**July 10, 2024**
**Synthesis report**

# Cloud sensor data validation with manually labelled all-sky images and weather measurements

Alessandro Marocco[1]; Supervisor: Noora Partamies [2]

**Abstract**
The Aurora Cloud Sensor III from Aurora Eurotech estimates the clarity of the sky, linked to cloud coverage, from temperature measurements. This sensor has been settled at the Kjell Henriksen Observatory (KHO) and is owned by the University College London. The validation of the cloud sensor data, between November 2019 and February 2020, is conducted with the use of manually labelled all-sky images from an optical camera at KHO and cloud cover index from the Svalbard Airport station operated by the Norwegian Center for Climate Services. The use of a clarity threshold value of 56 to distinguish clear and cloudy sky provides a 90.1% consistency with the manually labelled all-sky images. The clarity variable and the cloud coverage index present a mean differences of about 22% which can be explained by the localization and sampling. The co-evolution of the time series of clarity, labelling and cloud cover index plus the look of images corresponding to certain range of clarity also support the accuracy of the cloud sensor. The use of the 56 clarity threshold for another period of time (January and February 2019) when labelled images are also available provides similar consistency, about 91 %, and is also the best clarity threshold for maximizing consistency. A timetrend analysis of the clarity values over the years reveals a slight trend of reduction of about 4.18 in clarity between 2016 and 2024 but does not seem to affect the polar night period (November to February) too much, which keeps a bimodal distribution of clarity with an average clarity value between 41.78 and 53.14. Determination of the clarity threshold value can be carried out with a smaller labeled data set than the one used (25 739 labels), as random sampling has shown, with a similarly determined clarity threshold value for a 500-fold smaller labeled data set.

**Keywords**
Cloud sensor - Clarity - All sky images - Cloud cover - Kjell Henriksen Observatory

[1] *Département de Géosciences, École Normale Supérieure, Paris, France*
[2] *Department of Arctic Geophysics, University Center in Svalbard, Longyearbyen, Norway*

## Contents



## 1. The cloud sensor

The cloud sensor used at the Kjell Henriksen Observatory (KHO) is the Aurora Cloud Sensor III manufactured by Aurora Eurotech and is owned by the University College London. It is installed on the south-west side of the building and is slightly tilted (about 20°) toward south-east in order to allow rain to run off (Figure 1). The field of view of the sensor is 90° compared to the 180° field of view of the all sky image. So, the sensor collect information corresponding to a circle centered 20° south-east of the zenith on the all sky image.

In order to determine a value of clarity, the sensor measures the far temperature in the sky $T_{\text{sky}}$ and the air temperature $T_{\text{air}}$. Then, the clarity value is simply calculated by subtracting the air temperature by the sky temperature: $\boxed{\text{Clarity} = T_{\text{air}} - T_{\text{sky}}}$. Indeed, the cloud temperature is much higher than the temperature of the far sky. So, when clouds are covering the sky the measured sky temperature raises so the clarity value calculated decreases by definition.

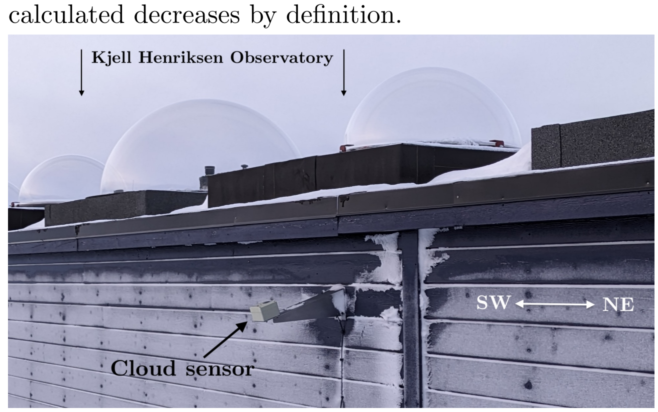


**Figure 1.** Position of the cloud sensor on the Kjell Henriksen Observatory's building

There are also a ceramic rain sensor and a sensor of light included for additional measurements.

## 2. Data presentation

The cloud sensor data used for the validation are from November 2019 to February 2020 with a value every minute. Since the value of clarity is simply a difference of temperatures, its unit is [°C] which doesn't has any interpretative sense as such. The range of the clarity value is between 9.7 and 82.4 with an average of 53.4. The histogram of the clarity values (Figure 2) is bimodal at approximately 35 and 65 that can be respectively attributed to a most common clear sky and cloudy sky.

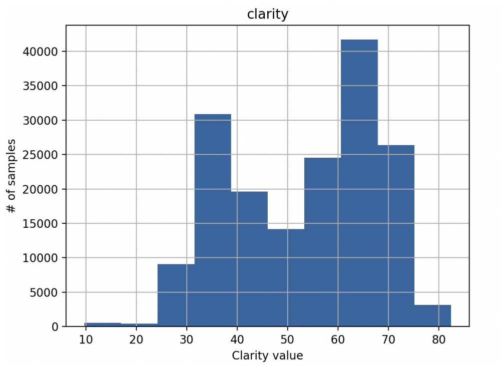


**Figure 2.** Histogram of clarity from the cloud sensor (November 2019 to February 2020)

Therefore, manualy labelled images are used over the same period (November 2019 to February 2020) from the Sony A7s All Sky Color Camera at KHO. The labelling is binary: "Cloudy" and "Clear" and made approximately every 6 minutes. The histogram of the label data (Figure 3) shows slightly more cloudy images than clear images.

