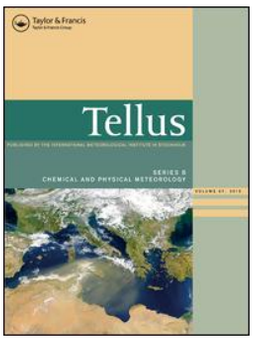

Tellus B: Chemical and Physical Meteorology

ISSN: (Print) 1600-0889 (Online) Journal homepage: http://www.tandfonline.com/loi/zelb20

'How to find bananas in the atmospheric aerosol': new approach for analyzing atmospheric nucleation and growth events

Jost Heintzenberg, Birgit Wehner & Wolfram Birmili

To cite this article: Jost Heintzenberg, Birgit Wehner & Wolfram Birmili (2007) 'How to find bananas in the atmospheric aerosol': new approach for analyzing atmospheric nucleation and growth events, Tellus B: Chemical and Physical Meteorology, 59:2, 273-282

To link to this article: https://doi.org/10.1111/j.1600-0889.2007.00249.x

Published online: 18 Jan 2017.

Tellus (2007), 59B, 273–282
Printed in Singapore. All rights reserved

TELLUS

'How to find bananas in the atmospheric aerosol': new approach for analyzing atmospheric nucleation and growth events

By JOST HEINTZENBERG*, BIRGIT WEHNER and WOLFRAM BIRMILI, *Leibniz-Institute for Tropospheric Research, Permoserstr. 15, 04318 Leipzig, Germany*

(Manuscript received 23 October 2006; in final form 8 January 2007)

ABSTRACT

We have devised a new search algorithm for secondary particle formation events, or 'nucleation events' in data sets of atmospheric particle size distributions. The search algorithm is simple and based on the investigation of 18 integral parameters of the particle size distribution, three of which were found to be most relevant for identifying nucleation events. The algorithm is tested using long-term size distribution data sets of high-size resolution observed at Melpitz, Hohenpeissenberg, and Leipzig, Germany, and Beijing, China, thereby covering a wide range of clean and polluted conditions. By specifying the particular training sets, the method can be used by other researchers with different data sets or different research goals. The same search approach could be applied to identify and analyze other systematic changes in size distribution such as during frontal passages or sand storms. As an example application of the new algorithm, the 50 strongest nucleation events ('bananas') at each of the four sites are analyzed statistically in terms of average changes of integral parameters of the particle size distribution.

1. Introduction

The formation of new atmospheric nuclei from supersaturated vapours is of essential relevance to the balance of particle number in the atmosphere. Historically, the first direct evidence for atmospheric new particle formation was obtained through air ion measurements. In an attempt at surveying the complete particle size distribution, Junge (1955) identified relative maxima in the number-size distribution of the continental aerosol below 20 nm diameter in the ion-spectra of Israel and Schulz (1932). At that time these maxima were ascribed to individual particle sources leading to 'line spectra' in the submicrometer aerosol. Using ion spectrometers with better temporal resolution Misaki (1964) and Abel et al. (1969a,b) identified distinct concentration maxima in the diameter range below 50 nm, which showed a trend towards larger particle sizes with time. In the late 1980s, the technical developments of differential mobility particle sizers and ultrafine condensation particle counters allowed quantitative measurements of particle size distributions down to a nuclei size of 3 nm (Winklmayr, 1987; Stolzenburg and McMurry, 1991; Winklmayr et al., 1991).

*Corresponding author.
e-mail: jost@tropos.de
DOI: 10.1111/j.1600-0889.2007.00249.x

Systematic long-term observations of particle size distributions in Finland and Germany led to the discovery and systematic study of events of particle nucleation and growth in the continental boundary layer (Mäkelä et al., 1997; Birmili, 1998). Since then, particle formation events have been identified at numerous rural continental and urban sites mostly during daytime (Kulmala et al., 2004b), and it has been recognized that these particle formation events are driven by the photochemical production of gaseous precursors (Boy and Kulmala, 2002; Birmili et al., 2003; McMurry et al., 2005). Observations of boundary-layer particle formation events usually comprise the following features: (1) rapid increase in particle number concentration (<20 nm) as a consequence of nucleation and subsequent particle growth into detectable sizes, (2) further particle growth within the size range of the instrumentation, as a consequence of continuing condensation of vapors and (3) gradual decrease in particle number concentration within several hours as a result of coagulation and mixing with other air masses. These occurrences of particle formation and subsequent growth have informally been dubbed 'banana events', due to their appearance in two-dimensional contour plots of particle number concentration as a function of time and particle diameter.

Detailed methods have been developed to analyze experimentally observed particle formation events with respect to the particle formation rates (Kulmala et al., 2001; Kerminen and Kulmala, 2002), and condensational growth rates (Birmili and

Wiedensohler, 2000; Dal Maso et al., 2005), and the condensational sink for condensable vapors (Mäkelä et al., 2000; Stanier et al., 2004). Owing to their thermodynamic properties, the binary sulfuric acid/water (Kulmala et al., 1995; Pirjola et al., 1998), and the ternary sulfuric acid/water/ammonia system (Kulmala et al., 2000) have been singled out as prime candidates to explain atmospheric particle formation events, although the participation of low- volatility oxidation products of biogenic precursors, such as monoterpenes (Kavouras et al., 1998) or other organic vapours (Berndt et al., 2005), as well as ion-assisted processes (Kulmala et al., 2004a; Yu and Turco, 2000) remains unclear. It has generally been recognized that the coagulation rate of newly formed particles onto the pre-existing aerosol particle population can determine whether a burst of particles can be observed >3 nm or not (Kerminen et al., 2001; McMurry et al., 2005).

An important approach to identifying the precursor gases, aerosol dynamical processes and degree of meteorological control involved in atmospheric particle formation is the statistical analysis of time-series, ideally spanning over several years. Often, a combination of atmospheric parameters, notably including radiation, precursor, condensational sink and atmospheric stability parameters may be found to describe the atmospheric conditions, which are favourable to particle formation (Clement et al., 2001; Birmili et al., 2003; Stanier et al., 2004).

