https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

Automatic detection of laminated clouds in lidar data

Emily M. McCullough[1], Robin Wing[2], and James R. Drummond[1]

[1]Department of Physics and Atmospheric Science, Dalhousie University, 1453 Lord Dalhousie Drive, PO Box 15000, Halifax, NS, B3H 4R2, Canada
[2]Leibniz Institute for Atmospheric Physics, Schloßstraße 6, 18225 Kühlungsborn, Germany

Correspondence: Emily M. McCullough (e.mccullough@dal.ca)

Abstract.

In the Arctic, days with clouds containing laminations (layers), as identified by visual inspection of lidar data (McCullough et al., 2019), are strongly correlated with ground-level precipitation (McCullough et al., 2021). This provides insight into the cloud conditions and processes involved in the production of snow. We now present a computationally-inexpensive computer algorithm that (1) automatically identifies laminated cloud features in time-altitude arrays of lidar data and (2) classifies the sky scene according to the presence of laminated clouds. A discretized residual method is used.

Algorithm validation uses the 2016-2019 dataset from McCullough et al. (2019): 532 nm backscatter, up to 5 km altitude, measured by the Canadian Network for the Detection of Atmospheric Composition Change (CANDAC) Rayleigh-Mie-Raman lidar (CRL; 1 min, 7.5 m resolution) at Eureka, Nunavut (79.6° N, 85.6° W). Sky scene classification at daily resolution is replicated with accuracy = 0.86. Correlations with snow are nearly identical: r = 0.62 (algorithm) and r = 0.63 (visual).

The computer algorithm is then employed at hourly resolution (r = 0.50), and for regions below and above 1 km separately. Correlations with snow are largest when laminations occur simultaneously above and below 1 km (r = 0.44), followed by laminations exclusively above 1 km (r = 0.28); there is only a very small correlation for laminations exclusively below 1 km (r = 0.09). Therefore, cloud laminations above 1 km altitude are key lidar indicators of the production of precipitating snow detected at ground level.

1 Introduction

Mixed-phase Arctic clouds contain inhomogeneities, including the separate spatial distributions of water and ice (Whale, 2018; Verlinde et al., 2007, 2013; von Terzi et al., 2022). Traditional ice particle production processes such as depositional growth via the Wegener–Bergeron–Findeisen (WBF) process are favoured in the dendritic growth layer between -12 and -20 °C (Korolev et al., 2007). Elsewhere in the cloud, secondary ice processes, including splintering via the Hallet-Mossop process (Field et al., 2017; Hallett and Mossop, 1974; von Terzi et al., 2022), have long been considered to be effective within narrow temperature ranges. Collisional breakup (Hobbs and Rangno, 1998, 1990, 1985) has also been implicated in discrepancies between the number of ice-nucleating particles and ice particle number concentration over a broader range of temperatures (von Terzi et al., 2022; Kanji et al., 2017; Georgakaki et al., 2022; Phillips et al., 2018). As more than 70 percent of global precipitation is

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

generated through ice phase processes (Heymsfield et al. (2020); Field and Heymsfield (2015); Mülmenstädt et al. (2015), von Terzi et al. (2022)), the distributions of ice particles and water droplets within the cloud are particularly relevant to precipitation-production processes occurring within these clouds. Therefore, accurately representing cloud inhomogeneities has the potential to improve weather forecasting (Barrett et al., 2017a, b; Vassel et al., 2019).

Previously, we have shown that laminated clouds, consistent with an interpretation of alternating layers of ice and water, on the order of 7.5 to 30 metres thick, are common at Eureka (McCullough et al., 2019). Alternating layers of randomly-oriented and/or preferentially-oriented ice particles are alternative reasonable hypotheses. The laminated clouds were identified by visual inspection of 532 nm attenuated backscatter plots. Detection criteria included: three or more bright layers, approximately horizontal, stacked within 150 m of each other within a cloud for at least 30 minutes. We have further shown that days containing these layered clouds strongly correlate with ground-level precipitating snow conditions (McCullough et al., 2021). Nonlaminated days were subclassified as nonlaminated cloudy, nonlaminated clear, or omitted from the study due to insufficient high-quality measurement coverage of the time and space in the daily array. Because the manual visual process can be subjective and definitely slow, we now expand upon this analysis and develop an objective, automated computer method to replace the visual method.

Visual inspection of lidar profiles and scenes is a reliable and trusted interpretation method that has been used for many applications in the atmospheric science community (e.g. Shupe (2007); Sedlar et al. (2021); Vassel et al. (2019); Schäfer et al. (2020)). Schäfer et al. (2020) recently used visual inspection to categorize noctilucent clouds into subtypes based on cloud layer type and whether they exhibit substructure. While reliable and superior for individual images, visual inspection is too time-intensive for large datasets. Therefore, a computer approach is highly desirable; however, we were unable to find a lidar computer algorithm that can search for fine layers within clouds without other supporting data. Therefore, here we construct and demonstrate a computer-based solution that can be applied to CRL and other lidars worldwide; the laminated cloud features can be automatically detected in attenuated backscatter data.

Section 2 describes the measurement datasets. Section 3 describes factors considered when selecting the computer lamination detection program algorithm. Section 4 describes the algorithm itself and explains how non-laminated days are subclassified. We use the residuals method, previously used in gravity wave analyses (Sica and Russell, 1999; Duck et al., 1998). A "background state" containing variations not of interest is subtracted from the original data leaving only variations of interest. Section 5 is the validation study for the algorithm. Section 6 increases the study time resolution from daily to hourly. This leverages the computer method's processing speeds and extends the study of the laminations themselves without altering their identification criteria. This higher time resolution matches the nearly hourly availability of ECCC weather reports used for the snow correlation portion of the study. Section 7 extends the study by classifying two altitude regions separately: 250 m - 1000 m and 1-5 km, investigating whether there is a difference in correlation between precipitating snow and laminated clouds across these two altitude ranges. Section 8 discusses some hypotheses regarding the physical interpretation of the laminations. Section 9 summarizes and presents the conclusions of this paper.

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

2 Measurements used

2.1 CRL lidar

The ground-based, remotely-operated, zenith-pointing CRL lidar (Canadian Network for the Detection of Atmospheric Change (CANDAC) Rayleigh-Mie-Raman lidar) has operated since 2007 at the Polar Environment Atmospheric Research Laboratory (PEARL; Eureka, Nunavut in the Canadian High Arctic, 79.6° N, 85.6° W). The transmitted wavelengths are linearly-polarized at 355 nm and 532 nm. Backscattered light is received by a 1-m Dal-Kirkham telescope and directed into an 8-channel (7 wavelengths, with linear depolarization capability at 532 nm) photomultiplier tube polychromator (Nott et al., 2012; McCullough, 2015; McCullough et al., 2017).

The dataset used in this paper is identical to that used in McCullough et al. (2021), which uses only the 532 nm wavelength because this is the most commonly measured quantity for the CANDAC Rayleigh-Mie-Raman Lidar (CRL) during our 2016-2019 study period. CRL's other measurement channels, including 355 nm backscatter and 532 nm depolarization, also show evidence of the laminations, as shown in McCullough et al. (2019). However, these products are available on fewer measurement days, and generally have lower signal-to-noise ratios than the 532 nm channel.

Standard CRL processing routines were applied to the 532 nm data as described in McCullough (2015) and McCullough et al. (2017): saturation and deadtime correction, dark counts removal, background subtraction, and gluing of the photon-counting and analogue signals into a combined vertical profile for each minute. The standard-processed lidar signals were multiplied by the square the returns' altitude, and the base-10 logarithm of the result was taken. The result is log(range-scaled signal) in units of log(m^2MHz), or the logarithm of the attenuated backscatter. This product initiates the algorithm described in Section 4.

We used the attenuated backscatter product as the starting point for the computer algorithm because it provides excellent lamination visibility by eye and requires the least data (pre-)processing, thereby minimizing instrument and undesirable atmospheric effects.

The visual sky scene classification method treats colour plots of lidar data as images. Hence, we refer to time-altitude arrays of lidar data as images, and individual array elements as pixels.

2.2 Eureka Weather Station meteorological reports

Meteorological observations of surface weather at the Environment and Climate Change Canada (ECCC) Eureka Weather Station are reported hourly, 22 to 24 hours per day. Weather observers make these observations at the Eureka Weather Station instrument compound, located 180 m to the southeast of the CRL lidar.

The standard ECCC weather observation types were grouped into six categories as follows: (1) Precipitating Snow (including snow, moderate snow, snow showers, snow grains, snow pellets), (2) Blowing Snow, (3) Rain (including rain, rain showers, moderate rain, drizzle, freezing drizzle, freezing rain), (4) Ice Crystals, (5) Fog, (6) Freezing Fog. We excluded observations between 2016 and 2019 with fewer than two reported occurrences during the study period and in the following categories: blowing sand, dust, haze, ice pellets, ice pellet showers, smoke, no value, N/A., clear, mainly clear, cloudy, and mostly cloudy.

Arrays were created at daily and hourly resolutions to indicate the presence or absence of each weather type.

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

3 Factors influencing the choice of a computer algorithm

The input to our algorithm is a set of daily log(attenuated backscatter) arrays, in which each array element, or pixel, corresponds to approximately 1 minute by 7.5 m in altitude. The classification of these time-altitude arrays is computationally analogous to an image classification problem. 1-D lidar profile classification methods, such as those used by Farhani et al. (2021), are not suitable because they do not account for temporal correlations of the laminated signals. Developing computer methods to identify image patterns is an active area of research, but human vision remains the gold standard for identifying features in complicated images. It is used for training sophisticated computer methods (e.g., Pinto et al. (2008), Gulshan et al. (2016), Baduge et al. (2022), von Ahn and Dabbish (2004), Jafarinaimi (2012), Heinle et al. (2010)). Lu and Weng (2007) provide a review of computer classification approaches as applied to satellite images of ground cover, which also applies to lidar sky scenes. The complexity of the log(attenuated bacscatter) arrays and the variation in the appearance of layers and laminated regions pose a challenge for applying existing image classification algorithms. Hence, we have developed our own algorithm that attempts to deal with several well-known difficulties:

(1) Invariant object recognition - the ability to recognize that an object or feature is the "same thing" even if its appearance changes under different circumstances (Karimi-Rouzbahani et al., 2017b). The visual appearance of our feature of interest can vary widely based on (a) Lamination geometry: Layers can be stacked from three to many tens in number, can rise and fall slightly in altitude, can split into two, and can merge. Hence, pattern-matching approaches (e.g. Lahiri et al. (2011), Ridder et al. (2017)) are not appropriate. Schäfer et al. (2020) reached the same conclusion. They resorted to visual classification to identify small-scale structures (100 m thick, less than 30 minutes in duration, at 10-minute intervals) in noctilucent clouds, which are attributed to gravity waves that vary the altitude of cloud layers. These layers are visually similar to the laminations reported here, and their host clouds share the characteristic of having a wide variety of individual patterns and subpatterns in the layering structure. (b) Illumination: Lidar images contain a large dynamic range, and the features in them are not consistently illuminated. Clouds are much brighter than a clear sky and usually brighter than aerosol layers. The brightness of any lamination is influenced by instrumental factors (laser power, etc.), attenuation by underlying atmospheric features, as well as the scattering characteristics of the layer of interest itself. Therefore, the relative brightness of a layer with respect to its local background is inconsistent, and classic image segmentation thresholding algorithms will fail under these conditions. (c) Partial occlusion: Lidar images include columns with low or no laser power reaching the lamination altitude. Therefore, gaps appear in the array's laminations. Accounting for such partially-occluded targets is a challenge for computer vision algorithms as compared to human vision (Karimi-Rouzbahani et al. (2017a, b) and references therein, including Pinto et al. (2008), Tarr and Bülthoff (1998),Tjan et al. (1995), Tarr et al. (1998)).

