Potential source regions and processes of the aerosol in the summer Arctic

J. Heintzenberg\textsuperscript{1,2}, C. Leck\textsuperscript{2}, and P. Tunved\textsuperscript{3}

\textsuperscript{1}Leibniz-Institute for Tropospheric Research, Permoser Str. 15, 04318 Leipzig, Germany
\textsuperscript{2}Department of Meteorology, Arrhenius Laboratory, Stockholm University, 10691 Stockholm, Sweden
\textsuperscript{3}Department of Applied Environmental Science, Stockholm University, 10691 Stockholm, Sweden

Received: 22 January 2015 – Accepted: 6 March 2015 – Published: 20 March 2015

Correspondence to: J. Heintzenberg (jost@tropos.de) and C. Leck (lina@misu.su.se)

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

Sub-micrometer particle size distributions measured during four summer cruises of the Swedish icebreaker Oden 1991, 1996, 2001, and 2008 were combined with concurrent gas data, back trajectories and daily maps of pack ice cover in order to investigate source areas and aerosol formation processes of the boundary layer aerosol in the central Arctic. With a clustering algorithm potential aerosol source areas were explored. Clustering of particle size distributions together with back-trajectories delineated five potential source regions and three different aerosol types that covered most of the Arctic basin. Clustering the open water information along the trajectories clearly separated marine from pack ice aerosol. Clustering the two subpopulations above and below the 50% value of average open water during the course of the trajectories yielded further details about marine and pack ice aerosol sources. In the pack ice subpopulation three potential source areas over the pack ice with very low percentages of open water from which aged aerosol with only one major mode around 40 nm derived. A significant difference between newly formed and aged aerosol over the pack ice became clear when comparing the course of open water under the trajectories for these two aerosol types. In both subpopulations the air had spent ten days over pack ice with less than 50% open water while traveling over ever more contiguous ice. Trajectories connected with high concentrations of newly formed small particles, however, experienced more open water during the last four days before arrival in heavy ice conditions at Oden. Thus we hypothesize that both, long travel times over the more contiguous ice, combined with more open water conditions during the last days before air mass arrival were an essential factor controlling the formation of ultrafine particles over the central Arctic pack ice. A comparison the Oden data with summer size distribution data from Alert, Nunavut and Mt. Zeppelin, Spitsbergen confirmed the Oden findings with respect to particle sources over the central Arctic.
1 Introduction

The investigation of the summer aerosol over the central Arctic Ocean began with the first Swedish Arctic icebreaker expedition (*Ymer*-80) in 1980 (Lannefors et al., 1983) followed up later in a series of four international ice-breaker expeditions to the summer central Arctic Ocean on the Swedish icebreaker *Oden* in the years 1991 (Leck et al., 1996), 1996 (Leck et al., 2001), 2001 (Leck et al., 2004), and 2008 (Tjernström et al., 2014).

As illustrated in Fig. 1, several hypothesized sources may contribute to the aerosol population over the central Arctic Ocean, and thus to the formation of low-level stratiform clouds and their effects on the surface energy balance. We observed long-range transported biomass burning or pollution plumes in helicopter profiles. These always occurred in the free troposphere well above the top of the boundary layer (Kupiszewski et al., 2013). Whether these particles can influence the aerosol in the boundary layer and thus the formation of low-level stratiform clouds over the pack ice in summer is an ensuing question. To address this question it is necessary to consider the vertical structure of the troposphere over the high Arctic in summer. The high Arctic boundary layer is typically a shallow well-mixed layer at the surface typically only a couple hundred meters deep, capped by a temperature inversion (Tjernström et al., 2012). At times the inversion may be strong, such as when there is substantial advection of warmer air from lower latitudes, while the free troposphere is stably stratified. In contrast to more southerly latitudes, deep convection, which could enhance mixing across the whole troposphere, does not occur other than possibly in frontal zones associated with passing weather systems. These characteristics limit the mixing of particles from elevated plumes in the free troposphere into the boundary layer. The only mechanism that can bring elevated plumes down to the inversion is large-scale subsidence, which is a very slow process. While free troposphere pollution plumes were observed we did not find evidence of any light-absorbing carbon particles at the point of measurements (25masl) onboard the icebreaker during any of the four cruises. The only available
light absorbing surface aerosol measurements over the summer pack ice indicated extremely low concentrations on the order of a few nanograms of black carbon per cubic meter (Heintzenberg, 1982; Maenhaut et al., 1996).

Thus long-range transported biomass burning or pollution plumes are not considered to have any direct significant contribution to the aerosol population in air layers close to surface (Kupiszewski et al., 2013). Instead, transport of precursor gases and marine biogenic particles from the marginal ice zone (MIZ) or locally from open leads over the pack ice has been found to result in raised concentrations of accumulation mode particles within the high Arctic boundary layer (Heintzenberg et al., 2006; Chang et al., 2011; Heintzenberg and Leck, 2012; Kupiszewski et al., 2013; Hellén et al., 2012; Nilsson and Leck, 2002; Leck et al., 2013). This may involve both direct emissions of primary larger accumulation mode marine particles, as well as growth of smaller particles via two processes, namely heterogeneous condensation and aerosol cloud processing.

The fact that near-surface airborne aerosol, as well as low-level cloud and fog droplets, contained the same type of polymer gel material as found in the open-lead surface microlayer (Gao et al., 2011; Leck et al., 2013; Orellana et al., 2011; Bigg et al., 2004) supports the hypothesis of a local aerosol source within the pack ice. The presence of bubbles in the water column (Norris et al., 2011) provides a plausible mechanism for getting surface material airborne. However, direct measurements of aerosol number concentration fluxes from the local open lead, (Held et al., 2011b), have failed to explain near-surface airborne aerosol concentrations observed simultaneously on the nearby Oden. Even though the pack ice open lead was a net source of aerosol particles, the snow surface on the surrounding ice was a net sink. Considering the regional ice fraction, this suggests that the surface as a whole may have been a net sink of aerosols in terms of total number concentration. However, a statistical analysis of aerosol observation from the four Arctic experiments on Oden, (Heintzenberg and Leck, 2012), suggests particle sources in the innermost Arctic. This inconsistency be-
between direct observations of local aerosol flux and statistical interpretations of aerosol properties and concentrations remains an important issue to resolve.

The same biological processes, providing the gel material over nearby open leads, could also be active upwind of Oden, from near the MIZ and the open ocean beyond. If mixed through the deeper atmospheric mixed layer over the open water (ca 500 m), these particles could be advected in over the central Arctic on top of the shallow local boundary layer (Tjernström et al., 2012), while efficient scavenging processes associated with low clouds and fog near the MIZ (Nilsson and Leck, 2002; Heintzenberg and Leck, 2012) may explain the very low near-surface aerosol concentrations. Aerosol particles or their precursors in the upper layer could potentially be advected over long distances and later be entrained into the local boundary layer through the cloud top by cloud induced mixing (e.g., Shupe et al., 2013).

Heterogeneous condensation and aerosol cloud processing occurs when the oxidation products of dimethyl sulfide (DMS) released by phytoplankton advected from open waters south of and along the marginal ice edge, (Leck and Persson, 1996a), condense on non-activated particles which then are incorporated into cloud droplets. In the latter droplets liquid-phase oxidation of absorbed gases can add further material to the droplet constituents. Once cloud droplets evaporate they leave behind raised concentrations of accumulation mode particles, grown via the two processes described. In the process, the bimodal particle size distribution characteristic of cloud-processed air is created (Hoppel et al., 1994). Boundary layer clouds have also been found to reduce accumulation mode particles occasionally as few as 1 cm$^{-3}$, due to the high capability of these particles to act as CCN, and their uptake via cloud droplets, resulting in wet deposition of accumulation mode particles, as suggested by Lannefors et al. (1983), Bigg et al. (2001), and Mauritsen et al. (2011).

