Response to Anonymous Referee #1
Received and published: 13 August 2018

Journal: ACP
Title: A new Description of Probability Density Distributions of Polar Mesospheric Clouds (PMC)
Author(s): Uwe Berger et al.
MS No.: acp-2018-642
MS Type: Research article
Special Issue: Layered phenomena in the mesopause region (ACP/AMT inter-journal SI)

Introductory remarks:
We greatly appreciate the comments from the reviewer. We have taken his/her suggestions for improvements into account when preparing the revised version of the manuscript. In the following we respond to the reviewer's comments point by point.
We have marked the changes in the tracked version of the manuscript. Author responses are in italics. Line numbers refer to the first paper version. In the new tracked version deleted sequences are marked red. New text is marked in blue.
We want to thank the reviewer for the detailed review with many useful ideas and suggestions which, we think, have significantly increased the quality of the manuscript.

GENERAL SUMMARY AND COMMENTS:

This paper presents a new analysis of the statistical behavior of various polar mesospheric cloud (PMC) properties, using observational data from the ALOMAR lidar in Norway. The authors show that the commonly used g-distribution, which prescribes an exponential dependence for the cumulative probability of a given parameter value, is not adequate for the full range of all PMC quantities determined from lidar data. A revised probability density function called the Z-distribution is developed that has two free parameters (scale and shape), and simplifies to the g-distribution when shape =1. The new function provides a better representation of quantities such as ice mass density and ice radius, and is also shown to enable the creation of "artificial" data from one parameter into a second parameter given Z-distributions of both quantities.

This work is a valuable extension of the original g-distribution concept. However, it is not clear that the proposed applicability to long-term trend studies is justified. Some suggestions and comments related to specific items are provided below.

SPECIFIC COMMENTS:

1. p. 2, line 1: Since the monograph cited here may not be easily accessible for many readers, I suggest reproducing the relevant figure (with permission) to help introduce the basic concept of the g-distribution.
We agree but given the number of figures already in the manuscript we see no space for more figures. The basic concept of g-distribution is discussed in detail with respect to ALOMAR data in section 3, 3.1.1, 3.1.2, and 3.2 (a total of approximately 6 pages including 3 figures). Also Appendix A summarizes some general properties of the g-function. Please keep in mind that the Thomas-Monograph-Paper is cited to reference correctly the first publication about

exponentially distributed (g-function) PMC data. During the last two decades many people use this g-function approach in analyses of PMC data.

2. p. 3, line 8: This choice obscures the inter-annual variability in PMC behavior that has been demonstrated in many previous studies (e.g. Rong et al. [2014], DeLand and Thomas [2015], Fiedler et al. [2017]), which can represent a factor of two variation for the slope of the g-distribution at the latitude of ALOMAR. How does this averaging affect the applicability of the results derived later to any specific PMC season?
Comment:
We had written at the end of section 2 (ALOMAR data description):
"In this paper we will analyze the climatology of all ice seasons from 2002 until 2016 merging all 15 seasons to one data record. Within this combined data set we then get a total number N of 8,597 observations which is sufficiently numerous in order to avoid too large statistical irregularities in a frequency histogram of the data."

Hence we treat the entire period 2002-2016 to get a larger number of data which improves the statistics of frequency rates. If one subdivides the total period into single seasons, we have to carefully test and analyze the single season numbers. Such a work is in preparation, and we think that these results will be presented in a subsequent publication.

Concerning your second comment that g-slopes sometimes differ from season to season which can represent a factor of two for the slope. Exactly, that point makes us suspicious. How do you know that in all these seasonal data analyses a g-approach is justified? We think this is not true, and therefore the slope-results should be considered with caution. Again we mention that we will address this task in a subsequent publication in the near future.

3. p. 3, line 10: Is the three-color mode of operation used less often? Other papers discussing ALOMAR lidar measurements talk about 15-minute binning (e.g. Fiedler et al. [2017]), so I would actually expect many more individual profiles to be available from 15 years of data.
Yes, this is true. One color measurements in the green laser line have been performed since 1997 and are much more frequent compared to the three-color measurements. Both techniques use integration times of 15 minutes. BUT for our DERIVATION of a new Z-pdf presented in this paper we wanted to use only three-color measurements because only those deliver SIMULTANEOUS measurements of max. backscatter, radius, number density, and ice water density at the height of beta-max. These simultaneous measurements are necessary to calculate correlation and regression plots that help to understand the physical meaning (power law dependence between parameters) of the new Z-pdf. Nevertheless, the Z-pdf can be constructed by a single data sample itself, see examples in section 5.3.

4. p. 3, line 31: I'm not sure I understand how the standard deviation can be equal to the mean with a threshold of zero. Statistically, one sigma should not encompass all data smaller than the mean, but with the definition given in line 29, you are not allowing for negative values of x. So if (mean – st.dev.) = 0, where does (mean – 2*sigma) fall?
Yes, this is true. Mean, variance, and ... can be always calculated for given distribution (Keyword: Moments of a pdf). See the mathematical derivation in Appendix A. Please keep in mind that all of our 4 ice parameters beta_max, r, n, and IMD are physical parameters with values larger than zero. Hence a beta-max value of zero does not exist, same argument holds for a case of negative beta-max. They do not exist!. The same is also true for n, IMD, and radius.

Please note, that this general property of a g-function, the mean is equal to the standard deviation, is used for our g-function test, see Eq 4.

5. p. 5, line 6: I'm not sure that the term "obviously" is appropriate. The fit line in Figure 2a does go through the data, but the fluctuations between y = 40-90 look comparable in magnitude to those between x = 20-40 in Figure 1a. The roll off at y < 40 in Figure 2a is more significant to me, and it suggests that using a higher threshold (e.g. 40) would yield a satisfactory fit.

