My report consist of two parts, one general one dealing with the procedure and a second one dealing with the paper itself.

1) **Preface regarding my personal feeling about the refereeing procedures:**

I personally think the open review system deserves some improvements at least regarding the officially approached referees. I have talked with quite a few colleagues who feel similarly.

I think that the official review system should stay anonymous. Scientists who have worked a lot with refereeing systems know that usually the most fair process would involve a double blind review (although , I have to admit, referees are often able to guess who the author is from citations and the field of work etc.). The second best system is a single blind review procedure which is usually applied. I have no idea what the benefits of an open review system with official journal referees should be (although a discussion forum for the general public is fine). The only reason I can guess is that scientists want to gain rewards for being officially attributed to a review process. This should give rise to a huge sigh: is our scientific system already in such a bad shape that scientists do not only want to gain reputation by their own papers but also by writing open reviews about others as well (if they disagree with something they anyhow always have the chance to write a comment later on under their name). Is this the next step in the publish or perish game. I personally think this is definitely the wrong direction and it contradicts the fundamentals of the well established anonymous scientific peer review system.

Of course the guidelines from this journal argue that referees may stay anonymous, but why is it offered at all (which may be read as a useful suggestion) to do it openly. It always happened in the past that occasionally referees directly contacted authors to discuss things further on, but this was their own decision and not suggested by the publisher.

One major disadvantage I also see for the officially approached referees is potential prejudice when seeing all these other comments on the website already before their own report is prepared – in particular if there are comments from so-called leading experts in the field writing under their name. I think this need not but may even pose an ethical issue. This is bad publishing policy to my opinion if the publisher wants to get absolutely independent reports without any prejudice.

My own reaction in this particular case: once I heard about the procedure I did not log in to the articles website (although uselessly reminded many times to do so to see new comments) before my own report was ready. This means that I also did not change anything in my report according to other comments. So some comments may be dublicate to others or some unclear items to me have been discussed and resolved already but I do not care for the sake of judging this manuscript totally independent.

I really would suggest to the publisher to go back to a completely anonymous system. In a second step you may start a public discussion with personalized comments but this should only be an addition after the regular refereeing is done.

2) Comments on the paper in question

This presents an extensive study with a camera system for first a long term recording of a sky region close to the sun and second an automated halo detection algorithm. The underlying motivation is to characterize ice crystals in cirrus clouds.

Overall the paper is nicely structured. It describes the used apparatus and analysis algorithms more or less in detail. I suggest to add a bit more discussion to related work, in particular when describing the halo model. I do have quite a few comments which may help to improve the paper. My method of referring usually involves first a brief reading – rather a scanning of the content – and then a careful reading where I mark sections and write comments in the margins. During a third round I summarize all my comments. Therefore I do not distinguish between serious comments or minor grammar corrections here, I rather start at page one and work forward. The authors should be able to see what is more important and what is just a minor comment.

I expect that the authors will most probably adequately deal with my comments, therefore I assume that the paper will be acceptable after a revision. However, I would like to see the paper again. So my official recommendation is: some amendments needed

Page 1 line 2: when mentioning sundogs you should first mention the scientific description parhelia and then maybe the non-technical term sundog.

Page 1, lines 4-8: As I understand the system was mostly in Munich but also for a 6 week period in the NL. In the text, the Cabauw period is mentioned once, but later on it is always stated that the measurements in Munich are discussed from 01/2014 to 06/2016. I propose to explicitly state which periods were used and which excluded for the Munich measurements.

Page2, top: I miss other refs. For example you use some old ones, but why not also Pernter Exners excellent book. Concerning general refs: you only mention Tape94. Tape has later written another interesting book on halos as well: Atmospheric halos and the search for angle X, W. Tape, J. Moilanen, AGU (Am Geophys. Union) 2006. Also, it is rather odd that you refer to the famous book by Minnaert with the year 1993. This is just a new translation of the much older NL book with the first edition dating back to 1937. Newcomers to the field might assume Minnaert is a contemporary scientist, please correct ref. by adding e.g. the original information.

Page 2, line 7: Probably you assume that everybody already knows about the 22° halo with respect to the circumscribed halo / tangent arcs. My personal experience is that most people just know about the 22°halo and have problems in understanding the differences to the other ones. Maybe just give short explanations in one or two sentences describing the differences (sun elevation). And it also seems to me that you use tangent arcs and circumscribed halo synonymously, if so: please make a respective statement somewhere.

Page 3, line 32: maybe clarify …. This implies that the <u>most important</u> <u>recorded (?)</u> halo …

Page 3, line 6: I miss the main results about halo frequencies in Germany from AKM. Later on you give similar results from Sassen in the US, you discuss the Munich results in Bavaria, so why not add results reported from all over Germany as well? Meteorological conditions in Germany should be closer to yours than the ones in the US.

Page 4 line 9: here you suddenly also mention tangent arcs, previously only 22° halo and circumscribed halo (see above). As mentioned above, briefly discuss all relevant halo features which may be observable with your equipment and then explain why you mainly focus on 22° halos and that / under which conditions your data may contain misinterpreted circumscribed halos / tangent arcs.

