A statistical analysis of North East Atlantic (submicron) aerosol size distributions

M. Dall’Osto¹,*, C. Monahan¹, R. Greaney¹, D. C. S. Beddows², R. M. Harrison², D. Ceburnis², and C. D. O’Dowd²

¹School of Physics and Centre for Climate and Air Pollution Studies, Ryan Institute, National University of Ireland Galway, University Road, Galway, Ireland
²National Centre for Atmospheric Science, School of Geography, Earth & Environmental Sciences, Division of Environmental Health & Risk Management, University of Birmingham Edgbaston, Birmingham, B15 2TT, UK
*now at: Institute for Environmental Assessment and Water Research (IDÉA-CSIC), Barcelona, Spain

Received: 30 June 2011 – Accepted: 12 July 2011 – Published: 2 August 2011
Correspondence to: M. Dall’Osto (manuel.dallosto@gmail.com)
Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

The Global Atmospheric Watch research station at Mace Head (Ireland) offers the possibility to sample some of the cleanest air masses being imported into Europe as well as some of the most polluted being exported out of Europe. We present a statistical Cluster analysis of the physical characteristics of aerosol size distributions in air ranging from the cleanest to the most polluted for the year 2008. Data coverage achieved was 75% throughout the year. By applying the Hartigan-Wong $k$-Means method, 12 Clusters were identified as systematically occurring and these 12 Clusters could be further combined into 4 categories with similar characteristics, namely: coastal nucleation category (occurring 21.3% of the time), open ocean nucleation category (occurring 32.6% of the time), background clean marine category (occurring 26.1% of the time) and anthropogenic category (occurring 20% of the time) aerosol size distributions. The coastal nucleation category is characterised by a clear and dominant nucleation mode at sizes less that 10 nm while the open ocean nucleation category is characterised by a dominant Aitken mode between 15 nm and 50 nm. The background clean marine characteristic is a clear bimodality in the size distribution, although it should be noted that either the Aitken mode or the Accumulation mode may dominate the number concentration. By contrast, the continentally-influenced size distributions are generally more mono-modal, albeit with traces of bi-modality. The open ocean category occurs more often during May, June and July, corresponding with the N. E. Atlantic high biological period. Combined with the relatively high percentage frequency of occurrence (32.6%), this suggests that the marine biota is an important source of new aerosol particles in N. E. Atlantic Air.

1 Introduction

The parameters of the atmospheric aerosols are poorly characterized in global climate models. Particularly uncertain are the ones influencing the radiative balance and...
properties of clouds, such as the number size distribution, chemical composition and particle mixing state. The lack of a proper representation concerning size distributions of the aerosol in global and regional models is a major reason why direct and the indirect climate effects (Twomey, 1974) of the atmospheric aerosols constitute the largest uncertainty in our present understanding of the anthropogenic climate forcing (IPCC, 2001). Atmospheric aerosol particles span over several orders of magnitude in diameter ($D_p$), from a few nanometer to hundreds of micrometer. Small particles, in particular nucleation mode (typically with $D_p < 10$ nm) and Aitken mode (10 nm $< D_p < 100$ nm) particles contribute little to total particulate mass in background air; however, they contribute significantly to surface area and dominate particle number concentration. The surface area contributes to the overall aerosol condensation sink and optical properties (influencing the direct aerosol radiative effect) while number concentration influences cloud droplet concentrations (contributing to the indirect aerosol radiative effect). The shape of the aerosol size distribution is influenced by air mass origin which determines difference aerosol sources and evolution processes (Dall’Osto et al., 2010). Studies such as that of Dall’Osto et al. have focused on characterization via air mass origin over short periods of weeks to months rather than quantifying the frequency of occurrence of each size distribution type over time scales of the order of a year (Charron et al., 2007; Costabile et al., 2009).

Similarly, a number of studies have already focused on size distributions of particles detected at Mace Head, mainly focusing on marine aerosol constituents. Typical marine aerosol size distributions sampled at the coastal Mace Head research station showed low particle number concentrations with a sub-micron marine size distribution characterised by a bi-modal shape, with an accumulation mode centred at 200 nm and a fine mode at 40 nm, indicative of a cloud-residual accumulation mode produced by the in-cloud growth of activated fine mode particles (Hoppel et al., 1994). A study on the seasonality of the clean marine particle size distributions was reported by Yoon et al. (2007), showing the aerosol size distribution modal diameters for different seasons: 0.031 µm in winter and 0.049 µm in summer for the Aitken mode and 0.103 µm in...
winter and 0.177 µm in summer for the accumulation mode, respectively. By contrast, air masses affected by anthropogenic pollution sampled at Mace Head during anticyclonic periods and conditions of continental outflow Aitken and accumulation mode were enhanced by a factor of 5 compared to the marine sector (O’Dowd et al., 2001; Coe et al., 2006).

As part of the EUCAARI (European Aerosol Cloud Climate and Air Quality Interactions) Integrated Project, one of the major activities was to conduct and analyse size distribution measurements at a range of European supersites (Kulmala et al., 2009). The EUCAARI intensive measurement programme was throughout the full year of 2008. As one of the 12 atmospheric supersites EUCAARI supersites, Mace Head (Ireland) is uniquely located on the interface between the NE Atlantic and Europe, thus enabling sampling of both the cleanest air entering into Europe along with some of the most polluted air being exported out of Europe into the N Atlantic (Jennings et al., 2003; O’Connor et al., 2008).

Asmi et al. (2010) presented a combined statistical analysis of all 12 EUCAARI supersites for the year 2008 in terms of means and extremes of the number concentration size distributions encountered.