(2) Segmentation - the ability to tell that two items are separate, even if they are close together in the image. Given that the laminations can be very close together in altitude, merge into a thicker layer, or split into two thinner layers, segmentation is an essential capability for identifying the number of laminations within a vertical distance in lidar images. As the laminations of interest are at the same vertical spatial scales as the noise in the images - the resolution limit of the lidar data, 7.5 m - segmentation of targets in clutter becomes a significant computational challenge (Ghodrati et al., 2014).

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

(3) Context recognition and the ability to ignore irrelevant features - Current computer image classification methods assume that there is no competing non-target information in the picture (Paymode and Malode, 2022). Each lidar image can contain any number of irrelevant atmospheric features, both outside and inside host clouds, including other clouds, precipitation fall streaks, and aerosols. While complicated backgrounds provide important context that humans use as a diagnostic feature in complex discrimination tasks (Praß et al. (2013), Nielsen et al. (2006)), as noted in Karimi-Rouzbahani et al. (2017b), they pose a significant challenge for computer algorithms. Even the identification of clouds as separate from aerosol features in lidar plots is frequently accomplished by analyzing the feature morphology and behaviour by eye, rather than the signal threshold, unless a multi-instrument algorithm such as those operated by Cloudnet (Illingworth et al., 2007) or MPLNET (Welton et al., 2001) is employed (Radenz et al., 2021; Illingworth et al., 2007; Bühl et al., 2016; Kalesse-Los et al., 2022). Heinle et al. (2010) developed a k-nearest-neighbour statistical cloud type classification algorithm for all-sky camera images. However, it is not robust to images containing irrelevant features (e.g. raindrops, sunlight, or multiple cloud types), and it is unsuitable for our project without significant data modification and pre-processing. Additionally, our measurement dataset may be too small to allow sufficient training images (Bianconi et al., 2021).

(4) Computational time - Supervised and unsupervised machine learning methods are frequently computationally intensive. A wavelet-like approach similar to Reichert et al. (2023), who used a combination of Gabor filters (e.g. Jann (2017)) for the laminations, is too computationally intensive. In the future, it may be possible to incorporate a wavelet approach into the algorithm.

A final consideration in our algorithm choice was our desire to understand which image features were used in their classification, so that we can explore any physical meteorological implications that may arise. Deep neural networks, for example, implicitly optimize their selection of image features for classification. These may include features that are unknown to or unrecognized by researchers (Bianconi et al., 2021). Determining which features a neural network has used in a successful classification is an area of active research (Gulshan et al., 2016; Baduge et al., 2022).

Given these considerations, an explicit computer procedure, without hidden layers, seemed appropriate for this laminations project, and we have found a successful method using conventional programming techniques that runs on a consumer-grade laptop.

4 Computer Algorithm

4.1 Identification of laminated cloud regions

The input data for the algorithm is a daily array, A1, of the base-10 logarithm of the range-scaled attenuated backscatter from CRL, spanning 250-5000 m in altitude, as shown in Fig. 1 (a). They have already been pre-processed with standard CRL correction procedures to account for deadtime corrections and background subtraction, as described in McCullough et al. (2019). The altitude resolution is 7.5 m. Each profile is integrated for 1 minute before being recorded, which takes about 5.4 seconds per file. Therefore, there are 1.091 minutes per profile (55 files per hour), and each pixel represents 1.091 minutes x 7.5 m altitude. The measurements are not co-averaged in time or altitude.

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

A cloud mask (black contour in Fig. 1 (a)) is determined by a threshold value that varies between 7.25 and 8.25 in units of log(m^{2}MHz) during the four-year study period. On the date shown, this threshold was 8.25. Visual inspection of cloud mask contours superimposed on a sample subset of attenuated backscatter colour plots confirmed that the threshold's value is appropriate. Future development of an automated cloud mask would improve the algorithm. Finally, there will never be a perfect threshold, as optically thin clouds and relatively optically thick aerosols can produce identical attenuated backscatter values.

$A1$ is filtered with a moving median in altitude with a window frame length of 25 pixels (187.5 m), and filtered separately in time with a frame length of 51 pixels (55.6 minutes). The resulting filtered arrays are averaged to create $A1_{smoothed}$. In this study, $A1_{smoothed}$ is considered the background state of the local sky. It includes information such as cloud presence and variations within the cloud, which are spatially larger than the laminations we seek.

The first residuals array $R1$ = $A1$ - $A1_{smoothed}$ is shown in Fig. 1(b). The quasi-horizontal stripes, the focus of this investigation, are more clearly visible to the naked eye than in Fig. 1(a). However, vertical features, such as fall streaks, and horizontal features, such as cloud edges, remain. The fall streaks, in particular, are bright and visually interrupt the stripes at the colour scale shown; however, the stripes can be seen to continue through the fall streaks if the colour scale is changed.

$R1$ is now smoothed using a moving median filter with a frame length of 15 altitude pixels and 1 time pixel, producing $R1_{smoothed}$. Fig. 1(c) shows the second residuals array $R2$ = $R1$ - $R1_{smoothed}$. The fall streaks are suppressed, and the laminations are clear.

To allow for simpler image processing techniques, the $R2$ array is discretized to produce $R2D$ using thresholds which are determined relative to σ, the standard deviation of in-cloud $R2$ values, excluding NaNs, as follows:

$$R2D = \begin{cases} -1 & R2 \leq -0.2\sigma \\ 0 & -0.2\sigma < R2 < 0.2\sigma \\ 1 & R2 \geq 0.2\sigma \end{cases}$$

These thresholds determine which of the scene's pixels may belong to the stripes we are seeking to identify laminated features. The $R2D$ = -1 pixels have signal values significantly below the mean cloud values and may be part of "negative stripes". Similarly, $R2D$ = 1 pixels may be part of "positive stripes".

The threshold σ can be calculated from $R2$ using either all pixels or only in clouds. Most days, the thresholds are similar, but on an otherwise clear day with a relatively small, bright cloud containing laminations, they differ. The in-cloud-only σ values are more appropriate for lamination detection, and a threshold $\pm 0.2\sigma$ provides a reasonable tradeoff between rejecting noise and including real stripes.

We use a connected-component analysis to identify $R2D = \pm 1$ pixels which correspond to positive or negative stripes rather than noise. Connected components are defined as contiguous groups of adjacent pixels sharing the same value. Any connected component of pixels with value $R2D = 0$ that is entirely enclosed by a region of $R2D$ which is entirely $+1$ or -1 is reassigned to the value of the surrounding region, effectively filling "holes" within otherwise continuous stripes. Furthermore, all connected components with a duration shorter than 10 pixels (i.e., approximately 10 minutes) are excluded from further

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

analysis. The final result is a classification array $S1$ (see Fig. 1(d)), in which positive and negative stripes are assigned values of 1 and -1, respectively, and all remaining pixels are set to 0.

To suppress spurious detections, we identify connectivity solely on the shared edges (i.e., top, bottom, left, or right), excluding diagonal (corner) adjacency. This choice affects the sensitivity of the method: some fine-scale features that are only one pixel in vertical extent and exhibit drift in altitude over time, appearing as diagonals in the $R2D$ array, are excluded, even if they are visually apparent.

A minimum of three stripes within 150 metres of altitude, corresponding to 20 vertical pixels, as specified by the lamination identification criteria given by McCullough et al. (2019). Each pixel at time index t and altitude index z, $S1(t, z)$, is tested to determine whether the lower and upper edges of at least three stripes of the same sign are detectable within 22 pixels, including and immediately above the pixel of interest. If so, all pixels $S1(t, z)$ to $S1(t, z + 21)$ are labelled as "laminated pixels" with a value of 1 in the laminated pixel array $L1$, Fig. 1(e), and non-laminated pixels have a value of 0. Twenty-two, and not 20, pixels are included in this range so that the bottom *edge* of the bottom stripe and the top *edge* of the top stripe may be detected within the 20 pixel range.

Morphological closing is performed on $L1$ to produce $L2$. A rectangular structural element with structural parameters (2,4) is first used to dilate, and then erode, the array, filling in small holes in regions with densely laminated pixels. Unlike smoothing procedures, this tends to preserve the onset and end time of each group of laminated pixels.

Following morphological closure, the connected components of $L2$ are determined, again requiring connections to be along pixel edges. These components are tested against a 30-minute duration criterion for laminated regions, and only those components exceeding 30 minutes (27 pixels) are retained in array L3, with all other pixels set to 0.

The cloud mask is now used to mask non-cloud pixels (0) and produce $L4$. The duration check is then repeated to retain only those components of $L4$ that exceed 30 minutes (27 pixels) as the new array $L5$, shown in 1 (f).

The laminated regions identified in $L5$ are the primary desired output of the automated layer detection computer program: Laminated cloud regions that meet all the requirements set out in McCullough et al. (2019) for stripe number, vertical extent, duration, and location within a cloud. As noted above, this computer method will exclude the narrowest laminations visible if the contributing stripes are not perfectly level over the entire 10-minute test period for individual stripes. Therefore, a few laminations are missed at the finest scales.

