Referee #1

This is an interesting paper on cloud elimination or cloud flagging (CF) techniques used for sun-photometric measurements. The paper contributes towards improving such measurements. It fits the scope of the journal.

Introduction
Before starting describing the CF methods I would start the paper with some sentences like the paragraph below to show the importance of this study in combination with AOD related research and use:
AOD is the most comprehensive aerosol parameter for radiative forcing studies. Surface based AOD measurements are conducted from various surface base networks (e.g. aeronet, gaw-pfr, skynet) (e.g. Holbern et al., Nakajima et al., 2020 AMT). The series are used for local, short term or long term aerosol studies and for satellite validation. One of the main problems of such measurements is the fact that they can not be conducted under cloudy conditions at least when present in the detector-sun path of photons. For that case there are algorithms that are used in order to eliminate the possibility of cloud-present measurements to be included in the AOD data series. Such algorithms contribute substantially to the quality of AOD data worldwide.

Agreed, thank you for the suggestion.
The first paragraph (lines 9-16) now reads
Sun photometry is one of the longest employed and robust measurement techniques for total column aerosol optical depth (AOD) retrieval (Holben et al., 1998, 2001). AOD is the most comprehensive aerosol parameter for radiative forcing studies and serves as the ground-truth for validation of satellite data. Various surface based networks such as AERONET (Holben et al., 2001), SKYNET (Takamura and Nakajima, 2004), and the WMO global atmosphere watch programme GAW-PFR (Kazadzis et al., 2018) conduct measurements of AOD at a high time resolution, providing data for local short and long term aerosol studies. As the calculation of AOD from photometer measurements is based on the assumption of a cloud-free path between instrument and sun, the identification and removal of cloud contaminated data is one of the most important prerequisites for high qualitiy AOD data.

Authors use the PFR instrument in their analysis but the introduction is mainly for aeronet/cimel. The latest publication describing the PFR CF algorithm can be found here: https://gi.copernicus.org/articles/7/39/2018/
Also it is essential to mention the PFR algorithm more explicitly as it differs in some aspects from the aeronet. In general different networks are using slightly or more different algorithms for CF. For example in https://acp.copernicus.org/articles/18/3185/2018/
there is a comparison of different CF algorithms at synchronous AOD measurements from different instruments/networks.

Added *Kazadzis et al., 2018* (line 19).
Added *For our site, we use a limit of 0.02 if AOD is lower than 0.2, otherwise 0.03* (line 70)

In general the methodology is based on one minute data as derived from PFR. If the authors want to generalize the method being important for other AOD measuring networks some discussion on the measurement frequency vs method quality has to be presented. This is because for most other than PFR instruments, measurements are more than 1 minute apart increasing the possibility of cloud contamination in N consecutive measurements.

This is a good point.
We’ve added an analysis of subsampled data (lines 163-172) as well as two figures (7 and A2) for

illustration.
Finally, we investigate the performance of the proposed method at lower time resolution. We subsampled the timeseries to 5, 10, and 15 minutes and analyzed the resulting data with the same settings for the algorithm. Key parameters of the resulting d20 distribution from the whole 10-year data record are shown in Figure 7. The coarser the time resolution, the higher minimum and median d20 values and interquartile range. This is mainly due to the overall density of points decreasing, thus increasing the mean distance to the nearest neighbours. Furthermore, the values of the first time derivative will be lower, so the relative weight of this dimension decreases.
To account for the changes in density, there are two possible adjustments: lower the number of nearest neighbours, or set a higher cloud flagging threshold. As an example of the latter, we show the timeseries on 12th of March 2020 (same as in Figure 3) at four different time resolutions in Figure A2. With increasing the d20 threshold to 0.019, 0.027, and 0.042 respectively for 5, 10, and 15 minute resolution, a very similar flagging behaviour can be achieved at all time resolutions.

The Angstrom parameter a: It is a very good proxy for cloud flagging. However its variability depends also in AOD. Low AOD measurement days lead to much more "sensitive" and variable angstrom a than the ones with higher AOD. In low AOD days small but real AOD variability in combination with AOD measurement uncertainties can lead to high angstrom fluctuations.
Dust AOD variability could be an issue. Cuevas et al., discuss the 1 minute variability https://amt.copernicus.org/articles/12/4309/2019/amt-12-4309-2019.pdf in this paper and the supplement. The example presented here is a good example but probably some discussion should be included based on the above mentioned 10 year time series of AOD cases.

Indeed, the variability in the Angstrom parameters increases with decreasing AOD. But since the Angstrom parameters are only 2 out of 4 dimensions and the variation of AOD in these cases is also lower, so the algorithm is relatively robust for these low AOD scenarios. We have included examples in the appendix of low and high AOD days, as well as a day with Saharan dust and hence high AOD in Figure 3.

Main effect of the non correct CF in an AOD series is the data cloud "contamination"that leads to a systematic higher instant AOD values but also affects daily, monthly AOD averages. Such systematic effects could have an impact on long term series statistics and much more to trend analysis of AOD related changes. I think this could be mentioned in the conclusions. It is an aspect that methodologies such as the paper presents, contribute towards better quality results.

Yes, thank you for the suggestion. However, it should be noted that there is a (maybe counterintuitive) possibility that the daily mean AOD can also increase when more thin clouds are flagged (if there is a real change in AOD over the day, but thin clouds only occur during the period of lower AOD).
Added (lines 181-184)
[…] optically thin clouds, such as cirrus and contrails, which lead to systematic bias of higher instant AOD values. Reducing this bias contributes to an improvement of long term statistics and trend analysis of aerosol conditions.

Referee #2

The manuscript describes a cloud flagging algorithm for a PFR radiometer deployed in Innsbruck. It compares a traditional filtering based on temporal AOD variations with the proposed method based on clustering. The approach is interesting and fits the scope of the journal. However the methods are insufficiently described, and the applicability is questionable or at least not supported by the information provided in the paper. I recommend a major revision is made before it can be published.

General comments

1. The introduction needs to be improved, with more detailed description of the GAW-PFR network and related references. Other cloud flagging algorithms could be mentioned, as well as the main principles of them (temporal AOD variations, spectral changes/Angstrom exponent, etc.). The AOD spectral derivatives and curvature used later, suggest including reference to O'Neill papers on this matter. Overall, references are missing to support certain statements throughout the text (see specific comments).

Added *Takamura and Nakajima, 2004*, *Kazadzis et al.,2018* (line 13), and *O'Neill et al., 2001, 2003* (line 57)

2. The "multiplet" method is not at all described (lines 60-65). First, the authors should provide how many measurements in what time interval are analyzed. And they should also provide the threshold value for cloud flagging. Second, the Smirnov (2000) algorithm is much more complex than the "triplet" criterion. Several other steps are given(on the second temporal AOD derivative, on the AOD standard deviation for the entire day, etc.). The description provided by the authors is too simplified. You need to be much more specific and detailed.

Added (lines 71-73)
Further limits are set on the standard deviation of AOD within a day and the second time derivative of the timeseries. Out of these parameters, the multiplet criterion is the most relevant (more than 99% of the flagged points) [...]

3. The applicability of the method is highly questionable, for two reasons. First, the threshold value used to flag the data (0.012 for d20) seems to be based on the analysis of few days only. This is not acceptable in my view. Many circumstances, simply other AOD level or other aerosol type, can change the variability of the AOD or the Angstrom exponent. This is not considered at all in the discussion. I would recommend analyzing the four variables and the d20 distance in different situations (high and low AOD, different aerosol types, season, etc.) to derive a more robust criterion.

This is a justified concern which we share, but which is at the very root of the problem: to find the best threshold in a continuous distribution of d20 values (not a binary one clearly distinguishing between cloudy and clear) such that as many cloudy and as little cloud free values as possible are flagged. We used a number of clear sky days (about 150) as a starting point for a reference value and indeed fine tuned / checked according to benchmark days for different situations (high and low AOD, different aerosol types, season, etc.), see figures 3 and A1.
Even a human observer will have to decide at which point in the transition he deems the sun obstructed when a cloud is moving. Furthermore, the real positives/negatives, especially in this grey zone, are unknown – this is the reason this kind of algorithm is needed in the first place.
While we would also prefer a more objective and deterministic way of choosing the threshold, we believe that in principle there is no real alternative here to improve the threshold determination.

Second, this method is restricted to one site. It performs apparently well in Innsbruck, but the authors should not pretend that it can be expanded as is to other sites (or instruments). We have no idea how it would perform in an Amazonian site with highly variable biomass burning aerosol, for instance. Claiming that the method can improve the Smirnov (2000) algorithm (line 145) in cloud flagging is too much to say. Suggesting that it can work as good as Giles (2018) with less input information (line 146), is also too much to say. Those algorithms have been tested over a huge database and thoroughly describe the difficulties and the compromise that needs to be taken in an operational algorithm.

There is no reason why the algorithm should not perform at sites with similar aerosol conditions (Europe/North America). For more dissimilar sites it depends whether the density of points in clear conditions is dissimilar enough from cloudy conditions. The volcano example is the closest we'd have to such conditions and does not suggest that the algorithm cannot be applied here.
We do not claim to improve Smirnov, but employ a completely different algorithm to solve an open issue in the time-series analysis, which is detection of thin clouds.
Aureole scan provides admittedly more information, but is not available everywhere, and takes a lot of time. No comparison with the performance of Giles (2018) can be done, and we did not intend to suggest a better performance.