In this paper, aerosol size distributions are analyzed by using k-means Cluster analysis (Beddows et al., 2009). All particle size distributions are considered, and not filtered by any other criteria (i.e. wind direction, particle number or black carbon concentrations). Different states of the aerosol were determined by using a novel application of Cluster analysis, which uses the degree of similarity and difference between individual observation to define the groups and to assign group memberships. One advantage of this clustering method over providing an average of aerosol size distributions is that it does provide a specific number of size distributions which can be compared across different time periods (Beddows et al., 2009). Accordingly, the final Cluster centres reflect particle number size distributions representative of each cluster. Some examples of particle size distributions Cluster analysis for substantial SMPS datasets can be found in the literature. Similar approaches have previously been used: for example, Tunved
et al. (2004) showed an analysis resulting in eight different Clusters capturing different stages of the aerosol lifecycle. Additionally, Charron et al. (2007) presented an examination of the source signature or origin signature represented by particle number size distribution and modal diameters measured at a rural receptor size in southern England while Beddows et al. (2009) was able to reduce the complexity of the different rural, urban, and curbside atmospheric particle size data according to the temporal and spatial trends of the clusters. Whilst a number of intensive field studies have been focusing on average monthly datasets, clustering analysis on year long particle size distributions measurements are scarce (Engler et al., 2007; Kivekäs et al., 2009; Venzac et al., 2009).

2 Methodology

2.1 Location

The Mace Head Atmospheric Research Station is located in Connemara, County Galway on the Atlantic Ocean coastline of Ireland at 53°19′36″ N, 9°54′14″ W and offers a clean oceanic sector from 190° through west to 300°. Meteorological records show that on average, over 60% of the air masses arrive at the station in the clean sector (Jennings et al., 2003). Air is sampled at 10 m height from a main community sampling duct positioned at 80–120 m from coast line depending on tide. More information about the station can be found in O’Connor et al. (2008). Clean air is generally characterized by black carbon mass concentrations of 50 ng m⁻³ or less.

2.2 Instrumentation

The on-line aerosol analysers sampled from a 10 m height 10 cm diameter laminar flow community sampling duct with a 50% size cut at 3.5 µm at wind speed of about 10 m s⁻¹ (Dall’Osto et al., 2010). Total particle concentrations at sizes larger than 3
and 10 nm diameter were measured using a Thermo Systems Inc. (TSI) Condensation Particle Counter (CPC) 3025 and 3010, respectively. Size distributions were measured using a TSI nano-Scanning Mobility Particle Sizer (SMPS) between 3 and 20 nm, scanning every 30 s, and a standard SMPS operating 10-min size distribution scans between 20 and 500 nm (Wang and Flagan, 1990).

Total particle concentrations were measured using a Condensation Particle Counter (CPC) bank consisting of a standard Thermo Systems Inc. (TSI) CPC 3025A with a 50% particle diameter size cut-off at 3 nm and a TSI CPC 3010 with a 50% size cut at 10 nm. Data acquisition time resolution for the CPC bank was set to 1 Hz. The CPC3025 was diluted by a factor of 16.1 to avoid saturation of the CPC total number concentration in excess of 100 000 cm$^{-3}$ (Yoon et al., 2005). It is possible to detect number concentration of ultrafine particles up to $1.7 \times 10^6$ cm$^{-3}$ with the aid of this calibrated dilution system. Throughout this paper $N_d > 3$ nm and $N_d > 10$ nm will indicate TSI 3025 and TSI 3010 particle number concentrations, respectively.

Basic meteorological parameters such as wind speed (WS), wind direction (WD), relative humidity (RH), temperature ($T$), atmospheric pressure, precipitation, global radiation, UV-radiation are measured at the 10 m height level, with some of the measurements duplicated at 22 m. Aerosol scattering coefficient measurements were performed by a TSI Inc. three-wavelength integrating nephelometer (Bodhaine et al., 1991; Heintzenberg and Charlson, 1996). Aerosol absorption (and Black Carbon mass) was measured using both a McGee Scientific Aethalometer AE-16 and a Multi-Angle Absorption Photometer (MAAP). Furthermore, a TEOM instrument (PM$_{2.5}$) was also deployed.

### 2.3 Clustering method

The available SPMS data (five minutes resolution) were averaged in hourly bins for the analysis. For the year 2008, the data coverage was 88% (1 January and 18 November) but during some of the hours within this period the measurements were not available. The overall coverage for the year 2008 was 75%. The 6578 SMPS size distributions...
obtained at one hour resolution were then subsequently normalised by their vector-length and Cluster analysed (Beddows et al., 2009). The use of Cluster analysis was justified in this work using a Cluster Tendency test, which calculated a Hopkins Index of 0.20. The choice of k-means clustering was made from a selection of the partitional Cluster packages (Beddows et al., 2009). The Dunn-Index for the results of the K-means analysis for different Cluster numbers showed a clear maximum for 12 clusters, some of which belonged only to specific times of the day, specific mechanisms as well as specific seasons.

3 Results

3.1 Overview of meteorological conditions of year 2008

The main meteorological data for the year 2008 are presented in Fig. 1 as monthly averages (December is not included in the analysis as SMPS measurements were not available for this month). The month of July and August were found to have the highest temperature (15 ± 1°C, both), with the lowest temperatures occurring January to March. However, when considering atmospheric pressure and WS, the month of August showed lower values when compared with other summer months, somehow creating a discontinueum between the months of July and September. WS was found to be the highest during the winter months, with average values of 9.7 ± 8 and 9.6 ± 8 m s\(^{-1}\) over January and February, respectively. Monthly averaged WD were mainly associated with the clean marine sector (180–300°), whereas the month of May was most associated with Easterly wind (140 ± 75°).

3.2 Air mass back trajectories

Back trajectories of the air masses arriving at Mace Head were calculated for 00:00, 06:00, 12:00 and 18:00 UTC for each day of year 2008, depicting the path taken by
the air mass reaching the sampling site over the previous five days. The back-trajectories were calculated using the on-line HYSPLIT model developed by the National Oceanic and Atmospheric Administration (NOAA) (Draxler et al., 1997). The meteorological air mass classification was used following the method described in Dall’Osto et al. (2010), where five synoptic period types were defined: cP (Continental European), cmP (Continental-Marine), mP (Polar Marine), mT (Tropical Marine) and mA (Arctic Marine). On average, the predominant air mass type was found to be mP (42 ± 15 %), followed by mA (22 ± 10 %), mT (19 ± 8 %), cP (12 ± 11 %) and cmP (5 ± 8 %) when considering all year 2008. Figure 2 shows for each month of 2008 the percentage occurrence of each of the five different air masses. When considering different seasons, spring was associated with cmP, cP and mA air masses relative to the other types. By contrast, mP air masses dominated during winter months. The month of May was found to be the most affected by anthropogenically influenced air masses (cmP and cP together contributing up to 52 %), whilst the month of August and October were found the least affected by such air mass types.