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

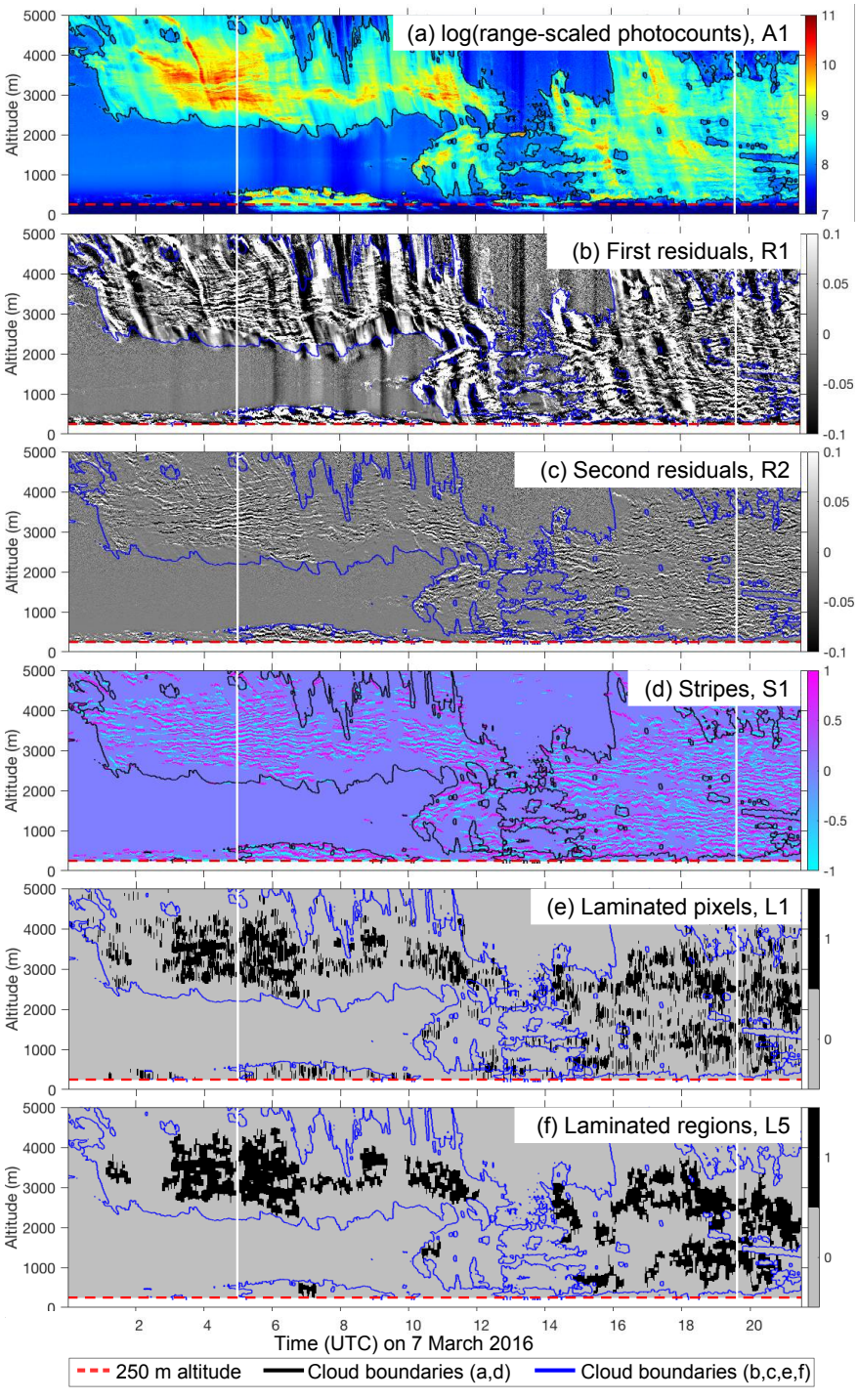


Figure 1. Procedure for identifying laminations on 7 March 2016. (a) Input array A1, photocounts in units of log(m^2MHz), (b) First residuals array R1, (c) Second residuals array R2, (d) Positive (magenta) and negative (cyan) stripes in array S1, (e) Laminated pixels array L1, (f) Final array, laminated regions after morphological closing, cloud mask, and duration thresholding, L5.

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

4.2 Classification of lidar sky scenes

Using array $L5$ for all days in the dataset, we classify the status of each day's sky scene laminations.

As in McCullough et al. (2019), "Laminated" days are identified first. Any day containing at least one qualifying laminated region in $L5$ is classified as Laminated, corresponding to Class 2 in McCullough et al. (2019).

Days in the CRL dataset showing no laminations do so for one of two reasons: either there were no laminations in the sky, or CRL measurements were insufficient to detect them. The sub-classification of non-laminated days uses the $A1$ and cloud mask arrays.

"Incomplete" days, with no CRL data for some or all hours of the day, correspond to Classes 1, 4, and 5, in McCullough et al. (2019), which are combined for the present paper. The Incomplete flag is used to check for missing or discontinuous measurements that exceed the lamination patch duration threshold of 30 minutes. In principle, a qualifying set of laminations could have been present in the portion of the sky not imaged by the lidar during the time the lidar was not measuring.

"Obscured" days have sufficient measurement profiles, but have a portion of the sky scene that is uninterpretable due to low signal for a duration and vertical extent exceeding the lamination patch duration threshold. This corresponds to Class 3 as described by McCullough et al. (2019). It includes days for which lower-altitude atmospheric features fully attenuate returns from higher altitudes, as well as days with poor laser power, which are too noisy at the top of the profiles. This is implemented by allowing no more than 25% of the pixels in the top 150 m of measured columns to have NaN values. The visual method lacked strict numerical criteria for this test.

The remaining measured days are classified as either "Nonlaminated Cloudy" if the cloud mask shows at least one cloud exceeding the threshold lamination patch size (even if the day is not predominantly cloudy), or "Nonlaminated Clear" for all remaining days not meeting that threshold (even if they contain a little cloud, or several little clouds, each not exceeding the threshold size).

5 Validation of the Algorithm

As discussed above, the algorithm correctly identifies the majority of the Laminated days without falsely identifying an undue number of days as Laminated. Two metrics are used to expand this result: (1) the ability to replicate the sky scene classifications that were published in McCullough et al. (2019, 2021) with the visual method and (2) verifying the overall conclusions drawn from correlation studies between lidar sky scene and the ECCC surface weather from those papers.

5.1 Sky Scene Classification Validation

A confusion matrix, Fig. 2(a), is used to understand the algorithm's output and to identify which specific types of sky scenes were and were not well classified. The visual classifications are the known values, and the computer classifications are the predicted values.

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

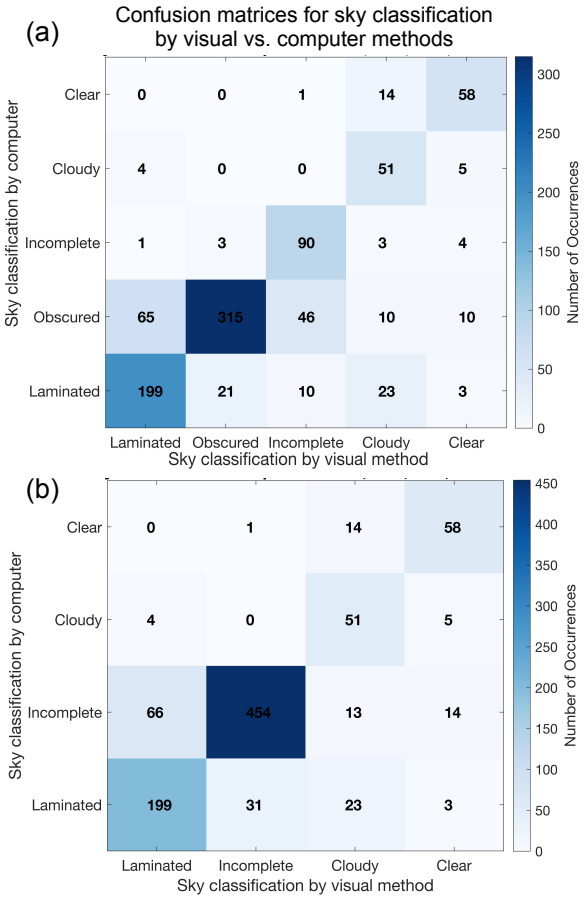


Figure 2. A confusion matrix comparing computer and visual sky classification results at daily resolution. (a) has all the classes and (b) combines Obscured and Incomplete classes.

Some non-zero off-diagonal elements, which could confuse the Obscured and Incomplete classes, were anticipated with the computer algorithm. For example, considering a particular clear sky day with both some hours of missing data (contributing to an Incomplete classification) and low laser power (contributing to an Obscured classification), a visual observer would have classified the day as Incomplete as there were clearly no lower altitude clouds which might be responsible for the low signal returns near 5 km. However, the computer would first test for the number of NaN pixels at high altitude, and would immediately classify the day as Obscured, regardless of the cause. Therefore, we combine these classes when formally evaluating the algorithm's performance, as shown in Fig. 2(b).

Table 1 details the classification results and some commonly used diagnostic values. Using the equations from Chicco et al. (2021) and Saito and Rehmsmeier (2015), as shown in the table row labels, we calculate diagnostic values for each class.

For the first four columns of Table 1, we used a classwise one-versus-all method based on Fig. 2(b). We assess for each class separately: True positive, TP, the number of scenes correctly identified as belonging to that class; True negative, TN,

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table 1. Diagnostic values comparing the computer to the visual classification results using a one-versus-all method. True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) values are derived from the confusion matrix shown in Fig. 2(b) and are used to calculate all other quantities according to the equations in the row labels. The first four columns of values are provided for each sky classification. The fifth column, Macro, and the sixth column, Micro, are combined metrics for all classes.

	Laminated	Incomplete	Cloudy	Clear	Macro	Micro
TP	199	454	51	58	-	762
TN	610	357	826	841	-	2634
FP	57	93	9	15	-	174
FN	70	32	50	22	-	174
Support $TP + FN$	269	486	101	80	936	3744
Proportion $\frac{TP + FN}{TP + TN + FP + FN}$	0.287	0.519	0.108	0.0855	1	0.25
Precision $\frac{TP}{TP + FP}$	0.777	0.830	0.850	0.795	0.813	0.814
Recall $\frac{TP}{TP + FN}$	0.740	0.934	0.505	0.725	0.726	0.814
Specificity $\frac{TN}{TN + FP}$	0.915	0.793	0.989	0.982	0.920	0.938
Accuracy $\frac{TP + TN}{TP + TN + FP + FN}$	0.864	0.866	0.937	0.960	0.907	0.907
F1-score $\frac{2TP}{2TP + FP + FN}$	0.758	0.879	0.634	0.758	0.757	0.814
Pearson's r (or MCC) $\frac{(TP)(TN) - (FP)(FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$	0.664	0.738	0.626	0.738	0.691	0.752

the number of scenes correctly identified as not belonging to that class; False positive, FP, the number of scenes incorrectly labelled as belonging to that class; False negative, FN, the number of scenes incorrectly labelled as not belonging to that class. There are a total of n = 936 samples (measurement days classified). The fifth column, Macro, is a combined metric for all classes. Macro averaging gives each sky class equal weight by taking the unweighted mean of all individual values already calculated in the columns to the left; however, it does not apply to TP, TN, FP, and FN. The sixth column, Micro, gives equal consideration to each sample in the averaging process. New Micro TP, TN, FP, and FN values are calculated as the sums of the individual class values and are used in all subsequent values in the Micro column of the table.

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.