Page 4 line 10: you mention Sassen´s US results. Are the German AKM results consistent, i.e. do you expect the same order for the frequency of different observed halo types?

Page 5 line 11: 30/110=0.27 is definitely closer to 1/4 than to the mentioned 1/3.

Page 7, line 9: I would add quadratic when describing the pixels.

Page 8, Fig. 5: I miss information of how you made sure that your halo pixels were not saturated. Or did you exclude images with saturated pixels?

Page 8 Sect. 2.3: I wonder whether you also found some odd radius halos. If so, your choice of angular intervals may not always work. Or did you also test for smaller angular features with different intervals? And do you have a plot frequency of halo observations versus halo angles?

Table 2: I miss a more thorough discussion, for example based on your own simulations. Are these experimental results in agreement with your theoretical expectations?

Table 2: Of course the red inner edge was expected. Did you observe this always? Table 2 only refers to a single halo.

Page 9: concerning pointing accuracy: if you observe halos in all segments, you could use combinations of sectors 1+4, 2+5 and 3+6 to better determine the center. Would that improve your accuracy?

Page 10, line 4: A detailed description … refers to only 9 lines of text. This does not seem very detailed.

Page 10 referring to Sect. A and B should be referring to Appendix A and Appendix B

Caption Fig. 7: says it is the same as Fig. 5: Fig 5 showed R, G , and B. Here there is only one curve. Which one? Or is it something like 1/3(R+B+G) ?

Page 12, line 2 from bottom: was repeated 100 times. Question: 100 times for the exact same 30% ?

I wonder how many halo events were responsible for your halo images. For example if the same cloud gives rise to a 20 minute long halo display, some of your images refer to the same event, i.e. similar ice crystals within the same cloud. It would also be interesting to know the percentage of time a single halo producing cirrus is giving rise to halos. This could also change your statement how many clouds did not produce halos at all. Did you investigate whether there were reproducible differences between the halos of different cirrus clouds from different days (or e.g. if there were some with odd radius halos).

Page 13: you mention it indirectly but I would point it out more clearly: your test was done by also visually classifying the images. There is of course one problem: the eye may not be good to detect halos on still images. It is much better comparing differences in time sequences.

How have you done it? Looking at time lapse sequences or looking at stills? Can you comment?

Please comment why you have always used a 2σ deviation rather than 1σ or 3σ.

You mention the false positive or false negative results from your classifier. I expect a discussion. What have you observed when visually counterchecking those images. Could you find some reasons for the results, e.g. some brightly illuminated cloud fractions or shadows or ..?

Later on you mention sundogs: I assume for parhelia it would have been better if the segment boundaries would have been rotated such that they would fall within a single segment rather than being split up between 2.

Page 13, line 18: no discrimination so far, fine! But please comment whether you think that it will be easily possible to distinguish between the 22° halo and tangent arcs/circumscribed halo.

Chapter 4 and your sensitivity study: Nice model but in principle not new. Therefore I miss reference to other relevant work. There are many people who have e.g. applied less sophisticated Monte Carlo methods to halos, also including multiple scattering on halos (I am sure, a proper literature search will show up several papers). There have also been some studies on visibility of halos with respect to cloud optical thickness from the Atmospheric Optics community (see e.g., Bull. Am. Met. Soc. **89**, 471-485 (2008) or AppOpt47/34, H157 (2008)). I do expect a discussion either why you have not mentioned other simulations or you should add some of them and compare your work to their models.

Page 15, line 12: skip arc.

Why did you only use columns, why did you not use any plates in your simulations? Whenever you observe sundogs, you need the plates as well!

Why did you use the spectral albedo for grass. Munich is green I guess, but the institute is probably still in the middle of the town and not only surrounded by parks.

Page 16: it may be easier to understand and/or helpful to visualize your condition HR=1.0 for a treshhold which can be easily done.

Page 16, line 14: you mention that in multiple scattering HR decreases. This is plausible, but please also give reference. Or did you only intend to give a qualitative plausibility statement?

Sect. 5 Conclusions: I miss somewhere – not necessarily here – a discussion about potential reasons for the observed differences in halo frequency observations Sassen/AKM/your work. For example discuss meteorological differences in cloud formation due to different climates etc. You only mentioned a little bit on pages 5,6.

I assume – if extending the halo algorithm to parhelia, the color separation may be an additional criterion for distinguishing halo types.

Outlook: maybe say a few words what may in principle be possible. Can you imagine using a wide angle lens and also detect the colorful circumzenithal arcs which also have a reasonably high observation frequency in Germany?

Fig. 11 discussing your decision tree: In the paper you mention and show figures with HR around 1.1 to 1.15, but your criteria in the tree give numbers much lower (1.01 or so). Please comment. You may want to discuss what typical brightness variations under daylight conditions are considered to be easily perceived by the human eye.