New particle formation (nucleation) occurred about 15% of the observed time period (Karl et al., 2013). However, these events often occur as a simultaneous increase of particle number concentrations in the <10 nm and 20–50 nm size ranges, and not as the prototypical “banana growth” (e.g., c.f. Kulmala et al., 2001). Conventional nu-
cleation paradigms (Karl et al., 2012) fail to explain this phenomenon. Simultaneous concentration increase at several discrete sub-micrometer particle sizes could be due to vertical mixing of air from different levels above the surface, with different particle size distributions coming from different source regions. An alternate hypothesis explaining this could be fragmentation and/or dispersion of primary marine polymer gels, 200–500 nm diameter in size, into the nanogel size fractions down to a few nanometer polymers (Karl et al., 2013; Leck and Bigg, 2010); this appears to be consistent with the finding of a particle source in the central Arctic being most pronounced in the smallest particles sizes below 26 nm in diameter (Heintzenberg and Leck, 2012). Fragmentation would be promoted with exposure to ultraviolet light (Orellana et al., 2011) and long travel times over the pack ice. Leck and Bigg (2010, 1999) also suggested that disruption of particles by electric charge, such as electrospinning (Reneker and Chun, 1996), might provide an appropriate fragmentation mechanism. This appears to be consistent with observation since it would be favored by evaporation of cloud or haze drops (e.g., Heintzenberg et al., 2006). Fragmentation hypotheses may also explain why only a few percent of the observed total particle number variability was explained by the direct measurements of particle number fluxes (Held et al., 2011a).

While the four expeditions provided a wealth of new observations and understanding of the system of low-level clouds, their formation, and their effects on the boundary-layer and surface energy balance over the Arctic pack ice area, the ultimate partitioning of aerosol particles among potential source regions and processes remains elusive. The present paper continues the analysis of the aerosol data from the four Oden cruises with a focus on the above discussed potential source regions and related aerosol formation processes. The ship positions during the cruises shown in Fig. 2 indicate that the measured data only cover a small part of the European Arctic sector. However, with back trajectories the data coverage can be extended over the whole Arctic basin. This approach was first followed with aerosol data measuring during the Ymer-80 expedition by Jaenicke and Schütz (1982) and with Norwegian Arctic aerosol data by Heintzenberg and Larsen (1983). For the present study back trajectory information
was complemented with daily maps of ice concentrations. Sections 2.3 and 2.4 gives more details. For the combination of aerosol and information of air origin a dedicated cluster algorithm was developed, tested with trace gas data from the four *Oden* cruises, and applied to meteorological and aerosol data. For a test of the clustering algorithm the aerosol database was complemented with the data on atmospheric dimethyl sulfide (DMS(g)) concentrations taken during all four cruises (Leck and Persson, 1996b; Kettle et al., 1999).

Moreover, to date, 23 years after the first *Oden* expedition there are still no other surface aerosol data from the central Arctic to compare with. The nearest land stations are Mt. Zeppelin, Spitsbergen and Alert, Nunavut. The present paper therefore also make an attempt to use the size distributions taken on *Oden* and the clusters derived with them and connect them with size resolved aerosol number data and trajectories from these two land stations.

2 Experimental data

2.1 Sampling conditions on icebreaker *Oden*

An identical sampling manifold was utilized in all four icebreaker expeditions upstream of all gas phase and aerosol measurements. The sampling manifold extended at an angle of 45° to about three meters above the container roof of the laboratory container on *Odens*’ 4th deck to optimize the distance both from the sea and from the ship’s superstructure. The height of the sampling manifold was ~ 25 m a.s.l. and consisted of two masts (PM$_1$: Diameter $< 1 \mu$m and PM$_{10}$: Diameter $< 10 \mu$m), with one additional sampling line for volatile organic compounds including DMS. Direct contamination from the ship was excluded by using a pollution controller. Provided that the wind was within $\pm 70^\circ$ of the direction of the bow and stronger than 2 m s$^{-1}$, no pollution reached the sample inlets. Further details of the location and properties of air intakes and instru-
ments, position on the ship, pumping arrangements and precautions to exclude contaminated periods can be found in Leck et al. (2001) and in Tjernström et al. (2014).

2.2 Data collected onboard Oden

2.2.1 Gas data

As compared to the 1271 hourly DMS values, which were concurrent with contamination-free aerosol data a total of 2035 h of DMS data were available in the four cruises for clustering.

During the expedition in 1991, integrated samples of DMS were analyzed by a Gas Chromatograph (GC)-Flame Photometric Detection (FPD) system where glass-fiber-wool cold-trap was used in the pre-concentration step. The sampling duration was 20 min (Persson and Leck, 1994). During the three subsequent cruises in 1996, 2001 and 2008, DMS was automatically collected with a time resolution of 15 min and pre-concentrated in the following two steps: first, a gold trap (gold wire in a Pyrex glass tube) for collection, and second, a (TENAX®) medium to achieve a sharp injection of the analyte into the GC-FPD. To remove atmospheric oxidants prior to collection, a high-capacity scrubber based on 100 % cotton wadding was used (Persson and Leck, 1994) in all four cruises. The overall accuracy, valid for both GC-FPD methods described above, was within ±12 % with a detection limit of 0.045 nmol m⁻³.

To further improve on time resolution, we added a Proton Transfer Reaction Mass Spectrometer system (PTR-MS) (Lindinger and Hansel, 1998) during the 2001 cruise with a sampling frequency of 2 min and in the 2008 experiment DMS was measured a PTR-TOFMS (Aerosol Time of Flight Mass Spectrometer) built at Innsbruck University. The PTR-TOFMS was calibrated by applying a dynamically diluted DMS gas standard (Apel & Riemer Environmental Inc); zero-calibrations were performed every 2–6 h using catalytically scrubbed air. The sampling frequency of the PTR-TOFMS system was 1 min. The instrument is described in detail in (Graus et al., 2010). For the benefit
of time resolution using the PTR systems, the detection limit was increased by a factor of ten to 0.45 nmol m\(^{-3}\). 1 nmol m\(^{-3}\) = 22.4 ppt(v) at 0°C and 1013.25 mbar.

### 2.2.2 Aerosol data

The *Oden* database is essentially the same as described in Heintzenberg and Leck (2012) with 2645 h of sub-micrometer particle number size distributions between 5 and 560 nm diameter. Tandem Differential Mobility Particle sizers (TDMPS) were used to measure the number size distributions of dry sub-micrometer particles used pairs of very similar differential mobility analyzers (DMAs) in all years. The TSI 3010 counters used in the DMAs were size and concentration calibrated against an electrometer and the TSI 3025 counters for particle sizes below 20 nm diameter in the standard way (Stolzenburg, 1988). In 1996 a second, modified TSI 3010 was utilized to extend the data from 20 to 5 nm instead of a TSI 3025. The harmonized size range for all cruises comprised 36 channels, which were spaced in equidistant fashion on a logarithmic scale. Before taking hourly averages the data had been cleaned thoroughly for possible pollution from the ship (cf. Heintzenberg and Leck, 2012, for details). In 1991 the Arctic part of the cruise covered the time from 18 August through 26 September. In 1996 the icebreaker stayed in the pack ice region from 26 July to 4 September. The corresponding period in 2001 was 10 July through 25 August, and in 2008 4 August through 5 September. A total of 2645 h of aerosol data remained (cf. Table 1)

### 2.3 Aerosol data from Arctic land stations

During the last two cruises in 2001 and 2008 sub-micrometer size distribution measurements were taken at the observatory on Mt. Zeppelin 78.9°N, 11.86°E; elevation 474 m.a.s.l.) (Tunved et al., 2013). For comparison with the *Oden* data 1968 hourly average number size distributions in 20 diameter channels from 20 to 600 nm were available with concurrent five-day back trajectories.
The Dr. Neil Trivett Global Atmosphere Watch (GAW) Observatory at Alert, Nunavut (82.5° N, 75° W; elevation 210 m a.s.l.) is the only other site close to the central Arctic with comparable aerosol measurements, i.e. regular sub-micrometer particle size distribution measurements since 2011 (Leaitch et al., 2013). Thus, no Alert size distributions are available during any Oden cruise. Instead, Alert data during the core month August of the Oden cruises will be utilized for comparison. Specifically, we have 1517 hourly average size distributions in 54 channels between 10 and 500 nm during the Auguts of 2011, 2012, and 2013 with concurrent five-day back trajectories arriving at 250 m over Alert.