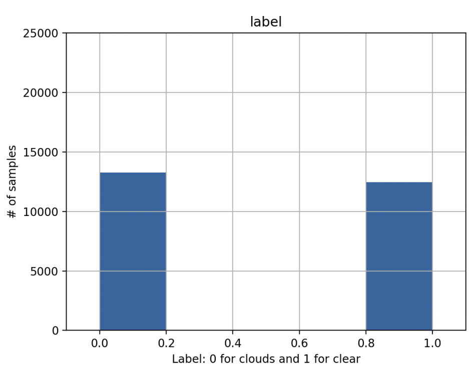


**Figure 3.** Histogram of manual labelled images (November 2019 to February 2020)

Also, the data from the weather station at Svalbard Airport has been recovered from the Norwegian Centre for Climates Services for the same period (November 2019 to February 2020) and a measurement every 3 hours is available. The measurement corresponds to a ponctual meteorologist evaluation which relies on two celiometers suggestion for total cloudiness. These data are for another localization than KHO but provides a source external to UNIS. The cloud coverage index represents the fraction of the sky cover by clouds on a scale of 0 to 8. The maximum value meaning that all the sky is covered by clouds and the minimum value meaning that the sky is all clear. A conversion of these values is realized in order to matches with our data of clarity and label by normalizing and inverting the scale: 0 to 8 becomes 1 to 0. The histogram of the cloud cover index is shown in Figure 4. There is also a bimodal distribution that can be attributed to a most common clear sky and cloudy sky.

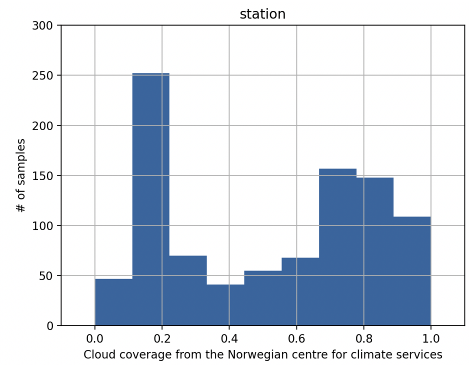


**Figure 4.** Histogram of the cloud cover index from the Norwegian centre for climates services at the Svalbard Lufthavn(November 2019 to February 2020)

## 3. Thresholding method

As explained in the user guide of Aurora Eurotech, the values of clarity are not universal and thresholds have to be set in order to distinguish between cloudy and clear skies. For example, the thresholds determined in the rural UK by Aurora Eurotech are 28 clarity value for the transition Clear / Cloudy and 15 clarity value for the transition Cloudy / Very Cloudy. These thresholds are not at all suitable with our data in Svalbard, notably because of the extended range of values.

In order to choose the best threshold value possible according to our validation set which are considered to be our labelled images here, every threshold value of clarity between the minimum and the maximum of clarity are tested and then the consistency (i.e. the correspondence between the clarity after thresholding and the labelled images) is compute (Figure 5.)

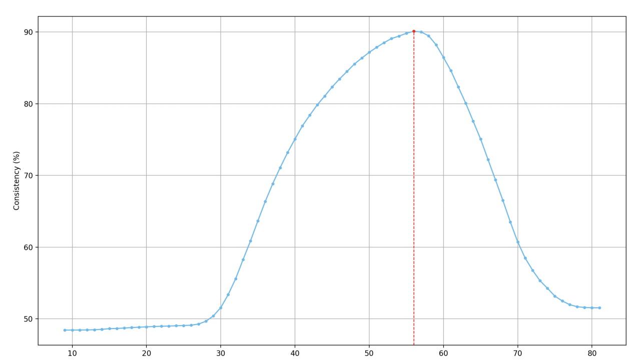


**Figure 5.** Consistency between the clartiy after thresholding and the labelled images for each threshold value of clarity (November 2019 to February 2020).

The best clarity threshold according to our calculations is 56 with approximately 90.1 % correspondence to the manual labelled images. From that value, the clarity data can be categorized in Cloudy (0) and Clear (1) in order to be compared with the manual labelled images which are also binary. The histogram of the clarity categorized as such is shown in Figure 6.

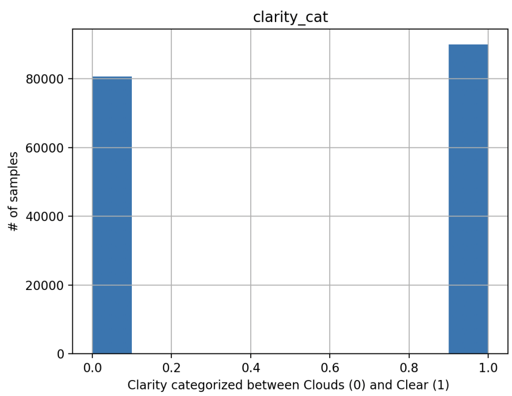


**Figure 6.** Histogram of the clarity binary categorized between cloud and clear (like manualy labelling) with a thresholding value of 56.

## 4. Validation

The validation of the cloud sensor data is realized with both validation data set available which are considered at that time as the exact data. Then, time series can be compared as the same time on an unique plot.

### 4.1 Cloud sensor and labelled images

As previously said, the consistency between the clarity and the manual labelled images for a threshold value of 56 is 90.1%. It means that only 9.9% of the data are misclassified according to the manual labelled images, which is a fairly low number. These misclassified images with such threshold can be seen in the following histograms (Figure 7). In the first histogram, the value of the manual labelling (0 or 1) is shown for clarity categorized as 0 (on the left) and as 1 (on the right). In the case of perfect correspondence, we would expect only 0 value of manual labelling for clarity categorized as 0 meaning that every image categorized as Clouds with the clarity thresholding has also been categorized as Clouds by the manual labelling. The same reasoning applies for the Clear (1).

