1. Introduction

We would like to thank the reviewer for suggestions to improve the manuscript. It has been modified according to the comments. Detailed replies on the comments are given below; comments of the reviewer are given italicized.

2. Comments

1. P15453. The title has to be changed following the previous remark.

The title was changed to:

> Ground-based lidar measurements from Ny-Ålesund during ASTAR 2007

2. P15456-l5. Lidar are not necessary blinded in presence of thick aerosol or cloud layers. Could-you give a limit value in term of optical thickness for the lidar used in the frame of this paper?

This paragraph was changed to:

> As observational data in the Arctic are sparse and model parameterizations are usually tested against older campaign data, we used the special opportunity to run the KARL lidar (Ritter et al., 2004) between 1 March and 30 April whenever possible. Thick low level clouds would have blinded the detectors so the lidar was not switched on when the sky was overcast. With KARL 145 h of evaluable lidar data were gathered, while the MPL yielded 47 days of lidar data.
> This allows us to give a statistical overview of the Arctic spring time in the year 2007. Due to the high inter annual variability of the Arctic climate (Eckhardt et al., 2003) this analysis is still only a snapshot, observations of other years might differ substantially.

The information on multiple scattering was moved to section 3.1.1:

> KARL can only be operated when the backscattered fraction of the light is not too strong in order not to damage the photomultipliers. This inhibits the evaluation of optically thick clouds with high backscatter, especially in the lower altitudes as the dynamic range of lidar return signals is inversely proportional to the distance z^2. In total almost 150 hours of lidar data were collected with KARL. About four hours of data could not be evaluated due to low laser power, optical adjustments and multiple scattering at clouds with an optical depth above 0.55.
> These data sets were excluded from this study. Hence, we restrict ourselves to clear sky conditions and clouds with low optical thickness and neglect multiple scattering in this study. About 145 hours of trustful data remained.

3. P15456-l28. Could you detailed why liquid water is present at the cloud top?

Liquid water on cloud top is characteristic for mixed phase clouds which occur frequently in the Arctic. The presence of liquid water is underestimated in current climate models at low temperatures (Gayet et al., 2009, this issue). Further examples of mixed phase clouds with proper references for their structure and importance are analysed in Lampert et al. "Observations of boundary layer, mixed-phase and multi-layer Arctic clouds with different lidar systems during ASTAR" in this issue. We add this reference to clarify this point.

4. P15457-l21. Dense clouds do not systematically affect the photomultiplier but lead to some difficulties to inverse the lidar signal.

This is true. However, in our system especially the 1064nm channel and the 532nm channels in both depolarization directions can go into saturation if thick clouds lower than 1.5km are present. This was the main constrain that limited our observation time. To not stress the APD (1064nm) and PMTs the whole system was switched off in these cases. Dense clouds above 4km with optical depths of more than 0.55 (roughly corresponding to a BSR>50 at 532nm) would have led to multiple scattering and are excluded in this study. These figures are given in the new version to clarify this point.

5. P15457-l23. How do you assess the multiple-scattering effect on the data?

A cloud of optical depth 0.4 at 9km altitude did not produce multiple scattering neither and was also included (14 April 2007). On the other hand, a cloud with optical depth of 0.8 at 3km altitude was excluded (22 April 2007). So in total an optical depth of 0.55 is the maximum included in our study. The optical depths values given here were calculated according to the transmittance method:

Wei-Nai Chen, Chih-Wei Chiang, and Jan-Bai Nee, "Lidar Ratio and Depolarization Ratio for Cirrus Clouds," Appl. Opt. 41, 6470-6476 (2002).

In total only 4 hours of observation time (out of 150 hours) had to be cancelled for this study. This includes low laser power, misalignment of the optical set-up and very few situations of multiple scattering. Our system as it was in 2007 reacted more sensitive to blinding of the detectors at the bottom layer of low clouds - this is the main reason why only 150 hours of observation time could be derived. The information of the limiting optical depths and the missing data will be included in the manuscript.

