

REPORT SERIES IN AEROSOL SCIENCE

N:o 127 (2011)

STATISTICAL STUDIES OF ATMOSPHERIC NANOPARTICLES
AND THEIR PRECURSORS

SANTTU MIKKONEN

*Department of Applied Physics
Faculty of Science and Forestry
University of Eastern Finland
Kuopio, Finland*

Academic dissertation

To be presented, with the permission of the Faculty of Science and Forestry of the University of Eastern Finland, for public examination in auditorium L21, Snellmania building, on October 28th 2011 at noon.

Kuopio 2011

Author's address: Department of Applied Physics
University of Eastern Finland
P.O.Box 1627
FIN-70211 Kuopio, Finland

Supervisors: Professor Ari Laaksonen, Ph.D.
Department of Applied Physics
University of Eastern Finland
Kuopio, Finland
and
Finnish Meteorological Institute
Helsinki, Finland

Professor Jari Kaipio, Ph.D.
Department of Applied Physics
University of Eastern Finland
Kuopio, Finland
and
Department of Mathematics
University of Auckland
New Zealand

Sami Romakkaniemi, Ph.D.
Department of Applied Physics
University of Eastern Finland
Kuopio, Finland

Reviewers: Associate Professor Ilona Riipinen, Ph.D.
Department of Applied Environmental Science
Stockholm University
Stockholm, Sweden

Tatu Anttila, Ph.D.
Finnish Meteorological Institute
Helsinki, Finland

Opponent Miikka Dal Maso, Ph.D.
Department of Physics
University of Helsinki
Helsinki, Finland

ISBN 978-952-5822-50-2 (printed version)
ISSN 0784-3496
Unigrafia Oy
Helsinki 2011

ISBN 978-5822-51-9 (PDF version)
<http://www.atm.helsinki.fi/FAAR/>
Helsinki 2011

Acknowledgements

The research for this thesis was carried out at the Department of Applied Physics in the University of Eastern Finland. Finishing the study took apparently quite a long time since the department has been lead by four different persons within the time. I would like to express my gratitude to all of them for providing me with the working facilities.

Three persons have earned my gratitude by supervising and guiding my work. First one is Prof. Ari Laaksonen, who introduced me to aerosol science and gave me the possibility work in such a remarkable research environment as the Aerosol physics group in Kuopio is. Second one is Prof. Jari Kaipio, who accepted me as a member of the staff of the Department of Applied Physics, even though I did not use Matlab, and has guided me into scientific thinking with discussions in non-formal facilities. Third member of my supervisor group was Dr. Sami Romakkaniemi whose effort in completion of this thesis was beyond a price. Sami put me through my paces while I was writing the two last articles included in this thesis and the thesis itself. Our morning coffee breaks and discussions about science, life, the universe and everything have also been priceless.

I want to thank all my current and past co-workers in the aerosol science society in Kuopio, especially Aerosol physics group in the Department of Physics and the researchers in FMI Kuopio unit. All co-authors of the studies included in this thesis are highly acknowledged. The work could not be done without multidisciplinary co-operation. Special thanks go to Prof. Kari Lehtinen, Doc. Jorma Joutsensaari and Dr. Amar Hamed who have influenced to all my papers with their expertise. I want to thank Associate Prof. Ilona Riipinen and Doc. Tatu Anttila for carefully reviewing of this thesis and making helpful comments on it. I gratefully thank Prof. Jim Smith for helping me to improve the language of the thesis and the last two papers.

My greatest gratitude goes to my family and friends. My parents and my sister have given me the building blocks of my life, encouraged me to find my own way in life and helped me to reach this far. I have a bunch of great friends who have grown up with me and supported me in all the turning points of my life and I want to thank them for that. I want to thank my sons Eetu and Leevi for taking my mind out of the work. I am truly fortunated to have such wonderful kids. Finally, I want thank my beloved Emma for caring, loving, understanding my oddities and supporting me at every turn.

Abstract

Aerosols affect our everyday life in many ways. Understanding the processes controlling the formation and growth of the aerosol particles is essential in order to estimate their effect on the global climate, air quality and human health. A large number of measurements on the composition of the atmosphere have been made, and are made all the time, but large amounts of the collected data still requires further analysis. Appropriate methods are needed for the analysis of large datasets, and statistical methodology offers the tools for finding the properties of the data. In this work we studied the usability of several different statistical methods for extensive measurement data. We applied the tried and tested methodologies for the studies included in this thesis, namely discriminant analysis for new particle formation (NPF) event classification, mixed models for parameterizing the number concentration of aerosol particles and nonlinear regression for constructing proximity measure for sulphuric acid concentration.