3.3 SMPS clustering

K-means analysis of the SMPS aerosol size distributions gave 12 Clusters whose frequency varied between 2.8 % and 15.2 % (Fig. 3), none of which dominated the overall population. Table 1 summarises the main features of the 12 Clusters and their particle size distributions are presented in Fig. 4a–d. Further information related to each Cluster is provided in Figs. 4–7 and Tables 2–4. In this section average meteorological data (WS, WD, RH, Temp.) are also presented for each SMPS cluster. Moreover, particular emphasis is given in describing unique physical and chemical properties as well specific temporal profiles peculiar of a given cluster.

– Cluster 1 (8.9 %): Figure 4a shows size distribution of Cluster 1, revealing a strong nucleation mode in the smallest detectable SMPS size bin at about 4-5 nm, as well as an Aitken mode (about 60 nm) and accumulation mode (at about
200 nm). WS and WD were $7.0 \pm 3 \text{ m s}^{-1}$ and $236 \pm 98^\circ$, respectively (Fig. 5a and b); temperature of $13.3 \pm 3^\circ \text{C}$ and RH of $75 \pm 12\%$, respectively (Fig. 5c and d); with the highest average temperature of all clusters. This Cluster presented the second highest average particle number concentration ($N_D > 3 \ 23352 \text{ cm}^{-3}$ and $N_D > 10 \ 1905 \text{ cm}^{-3}$) and also the highest difference between the two CPCs ($N_D > 10 - N_D > 3$, Table 3).

- Cluster 2 (3.8\%): Figure 4a presents a similar distribution to Cluster 1, but the nucleation mode is shifted towards larger diameters (6–8 nm) whilst both the aitken and accumulation modes are found similar to Cluster 1. WS and WD were $5.7 \pm 3 \text{ m s}^{-1}$ and $219 \pm 96^\circ$, respectively; the temperature was $11.4 \pm 4^\circ \text{C}$ and the RH was $75 \pm 13\%$. $N_D > 3$ revealed the highest particle number concentration (45 574 cm$^{-3}$) as well as the highest difference $N_D > 3 - N_D > 10$ (42 675 cm$^{-3}$).

- Cluster 3 (8.6\%): Figure 4a shows the third clear nucleation size distribution, with a larger nucleation mode at about 9-11 nm relative to Cluster 1 and Cluster 2. WS and WD were $5.7 \pm 3 \text{ m s}^{-1}$ and $193 \pm 96^\circ$ (respectively), the temperature was $11.5 \pm 3^\circ \text{C}$ and the RH was $78 \pm 13\%$ (respectively) representing similar pattern of Cluster 2. This Cluster presented similar high particle number concentrations to Cluster 1 and Cluster 2 and therefore it was merged in the subgroup of particle size distributions attributed to coastal nucleation events (Fig. 4a).

- Cluster 4 (12.6\%): This Cluster (Fig. 4b) showed a nucleation mode at about 20 nm, significantly larger than the previous 3 Clusters attributed to coastal nucleation events. An accumulation mode at about 200 nm is also seen. WS and WD were $6.38 \pm 3 \text{ m s}^{-1}$ and $205 \pm 86^\circ$, the temperature was $11.4 \pm 3^\circ \text{C}$ and the RH was $81 \pm 12\%$. Furthermore, the particle number concentration was found to be high ($N_D > 3 \ 13733 \text{ cm}^{-3}$ and $N_D > 10 \ 2249 \text{ cm}^{-3}$) but the difference is not as high as the ones of Clusters 1 to 3.
Cluster 5 (11.0 %): Figure 4b shows a broad mode at about 30 nm for this cluster, as well as a smaller accumulation mode at about 200 nm. WS and WD were $6.6 \pm 3 \text{ m s}^{-1}$ and $216 \pm 86^\circ$ (respectively), the temperature was $11.6 \pm 3^\circ C$ and the RH was $83 \pm 12\%$ (respectively). This Cluster was found to be similar to Cluster 4 and therefore merged into the same subgroup shown in Fig. 4b. Merged with Cluster 4, it showed the lowest average PM$_{2.5}$ concentration ($7.8 \pm 3 \mu g m^{-3}$, see Table 3d).

Cluster 6 (12.5 %): This Cluster (Fig. 4c) presents a broad aitken mode at about 40 nm and an accumulation mode shifted towards bigger diameter at about 250 nm. WS and WD were $6.0 \pm 3 \text{ m s}^{-1}$ and $184 \pm 105^\circ$, the temperature was $9.3 \pm 3^\circ C$ and the RH was $84 \pm 12\%$. Given the broad particle size distribution spanning across the Aitken and the accumulation mode, this Cluster is different from the ones previously described and therefore it is presented in Fig. 4c.

Cluster 7 (15.2 %): Figure 4d shows much smaller particle concentration relative to previous clusters, with both the Aitken and accumulation mode shifted towards smaller diameter (30 nm and 150 nm, respectively). WS and WD were $9.0 \pm 4 \text{ m s}^{-1}$ and $240 \pm 63^\circ$, temperature of $9.1 \pm 3^\circ C$ and RH of $84 \pm 10\%$, with the lowest average temperature among all clusters. It was associated with the most westerly wind detected (Fig. 5b, $241 \pm 30^\circ$) and second strongest WS among all Clusters (Fig. 5a, $9 \pm 4 \text{ m s}^{-1}$). The BC concentration was found the third lowest among all Clusters (Table 3e). This Cluster was also peculiar as presented the lowest particle number concentration ($N_D > 3 \ 892 \pm 90 \text{ cm}^{-3}$ and $N_D > 10 \ 282 \pm 40 \text{ cm}^{-3}$) among all 12 classifications.