5.1.1 Laminated sky scene class results

Of all the assessment metrics included, the most important for our application are those related to the algorithm's ability to identify and classify Laminated days correctly. The visual classification identified 269 days as Laminated and 667 days as not laminated (Incomplete, Obscured, Clear, and Cloudy). The computer algorithm correctly classified 199 Laminated days (true positive, or TP), and 610 non-laminated days (true negative, or TN). It incorrectly identified 127 days, including 57 false positives, and 70 false negatives, which is discussed in Sec. 5.1.2.

"Support" refers to the number of samples identified visually as belonging to a particular class, and "support proportion" expresses the support as a fraction of the total samples. The differences in support (and support proportion) across classes indicate that our data is unbalanced. That is, we are testing unequal numbers of each class. For instance, the number of Laminated samples is only 30% of the total. This must be considered when interpreting the diagnostic values in the final three rows and two columns of the table.

Precision, or positive prediction value, indicates the classifier's success at avoiding false positives. If precision is high, then most predicted positives are truly positive. Precision = 0.78 for the Laminated class.

Recall, or sensitivity, or true positive rate, is the fraction of the positive samples correctly identified by the classifier. High recall implies that the classifier misses only a few positive cases. Recall = 0.74 for the Laminated class.

Specificity, or true negative rate, is the fraction of negative samples correctly identified as negative. Specificity = 0.92 for the Laminated class.

Accuracy assesses how well, overall, days were correctly classified as belonging to, or not belonging to, that particular class. The Laminated class has an accuracy = 0.86, indicating a successful classification. This metric is less helpful for imbalanced data (such as ours), as high accuracy can result from a poor classification scheme, e.g. if all observations fall into the majority class (Chicco and Jurman, 2020; Chicco et al., 2021).

F1-Score is based on the harmonic mean of Precision and Recall. F1-Scores also include a parameter, β, that weights the relative importance of precision and recall in the assessment. As beta increases, recall is prioritized over precision. By setting $\beta = 1$, as we do in Table 1, the precision and recall are weighted equally. F1-Score $= (1+\beta^2)(2)(\text{Precision})(\text{Recall})((\beta^2)(\text{Precision})+(\text{Recall})))^{-1}$.

F1-Scores are appropriate for imbalanced datasets, and are most useful when the positive class is the priority, as is the case for the Laminated class. One drawback to the F1-Score is that its value changes if the positive and negative classes are reversed. F1-Score also does not account for the number of TN samples, which can be an important factor in evaluating a classification algorithm. A further drawback to F1-Scores is that they can be high even if only one of TP or TN is large and the other is small. The F1-Score for the Laminated class is 0.76, indicating a successful classification.

Pearson's r is the Pearson product-moment correlation coefficient between the actual and predicted values. It is also known as the Matthews Correlation Coefficient (MCC), and is often encountered in machine learning literature. It ranges from -1 to 1. All four elements of the confusion matrix must have successful outcomes for Pearson's r to have a high value. It is only

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

high if both the majority of positive and the majority of negative samples are correctly predicted (Chicco et al., 2021). For the Laminated class, r = 0.66, which indicates a successful classification.

A significant advantage of r as an assessment metric is that it highlights problems with a classifier for datasets with any degree of support imbalance: Positively imbalanced (more positive samples than negative ones) classes, balanced classes (equal number of positive and negative samples), and negatively imbalanced classes (more negative samples than positive ones, as is the case for the laminated class). For any degree of support imbalance, the degree of bias in the output value of r is smaller than the bias induced in both accuracy and F1-score (Luque et al., 2019).

If r is close to 1, so will the F1-Score and accuracy; however, the converse is not always true. Therefore, while F1-Score and accuracy can provide more optimistic results, r may be more suited to identifying deficiencies in the classification scheme (Chicco and Jurman, 2020). In our case, the r value (0.66) for the Laminated class is, as expected, lower than both the accuracy (0.86) and F1-Score (0.76), but still a high value. Therefore, we are confident that our computer algorithm can detect laminations, but we should also dissect the mismatches.

5.1.2 Laminated misclassifications

For the 70 false negatives (FN, missed detections of laminations), we examine the plots for each such day, and the reasons for the false negatives can be mostly grouped into: (1) Thin lamination days. The computer can only detect the thinnest 1-pixel-thick laminations (7.5 m) if all three required stripes remain at the exact altitudes, not even varying by 1 pixel, for the entire required 10-pixel stripe duration. The majority tend to drift in altitude, and the computer algorithm does not detect them. (2) Days having many diagonal stripes, which happen to include at least one time bin where the stripe is only 1 pixel thick have the same limitations in the code as just described. (3) Large variations in brightness between different clouds during the course of the day can cause the daily in-cloud σ value not to be locally representative for any particular cloud, thus making stripes impossible to distinguish. (4) Days with high aerosol loads, which make the cloud thresholds for the cloud mask inappropriate, meaning that parts of clouds are excluded (cutting off legitimate stripes) and/or aerosol regions are interpreted as clouds (thus changing the σ value used for thresholding positive and negative stripes). (5) Days for which the overall brightness of the lamination host cloud varied considerably on short timescales. In this situation, the $R1_{smoothed}$ array fails to capture sufficient temporal variation in the local background that hosts the stripes. Hence, although a person can visually see that the stripe pixels are markedly brighter than the pixels above and below, they are not sufficiently bright in comparison to the host cloud as a whole to meet the stripe threshold. This creates gaps in the stripes, preventing them from meeting the duration threshold.

The 57 false positives (FP, in which the computer detected laminations but the human did not) can nearly all be understood by a single atmospheric situation: these “false” detections happen most often at low altitudes, partially inside the lidar’s overlap region, where the receiver does not fully capture some backscatter from the laser beam. Because the array $A1$ is not overlap corrected, a human would adjust the colour scale on each plot to account for the clouds near the ground, which have lower signal levels, but the computer cannot do that. When we change the colour scale for the FP days, we can see the laminations by eye which the computer has identified. Although these low-altitude laminations match the mathematical criteria, the original visual observers would not have considered laminations in this region to be within a classic “cloud”, as noted by McCullough

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

et al. (2019, 2021). The computer criteria correctly identify the stripes, but they are visually and morphologically different from the laminations identified in clouds at higher altitudes. These lower-altitude stripes tend to be thicker and less variable in altitude over time than the laminations in clouds aloft. Additionally, there are usually only 3-4 such stripes in the region below 1 km. In contrast, stripes that make up the laminations aloft tend to occur in sets that are much more numerous. This group of FP would be reduced if the algorithm excluded such features.

A final group of FP includes days with significant in-cloud signal variations. There, the laminations identified by the computer in the $L5$ array were not identified by visual observers in the $A1$ arrays. However, the same visual observers can detect the laminations clearly in both the $R1$ and $L5$ arrays (which were not used in McCullough et al. (2019, 2021)). Therefore, we consider these "false" positives by the computer to reflect the success of the algorithm - the computer is exceeding the visual method used in previous work.

On most days, the laminations are correctly identified as present or absent, and we can adequately explain the reasons for non-matching classifications. We conclude that the computer algorithm is valid for identifying laminations in lidar data. We also note that this algorithm is inferior to human vision assessment for the thinnest laminations (7.5 m thick), and superior in conditions of highly variable cloud backscatter values.

5.1.3 Non-Laminated Days

To use this dataset for correlation studies with weather, we require the non-laminated days be correctly subclassified. We evaluate the non-laminated classes individually in columns 2, 3, and 4 of Table 1 in the same method described above for the Laminated class. All individual correlations have values $r > 0.62$, signifying that these values are valid.

5.1.4 Results of the classification scheme

The final two columns of Table 1 present the combined metrics for the classification scheme as a whole, including all classes. The Micro and Macro scores are defined, e.g., in Hinojosa Lee et al. (2024). The Macro Average column averages the individual class scores from the first four columns, and the macro score gives equal weight to each class. Macro-Accuracy = 0.91, Macro-F1-Score = 0.76, Macro-MCC = 0.70. The Micro Average row uses TP, TN, FP, and FN as the sum of the individual class values. Then, all subsequent values in the column are calculated using those cumulative values. Therefore, the micro average gives equal weight to each of the 936 samples in our dataset. Micro-Accuracy = 0.91, Micro-F1-Score = 0.82, and Micro-MCC = 0.75; all have high scores. Given these results, we assert that our computer algorithm is valid for identifying laminations in lidar data and that our classification scheme is valid.

5.2 Sky-Weather correlation validation

Using the computer classified daily results, having demonstrated their validity, we now test the correlation between lidar sky scenes (Laminated, Nonlaminated Cloudy, and Nonlaminated Clear) and ECCC reported weather (Precipitating Snow, Rain, Blowing Snow, Ice Crystals), as in McCullough et al. (2019, 2021). The correlation test includes 389 days, as only days

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

classified as Laminated, Nonlaminated Cloudy and Nonlaminated Clear are considered. Incomplete and Obscured days are omitted.

The correlation results are presented as colour-coded correlograms in Fig 3. The colours correspond to the value, r, of the Pearson product-moment correlation coefficient (the covariance of two random variables divided by the product of the standard deviation of each variable X and Y, as in: r = $cov(X,Y)/\sigma_X * \sigma_Y$). We continue to follow the convention suggested by Cohen (1988) for qualitative descriptors: r$\leq\sim 0.1$ is small, r$\geq\sim 0.5$ is large, and the middle range is medium. Working with a 95% (2σ) confidence interval, we consider r values with p ≤ 0.05 to be significant. Insignificant values are identified by a large 'X' superimposed on the affected correlogram cells (there are no insignificant values in 3, but some do exist in the 6 correlogram). Each cell of the confusion matrix includes the r value, and a small table: TP at the upper left, TN at the lower right, FP at the upper right, and FN at the lower left. Other metrics, such as accuracy and F1, can be directly calculated from the correlogram. The correlation coefficient r is the appropriate choice for a single metric in our data situation, given the negatively imbalanced dataset, as discussed above and will be used for further discussion.

The computer-classified dataset aligns with the results from the visually classified dataset presented in McCullough et al. (2021). The strongest amplitude correlations are those related to precipitating snow. There is a strong positive correlation between precipitating snow and Laminated days (r = 0.62, compared to r = 0.63 from the visual method). There is a medium strength negative correlation between precipitating snow and Nonlaminated Cloudy days (r = -0.36, compared to r = -0.37 from the visual method) and a slightly stronger medium strength negative correlation between precipitation snow and Nonlaminated Clear days (r = -0.41; compared to r = -0.40 from the visual method). No other correlation is larger than r = $\pm$ 0.24 in amplitude (compared to r = $\pm$0.23 from the visual method).

Given that the correlations between lidar sky scene classification and ECCC-reported weather are virtually identical whether using the visual or the computer method, this gives further confidence that the algorithm is valid for our purposes.

6 Increase to Hourly Resolution

The ECCC weather reports are produced hourly. Replicating the visual method study at this resolution is too time-intensive, however, with the computer algorithm, the hourly analysis is possible.