First paragraph in the conclusions now reads (lines 174-179)
We presented a new approach for flagging cloud contaminated data points from sun photometer measurements by treating them as outliers/region of low density in a four-dimensional space. Our routine only needs one semi-empirically derived threshold and direct sun measurements for assigning a cloud flag. The method tackles shortcomings of the currently employed routine based on Smirnov et al. (2000) in the presence of optically thin clouds, such as cirrus and contrails, which lead to systematic bias of higher instant AOD values. Reducing this bias contributes to an improvement of long term statistics and trend analysis of aerosol conditions.
Last paragraph now reads (lines 203-209):
So far we have tested the algorithm only for our instrument in Innsbruck. It performs well in different cloud and aerosol conditions, as shown in figure 4, and is able to alleviate AOD bias in the presence of thin clouds. For the application at other 205 measurement sites, the time resolution of the data needs to be considered, as lower measurement frequency leads to lower data density, and therefore higher mean distances between points (figures A2 and 7). Nonetheless, adaptations regarding the number of nearest neighbours, the relative weight of the different dimensions, or the d20 threshold can be easily done to optimize cloud detection with other instruments as well.

4. The validation of the method needs independent data, such as all-sky camera, lidar, ceilometer, solar radiation or even a human observer. The all-sky camera is mentioned in the paper but nothing is shown and no systematic analysis seems to be used for analyzing the performance of the clustering routine. The fact that more data points are removed is not sufficient. This can be easily achieved by any algorithm by using stricter thresholds.
The problem with all the suggested independent data is that the frequency of measurements is much lower, and they do not necessarily measure in the same path as the sun photometer. Furthermore, there are uncertainties in distinguishing cloud-contaminated from cloud-free datapoints as well, so it cannot serve as reference in a systematic analysis. However, clouds have clear spectral signature (as seen in the Angstrom parameter plots) which can be determined with sun photometer data.
We have added Figure 2.

Yes, stricter thresholds will lead to removal of more points, shifting the mistakes from false negatives to false positives, but our algorithm does not merely remove more, but also different datapoints, as seen in Figure 4. Again, while some datapoints can be identified as clear/cloudy without a doubt, there are many in the grey transition zone of a cloud forming or passing in the path. Where the cut is set exactly in the continuous transition depends on the kind of error one prefers to avoid.
Added (lines 99-104)
As can be seen from the d20 timeseries in figure 3, there is no clear distinction of two states (cloudy/clear), but rather a continuous spectrum of values that has to be divided to best fit the two categories. A lower threshold value will classify the ambiguous points as cloudy, but also risks a higher number of false positives and therefore lower overall data retention, which matters for error of mean values calculated from the data. Similarly, a higher threshold will cause more false

negatives, i.e. cloud contaminated points to be identified as clear. We set the d20 threshold to 0.012 considering these aspects on our clear reference and benchmark days.

Specific comments

Line 22: good place to add more information and references about GAW network (Wehrli, Kazadzis), and maybe Skynet cloud-screening (Takamura).

Done (line 12)

Line 23: do you think your algorithm could be used for lunar measurements too? This is emerging technique and worth mentioning, since your algorithm does not require any additional measurement, only AOD.

That is an interesting point. There is no obvious reason it should not work. However, it sounds like even more of a stretch than just applying it to different measurement sites.

Line 25: what happens if there are cirrus or contrails during most of the day? What happens if there is a change of aerosol type during the day?

Changed to *The main idea is that aerosol optical depth and microphysical properties [...] show little and slow variation within a day, while clouds introduce outliers and stronger fluctuations in these parameters.*

Line 51: I guess you use all 4 wavelengths, but please specify.

Done (line 56)

Line 52: this a good place to cite O'Neill papers on AOD spectral derivatives.

Done (line 57)

Line 61: as mentioned in the general comments, Smirnov's paper includes a long list of criteria. Please be more specific.

Added (line 72)

Further limits are set on the standard deviation of AOD within a day and the second time derivative of the timeseries.

Line 63: please provide the time interval for the multiplet and total number of measurements. The PFR raw acquisition time is 2 seconds if I remember correctly. Do you use 1 minute averages, standard deviations, how do you look at fluctuations over de 5 minutes period? Please provide details, including the threshold used for cloud flagging.

Added (line 41)

with an acquisition time of about 2 seconds

Line 74: the division by 10 to make the parameters comparable in magnitude is somewhat arbitrary. This is fine if it works, but couldn't you try some kind of mathematical normalization?

There are two common normalization strategies:

1) By mean and standard deviation

Given that AOD (nor the other parameters) is not by any means normally distributed, this is hardly applicable. There is not even consensus whether to use arithmetic or geometric means when it comes to averaging AOD over 1 day/month…

2) Mapping onto a [0,1] interval

This would kind of defeat the purpose of finding outliers with our method as it changes the density of the data. Of course it can be done, when using it on the whole timeseries – ultimately this means multiplying by a fixed factor (to have intervals of the same length, it is not necessary to shift the lowest value to 0 as only distances are analyzed), which is what we already are doing. This can be done, but should not significantly alter the outcome.

The division by 10 can be seen as a weighting function rather than a normalization of the different dimensions.

Added (line 82)

and therefore of equal weight in calculating the distance,

Line 74-75: the sentence about the finite difference units is hard to understand. Please rewrite.

Added (line 84)

(i.e. the value is divided by 12 when t is in hours)

Line 80: do you mean high AOD ó high number of data?
Changed (line 90)
retention of a high number of data points
Lines 81-84: the procedure to derive the d20 threshold seems arbitrary ("...is further fine tuned using days on which the Multiplet routine fails...") and is not explained at all. The paper must describe this method in detail so that others can reproduce it. Moreover, as explained in the general comment 3, few clear days are in my view not sufficient to derive the threshold. A robust statistics over a large sample of data would be desirable. The description should also reveal the difficulties (too strict threshold removes too many good data, too loose threshold allows too many cloud-contaminated data...and so on). And such analysis requires ancillary information to assess what data points are clear and which are cloud contaminated.
See general comment #2 and #4.
Added (lines 99-104)
As can be seen from the d20 timeseries in figure 3, there is no clear distinction of two states (cloudy/clear), but rather a continuous spectrum of values that has to be divided to best fit the two categories. A lower threshold value will classify the ambiguous points as cloudy, but also risks a higher number of false positives and therefore lower overall data retention, which matters for error of mean values calculated from the data. Similarly, a higher threshold will cause more false negatives, i.e. cloud contaminated points to be identified as clear. We set the d20 threshold to 0.012 considering these aspects on our clear reference and benchmark days.
Line 94: the airmass limitation was not mentioned until this point. It is one of the criteria used in Smirnov's algorithm. As explained above, you need to be more specific and detailed in describing the algorithms.
Added (line 67-68)
Datapoints for which the airmass exceeds a value of 6 are considered as cloudy.
Figures 2 and 3: please enlarge fonts in the axe's labels.
Done
Figure 3 caption: better say "solid line" than "black line" (the lines are actually in whitecolor)
Done
Line 110: maybe add reference about Saharan dust over Austria
Done (line 130)
Line 114: add reference about Eija volcanic eruption over Austria.
Done (line 134)
Line 118: what criteria or ancillary data are used for a manual screening? Note that readers need to know how to reproduce your results.
As it is inspected by a human, the criteria cannot be put down as specific numbers. Furthermore, there was no manual screening done, it was just mentioned that it would be possible.
Line 121: maybe add reference about sunshine hours
Done (line 142)
Line 127: wrong subscript, the longest wavelength is 870nm, I guess.
Done (line 149)
Line 137: 0.001 is the uncertainty is for the reference instruments if Langleys are made at Mauna Loa (Toledano 2018; Kazadzis 2018). For a field instrument (side-to-sidecalibration), the uncertainty is rather 0.5-1% in calibration constant or 0.005-0.01 in AOD.
Done (line 158)
Line 145: the data shown in the paper do not support this strong assessment in the conclusions.
See general comment #3
Line 162: the sentence about the results in high latitude sites is speculative and not supported by any data in the manuscript.
See general comment #3
We removed the mentioned high latitutes.
Line 165: please rewrite the sentence, it's hard to understand.

Now reads (lines 196-198)

Nonetheless, it can be used for real-time analysis as well, given that erroneously cloudy points can be corrected to clear when more data becomes available, but not the other way round (i.e. points identified as clear once will be labelled as clear regardless of additional measurements).

Line 170: “association”: do you mean correlation, covariance, ..?