2.4 Back trajectories

Three dimensional back trajectories have been calculated for the three different receptor sites used in this study: to the moving receptor ship Oden arriving at 100, 500 and 1000 m above sea level (a.s.l.), to the Zeppelin observatory located at the Zeppelin mountain near Ny Ålesund, Svalbard, at 474 m and to Alert at 250 m. The trajectories have been calculated backward for 10 days using the HYSPLIT2 model (Draxler and Rolph, 2003) with meteorological data provided by NCEP/NCAR project for years 1991–1996 (for more information consult http://www.esrl.noaa.gov/psd/data/gridded/data.nmc.reanalysis.html). For 2008 we applied the HYSPLIT4 model with GDAS data (Global Data Assimilation System). More information about the GDAS dataset may be found at Air Resources Laboratory (ARL), NOAA (http://ready.arl.noaa.gov/), where meteorological data also can be downloaded.

Geographic results were plotted on maps of stereographic projection centered on the North Pole. These maps were covered with a grid of 35 × 39 geocells, in which the passage of trajectories or the occurrence of other results of this study were counted. Figure 2 shows that the geographical region covered by the back trajectories extends to and partly beyond the pack ice limits of the studied summers. We are aware of the limitations in trajectory accuracy. On one hand the data sparse Arctic region limits the validity of the meteorological fields on which the trajectory calculations are based. On
the other hand, out to the nearest continental borders the meteorological setting, surface conditions and the resulting atmospheric fields in the central Arctic are relatively simple. Figure 9 in Leck and Persson (1996b) shows an example where the trajectories were able to resolve an influence of the settlements Barentsburg and Longyearbyen on Spitsbergen in the measurements onboard Oden which was located near the North Pole. If we assume some 30% position uncertainty relative to the trajectory length (on average 3000 km for a ten-day back trajectory) this will in general not allow us to differentiate between distant regions such as Beaufort Sea, Chukchi Sea, and Laptev Sea outside the pack ice. A distinction between these sees and Kara Seas is however possible. The meteorological information calculated along the trajectories was utilized in the analysis.

2.5 Ice data

Daily ice concentrations were taken from the NSIDC database (https://nsidc.org/data). Due to the orbits of the ice sensing satellites the area north of about 86° N was not covered. We assumed 100% ice cover in this area. The ice data were interpolated for each hour along all back trajectories because the maps of ice concentrations for the four cruises given in Fig. 3 clearly show that not only did the extent of the sea vary considerably over the 17 years time of the whole data set but also strongly within the study area. As integral parameters the average sum of open water in percentage of each back trajectory were calculated and will be referred to: (a) OS5, % (shorter than five days before arrival at Oden), (b) OG5 % (greater than five days). From the cruise-average gridded ice concentrations rough average ice limits were calculated for each cruise. For that purpose contiguous lines of 10% ice concentrations north of 76° N were formed and added to Fig. 2 and to maps of individual cruise years.
3 Clustering approach of aerosol and trajectory data

The basic clustering algorithm has been introduced in Heintzenberg et al. (2013). For the present study the input aerosol data were pre-processed with the common Standard Normal Variate (SNV) transformation by subtracting their respective grand average and dividing them by their respective standard deviations. The same SNV transformation was applied to the trajectories after projecting them onto a stereographic map centered on the North Pole. The clustering algorithm collects the clustered data in up to nine clusters based on different input information or coordinates:

- $x$, $y$, and/or height information of the projected trajectories,
- Percentage open water along the projected trajectories, and
- Particle number size distributions.

The algorithm can utilize any combination of these three sets of clustering coordinates, i.e. the projected horizontal coordinates of the trajectories or their combination with their height coordinates and/or open water information could be clustered but also their height coordinates alone. In each case the resulting clusters of aerosol properties are calculated if available. Vice versa, aerosol properties could be clustered and for each of such clusters the resulting trajectory clusters are calculated. Finally, clusters can be sought based on aerosol, trajectory, and ice information.

The search algorithm is constrained by the four parameters $N_{\text{init}}$, $X_{\text{av}}$, $P$, and $C_{\text{fin}}$. $N_{\text{init}}$ sets the initial minimum number of members, i.e. hourly data points that any cluster is required to have before further processing. The parameter $X_{\text{av}}$ is defined according to

$$X_{\text{av}} = \frac{\sum_{j=1,N_i} \sum_{k=1,m} (x_{j,k} - \bar{x}_{j,k})^2}{m \cdot N} \tag{1}$$

$m$ is the number of coordinates to be clustered. If particle size distributions are clustered $m$ corresponds to the number of diameters. $N$ is the number of members in the...
respective cluster. \( x_{j,k} \) is the coordinate \( k \) of cluster member \( j \) and \( \bar{x}_{j,k} \) is the corresponding average cluster coordinate. The average distance of cluster members from a cluster average coordinate stays below a set upper limit of \( X = X_{av} \). The similarity of cluster members can be improved by eliminating outliers in order of their distance from the cluster average.

Initially, the algorithm allows the input data to be segregated into a maximum of nine clusters. The algorithm will then eliminate the cluster with the maximum value of \( (x_{j,k} - \bar{x}_{j,k})^2 \) for any \( j \) and \( k \) until the number of cluster members is reduced to \( P (P \leq 1) \) times the initial number of members. Finally, the clusters will be compared to each other in order to eliminate cluster \( i \) with the minimum difference \( X' \) of average coordinates from any other cluster \( j \)

\[
X' = \frac{\sum_{k=1,m} (\bar{x}_{i,k} - \bar{x}_{j,k})^2}{m}
\]

until a given final number of clusters \( C_{fin} \) is reached. The non-sequential cluster numbering in the results discussed below reflects this elimination process, i.e. any cluster number missing in the results was eliminated in this process.

Tests of the cluster algorithm with Arctic 10 day trajectories only yielded clusters with very few members. Meteorologically this finding is easily understood: After a short time very little similarity in air pathways extending over ten days can be expected. Consequently, we limited all clustering experiments involving trajectories to five days. In Fig. 2 we see that the Oden cruises mainly covered the European plus western Russian sector of the inner Arctic. The trajectory coverage in Fig. 2 also shows that air from the longitudinal sector opposite to the Oden tracks, i.e. longitudes from about 150 to about 230° partly took more than five days to reach Oden. Due to meteorological variability, transport pathways from this sector to the measuring point were less similar than in other Arctic sectors and within five days the clustering algorithm could not often find many similar trajectories. Thus, in order not to miss potential source regions in this sector a conventional longitudinal sector cluster named “LC” was added to the...
algorithm that combined all unclustered data, the back trajectories of which had spent at least three days in this sector.

Any aerosol clustering experiment lies in between two extreme approaches. In the first one as many members as possible with somewhat similar properties are combined in each cluster, trying to cover the total data set as completely as possible with as few clusters as possible. Considering aerosol dynamics and the multitude of atmospheric processes much information will be lost in this approach. The other extreme clustering approach would attempt to be as specific as possible considering either air history and properties or aerosol properties in order reveal as much information as possible about potential aerosol source regions and formation processes. For the present study the clustering was directed towards the second extreme while trying to maintain sufficient coverage and statistical relevance in order to allow general conclusions.

For the geographic spread of the trajectories of any derived Cluster $i$ the metric $X_i$ is defined as

$$X_i = \frac{1}{n_j} \sum_{j=1}^{N_i} n_j$$

with $n_j$ being the number of trajectory hits in any of the $N_i$ geocells that are being crossed by trajectories of the respective cluster. The wider (and less regionally specific) the trajectory distribution of a cluster is the larger becomes $X_i$, and the more trajectories pass through any one cell, the narrower the spread becomes. Taken over all cruise years the 5 day back trajectories cover a total of 554 geocells. The corresponding number for 10 day trajectories is 870. Thus, with $X_i$ the fraction of possible geocells covered by the trajectories of any cluster $i$ can be visualized.