6.1 Method

We follow the above daily procedure through to identifying laminated regions in array $L5$ (see Fig. 1g). The day is then divided into hourly bins, each labelled by the end time of the corresponding full hour, as in the ECCC reports.

Because we know that the computer algorithm is conservative in identifying the thinnest laminations, and the starts and ends of laminated regions are often missed, we consider even a small amount of laminated cloud identified within a particular hour to be meteorologically relevant to that hour. Thus, an hour of interest is classified as Laminated if a laminated region (which by our criteria exceeds 30 minutes duration) persists for at least 10% of that minimum duration (i.e., 3 minutes out of the 30-minute threshold) within that particular hour. The balance of the laminated region then exists in the hour preceding or

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

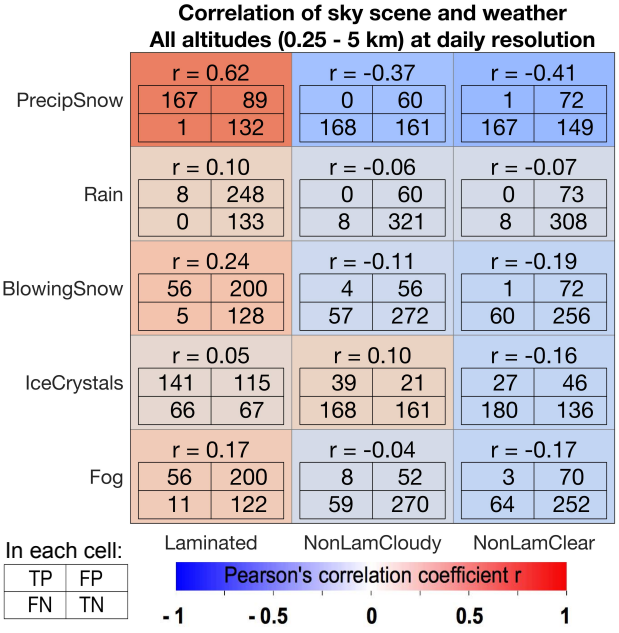


Figure 3. Daily resolution: Correlogram showing the correlation between computer classified sky scenes (laminated, non-laminated cloudy, and non-laminated clear days) and ECCC weather reported at ground level. Each cell is coloured according to its correlation coefficient value r. The one-versus-all confusion matrix in each cell takes ground-based weather as the "truth", with TP, TN, FP, and FN values arranged as in the legend at lower left.

following the hour of interest. For example, consider two adjacent hourly bins, Hour 1 and Hour 2: a laminated region that begins in Hour 1 and continues into Hour 2 must meet the 3-minute threshold in each hour independently for both hours to be classified as Laminated.

Computers can identify cloud locations more easily than laminations, and the inclusion criteria are stricter. For non-laminated hours, the cloudiness classification requires that at least 40% of the minimum 30-minute duration for a cloud (i.e., 12 minutes) be present entirely within the hour of interest for it to be considered Nonlaminated Cloudy. If the threshold is set too low, many hours which are predominantly clear would be classified as "cloudy". If the threshold is set too high, a cloud that spans two hour bins may be rejected.

Obscured and Incomplete criteria are assessed for each hour independently, with the full duration criteria required within an hour of interest for that hour to receive one of these classifications.

The result is shown in Figure 4 and includes 10825 hours of sky scene and weather observations. Some points about the result are: 1. Hours are omitted if they are classified as Incomplete or Obscured, or if there is no coincident ECCC weather report for that hour. 2.Data are omitted in the daily analysis if the day as a whole was classified as Incomplete or Obscured, or if the laminations occurred during an hour with no reported ECCC weather. 3. Lamination and precipitating snow events which were separated by many hours can give a positive correlation at daily resolution, despite not happening simultaneously. This is

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

unlikely at hourly resolution, as hour-by-hour coincidence is required. 4. The delay between the occurence of a precipitation-generating sky condition and the time when the weather is reported at the surface, may would not show positive correlations in our test. A dataset without gaps would allow us to identify such lag behaviour.

6.2 Results

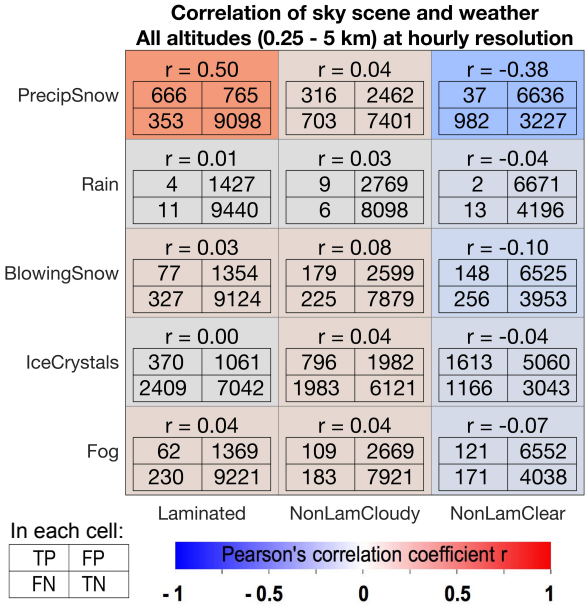


Figure 4. Hourly resolution: Correlogram showing the correlation between computer classified sky scenes and ECCC weather reported at ground level. The details within each cell are given in the caption of Fig. 3.

Considering the precipitating snow class, correlations with laminations are stronger (positive and negative) than those with other weather conditions (see Fig. 4), but the hourly correlations are weaker than the daily correlations (r=0.62 daily to 0.50 hourly). Nonlaminated Cloudy has changed sign from strongly negative to weakly positive (r=-0.37 daily to r=+0.04 hourly), and Nonlaminated Clear has decreased in strength (r=-0.41 daily to r=-0.38 hourly). No other classifications correlate ($r \leq 0.1$).

The change in sign for Nonlaminated Cloudy correlations from moderately negative to a very small positive can be partly explained by several known limitations of the computer lamination identification algorithm: (1) The thinnest laminations are not included in the identification, which leads to those hours being binned hourly as Nonlaminated Cloudy. (2) The algorithm is very conservative in identifying the start and end of the laminated regions. The edges of some regions are missed, thus some hours that actually have laminations are classed as Nonlaminated Cloudy.

When there is a time lag (positive or negative) between the onset of laminations and the precipitating snow, the bins identifying laminations may lead or lag those identifying with snow, thereby reducing the correlation coefficient. We conducted an

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

auto-correlation analysis and found no lag; however, since the data are discontinuous, a rigorous study was not possible, and a visual analysis at hourly resolution may be required.

At daily resolution, the visual and daily computer method criteria are closely comparable, but at hourly resolution, using the computer algorithm, the likelihood of false negative laminations increases. Given a hypothetical day with laminations during all 22 weather-observed hours, we would expect 1 true positive at daily resolution, 22 true positives at hourly resolution. However, suppose the algorithm detects the laminations during only one of the hours. In that case, the daily resolution result is the same (1 true positive), and the hourly result would be one correctly classified hour (Laminated, true positive) and 21 incorrectly classified hours (Nonlaminated Clear or Nonlaminated Cloudy, false negatives). Therefore, optimizing detection criteria for laminated pixels is very important at the hourly resolution.

7 Altitude separation

Radenz et al. (2021) have demonstrated that for similar temperatures, cloud layers that are coupled to the aerosol-rich boundary layer show larger and more rapid ice formation than de-coupled clouds (cloud base above 2 km); most aerosols are confined below 2 km at their non-Arctic study site, but they cite similar Arctic results from Achtert et al. (2020). Shipborne Arctic lidar measurements by Griesche et al. (2021) indicated that for mixed-phase clouds above -10°C, heterogeneous ice formation occurs 2 to 6 times more often when the cloud layer is coupled to the surface, than when it is not. Below -15°C, this difference is reduced. This finding is consistent with the hypothesis that continental aerosols are the dominant ice-nucleating particles in the Arctic free troposphere.

We now extend our analysis to different altitude regions and find that the sky scenes from CRL often show quite different features below 1 km than above that altitude. Further, the computer algorithm shows a bias toward including more laminations at lower altitudes than the visual method does, and these laminations tend to be thicker and have a different visual appearance than those at higher altitudes. We analyzed two separate altitude regions: “Low altitude” (250-1000 m) and “High altitude” (1000-5000 m).

7.1 Method

For this analysis, the identification of Obscured regions was done by examining the top 150 m of each altitude region: 4.85 - 5 km for the upper altitude region and 0.85 - 1 km for the lower altitude region.

7.2 Results: Low and High altitude sky scene classification

Fig. 5 presents the hourly resolution correlations between the sky scene and ECCC weather for 250 m to 1 km in panel 5(a) and 1 km to 5 km in panel 5 (b).

Precipitating snow is the only type of weather showing any significant correlation with sky classification. The highest positive correlation is between laminations and snow at high altitudes (r = 0.47), and weaker correlation at low altitudes (r = 0.3). Nonlaminated Cloudy skies show an interesting result, where high altitudes still show no correlation with snow, but low

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

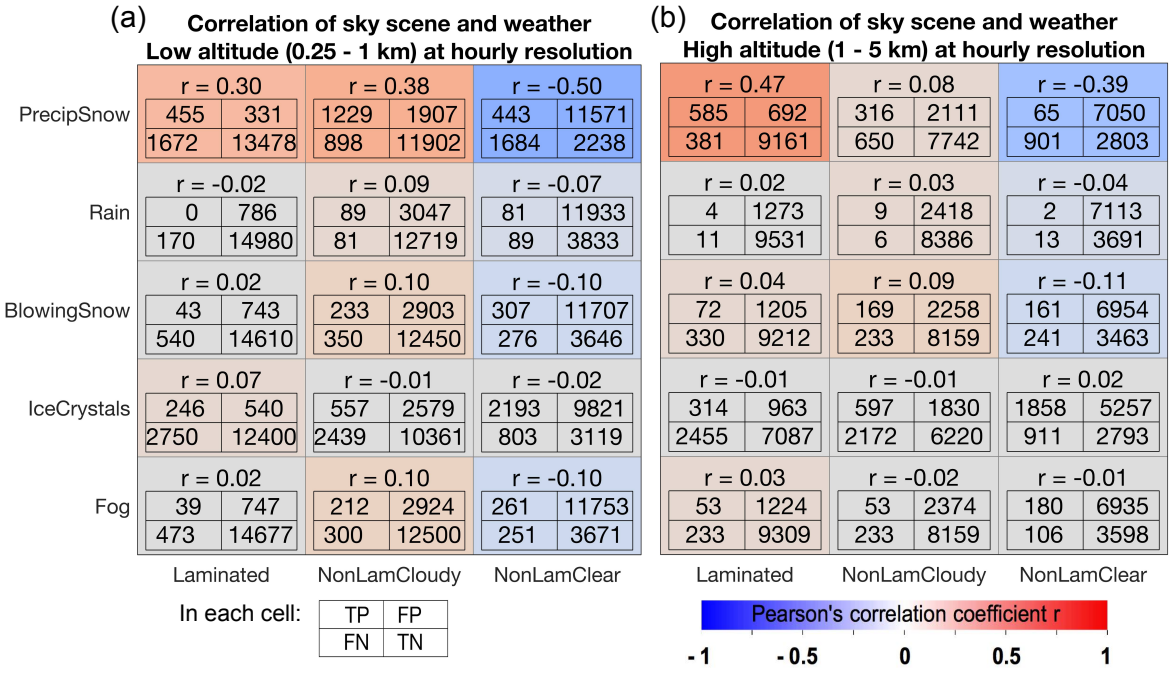


Figure 5. Correlograms showing the correlation between computer-classified sky scenes and ECCC weather reported at ground level. (a) Hourly, low altitude (b) Hourly, high altitude. Confusion matrices in each cell have the same style as in Fig. 3

altitudes show a moderate correlation (r = 0.38). Possibly, snow has been formed in higher-altitude laminated clouds, rather than being produced locally below 1 km (see Fig. 6).