Mutual conditional information (details in reference) as mentioned in the sentence (line 200)

Reducing cloud contamination in AOD measurements

Verena Schenzinger[1] and Axel Kreuter[1, 2]

[1]Institute for Biomedical Physics, Medical University Innsbruck, Innsbruck, Austria
[2]Luftblick OG, Innsbruck, Austria

Correspondence: Verena Schenzinger (verena.schenzinger@i-med.ac.at)

Abstract. We propose a new cloud screening method for sun photometry that is designed to effectively filter thin clouds. Our method is based on a k-nearest neighbour algorithm instead of scanning timeseries of aerosol optical depth. Using ten years of data from a precision filter radiometer in Innsbruck, we compare our new method and the currently employed screening technique. We exemplify the performance of the two routines in different cloud conditions. While both algorithms agree on the classification of a datapoint as clear or cloudy in a majority of the cases, the new routine is found to be more effective in flagging thin clouds. We conclude that this simple method can serve as a valid alternative for cloud detection, and discuss the generalizability to other observation sites.

1 Introduction

Sun photometry is one of the longest employed and robust measurement techniques for total column aerosol optical depth (AOD) retrieval (Holben et al., 1998, 2001). AOD is the most comprehensive aerosol parameter for radiative forcing studies and serves as the ground-truth for validation of satellite data. ~~As it~~ Various surface based networks such as AERONET (Holben et al., 2001), SKYNET (Takamura and Nakajima, 2004), and the WMO global atmosphere watch programme GAW-PFR (Kazadzis et al., 2018) conduct measurements of AOD at a high time resolution, providing data for local short and long term aerosol studies. As the calculation of AOD from photometer measurements is based on the assumption of a cloud-free path between instrument and sun, the identification and removal of cloud contaminated data is one of the most important prerequisites for high qualitiy AOD data.

The most widely used algorithm by Smirnov et al. (2000) sets a threshold on the temporal variation of AOD, assuming a higher variation in the presence of clouds~~.~~, as one of its flagging criteria. It was developed and employed in the AERONET network, as well as adapted for GAW-PFR (Kazadzis et al., 2018). While this method reliably flags thick clouds, ~~there is a limit to detect~~ detection of optically thin clouds exhibiting small AOD changes, i.e. below the threshold, is not possible. This limitation introduces a bias towards higher AOD (Chew et al., 2011; Huang et al., 2011), as well as a bias in the Angstrom parameters, which indicate particle size.

To remedy this problem without additional manual quality control, Giles et al. (2018) revised the Smirnov et al. (2000) algorithm (see table 2 of Giles et al. (2018) for specifics). They include aureole scans in their cloud screening routine which utilize the increased forward scattering behaviour of thin clouds for their identification. This is suitable for instruments that measure sky radiance in addition to direct sun. However, the procedure takes time and is only viable within AERONET who employ these instruments, whereas it is not applicable for precision filter radiometers operated within other networks, such as ~~the WMO global atmosphere watch programme (GAW)~~GAW-PFR.

Therefore, we developed a new algorithm that can identify thin clouds, and works with direct sun measurements only. The main idea is that aerosol optical depth and microphysical properties (represented by the Angstrom parameters, Gobbi et al. (2007)) ~~stay roughly constant~~ show little and slow variation within a day, while clouds introduce outliers and stronger fluctuations in these parameters.

Instead of scanning the timeseries of these variables, we examine their density in a four-dimensional space with a k-nearest-neighbours algorithm. This principle is well established in the fields of machine learning and data mining as an efficient way to identify outliers in data (Ramaswamy et al., 2000). In the context of AOD measurements, clear sky will lead to regions of high density/little distance between points, whereas clouds will result in less dense regions/outliers.

2 Data and Methods

2.1 Instrument and Raw Data

We use a precision filter radiometer (PFR) developed by the PMOD WRC Davos for the GAW network (Wehrli, 2005) with four channels (368nm, 412nm, 501nm, 864nm) and a field of view of 1.2°. The instrument is set up on top of a ten-storey university building in Innsbruck, Austria (47° 15’ N, 11° 24’ E). Our operational guidelines are based on the ones of GAW and the instrument is calibrated by PMOD, but our site runs independently of the network. Additional to the minutely PFR reading with an acquisition time of about 2 seconds, we measure the air temperature and pressure at the site and monitor the overall cloud conditions with an all-sky camera taking pictures every 10 minutes.

2.2 Processing and filtering

First steps in quality control include the removal of data points where any of the four voltages is negative. Furthermore, flags are introduced if the sun tracker records a value higher than 15”, and an ambient temperature above 310 K.

After the initial filtering, aerosol optical depth (AOD) is calculated from the voltage measurements. Starting from the Lambert-Beer law, the aerosol optical depth is derived as:

$$\tau_a(\lambda) = m_a^{-1}\left[\ln\left(\frac{V_0(\lambda)}{V(\lambda)R^2}\right) - \sum_o m_o \tau_o(\lambda)\right] \quad (1)$$

where $V(\lambda)$ are the measurements, $V_0(\lambda)$ the calibration factors, R the Sun-Earth distance, and m the airmass in the path between instrument and sun.

Several atmospheric constituents contribute to the optical depth, one of which are aerosols (indicated by subscript a). Other factors (subscript o) which are taken into account in our calculation are Rayleigh scattering (Kasten and Young, 1989; Bodhaine et al., 1999), Ozone (Komhyr et al., 1989), and NO_2 (Valks et al., 2011). We use climatological values for O_3 and NO_2, and temperature and pressure measured on site. Resulting unphysical values (negative or infinite) of the aerosol optical depth at any wavelength are discarded.

At each timestep, we perform a linear and a quadratic fit (indicated with subscripts l and q respectively) to $\tau_a(\lambda)$ at all 4 wavelengths to derive the Angstrom parameters ~~(Ångström, 1929, 1964; King and Byrne, 1976)~~(Ångström, 1929, 1964; King and Byrne, 1 .

$$\ln\left(\tau_a(\lambda)\right) = \ln(\beta_l) - \alpha_l \ln(\lambda) \tag{2}$$

$$\ln\left(\tau_a(\lambda)\right) = \ln(\beta_q) - \alpha_q \ln(\lambda) + \gamma_q \ln(\lambda)^2 \tag{3}$$

The spectral slope α_l of the linear fit and the spectral curvature γ_q of the quadratic fit are used in further analysis and referred to without the subscript hereafter.

2.3 Cloud flagging

The next step in quality control of the data is the flagging of potentially cloud-contaminated datapoints. Figure 1 shows the basic principle of the presently employed scheme and the proposed new method.

Currently, our operational routine is based on the criteria laid out in Smirnov et al. (2000), with some minor adaptions due to the higher measurement frequency according to Wuttke et al. (2012). ~~The~~ Datapoints for which the airmass exceeds a value of 6 are considered as cloudy. For lower airmass, the main criterion for filtering datapoints is the difference between the maximum and minimum AOD value within a multiplet of consecutive datapoints (Smirnov et al. (2000) uses a triplet, whereas we look at a quintuplet) which cannot exceed a set value. For our site, we use a limit of 0.02 if AOD is lower than 0.2, otherwise 0.03. This threshold is balanced to filter clouds while retaining real AOD variations. ~~As this~~ Further limits are set on the standard deviation of AOD within a day and the second time derivative of the timeseries. Out of these parameters, the multiplet criterion is the most relevant ~~for flagging,~~ (more than 99% of the flagged points), so we will refer to ~~it~~ the currently employed method as "Multiplet routine" hereafter.

Instead of step-wise scanning timeseries, our new routine performs one calculation for all currently available datapoints. We use a k-nearest neighbours algorithm to establish the 20 closest points $\{P_1, P_2, ..., P_{20}\}$ for each of our measurements P_0 in a four dimensional space. Then the mean euclidian distance between P_0 and its neighbours is calculated (referred to as d_{20}) and P_0 is identified as cloudy if this distance exceeds a threshold. This method is usually used to identify clusters of data points, hence we will call it "Clustering routine".

The dimensions used are the aerosol optical depth at 501nm, its first derivative with respect to time, and the two Angstrom parameters α and γ. The first two cover temporal variations of one wavelength, the second two changes in the spectrum. To ensure that these parameters are comparable in order of magnitude, ~~the~~ and therefore of equal weight in calculating the distance, the Angstrom parameters derived from equations 2 and 3 are divided by a factor of 10. Furthermore, the finite-difference time derivative of the AOD $\frac{\Delta\tau}{\Delta t}$ is used in units of 1 per 5 minutes (i.e. the value is divided by 12 when t is in hours), analoguous to checking the AOD variation within a quintuplet of minutely measurements.

Points affected by a track error will not be considered in the set of possible nearest neighbours. If this leads to less than 20 ~~measurements are valid~~ valid measurements, the number of nearest neighbours k will be reduced accordingly down to a minimum value of 5 points in real-time analysis. To account for the lower number of nearest neighbours, the calculated distance is then multiplied by $\frac{20}{k}$ to make it comparable to the original d_{20} measure. Similarly, if the number of datapoints identified as clear on one particular day is lower than 30, the Clustering routine is re-run with 10 nearest neighbours during post-processing to ensure ~~high data retention~~retention of a high number of data points.

