1. Introduction

We would like to thank the reviewer for his careful reading and helpful comments. The manuscript has been modified according to the suggestions proposed by the reviewer. Detailed replies on the comments are given below; comments of the reviewer are given italicized.

2. Major comments

2.1. Page 15455, Line 24. It is stated that aerosol removal processes depend critically on the amount of water vapor and atmospheric trace gases. This sentence should be made more scientific accurate. Aerosol removal occurs by dry and wet deposition.
Dry deposition depends on turbulence transfer and proximity to the ground, while wet deposition depends on cloud nucleation and precipitation scavenging.

We changed the paragraph as follows:

> While our knowledge about the pollution pathways into the Arctic improved significantly in recent years (Law and Stohl, 2007) a realistic aerosol forecast in this remote region is still not possible as aerosol removal processes, e.g. wet and dry deposition are insufficiently known. Wet deposition depends critically on the amount of water vapor and atmospheric trace gases whereas dry deposition depends on turbulent transfer and mixing which leads to lifetimes as long as 30 days for particles around 0.2μm size in the dry and stable wintertime Arctic troposphere (Korhonen et. al., 2008, Shaw, 1995).

2.2. Page 15457, Line 22. Cases with multiple scattering were removed. The number of removed hours or cases should be stated. Also indicate how many of these cases were thick cirrus compared to water clouds. What is the optical depth cut-off used to reject cases based on multiple scattering? This value or something similar should be provided.

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, see comment 2.6.

2.3. Page 15458, Line 22. Low level Arctic temperature inversions are described as “complicated to describe”. Why do you say that? The classical Arctic inversion is forced by strong radiative cooling of the surface. Also in Figure 1 it is not possible to see the inversions using the temperature contour plot. It might be better to show a contour plot of the temperature lapse rate since that it the main function of the figure. Also I would say that defining a surface temperature inversion by a temperature difference of only 0.25 K is too weak. Why did you use 0.25 K? Also the contour interval should be stated in the figure caption.

The figure was partly changed according to the referee’s advice. Surface temperature inversions of 0.5K instead of 0.25K were plotted, showing that the frequency of occurrence was much higher in March than in April. Using temperature lapse rate instead of temperature itself seems a good idea, but unfortunately the data are not more clearly arranged, also the additional information on the actual temperatures would be lost.

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.

2.4. Figure 2. Wouldn't it be more useful to show a contour plot of the relative humidity with respect to ice? Also the contour interval might be too small which is resulting in a "crushing" of the lines. Also state the contour interval in the figure caption.

We show the relative humidity with respect to water as this is the quantity measured by the radiosonde. The intervals in the plot were enhanced to 15% for better readability.

2.5. -Page 15460, Line 10. The PEP-Tracer model is used for computing back trajectories, but later in the paper the HYSPLIT model is used. Why are two models in use? I like the idea of using both models but they should be run for all cases and compared to access variability caused by the model type.

The cluster analysis was done using our own model, which could easily be applied to the demands of the paper and was written especially for the Polar Regions (no singularity at pole). Additionally, PEP-Tracer uses ensemble calculations and hence features error estimation. As the PEP-Tracer model does not contain any information on precipitation, we prefer to use the well known HYSPLIT model whenever comparing to certain probed air masses. However, a comparison of both models for the case studies is a good advice and will be included in the revised manuscript at the beginning of section 4.3:

> The trajectory calculations in this section were performed with the PEP-Tracer and the HYSPLIT model (Orgis, 2009, Draxler and Hess, 1998). The first one uses ensemble calculations and hence provides uncertainty information; the latter one was used to obtain information on precipitation.

Additionally, the plots in Sec. 4.3 will be redone to include information from the PEP-Tracer calculations.

2.6. Section 3. I found it confusing keeping track of which lidar is used in all the results and retrievals and why. The beginning of section 3 should have a short introductory paragraph clearly stating how the two lidars are being used, how they complement each other, what are the shortcomings of each lidar. Also there should be some comparison of perhaps the BSR from each lidar separately for a given profile to see how they compare. I realize that some of this information is scattered throughout the paper, but it would help the reader to have it all together at beginning of this section.

According to this comment we moved some introductory information to the beginning of section 3:

There are measurements of two different lidar systems available, the Koldewey Aerosol Raman Lidar (KARL) and the micro-pulse lidar (MPL). For the cloud statistics, the complete available dataset of the two lidar systems was analyzed. Both systems are described in section 3.1. The MPL data (1368h) were used for a comparison with the temperature and relative humidity profiles obtained with the proximate balloon sounding (Sect. 3.3.1) and in a second analysis to retrieve a general overview in terms of cloud altitudes (Sect. 3.3.2). The KARL data (145h) were examined to separate between aerosol and cloud events and to analyze trends in depolarization and backscatter values depending on altitude and corresponding backward trajectory.