See comment 2.

6. P15458-l7. The reference on the Vaisala radiosondes is recent. Do exist original references?

There is a significant difference between humidity sensors on the RS80 and RS90/92 radiosondes. The quote of Miloshevich 2006 is in our point of view appropriate to analyse their trustability. The quote of Treffeisen 2007 describes the evaluation and usage of the Vaisala sondes at our site. Maybe the most original quote is a link to the Vaisala homepage (which may vary in time): http://www.vaisala.com/weather/products/rs92.html. We will encourage all readers in case of questions, remarks and cooperations to contact our institute.

7. P15458-l13. It is less reliable, but of how much?

Miloshevich, L. M., H. Vömel, D. N. Whiteman, B. M. Lesht, F. J. Schmidlin, and F. Russo (2006), Absolute accuracy of water vapor measurements from six operational radiosonde types launched during AWEX-G and implications for AIRS validation, J. Geophys. Res., 111, D09S10, doi:10.1029/2005JD006083

compared different water vapor sensors and detected four potential problems with the Vaisala sensors: calibration, a temperature dependent offset, a time lag effect and - during sunlight - a sensor heating effect.
A summary of the accuracy of the RS92 can be found in their Table 3. The worst accuracy was observed at lowest temperature, below -50°C. However, our statement of an overall precision of 5% in the RH retrieval is valid. At higher temperatures the error in RH is closer to 2-3%. We suggest clarifying this point by rewriting:

> The relative humidity RH can be measured between 0 and 100% with a resolution of 1% and an accuracy of 5% at -50°C - the colder the temperature the larger becomes

> the insecurity (Miloshevich et al., 2006, Währn et al., 2004). More details on radiosounding at the Koldewey station can be found in Treffeisen et al. 2007.

8. P15458-l22. Why is it complicated? Is the threshold of 0.25 important enough to define a temperature inversion? Perhaps, it will be interesting to show a temperature profile on Fig. 1.

The figure was partly changed according to your advice. Surface temperature inversions of 0.5K instead of 0.25K were plotted, showing the frequency of occurrence being much higher in March than in April. Temperature profiles are shown in other plots for the case studies in section4.

The paragraph was changed as follows:

> A frequent Arctic phenomenon are low level temperature inversions, which are forced by strong radiative cooling of the surface and inhibit the mixing of the air in the lowermost troposphere with that of the overlying free troposphere. Thus they play an important role in the dynamics of the Arctic planetary boundary layer (Kahl, 1990). We analyzed the occurrence of inversions below 6km altitude using the 71 obtained temperature profiles in the original resolution of 5s read-out, which equates to a vertical resolution of about 25m. The algorithm added up the temperature difference between two adjacent height steps as long as it was positive.
> In Fig. 1 temperature inversions of more than 2K are marked with blue dots. In March they were observed frequently, declining in April. The white dots mark the surface-based temperature inversions above 0.5K, whose inversion base was below 25m (Kahl, 1990), lower temperature differences were neglected. 13 out of the 15 surface-based inversions were observed in March, including the four events with surface-based inversions stronger than 2K.

9. P15460-l13. Could you justify the choice of the 3 pressures levels for the backtrajectories?

The pressure levels and calculation times used are data which can be exactly extracted from the ECMWF archive without further interpolation to other times/heights. 850hPa is equivalent to boundary layer heights, 700hPa is the lower troposphere where Arctic Haze usually occurs, 500hPa is the mid free troposphere, which is also of interest.

10. P15460-l15. How do you choose the clusters? What is the criterion? What happens if a trajectory crosses several clusters?

For the clustering of the trajectories we used a non-hierarchical cluster method, provided by the MATLAB software. The tool is named "kmeans" and compares the similarity of values in the Euclidian space. Therefore the trajectories were transformed to the Cartesian grid and clustered by x- and y-position. The trajectory altitude was neglected for a better comparison with Eneroth et al.. If the trajectory passes several clusters, the program chooses the cluster with the smallest difference.