To establish the threshold for possible cloud contamination, we calculate the distribution of d_{20} on about 150 clear days. We estimate a limit from this continuous distribution, which is further fine tuned ~~using days on which the Multiplet routine fails to identify thin clouds~~ on benchmark days. These were selected as representatives of different sky conditions and examples of unidentified thin clouds by a human observer of the AOD timeseries, α-γ diagrams, and sky camera reference. An example (12th March 2020) is given in ~~figure 3~~figures 2 and 3, with additional examples in figure A1. We show the four dimensions of our space, as well as the resulting d_{20} of our datapoints ~~. From these days, we~~ and the sky camera pictures for better illustration of the timeseries. Both algorithms pick up the thin clouds around 9:15am, but only Clustering determines some smaller contrails between 9:30 and 10:00 as cloudy.

As can be seen from the d_{20} timeseries in figure 3, there is no clear distinction of two states (cloudy/clear), but rather a continuous spectrum of values that has to be divided to best fit the two categories. A lower threshold value will classify the ambiguous points as cloudy, but also risks a higher number of false positives and therefore lower overall data retention, which matters for error of mean values calculated from the data. Similarly, a higher threshold will cause more false negatives, i.e. cloud contaminated points to be identified as clear. We set the d_{20} threshold to 0.012 considering these aspects on our clear reference and benchmark days.

3 Results and Discussion

To assess the performance of the Clustering routine, we will compare it to the Multiplet routine, using the last 10 years of measurements (2010-2019), with 3330 days of measurements in total. Of these days, 1906 are found to have clear datapoints by at least one routine.

To exemplify the similarities and differences of the two routines, figure 4 shows days with different cloud and aerosol conditions: clear, intermittent thick clouds, intermittent thin clouds, a combination of passing thick and thin clouds, Saharan dust, and volcanic ash. Depicted are the timeseries of AOD at 501nm as well as a scatterplot of the Angstrom parameters α and γ for five days. Additional examples are shown in figure A1.

On a clear day, the routines agree ~~unsurprisingly very well~~very well, as expected. Clustering retains more points at the beginning and end of the day, which get picked up by limiting the airmass in the Multiplet routine. On the other hand, some slight outliers in α and γ get flagged by Clustering. The difference in daily mean is smaller than the measurement error.

When thick clouds are passing, with just short intervals of clear sky in between, Multiplet hardly identifies these as such. As Clustering takes all available data into account, it can assign points as clear even if the immediately preceeding and consecutive

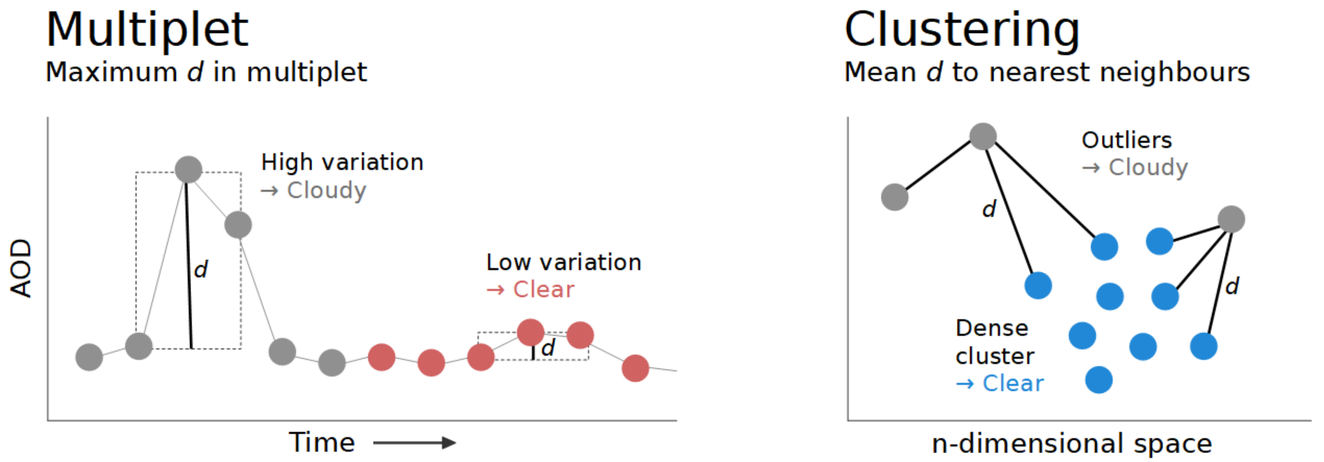

Figure 1. Schematics of the cloud screening methods. Left: The Multiplet method evaluates the difference between maximum and minimum AOD value of a set number of consecutive datapoints. Right: The Clustering method calculates the mean distance to k nearest neighbours in an n-dimensional space. Points for which a certain threshold of the respective measure d (indicated with the solid black lines) is passed will be identified as cloudy. These points are coloured in grey, wheras the clear points are coloured red/blue for Multiplet/Clustering.

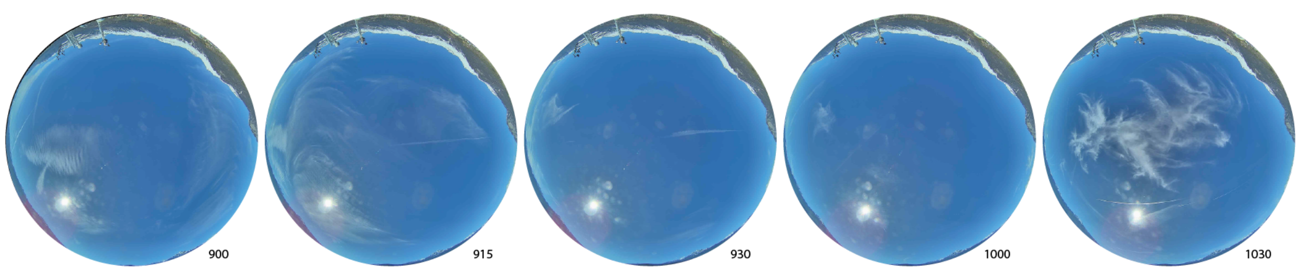

Figure 2. Sky camera pictures at 15min time resolution (time in UTC) for the timeseries shown in Figure 3.

point are deemed cloudy. Despite flagging less points, Clustering lowers the daily mean τ_{501} by 0.008 in this case, which is of similar magnitude as our measurement error.

On a day with lots of thin clouds (mainly contrails), the differences between the two routines are pronounced: a few relatively high AOD points in the morning (around 8am) pass Clustering, as do points during midday (between 10am and 12am). These points, which are spectrally very similar, are indeed cloud free, as confirmed by pictures of the sky camera. Multiplet however, filters less points as cloudy, which show cloud contamination as a decrease in fine mode fraction in the α-γ plane. For this day, the Clustering-Multiplet difference of daily mean τ_{501} is -0.027, which is the order of possible bias of Multiplet reported by Chew et al. (2011).

Another example of Clustering being more rigorous in cloud flagging can be seen on the day labelled with "Various Clouds". There were several optically thick clouds passing, which get identified correctly, but neither their thin edges nor the optically