3.1.1 Lidar systems

The micro pulse lidar (MPL) is a compact, continuously operating lidar system (Welton and Campbell, 2002, Spinhirne, 1993), that is running on a twenty-four hour operation basis at AWIPEV research station and maintained by the base personnel since June 2003. It uses a Nd:YLF laser (λ=523.5nm) and a Schmidt-Cassegrain telescope with 20cm diameter for laser transmission and receiving. The main parameters of the system are listed in Shiobara et al. (2003). The measured backscatter profiles cover a range of 60km with a vertical resolution of 30m and a temporal average of 1min. The system was working continuously over the whole 2 months period with some data losses (altogether 14 days) due to snow on the window, which has to be removed manually. For further analysis the data from 1 March until 30 April was averaged to means of 10min, background corrected and cut above 21km.

The Koldewey Aerosol Raman Lidar (KARL) is a Nd:YAG based Raman lidar. It measures the elastically backscattered light in three wavelengths (355nm, 532nm and 1064nm) as well as the N_2 shifted lines from the 2nd and 3rd harmonic. Additionally, a water vapor channel at 407nm as well as depolarization at 532nm is recorded. From the two N_2 shifted lines 387nm and 607nm, the extinction coefficient can be determined, making KARL a “3+2” wavelengths Raman lidar. In the ASTAR 2007 configuration the field of view of the 30cm mirror was 0.83mrad. In November 2006 a new laser (Spectra Pro 290-50), which works at 50Hz and yields more than 10W power at 355nm and 532nm and 20W at 1064nm, was installed in the lidar, increasing the energy output by a factor of 5 which significantly improved the data quality (Ritter et al., 2008). After beam widening the laser had an effective divergence of 0.5mrad. With 10min integration time and 60m height resolution the elastic wavelengths are easily evaluable up to 25km altitude in daylight conditions.

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.

The available KARL data set is much smaller than the MPL’s data set and a statistical evaluation of cloud occurrence in the strict sense is impossible.

> However, we analyzed the frequency of occurrence of different atmospheric structures and classified them according to BSR and VD.

2.7. Page 15462, Line 6. It is stated that the quotient in Eqn. 2 is normalized in clear air in 14 km altitude. This is confusing sentence that should be better explained.

The sentence was changed to the following:

> The quotient is normalized in the aerosol and cloud free stratosphere in 14km altitude assuming a background value of 1.4% which occurs due to Rayleigh scattering.

We think that this is justified, as the lidar profiles in the whole stratosphere closely follow the density profile. In volcanic quiescent periods (as in spring 2007) the tropopause / lower stratosphere region is free of aerosols. Only PSCs at temperatures around 195K occur sporadically in winter time.

2.8. Page 15463, Line 17. It says "For each altitude interval beginning at the surface, values above 0.1". What is this 0.1? It seems to be referring to BSR but BSR needs to be greater than 1. Please clarify.

The 0.1 is the difference between two adjacent BSR values; this is well above the noise limit and hence corresponds to real changes in the optical properties. The sentence was changed to enhance clarity:

> 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.

2.9. Section 3.1.1. What is meant by "Comparison with radiosondes"? The lidars and radiosondes are giving different variables. It is not a comparison unless the Raman channels were being used to determine a temperature and water vapor profile which is not the case in this paper.

The sections title was changed to:

> Correlation of radiosonde and MPL data

2.10. Page 15466, Line 25. Why is it expected that the backscatter ratio is decreasing with altitude?

On a clear day, the aerosol load near the surface in the PBL is higher compared to the free troposphere. It is our experience from KARL lidar data during the last years that apart from clouds (which can obviously be separated from aerosol) and "direct aerosol events" which can directly be connected to source regions via trajectory analyses (pollution events, biomass burning, Kasatochi volcano) the remaining background aerosol always decreases with altitude. This means that the optical influence of aerosol decreases stronger than the air density with altitude. Maybe this is a common feature for a site far away from significant source regions. We change the sentence into:

> At our Arctic site generally the backscatter ratio decreases with altitude.

2.11. Page 15466, Line 29. It is not clear how you except to use BSR to verify a climate model. Either explain this fully or leave it out. Perhaps it should go into the conclusions section.