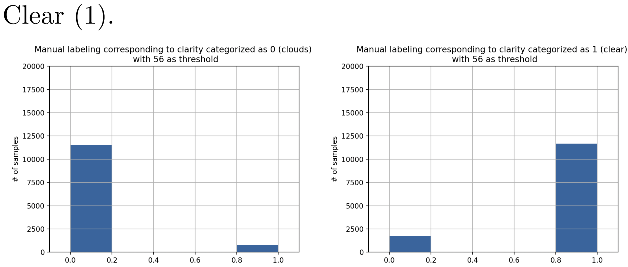


**Figure 7.** Labelling histograms for each category of clarity: 0 (on the left) and 1 (on the right)

It is therefore interesting to note that there are only few numbers of clarity categorized as Clouds that are categorized as Clear by manual labelling and vice versa, as the 9,9 % suggests. But, it seems that the clarity categorized as Clear has at least twice more numbers of misclassified than for Clouds ones.

It is also possible to plot the values of the clarity from the cloud sensor as histogram for each image label: Clouds and Clear (Figure 8). The best threshold value of clarity (56) is also shown. The two histograms are quite well separated. There is no clear skies labelled bellow a clarity value of approximately 40. Once again, it is possible to see that there is more misclassified cloudy sky as clear than clear sky as cloudy because there is more data of cloudy sky above the threshold than data of clear sky bellow the threshold.

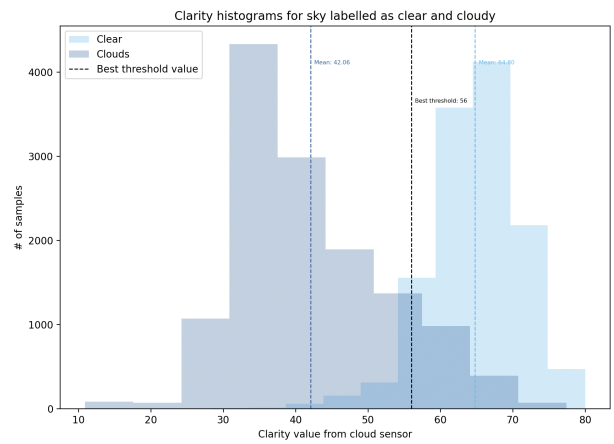


**Figure 8.** Histogram of clarity values for data manually labelled as clear sky and cloudy sky.

Moreover, the all sky images corresponding to the extreme values of clarity can be seen in the Extended figures part (Figure 21 and Figure 22) to get an idea of what the images look like for these specific values of clarity. The value of clarity and the label of each image is written on it. No anomalies detected among these images, the highest clarity is indeed corresponding to clear sky and lowest clarity to cloudy sky.

Also, a video compiling an entire day (08/02/2020) has been made to assess the progressive evolution of clarity with the sky. In this video, it is also possible the see intermediate values of clarity and the all sky images associated to.

### 4.2 Cloud sensor and weather station

The cloud cover index is also used to validate the cloud sensor. First, the clarity values are normalize using the minimum and the maximum values. Then, by using the normalized and reversed cloud cover index, the mean differences between clarity and the cloud cover index, both now in a scale between 0 and 1, is computed. A mean difference of approximately 0.22 is found. It means that the cloud sensor measurement is, in average, 0.22

far away from the cloud cover index, on a scale of 0 to 1. This value is not ideal but the weather station used is located at the Svalbard Airport which is far from KHO.

Another interesting visualization of the data is obtained by making an histogram of the clarity values for each cloud cover code between 0 and 8 (Figure 9). The pic of the histogram is expecting to shift from low clarity values to high clarity values when the cloud cover index decreases. The general trend is indeed following what is expected because the mean of clarity is increasing when the cloud cover index is decreasing but it is more disorganised for intermediate cloud cover index (4 and 5). In contrast, high cloud cover index (6, 7 and 8) has lowest clarity values and low cloud cover index (0, 1, 2 and 3) has higher clarity values. These results may be explained because when extreme events occurs (very clear sky or very cloudy sky), it affects a larger spatial region in the same way (KHO and Svalbard Airport). On the contrary, intermediate events when the sky is between clear and cloudy corresponds to more dynamic weather in time and localization so an intermediate clear sky or intermediate cloudy sky at the Svalbard Aiport is not inconsistent with any type of cloudiness in KHO. It means that the range of clarity values for this kind of event with intermediate cloud cover index at Svalbard Airport may be wider at KHO and still be consistent.

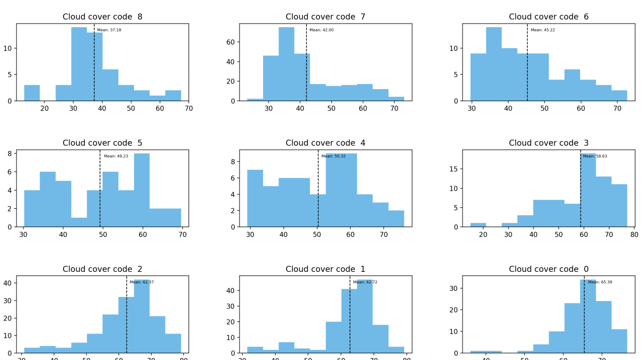


**Figure 9.** Histogram of clarity values corresponding for each cloud cover index from the Svalbard Airport station (0 to 8)

## 4.3 Time series

In order to compare the variables measured and to estimate the accuracy of the cloud sensor, time series are computed. The time series of the clarity is plot with the best threshold clarity value calculated earlier. The manually labelled images are also put in the same plot with binary value (0 for Clouds and 1 for Clear) as well as the normalized and reversed cloud cover index is added. The clarity is the only variable that is measure within a continue range and not binary (0 and 1) or categorized (0, 1, 2, 3, 4, 5, 6, 7 and 8).