In subsequent maps potential source regions are identified by different colors. However, each geocell can only have only one color per map. Thus, as a measure of overlapping regions the parameter $P_{\text{unique},i}$ is calculated as a parameter quantifying the uniqueness of the geographic area of Cluster $i$. $P_{\text{unique},i}$ is the sum of the $N_i$ geocells that are being crossed by trajectories of Cluster $i$ but of no trajectories of any other
 cluster; the sum is divided by $N_i$ and reported in percent. To sharpen this parameter only geocells that have been passed by a minimum number of trajectories (usually 25) are being counted. Assuming independent trajectory distributions 25 hits per geocells would correspond to a 25 % uncertainty.

The quality of the particle size distributions in the derived clusters is described in two ways. With $\chi_{j,k} = \frac{dn(d\log D_p)/d\log(D_p)}{dn_j,k}$ being the differential number concentration of cluster member $j$ at diameter $k$ and $\overline{dn_{j,k}}$ being the arithmetic cluster average of $dn_{j,k}$ the similarity of particle size distributions can be quantified for each cluster $i$. Additionally, in the graphical display of cluster-average size distributions the size-dependent standard deviations of the cluster averages are shown.

4 Test of the clustering with DMS

Leck and Persson (1996a, b) reported the evidence for the existence of a substantial DMS source located at the fringe of the central Arctic Ocean just along the MIZ, which is released to the atmosphere from the upper most ocean. This as a result of the melting ice, which is favorable for the production of the DMS precursor dimethyl-sulfoniopropionate, released by the marine microbial food web. By using a three-dimensional numerical model it was clearly shown that DMS(g) is advected over the pack ice in plumes originating from the source at the ice edge or in the adjacent sea just south there of (Lundén et al., 2007), the Barents and Kara Seas being in particular strong source areas.

The seasonal decrease of about 30 % per week for the DMS in water and air, between mid July and end of September, was found to reflect the spatial and seasonal changes of the biological activity in the bordering seas adjacent to the mostly ice covered central Arctic Ocean. The overall seasonal decline was reduced to about 20 % per week if photochemical losses during advection over the pack ice were taken into account. DMS(g) undergoes oxidation by the hydroxyl radical (OH, daytime) and the nitrate radical (NO$_3$, nighttime) (Karl et al., 2007) With the use of a pseudo-Lagrangian
approach and an analytical model Nilsson and Leck, 2002 estimated the DMS(g) photochemical turnover time to ca 2.4 days. In all, the above observations showed that over the pack ice area, local contribution to the atmospheric DMS concentrations can be assumed to be negligible compared to that advected from the marine source south along the ice edge.

The 5 day back trajectories (vertical dimension excluded) of all 2035 h in all four cruises with DMS(g) data were clustered in experiment “DMS”. Run parameters of this and all other clustering experiments are listed in Table 2. Four well-separated trajectory clusters were found. An additional fifth cluster “LC” comprised the unclustered data in the longitudinal sector defined in Sect. 3. On average over all years the five clusters cover 22 % of the DMS data. Key data of the five clusters are collected in Table 3. The regional distribution of the trajectories in the clusters is plotted in Fig. 4. The trajectories of the two DMS(g) clusters with highest median DMS(g) values, (Clusters 2 (red), and 3 (yellow)), clearly point towards highly source rich ice-free areas of Greenland Sea, and Barents Sea also identified by Leck and Persson (1996a, b) and Lundén et al. (2007). Median DMS concentrations in the five clusters are listed in Table 3. The high average percentages of open water under the related trajectories, (parameter OS5 in Table 3), corroborate these results. The remaining Clusters 7, 8, and “LC” exhibited low DMS(g) values together with low percentages of open water. The cluster test with DMS(g) identified realistic potential source regions in the MIZ and adjacent open waters lending confidence in searching for other potential source regions of the surface aerosol over the Arctic summer pack ice.

5 Regional distribution of potential aerosol source areas

Encouraged by the results of the test of the clustering algorithm introduced in the previous section we sought clusters of similar parameters in our total data set covering 2645 h in four Arctic summers. The combination of horizontal trajectory information and particle size distribution was segregated into five clusters covering 25 % of all hourly
The regional trajectory distributions of these clusters are plotted in Fig. 5 together with the average size distributions of the clusters. The trajectories of the five clusters cover different areas of the central Arctic and the open waters of the adjacent Arctic seas. All clusters except the longitudinal sector Cluster “LC” have geographic coverage values $X_i$ of $\approx 20$ or less. The uniqueness parameter $P_{\text{unique},i}$ of the five clusters are 93, 66, 35, 44, and 40 %, respectively, i.e. only 7 % of the geocells of Cluster 1 are hit by trajectories of other clusters whereas 65 % of the geocells of Cluster 4 are passed by trajectories of other clusters as well. Low-level advection of air from the open waters of the Barents Sea ($OS5 = 65$ %) yields the typical bimodal marine size distributions found in Cluster 1 (Heintzenberg et al., 2004). Its median total number of 110 cm$^{-3}$ is lower than the typical 250 cm$^{-3}$ found for remote marine regions in lower latitudes (Heintzenberg et al., 2004). Cluster 3 with potential source region over the Kara Sea has a similar bimodal shape, albeit with a much lower median total number of 37 cm$^{-3}$. The small tail of the average number size distribution of Cluster 3 towards the lower size limit indicates the occurrence of new particle formation in its potential source region, which is largely ice covered ($OS5 = 26$ %). Cluster 4 stems from a potential source region north of Greenland and around the North Pole with extremely low values of open water ($OS5 = 7$ %, with the caveat of limited satellite coverage). In its median total number of 60 cm$^{-3}$ the accumulation mode comprises but a small shoulder. The more distant Cluster 5 is located in the pack ice covered region of the Beaufort Sea and the Canadian archipelago ($OS5 = 16$ %). While the average particle number size distribution associated with Cluster 5 is similar in shape to that in Cluster 4 the median total number in this cluster is 80 cm$^{-3}$ with somewhat higher median concentrations of particles below 10 nm referred to as ultrafine particles. The average particle size distribution of Cluster “LC” differs strongly from that of the other clusters: Not only is the total number about twice as high as in any of the other clusters but these high numbers also are found at small diameters than in any of the other clusters, i.e. largely below 30 nm. Median open water percentages below 30 % in Cluster “LC” clearly demonstrate that the air masses with such high concen-
trations of ultrafine particles have spent long times of the pack ice. In general, median open water percentages in cluster experiment “All aerosol” differed strongly in between the clusters indicating that the amount of open water may be a controlling factor on the particle size distributions measured on *Oden*.

Besides the trajectories the ice maps yield the only system parameters that cover the whole Arctic basin. Figures 2 and 3 clearly show that both, the limits and internal variability of the Arctic pack ice varied strongly. In order to explore this controlling factor we clustered the open water information along the trajectories for the total data set in the experiment “Open water” and found two groups of clusters each with systematic differences of open water percentages in between the groups. Figure 6a collects all trajectories of the group “Marginal ice” whereas Fig. 6b comprises the trajectory distribution of the group “Pack ice”. With some overlap in the marginal ice zones reaching from Greenland to the Laptev Sea the geographic regions of the two subpopulations are largely complementary.

In group one, named “Marginal ice”, comprising Clusters 4, 5, 7, and 8 (cf. Table 3), the median values of OG5 and OS5 were 95 and 53 %, respectively whereas the corresponding values of OG5 and OS5 were 27 and 25 %, respectively in the group two, named “Pack ice”, comprising Clusters 1, 2, 3, and 6. The clusters of the open water experiment will not be considered in detail further down. Instead a reclustering within the two groups will be discussed next.