Such statistical analyses require, in general, a classification of a given atmospheric data set into particle formation events, and non-event periods. Furthermore, particle formation events are often classified with additional criteria, such as particle formation intensity, the question whether additional particle growth is observable or not, the degree of interference with anthropogenic pollution or whether the observed aerosol dynamical features can be tracked in their temporal development in a continuous manner over many hours or not. Since these additional classification criteria are relevant to a different degree in different atmospheres, the data evaluation practices among different researchers have consequently been variable, and included various approaches to characterize particle formation events according to aerosol parameters derived from size distribution observations. The parameters derived include, among others, particle number concentration below a certain diameter, apparent particle formation rate, particle growth rate, duration of an event or the smoothness of the observed temporal evolution of an event (Birmili et al., 2003; Stanier et al., 2004; Dal Maso et al., 2005). A classification scheme may be limited, in that, it has been developed to categorize observations at a single or a few sites. Another limitation may be that the selection of events is performed manually or visually, and does not obey numerically defined or reproducible criteria. As a result, different researchers may arrive at different selections of particle events for a given data set, which makes comparisons between different analyses difficult.

Moreover, in the rapidly growing body of high-quality atmospheric aerosol size distribution, data manual and subjective data evaluation procedures turn out to be tedious, erroneous, and therefore less suitable for systematic statistical analyses. To date, no objective algorithm has been reported which would be able to classify a variety of given atmospheric data sets into particle formation events, and non-events on the basis of objective criteria. Therefore, in the present study, an objective methodology has been developed, which relies on integral parameters of the first three moments of the particle size distribution without inherent assumptions concerning the shape of the size distribution. The usefulness of the new methodology is demonstrated with exemplary results from size distribution observations at two rural and two urban settings.

2. Database

The methodology to identify new particle formation events in extended time-series of boundary layer particle size distribution data was developed with the aid of four existing data sets, three of which are in a state of ongoing collection. The selection of data sets include continuous, ground-based observations of particle number size distributions at the following sites: Hohenpeissenberg (rural, Southern Germany, Global Atmosphere Watch station of the German Meteorological Service DWD); Melpitz (rural, in the Easter German lowlands); Leipzig (urban background, Germany); Beijing (urban background, China). Each of these data sets covers a continuous measurement period of at least two years. Table 1 compiles the details of the respective data sets used, and related references.

In Fig. 1, exemplary particle formation events are shown for each of the four sites considered in the present study. The cases shown were selected with respect to (i) the nucleation mode concentrations being as high above the site-typical background levels as possible and (ii) the nucleation mode diameter increasing as continuously and smoothly as possible with time. Figure 1 also illustrates three of the moment-based size distribution parameters discussed below.

Importantly, all data sets were collected using identical, or equivalent instrumentation: Twin differential mobility particle sizers (TDMPS, Birmili et al., 1999) were used to record size distributions between 3 and 800 nm at controlled sheath air relative humidities below 10%. The high- size ($\approx$40 logarithmically spaced size channels) and time resolution (10–20 min) of the TDMPS allows to detect and track clearly new particle formation events and growth.

Originally, all data were stored in daily files. For two reasons, the individual day-files of size distributions were combined after quality control into contiguous time-series at each station. One, time windows at any time of the day could be searched for given types of events. Two, the concept of the search algorithm to be described below requires a normalization of size distributions in any given window with respective medians taken over that time window plus a given time period before, which may extend into adjacent days.

Table 1. Database of number-size distributions (3–800 nm) used in the present study

Site	ID	Type	Period	Time resolution (min)	Reference
Melpitz, Germany	MEL-1	Rural	25 March 1996 to 14 August 1997	15	(Heintzenberg et al., 1998; Birmili and Wiedensohler, 2000)
Melpitz, Germany	MEL-2	Rural	22 July 2004 to 30 April 2006	20	(Engler et al., 2006; Wehner et al., 2005)
Hohenpeissenberg, Germany	MOH	Rural	01 April 1998 to 04 September 2000	10	(Birmili et al., 2000, 2003)
IfT, Germany	IFT	Urban	13 February 1997 to 30 April 2006	10	(Wehner and Wiedensohler, 2003)
Beijing, China	BEI	Urban	03 March 2004 to 30 April 2006	10	(Wehner et al., 2004)

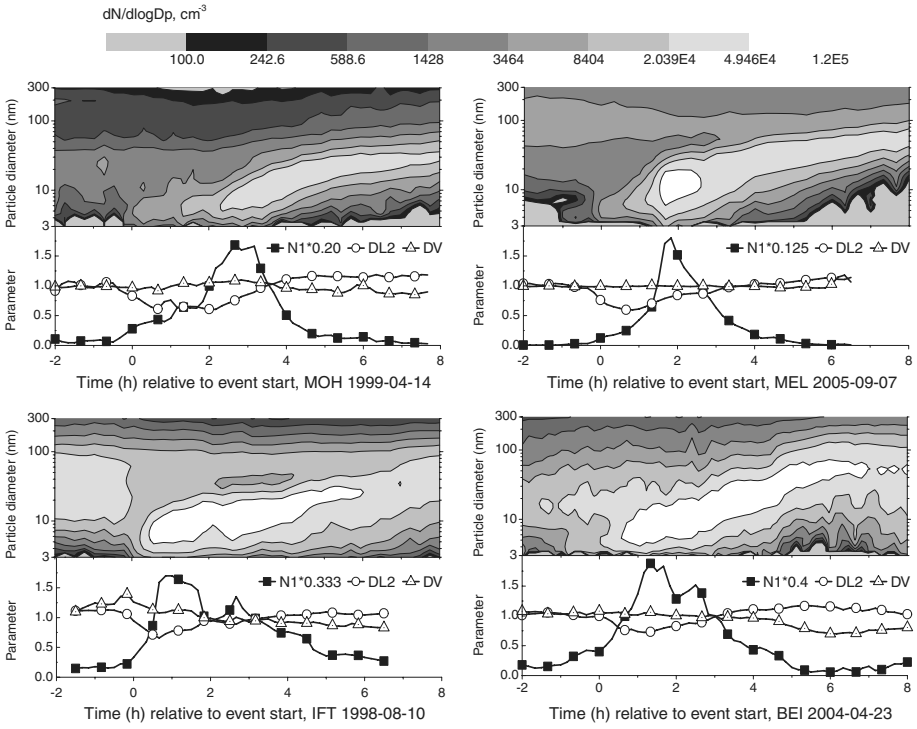

Fig. 1. Examples of particle formation in boundary layer air, each being representative for one of the four observational data sets at Hohenpeissenberg (MOH, top left), Melpitz (MEL, top right), IFT (bottom left) and Beijing (BEI, bottom, right) used in this work. Shown are the evolution of the particle size distribution (bananas), and the parameters 5, 13 and 18, (*cf*. Table 2), which are used for the identification of particle formation
events.