Done, we delete 'Obviously' and replace the sentence with: We show in the following that the data points have not a precise linear shape.

A comment to a larger threshold of y>40: Such a tendency of choosing a special interval where linearity is more or less true can be often seen in literature. This is what we call "at least piecewise exponential shape", see introduction. Unfortunately, such an 'arbitrary' selection procedure is inadmissible in the sense of honest statistics.

Remember that we have chosen a threshold for beta_max of three, and the value of three is even a conservative estimate. Have a look at Fig 3a, the regression plot between beta-max (x) and ice mass density (y). You will see that the mean regression line (solid line) has a value of y=20 for x=3. Hence the y threshold is 20. And even for this case, a lot of data points have been canceled by the condition x>3 and y>20.

6. p. 5, lines 17-18: Regarding "larger discrepancies", see comment #5.

Yes, for ice radius and number density, exponential fits get even worse, see Fig 2c-f.

7. p. 7, lines 3-4: MBS is a first-level measured quantity, whereas IMD, R, and n are derived based on various assumptions. Does this "failure" say something about the functional forms used to create the latter group of products?

We think that this has no influence on the group of products.

8. p. 8, lines 2-3: This figure uses data well below the previously defined fit threshold for both MBS and IMD. Does the result change if MBS > 3, IMD > 20 are required as specified for Figures 1 and 2?. What about IMD > 40, as suggested in comment #5?

This figure uses ALL data that has been detected. As already said, a threshold of 3 is a conservative estimate for lidar sensitivity. Since summer 2002 there has been further development of the lidar system at almost every year. So, sensitivity became better and better.

Does the result change if MBS > 3, IMD > 20 is used or other thresholds? *We performed several numerical tests which show the following: regression points, mean, and median will always change when introducing different thresholds. But (c,d)- values change only slightly within a few percent (MBS > 3, IMD > 20).*

9. p. 13, lines 15-16: The first two derived threshold values are close to those given in Section 3.1.2. Is the third value a maximum?

The third value is n_th = 662 cm-3, see discussion of n_ice in this section where we had given a description:

…The sample of ice number density shows a completely different behavior with a slope parameter that is negative with b= - 0.819. The physical meaning is that the parameter ice number density is negatively correlated with all other ice parameters. For example, large ice numbers n correspond to small ice radii, IMD and MBS values. As a consequence this leads to a threshold of n in the reverse direction, that is from large values to small values defined by n < n_th =662cm −3 …

10. p. 13, line 33: This statement seems to connect back to lines 23-24 on this page. Isn't it circular reasoning to say that they agree?
At lines 23-24 we discuss max. backscatter. At line 33 we discuss ice mass density (IMD). Here we note that also numerical values of mean, median, and standard deviation for IMD (and r and n) agree almost perfectly. We think that these hints are justified.

11. p. 14, lines 5-6: This statement is physically plausible for radius. It seems reasonable for IMD which is proportional to r^3. Not sure about MBS, because it seems like large density could overcome the dependence on r (but is this true if MBS is proportional to r^6?).
We calculated all correlations, and in fact this is the result.
As you say it is plausible for radius and IMD. IMD is proportional to n (radius^3).*
MBS is proportional to n (radius^5.8). You see the similarities. For that reason both pairs (n,MBS) and (n,IMD) are negatively correlated.*

12. p. 18, lines 17-19: I'm still not convinced that the parameters derived from a multiseason collection of data are valid to use for this type of "synthetic" data calculation with a smaller subset of original IWC data, based on the previous comments about interannual variations.
The larger the data sample the better the statistics. The multi-season data sample describes the general properties of the general frequency distribution without season-to-season changes. As we described before we plan to write a second paper in the near future that investigates the step from a decadal period to a single season.

13. p. 20, lines 10-13: Is this statement saying that the uncertainty in the retrieval assumptions is large enough to justify the difference in b? What level of agreement would be needed for confidence?
The conclusion is that the application of simplified assumptions used in the analytical example (for example Gaussian distributed ice particles at the height of maximum backscatter, constant ice particle number or spherical shape of ice particles) do not reproduce a power constant d that results from the lidar observations. This is a systematic difference that can't be explained by statistical errors.
We modified these sentences with:
***The** power constant (d=1.46) derived from the shape parameters b x and b y of the Z distribution analysis of real ALOMAR IBS and IMD data is **significantly** different from the power estimate (d=1.93) belonging to the analytical example that necessitates various assumptions, e.g Gaussian distributed ice particles at the height of maximum backscatter, constant ice particle number or spherical shape of ice particles. Hence, we conclude that the determination of shape parameters b from a Z-distribution analysis **of observational data** therefore provides …*

14. p. 20, lines 15-17: This goal would require quantitative answers to comment #13 in order to be able to identify such changes. It also goes back to comment #2 regarding the question of how these fits behave with different individual years of data, and discussing how much noise increases with the reduction in the number of samples.
As we mention this might be a future goal. We will try to address this point in a future paper.

TYPOGRAPHICAL ERRORS

p. 5, lines 14-15: “unequal” could be “not equal to”.
Done

p. 5, line 15: “unequal” could be “not equal to”.
Done

p. 8, line 26: “allows to” should be “allows us to”.
Done

p. 9, line 9: “particulary” should be “particularly”.
Done

p. 14, line 8: “tale” should be “tail”.
Done

p. 15, line 7: “Have in mind” could be “Please keep in mind”.
Done

p. 16, line 11: “outcome” should be “outcomes”.
Done

p. 24, line 12: “stimulus” could be “stimulating”.
Done