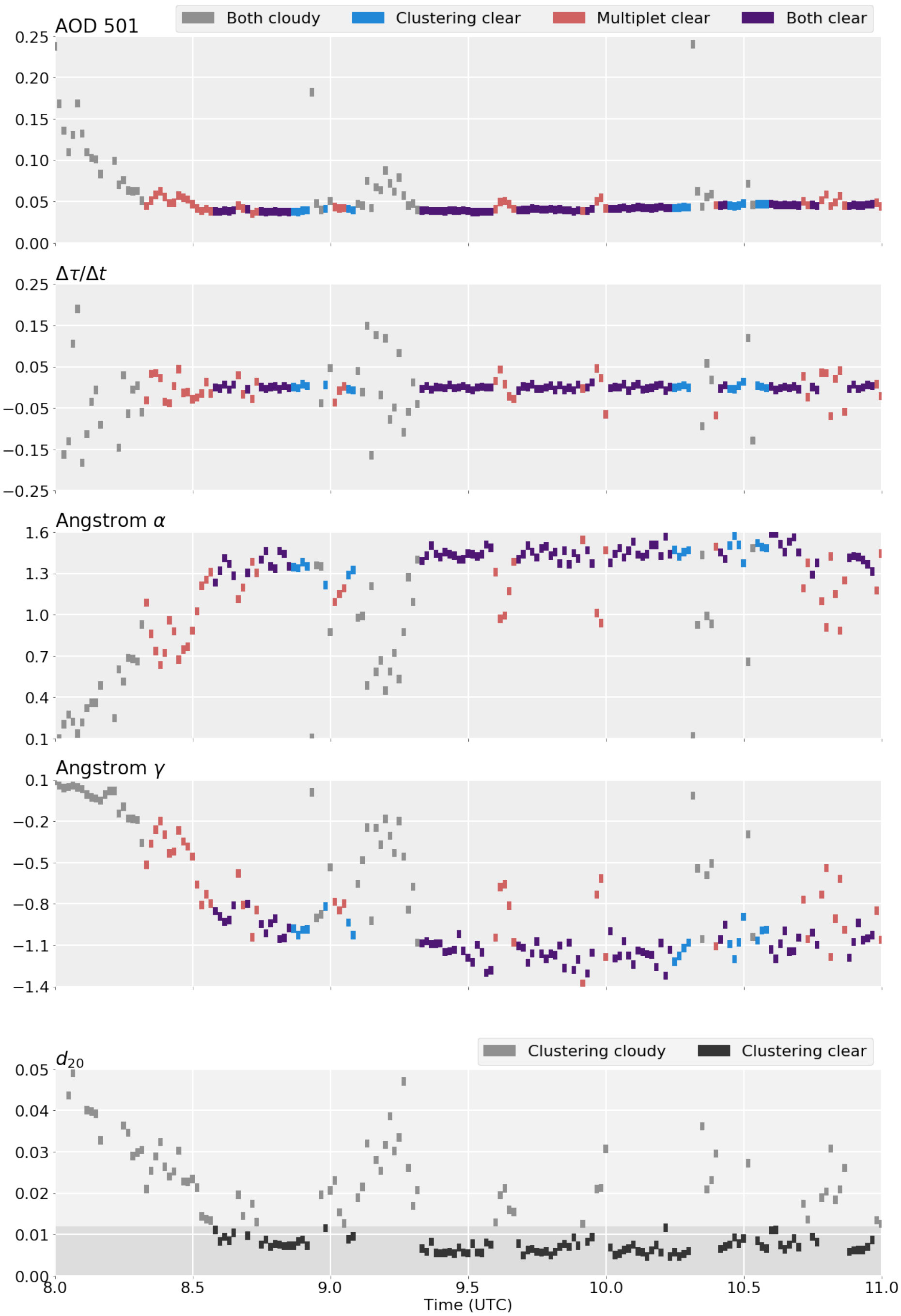

Figure 3. Three hours of an example day (12[th] March 2020) to illustrate the Clustering method: Timeseries of the four dimensions used, as well as the timeseries of the distance measure d_{20}. Colour of the rectangles codes for the flagging of the datapoint (see legend). For the distance measure d_{20}, points below the threshold, i.e. categorized as clear by Clustering, are coloured dark grey, cloudy points in light grey.

thin clouds on that day get picked up by the Multiplet routine. This day gets correctly eliminated by Clustering despite Multiplet marking 89 datapoints as clear.

Occasionally, Saharan Dust can get transported to Austria ~~.~~ (e.g. Ansmann et al. (2003)). Despite unusually high AOD, both routines correctly identify most of the data as cloud free. Daily mean τ_{501} is slightly lower (-0.005) when using Clustering, but this is still of the order of the calibration error.

One very unusual event is depicted last: after the eruption of Eyjafjallajökull in Iceland in April 2010, its ash plume was dispersed over Europe (Schäfer et al., 2011). It exhibits high AOD, and similar particle radii and fine mode fraction as Sahara dust. Clustering flags more data due to the high variation in AOD with time, but still retains data in the afternoon after about 1pm. Unfortunately, we do not have pictures available to estimate whether the data in the morning was cloud free, and should therefore be retained. Clustering lowers daily mean τ_{501} significantly, leading to -0.057 absolute and -12% relative difference. However, such an event is rare enough to be manually cloud screened, if necessary.

Overall, the Clustering routine flags more data than the Multiplet routine, albeit not necessarily the same data points. A more detailed comparison can be seen in figure 5. The Multiplet routine identifies about 47.6% of datapoints as cloudy, Clustering about 50.5%, which is a realistic value considering the amount of sunshine hours Innsbruck receives on average (Stadt Innsbruck, 2019).

As the main objective of the new algorithm was to filter thin clouds which previously passed the quality criteria, a higher number of flagged datapoints overall is expected. On the other hand, Clustering can flag isolated outliers without flagging the preceding and succeeding points of the multiplet, which lowers the number of flagged points. In 88% of all cases, the two methods agree in the (non-)assignment of a cloudflag. Nonetheless, about 10% of the data deemed cloudy by Multiplet is not flagged by Clustering, whereas 15% of the data passing the Multiplet criteria is identified as cloudy by Clustering.

The mean AOD values of all clear points based on Multiplet flagging are $\bar{\tau}_{368} = 0.19$, $\bar{\tau}_{412} = 0.16$, $\bar{\tau}_{501} = 0.13$, and ~~$\bar{\tau}_{368} = 0.05$~~$\bar{\tau}_{862} = 0.05$. The respective values based on Clustering do not differ significantly, which is partly due to the low number of datapoints on which the routines disagree.

On daily timescales, Clustering eliminates 169 days for which Multiplet would still find valid datapoints. On the other hand, there are only 10 days where the opposite is the case. Nonetheless, there are more than 1000 days without clear data in the 10-year record. The number of datapoints on the days which are disregarded by Clustering ranges between 1 and 89. Most of these days would therefore not be considered in further analysis in other measurement networks (Kazadzis et al., 2018; Giles et al., 2018) either. Furthermore, as shown in Figure 4, some of these days should be eliminated as they are indeed cloud contaminated.

Clustering leads to lower daily mean AOD on about 63% of the days (Figure 6). The mean difference is -0.0029 for $\bar{\tau}_{501}$, which is of the order of the calibration error (~~0.001~~ 0.005 to 0.01, depending on wavelength and airmass). However, on particular days this difference can range from -0.08 to 0.04 in absolute numbers, or -62% to +27% relative to the values based on Multiplet screening. Similarly, Clustering leads to higher mean α on 67% of the days. Averaged over ten years of data this leads to an increment of $\bar{\alpha}$ by 0.02. In extreme cases, the difference can be as high as +0.54. Both distributions are indicative of Clustering flagging thin clouds which Multiplet cannot properly detect.

Finally, we investigate the performance of the proposed method at lower time resolution. We subsampled the timeseries to 5, 10, and 15 minutes and analyzed the resulting data with the same settings for the algorithm. Key parameters of the resulting d_{20} distribution from the whole 10-year data record are shown in Figure 7. The coarser the time resolution, the higher minimum and median d_{20} values and interquartile range. This is mainly due to the overall density of points decreasing, thus increasing the mean distance to the nearest neighbours. Furthermore, the values of the first time derivative will be lower, so the relative weight of this dimension decreases.

To account for the changes in density, there are two possible adjustments: lower the number of nearest neighbours, or set a higher cloud flagging threshold. As an example of the latter, we show the timeseries on 12th of March 2020 (same as in Figure 3) at four different time resolutions in Figure A2. With increasing the d_{20} threshold to 0.019, 0.027, and 0.042 respectively for 5, 10, and 15 minute resolution, a very similar flagging behaviour can be achieved at all time resolutions.

4 Conclusions

We presented a new approach for flagging cloud contaminated data points from sun photometer measurements by treating them as outliers/region of low density in a four-dimensional space. Our ~~method successfully tackles the problem of the Multiplet routine by Smirnov et al. (2000) of not detecting~~ routine only needs one semi-empirically derived threshold and direct sun measurements for assigning a cloud flag. The method tackles shortcomings of the currently employed routine based on Smirnov et al. (2000) in the presence of optically thin clouds, such as cirrus and contrails, which lead to systematic bias of higher instant AOD values. Reducing this bias contributes to an improvement of long term statistics and trend analysis of aerosol conditions.

~~In contrast to the new routine employed by AERONET (Giles et al., 2018), which introduces more quality parameters (ten in total) and requires an aureole scan, our routine only needs one semi-empirically derived threshold and direct sun measurements for assigning a cloud flag.~~ While fewer datapoints are retained overall, which is expected from being able to filter thin clouds, the Clustering routine does not just flag more, but different datapoints (figure 5). As there is an ambiguity in the transition between humidified aerosols and clouds (Koren et al., 2007), an exact discrimination between false positives and negatives for either routine is not possible. Nonetheless, the new routine leads to lower AOD and higher α in the long term mean, which indicates a reduction of cloud contamination bias.

Detailed comparison with the previously employed cloud screening routine showed that both methods agree in their classification for the vast majority of cases (figure 5). Still, Clustering reduces mean AOD for most of the days in our testing period (figure 6). The daily mean AOD at 501nm averaged over the last 10 years is lowered by 0.0029, which is comparable to instrument precision (Wuttke et al., 2012). However, on single days Clustering reduces daily mean by more than 0.02 (up to 0.08), which is the same magnitude as reported as bias of the Multiplet routine by Chew et al. (2011), and exceeds the error of the instrument and trace gas optical depth. Together with specific example days (~~figure 4~~ figures 4 and A1), this supports the notion that clustering corrects some cloudy points of the Multiplet routine to clear, while flagging some of its erroneously clear points as cloudy. The small difference in the long term mean is partly due to the specific cloud conditions in Innsbruck, and could