The main conclusions of the work presented in this thesis are: 1) New particle formation is frequent phenomenon in semi polluted and polluted areas such as San Pietro Capofiume, Italy, Melpitz and Hohenpeissenberg, Germany, and conditions favouring new particle formation can be characterized with measured atmospheric variables, 2) The concentration of 50 nm sized particles can be predicted with in-situ meteorological and gas phase parameters by using a multivariate mixed effects model. 3) New particle formation was found to be important source of Cloud Condensation Nuclei (CCN): probability of a NPF event was found to be a significant predictor for the number concentration of CCN-sized particles. 4) Decreasing sulphur dioxide concentration was found to decrease the sulphuric acid concentration in the atmosphere. We were able to introduce a universal proximity measure for sulphuric acid concentration. CCN concentration was found to increase when sulphur dioxide concentration was decreased.

The work presented in this thesis increases our understanding of the formation and growth processes of new particles. The trace gases and meteorological variables that participate directly in the processes or indicate the conditions needed for the processes to take place are specified. The study also gives tools for the global atmospheric models to get more precise results without the need for increased computational resources.

Contents

Acknowledgements	1
Abstract	4
List of publications.....	6
1 Introduction	7
2 Sources of atmospheric aerosol particles	12
2.1 New particle formation in the atmosphere	14
2.2 The role of sulphuric acid in new particle formation.....	16
2.3 Observed formation of new aerosol particles	18
2.4 Growth of particles to Cloud Condensation Nuclei (CCN)	20
2.5 Modelling the new particle formation and growth processes.....	22
3 Methods.....	23
3.1 Measurements	23
3.2 Statistical methodology.....	28
3.2.1 Discriminant analysis	28
3.2.2 Mixed Models.....	30
3.2.3 Nonlinear regression analysis	35
4 Review of papers and author's contribution.....	38
5 Conclusions	40
6 References	43

List of publications

This thesis consists of introductory review, followed by four research articles. The papers are reproduced with the kind permission of the journals concerned.

- I. Mikkonen, S., Lehtinen, K. E. J., Hamed, A., Joutsensaari, J., Facchini, M. C. and Laaksonen, A. (2006) Using discriminant analysis as a nucleation event classification method. *Atmos. Chem. Phys.*, 6, 5549-5557.
- II. Hamed, A., Birmili, W., Joutsensaari, J., Mikkonen, S., Asmi, A., Wehner, B., Spindler, G., Jaatinen, A., Wiedensohler, A., Korhonen, H., Lehtinen, K. E. J., and Laaksonen, A. (2010) Changes in the production rate of secondary aerosol particles in Central Europe in view of decreasing SO₂ emissions between 1996 and 2006. *Atmos. Chem. Phys.* 10 1071-1091.
- III. Mikkonen, S., Korhonen, H., Romakkaniemi, S., Smith, J. N., Joutsensaari, J., Lehtinen, K. E. J., Hamed, A., Breider, T. J., Birmili, W., Spindler, G., Plass-Duelmer, C., Facchini, M. C., and Laaksonen, A. (2011) Meteorological and trace gas factors affecting the number concentration of atmospheric Aitken (D_p=50 nm) particles in the continental boundary layer: parameterization using a multivariate mixed effects model. *Geosci. Model Dev.* 4 1-13.
- IV. Mikkonen, S., Romakkaniemi, S., Smith, J. N., Korhonen, H., Petäjä, T., Plass-Duelmer, C., Boy, M., McMurry, P. H., Lehtinen, K. E. J., Joutsensaari, J., Hamed, A., Mauldin III, R. L., Birmili, W., Spindler, G., Arnold, F., Kulmala, M., and Laaksonen, A. (2011) Statistical proxy for sulphuric acid concentration. *Atmos. Chem. Phys. Discuss.*, 11, 20141-20179.