The sentence will be removed.

2.12. Page 15468, Line 5. Why is it not possible that water clouds could decrease the VD to below 1.2?

Water clouds can decrease the VD to below 1.2 but the presence of water droplets would be characterized by a BSR larger than 1.2. To reduce confusion, the sentence was removed from the manuscript.

2.13. Page 15468, Last Paragraph. There is no discussion in this paragraph about the particle type associated with the C4 category even though C4 makes up about a third of the cases below 5 km. There is some speculation about C4 in the Conclusions on page 15477 but this should be moved forward to this section.

You are right, C4 should be mentioned here and the paragraph was changed to:

> The statistical results for the different altitude intervals are given in Table 4. Within the analyzed 145h of data, the fraction where neither enhanced backscatter nor volume depolarization was observed (C1) increases with height as well as the ice cloud fraction (C7–C10) up to 8km. Depolarization without noticeable backscatter (C4) makes up a third within the lowest 6km. Water clouds (C2), aerosols (C5/C6) and water clouds with a certain ice cloud fraction (C3–C6) decrease with height. With reference to the total number of detected clouds, pure water clouds (C2 1–7 km) account for 2.8%, mixed-phase clouds and aerosols (C3–C5 1–7 km) account for 70.0% and the ice cloud fraction (C3–C5 7–12 km and C7–C10 1–12 km) contributes to 27.2%.

2.14. Page 15469, Line 7. It says here that the fraction of clear sky measurements is lower from the Canadian Arctic while on page 15461, line 5 it says that trajectories from northern Canada are very clean. Explain this contradiction.

The term ‘clean’ air refers to air masses which contain a low amount of pollution but can contain ice or water particles. The term ‘fraction of clear sky measurements’ refers to measurements with no particle detection, e.g. BSR<1.2.

2.15. Page 15470, Line 7. It is stated that in one of the cases there is indication of a “subvisible water cloud layer”. I am not aware of this type of cloud being reported in the literature. For a water cloud to be subvisual it would require quite an extraordinary situation of limited condensation which does not seem possible. The authors should remove this phrase or provide a detailed account of this new type of cloud and strong evidence that it truly is a subvisible liquid cloud and not a haze layer. This is also referred to on page 15471, line 21 and page 15476, line 19.

We are aware that our statement of subvisible water layers at low temperatures is something unusual (cf. Intrieri, J. M. and Shupe, M. D.: Characteristics and radiative effects of diamond dust over the western Arctic Ocean region, J. Clim., 17(15), 2953–2960, 2004).
However, with March 8 and April 7 we present here two such subvisible clouds, a third one is discussed in Lampert et al, in this issue. In all cases the lidar clearly sees a short living (life time: order from minutes to few hours) increase of backscatter with low depolarization just in altitudes, where the radiosonde sees a supersaturation of humidity due to the temperature inversion (min temp. around 260K here or even somehow lower in Lampert et al). Arctic Haze has, according to our experience, clearly a higher depolarization, is accompanied by a

local minimum of humidity and, due to the distance of its source regions, is a longer living phenomenon (time scale: several hours to even days). We are aware that current climate models underestimate the existence of liquid water in the Arctic - see for example Verlinde, J. et al., 2007: The mixed-phase Arctic cloud experiment, *Bull of the Americ. Met. Soc*, **88**, pp. 205-221
We would like to keep the case March 8 in the manuscript to present these observations to a broader community. At the moment, unfortunately, we do not have more than these few examples; here definitely any additional information on number of aerosol / CCN / IN as well as temporal variation of temperature and humidity would be valuable.

2.16. Page 15473, Line 25. The inversion of lidar data from 2 wavelengths is used to retrieve the real and imaginary part of the refractive index, the number concentration, the effective radius and the modal width. It seems that this result is not possible from only 2 wavelengths. Boeckmann (2001) whom they cite states that at the very minimum one requires 3 backscatter wavelengths and one Raman channel. The retrieval algorithm inputs need to be explained in more detail.

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)

2.17. Did the lidar operate at the same time as aircraft over-flights as part of ASTAR 2007? If so would it be possible to validate the particle types by using aircraft measurements?