Cluster 8 (9.0 %): This Cluster (Fig. 4b) was found similar to Cluster 5, with modes slightly bigger at about 50 nm and 200 nm (aitken and accumulation, respectively). WS and WD were $6.7 \pm 3 \text{ m s}^{-1}$ and $202 \pm 90^\circ$, temperature of $11.5 \pm 3^\circ C$ and RH of $86 \pm 9\%$, also reflecting similar averages of Cluster 4 and 5.
Cluster 9 (4.9 %): Figure 4c shows a broad aitken mode for this Cluster type at about 60 nm, and also a broad mode at higher diameter for the accumulation mode (250 nm). WS and WD were $6.3 \pm 3$ m s$^{-1}$ and $186 \pm 90^\circ$, the temperature was $11.6 \pm 3^\circ$ and the RH was $87 \pm 11\%$. This Cluster presented the second highest BC concentration ($207 \pm 218$ ng m$^{-3}$, Table 3e).

Cluster 10 (2.8 %): This Cluster shows the highest loading of accumulation mode particle concentration, with two modes peaking at about 100 nm and 300 nm, respectively (Fig. 4c). WS and WD were $5.6 \pm 3$ m s$^{-1}$ and $155 \pm 82^\circ$, respectively whilst the average temperature was $10.4 \pm 3^\circ$C and the average RH was $86 \pm 11\%$. Furthermore, it presented the highest BC concentration (almost two-fold increase over all Cluster of $573 \pm 540$ ng m$^{-3}$ – Table 3e), and also highest PM$_{2.5}$ concentration ($13.3 \pm 8$ µg m$^{-3}$ – Table 3d).

Cluster 11 (3.1 %): Figure 4d shows that this Cluster exhibited very low particle average particle number size distributions, with two clear modes at about 60 nm and 200 nm, respectively, showing the typical clean marine bimodal size distribution. WS and WD were $8.4 \pm 3$ m s$^{-1}$ and $229 \pm 53^\circ$ (respectively), whereas the average temperature was found to be $10.9 \pm 3^\circ$C and the average RH 86 ± 10%. The $N_D > 3$ and $N_D > 10$ average concentrations were found the second lowest of all Clusters ($764 \pm 80$ cm$^{-3}$ and $321 \pm 40$ cm$^{-3}$, respectively). Whilst the CPC was the second lowest, the PM$_{2.5}$ concentration ($11.5 \pm 9$ µg m$^{-3}$) was found the third highest and likely to be due to coarse sea salt particle contribution (Table 3).

Cluster 12 (7.8 %): This Cluster (Fig. 4d) shows a broad aitken mode merged with the accumulation mode at about 100 nm. A second shoulder at about 350 nm is also seen, not seen in any of the other 12 clusters. The average WS and WD were $11.4 \pm 3$ m s$^{-1}$ and $237 \pm 53^\circ$, the temperature was $9.0 \pm 3^\circ$C and the RH was $83 \pm 10\%$. $N_D > 3$ and $N_D > 10$ average concentrations were found the third lowest of all Clusters ($773 \pm 80$ cm$^{-3}$ and $379 \pm 40$ cm$^{-3}$, respectively), whilst PM$_{2.5}$ concentration ($11.5 \pm 9$ µg m$^{-3}$) was the second higher, reflecting the same feature of...
Cluster 11 with CPC and PM values being anti correlated. In other words, bimodal size distributions associated with the lowest particle number concentrations were found to have the highest values of particle PM$_{2.5}$ mass.

4 Discussions

Based on the similarity of the SMPS Clusters reported in the previous section, we created four categories (each one composed of three unique SMPS clusters) representing the 12 $k$-means SMPS clusters. The similarity among each group can be seen in Fig. 4: “coastal nucleation” in Fig. 4a (Cluster 1, 2 and 3), “open ocean nucleation” in Fig. 4b (Cluster 4, 5 and 8), “Background clean marine” in Fig. 4c (Cluster 7, 11 and 12) and “anthropogenic” in Fig. 4d (Cluster 6, 9 and 10). It is important to note that the different aerosol size distributions Clusters were merged in these four groups not only upon their similar size distributions among each other (see Fig. 4a–d) but also by considering strong correlations with other physical and chemical parameters obtained with other instruments and discussed individually for each group in the next sections. Furthermore, meteorological and air mass trajectories data analysis were also considered and those are described below for both individual SMPS Clusters and group of SMPS clusters. The average aerosol size distributions of the four aerosol categories (obtained by averaging the SMPS Clusters of each individual category) are presented in Fig. 4e. Clear differences can be seen: a dominant nucleation mode at sizes less that 10 nm for the coastal nucleation category, a dominant Aitken mode between 15 nm and 50 nm for the open ocean nucleation category, a clear bimodality in the size distribution for the background clean marine and a generally more mono-modal one for the continentally-influenced size distributions. In the following sections each individual SMPS Cluster belonging to the four aerosol categories is discussed. Whilst the general physical and chemical properties of the SMPS Clusters belonging to each of the individual category (Fig. 4e) are broadly similar, their temporal trends and their diurnal profiles are at times different (due to meteorological data as well as marine biota...
availability) and therefore each individual Cluster and not only the averaged information of each aerosol category is presented. Specifically, Sect. 4.1 gives particular emphasis on unique features of each of the four aerosol size distribution groups, whereas in Sect. 4.2 only the seasonality trends of different aerosol size distributions sampled at Mace Head among the year 2008 are discussed.