The overall time series (Figure 11) shows that the clarity is fairly consistent with the manual labelling and with the cloud cover index. Indeed, most of the time, especially during extreme events, the manual labelling is coherent with the clarity showing Clouds label when the clarity is low and Clear labels when the clarity is high. Then, the cloud cover index at the Svalbard Airport seems to have similar trend as the clarity time series. But the time scale is too large to clearly see the dot points of labelling and the variation of cloud cover index with these plot options and it is necessary to zoom in.

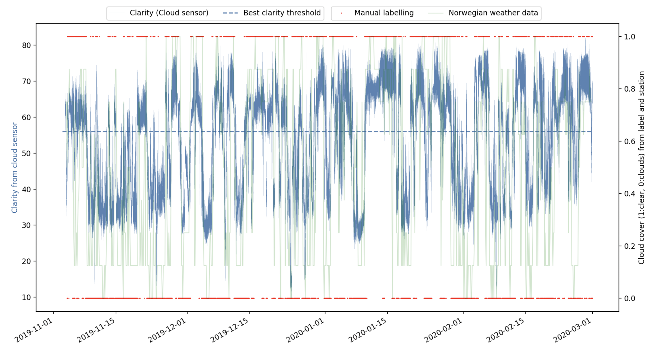


**Figure 10.** Time series of the clarity (cloud sensor) in blue with the best threshold clarity value (56) in line, labelling category (manually) in red and cloud cover index (Norwegian Centre for Climate) in green between November 2019 and February 2020.

As an example, we can zoom into the same day used to make the video compilation of images (08/02/2020). It can also be interesting to watch the video with this plot on the side. First of all, the best threshold value is working well because almost every value of clarity is under the threshold when the label is Clouds and above the threshold when the label is Clear. Then, the trends of the clarity and the cloud cover index are similar because when the clarity increases the cloud cover index also roughly increases and vice-versa. Nevertheless, the correspondence isn't perfect but it is also because there is a value of cloud cover index only every 3 hours.

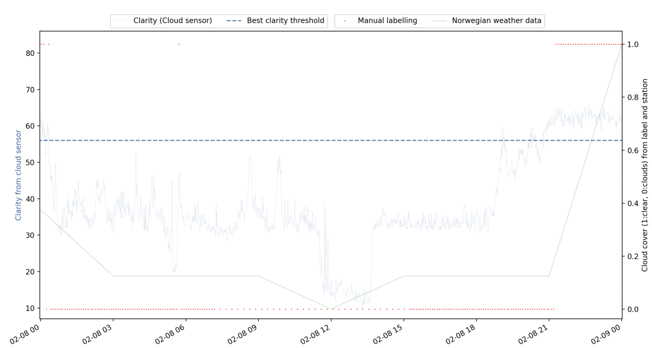


**Figure 11.** Time series of the clarity (cloud sensor) in blue with the best threshold clarity value (56) in line, labelling category (manually) in red and cloud cover index (Norwegian Centre for Climate) in green for the day 08/02/2020

Other zooms into the main plot can be seen in the Extended figures part (Figure 23 and Figure 24), for different time scales, in order to appreciate more the consistency between the different time series.

## 5. Testing

Manual labelled images are also available for January and February 2019 which makes it a valuable testing data set for the best clarity threshold determined for November 2019 to February 2020. The histograms for the testing data set (Figure 12) present similar characteristics with a bimodal distribution but with more cloudy skies than clear skies compared to November 2019 to February 2020.

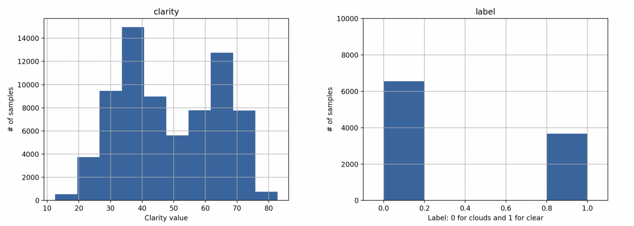


**Figure 12.** Clarity (on the left) and label (on the right) histograms (January and February 2019)

The time series of clarity, labels and cloud cover index for January and February 2019 are also plot (Figure 13).

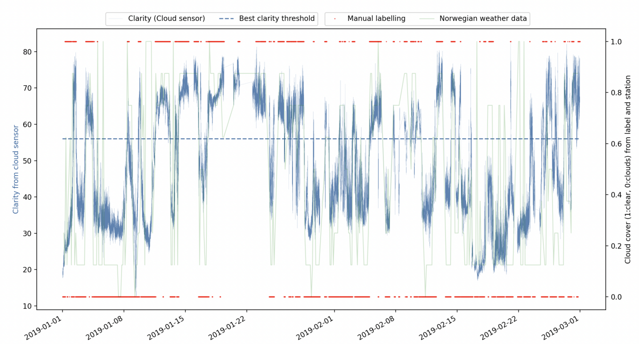


**Figure 13.** Time series of the clarity (cloud sensor) in blue with the threshold clarity value (56) in line, labelling category (manually) in red and cloud cover index (Norwegian Centre for Climate) in green for January and February 2019.