3. Sensitivity analysis and training

Concentration levels of particles and precursor gases, as well as the average meteorological conditions vary strongly in between the four sites. We used the Melpitz data as a first test bed because the new particle formation events in that atmosphere are disturbed the least by local anthropogenic sources as well as terrain effects: First, the Melpitz site is surrounded by flat country in all directions, so that there are no orographic effects interfering with the evolution of new particle formation events, such as at the Hohenpeissenberg mountain site. Second, Melpitz is in rural countryside; unlike IfT and Beijing, it is therefore not immediately influenced by urban aerosol sources, whose emissions would interfere with the evolution of a secondary new particle formation event.

The essential basis for the formation event selection are a number of integral parameters calculated from the particle number size distribution, which are explored for their sensitivity to the occurrence and evolution of nucleation and growth events. The first three moments of the size distribution were employed, together with the related characteristic diameters. By calculating these integral parameters not only for the total size range of the TDMPS but also in two subranges, a total of 18 parameters P_i, $i = 1, 18$ were derived, which are listed in Table 2.

Table 2. Integral parameters of the submicrometer particle size distribution explored in the sensitivity analysis. $D_{low} = 3$ nm, $D_{high} = 800$ nm, D_1 = upper diameter limit of subpopulation 1, D_2 = upper diameter limit of subpopulation 2. $D_{m,j}$ = Median diameter of moment j, with j = length, surface or volume

No.	ID	Parameter	Unit	Definition
1	NT	Total number	cm^{-3}	$\int_{\log D_{low}}^{\log D_{high}} \frac{dn(d\log D_p)}{d\log D_p} d\log D_p$
2	LT	Total length	nm cm^{-3}	$\int_{\log D_{low}}^{\log D_{high}} \frac{dn(d\log D_p)}{d\log D_p} D_p d\log D_p$
3	ST	Total surface	nm^2 cm^{-3}	$4\pi \int_{\log D_{low}}^{\log D_{high}} \frac{dn(d\log D_p)}{d\log D_p} \left(\frac{D_p}{2}\right)^2 d\log D_p$
4	VT	Total volume	nm^3 cm^{-3}	$\frac{4}{3}\pi \int_{\log D_{low}}^{\log D_{high}} \frac{dn(d\log D_p)}{d\log D_p} \left(\frac{D_p}{2}\right)^3 d\log D_p$
5	N1	Number $\leq D_1$	nm^{-3}	As in parameter 1 but for $D_{high} = D_1$
6	L1	Length $\leq D_1$	nm cm^{-3}	As in parameter 2 but for $D_{high} = D_1$
7	S1	Surface $\leq D_1$	μm^2 cm^{-3}	As 3 but for $D_{high} = D_1$
8	V1	Volume $\leq D_1$	μm^3 cm^{-3}	As in parameter 4 but for $D_{high} = D_1$
9	N2	Number $\leq D_2$	cm^{-3}	As in parameter 1 but for $D_{high} = D_2$
10	L2	Length $\leq D_2$	nm cm^{-3}	As in parameter 2 but for $D_{high} = D_2$
11	S2	Surface $\leq D_2$	μm^2 cm^{-3}	As in parameter 3 but for $D_{high} = D_2$
12	V2	Volume $\leq D_2$	μm^3 cm^{-3}	As in parameter 4 but for $D_{high} = D_2$
13	DL2	Median length diameter $\leq$ D2	nm	$\int_{\log D_{low}}^{\log DL2} \frac{dn(d\log D_p)}{d\log D_p} D_p d\log D_p = 0.5 \int_{\log D_{low}}^{\log D_2} \frac{dn(d\log D_p)}{d\log D_p} D_p d\log D_p$
14	DS2	Median surface diameter $\leq$ D2	nm	As in parameter 13 but for particle surface
15	DV2	Median volume diameter $\leq$ D2	nm	As in parameter 13 but for particle volume
16	DL	Median length diameter	nm	$\int_{\log D_{low}}^{\log DL} \frac{dn(d\log D_p)}{d\log D_p} D_p d\log D_p = 0.5 \int_{\log D_{low}}^{\log D_{high}} \frac{dn(d\log D_p)}{d\log D_p} D_p d\log D_p$
17	DS	Median surface diameter	nm	As in parameter 16 but for particle surface
18	DV	Median volume diameter	nm	As in parameter 16 but for particle volume

We emphasize that these 18 parameters serve primarily to identify particle formation events. Some of them have no direct physical significance. We have ignored aerosol-related process parameters, such as the sink rate for condensing vapours ('condensational sink', *CS*). The parameter *CS* is an integral parameter that integrates over the size distribution between the weights of the first and second moment (*cf*. e.g. Dahneke's formulation in Seinfeld and Pandis, 1998), and can be readily derived from a size distribution. However, we found that the additional use of CS did not improve the algorithm's ability to detect particle formation events compared to the set of the existing 18 parameters, so it was not considered in the search algorithm. Most likely, this is a result of CS being a size distribution moment in between moments one and two, which are already considered in the set of 18 parameters.

The choice of the upper diameter limits D_1 and D_2 of two subranges of the particle size distribution needs explanation. In order to obtain maximum sensitivity to nucleation events, D_1 should be a size threshold as low as possible because the formation events are caused by stable clusters formed at sizes of approximately 1 nm, which need to grow into the detectable size range >3 nm prior to the visible beginning of a formation event. To motivate concrete values for D_1, we analyzed the typical particle number size distributions during particle formation events (*cf*. Fig. 2). Choosing a D_1 close to the lower limit of the TDMPS

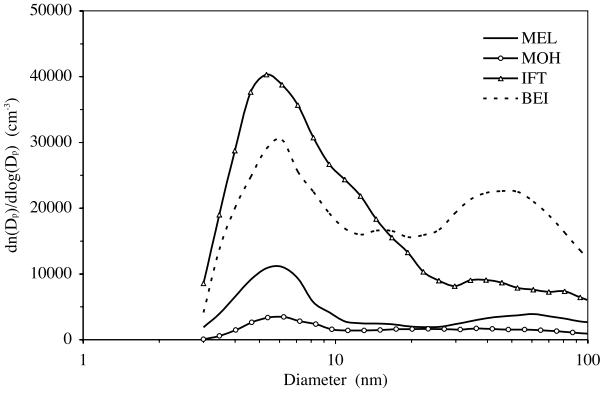

Fig. 2. Median size distributions below 100 nm particle diameter D_p for all training events at the four sites given in Table 1 during the time between the start and the end of the first hour of the respective events.

system would lead to an increased statistical uncertainty in the related integral parameters (5)–(8). An increasing D_1, on the other hand, would reduce the connection of the integrated concentration values to the most recently nucleated particles. Figure 2 demonstrates that the typical nucleation mode diameter during the time of peak concentrations was between 5 and 6 nm. Consequently, D_1 had to be as large as possible and as low as necessary to meet the above two arguments. We selected the value $D_1 =$ 10 nm for this study, which is in agreement with a previous analysis of nucleation events in Beijing (e.g., Bauer, 2006), and close to the value 11 nm employed for Hohenpeissenberg data in Birmili et al. (2003).

