. Tests show that the scoring of sky type and halo presence is about 90% reliable.

Data on the annual variation in sky type and annual distribution of ice halo appearances have been presented. In order for an ice halo to form, smooth crystal habit must be represented in the atmosphere. We find that this crystalline habit peaks during March and April for the SGP site.

Further work will address:

- An analysis of the complete SGP record, as well as NSA, and ENA records. We will find insight into temporal and geographical distributions of ice halos and their relation to cirroform clouds.
- Such an analysis will be significantly strengthened with the inclusion of ceilometer and Lidar data, depending on availability.
- The usefulness of other radiative measurements for the analysis will be explored.

5. Acknowledgement

Data were obtained from the Atmospheric Radiation Measurement (ARM) Program sponsored by the U.S. Department of Energy, Office of Science, Office of

Biological and Environmental Research, Climate and Environmental Sciences Division. The work was supported by The Undergraduate Research Opportunities Program (UROP) at the University of Minnesota, as well as a grant to the University of Minnesota, Morris from the Howard Hughes Medical Institute through the Precollege and Undergraduate Science Education Program. SB wishes to thank the University of Minnesota-Morris for the generous one-semester release from teaching obligations, allowing the completion of this work.

References

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

Boyd, S., Sorenson, S., Richard, S., King, M., and Greenslit, M.: Analysis Algorithm for Sky Type and Ice Halo Recognition in All-Sky Images, under review at Atmosph. Meas. Techn., manuscript number amt-2018-401, submitted November 15 2018, 2018b.

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.

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.

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.

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.

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.

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