The thresholding with the 56 clarity value is then applied to the testing data set and a consistency of about 91 % is obtained between clarity and label. This value is high and very close to the 90.1 % obtained for November 2019 to February 2020. Moreover, 56 is also the best clarity thresholding value possible over the testing period (Figure 14).

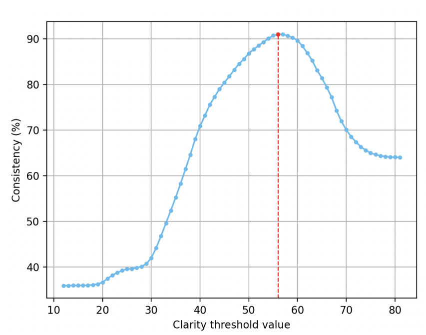


**Figure 14.** Consistency between the clartiy after thresholding and the labelled images for each threshold value of clarity (January and February 2019).

## Timetrend

The cloud sensor data may be affected by trend over time which can affect the robustness of the thresholding value. For example, the rising of average temperatures, which is most intense in the Arctic (ref?), should affect the calculation of the clarity value because it is directly linked to temperature. In this case, the threshold may be less accurate for time-distant data from 2019/2020 and frequent calibration with labelled data may be needed.

Timeseries of clarity are available between January 2016 and February 2024 (Figure 15). The data is affected by gap and drop-outs at certain times. However, a linear regression of the whole period is made in order to identify any timetrend. The coefficient of determination of these cyclic data is of course low ($R^2 \approx 0.00685$) and a linear model is not meant to fit it. The slope of the clarity linear regression is about -0,0014301 per day, which represents a reduction of $\approx 4.18$ in clarity between 2016 and 2024 (8 years). It appears that there is a general tendency over year with a reduction of the clarity which means either high altitude atmosphere temperature $T_{sky}$ higher or ground air temperature $T_{air}$ lower according to the definition of clarity.

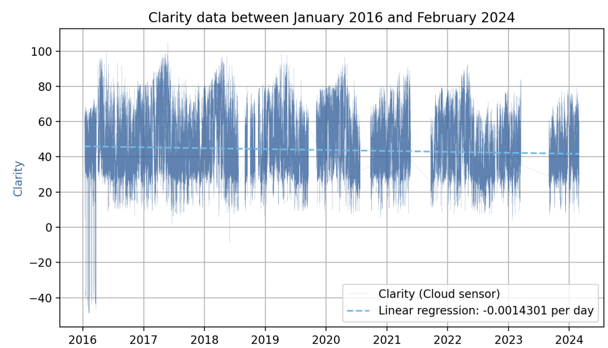


**Figure 15.** Time series of the clarity between January 2016 and February 2024

The thresholding method has been applied to winter months while the linear regression has been made over the whole data including all seasons and outliers (gaps and drop-outs), so a closer look is needed to identify a timetrend affecting the quality of our threshold. Only the clarity from November to February for each year between 2016 and 2024 are represented on the Figure 16. Two time series has gaps (2018/2019 and 2023/2024 in red) and the mean for each time series is represented with a dashed line. The range of the values is similar each year (aproximately 20 to 80) even if it seems to get lower for the recent years (2022/2023 and 2023/2024). The mean value of clarity for each winter varies between 41.78 and 53.14

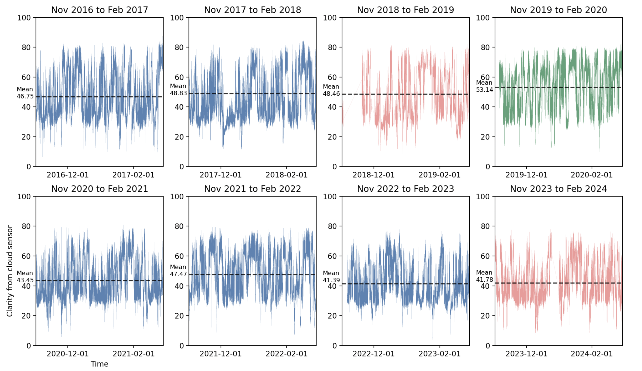


**Figure 16.** Timeseries of the clarity from November to February for each year between 2016 and 2024 (red time series contains gap and green time series was used for the thresholding method)

The histograms of the clarity from November to February for each year between 2016 and 2024 (Figure 17) show a bimodal distribution. The mean value of clarity for each winter is always between the two modes of the histogram distribution. Most of the winters has more cloudy than clear samples

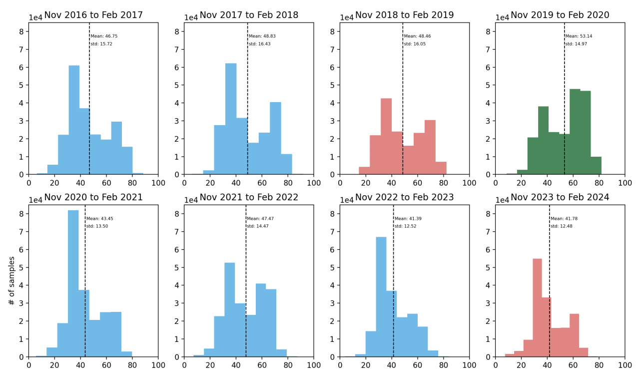


**Figure 17.** Histograms of the clarity from November to February for each year between 2016 and 2024

## Sampling

Determining the clarity threshold involved the use of more than 25 000 manually labelled images. This section looks at the possibility of using less data to determine this threshold value. And how much accuracy may be lost by using less data. This question is particularly important if one wants to determine a threshold for another period of the year, for example summer, and thinks about how many labelled data may be needed.