4.1 Classification and description of groups

4.1.1 Coastal nucleation (Coast. N. type)

Cluster 1, 2 and 3 present together a strong aerosol nucleation mode below 10 nm in size (Fig. 4a). Whilst there is a progressive increase in the nucleation mode size among the 3 clusters, the aitken and accumulation mode sizes do not change much. These three represented 21.3 % of the total size distributions classified, implying about one fifth of the time Mace Head is under the influence of coastal nucleation events. Cluster 1 and 2 showed the highest CPC particle number concentrations and also the highest difference $N_D > 3 - N_D > 10$ (see Table 3a–c). Cluster 3 showed the lowest CPC concentrations among the 3 coastal nucleation Clusters ($8543 \text{ cm}^{-3}$ and $2840 \text{ cm}^{-3}$ for $N_D > 3$ and $N_D > 10$, respectively) and the largest nucleation mode at about 9 nm, suggesting this mode is more related to the growth of already nucleated particles. Cluster 1 and Cluster 2 showed the lowest average RH ($75.1 \pm 10 \%$ and $74.9 \pm 10 \%$, respectively) of the 12 Clusters as well as the highest average temperature, suggesting solar radiation was strongly involved in the process (Fig. 5). A very interesting feature of these 3 Clusters relative to all the others classified is their strong diurnal profile variations (Fig. 6). It is important to stress that these 3 Clusters were the only ones showing a clear diurnal trend spiking during the day (Fig. 6). As expected, there is a progressive shift in the maximum peak of the diurnal profile of Cluster 1, 2 and 3 as the hour of the day increases. This can be explained as time during the day passes there is a shift in the size of the nucleation mode from Cluster 1 to Cluster 3 (4–5 nm, 6–8 nm and 9–11 nm for Cluster 1, 2, 3, respectively). Cluster 1 is most likely associated with
the initial stage of a nucleation event occurring at low tide while Cluster 2 and 3 are representing later stages of the nucleation events when particles have grown to larger sizes during mid-day hours before advecting to Mace Head. Our study is in line with the ones of O’Dowd et al. (2002a, b), which reported that “Type I” showed the highest concentrations at the size below 5 nm (Cluster 1, this study), whereas in the “Type II” cases, the maximum concentration peak lies at 6–8 nm (Cluster 2, this study). In “Type III” case the particle growth arising from below 10 nm to sizes above 20 nm was observed (Cluster 3, this study).

4.1.2 Open ocean nucleation and growth (Op. Oc. Nucl. type)

The importance of particle formation events in marine air has not been as clearly established to date. While nucleation appears to be a frequent phenomenon in many coastal areas (Yoon et al., 2006; O’Dowd and de Leeuw, 2007; Modini et al., 2009), in open ocean environments new particle formation events have only occasionally been observed (Ehn et al., 2010). Dall’Osto et al. (2010) and O’Dowd et al. (2010) reported open ocean nucleation events detected at the Mace Head station, without reporting indications on time occurrence of such unique events relative to the total size distributions sampled. In our results, Cluster 4, 5 and 8 represented 32.6 % of the classified SMPS spectra, suggesting that about a third of the time the Mace Head atmospheric station is under the influence of open ocean nucleation events. Moreover, these Clusters were found to be associated with the lowest BC concentrations as well as lowest PM\textsubscript{2.5} concentrations (see Table 3d–e, respectively). These Clusters did not present a strong correlation with solar radiation (see Fig. 5f). By contrast, about 40 % of the time Cluster 4, 5 and 9 occurred during night time. It should be noted that our classification can be compromised by the possibility that part of the Cluster 4 can be due to a sub-Cluster of coastal nucleation events (such as type III described in O’Dowd et al., 2002a). However, because of the difference in the diurnal profile (see Fig. 6) we can exclude a strong influence of coastal nucleation events to Cluster 4.
4.1.3 Anthropogenic influence (Anth. type)

Clusters 6, 9 and 10 (Figure 4c) show typical particle size distributions affected by anthropogenic pollutants, confirmed by the fact that the BC average concentrations associated with these three Clusters were the highest among the 12 classified Clusters (Table 3e). Furthermore, Clusters shown in Fig. 4c presented average wind direction linked with inland advections (Fig. 5b). Specifically, Cluster 10 was associated with the highest BC loading (573 ± 60 ng m\(^{-3}\)), whereas Cluster 9 and Cluster 6 were the second and third highest (207 ± 20 ng m\(^{-3}\) and 178 ± 19 ng m\(^{-3}\), respectively – see Table 3e). Cluster 10 presented also the highest PM\(_{2.5}\) concentrations (16.5 µg m\(^{-3}\), Table 3d). Among the three clusters, Cluster 6 was associated with lower temperatures (9.4°C, Fig. 5d), and its seasonality pattern is described in the following section.

Table 1 shows that Cluster 9 and Cluster 10 have a strong correlation with cP and cmP air masses, with Cluster 10 showing a 40% concurrence with cP air mass type. Cluster 10 was also associated with the strongest scattering signal among all Clusters (Table 3f). Nephelometer scattering data for all the 3 wavelengths shows a contrasting trend seen for the Cluster 10 relative to all others, showing higher values for shorter wavelengths and, therefore, suggesting dominant contribution of submicron particles (Dall’Osto et al., 2010).

4.1.4 Background clean marine

Cluster 7, 11 and 12 represented 26.1% of the total SMPS spectra and were characterised by the lowest particle number concentrations among all clusters. These three Clusters were also characterised by similar wind direction and wind speed properties, with the mostly westerly and the strongest average wind speed among all Clusters (Fig. 5a and b). Among this group, Cluster 7 and Cluster 12 were found to be similar, associated with cold temperatures and low average atmospheric pressure, whilst Cluster 11 showed opposite trends and occurring mainly during summer months (Fig. 5e). Clusters associated in this background clean Marine type group were mostly
detected with mP air masses; with Cluster 12 associated up to 59% of the times with this air mass type (Table 1). Our study shows similar results to the ones reported by Yoon et al. (2007), and further discussions on the seasonality of SMPS Clusters associated with this type are provided in the next section. Aerosols described by these SMPS Clusters were also characterised by high scattering efficiency (Table 3f), reflected again in the high PM$_{2.5}$ mass loadings (Table 3e). For example, Cluster 11 and 12 were the second and third (respectively) Cluster associated with the highest scattering signal (Table 3f), regardless of the fact that those two Clusters were associated with very low particle number concentrations (Table 3a and b). It is worth noting that Clusters 7 and 11 were similar to the clean marine air particle spectra reported by O’Dowd et al. (2004) representing low and high biological activity over the Northeast Atlantic. However, the Cluster 12 is a unique Cluster exhibiting large accumulation mode size (with an accumulation mode of average diameter of 372 nm, Fig. 4d), likely originating via sea spray at high wind speed. Recently, Sparklen (2007) used a statistical synthesis of marine aerosol measurements from experiments in four different oceans, predicting a bimodal size distribution that agrees well with observations as a grand average over all regions, but large regional differences were found. Notably, observed Aitken mode number concentrations were more than a factor of 10 higher than in the model for the N Atlantic indicating the importance of organic matter in this region. Beside different size distributions associated with different seasons, our study suggests that not only one bimodal size distribution but two distinct types of aerosol size distributions are present during winter months.