The upper limit of the second subpopulation D_2 should encompass the maximum diameter range of the number size distribution that may be affected by the new and evolving nucleation mode. An integration of the size distribution up to D_2 includes all particles up to D_1, but may also include larger particles that have already grown significantly at the time of observation. In continental aerosol number size distributions, a relative concentration minimum is frequently found between 20 and 30 nm, thereby demarking the nucleation and Aitken modes (e.g. Birmili et al., 2001; Heintzenberg et al., 2004). A drawback of the TDMPS technique is that in this size range, an instrumental overlap between the two subsystems (Differential Mobility Analyzer, DMA; Ultrafine DMA, UDMA) may cause inferior data quality for some data periods. In order to reduce the uncertainties associated with the overlap region, a value of $D_2 =$ 20 nm was chosen for the present study, so that the second subpopulation only refers to the size distribution section collected by the UDMA.

Based on data set MEL-2 (*cf*. Table 1) a training set of 46 manually selected days with emphasized nucleation and growth events was compiled, covering the time period 25 July 2004 to 20 September 2005. These event days were determined visually and represent a comprehensive, though subjective selection of formation events. With an analysis of these training days, the most sensitive and most specific of the 18 parameters were to be found for their subsequent utilization in the search algorithm. The values of the 18 integral parameters (*cf*. Table 2) at each time step during a search window of each training day, (07:00 to 18:00 local time, beginning roughly with sunrise), were divided by their respective average value during the reference period (18:00 of the preceding day to 18:00 of the training day). This normalization by reference values was introduced to be able to distinguish nucleation events independently of the absolute number concentrations of the aerosol.

The start time of a particle formation is now defined as the time, at which $N1/N1_{\mathrm{ref}}$ (*N1* being parameter 5 in Table 2 integrating the number size distribution up to the upper limit D_1) reaches the value of one during the search window of any of the training days. Thus, the temporal evolutions of all other integral parameters were analyzed on this relative timescale defined through the evolution of $N1/N1_{\mathrm{ref}}$.

The choice of *N1* becomes clear when analyzing the median relative temporal evolution of all 18 parameters taken over the whole training set. In Fig. 3, their relative change G_{i} around the event starting time $t_0 = 0$ (defined by *N1*) is displayed. G_{i} is calculated by

$$G_i = \frac{\left(P_i^{t_{0+1}} - P_i^{t_{0-1}}\right)}{(t_{0+1} - t_{0-1})}, i = 1, 18.$$

The time steps i depend on the temporal resolution of the aerosol measurements. For the data sets of the present study, this resolution is given in Table 1. The largest increase about $t_0 = 0$ occurs in *N1* (5), the largest decrease in *DL2* (13) whereas DV (18) exhibits the greatest stability about the start of the training events. Thus, these three parameters exhibit the most widely differing temporal change at the start of an event. Corresponding calculations with training sets from the other three sites, (not shown here), gave very similar results.

Consequently, the relative temporal evolution during nucleation and growth events of the three key aerosol parameters (5, 13, 18) were analyzed in more detail. In the search algorithm

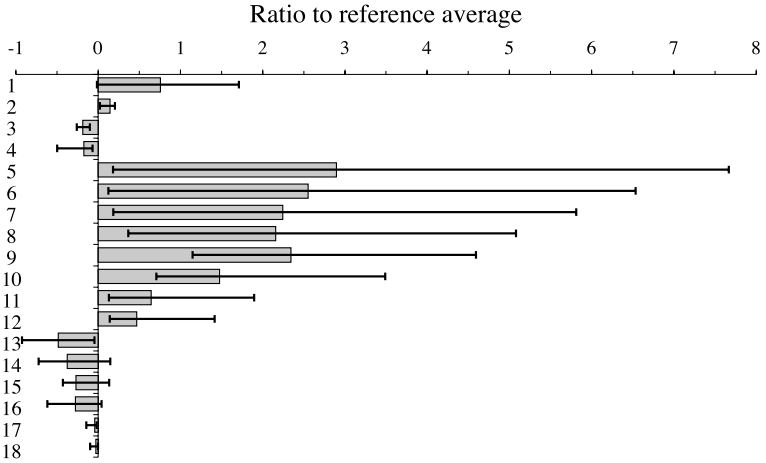

Fig. 3. Relative rate of change of the 18 integral parameters (top to bottom left of vertical axis) in Table 2 around the event starting time $t = 0$, as determined from parameter 5 (*N1*, particle number $\leq$ 10 nm diameter). The analysis is based on a training set of 46 nucleation events in data set MEL-2. The error bars correspond to the 25 and 75 percentiles.

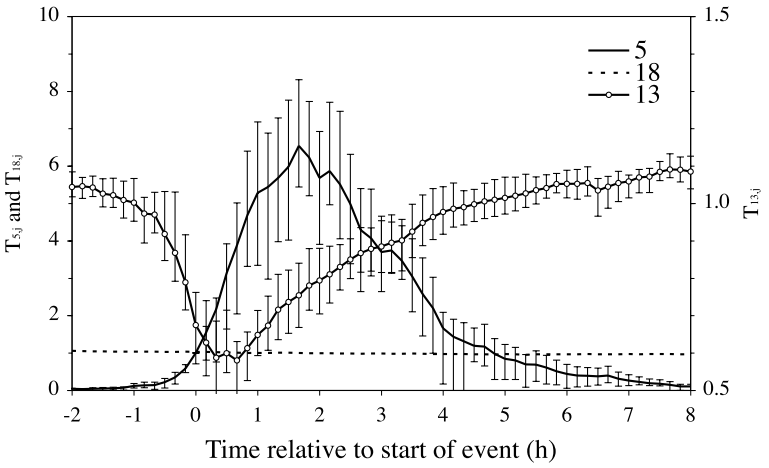

Fig. 4. Median temporal evolution of the relative integral parameters $T_{5,j}$ (providing the relative time base of time steps *j*), $T_{13,j}$ and $T_{18,j}$ in the Melpitz training set of 46 nucleation events (*cf.* Table 2 for parameters). The error bars correspond to the 25 and 75 percentiles. Error bars at curve 18 are too small to be displayed.

these relative parameters are named $T_{i,j}$ with $i = 5$, 13, 18 and j being the index of time steps of the time-series from 2 hr before until 8 hr after t_0. Their median temporal evolution at Melpitz is displayed in Fig. 4 together with 25 and 75 percentiles. Within 2 hr after the start of an event, *N1* increases to six times its reference value decreasing after that to the reference level by hour 5. *DL2* starts to decrease somewhat before *N1* increases significantly and reaches its minimum value of about 60% of the reference value already within the first hour of the event with subsequent slow recovery to the reference level by hour 4. The volume median diameter *DV* (18) stays nearly constant throughout the analyzed 10 hr. The percentile error bars in Fig. 4 indicate that most of the interevent variability occurs during the first 3 hr of the events.