As in experiment “All aerosol” we clustered horizontal trajectory information and particle size distributions in the subpopulation “Marginal ice”, comprising 787 h after constraining the data input by requiring both, OG5 and OS5 being greater than 50 %. 26 % of the data were collected in the four clusters displayed in Fig. 7 in terms of their average number size distributions together with the geographic distributions of the respective cluster trajectories. Clusters 1 and 2 exhibit typical bimodal marine size distributions as already found in clusters 1 and 3 in experiment “All aerosol” (cf. Fig. 5), albeit with more distinct potential source regions (cf. parameter $P_{\text{unique}}$ in Table 3) over the open waters of Kara Sea and Laptev sea (Cluster 1), and North Atlantic and Bar-
ents Sea (Cluster 2). The high median DMS concentrations in the latter cluster (cf. Table 3) reflect the highly productive open waters of the respective source region.

The trajectories of Cluster 3 come from the northern part of Greenland and the average trajectory height of 1600 m during the last five days before trajectory arrival clearly point towards a free tropospheric origin of this cluster. This character is also reflected by its average number size distribution in Fig. 7c, which is essentially monomodal with its peak around 40 nm. The relatively high number concentrations, (median \( N_{TO} = 130 \text{ cm}^{-3} \)), indicate a long-range transport of polluted air in the free troposphere (e.g., Leaitch and Isaac, 1991; Parungo et al., 1990).

The average number size distribution of Cluster 4 is shown in Fig. 7b and reflects another special case of input of polluted air into the pack ice region. The small diameter of 22 nm of its main peak indicates a rather fresh aerosol generated in the air mass that passed over Spitsbergen. Due to the low average trajectory travel height of some 500 m the air seemingly picked up a small accumulation mode around 150 nm. A more detailed analyses of a similar case is discussed in Bigg et al. (1996) and Leck and Persson (1996b).

Next, we discuss in more detail the subpopulation “pack ice” (cf. Fig. 6b). Clustering the open water information along the trajectories yielded the four Clusters 1, 2, 3, and 6 which according to their OS5 values clearly were associated with the inner pack ice region (cf. Table 3). Their cluster-average size distributions and respective geographic trajectory distributions are displayed in Fig. 8. The monomodal size distribution with low total numbers of Cluster 1 in Fig. 8a strongly reminds of the aged aerosol in Clusters 4 and 5 in experiment “All aerosol”. Cluster 1 practically covers the whole pack ice region in Fig. 6b, i.e. this type of aged aerosol may appear all over the inner Arctic. While the geographic distributions of Clusters 2, 3, and 6 largely are located in the same inner pack ice region their size distributions in Fig. 8b look very different. Several peaks below 50 nm appear with high number concentrations (up to 900 cm\(^{-3}\) at sizes down to the lower diameter limit of the instruments). This type of aerosol strongly likens that of Cluster LC in the experiment “All aerosol” (cf. Fig. 5).
Finally we explored in greater detail the large geographic region of Cluster 1 in experiment “Open water” by reclustering its 636 h of aerosol data with the information of horizontal trajectories and size distributions in experiment “Pack ice low”.

The results are plotted in terms of average size distributions and geographic distributions of trajectories in Fig. 9. Similar to Clusters 4 and 5 all clusters of the experiment “Pack ice low” have one main number peak around 40 nm and a varying second mode around 100 nm which may indicate some cloud processing. The similarity in size distribution while being associated with different potential source regions is due to the fact that the prescribed tight ice conditions occurred in different areas of the pack ice in different years (cf. Fig. 3).

6 Comparison with the nearest land stations

As pointed out in Sect. 1, 23 years after the first Oden expedition there are still no other surface aerosol data from the central Arctic to compare with. The nearest land stations are Mt. Zeppelin, Spitsbergen, (78.9° N, 11.86° E), and Alert, Nunavut, (82.5° N, 75° W). In this section the size distributions taken on Oden and the clusters derived with them and with back trajectories will be connected with aerosol data and trajectories from these two land stations.

6.1 Comparison Oden/Mt. Zeppelin

For a first comparison of particle size distributions observed at the location of the icebreaker Oden and at Mt. Zeppelin during the summers of 2001 and 2008, and the back trajectories to Mt. Zeppelin were employed. The closest points with distances less than 360 km between a trajectory point to the concurrent position of the icebreaker were sought along each trajectory. A total of 296 h fulfilled this condition with an average travel time between Oden and Mt. Zeppelin of 36 h and an average minimal distance between back trajectory and Oden of 177 km. Size distributions measured on Oden at
the time of minimal distance were compared to size distributions measured on Mt. Zeppelin at the time of trajectory arrival. Figure 10 gives the statistics of this comparison in terms of 25, 50, and 75 % percentiles.

Absolute concentration levels, and the shapes of the size distributions with their main peaks roughly compare at the two points, encouraging further investigations. In all three percentiles a similar systematic change is apparent in Fig. 10. During the travel from the more central pack ice covered Oden area to Mt. Zeppelin concentrations decreased at all diameters larger than some 30 nm, which could be due to cloud scavenging in the marginal ice zone.

Encouraged by this statistical comparison of trajectory-connected data at the two stations cluster experiment “Oden-Zeppelin” was set up, clustering the combined particle size distribution data from the two stations. For this experiment the size distributions on Oden and at Mt. Zeppelin had to be harmonized. The Zeppelin data of the two larges channels (501 nm, and 631 nm) were interpolated at the largest Oden diameter of 570 nm. All Oden data were interpolated at the more coarsely spaced Mt. Zeppelin channels between 20 and 570 nm. This harmonization yielded size distributions with 15 common diameter channels plus 11 channels from 5.1 to 20 nm that were only measured on Oden set to “missing data” at Mt. Zeppelin and were not utilized in the clustering algorithm. Averages of three very different clusters of combined size distributions are shown in Fig. 11.

Despite our disregarding in this cluster analysis any direct trajectory connection we derive quite similar cluster-average size distributions in terms of shape and absolute concentrations. Because of its lower size limit the Mt. Zeppelin instrument could not detect freshly formed ultrafine particles. However, the steep rise towards 20 nm in Cluster 2 of the Mt. Zeppelin data in Fig. 11 is in good agreement with the right flank of the main peak about 15 nm that only shows up in the Oden data.

For each cluster the geographical distribution of five-day back trajectories were calculated. For the Oden data in the clusters we utilized trajectories at 500 m arrival height in order to be more compatible with the Mt. Zeppelin trajectories arriving at 474 m. Com-
mon potential source areas were explored by plotting the average relative occurrence of trajectory points only in geocells that were hit by back trajectories to both stations. Figure 12 presents the geographical distribution of jointly hit geocells for the three clusters in Fig. 11. The high standard deviations of the Oden data in cluster 1 below 20 nm indicate the rather episodic occurrence of ultrafine particles. Thus, we separated two cases of potential source areas for cluster 1 in the Oden data, one for all cases with number concentrations below 10 nm (N10) = 0 cm$^{-3}$ measured on Oden, and one for all cases with N10 > 1 cm$^{-3}$.

The cases of newly formed ultrafine particles were only connected with air masses from the central Arctic. Except for one geocell north of Nordaustlandet, Svalbard Cluster 2 with its main peak around 15 nm also was connected with air from the central Arctic. Only cluster 4 had back trajectories leading out of the pack ice limit into the North Atlantic.

6.2 Comparison with Alert, Nunavut

With the cluster experiment “Oden-Alert” commonalities were sought in the shape of the size distributions measured on Oden and at Alert. For this exercise the data from both sites had to be harmonized in a fashion similar to the corresponding exercise with Mt. Zeppelin data. The higher resolution Alert data were interpolated at all possible diameters of the Oden data (11 to 435 nm). The interpolation yielded aerosol data at 34 common diameters, which could be clustered. Disregarding the fact that they were not synchronized we pooled the harmonized data from both sites into a set of 4877 h of size distributions for the clustering. With the run parameters listed in Table 2 31% of the set were sorted into three clusters of similar shapes of size distribution. Figure 13 presents average size distributions at both sites for these three clusters. In these cluster averages the Oden data extend the distributions to diameters between five and 11 nm. Clusters 1 and two have bimodal shapes albeit with the Aitken mode diameter of Cluster 1 being about 10% smaller than that of Cluster 2. Cluster 3 with
highest number concentration has only one mode in the Alert size range with its peak between 40 and 50 nm.