11. P15460-l25. Are the observed differences when comparing to the work of Eneroth et al. due to the weak statistical representativeness? We can see this with the differences between this work and the total ASTAR campaign (P15461-l1).

We don't think that it is due to statistical representativeness because all of our clusters are filled with a sufficient number of trajectories. The observed differences are probably due to interannual differences in the large-scale meteorological condition. In Fig. 1 in Eneroth et al. one can see the large standard deviations of the monthly distribution of the different clusters, which in some cases exceeds 100%.

Quinn et al 2007, Tellus B 59/1 and ref. therein report a correlation between Arctic Haze occurrence (over America) and the Arctic oscillation. We have not seen such a simple correlation in our data in Spitzbergen, but a high interannual variability of Arctic Haze. Moreover, the time of the ASTAR2007 (March 26 to April 18) campaign was unfortunately too short to allow a statistical meaningful sample on a general scale as there were no obvious Arctic Haze event during the campaign.

12. P15461-l13. How the clouds and aerosols contributions are separated in using the measurements of the KARL instrument?

The separation of clouds and aerosols is described in section 3.4.2 of the paper.

13. P15462-l13. It is not exact, this is the inverse. Re-examine the definition of the color ratio.

Unfortunately there are several definitions of the color ratio in the literature. Sometimes simply the ratio β^{aer}(short wavelength) / β^{aer} (long wavelength) is used. In this case color ratio=1 refers to large, a high value to small particles.
Beyerle, G., Gross, M. R., Haner, D. A., Kjome, N. T., McDermid, I. S., McGee, T. J., Rosen, J. M., Schäfer, H-J., Schrems, O.(2001). A Lidar and Backscatter Sonde Measurement Campaign at Table Mountain during February-March 1997: Observations of Cirrus Clouds, Journal of the Atmospheric Sciences, Vol.58, 1275-1287.

Sometimes a definition similar to the Angström Exponent for the extinction is used (frequently called color index) $CI= - \ln (\beta^{aer} (\lambda_1) / \beta^{aer} (\lambda_2)) / \ln (\lambda_1/\lambda_2)$
However, if the referee doesn't mind, we would like to keep our definition. In this way the results are directly comparable to Lampert et al. in this issue.

14. P15462-l17. A reference is needed for the Angström exponent. Angström, A., The parameters of atmospheric turbidity. Tellus 16, 64-75, 1964.

The quote of Angström will be included, thanks!

15. P15462-l20. The uncertainties on the Angström exponent are necessary mainly with low values of extinction or backscatter coefficient (cf. Hamonou et al., 1999). Hamonou, E., P. Chazette, D. Balis, F. Dulac, X. Schneider, E. Galani, G. Ancellet, and
A. Papayannis, Characterization of the vertical structure of Saharan dust export to the Mediterranean basin, J. Geophys. Res., 104, 22257– 22270, 1999.

This is true. The quote will be added in the manuscript.

16. P15463-l1. It is not true for both the cloud top and thickness when lidar is located on the surface.

The paragraph was changed:

> Multiple scattering causes the lidar return signal from clouds to increase with increasing receiver field of view (FOV) (Eloranta, 1998) and optical thickness of the cloud (Hu et al., 2006; Cho et al., 2008). For the MPL multiple scattering can be neglected as it has no significant influence on the qualitative cloud base detection. KARL statistics refer to rather clear conditions, a few data sets with clouds showing multiple scattering (approx. at BSR>50 and AOD<0.55 in 5km altitude) have been removed.

See comment 5.

17. P15463-l6. Another reference for the MPL is Spinhirne et al. (1993). Spinhirne, J. D., Micro pulse lidar, IEEE Trans. Geosci. Remote Sens., 31, 48– 55, 1993.

The quote will be added in the manuscript.