Encouraged by the clearly defined description of events at Melpitz by means of three key aerosol parameters, corresponding training sets of well defined nucleation and growth events at the other three sites were compiled manually. At Hohenpeissenberg, 23 event days dating from 30 April 1999 to 30 June 2000 comprised the training set. At IfT, the training set covered 29 days dating from 13 April 1997 to 17 April 2003. In Beijing, 50 event days dating from 07 March 2004 to 29 January 2006 were included in the training set. The training events were again selected after the two criteria (a) high nucleation mode concentration and (b) temporal evolution of the geometric mean diameter of the nucleation mode. Median relative temporal evolutions of the three key parameters at all sites are plotted in Fig. 5. At all sites, the temporal developments of the key aerosol parameters is strikingly similar. At the two rural sites (MEL, MOH), high particle numbers (*N1*) $\leq$ 10 nm tend to be sustained over longer times than at the urban sites. In particular, at Beijing *N1* does not reach as high values as at the other sites. However, *N1* also drops off more slowly at Beijing than at the other sites. Median length diameters decrease most at the two rural stations (*cf*. Fig. 5b), while median volume diameters of the submicrometer aerosol remains practically constant on average at all stations, possibly with a minor trend downward during the first 5 hr of the training events (*cf*. Fig. 5c).

4. The search algorithm

The primary parameter spaces of the search algorithm are two time windows: (a) the reference time window (usually 18:00 of the preceding day to 18:00 of a search day), and (b) the search time window (usually 07:00–18:00 of a search day). Both time windows were chosen empirically for the present study and could be varied according to the setting for which the algorithm is applied. The three key integral parameters P_{ij} are normalized by their respective average values taken over the reference time window of the given search day just as the $T_{i,j}$ were for the training events. The search algorithm then compares the relative time-series of the three key aerosol parameters $P_{i,j}$ during the search window of the given search day with the $T_{i,j}$ during the training days.

The algorithm steps through all possible time steps t_x of the search window for which TDMPS data at a preset number n of time steps before and after t_x are available. The quality of coincidence of the *3n* normalized aerosol parameters and the respective training values is tested by calculating a penalty function *PF* at each time t_x of the search window. This penalty function is defined as

$$PF = \frac{1}{3n} \sum_{j=1}^{n} \sum_{i=1,3} abs \left(\frac{P_{i,j} - T_{i,j}}{T_{i,j}} \right),$$

with j being the index of the n time steps of the time-series. *PF* is calculated for all possible t_x of a search window. If the minimum of *PF* falls below a preset upper limit then the corresponding time period of the search window is interpreted as an event of the sought-after type. By specifying the number of time steps before and after t_x the part of the training curves $T_{i,j}$ against which

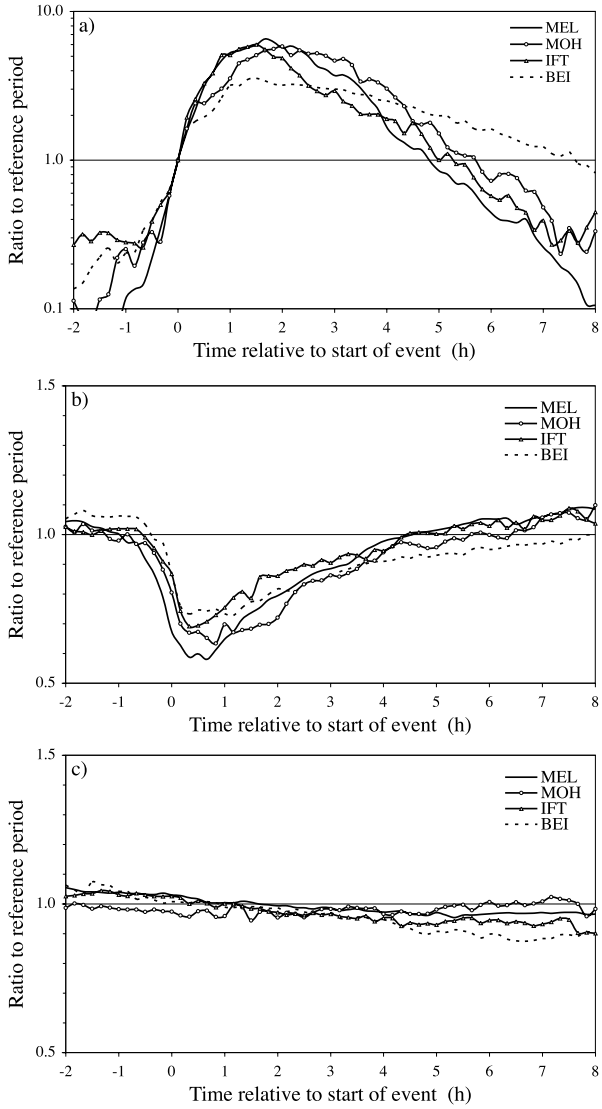

Fig. 5. (a) Median temporal evolution of the integral parameter 5 (particle number $\leq$ 10 nm) relative to its average value taken over the time period 18:00 of the event-preceding days to 18:00 of the event days in the training sets at Melpitz (MEL), Hohenpeissenberg (MOH), IfT and Beijing (BEI); (b) Median temporal evolution of the integral parameter 13 (median length diameter $\leq$ 20 nm) relative to its average value taken over the time period 18:00 of the event-preceding days to 18:00 of the event days in the training sets at Melpitz (MEL), Hohenpeissenberg (MOH), IfT and Beijing (BEI); (c) Median temporal evolution of the integral parameter 18 (median volume diameter) relative to its average value taken over the time period 18:00 of the event-preceding days to 18:00 of the event days in the training sets at Melpitz (MEL), Hohenpeissenberg (MOH), IfT and Beijing (BEI).

the actual time-series is tested will be specified. By weighing *PF* additionally with the maximum of *N1* during any nucleation event, the strongest ones can be identified.