The geographic distribution of back trajectories for the three clusters in experiment “Oden-Alert” was collected in Fig. 14. Only geocells that are hit by back trajectories from both sites are marked. Additionally, two subpopulations of trajectories were formed. For Fig. 14a only data without any particles less than 10 nm measured on Oden were utilized. No joint geocells occurred for Cluster 2 in this subpopulation. The joint geocells for Clusters 1 and 3 cover most of the central Arctic with branches into the open water areas of the Eurasian Arctic sectors from the Fram Strait to the Laptev Sea. In Fig. 14b only cases with N10 > 1 cm$^{-3}$ were collected. For Cluster 1, into which typical bimodal marine size distributions were sorted the geographic distribution of potential source areas did not change much in Fig. 14b. Figure 14 indicates, however, that even in this type of air new particle formation was recorded on Oden. Cluster 3 with the strongest cases of new particle formation were focused onto the central Arctic when N10 on Oden was greater than 1 cm$^{-3}$. Also, joint cells of Cluster 2 with its main mode around 30 nm appeared over the ice covered area between Greenland and the North Pole in Fig. 14b.

7 Synopsis and conclusions

The present paper continues the analysis of the aerosol data from the four summer cruises of the Swedish icebreaker Oden in 1991, 1996, 2001, and 2008 with a focus on the potential source regions and related aerosol formation processes as illustrated in Fig. 1. While the four cruises provided a wealth of new observations there appears to be an inconsistency when comparing direct observations of a local particle flux from an open lead$^1$ (Held et al., 2011a) suggesting the pack ice area to be a net sink

$^1$The high Arctic open leads can be described as ever-changing open water channels comprising 10–30% of the ice pack ice area, ranging from a few meters up to a few kilometers in width.
of aerosols, to statistical interpretations of aerosol concentrations (Heintzenberg and Leck, 2012), which suggests the innermost Arctic to be a source of sub-micrometer particles. Further support of the latter findings relates to the fact that near-surface airborne aerosol, as well as low-level cloud and fog droplets, contained the same type of polymer gel\(^2\) material as found in the open-lead surface microlayer (Gao et al., 2012; Leck et al., 2013; Orellana et al., 2011; Bigg et al., 2004; Leck and Bigg, 2005).

When comparing the course of open water under the trajectories in this study for the two aerosol types, i.e. Cluster 1, 2, 3, and 6 in the experiment “Open water” (Fig. 8) with Clusters 2, 3, and 8 of the experiment “Pack ice high” (cf. Fig. 8b) significant differences between newly formed and aged aerosol over the pack ice become clear which lends further support to the findings of particle sources over the inner Arctic.

In both subpopulations the air had spent ten days over pack ice with less than 50% open water while traveling over ever more contiguous ice. Trajectories connected with high concentrations of newly formed small particles, however, experienced more open water during the last four days before arrival in heavy ice conditions at Oden. Thus we hypothesize that both, long travel times over the more contiguous ice, combined with more open water conditions during the last days before air mass arrival were an essential factor controlling the simultaneous occurrence of high number concentrations at several discrete particle sizes in the <10 and 20–50 nm size ranges over the Arctic pack ice. An hypothesis fitting with this chain of events could be fragmentation and/or dispersion of primary marine polymer gels, 200–500 nm diameter in size, into the nanogel size fractions down to a few nanometer polymers (Karl et al., 2013; Leck and Bigg, 2010). Fragmentation was suggested previously to be favored by evaporation of cloud or haze drops and promoted by long travel times over the pack ice (e.g.,

\(^2\)Marine gels or polymer gels are produced by phytoplankton and biological secretions of sea ice algae at the sea–air interface. The polymer gels are made up of water-insoluble, heat resistant, highly surface-active and highly hydrated (99% water) polysaccharide molecules spontaneously forming 3-dimensional networks inter-bridged with divalent ions (Ca\(^{2+}\)/Mg\(^{2+}\)), to which other organic compounds, such as proteins and lipids, are readily bound.
Heintzenberg et al., 2006). The fragmentation hypotheses appears to be consistent with the findings of a polymer gel source at the air–sea interface (Leck et al., 2013; Orellana et al., 2011; Bigg et al., 2004; Leck and Bigg, 2005; Gao et al., 2012) and may also explain why only a few percent of the observed total particle number variability at the ship was explained by the direct measurements of particle number fluxes (Held et al., 2011a). Based on past and present results we conclude the inner most Arctic to be a source of sub-micrometer particles.

Even though the Alert data had been taken in later years they still confirm the findings from the other sites with respect to particle sources over the central Arctic (cf. Figs. 13 and 14). Also, our comparison with Spitsbergen data clearly identified similarities in the structure of the size distributions and, again, pointed towards particle sources in the inner Arctic (Figs. 11 and 12). There appears to be an inconsistency when comparing observations of small particle formation over the inner Arctic where conventional nucleation paradigms (Karl et al., 2012) fail to explain this phenomenon and those south of the pack ice. Previously reported results from Alert in spring, (Leaitch et al., 2013), and on Mt. Zeppelin, Spitsbergen in early summer, (Engvall et al., 2008), showed nucleation events followed by subsequent prototypical “banana growth” (e.g., c.f. Kulmala et al., 2001), which the authors explained by solar radiation in concert with the presences of precursor gases and attendant low condensational sinks. Possible reasons for the inconsistency with the data collected during the four icebreaker expeditions could be that the DMS source and photochemical sink generating the precursor gases for nucleation and early growth is both seasonal and temperature dependent (Leck and Persson, 1996a, b; Kerminen and Leck, 2001; Karl et al., 2007, 2012). Given that, perhaps the main difference between the studies concerns how efficiently nucleation and growth of particles resulting from DMS oxidation are predicted by the choice of model and lack of observations to constrain the model assumptions. The Spitsbergen data on Mt. Zeppelin now cover more than 12 years. In a follow-up study we will pursue the search for potential source regions of the Arctic aerosol with this larger data set.
With a clustering the open water information along the trajectories a clear separation of marine vs. pack ice aerosol was achieved. Then the total data set was divided into two subpopulations above and below the 50% value of average open water during the course of the trajectories. The two constrained data sets were investigated further for potential source regions of pack ice and marine aerosols by clustering their horizontal trajectory components. In the marine aerosol this clustering yielded two main source regions over Laptev and Kara Seas, the aerosol showing bimodal features (cf. Fig. 7a). Beyond that two special cases emerged in the marine aerosol: One, covering polluted North Atlantic air that had passed over Svalbard (cf. Fig. 7b), and two, free tropospheric air that had crossed Greenland before arriving at Oden (cf. Fig. 7c).

The subpopulations below the 50% value of average open water during the course of the trajectories indicated two different aerosol types in addition to the case of small particle formation discussed above: Bimodal marine aerosol from the marginal ice zone and open seas around the pack ice (cf. Figs. 5a and 9c) and an aged aerosol that also occurred frequently over the pack ice (Figs. 5c, 8a, and 9). For the former case this may involve both direct emissions of larger polymer gel accumulation mode particles, as well as growth of smaller particles via two processes, namely heterogeneous condensation and aerosol cloud processing in which the bimodal particle size distribution characteristic of cloud-processed air is created (Hoppel et al., 1994). Previous studies in the same area and season (Heintzenberg et al., 2006; Chang et al., 2011; Heintzenberg and Leck, 2012; Kupiszewski et al., 2013; Hellén et al., 2012; Nilsson and Leck, 2002; Leck et al., 2013; Leck and Persson, 1996b; Leck and Bigg, 2005) have shown raised concentrations of accumulation mode particles within the high Arctic boundary layer which the authors attribute to sources upwind Oden: transport of precursor gases and marine biogenic particles from the MIZ or locally from open leads over the pack ice. Previous reported result of individual particles by Bigg and Leck (2001, 2008), Leck et al. (2002), and Leck and Bigg (2005a, b, 2010) collected over the pack ice however have failed to find evidence of sea salt particles of less than 200 nm in diameter. Larger
super-micrometer contained a varied and appreciable organic component shown to be polymer gels but also significant amounts of sodium chloride (Leck et al., 2002, 2013). The frequent occurrence of the aged aerosol (Figs. 8a and 9) belonged to the sub-population in which the air had spent ten days over pack ice with less than 50% open water while traveling over ever more contiguous ice (cf. Fig. 15) but had experienced less open water during the last four days before arrival at Oden relative to the sub-population newly formed particles (cf. Fig. 8b). The noted relative losses of the accumulation mode can be explained by an efficient scavenging processes associated with low clouds and fog near the MIZ and during the first days of advection over the pack ice (Nilsson and Leck, 2002; Heintzenberg and Leck, 2012). The loss in the sub-Aitken mode particle sizes would have resulted from coagulation processes most efficient and thus most realistic when involving clod/fog droplets (Karl et al., 2012).