4.2 Overall seasonality of different particle size distributions clusters

Figure 7 shows the seasonality of each SMPS cluster, represented by the occurrence of each Cluster during each month considered in this study. The 12 SMPS particle size distribution Clusters showed very different seasonality due to multiple reasons, including different meteorology and different biological ocean activity throughout the year as well as different anthropogenic influences over time. Whilst the seasonality
could be described for the 4 groups presented in Sect. 4.1, some Clusters belonging to the same group showed a different seasonality and therefore each Cluster is presented and discussed individually.

Within coastal nucleation Clusters (1, 2, 3), Cluster 1 and Cluster 2 peaked during spring and autumn (Fig. 7). Cluster 1 seems to peak more during late summer times (July–September), whereas Cluster 2 shows its highest frequency during spring months (March–May). Cluster 3 did not present a clear seasonality, somehow in between Cluster 1 and Cluster 3. The concentration of precursor gases is related to the width of tidal zone exposed to the atmosphere, and in turn, tidal height. In addition to these two factors, the origin of air masses arriving at the measurement location is also particularly important, especially at Mace Head (O’Dowd et al., 2002a) since this also determines which tidal areas the air is inter-acting with. The seasonal variation of the number of event days and event duration show a clear seasonal cycle, with the more frequent occurrence in spring and autumn, and the rarest in the winter season. Our study confirms earlier studies where it was found that the number of event days for summer is relatively lower than for spring/autumn, mainly due to the amount of precursor gases emitted from marine algae during low tide (O’Dowd et al., 2002a, b).

Cluster 4 (open ocean nucleation events) by contrast showed a very different trend, peaking mainly during the months of May and June and supporting the pioneering study of O’Dowd et al. (2010). We conclude that not only the diurnal profiles but also the seasonality patterns are different between the coastal and open ocean nucleation events. The other two Clusters belonging to open ocean nucleation events (Cluster 5 and Cluster 8) peaked during summer. The higher occurrence of Cluster 4 during May and June is likely to be related with the higher probability of having an open ocean nucleation events during the high biogenic activity months, whereas the occurrence of Cluster 5 and 8 (with larger Aitken modes, Fig. 4b) during warmer summer months is likely to be due to photochemically enhanced secondary aerosol production.

Regarding the Clusters affected by anthropogenic aerosols (higher value of BC), Cluster 9 showed a strong seasonal trend spiking during summer months (Fig. 11).
remaining Clusters belonging to anthropogenic size distributions did not show clearly discernable seasonal trend.

Finally, when considering background clean marine size distributions, Cluster 7 and 12 tended to occur more frequently during winter and characterising clean marine air masses advecting to Mace Head research station during the colder months. The lower contribution of clean marine events to the total number of events for autumn months is most likely due to the more frequent occurrence of high pressure weather systems during these months, thus bringing more continental air to Mace Head; hence the number of clean marine events are reduced (Yoon et al., 2006).

The same data presented in Fig. 7 are now represented as normalised probability of the 12 Clusters for each month in Table 4. In other words, we now look at each month of year 2008 and not at each Cluster to investigate which SMPS Clusters dominate each month. By looking at Table 4, a clean pattern can be seen for some of the months. For example, Cluster 7 and 12 were dominant during winter months (January, February, March, October and November). Cluster 8 and 9 were instead typical in June, July and September. Perhaps the most interesting temporal feature of the 12 SMPS Clusters lays in the sub-group of the background marine size distributions (Fig. 4d). Whilst Cluster 7 and Cluster 12 occurs mainly during winter months, Cluster 11 tends to be more predominant during warmer months. Yoon et al. (2007) reported that the aerosol light scattering coefficient showed a minimum value of 5.5 Mm$^{-1}$ in August and a maximum of 21 Mm$^{-1}$ in February. This seasonal variation was due to the higher contribution of sea salt in the MBL during North Atlantic winter. By contrast, aerosols during late spring and summer exhibited larger angstrom parameters than winter, indicating a large contribution of the biogenically driven fine or accumulation modes. Our study shows similar trends, with clean marine aerosols dominating winter months, and secondary marine aerosols as well as anthropogenic ones dominating the summer months. Yoon et al. (2007) furthermore reported the clean marine aerosols with the large organic fraction appear to be physically larger than the aerosols without (the former occurring mainly during summer times during high oceanic biological activity,
the latter during winter). Laboratory studies on bubble-mediated aerosol production (Sellegri et al., 2006) revealed that the aerosol distribution resulting from the use of surfactant-free seawater comprised three modes: (1) a dominant accumulation mode at 110 nm; (2) an Aitken mode at 45 nm; and (3) a third mode, at 300 nm, resulting from forced bursting of bubbles. The forced bursting occurs when bubbles fail to burst upon reaching the surface and are later shattered by splashing associated with breaking waves and/or wind pressure at the surface. However, the more complex submicrometer spectral structure that is significantly affected by salinity, temperature and surfactant concentration suggested that more detailed studies are required for laboratory generated sea-spray aerosol (O’Dowd and DeLeew 2007).