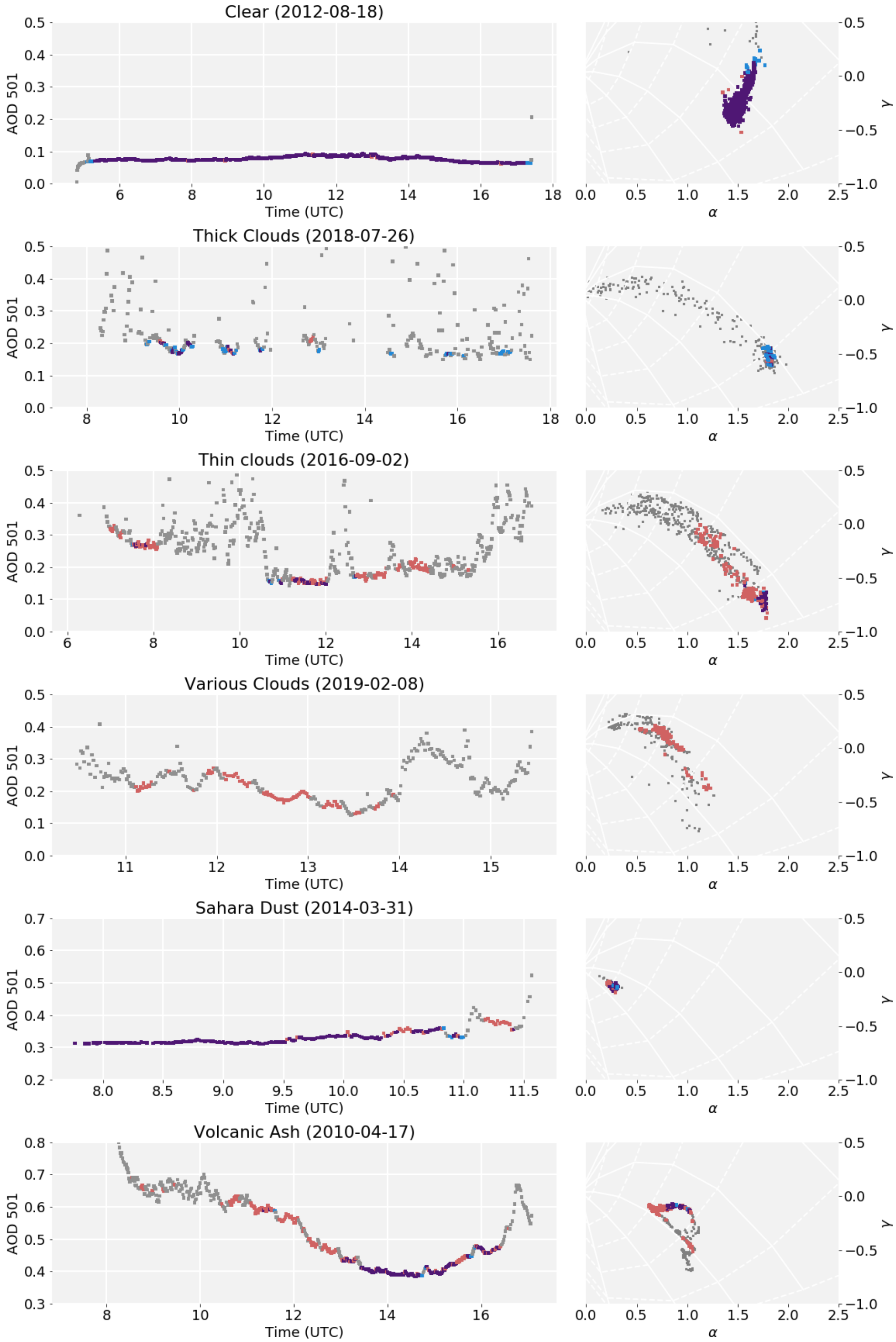

Figure 4. Comparison of cloud flagging routines on selected days with different cloud conditions as signified in the respective title. Colours as in Figure 3. Left: Timeseries of AOD at 501nm, Right: α-γ-plots. The ~~black~~ white solid lines show different particle radii, the white dotted lines different fine mode fractions, grid adapted from Gobbi et al. (2007). Note the different x-axis scales for the timeseries.

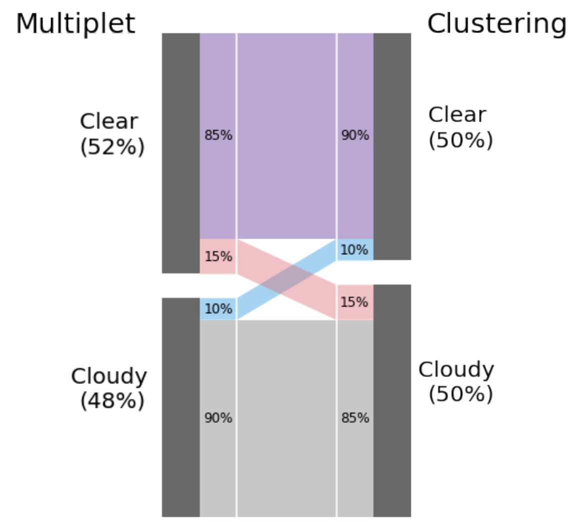

Figure 5. Comparison of flagging by the two routines. The height of each area is proportional to the total number of datapoints in each category. Grey: Both routines classify as cloudy, Red/Blue/Purple: Multiplet/Clustering/Both classify as clear.

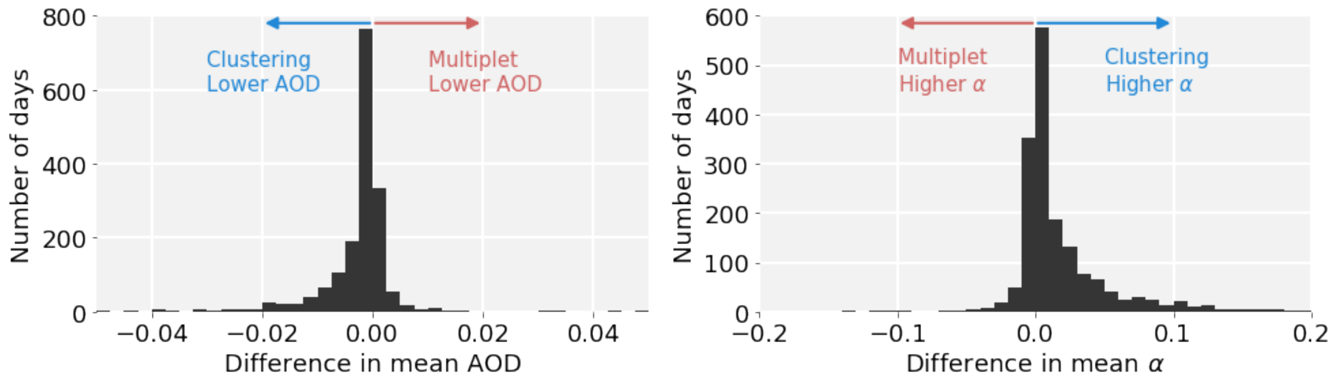

Figure 6. Histogram of the Clustering-Multiplet difference in daily mean of AOD at 501nm (left) and of α (right). Negative/positive values mean that the daily mean is lower/higher when screened by Clustering.

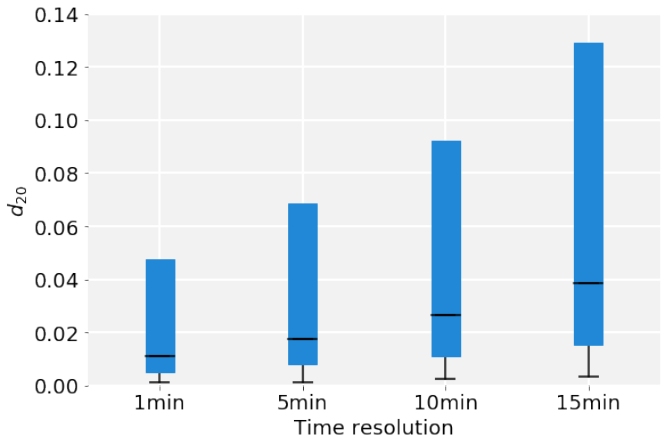

Figure 7. Comparison of the distribution of d_{20} for different time resolutions over 10 years. The bars extend from the 25$^{\text{th}}$ to the 75$^{\text{th}}$ percentile, the minimum and median of the distributions are indicated in black. Note that the maxima are higher than the graph range and therefore not shown.

therefore be much larger in ~~high latitude regions with low mean AOD and~~ regions with higher prevalence of thin clouds.

Due to the nature of the Clustering routine, it needs ~~a minimum number of measurements~~ at least k measurements to serve as possible nearest neighbours. In our case, we chose $k = 20$, though dynamic adaptions can be made if ~~these are not~~ there are less points available. As ~~its accuracy~~ the accuracy of the algorithm increases with a higher number of datapoints, it is ideal for post-processing. ~~This does not diminish its usefulness~~ Nonetheless, it can be used for real-time analysis ~~, as~~ as well, given that erroneously cloudy points can be ~~cleared later~~ corrected to clear when more data ~~is~~ becomes available, but not the other way round ~~.~~ (i.e. points identified as clear once will be labelled as clear regardless of additional measurements).

While the four dimensions considered in the Clustering routine account for variations of one specific wavelength and in the spectrum, the question arises whether these can be reduced even further. Especially γ, which has the highest error of the variables (Gobbi et al., 2007), might be a candidate. Initial independence tests using mutual conditional information as measure (Runge et al., 2019) show a strong association of α and γ. However, outliers in γ can appear independently of α, which is why we kept γ as a dimension and therefore data constraint.