As the algorithm is intended for the analysis of large time-series, it includes the statistical evaluation of all 18 integral aerosol parameters during the events that have been identified by the search. From these results, nucleation rates, particle growth curves and volumes of condensing vapours are calculated.

5. Example applications

The application of the described training procedure and event search generates results that can be used to address a variety of scientific questions related to the circumstances and the origin of particle formation events. The systematic statistical synopsis of the aerosol database introduced here, in combination with meteorological and precursor gas data will be the subject of a separate paper. Here, only exemplary results are compared for the four rural and urban sites. For this comparison, an average evolution of the three search parameters *N1*, *DL2* and *DV* is used, i.e. the same type of event is searched for at all four sites. By requiring a certain threshold of *N1* to be exceeded during each event, the search results can be restricted to the most intensive events. In the examples shown here, the thresholds for the penalty function *PF* were lowered individually at each station until a number of about 50 events were found (confirmed by visual inspection of the time-series).

In Table 3 characteristic parameters of these 50 events at each site are displayed. Average start and end times differ by $\pm$1 hr and the time at which a minimum in median length diameter is reached varies between half an hour and 1.5 hr after the nominal start of the events. It is noteworthy that particle formation events start at IfT, the urban site, significantly later than at the rural site Melpitz. The reason for this observation is not explained yet. Our current speculations are that effects of particle coagulation, or urban boundary layer meteorology may be involved, but these are lacking proof. Average number, surface and volume $\leq$ 20 nm increase at all sites during the events. Interpreting the change in *N2* as nucleation rates yields values between 1.5 (at Hohenpeissenberg) and 18 cm^{-3} s^{-1} at IfT. The Hohenpeissenberg value is

Table 3. Characteristic parameters of the 50 strongest nucleation and growth events at the four sites given in Table 1. H_{start} is the average local decimal start hour, H_{end} is the average local decimal hour 4 hr after the start of an event. H_{min} is the average local decimal hour, at which the median length diameter $\leq$ 20 nm reaches its minimum during an event. Parameters ΔN2, ΔS2, ΔV2 are the average changes of the respective integral parameters (cf. Table 2) per second from the start of the event to the maximum in N2. GR2, GR3 are average growth in nm h^{-1} of surface and volume $\leq$ 20 nm particle diameter after H_{min}, respectively

Location	H_{start}	H_{min}	H_{end}	ΔN2	ΔS2	ΔV2	GR2	GR3
MEL	10.0	0.7	14.3	6.1	812	4510	1.5	1.1
MOH	10.5	1.4	14.9	1.5	141	675	1.1	0.8
IFT	11.8	0.5	16.1	18	231	6700	0.6	0.2
BEI	9.8	1.0	14.1	6.7	847	3390	0.4	0.3

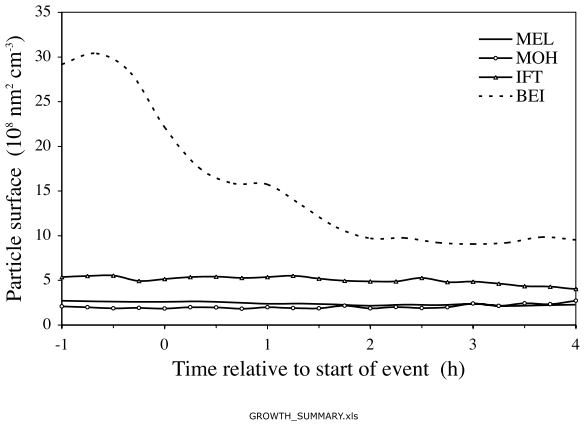

Fig. 6. Median temporal evolution of the integral parameter 3 (total surface diameter) during the 50 strongest events at Melpitz (MEL), Hohenpeissenberg (MOH), IfT and Beijing (BEI).

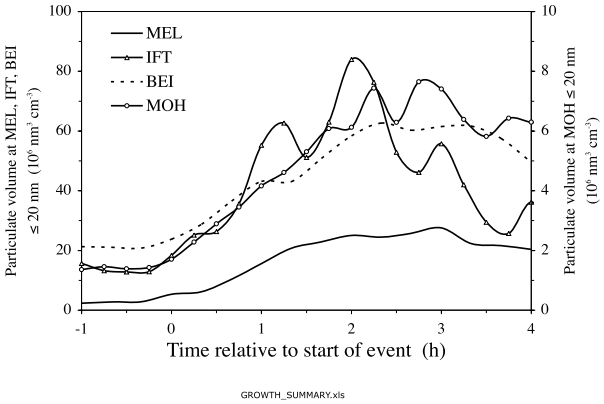

Fig. 7. Median temporal evolution of the integral parameter 12 (particle volume $\leq$ 20 nm diameter) during the 50 strongest events at Melpitz (MEL), Hohenpeissenberg (MOH), IfT and Beijing (BEI).

in the range of values provided by Birmili et al.'s (2003) analysis for the formation rates of 3–11 nm particles. For Beijing and IfT, no growth rates have been published yet. The volume increase $\leq$ 20 nm diameter can be translated into a number of molecules accreted in the condensed phase. Assuming this condensed phase to consist of concentrated H_2SO_4 yields values between 7.000 (Hohenpeissenberg) and 70.000 (IfT) molecules cm^{-3} s^{-1} as condensation rate.

Systematic differences in the temporal evolution of these 50 events at each of the four sites are illustrated in Fig. 6, which shows the median evolution of total particle surface (ST) from one hr before the start until 4 hr after the start of the formation events. It can be seen that at the two rural sites and IfT, total particle surface concentrations are all below 500 μm^2 cm^{-3} and exhibit, on average, little diurnal variation. At Beijing, total particle surface concentrations are much higher (700–2500 μm^2 cm^{-3}) because of intense local and regional pollution (Wehner et al., 2004). It is interesting to note that the average particle surface concentration decreases by one third during the hour prior to the event start, and by 70% during the entire 5-hr-period. Its sharp coincidence with the particle number increase near the event start suggests that the depicted evolution of the particle surface concentration is linked to the atmospheric processes causing the new particle formation events. Dramatic decreases in total particle surface concentration, which is controlled mainly by longer-lived accumulation mode particles, are often linked to air mass changes, such as horizontal advection or vertical mixing. A preliminary conclusion from these considerations is that the air masses establishing themselves during the particle formation event provide atmospheric conditions more favourable for nucleation than the ones prevailing before. The lower particle surface area during particle formation events is largely consistent with the process of increasing the equilibrium concentrations of condensable vapors, and the lifetimes of newly formed particles against coagulation losses.