The moderate negative correlations between precipitating snow and clear skies in both altitude ranges (r = -0.50 at low altitude, and r = -0.39 at high altitude) are consistent with the results for the entire altitude range 0.25 - 5 km together in Fig. 4.

460 Remaining comparisons in Fig. 5 have small correlations not exceeding r= ±0.11.

We next examine the various combinations of sky scenes that occur between low- and high- altitude regions, and relate them to the production of precipitating snow. Some combinations occur seldom, for example, a clear sky at high altitude simultaneously with laminations at low altitude (30 hours). Other combinations occur more often, e.g., clear sky simultaneously in both altitude regions (6678 hours). Table 2 shows the number of hours with each combination of sky conditions. Obscured

465 and Incomplete hours, and those with no weather report available, have been omitted.

Figure 6 shows the correlations present with each combination of a low and a high sky scene classification shown in Table 2. Only hours with valid sky scene classifications for both altitude regions independently, and a corresponding ECCC weather report are included and 10810 hours met these criteria, with 961 reporting snow (8.9% of the hours in this sample), and 9849 hours not reporting snow. There was no appreciable change in the results in this section when we repeated the correlation test

470 using a randomly selected 961-hour subset of the hours with no reported snow, to match the sample size of the hours reporting snow.

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table 2. Number of hours in the input dataset for Fig. 5 with each combination of high altitude sky scene classification and low altitude sky scene classification.

	High Laminated	**High NonlamCloudy**	**High NonlamClear**
Low Laminated	632	59	30
Low NonlamCloudy	349	698	405
Low NonlamClear	289	1670	6678

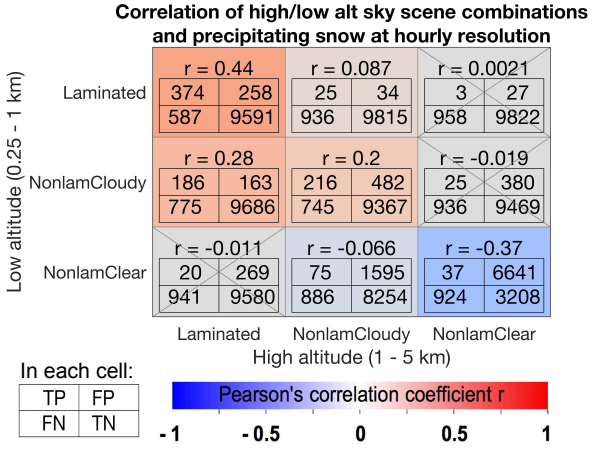


Figure 6. Correlograms showing the correlation between computer-classified sky scene combinations (high altitude sky classification on the x-axis and low altitude sky classification on the y-axis) and ECCC weather reported at ground level, considering only the presence or absence of precipitating snow. Confusion matrices in each cell have the same style as in Fig. 3. Presence of precipitating snow is considered the "truth", and the sky scene combination is considered the "prediction".

The two altitude ranges cannot be considered entirely independently, since precipitation and meteorology in one part of the atmospheric column affect other regions. Precipitating snow from above can seed more precipitation in lower clouds. When laminations are present at any altitude (1359 hours), nearly half are found in both the low- and high-altitude regions simultaneously (632 hours, or 47 % of such samples), and correlation with snow is highest in these cases (r = 0.44). Equally often, both altitudes are cloudy, but laminations are confined exclusively to the higher-altitude region (638 hours, 47 % of the samples), and the correlation with snow is moderate (r = 0.28). Laminations are seldom confined exclusively to the lower-altitude region (89 hours, 6.5 % of the samples), and in these cases, there is practically no correlation with snow (r = 0.09).

7.3 Discussion

The highest correlation with snow is with laminations present at all altitudes, dominated by the presence of laminations at high altitudes, influenced only negligibly by the addition of laminations at low altitudes. Given the stark visual differences between laminations identified at low and high altitudes, it is perhaps for good reason that the thicker low altitude stripes were

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

generally discounted by visual observers - these appear to be a distinct atmospheric feature, undergoing different processes. Therefore, the "false positive" lamination cases may be misidentifications of features from a phenomenon quite separate from the processes of the thinner, higher laminations which are related to the production of snow. In the future, refinements to the detection criteria (particularly at the hourly resolution) are needed in order to reject the thick, low-altitude layered features, if the focus is to be on laminations which are related to precipitating snow.

8 Physical interpretation of laminations, and outlook

The laminations discussed in this study, particularly their potential role in precipitation formation, represent a relatively recent observation from a data-sparse region of the atmosphere: the Arctic troposphere. Several questions direct our interpretations and future investigations.

We become interested in the spatial distribution of hydrometeors as they form within the clouds: Are more, or larger, hydrometeors preferentially formed in-situ within the confines of the positive layers? If so, this may indicate that the laminated structure precedes and predisposes the cloud to the production of snow. Conversely, are droplets and/or crystals formed uniformly throughout (at least part of) the cloud volume, with those of a particular phase or crystal shape then accumulating preferentially in layers as they fall? In the latter case, the formation and movement of precipitating snow would then be responsible for initiating the development of the laminated structure of the cloud.

We have some insight based on lidar measurements. Observational evidence from McCullough et al. (2019) suggest that the structure of the laminations may precede the condensation into cloud: There are examples of continuous individual stripes extending from one cloudy region through an area of lower overall backscatter into another cloudy region.

In the future, investigations into the dynamics of these laminations and the individual hydrometeors within them would benefit from comparisons with similarly layered cloud structures observed in other atmospheric regions, even if their microphysical formation mechanisms differ. For example, mesospheric noctilucent clouds (NLCs) studied by Schäfer et al. (2020) frequently exhibit morphologies strikingly similar to those described here. Both tropospheric and mesospheric laminations share several features: stacks of thin layers within a cloud, persistence exceeding one hour, rapid vertical variation, and both parallel and non-parallel layer motion. Such similar structures in two distinct atmospheric regimes suggest that thin laminae may be more widespread than currently recognized, perhaps overlooked due to insufficient measurement resolution.

Given that the visual approach becomes unwieldy for large datasets and that large datasets are necessary, this computer algorithm, broadly applicable to any visually similar phenomenon, provides an efficient tool for future investigations.

Measurements of supporting cloud and atmospheric properties at much higher spatial resolution will elucidate the differences in the atmospheric state between positive and negative stripes, both within and surrounding laminated clouds. Presently, density, temperature, and relative humidity measurements are commonly available at resolution far coarser than those required for such an analysis, e.g. from radiosondes. We anticipate that higher resolution data taken during intensive ground-based and aircraft campaigns, intended to be co-added to maximize signal-to-noise, may have additional applications at its native resolution for an investigation of laminated clouds and the background laminated structures they seem to inhabit.

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

9 Conclusions

We have developed a computer algorithm to automatically identify laminated cloud features in daily time–altitude arrays of log(attenuated backscatter) lidar measurements. It automatically detects and quantifies laminations, providing a new tool for investigating cloud microphysics in regions with sparse in situ measurements. To demonstrate its utility, we validated it using the results from our previous visual identification method.

In the first validation of the daily results, the algorithm achieved an accuracy of 0.86, an F1-score of 0.76, and a correlation coefficient of r = 0.66, comparable to the visual method. The second validation with the ECCC data confirmed that the strong correlation observed visually between laminations and snow was also present. The algorithm had a correlation coefficient of r = 0.62, while the visual method had r = 0.63.

When the analysis was extended to hourly resolution, a moderate correlation (r = 0.5) with laminated clouds and ground-level snow was found, compared to r = 0.62 daily.

Laminations above 1 km appear most relevant to the production of precipitating snow detected at the surface, given their much higher correlation (r = 0.28) as compared to the negligible correlations with laminations below 1 km (r = 0.09). For laminations that occur simultaneously below and above 1 km, there is a moderate correlation with snow (r = 0.44). With cloudy skies at all altitudes and with laminations confined exclusively below 1 km, the correlation with snow is negligible (r = 0.09). In contrast, if the laminations are confined exclusively above 1 km, the correlation with snow is much larger (r = 0.28).

The smallest scales the algorithm can resolve are the resolution limit of our lidar, 7.5 metres, and these are detectable only if the individual stripes making up the laminations each remain horizontal over a continuous 10-minute period. However, these fine features suggest that even finer measurements are needed.

These results underscore the importance of high-resolution lidar observations for studying fine-scale cloud structures are relevant to Arctic precipitation formation.

Code and data availability.

- **CRL data**: A standardized set of range-scaled photocount plots is available in the Borealis CRL Dataverse data repository (McCullough, 2021). The associated .mat files contain time, altitude, and range-scaled signal values, allowing the reader to reproduce the plots at any desired colour scale or aspect ratio. A spreadsheet identifying the visual method sky-scene classification for each day is also provided. Further CRL measurement data is available upon request from the corresponding author (e.mccullough@dal.ca).
- **ECCC data** is available at: https://climate.weather.gc.ca/historical_data/search_historic_data_e.html; Last verified 6 June 2020. The search parameter used is: Station Name “Eureka”.