So far we have tested the ~~routine~~ algorithm only for our ~~specific measurement site~~instrument in Innsbruck. It performs well in different cloud and aerosol conditions, as shown in figure 4, and ~~was shown~~ is able to alleviate AOD bias in the presence of thin clouds. ~~Adaptations~~ For the application at other measurement sites, the time resolution of the data needs to be considered, as lower measurement frequency leads to lower data density, and therefore higher mean distances between points (figures A2 and 7). Nonetheless, adaptations regarding the number of nearest neighbours, the relative weight of the different dimensions, or the d_{20} threshold can be easily done to optimize cloud detection with other instruments as well.

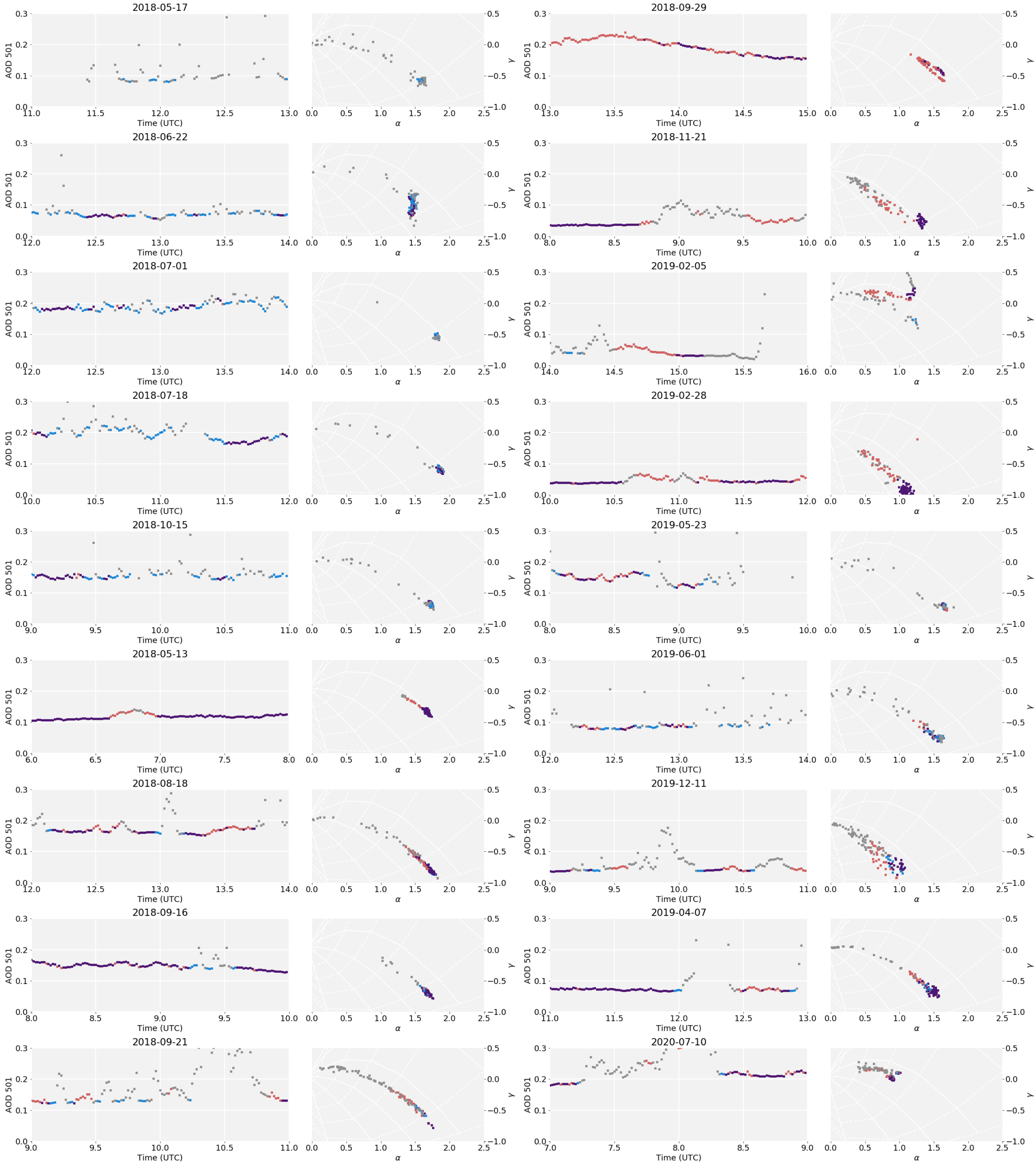

Figure A1. 2 hour exerpts from selected days, additionally to figure 4. The first five examples highlight cases where Clustering flags much less than Multiplet, the others show the performance in the presence of thin clouds. Both categories are ordered by date.

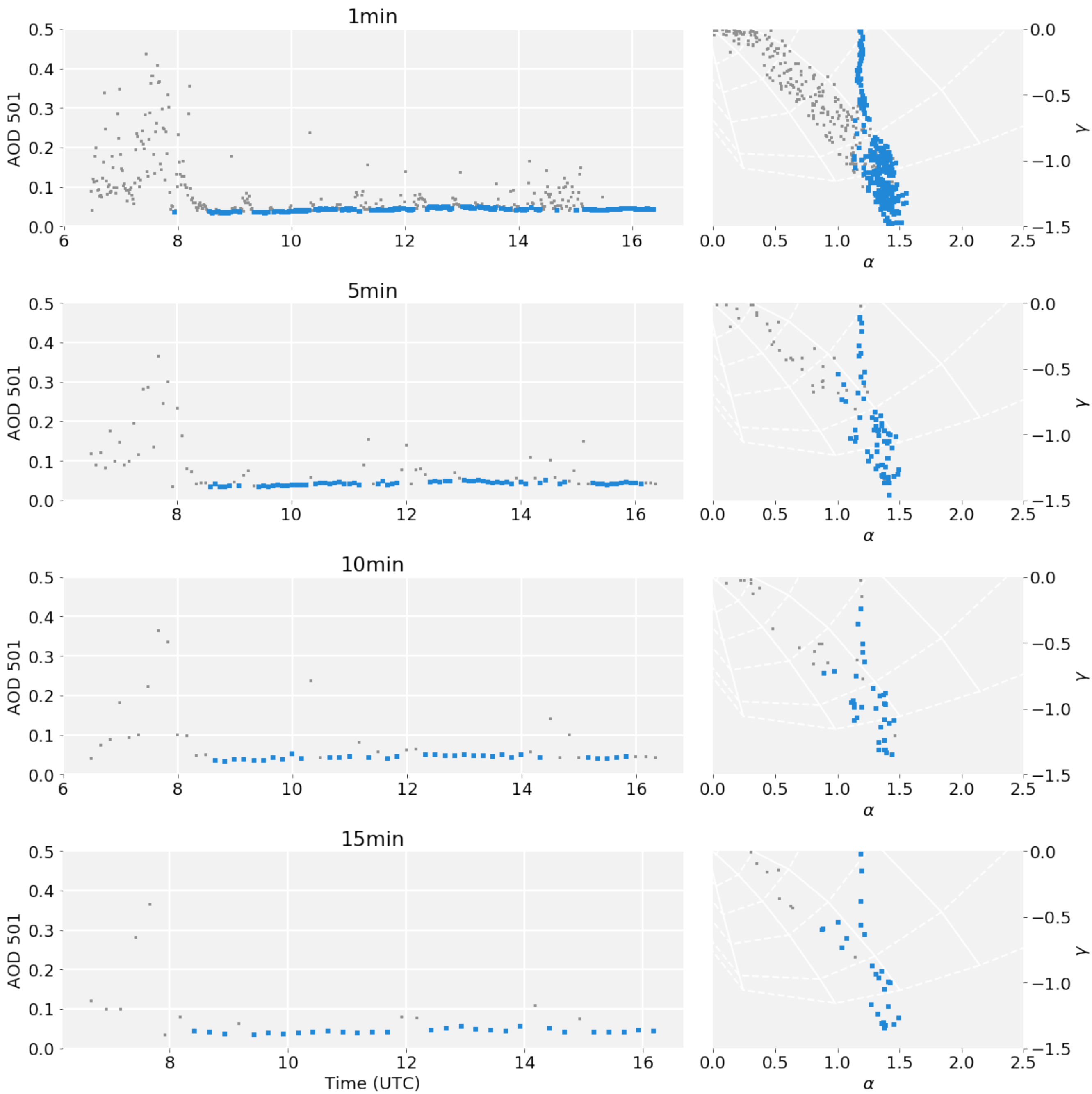

Figure A2. Time series of AOD at 501nm and α-γ plots at the original 1min resolution, as well as 5min, 10min, and 15min subsampling. Note that the initial datapoint was chosen randomly, so the 10 and 15 minute resolutions are not a further subset of the 5 minute resolution. Grey squares indicate cloudy points, blue clear ones. The cloud detection threshold was set to 0.019, 0.027, and 0.042 respectively for the lower time resolutions.

Author contributions. VS designed, implemented, and evaluated the Clustering algorithm. AK provided the raw data record. Both authors participated in writing, figure design, and interpretation of the results.

Competing interests. The authors declare that they have no conflict of interest.