Acknowledgements. We are most grateful to NSIDC for their providing Arctic sea ice data. In particular we are indebted to Sara, Lisa, and Terry from the NSIDC user service, who helped J. Heintzenberg with endless patience to understand the formalities of NSIDC’s ice data. Robert Leaitch very kindly processed and provided the Alert aerosol data for this study.

References


Table 1. Start and end date and time (UTC) of the hourly *Oden* aerosol data utilized in the present paper in 1991, 1996, 2001, and 2008, and the number of hourly averages after screening for possible pollution from the ship.

<table>
<thead>
<tr>
<th>Year</th>
<th>Start date and time</th>
<th>End date and time</th>
<th>Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>18 Aug 1991 16:00</td>
<td>26 Sep 1991 23:00</td>
<td>768</td>
</tr>
<tr>
<td>1996</td>
<td>24 Jul 1996 19:00</td>
<td>04 Sep 1996 23:00</td>
<td>581</td>
</tr>
<tr>
<td>2001</td>
<td>10 Jul 2001 00:00</td>
<td>25 Aug 2001 23:00</td>
<td>676</td>
</tr>
<tr>
<td>2008</td>
<td>04 Aug 2008 06:00</td>
<td>07 Sep 2008 17:00</td>
<td>620</td>
</tr>
</tbody>
</table>
Table 2. Run parameters of the cluster experiments. Constraint = constraints on data input to clustering algorithm. OW = percentage of open water along the trajectories. PSD = particle size distribution. $N_{init,t}$ = initial number of hours required in each cluster. $X_{av}$ = average distance of the cluster members from a cluster average of normalized coordinates (cf. Eq. 1). $P (P \leq 1)$ = outlier reduction factor to be applied to each cluster (cf. Sect. 3). $C_{fin}$ = number of clusters after eliminating clusters with smallest average distance from any other cluster (cf. Sect. 3).

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Constraint</th>
<th>$X$</th>
<th>$Y$</th>
<th>OW</th>
<th>PSD</th>
<th>$N_{init}$</th>
<th>$X_{av}$</th>
<th>$P$</th>
<th>$C_{fin}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMS</td>
<td>None</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>PSD</td>
<td>48</td>
<td>0.2</td>
<td>0.75</td>
<td>4</td>
</tr>
<tr>
<td>All aerosol</td>
<td>None</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>PSD</td>
<td>72</td>
<td>0.2</td>
<td>0.75</td>
<td>4</td>
</tr>
<tr>
<td>Open water</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>PSD</td>
<td>24</td>
<td>0.2</td>
<td>0.9</td>
<td>8</td>
</tr>
<tr>
<td>Marginal ice</td>
<td>OG5 and OS5 &gt; 50 %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>PSD</td>
<td>24</td>
<td>0.2</td>
<td>0.9</td>
<td>4</td>
</tr>
<tr>
<td>Pack ice low</td>
<td>Data from cluster 1 of “Open water”</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>PSD</td>
<td>48</td>
<td>0.7</td>
<td>0.5</td>
<td>3</td>
</tr>
</tbody>
</table>