  The three links to Station “Eureka A” all include data from dates applicable to this project. They link to sites with different Climate ID numbers assigned by the Meteorological Service of Canada, and the weather observations for these sites are all made near the CRL lidar. New Climate ID numbers are assigned when stations discontinue a particular type of observation, even if other observations continue at the exact location, under the same site name. The three links are:

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

- Climate ID 2401203. Most of the data used in this paper are from this site, located at -85.81 W, 79.99 N. Data are available from 2016 Feb 22 at 15:00 UTC through the last day of measurement included in this paper, 6 December 2019 and beyond. Date last verified in June 2020.
- Climate ID 2401208. Not used in this paper. No data are available for early February 2016. Most days have no recorded weather. Located at -85.81 W, 79.99 N.
- Climate ID 2401200. This paper uses data from this site only for dates on which it is unavailable from Climate ID 2401203. Data are available from Climate ID 2401200 for January 2016 through 25 February at 12:00 UTC. This record has more frequent observations than the record from Climate ID 2401203. However, each record holds fewer values (e.g., the version from Climate ID 2401203 may say "Blowing Snow, Fog" and have an entry only every three hours, while Climate ID 2401200 records something every hour, but for that particular entry lists only "Blowing Snow"). Located at -85.93 W, 79.98 N.

Until early 2016, weather observations and temperatures were recorded at or near the Weather Station building at Eureka. After that, some equipment was moved to the airport runway, which is 1.8 km to the Northeast and 73 m higher in elevation; however, weather observations are still made at the Weather Station. The Weather Station is closer to the CRL lidar, and at a similar altitude.

There is a mechanism for downloading monthly CSV files of hourly observations, and these are the data used for this project.

A spreadsheet of the weather conditions reported for each study day is included in the Dataverse data repository with the CRL data.

See https://climate.weather.gc.ca/glossary_e.html#climate_ID for further information about interpreting the historical climate data.

- **Code availability**: The MATLAB software for lamination identification and sky classification is available in the Zenodo code repository (McCullough, 2026).

Competing interests. At least one of the (co-)authors is a member of the editorial board of Atmospheric Measurement Techniques. The authors also have no other competing interests to declare.

Author contributions. EMM: Conceptualization, methodology, software, data curation, formal analysis, investigation, visualization, writing - original draft, writing - review and editing. RW: Formal analysis, writing - review and editing. JRD: Former Principal Investigator of PEARL laboratory. Funding acquisition, project administration, resources, methodology, supervision, writing - review and editing.

Acknowledgements. The Natural Sciences and Engineering Research Council, Environment and Climate Change Canada and the Canadian Space Agency support this research.

PEARL has been supported by a large number of agencies whose support is gratefully acknowledged: The Canadian Foundation for Innovation; the Ontario Innovation Trust; the (Ontario) Ministry of Research and Innovation; the Nova Scotia Research and Innovation Trust; the Natural Sciences and Engineering Research Council; the Canadian Foundation for Climate and Atmospheric Science; Environment and Climate Change Canada (ECCC; who also provided the radiosonde data); Polar Continental Shelf Project; the Department of Indigenous and Northern Affairs Canada; and the Canadian Space Agency. This work was carried out during the Canadian Arctic ACE/OSIRIS Validation Campaigns of 2016, 2017, 2018, and 2019. The Canadian Space Agency, Environment and Climate Change Canada, the Natural Sciences and

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

Engineering Research Council of Canada, and the Northern Scientific Training Program funded it. This project has also been supported by NSERC Discovery Grants and Northern Supplement Grants held by James R. Drummond, Robert J. Sica, and Kaley A. Walker, as well as the NSERC CREATE Training Program in Arctic Atmospheric Science (PI: Kim Strong). Thomas J. Duck was the initial Principal Investigator for CRL lidar and contributed helpful comments regarding early drafts of the paper.

In addition, the authors thank the following groups and individuals for their support during field campaigns at Eureka: PEARL site manager Pierre Fogal; Canadian Arctic ACE/OSIRIS Validation Campaign project lead Kaley A. Walker, CRL operators: Graeme Nott, Chris Perro, Colin P. Thackray, Jason Hopper, Shayamila Mahagammulla Gamage, Jon Doyle; Canadian Network for the Detection of Atmospheric Change (CANDAC) support scientist Alexey Tikhomirov; CANDAC operators: Mike Maurice, Peter McGovern, John Gallagher, Alexei Khmel, Paul Leowen, Ashley Harrett, Keith MacQuarrie, Oleg Mikhailov, and Matt Okraszewski; and the Eureka Weather Station staff. Thanks to members of CANDAC for their helpful discussions.

RW was partially supported by the Deutsche Forschungsgemeinschaft project 445400792 “Analyzing the Motion of the Middle Atmosphere Using Nighttime RMR–lidar Observations at the Midlatitude Station Kühlungsborn (AMUN)”.

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

References

Achtert, P., O'Connor, E. J., Brooks, I. M., Sotiropoulou, G., Shupe, M. D., Pospichal, B., Brooks, B. J., and Tjernström, M.: Properties of Arctic liquid and mixed-phase clouds from shipborne Cloudnet observations during ACSE 2014, Atmospheric Chemistry and Physics, 20, 14 983–15 002, 2020.

Baduge, S. K., Thilakarathna, S., Perera, J. S., Arashpour, M., Sharafi, P., Teodosio, B., Shringi, A., and Mendis, P.: Artificial intelligence and smart vision for building and construction 4.0: Machine and deep learning methods and applications, Automation in Construction, 141, 104 440, 2022.

Barrett, A. I., Hogan, R. J., and Forbes, R. M.: Why are mixed-phase altocumulus clouds poorly predicted by large-scale models? Part 1. Physical processes, Journal of Geophysical Research: Atmospheres, 122, 9903–9926, 2017a.

Barrett, A. I., Hogan, R. J., and Forbes, R. M.: Why are mixed-phase altocumulus clouds poorly predicted by large-scale models? Part 2. Vertical resolution sensitivity and parameterization, Journal of Geophysical Research: Atmospheres, 122, 9927–9944, 2017b.

Bianconi, F., Fernández, A., Smeraldi, F., and Pascoletti, G.: Colour and Texture Descriptors for Visual Recognition: A Historical Overview, Journal of Imaging, 7, https://doi.org/10.3390/jimaging7110245, 2021.

Bühl, J., Seifert, P., Myagkov, A., and Ansmann, A.: Measuring ice- and liquid-water properties in mixed-phase cloud layers at the Leipzig Cloudnet station, Atmospheric Chemistry and Physics, 16, 10 609–10 620, https://doi.org/10.5194/acp-16-10609-2016, 2016.

Chicco, D. and Jurman, G.: The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation, BMC genomics, 21, 1–13, 2020.

Chicco, D., Tötsch, N., and Jurman, G.: The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation, BioData mining, 14, 1–22, 2021.

Cohen, J.: Statistical Power Analysis for the Behavioural Sciences, Lawrence Erlbaum Associates, 2nd edn., 1988.

Duck, T. J., Whiteway, J. A., and Carswell, A. I.: Lidar observations of gravity wave activity and Arctic stratospheric vortex core warming, Geophysical Research Letters, 25, 2813–2816, 1998.

Farhani, G., Sica, R. J., and Daley, M. J.: Classification of lidar measurements using supervised and unsupervised machine learning methods, Atmospheric Measurement Techniques, 14, 391–402, https://doi.org/10.5194/amt-14-391-2021, 2021.

Field, P. and Heymsfield, A.: Importance of snow to global precipitation, Geophysical Research Letters, 42, 9512–9520, 2015.

Field, P. R., Lawson, R. P., Brown, P. R., Lloyd, G., Westbrook, C., Moisseev, D., Miltenberger, A., Nenes, A., Blyth, A., Choularton, T., et al.: Secondary ice production: Current state of the science and recommendations for the future, Meteorological Monographs, 58, 7–1, 2017.

Georgakaki, P., Sotiropoulou, G., Vignon, É., Billault-Roux, A.-C., Berne, A., and Nenes, A.: Secondary ice production processes in wintertime alpine mixed-phase clouds, Atmospheric Chemistry and Physics, 22, 1965–1988, 2022.

Ghodrati, M., Farzmahdi, A., Rajaei, K., Ebrahimpour, R., and Khaligh-Razavi, S.-M.: Feedforward object-vision models only tolerate small image variations compared to human, Frontiers in Computational Neuroscience, 8, https://doi.org/10.3389/fncom.2014.00074, 2014.

Griesche, H. J., Ohneiser, K., Seifert, P., Radenz, M., Engelmann, R., and Ansmann, A.: Contrasting ice formation in Arctic clouds: surface-coupled vs. surface-decoupled clouds, Atmospheric Chemistry and Physics, 21, 10 357–10 374, 2021.

Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J., Kim, R., Raman, R., Nelson, P. C., Mega, J. L., and Webster, D. R.: Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs, JAMA, 316, 2402–2410, https://doi.org/10.1001/jama.2016.17216, 2016.

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

Hallett, J. and Mossop, S.: Production of secondary ice particles during the riming process, Nature, 249, 26–28, 1974.

Heinle, A., Macke, A., and Srivastav, A.: Automatic cloud classification of whole sky images, Atmospheric Measurement Techniques, 3, 557–567, https://doi.org/10.5194/amt-3-557-2010, 2010.

Heymsfield, A. J., Schmitt, C., Bansemer, A., Gettelman, A., Field, P. R., Liu, C., et al.: Contributions of the liquid and ice phases to global surface precipitation: Observations and global climate modeling, Journal of the Atmospheric Sciences, 77, 2629–2648, 2020.

Hinojosa Lee, M. C., Braet, J., and Springael, J.: Performance Metrics for Multilabel Emotion Classification: Comparing Micro, Macro, and Weighted F1-Scores, Applied Sciences, 14, https://doi.org/10.3390/app14219863, 2024.

Hobbs, P. V. and Rangno, A. L.: Ice particle concentrations in clouds, Journal of the atmospheric sciences, 42, 2523–2549, 1985.

Hobbs, P. V. and Rangno, A. L.: Rapid development of high ice particle concentrations in small polar maritime cumuliform clouds, Journal of Atmospheric Sciences, 47, 2710–2722, 1990.

Hobbs, P. V. and Rangno, A. L.: Microstructures of low and middle-level clouds over the Beaufort Sea, Quarterly Journal of the Royal Meteorological Society, 124, 2035–2071, 1998.

Illingworth, A., Hogan, R., O'connor, E., Bouniol, D., Brooks, M., Delanoë, J., Donovan, D., Eastment, J., Gaussiat, N., Goddard, J., et al.: Cloudnet: Continuous evaluation of cloud profiles in seven operational models using ground-based observations, Bulletin of the American Meteorological Society, 88, 883–898, 2007.

Jafarinaimi, N.: Exploring the character of participation in social media: the case of Google Image Labeler, in: Proceedings of the 2012 IConference, iConference 2012, p. 72–79, Association for Computing Machinery, New York, NY, USA, ISBN 9781450307826, https://doi.org/10.1145/2132176.2132186, 2012.

Jann, A.: Detection of gravity waves in Meteosat imagery by grating cell operators, European Journal of Remote Sensing, 50, 509–516, 2017.

Kalesse-Los, H., Schimmel, W., Luke, E., and Seifert, P.: Evaluating cloud liquid detection against Cloudnet using cloud radar Doppler spectra in a pre-trained artificial neural network, Atmospheric Measurement Techniques, 15, 279–295, https://doi.org/10.5194/amt-15-279-2022, 2022.