At the two rural sites and at IfT total particulate surface concentrations during the 50 strongest events are well below 500 μm^2 cm^{-3} and show no dip before or during the events that might favour nucleation/growth. At Beijing, however, average pollution levels are much higher (Wehner et al., 2004) and the median factor of two decrease of particulate surface concentration [CS] during the hour before the onset of events generates conditions more favourable for nucleation. In the highly polluted atmosphere of Beijing, meteorological control appears crucial for new particle formation.

Figure 7 shows median evolutions of particulate volume $\leq$20 nm at the four sites illustrating that there can be differences in more than an hour in the time period during which particle growth is sustained at the different sites. Both, the length of the growth period and the linearity in growth are highest in the Beijing events, albeit neither nucleation rates nor growth rates.

6. Discussion and conclusions

In four very different rural and urban settings very similar temporal evolutions of the three key integral aerosol parameters number concentration below 10 nm particle diameter, median length diameter below 20 nm diameter and median volume diameter were found right before and during events of new particle formation. Based on these findings, an objective search algorithm has been developed, which utilizes given temporal developments of these three parameters to identify and analyze statistically events of particle nucleation and growth (bananas) in time-series of submicrometer number-size distributions. A comprehensive statistical synopsis of the existing aerosol database combined with meteorological and trace gas data will be the subject of a separate paper.

With the help of this new methodology, many questions can be put to the large databases of number-size distributions available to date, e.g., by requiring total surface or volume to follow given

courses before and during events, cases with certain nucleation and growth limitations can be identified and analyzed separately. In general, the algorithm can only find events that it has been trained for. In atmospheric time-series of high-resolution size distributions not only strong and weak nucleation events occur but also events that are interrupted by short-term local meteorological processes. The search algorithm cannot find all three aforementioned types of events at the same time. To cover both strong and weak events, the upper limit of the penalty function *PF* would have to be relaxed, which would cause many non-events, the temporal evolution of which partly follows the given training set, to contaminate the search results. Instead of that, a comprehensive event analysis of a given time-series should be performed independently with several training sets, one for strong events, one for weak events and one for interrupted events. In the latter case, one could disregard the fit of search days and training set during certain time windows of the total length of events.

The same search approach could be applied to identify and analyze other systematic changes in size distribution, such as during frontal passages or sand storms. In an inverted way, the algorithm can be used to clear time-series of events of secondary formation in source-receptor studies such as in Zhou et al. (2005).

The search algorithm reproduces to a certain extent the profiles that have been prescribed by the given training events. By describing the training data sets statistically, i.e., in terms of some of the 18 parameters, the method can be considered 'objective'' in a sense that other researchers could unambiguously reproduce the same results from the same data. Likewise, new, and corresponding training data sets will be required when analyzing other data sets by this method.

In summary, we have devised a new search algorithm for secondary particle formation events in data sets of atmospheric particle size distributions. The search algorithm is simple and based on the investigation of 18 integral parameters of the particle size distribution, three of which were found to be most relevant for identifying nucleation events. Importantly, the algorithm is able to identify particle formation events both, in clean and polluted air. By specifying the particular training sets, the method can be used by other researchers with different data sets or different research goals. Currently, the training data sets are only made available as numerical curves. It is an outlook for the future that the diurnal parameter curves of the training data sets are made available in parameterized form to allow large-scale meteorological models to include particle nucleation and growth.

7. Acknowledgements

We are strongly indebted to the highly competent IfT teams who collected and processed TDMPS data at Melpitz and Hohenpeissenberg and are still accumulating new data at the Melpitz, IfT and Beijing stations.

References

Abel, N., Jaenicke, R., Junge, C., Kanter, H., Prieto, P. R. G. and co authors. 1969a. Luftchemische Studien am Observatorium Izaña (Teneriffa). *Meteor. Rdsch.* **6**, 158–167.

Abel, N., Winkler, P. and Junge, C. 1969b. Studies of size distributions and growth with humidity of natural aerosol particles. Final Scientific Report Contract AF 61 (052) – 965, Max-Planck-Institut für Chemie, Mainz, Germany, pp. 104.

Bauer, S. 2006. *Partikelneubildung und Wachstum ultrafeiner Partikel in Peking, China. Diploma Thesis*, Universität Leipzig, Leipzig, Germany, pp. 69.

Berndt, T., Böge, O., Stratmann, F., Heintzenberg, J. and Kulmala, M. 2005. Rapid formation of sulfuric acid particles at near-atmospheric conditions. *Science* **307**, 698–700.

Birmili, W. 1998. *Production of new ultrafine aerosol particles in continental air masses*. Ph. D. Thesis, Universität Leipzig, Leipzig, pp. 107.

Birmili, W., Stratmann, F. and Wiedensohler, A. 1999. Design of a DMA-based size spectrometer for a large particle size range and stable operation. *J. Aerosol Sci.* **30**(4), 549–553.

Birmili, W. and Wiedensohler, A. 2000. New particle formation in the continental boundary layer: Meteorological and gas phase parameter influence. *Geophys. Res. Lett.* **27**(20), 3325–3328.

Birmili, W., Wiedensohler, A., Plass-Dülmer, C. and Berresheim, H. 2000. Evolution of newly formed aerosol particles in the continental boundary layer: a case study including OH and H_2SO_4 measurements. *Geophys. Res. Lett.* **27**(15), 2205–2209.

Birmili, W., Wiedensohler, A., Heintzenberg, J. and Lehmann, K. 2001. Atmospheric particle number size distribution in Central Europe: Statistical relations to air masses and meteorology. *J. Geophys. Res.* **106**(D23), 32005–32018.

Birmili, W., Berresheim, H., Plass-Dülmer, C., Elste, T., Gilge, S. and co-authors. 2003. The Hohenpeissenberg aerosol formation experiment (HAFEX): A long-term study including size-resolved aerosol, H_2SO_4, OH, and monoterpenes measurements. *Atmos. Chem. Phys.* **3**, 361–376.

Boy, M. and Kulmala, M. 2002. Nucleation events in the continental boundary layer: influence of physical and meteorological parameters. *Atmos. Chem. Phys.* **2**, 1–16.

Clement, C. F., Pirjola, L., dal Maso, M., Mäkelä, J. M. and Kulmala, M. 2001. Analysis of particle formation bursts observed in Finland. *J. Aerosol Sci.* **32**(2), 217–236.

Dal Maso, M., Kulmala, M., Riipinen, I., Wagner, R., Hussein, T. and co-authors. 2005. Formation and growth of fresh atmospheric aerosols: eight years of aerosol size distribution data from SMEAR II, Hyytiälä, Finland. *Bor. Env. Res.* **10**, 323–336.