The histogram (Figure 18) of the labelled images from November 2019 to February 2020 shows again that there is 51.55 % of cloudy labels and then 48.45 % of clear labels.

Eight random sampling of size 500 times smaller (≈ 50 samples of labelled images) among all the labelled images are realized. The cloudy labels in these eight samples varies between 43.14% to 60.78%. The histograms of the eight random sampling are on the Figure 19.

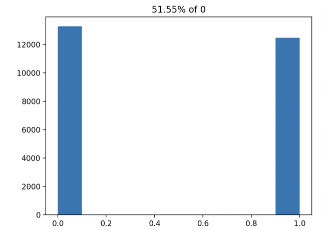


**Figure 18.** Histogram of manual labelled images (November 2019 to February 2020)

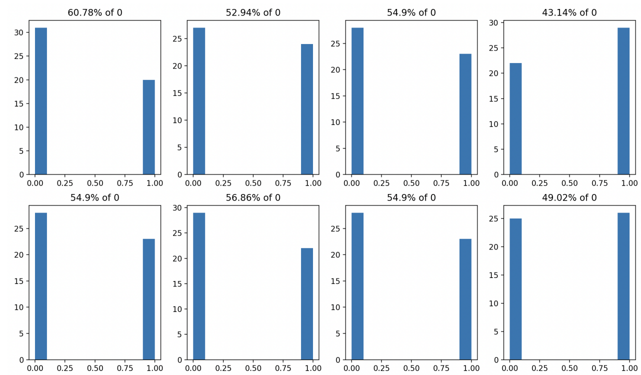


**Figure 19.** Histograms of the eight random sampling (500 times smaller) of the manually labelled images

Then, the thresholding method is applied for each of the eight sampling and the results are shown on the Figure 20. For each of the samples, the maximum of consistency is near the best threshold value determined for the whole labelled images (clarity value of 56). It means that even with a few labelling, it appears that it is still possible to get a fairly good thresholding value of clarity.

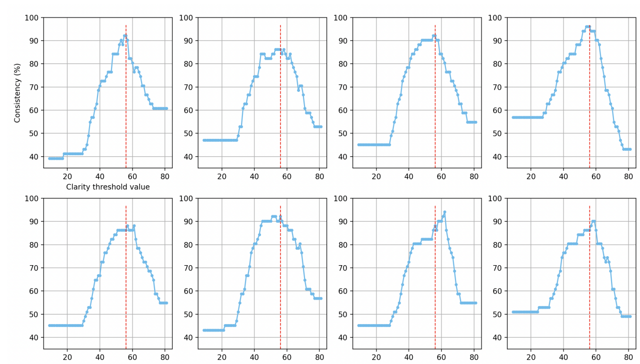


**Figure 20.** Consistency between the clartiy after thresholding and the labelled images for each threshold value of clarity for each of the eight random sampling (November 2019 to February 2020).

## References

Aurora Eurotech, 2009. Aurora Cloud Sensor User Guide.

Norwegian Centre for Climate Services [online] https://seklima.met.no (consultated on February 13).