Kanji, Z. A., Ladino, L. A., Wex, H., Boose, Y., Burkert-Kohn, M., Cziczo, D. J., and Krämer, M.: Overview of ice nucleating particles, Meteorological monographs, 58, 1–1, 2017.

Karimi-Rouzbahani, H., Bagheri, N., and Ebrahimpour, R.: Hard-wired feed-forward visual mechanisms of the brain compensate for affine variations in object recognition, Neuroscience, 349, 48–63, https://doi.org/https://doi.org/10.1016/j.neuroscience.2017.02.050, 2017a.

Karimi-Rouzbahani, H., Bagheri, N., and Ebrahimpour, R.: Invariant object recognition is a personalized selection of invariant features in humans, not simply explained by hierarchical feed-forward vision models, Sci Rep, 7, 14 402, 2017b.

Korolev, A. V., Isaac, G. A., Strapp, J. W., Cober, S. G., and Barker, H. W.: In situ measurements of liquid water content profiles in midlatitude stratiform clouds, Quarterly Journal of the Royal Meteorological Society, 133, 1693–1699, https://doi.org/10.1002/qj.147, 2007.

Lahiri, M., Tantipathananandh, C., Warungu, R., Rubenstein, D., and Berger-Wolf, T.: Biometric animal databases from field photographs: Identification of individual zebra in the wild, in: Proceedings of the 1st ACM International Conference on Multimedia Retrieval, ICMR'11, p. 6, https://doi.org/10.1145/1991996.1992002, 2011.

Lu, D. and Weng, Q.: A survey of image classification methods and techniques for improving classification performance, International Journal of Remote Sensing, 28, 823–870, https://doi.org/10.1080/01431160600746456, 2007.

Luque, A., Carrasco, A., Martín, A., and de las Heras, A.: The impact of class imbalance in classification performance metrics based on the binary confusion matrix, Pattern Recognition, 91, 216–231, https://doi.org/https://doi.org/10.1016/j.patcog.2019.02.023, 2019.

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.



McCullough, E. M.: A new technique for interpreting depolarization measurements using the CRL atmospheric lidar in the Canadian High Arctic, Ph.D. thesis, The University of Western Ontario, London, Ontario, Canada, http://ir.lib.uwo.ca/etd/3418, 2015.

McCullough, E. M.: Replication Data for: The relationship between clouds containing multiple layers 7.5 - 30 m thick and surface weather conditions, https://doi.org/10.5683/SP2/ST9YSB, 2021.

McCullough, E. M.: Code for Automatic Detection of Laminated Clouds in Lidar Data, https://doi.org/10.5281/zenodo.20144243, 2026.

McCullough, E. M., Sica, R. J., Drummond, J. R., Nott, G., Perro, C., Thackray, C. P., Hopper, J., Doyle, J., Duck, T. J., and Walker, K. A.: Depolarization calibration and measurements using the CANDAC Rayleigh–Mie–Raman lidar at Eureka, Canada, Atmospheric Measurement Techniques, 10, 4253–4277, https://doi.org/10.5194/amt-10-4253-2017, 2017.

McCullough, E. M., Drummond, J. R., and Duck, T. J.: Lidar measurements of thin laminations within Arctic clouds, Atmospheric Chemistry and Physics, 19, 4595–4614, 2019.

McCullough, E. M., Wing, R., and Drummond, J. R.: The Relationship between Clouds Containing Multiple Layers 7.5–30 m Thick and Surface Weather Conditions, Atmosphere, 12, 1616, 2021.

Mülmenstädt, J., Sourdeval, O., Delanoë, J., and Quaas, J.: Frequency of occurrence of rain from liquid-, mixed-, and ice-phase clouds derived from A-Train satellite retrievals, Geophysical Research Letters, 42, 6502–6509, 2015.

Nielsen, K. J., Logothetis, N. K., and Rainer, G.: Discrimination Strategies of Humans and Rhesus Monkeys for Complex Visual Displays, Current Biology, 16, 814–820, https://doi.org/https://doi.org/10.1016/j.cub.2006.03.027, 2006.

Nott, G., Duck, T., Doyle, J., Coffin, M., Perro, C., Thackray, C., Drummond, J., Fogal, P., McCullough, E. M., and Sica, R.: A remotely operated lidar for aerosol, temperature, and water vapor profiling in the High Arctic, Journal of Atmospheric and Oceanic Technology, 29, 221–234, 2012.

Paymode, A. S. and Malode, V. B.: Transfer learning for multi-crop leaf disease image classification using convolutional neural network VGG, Artificial Intelligence in Agriculture, 6, 23–33, 2022.

Phillips, V. T., Patade, S., Gutierrez, J., and Bansemer, A.: Secondary ice production by fragmentation of freezing drops: Formulation and theory, Journal of the Atmospheric Sciences, 75, 3031–3070, 2018.

Pinto, N., Cox, D. D., and DiCarlo, J. J.: Why is Real-World Visual Object Recognition Hard?, PLOS Computational Biology, 4, 1–6, https://doi.org/10.1371/journal.pcbi.0040027, 2008.

Praß, M., Grimsen, C., König, M., and Fahle, M.: Ultra Rapid Object Categorization: Effects of Level, Animacy and Context, PLOS ONE, 8, 1–10, https://doi.org/10.1371/journal.pone.0068051, 2013.

Radenz, M., Bühl, J., Seifert, P., Baars, H., Engelmann, R., Barja González, B., Mamouri, R.-E., Zamorano, F., and Ansmann, A.: Hemispheric contrasts in ice formation in stratiform mixed-phase clouds: disentangling the role of aerosol and dynamics with ground-based remote sensing, Atmospheric Chemistry and Physics, 21, 17 969–17 994, https://doi.org/10.5194/acp-21-17969-2021, 2021.

Reichert, R., Kaifler, N., and Kaifler, B.: Limitations in Wavelet Analysis of Non-Stationary Atmospheric Gravity Wave Signatures in Temperature Profiles, Atmospheric Measurement Techniques Discussions, 2023, 1–25, https://doi.org/10.5194/amt-2023-240, 2023.

Ridder, C., Baumgarten, G., Fiedler, J., Lübken, F.-J., and Stober, G.: Analysis of small-scale structures in lidar observations of noctilucent clouds using a pattern recognition method, Journal of Atmospheric and Solar-Terrestrial Physics, 162, 48–56, https://doi.org/https://doi.org/10.1016/j.jastp.2017.04.005, layered Phenomena in the Mesopause Region, 2017.

Saito, T. and Rehmsmeier, M.: The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets, PloS one, 10, e0118 432, 2015.

https://doi.org/10.5194/egusphere-2026-4312
Preprint. Discussion started: 5 November 2026
© Author(s) 2026. CC BY 4.0 License.

Schäfer, B., Baumgarten, G., and Fiedler, J.: Small-scale structures in noctilucent clouds observed by lidar, Journal of Atmospheric and Solar-Terrestrial Physics, 208, 105 384, 2020.

Sedlar, J., Igel, A., and Telg, H.: Processes contributing to cloud dissipation and formation events on the North Slope of Alaska, Atmospheric Chemistry and Physics, 21, 4149–4167, https://doi.org/10.5194/acp-21-4149-2021, 2021.

Shupe, M. D.: A ground-based multisensor cloud phase classifier, Geophysical Research Letters, 34, https://doi.org/https://doi.org/10.1029/2007GL031008, 2007.

Sica, R. and Russell, A.: Measurements of the effects of gravity waves in the middle atmosphere using parametric models of density fluctuations. Part I: Vertical wavenumber and temporal spectra, Journal of the atmospheric sciences, 56, 1308–1329, 1999.

Tarr, M., Williams, P., Hayward, W., and Gauthier, I.: Three-dimensional object recognition is viewpoint dependent, Nature Neuroscience, 1, 275–277, https://doi.org/10.1038/1089, funding Information: This research was supported by an Air Force Office of Scientific Research Grant. We thank Jay Servidea and Jaymz Rosoff for their assistance in running the psychophysical studies., 1998.

Tarr, M. J. and Bülthoff, H. H.: Image-based object recognition in man, monkey and machine, Cognition, 67, 1–20, https://doi.org/https://doi.org/10.1016/S0010-0277(98)00026-2, 1998.

Tjan, B. S., Braje, W. L., Legge, G. E., and Kersten, D.: Human efficiency for recognizing 3-D objects in luminance noise, Vision Research, 35, 3053–3069, https://doi.org/https://doi.org/10.1016/0042-6989(95)00070-G, 1995.

Vassel, M., Ickes, L., Maturilli, M., and Hoose, C.: Classification of Arctic multilayer clouds using radiosonde and radar data in Svalbard, Atmospheric Chemistry and Physics, 19, 5111–5126, https://doi.org/10.5194/acp-19-5111-2019, 2019.

Verlinde, J., Harrington, J. Y., McFarquhar, G. M., Yannuzzi, V. T., Avramov, A., Greenberg, S., Johnson, N., Zhang, G., Poellot, M. R., Mather, J. H., Turner, D. D., Eloranta, E. W., Zak, B. D., Prenni, A. J., Daniel, J. S., Kok, G. L., Tobin, D. C., Holz, R., Sassen, K., Spangenberg, D., Minnis, P., Tooman, T. P., Ivey, M. C., Richardson, S. J., C. P. Bahrmann, S. M., DeMott, P. J., Heymsfield, A. J., and Schofield, R.: The mixed phase Arctic cloud experiment, Bulletin of the American Meteorological Society, pp. 205–211, 2007.

Verlinde, J., Rambukkange, M. P., Clothiaux, E. E., McFarquhar, G. M., and Eloranta, E. W.: Arctic multilayered, mixed-phase cloud processes revealed in millimeter-wave cloud radar Doppler spectra, Journal of Geophysical Research: Atmospheres, 118, 13–199, 2013.

von Ahn, L. and Dabbish, L.: Labeling images with a computer game, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '04, p. 319–326, Association for Computing Machinery, New York, NY, USA, ISBN 1581137028, https://doi.org/10.1145/985692.985733, 2004.

von Terzi, L., Dias Neto, J., Ori, D., Myagkov, A., and Kneifel, S.: Ice microphysical processes in the dendritic growth layer: a statistical analysis combining multi-frequency and polarimetric Doppler cloud radar observations, Atmospheric Chemistry and Physics, 22, 11 795–11 821, https://doi.org/10.5194/acp-22-11795-2022, 2022.

Welton, E. J., Campbell, J. R., Spinhirne, J. D., and Scott III, V. S.: Global monitoring of clouds and aerosols using a network of micropulse lidar systems, in: Lidar remote sensing for industry and environment monitoring, vol. 4153, pp. 151–158, SPIE, 2001.

Whale, T. F.: Ice nucleation in mixed-phase clouds, Elsevier, 2018.