1 Introduction

There are small particles in the air floating everywhere around us. The size of these particles ranges from nanometers to tens of micrometers, and they are composed of wide variety of different chemical compounds. The mixture of these particles and the gas surrounding them is called aerosol. However, in the common language and in this thesis, aerosol refers to the particles themselves. Atmospheric aerosols can be divided into two groups defined by their origin: primary and secondary aerosols (see e.g. Seinfeld and Pandis, 2006). Primary particles are emitted directly to the atmosphere and some of the primary aerosols are formed anthropologically by different human activities like industry, transportation and home heating. However, most primary aerosol mass is of natural origin, for example sea salt, desert dust and organic aerosols like pollen and bacteria. Secondary aerosols are formed in the atmosphere by nucleation, i.e. gas-to-particle phase transformation, followed by condensation of vapours. Although aerosols are ubiquitous, their concentration may vary by several orders of magnitude between different areas. In very clean arctic areas concentrations may be as low as some tens of particles in a cubic centimetre (e.g. Koponen et al., 2003) but in polluted urban areas concentration may be as high as several hundreds of thousands particles in cubic centimetre (e.g. Mönkkönen et al., 2005).

Atmospheric aerosols have a great influence on the radiation balance of the Earth and atmosphere. Aerosols scatter solar radiation, which decreases the amount of radiation that reaches the ground and increases the albedo of the Earth (e.g. Zhang, 2010). This is called the direct effect of aerosols (e.g. Ramanathan et al., 2001). The indirect effect of aerosols is that they act as nuclei for cloud droplets thus inducing the formation of clouds. The number concentration of aerosols in the air affects the properties of the clouds. If the number concentration of cloud condensation nuclei is high there will also be more cloud droplets, and the cloud droplets become smaller as the amount of condensing water is almost independent on the aerosols. Smaller cloud droplets will make cloud optically thicker and thus increase albedo (First indirect effect, Twomey, 1977). Also, smaller cloud droplets will form drizzle less effectively and thus cloud lifetime may be increased (Second indirect effect, Albrecht 1989). Aerosols may also have warming effect on the atmosphere. Black Carbon (BC) is the main light-absorbing component of

aerosols (Seinfeld, 2008 and references therein). BC particles heat the surrounding atmosphere by absorbing energy from solar radiation.

The effect of aerosols on the climate has been recognized by the Intergovernmental Panel on Climate Change (IPCC). In their reports, including the most recent one (IPCC, 2007), an increasing number of aerosols is suggested to decrease the amount of solar radiation reaching the surface of the Earth through both the direct and indirect effects. This in turn has a cooling effect on climate. This effect competes with the warming caused by increased concentrations of atmospheric greenhouse gases, like carbon dioxide, methane, nitrous oxide, and the chlorofluorocarbons (CFCs).

The second major effect of aerosol relates to human health. The World Health Organization (2006) estimates that yearly 2 million people in the world suffer premature deaths because of air pollutants. The premature deaths caused by aerosols is just a tip of the iceberg, also other severe health effects including asthma, lung cancer, cardiovascular issues and birth defects are known to be related to poor air quality, primarily particulate air pollution. Pope and Dockery (2006) summarized the results of several different studies of health effects of aerosols, especially Particulate Mass (PM), i.e. the total mass of aerosol particles suspended in the air. They concluded that there exists evidence of PM-related cardiovascular health effects but recognized that much more research has to be done before all gaps in understanding the health effects of aerosols are filled. It is also suggested that particle number concentration is more important for health effects than total particle mass (Lighty et al., 2000). Nanometre sized particles, which have negligible contribution to PM but large number concentration in the atmosphere, have a larger surface to volume ratio and thus larger surface area compared to the mass and highest content of potentially toxic hydrocarbons among all PM sources. They can also penetrate deeper into lung tissue than fine particles (0.1 to 2.5 μm in diameter) or coarse particles, which have a diameter more than 2.5 μm . (Nel, 2005 and references therein).

The atmosphere of the Earth is an indisputably complex system with a huge number of different processes affecting aerosol that are not well understood. Four different sources of knowledge are used to increase the common understanding of the atmospheric processes affecting aerosol concentration: field measurements, laboratory experiments, theoretical calculations and modelling studies. So far, despite of remarkable research effort, extensive description of

nanoparticles and interaction mechanisms of the atmospheric processes, e.g. atmospheric nucleation or the subsequent growth processes of atmospheric nanoparticles, have not been attained. Even if the nature of some of the processes could be solved, the process in question may interact with other processes in so many ways that understanding the entirety of the process chain with all possible feedbacks is difficult. Atmospheric global scale computer models are applied to address this complexity. The models include physically based descriptions of different processes and thus include also feedbacks in order to estimate how the atmospheric processes really work. However, these models have severe limitations which hinder their ability to make detailed descriptions of the atmosphere. For example the models require so much computing power that given current knowledge a situation where every detail could be calculated with full accuracy is not foreseen (Dusek et al., 2006). Thus simple parameterizations of certain atmospheric processes are needed.