## Extended figures

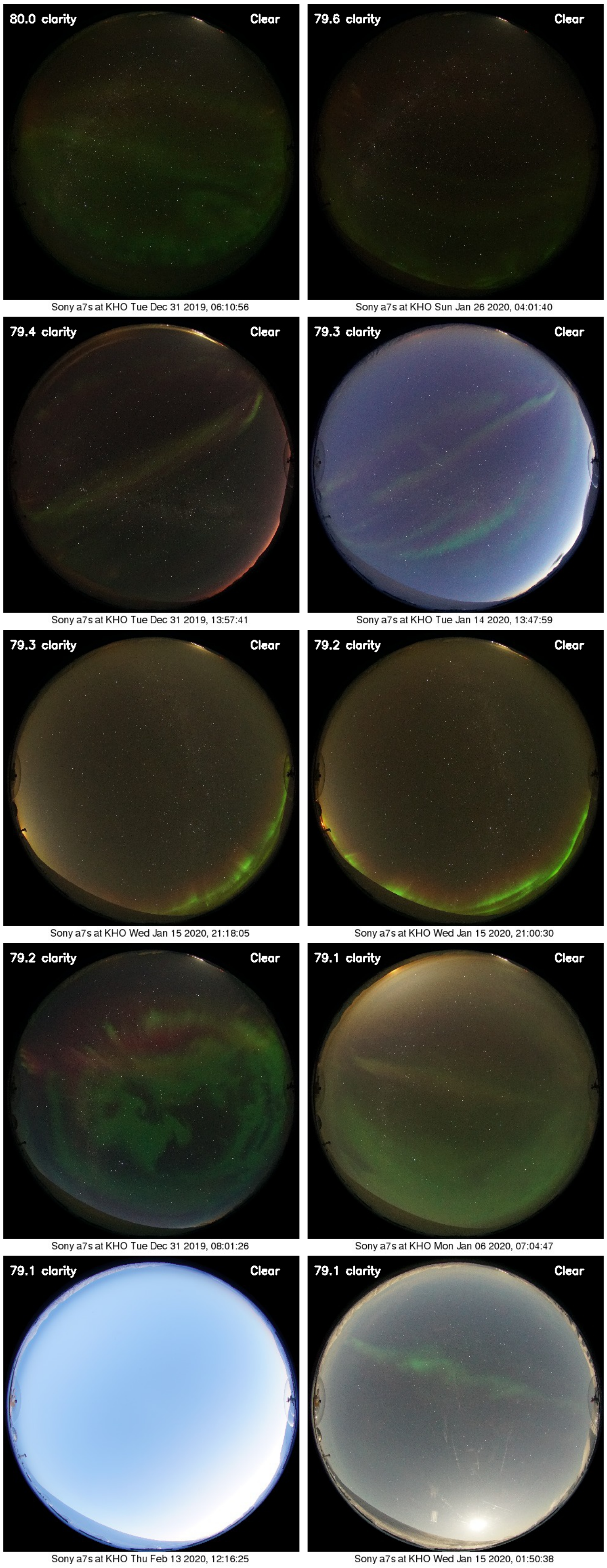


**Figure 21.** All sky images corresponding to the highest values of clarity from the cloud sensor

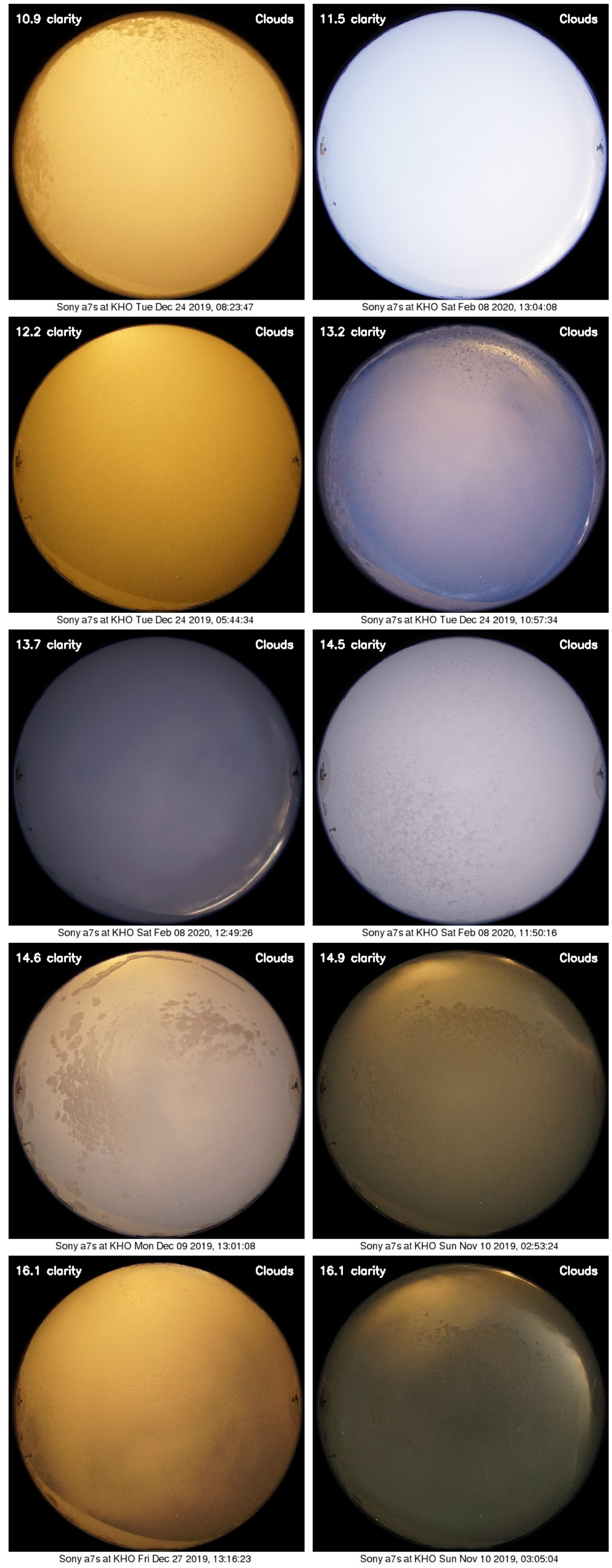


**Figure 22.** All sky images corresponding to the lowest values of clarity from the cloud sensor

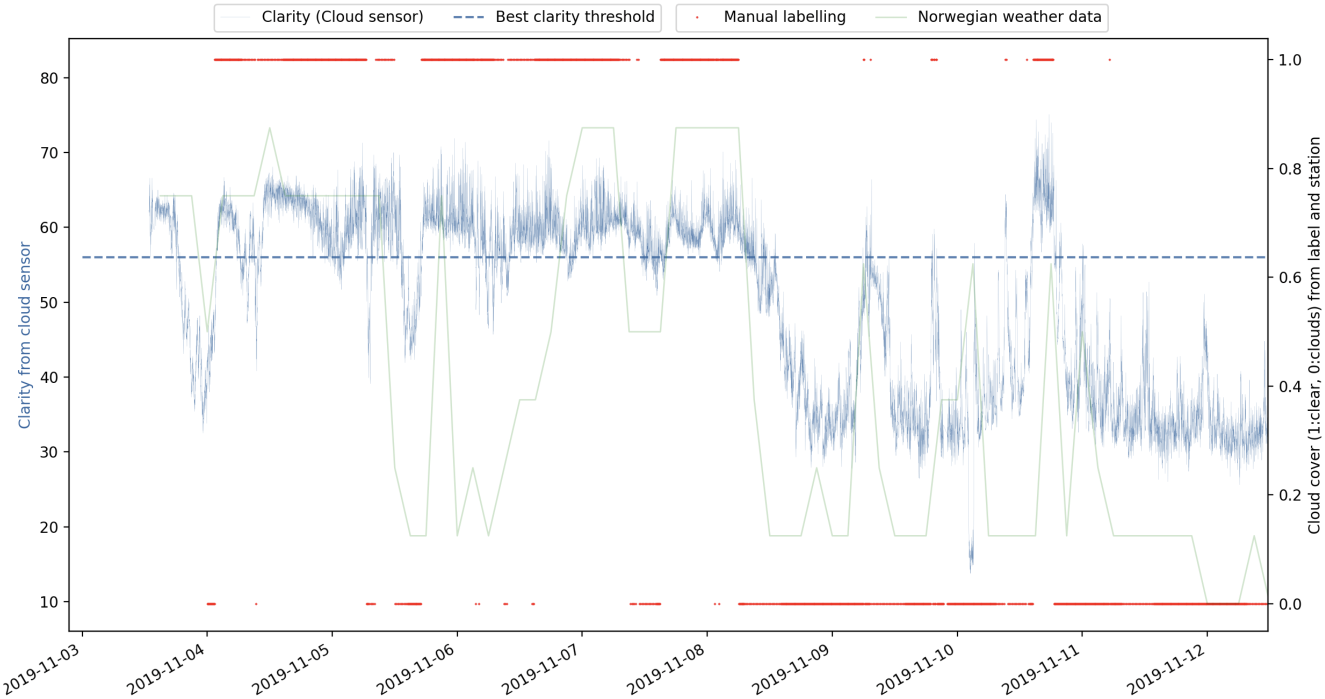


**Figure 23.** Time series of the clarity (cloud sensor) in blue with the best threshold clarity value (56) in