Our results shows that open ocean nucleation events (Cluster 4, 5 and 8) and background clean marine aerosol size distributions – characterized by an Aitken mode at 50 nm and an accumulation mode at 200 nm (Cluster 11) - dominated warmer months. By contrast, clean marine conditions during winter months present two different size distributions: a bimodal distribution with a mode at 30 nm and a mode at 200 nm (Cluster 7) and a broader bimodal distribution peaking at 100 nm and 350 nm (Cluster 12). The results display clear seasonal differences in the spectral modal properties suggesting differences in the NE Atlantic marine aerosol formation processes. Moreover – within winter months – two different Clusters dominate (Cluster number 7 and 12) with the former presenting a more numerous Aitken mode and the latter a more numerous accumulation mode (Fig. 4d). Both primary and secondary formation processes are likely to contribute to the observed physical and chemical size distributions but it is not clear as to what extent each process contributes at this stage, especially in the Aitken mode.

5 Conclusions

We analyzed SMPS data collected at Mace Head (Ireland) during the EUCAARI year (2008, 75 % data coverage). By applying k-means clustering, 12 groups of aerosol size
distributions were found which could be further grouped in four classifications: coastal nucleation events (21.3%), open ocean nucleation events (32.6%), background clean marine (26.1%) and anthropogenic (20%) aerosol size distributions. Kerminen et al. (2010) recently discussed the highlights of the EUCAARI project and future directions on atmospheric nucleation. Our results show that altogether coastal and open ocean nucleation events apportion for about 54% of the size distributions analysed during the year 2008, stressing the importance of secondary aerosol processes. During winter periods, when nucleation events are marginal, about 60% of the particle size distributions show a bimodal shape due to clean marine air masses with low particle number and BC concentrations. However, different types of background clean marine aerosol size distributions are detected among clean marine air masses, pointing out to both primary and secondary marine aerosol processes not yet fully understood.

**Acknowledgements.** This work was funded by the European Commission Framework Programme 6 EUCAARI Integrated Project, Ireland Higher Education Authority Programme for Research in Third Level Institutes – Cycle 4, Science Foundation Ireland, and the Irish Environmental Protection Agency. NOAA Hysplit is also acknowledged for providing air mass back trajectories.

**References**


A statistical analysis of aerosol size distributions

M. Dall’Osto et al.

11, 21677–21711, 2011


21699
Table 1. *K*-means Cluster analysis results and main features of the 12 Clusters obtained. The not defined (N.D.) Aitken modes are due to the predominance of the nucleation and/or accumulation modes in some of the aerosol size distributions clusters.

<table>
<thead>
<tr>
<th>Cluster number</th>
<th>%</th>
<th>Size type</th>
<th>Nucleation mode</th>
<th>Aitken mode</th>
<th>Accumulation mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8.9</td>
<td>Coastal nucleation</td>
<td>4.5</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>2</td>
<td>3.8</td>
<td>Coastal nucleation</td>
<td>6.5</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>3</td>
<td>8.6</td>
<td>Coastal nucleation</td>
<td>10.2</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>4</td>
<td>12.6</td>
<td>Open ocean nucleation</td>
<td>15.0</td>
<td>23</td>
<td>–</td>
</tr>
<tr>
<td>5</td>
<td>11.0</td>
<td>Open ocean nucleation</td>
<td>–</td>
<td>28</td>
<td>120</td>
</tr>
<tr>
<td>6</td>
<td>12.5</td>
<td>Continental polluted</td>
<td>–</td>
<td>42</td>
<td>–</td>
</tr>
<tr>
<td>7</td>
<td>15.2</td>
<td>Clean marine</td>
<td>–</td>
<td>30</td>
<td>165</td>
</tr>
<tr>
<td>8</td>
<td>9.0</td>
<td>Open ocean nucleation</td>
<td>–</td>
<td>40</td>
<td>165</td>
</tr>
<tr>
<td>9</td>
<td>4.9</td>
<td>Continental polluted</td>
<td>–</td>
<td>62</td>
<td>230</td>
</tr>
<tr>
<td>10</td>
<td>2.8</td>
<td>Continental polluted</td>
<td>–</td>
<td>112</td>
<td>–</td>
</tr>
<tr>
<td>11</td>
<td>3.1</td>
<td>Clean marine</td>
<td>–</td>
<td>52</td>
<td>212</td>
</tr>
<tr>
<td>12</td>
<td>7.8</td>
<td>Clean marine</td>
<td>–</td>
<td>112</td>
<td>372</td>
</tr>
</tbody>
</table>
Table 2. Occurrence (%) of the five different air mass back trajectories type described in Dall’Osto et al. (2010) for each of the 12 SMPS clusters. Air mass back trajectories were calculated every six hours (0 a.m., 6 a.m., 12 p.m., 6 p.m.) for the whole year of 2008.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Air mass back trajectories type</th>
<th>mP</th>
<th>mT</th>
<th>mA</th>
<th>cmP</th>
<th>cP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>30</td>
<td>6</td>
<td>46</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>19</td>
<td>9</td>
<td>48</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>42</td>
<td>7</td>
<td>42</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>35</td>
<td>28</td>
<td>32</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>45</td>
<td>13</td>
<td>33</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>42</td>
<td>7</td>
<td>35</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>54</td>
<td>16</td>
<td>25</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>40</td>
<td>20</td>
<td>26</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>33</td>
<td>18</td>
<td>18</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>15</td>
<td>13</td>
<td>15</td>
<td>17</td>
<td>40</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>41</td>
<td>39</td>
<td>16</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>59</td>
<td>24</td>
<td>6</td>
<td>0</td>
<td>11</td>
</tr>
</tbody>
</table>
**Table 3.** Average CPC particle number concentration for each of the 12 SMPS clusters: (a) \( N_D > 3 \) nm (CPC TSI 3025A) \((\text{cm}^{-3})\), (b) \( N_D > 10 \) nm (TSI CPC 3010) \((\text{cm}^{-3})\), (c) \( N_D > 3 \) nm–\( N_D > 10 \) nm \((\text{cm}^{-3})\), (d) average PM\(_{2.5}\) concentrations \((\mu \text{g m}^{-3})\), (e) average BC concentrations (MAAP) \((\text{ng m}^{-3})\) and (f) average scattering properties (nephelometer) for 3 wavelength (450, 550 and 700 nm) \((*1E^{-6} \text{ m}^{-1})\) for each SMPS cluster.