18. P15463-l15. For the first interval, close to the surface, do you consider the overlap factor as very important for the MPL? An assessment of the overlap factor could be found in Chazette (2003). Chazette, P., The monsoon aerosol extinction properties at Goa during INDOEX as measured with lidar, J. Geophys. Res., 108, doi10.1029/2002JD002074, 2003.

The first altitude interval was analyzed between 100m and 1km, which is in agreement with the paper you refer to. We will include this information in our manuscript:

> For ten altitude intervals of 1km width, the retrieved BSR profiles were analyzed to find cloud structures within the interval.
> Different thresholds for the difference between two adjacent BSR values were used which were determined conducting sensitivity studies. For each altitude interval beginning at the surface (the first interval was restricted to 100m–1km due to incomplete overlap, see Chazette, 2003), BSR differences above 0.1 in conjunction with increasing BSR values for at least three height steps or a single BSR peak difference of minimal 0.2 were needed to detect a cloud.

19. P15462-l15. How do you define the thresholds to detect the clouds on the lidar profiles? A sensitivity study is necessary to fixe the value.

See comment 18.

20. P15464-l1. The first sentence is not clear.

> For the following analysis the three subsequent MPL profiles after each balloon launch were evaluated since the weather balloons take about 30 min to reach the tropopause.

21. P15466-l10. In your discussion on the cloud fraction, you can also consider the reference of Berthier et al. (2008). Berthier, S., Chazette, P., Pelon, J., and Baum, B. (2008), Comparison of cloud statistics from spaceborne lidar systems, Atmos. Chem. Phys., 8, 6965–6977.

Thank you for the notice, the values retrieved in this paper for the mean cloud top altitude at about 80°N coincide very well with our cloud occurrence analysis for the MPL. This fact will be included in section 3.3.2:

> The frequency of cloud occurrence in the different altitude intervals is marked with the dashed green curve in Fig. 5. Profiles with snow on the window (14 days total) are not considered. Low clouds between 1 and 4km dominate, while there is another peak for the higher clouds between 7 and 8km. The mean cloud altitude is between 4 and 4.5km, which is in accordance with the mean cloud top height at 80°N retrieved by satellite measurements in Berthier et al. (2008).

22. P15466-l24. What is the overlap factor of the KARL system? What is the initial reference?

The KARL system was built up in 2001 and repeatedly updated since then. Therefore the valid reference for the ASTAR 2007 period is Ritter et al. 2008. Full overlap is reached at about 1000m, which is why we skipped the altitude interval 0-1km as stated on page 15467 in line 6.

23. P15467-l1. On the Figure 6, are the points with VD > 2% and BSR > 10 biased by the multiple scattering?

No, as explained in the new manuscript, all data with multiple scattering (BSR>50) are excluded from the study. We checked all lidar data and are sure not to depend on multiple scattering.

24. P15468-l1. How do you justify the threshold? With a sensitivity study?

No. Our thresholds are empirical values based on previous data analysis. In the cases of Arctic Haze in previous years (partly unpublished yet) we clearly see a depolarization between 2-5% whereas cirrus observations suggest that a separation into medium and high depolarization is justified. The backscatter ratio was categorized in three categories: no particle backscatter above the background level (BSR<1.2), moderate backscatter (1.2<BSR<2) and strong backscatter (BSR>2). The BSR=2 threshold is somewhat arbitrary. In total we have in our 10min/1000m resolution about 8600 data points to distribute into our ten classes. Hence, all classes are highly populated with the exception of the high depolarisation cases (in total 93) because highly depolarising layers are generally geometrically thin in the lidar and we face a result of our 1km altitude averaging here. Additionally the case C6 "thick aerosol" was underrepresented with 105 data points due to the generally clear conditions.
Behind the presented data of Table 3, there are about 106000 data points in the original resolution of 10min/60m. Note also that we are restricting ourselves here to the troposphere whereas even in 15km altitude the error in the BSR is about 0.05 and in the VD about 0.2% due to noise in the data.
All data were evaluated consistently, so systematic errors can be neglected for a relative classification. Hence we are very sure on our results. Some justification was added to the manuscript:

> Additionally, Fig. 8 gives an overview of the threshold values for the ten cases; they are empirical values based on previous data analysis. In the cases of Arctic Haze in previous years (partly unpublished yet) we clearly see a depolarization between 2-5% whereas cirrus observations suggest that a separation into medium and high depolarisation is justified. The backscatter ratio was divided into three categories: no particle backscatter above the background level (BSR<1.2), moderate backscatter (1.2<BSR<2) and strong backscatter (BSR>2).