line, labelling category (manually) in red and cloud cover index (Norwegian Centre for Climate) in green between the 3th and the 12th November 2019.

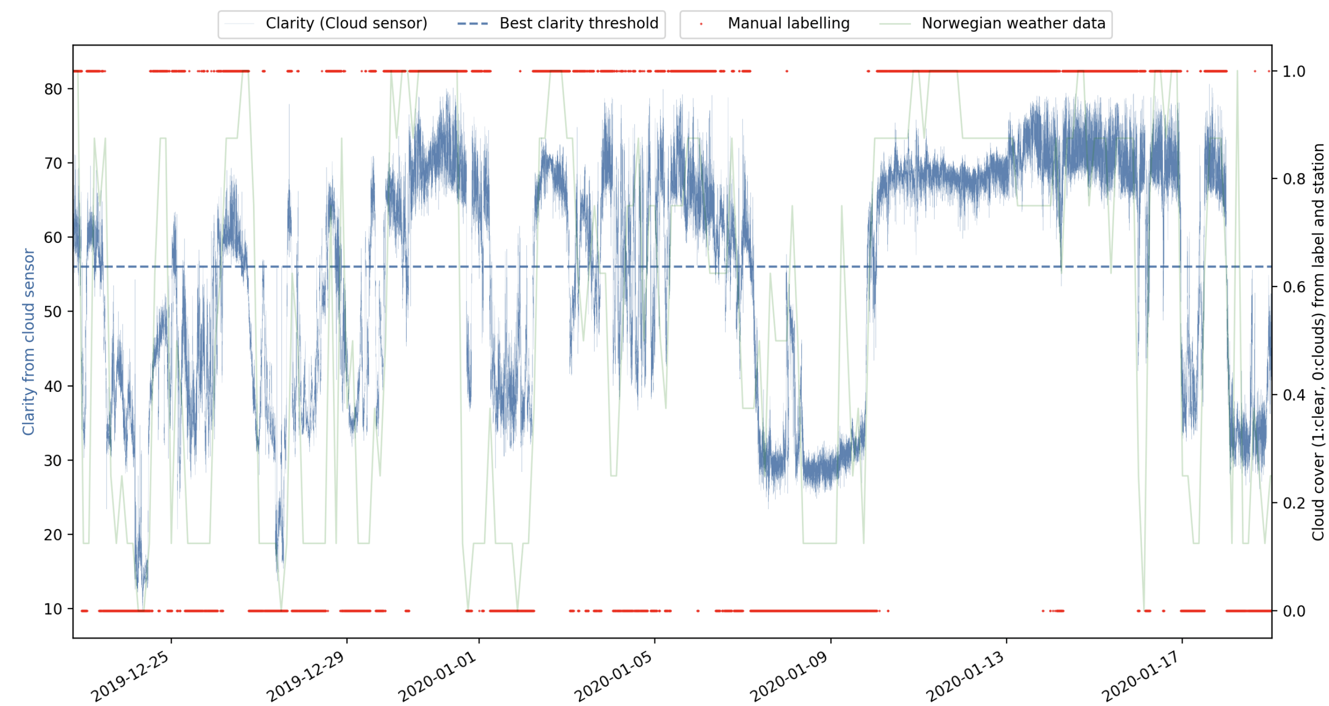


**Figure 24.** Time series of the clarity (cloud sensor) in blue with the best threshold clarity value (56) in line, labelling category (manually) in red and cloud cover index (Norwegian Centre for Climate) in green between the 3th and the 25 December 2019 and the 17 January 2020.