One of the key phenomena associated with the atmospheric aerosol system is the formation and growth of new atmospheric aerosol particles. This is due to the abovementioned properties of aerosols and the fact that several studies have shown that New Particle Formation (NPF) events make an important contribution to CCN concentrations (e.g. Lihavainen et al., 2003; Kerminen et al., 2005; Laaksonen et al., 2005; Spracklen et al., 2006, 2008). Formation of new particles in the atmosphere is a complex process including the nucleation of nanometre-size clusters from atmospheric vapours, the growth of these clusters to sizes at which the current measurement devices are able to detect, and the removal of growing clusters by coagulation with the pre-existing aerosol particle population. Kulmala et al. (2004) and Kerminen et al. (2010 and references therein) concluded that new particle formation takes place almost everywhere in the troposphere but several gaps in our knowledge regarding this phenomenon still exist. The exact particle formation mechanisms and the roles of the participating vapours in different environments are still unclear; researchers do not fully understand how aerosols form at the molecular level (Zhang, 2010). It is not even known whether atmospheric nucleation is dominated by a single nucleation pathway, or whether multiple different mechanisms are competing with each other (Kerminen et al., 2010). In addition, the importance of the physical factors influencing nucleation rate and growth of the nucleated particles is not clear and the roles of organic compounds and atmospheric ions in the nucleation process are still to be determined.

Sulphuric acid has been proposed to be the key component in atmospheric new particle formation (Weber, et al. 1996; Kulmala et al., 2004; Kerminen et al., 2010; Kuang et al., 2010; Sipilä et al., 2010) and a major contributor on the growth of freshly formed particles (Stolzenburg et al., 2005; Kulmala et al., 2006; Laaksonen et al., 2008a) along with ammonium salts and other organic compounds (Kuang et al., 2010; Smith et al., 2010). However, the exact roles of different components of the formation and growth processes of the aerosol particles are to be determined. The development of the instrumentation used in atmospheric measurements has progressed rapidly during the past decades, but until recently the devices have not been able to observe nucleation directly (Kulmala et al., 2007; Sipilä et al., 2010). The lack of a proper mechanistic understanding of atmospheric nucleation has made it difficult to develop reliable aerosol formation parameterisations for large-scale modelling frameworks.

Aerosol measurements produce large amounts of data and all associated properties and formation processes cannot be detected with simple physical data analysis methods. Physical properties of NPF have been widely studied, e.g. in Mäkelä et al. (1997) and in Dal Maso et al., 2005. However, the use of statistical methods has been rare in the investigation of NPF events and particle concentrations, although they are a powerful tool in the analysis of large measurement datasets. Hyvönen et al. (2005) introduced some statistical data mining methods to explain new particle formation at Hyytiälä station in Finland. However, their model is not as successful for data recorded in more polluted areas like Melpitz and Hohenpeissenberg, Germany and Po Valley, Italy (**Papers I, III**). Wehner et al. (2003) studied correlation between sub micrometer urban aerosols and meteorological and trace gas variables. They used time series analysis and principal component analysis in order to detect diurnal patterns and interdependencies of the measured variables with similar results as reported in **Paper III** and in Hamed et al. (2007). Some statistical predictions on aerosol size distribution have been made (e.g. Voutilainen et al., 2001) but **Paper III** was the first study in the author's knowledge trying to predict the concentration of certain sized particles with statistical methods. Later, Clifford et al. (2011) used Generalised Additive Model (GAM) for modelling the number concentration of outdoor, airborne ultrafine particles in Helsinki, Finland, but their results have not been tested in any other environment. A proximity measure for sulphuric acid concentration was introduced in Petäjä et

al. (2009) and in **Paper II** but a more thorough statistical analysis such as that described in **Paper IV** has not been made previously.

The purpose of this thesis is to both test how certain statistical methods are suited for different kinds of measurements and then use the most applicable methods in the analysis of aerosol measurement data. The statistical methods applied, like the mixed model structure used in **Paper III**, are able to take account of the variance-covariance structure of the data. In addition, statistical methods allow for utilizing additional information that is not provided by the current measurements. The methods used in this work are tested for generalization across wider set of measurements and their use outside the locations where the measurements were made is also evaluated. This thesis introduces a variety of different methods for use in different kinds of aerosol studies. **Papers I-IV** increase our understanding of the atmosphere and the processes involved in new particle formation and growth of the newly formed particles and **Papers I, III and IV** provide statistical parameterizations of processes which can be used in large scale models to reduce their computing time.