25. P15468-l4&5. The sentence is not clear.

The sentence was removed from the manuscript.

26. P15469-l7. This is a description but what is the scientific interest? Can such observations lead to specific parameterization?

Yes, we hope that observations of these kinds will help to constrain the life time and removal processes of Arctic haze events. Currently we face the problem that the life time of tropospheric aerosol in the Arctic winter (dry, cold, very reduced wet removal) seems to be really long (quote Korhonen et al. 2008 is included in the new version).
With life times of almost 30 days, it becomes difficult to constrain the origin and chemical alteration of the old aerosol. Hence our motivation was to check in a statistical way whether differences of aerosol or cloud properties with trajectories can be seen. While we are convinced about the method, more data from enhanced aerosol events, both from direct transport from Eurasia as well as from aged aerosol from the inner Arctic should be included

in the future.
Recently, the ECMWF published aerosol "forecast" data (Morcrette et al. 2009, JGR 114 Aerosol analysis and forecast in the European Center for Medium Range Weather Forecasts integrated forecast system: forward modelling).
In comparison with data from our station, the ECMWF predicted AOD values differed significantly from the observational data. As the Arctic is a climatologically sensitive region clearly an improved understanding of Arctic clouds and aerosol events is needed.

27. P15469-l18. (sub-title 4.1) If C4 is considered to be medium in term of VD, why C7 and C8 are not in the same class?

C4 and C1 do not show enhanced backscatter and hence are not part of or case studies, this fact has to be described in the beginning of section 4:

> Within this section case studies supporting our classification scheme are given using the KARL data of five particular days with interesting cloud and aerosol structures (cf. Table 3). The cases C1 and C4 are not considered as there is no enhanced backscatter.

The title of the subsection will be changed to:

> Low and medium depolarization C2, C3, C5 and C6

28. P15470-l28. Why do you choose LR = 18 sr? Add the unity.

The unit sr was added. As the layers below and above were clear the LR was chosen such as to achieve the same BSR above and below the layer, similar to the "transmittance method" see
W. Chen, C. Chiang, and J. Nee, "Lidar Ratio and Depolarization Ratio for Cirrus Clouds," Appl. Opt. 41, 6470-6476 (2002)

29. P15473-l18. The aerosol layer is close to 6 km and this layer is not included in the interval in Figure 13 (except in “all points”). It is then difficult to distinguish the aerosol layer from the cloud layers.

The layer was situated at 6km when we first detected it in the evening of 14 March. Figure 13 shows data between 11:08 and 12:05 UTC on 15 March, when the layer already descended to about 5 km altitude. The data points referring to the aerosol layer are mainly found in C5, while the cloud data points show either less backscatter or more depolarization.

30. P15473-l25. How the aerosol parameters have been retrieved? It is not clear.

The inversion was performed with the lidar data of three elastic and two Raman channels. To clarify this in the manuscript the following sentence was included in the beginning of section 4.3:

> For the analysis of these aerosol cases all three elastic and both inelastic (387nm, 607nm) channels had to be considered according to Ansmann et al. 1992.

A. Ansmann, U. Wandinger, M. Riebesell, C. Weitkamp, and W. Michaelis, "Independent measurement of extinction and backscatter profiles in cirrus clouds by using a combined Raman elastic-backscatter lidar," Appl. Opt. 31, 7113-7113 (1992)

31. The conclusion could be reduced.

Some parts of the conclusion were skipped. (page 15477, line 2-9)