<table>
<thead>
<tr>
<th>Cluster number</th>
<th>( N_D &gt; 3 ) nm</th>
<th>( N_D &gt; 10 ) nm</th>
<th>( N_D &gt; 3 \text{ nm–} N_D &gt; 10 ) nm</th>
<th>TEOM</th>
<th>MAAP</th>
<th>Neph. (450 nm)</th>
<th>Neph. (550 nm)</th>
<th>Neph. (700 nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>23 352</td>
<td>1905</td>
<td>21 447</td>
<td>10.8</td>
<td>126</td>
<td>15</td>
<td>13</td>
<td>29</td>
</tr>
<tr>
<td>2</td>
<td>45 574</td>
<td>2899</td>
<td>42 676</td>
<td>11</td>
<td>166</td>
<td>18</td>
<td>15</td>
<td>25</td>
</tr>
<tr>
<td>3</td>
<td>85 43</td>
<td>2840</td>
<td>57 03</td>
<td>10.2</td>
<td>141</td>
<td>15</td>
<td>13</td>
<td>31</td>
</tr>
<tr>
<td>4</td>
<td>13 733</td>
<td>2249</td>
<td>11 484</td>
<td>8.1</td>
<td>95</td>
<td>14</td>
<td>13</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>36 50</td>
<td>1061</td>
<td>25 89</td>
<td>7.8</td>
<td>89</td>
<td>16</td>
<td>15</td>
<td>31</td>
</tr>
<tr>
<td>6</td>
<td>29 03</td>
<td>1321</td>
<td>15 81</td>
<td>8.8</td>
<td>178</td>
<td>18</td>
<td>15</td>
<td>28</td>
</tr>
<tr>
<td>7</td>
<td>89 2</td>
<td>282</td>
<td>611</td>
<td>9.5</td>
<td>72</td>
<td>25</td>
<td>24</td>
<td>29</td>
</tr>
<tr>
<td>8</td>
<td>20 46</td>
<td>1097</td>
<td>949</td>
<td>8.4</td>
<td>142</td>
<td>18</td>
<td>16</td>
<td>33</td>
</tr>
<tr>
<td>9</td>
<td>23 64</td>
<td>1346</td>
<td>10 18</td>
<td>8.5</td>
<td>207</td>
<td>25</td>
<td>20</td>
<td>33</td>
</tr>
<tr>
<td>10</td>
<td>32 96</td>
<td>2074</td>
<td>12 23</td>
<td>16.5</td>
<td>573</td>
<td>74</td>
<td>57</td>
<td>49</td>
</tr>
<tr>
<td>11</td>
<td>76 4</td>
<td>321</td>
<td>443</td>
<td>11.5</td>
<td>89</td>
<td>31</td>
<td>29</td>
<td>35</td>
</tr>
<tr>
<td>12</td>
<td>77 3</td>
<td>379</td>
<td>394</td>
<td>13.3</td>
<td>157</td>
<td>41</td>
<td>38</td>
<td>38</td>
</tr>
</tbody>
</table>
Table 4. Monthly occurrence (% among each month) for all 12 clusters.

<table>
<thead>
<tr>
<th>Month of the year</th>
<th>Cluster number</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>January</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>February</td>
<td>0.3</td>
<td>3.1</td>
</tr>
<tr>
<td>March</td>
<td>3.2</td>
<td>5.0</td>
</tr>
<tr>
<td>April</td>
<td>2.8</td>
<td>4.7</td>
</tr>
<tr>
<td>May</td>
<td>4.3</td>
<td>4.8</td>
</tr>
<tr>
<td>June</td>
<td>3.6</td>
<td>2.2</td>
</tr>
<tr>
<td>July</td>
<td>9.0</td>
<td>2.8</td>
</tr>
<tr>
<td>August</td>
<td>5.3</td>
<td>2.4</td>
</tr>
<tr>
<td>September</td>
<td>7.2</td>
<td>3.2</td>
</tr>
<tr>
<td>October</td>
<td>2.0</td>
<td>0.6</td>
</tr>
<tr>
<td>November</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>
Fig. 1. Overview of meteorological parameter for each month of year 2008: (a) RH (%), (b) Temperature (°C), (c) WD (°), (d) WS (m s\(^{-1}\)) and (e) atmospheric pressure (mb).

Fig. 2. Occurrence (%) of the five different air mass back trajectories type described in Dall’Osto et al (2010) for each of the 12 months during the 2008 EUCARRI year.
Fig. 2. Occurrence (%) of the five different air mass back trajectories type described in Dall’Osto et al. (2010) for each of the 12 months during the 2008 EUCARRI year.
Fig. 3. Occurrence (%) of the 12 Clusters obtained by k-means analysis.
Fig. 4. (a) Cluster 1, 2 and 3 (coastal nucleation) (b) Cluster 4, 5 and 8 (open ocean nucleation) (c) Cluster 6, 9 and 10 (anthropogenic) and (d) Cluster 7, 11 and 12 (background clean marine). Please note difference in log scale for (a) $10^5$ (b) and (c) $10^4$ and (d) $10^3$. (d) shows the average size distribution for each of the four groups.
Fig. 5. Average values for each Cluster for (a) Wind Speed (m s\(^{-1}\)); (b) Wind Direction (°); (c) Relative Humidity (%); (d) Temperature (°C); (e) Atmospheric Pressure (mb) and (f) solar radiation (J m\(^{-2}\), right scale) and percentage of hourly average size distribution detected during night time (%). Clusters are grouped among the four different groups: Coastal nucleation (Coast.), Open Ocean nucleation (Op. Oc.), Anthropogenic (Ant.) and Background marine (Back.), respectively.
Fig. 6. Diurnal variation (3 h resolution) of the 12 clusters.
Fig. 7. Monthly occurrence (% among all 12 clusters) of SMPS clustering analysis.