Engler, C., Rose, D., Wehner, B., Gnauk, T., Brüggemann, E. and co-authors. 2006. Size distributions of non-volatile particle residuals (Dp $<$800 nm) at a rural observation site in Germany and relation to air mass origin. *Atmos. Chem. Phys. Discuss.* **6**, 5505–5542.

Heintzenberg, J., Müller, K., Birmili, W., Spindler, G. and Wiedensohler, A. 1998. Mass-related aerosol properties over the Leipzig Basin. *J. Geophys. Res.* **103**(D11), 13 125–13 136.

Heintzenberg, J., Birmili, W., Wiedensohler, A., Nowak, A. and Tuch, T. 2004. Structure, variability and persistence of the submicrometer marine aerosol. *Tellus* **56B**(4), 357–367.

Israel, H. and Schulz, L. 1932. Über die Größenverteilung atmosphärischer Ionen. *Meteor. Z.* **49**, 226–233.

Junge, C. 1955. The size distribution and aging of natural aerosols as determined from electrical and optical data on the atmosphere. *J. Meteor.* **12**, 13–25.

Kavouras, I. G., Mihalopoulos, N. and Stephanou, E. G., 1998. Formation of atmospheric particles from organic acids produced by forests. *Nature* **395**, 683–686.

Kerminen, V.-M., Pirjola, L. and Kulmala, M. 2001. How significantly does coagulation scavenging limit atmospheric particle production? *J. Geophys. Res.* **106**(D20), 24 119–24 126.

Kerminen, V.-M. and Kulmala, M. 2002. Analytical formulae connecting the "real" and the "apparent" nucleation rate and the nuclei number concentration for atmospheric nucleation events. *J. Aerosol Sci.* **33**, 609–622.

Kulmala, M., Kerminen, V.-M. and Laaksonen, A. 1995. Simulations on the effect of sulphuric acid formation on atmospheric aerosol concentrations. *Atmos. Environ.* **29**(3), 377–382.

Kulmala, M., Pirjola, L. and Mäkelä, J. M. 2000. Stable sulphonate clusters as a source of new atmospheric particles. *Nature* **404**, 66–69.

Kulmala, M., Dal Maso, M., Mäkelä, J. M., Pirjola, L., Väkevä, M. and co-authors. 2001. On the formation, growth and composition of nucleation mode particles. *Tellus* **53B**(4), 479–490.

Kulmala, M., Laakso, L., Lehtinen, K. E. J., Riipinen, I., Dal Maso, M., and co-authors. 2004a. Initial steps of aerosol growth. *Atmos. Chem. Phys.* **4**, 2553–2560.

Kulmala, M., Vehkamäkia, H., Petäjä, T., Maso, M. D., Lauri, A. and co-authors. 2004b. Formation and growth rates of ultrafine atmospheric particles: a review of observations. *J. Aerosol Sci.* **35**, 143–176.

Mäkelä, J., Dal Maso, M., Pirjola, L., Keronen, P., Laakso, L. and co-authors. 2000. Characteristics of the atmospheric particle formation events observed at a boreal forest site in southern Finland. *Bor. Env. Res.* **5**, 299–313.

Mäkelä, J. M., Aalto, P., Jokinen, V., Pohja, T., Nissinen, A. and co-authors. 1997. Observations of ultrafine aerosol particle formation and growth in boreal forest. *Geophys. Res. Lett.* **24**(10), 1219–1222.

McMurry, P. H., Fink, M., Sakurai, H., Stolzenburg, M. R., Mauldin, L. and co-authors. 2005. A criterion for new particle formation in the sulfur-rich Atlanta atmosphere. *J. Geophys. Res.* **110**(D22S02), doi:10.1029/2005JD005901.

Misaki, M. 1964. Mobility spectrum of large ions in the New Mexico semidesert. J. *Geophys. Res.* **69**, 3309–3318.

Pirjola, L., Kulmala, M., Bischoff, A., Wilck, M. and Stratmann, F. 1998. Effects aerosol dynamics on the formation of sulphuric acid aerosols, 5th International Aerosol Conference. *J. Aerosol Sci.*, Edinburgh, 12–18. September, pp. 825–826.

Seinfeld, J. H. and Pandis, S. N., 1998. *Atmospheric Chemistry and Physics*. John Wiley & Sons, New York NY, pp. 1326.

Stanier, C., Khlystov, A. and Pandis, S. 2004. Nucleation events during the Pittsburgh air quality study: Description and relation to key meteorological, gas phase, and aerosol parameters. *Aerosol Sci. Technol.* **38**(Suppl. 1), 253– 264.

Stolzenburg, M. R. and McMurry, P. H., 1991. An ultrafine aerosol condensation nucleus counter. *Aerosol Sci. Technol.* **14**, 48–65.

Wehner, B. and Wiedensohler, A. 2003. Long term measurements of submicrometer urban aerosols: Statistical analysis for correlations with meteorological conditions and trace gases. *Atmos. Chem. Phys.* **3**, 867–879.

Wehner, B., Wiedensohler, A., Tuch, T., Wu, Z. J., Hu, M. and co-authors. 2004. Variability of the aerosol number size distribution in Beijing, China: New particle formation, dust storms, and high continental background. *Geophys. Res. Lett.* **31**(22), L22108, doi:10.1029/2004GL021596.

Wehner, B., Petäjä, T., Boy, M., Engler, C., Birmili, W. and co-authors. 2005. The contribution of sulfuric acid and non-volatile compounds on the growth of freshly formed atmospheric aerosols. *Geophys. Res. Lett.* **32**, doi:10.1029/2005GL023827.

Winklmayr, W. 1987. *Untersuchung des ultrafeinen Aerosols in der urbanen Atmosphäre von Wien*. Ph.D. Thesis, Universität Wien, Wien, pp. 187.

Winklmayr, W., Reischl, G. P., Lindner, A. O. and Berner, A. 1991. A new electromobility spectrometer for the measurement of aerosol size distributions in the size range from 1 to 1000 nm. *J. Aerosol Sci.* **22**, 289–296.

Yu, F. and Turco, R. P., 2000. From molecular clusters to nanoparticles: Role of ambient ionization in tropospheric and aerosol formation. *J. Geophys. Res.* **106**(D5), 4797–4814.

Zhou, L., Kim, E., Hopke, P. K., Stanier, C. and Pandis, S. N., 2005. Mining airborne particulate size distribution data by positive matrix factorization. *J. Geophys. Res.* **110**(D7), D07S19.