Oden-Zeppelin

Concurrent time periods in 2001, 2008

| Oden-Alert        |              | x   |     |    | PSD | 48         | 0.05     | 0.90| 3        |
Table 3. Key data of the clusters of the cluster experiments. LC = Longitudinal cluster (cf. Sect. 3). \(X_i\) = Width of geographic coverage (vis. Eq. 4); \(P_{\text{unique}}\) = Parameter of the uniqueness of geographic coverage (%; vis. Section 3); ZAVT = Average height of trajectories during the last five days before arrival at Oden; OS5 = cluster-median open water (%) under the back trajectories during the last five days before arrival at Oden; DMS = cluster-median DMS concentration (nmol m\(^{-3}\)); NTO = cluster-median total particle number concentration (cm\(^{-3}\)); N10 = cluster-median particle number concentration below 10 nm diameter (cm\(^{-3}\)); N26 = cluster-median particle number concentration below 26 nm diameter (cm\(^{-3}\)); P24 = Median sum of precipitation along the last 24 h along the trajectories; P48 = Median sum of precipitation along the last 48 h along the trajectories; P5D = Median sum of precipitation along the last five days along the trajectories; n.a. = non applicable; n.d. = no data. Aerosol and gas values for experiment Oden-Zeppelin hold for Oden data only. ¹ Number of cells hit jointly by trajectories to Oden and to Mt. Zeppelin. ² At Oden trajectories with 500 m arrival heights were employed.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Cluster</th>
<th>(X_i)</th>
<th>(P_{\text{unique}})</th>
<th>OS5</th>
<th>ZAVT</th>
<th>DMS</th>
<th>NTO</th>
<th>N10</th>
<th>N26</th>
<th>P24</th>
<th>P48</th>
<th>P5D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6</td>
<td>14</td>
<td>65</td>
<td>50</td>
<td>50</td>
<td>0.5</td>
<td>n.d.</td>
<td>n.d.</td>
<td>n.d.</td>
<td>n.d.</td>
<td>n.d.</td>
<td>n.d.</td>
</tr>
<tr>
<td>All aerosol</td>
<td>1</td>
<td>16</td>
<td>93</td>
<td>65</td>
<td>200</td>
<td>3.1</td>
<td>110</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>19</td>
<td>66</td>
<td>26</td>
<td>444</td>
<td>1.2</td>
<td>37</td>
<td>0.5</td>
<td>6</td>
<td>0.3</td>
<td>0.4</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>15</td>
<td>35</td>
<td>7</td>
<td>480</td>
<td>0.6</td>
<td>61</td>
<td>1.3</td>
<td>8</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>21</td>
<td>44</td>
<td>16</td>
<td>400</td>
<td>0.4</td>
<td>81</td>
<td>0.1</td>
<td>30</td>
<td>0.6</td>
<td>1.2</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>21</td>
<td>18</td>
<td>88</td>
<td>110</td>
<td>2.1</td>
<td>170</td>
<td>0</td>
<td>23</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Open water</td>
<td>1</td>
<td>21</td>
<td>23</td>
<td>10</td>
<td>380</td>
<td>0.6</td>
<td>100</td>
<td>0.4</td>
<td>11</td>
<td>0</td>
<td>0.1</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>24</td>
<td>4</td>
<td>25</td>
<td>670</td>
<td>0.6</td>
<td>180</td>
<td>13</td>
<td>61</td>
<td>0</td>
<td>3.9</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>33</td>
<td>0</td>
<td>29</td>
<td>300</td>
<td>1.2</td>
<td>150</td>
<td>8</td>
<td>63</td>
<td>0</td>
<td>4.3</td>
<td>4.3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>31</td>
<td>2</td>
<td>96</td>
<td>460</td>
<td>0.4</td>
<td>160</td>
<td>0.6</td>
<td>13</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>69</td>
<td>13</td>
<td>93</td>
<td>540</td>
<td>0.9</td>
<td>210</td>
<td>1</td>
<td>18</td>
<td>0</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>20</td>
<td>6</td>
<td>25</td>
<td>390</td>
<td>0.8</td>
<td>130</td>
<td>60</td>
<td>64</td>
<td>0.9</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>21</td>
<td>18</td>
<td>88</td>
<td>110</td>
<td>2.1</td>
<td>170</td>
<td>0</td>
<td>23</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Marginal ice</td>
<td>1</td>
<td>18</td>
<td>100</td>
<td>57</td>
<td>490</td>
<td>0.7</td>
<td>50</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0.2</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>15</td>
<td>97</td>
<td>65</td>
<td>240</td>
<td>2.8</td>
<td>120</td>
<td>0</td>
<td>18</td>
<td>0</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>24</td>
<td>95</td>
<td>64</td>
<td>1600</td>
<td>0.4</td>
<td>130</td>
<td>0</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>21</td>
<td>92</td>
<td>86</td>
<td>550</td>
<td>1.5</td>
<td>80</td>
<td>2</td>
<td>19</td>
<td>0</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Pack ice low</td>
<td>1</td>
<td>8</td>
<td>44</td>
<td>6</td>
<td>510</td>
<td>0.7</td>
<td>50</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>8</td>
<td>62</td>
<td>10</td>
<td>240</td>
<td>0.9</td>
<td>60</td>
<td>0</td>
<td>10</td>
<td>2.3</td>
<td>3.6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>10</td>
<td>60</td>
<td>12</td>
<td>190</td>
<td>0.4</td>
<td>40</td>
<td>0</td>
<td>5</td>
<td>1.3</td>
<td>1.9</td>
<td>6.6</td>
</tr>
<tr>
<td>Oden-Zeppelin</td>
<td>1</td>
<td>104</td>
<td>56</td>
<td>n.a.</td>
<td>890</td>
<td>0.6</td>
<td>10</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>38</td>
<td>0 ¹</td>
<td>n.a.</td>
<td>900</td>
<td>0.8</td>
<td>60</td>
<td>2</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>51</td>
<td>0 ¹</td>
<td>n.a.</td>
<td>740</td>
<td>2.3</td>
<td>50</td>
<td>0</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>61</td>
<td>58</td>
<td>n.a.</td>
<td>1080</td>
<td>0.5</td>
<td>130</td>
<td>1.4</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Oden-Alert</td>
<td>1</td>
<td>190</td>
<td>162</td>
<td>23</td>
<td>710</td>
<td>0.9</td>
<td>30</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>70</td>
<td>0 ¹</td>
<td>29</td>
<td>640</td>
<td>1.2</td>
<td>110</td>
<td>5</td>
<td>70</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>170</td>
<td>41</td>
<td>25</td>
<td>910</td>
<td>0.5</td>
<td>90</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
**Figure 1.** Schematic view of the sources and transport mechanisms of aerosol particles over the summer Arctic pack ice, adopted from Kupiszewski et al. (2013).
Figure 2. Map of the working area of the present study: White: Cruise tracks during the four Oden expeditions in 1991, 1996, 2001, and 2008. Red symbol: North Pole. Dark grey geocells: area covered with at least 100 trajectory hits per geocells by 5 day back trajectories in all four cruises. Additional geocells in light grey are covered likewise by 10 day back trajectories. Colored lines: ten percent limit of sea ice cover north of 76°N estimated from average sea concentrations (https://nsidc.org/data) during each of the four Oden cruises.
Figure 3. Gridded average Arctic sea ice cover in % during the four *Oden* cruises in 1991, 1996, 2001, and 2008. Only cells with at least 100 ice pixels per cell are plotted. The “blind spot” of satellite data north of 86° N is assumed to have 100 % ice cover.
Figure 4. Regional distribution of five clusters of back trajectories with hourly DMS values of all Oden cruises. Cluster 2 = red; cluster 3 = yellow; cluster 7 = mocha; cluster 8 = cyan; longitudinal cluster “LC” = grey. The color saturation indicates the number of trajectory hits per geocells in percent. Only geocells with at least 25 trajectory hits are shown. The symbol indicates the North Pole.
Figure 5. Left: average particle number size distributions of the five clusters of horizontal trajectory coordinates, combined with particle size distributions. The clusters are separated into (a) marine, (b) pack ice high, and (c) pack ice low. Cluster 1 = blue; cluster 3 = yellow; cluster 4 = green; cluster 5 = magenta; cluster “LC” = copper. Error bars give one standard deviation about the cluster-average. Right: corresponding regional distributions of median total number concentrations, (NTO, cm$^{-3}$). The color saturation indicates the total number associated with the respective trajectory. Only geocells with at least 25 trajectory hits are shown. The symbol indicates the North Pole.
Figure 6. (a) Geographic distribution of trajectories of the subpopulations “Marginal ice”, and (b) geographic distribution of trajectories of all data in the subpopulation “Pack ice”.
Figure 7. Left: average particle number size distributions of the four clusters of horizontal trajectory coordinates, combined with particle size distributions in the subpopulation “Marginal ice”. The clusters are separated into (a) marine, (b) Spitsbergen and (c) Greenland. Error bars give one standard deviation about the group average. Blue = cluster 1; yellow = cluster 3; green = cluster 4. Right: corresponding regional distributions of median total number concentrations, (NTO, cm$^{-3}$). The color saturation indicates the total number associated with the respective trajectory. Only geocells with at least 25 trajectory hits are shown. The symbol indicates the North Pole.
Figure 8. Left: average particle number size distributions of the four clusters of horizontal trajectory coordinates, combined with particle size distributions in the subpopulation “Open water”. The clusters are separated into (a) pack ice low, (b) pack ice high. Error bars give one standard deviation about the group average. Blue = cluster 1; yellow = cluster 3. Right: corresponding regional distributions of median total number concentrations, (NTO, cm⁻³). The color saturation indicates the total number associated with the respective trajectory. Only geocells with at least 25 trajectory hits are shown. The symbol indicates the North Pole.
Figure 9. Left: average particle number size distributions of the four clusters of horizontal trajectory coordinates, combined with particle size distributions in the subpopulation “Pack ice low”. Error bars give one standard deviation about the group average. Blue = cluster 1; red = cluster 2; yellow = cluster 3. Right: corresponding regional distributions of median total number concentrations, (NTO, cm$^{-3}$). The color saturation indicates the total number associated with the respective trajectory. Only geocells with at least 25 trajectory hits are shown. The symbol indicates the North Pole.
Figure 10. 25, 50, and 75 % percentiles of trajectory-connected number size distributions taken during the Oden cruises in 2001 and 2008 on the icebreaker Oden and on Mt. Zeppelin, Spitsbergen.
Figure 11. Average number size distributions in three clusters of harmonized size distribution data taken on Oden and on Mt. Zeppelin during the Oden cruises in the summers of 2001 and 2008. There are no Mt. Zeppelin data below 20 nm diameter. For Cluster 1 standard deviations about the average Oden data are shown.
Figure 12. Geographic distribution of back trajectories for the three clusters in Fig. 11 with joint occurrences of at least 25 trajectory hits per geocells. Left: for Cluster 1 only cases without particles less than 10 nm measured on Oden were considered. Right: for Cluster 1 only cases with particle concentrations less than 10 nm > 1 cm$^{-3}$ measured on Oden were considered. The cluster coloring corresponds to that in Fig. 11. Colored lines: ten percent limits of sea ice cover north of 76° N estimated from average sea concentrations (https://nsidc.org/data) during the Oden cruises of 2001 and 2008.
Figure 13. Average number size distributions in three clusters of harmonized particle size distributions measured on Oden during all cruises and at Alert, Nunavut during the Augusts of 2011, 2012, and 2013. There are no Alert data below 11 nm diameter.
Figure 14. Geographic distribution of back trajectories for the three clusters in Fig. 13 with joint occurrences of at least 25 trajectory hits per geocells. Left: only cases without particles less than 10 nm measured on *Oden* were considered, (no joint geocells for cluster 2). Right: only cases with particle concentrations less than 10 nm > 1 cm⁻³ measured on *Oden* were considered. The colors correspond to those in Fig. 13.
Figure 15. Median open water percentages along the trajectories of clusters 1–3 in experiment “Pack ice low” (cf. Fig. 9) and those of clusters 1, 2, 3, and 6 in experiment “Open water” (cf. Fig. 8).