The closure between in situ and remote sensing instrumentations was indeed one of the aims of ASTAR 2007. However, the of characterizing Arctic Haze simultaneously by ground-based and airborne instruments could not be achieved during that campaign, mainly due to the lack of pronounced aerosol events. Only a few overflights of the airplanes over Ny-Ålesund were performed as the planes mainly flew far away to catch some aerosol and to characterize the atmosphere above the usually inaccessible open ocean and sea ice. The overview article of this special issue summarizes the overall situation.
We had only three contemporary overflights: on 3 and 11 April at very clear conditions: the overall optical depth was on both days 0.095 at 368nm and 0.065 at 532nm and the lidar did not see any clear aerosol layer so the aerosol load was too little to assure a trustful lidar inversion. The third overflight was on April 16 but the lidar was handicapped due to low clouds and had to be switched off. So, unfortunately a real comparison could not be performed. However, principally the inversion of remote sensing optical data can agree fairly well to in situ measurements of aerosol as was pointed out by

Wandinger, U; Müller, D; Böckmann, C; Althausen, D; Matthias, V; Bösenberg, J; Weiss, V; Fiebig, M; Wendisch, M; Stohl, A; Ansmann, A:* *Optical and microphysical characterization of biomass-burning and industrial-pollution aerosols from multiwavelength lidar and aircraft measurements
Journal of Geophysical Research. D. Atmospheres. Vol. 107, no. D21. 16 Nov. 2002

Due to this disadvantageous situation we excluded all aircraft activity from this paper. However, a real closure between remote sensing and in situ instrumentations will remain an important topic for future research.

3. Minor Issues

Page 15455, Line 1. This sentence has awkward grammar.

It was changed to:

> The Arctic is considered to be a sensitive indicator of climate change due to a large number of special interactions and feedback mechanisms (Curry et al., 1996). Especially aerosols and clouds have a significant influence on the solar and terrestrial radiation budget (Shupe and Intrieri, 2004; Quinn et al., 2007). Formation, evolution and dissipation of clouds as well as sources and sinks of aerosol particles are not yet entirely understood.
> Therefore, data collection by ground-based and airborne experiments in this remote region is important (Solomon and Qin, 2007).

Page 15455, Line 15. “Haze” should read “haze”.
Page 15456, Line 6. “a manpower” should read “manpower”.
Page 15456, Line 15. “Gulf Stream and hence warmer” should read “warm Gulf Stream”.
Page 15456, Line 18. “differing” should read “different”.
Page 15456, Line 19. “low” should read “less”.
Page 15456, Line 20. “higher values” should read “more”.
Page 15456, Line 22. “Even” should read “Also”.
Page 15457, Line 8. “parameters as” should read “parameters such as”.
Page 15457, Line 25. “inter annual” should read “interannual”.

Thank you for the corrections, they were included.

Page 15458, Line 6. Only 71 weather balloons were analyzed for a two month period. With twice daily launches there should be closer to 120 soundings. Why are there fewer sounders?

There is only one regular weather balloon each day at 11 UTC. Some additional balloons were launched during the ASTAR campaign and for the Ozon MATCH campaign.

Page 15458, Line 13. It is stated that relative humidity from the soundings are less reliable at cold temperatures. State at what temperature this becomes a major concern.

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.

Page 15464, Line3. “temporally following” should read “subsequent”.
Page 15464, Line 13. “can not” should read “cannot”.
Page 15465, Line 10. “fraction were” should read “fraction when”.

Comments will be included.

Page 15469, Line 18. Does the low and medium depolarization also include cases C3 and C5?

Yes, it does. This subsection title will be changed to

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

In the last preceding paragraph information on C1 and C4 was added:

> 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.

Page 15470, Line 1. There are no balloon soundings in Figure 10.

You are right; the reference to the figure will be removed.

Page 15470, Line 10. Typically water droplets can exist if the temperature is warmer than about -40C. Below 210K must be a misprint.

This was indeed a misprint and should be:

> Within the higher layer the temperature during the radio sounding was below 210K, prohibiting the existence of fluid water droplets.

Page 15474, Line 2. “insecurity” should read “uncertainty”.
Page 15475, Line 12. “approved” should read “confirmed”.
Page 15475, Line 19. “hydrophil” should read “hydrophilic”.
Page 15476, Line 8. “A statistics” should read “Statistics”.

Changes were adopted.

Page 15476, Line 14. “As never clear super saturation in clouds has been observed” is a confusing phrase. Please rewrite.

Rewritten sentence:

> From the comparison of MPL data with 54 radiosondes, we found some cases with relative humidity down to 60% sufficient for cloud existence especially in altitudes above 4km. Hence a dry bias of the RS-92 in extreme environments must be considered, similar to the findings of Rowe et al. (2008).

I suggest the word “statistical” be removed from the title.

The term will be removed.