References

Ångström, A.: On the Atmospheric Transmission of Sun Radiation and on Dust in the Air, Geografiska Annaler, 11, 156–166, https://doi.org/10.1080/20014422.1929.11880498, 1929.

Ångström, A.: The parameters of atmospheric turbidity, Tellus, 16, 64–75, https://doi.org/10.3402/tellusa.v16i1.8885, 1964.

Ansmann, A., Bösenberg, J., Chaikovsky, A., Comerón, A., Eckhardt, S., Eixmann, R., Freudenthaler, V., Ginoux, P., Komguem, L., Linné, H., Márquez, M. n. L., Matthias, V., Mattis, I., Mitev, V., Müller, D., Music, S., Nickovic, S., Pelon, J., Sauvage, L., Sobolewsky, P., Srivastava, M. K., Stohl, A., Torres, O., Vaughan, G., Wandinger, U., and Wiegner, M.: Long-range transport of Saharan dust to northern Europe: The 11–16 October 2001 outbreak observed with EARLINET, Journal of Geophysical Research: Atmospheres, 108, https://doi.org/https://doi.org/10.1029/2003JD003757, 2003.

Bodhaine, B. A., Wood, N. B., Dutton, E. G., and Slusser, J. R.: On Rayleigh Optical Depth Calculations, Journal of Atmospheric and Oceanic Technology, 16, 1854–1861, https://doi.org/10.1175/1520-0426(1999)016<1854:ORODC>2.0.CO;2, 1999.

Chew, B. N., Campbell, J. R., Reid, J. S., Giles, D. M., Welton, E. J., Salinas, S. V., and Liew, S. C.: Tropical cirrus cloud contamination in sun photometer data, Atmospheric Environment, 45, 6724 – 6731, https://doi.org/https://doi.org/10.1016/j.atmosenv.2011.08.017, 2011.

Giles, D., Sinyuk, A., Sorokin, M., Schafer, J., Smirnov, A., Slutsker, I., Eck, T., Holben, B., Lewis, J., Campbell, J., Welton, E., Korkin, S., and Lyapustin, A.: Advancements in the Aerosol Robotic Network (AERONET) Version 3 Database – Automated Near Real-Time Quality Control Algorithm with Improved Cloud Screening for Sun Photometer Aerosol Optical Depth (AOD) Measurements, Atmospheric Measurement Techniques Discussions, pp. 1–78, https://doi.org/10.5194/amt-2018-272, 2018.

Gobbi, G. P., Kaufman, Y. J., Koren, I., and Eck, T. F.: Classification of aerosol properties derived from AERONET direct sun data, Atmospheric Chemistry and Physics, 7, 453–458, https://doi.org/10.5194/acp-7-453-2007, https://www.atmos-chem-phys.net/7/453/2007/, 2007.

Holben, B., Eck, T., Slutsker, I., Tanré, D., Buis, J., Setzer, A., Vermote, E., Reagan, J., Kaufman, Y., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization, Remote Sensing of Environment, 66, 1 – 16, https://doi.org/https://doi.org/10.1016/S0034-4257(98)00031-5, 1998.

Holben, B. N., Tanré, D., Smirnov, A., Eck, T. F., Slutsker, I., Abuhassan, N., Newcomb, W. W., Schafer, J. S., Chatenet, B., Lavenu, F., Kaufman, Y. J., Castle, J. V., Setzer, A., Markham, B., Clark, D., Frouin, R., Halthore, R., Karneli, A., O'Neill, N. T., Pietras, C., Pinker, R. T., Voss, K., and Zibordi, G.: An emerging ground-based aerosol climatology: Aerosol optical depth from AERONET, Journal of Geophysical Research: Atmospheres, 106, 12 067–12 097, https://doi.org/10.1029/2001JD900014, 2001.

Huang, J., Hsu, N. C., Tsay, S.-C., Jeong, M.-J., Holben, B. N., Berkoff, T. A., and Welton, E. J.: Susceptibility of aerosol optical thickness retrievals to thin cirrus contamination during the BASE-ASIA campaign, Journal of Geophysical Research: Atmospheres, 116, https://doi.org/10.1029/2010JD014910, 2011.

Kasten, F. and Young, A. T.: Revised optical air mass tables and approximation formula, Appl. Opt., 28, 4735–4738, https://doi.org/10.1364/AO.28.004735, 1989.

Kazadzis, S., Kouremeti, N., Nyeki, S., Gröbner, J., and Wehrli, C.: The World Optical Depth Research and Calibration Center (WORCC) quality assurance and quality control of GAW-PFR AOD measurements, Geoscientific Instrumentation, Methods and Data Systems, 7, 39–53, https://doi.org/10.5194/gi-7-39-2018, 2018.

King, M. D. and Byrne, D. M.: A Method for Inferring Total Ozone Content from the Spectral Variation of Total Optical Depth Obtained with a Solar Radiometer, Journal of the Atmospheric Sciences, 33, 2242–2251, https://doi.org/10.1175/1520-0469(1976)033<2242:AMFITO>2.0.CO;2, 1976.

Komhyr, W. D., Grass, R. D., and Leonard, R. K.: Dobson spectrophotometer 83: A standard for total ozone measurements, 1962–1987, Journal of Geophysical Research: Atmospheres, 94, 9847–9861, https://doi.org/10.1029/JD094iD07p09847, 1989.

Koren, I., Remer, L. A., Kaufman, Y. J., Rudich, Y., and Martins, J. V.: On the twilight zone between clouds and aerosols, Geophysical Research Letters, 34, https://doi.org/10.1029/2007GL029253, 2007.

O'Neill, N. T., Eck, T. F., Holben, B. N., Smirnov, A., Dubovik, O., and Royer, A.: Bimodal size distribution influences on the variation of Angstrom derivatives in spectral and optical depth space, Journal of Geophysical Research: Atmospheres, 106, 9787–9806, https://doi.org/https://doi.org/10.1029/2000JD900245, 2001.

O'Neill, N. T., Eck, T. F., Smirnov, A., Holben, B. N., and Thulasiraman, S.: Spectral discrimination of coarse and fine mode optical depth, Journal of Geophysical Research: Atmospheres, 108, https://doi.org/https://doi.org/10.1029/2002JD002975, 2003.

Ramaswamy, S., Rastogi, R., and Shim, K.: Efficient Algorithms for Mining Outliers from Large Data Sets, SIGMOD Rec., 29, 427–438, https://doi.org/10.1145/335191.335437, https://doi.org/10.1145/335191.335437, 2000.

Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., and Sejdinovic, D.: Detecting and quantifying causal associations in large nonlinear time series datasets, Science Advances, 5, https://doi.org/10.1126/sciadv.aau4996, 2019.

Schäfer, K., Thomas, W., Peters, A., Ries, L., Obleitner, F., Schnelle-Kreis, J., Birmili, W., Diemer, J., Fricke, W., Junkermann, W., Pitz, M., Emeis, S., Forkel, R., Suppan, P., Flentje, H., Gilge, S., Wichmann, H. E., Meinhardt, F., Zimmermann, R., Weinhold, K., Soentgen, J., Münkel, C., Freuer, C., and Cyrys, J.: Influences of the 2010 Eyjafjallajökull volcanic plume on air quality in the northern Alpine region, Atmospheric Chemistry and Physics, 11, 8555–8575, https://doi.org/10.5194/acp-11-8555-2011, 2011.

Smirnov, A., Holben, B., Eck, T., Dubovik, O., and Slutsker, I.: Cloud-Screening and Quality Control Algorithms for the AERONET Database, Remote Sensing of Environment, 73, 337–349, https://doi.org/https://doi.org/10.1016/S0034-4257(00)00109-7, 2000.

Stadt Innsbruck: Monats- und Jahressummen der Sonnenscheindauer seit 1906, https://www.innsbruck.gv.at/data.cfm?vpath=redaktion/ma_i/allgemeine_servicedienste/statistik/dokumente38/meteorologischebeobachtungen/monatsjahressonnenscheindauerseit1906pdf, 2019.

Takamura, T. and Nakajima, T.: Overview of SKYNET and its activities, Optica Pura y Aplicada, 37, 3303–3308, 2004.

Valks, P., Pinardi, G., Richter, A., Lambert, J.-C., Hao, N., Loyola, D., Van Roozendael, M., and Emmadi, S.: Operational total and tropospheric NO_2 column retrieval for GOME-2, Atmospheric Measurement Techniques, 4, 1491–1514, https://doi.org/10.5194/amt-4-1491-2011, https://www.atmos-meas-tech.net/4/1491/2011/, 2011.

Wehrli, C.: GAWPFR: A Network of Aerosol Optical Depth Observatioins with Precision Filter Radiometers, in: WMO/GAW experts workshop on a clobal surface-based network for long term observations of column aerosol optical properties, edited by Baltensperger, U., Barrie, L., and Wehrli, C., no. 162 in GAW Report, pp. 36–39, 2005.

Wuttke, S., Kreuter, A., and Blumthaler, M.: Aerosol climatology in an Alpine valley, Journal of Geophysical Research: Atmospheres, 117, https://doi.org/10.1029/2012JD017854, 2012.