The main objectives of the work presented in this thesis include

1. Investigating factors affecting new particle formation and constructing a statistical model which is able to predict the NPF days with measured meteorological and trace gas variables (**Paper I**).
2. Studying growth of freshly formed particles and constructing parameterization for 50nm sized particles with measured meteorological and trace gas variables for the use of large scale atmospheric models (**Paper III**).
3. Investigating the role played by newly formed particles in the formation of cloud condensation nuclei (**Papers II, III**).
4. Studying sulphuric acid formation in the atmosphere and the role of sulphur dioxide in the formation process as well as constructing a proximity measure for the sulphuric acid concentration (**Papers II, IV**).

2 Sources of atmospheric aerosol particles

A large fraction of aerosol is of primary origin. Natural sources, i.e. those fully unaffected by human activities, produce aerosol particles. Such sources are for example: natural and semi-natural vegetation, biomass burning and forest fires, windblown dust, volcanoes and sea salt. Even though natural sources are responsible of most of the primary particle production, most of the fine and ultrafine aerosol materials that concern public health agencies derive from human activities (Kennedy et al., 2007).

Determining the origin of the aerosol particles is usually not so straightforward. Sometimes it is possible to detect the primary emissions from particle size distribution measurements in standard atmospheric conditions. For example Figure 1. from **Paper II** shows average size distributions for strong NPF event and nonevent days separately. The peak in nonevent days in 1996-1997 data around 15 nm was most probably caused by diesel emissions from the nearby roads.

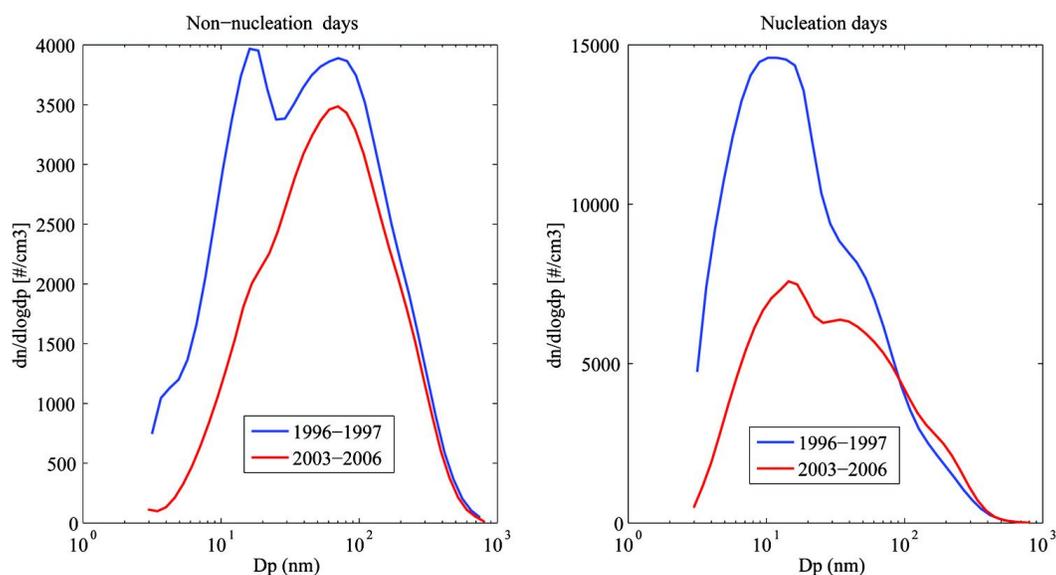


Figure 1. Average size distributions for nonevent and NPF event days in Melpitz, Germany during two measurement periods (**Paper II**)

Anthropogenic pollution originates mostly from stationary combustion sources like burning coal, fuel oil, biomass and waste, and from internal combustion engines burning gasoline and diesel (Lighty et al., 2000). Thus the primary aerosol particles and gaseous pollutants such as SO₂,

NO_x , O_3 and CO are highly correlated, which makes it possible to use measured pollutant levels as covariates in the prediction of the concentration of the aerosol particles as made in **Paper III**.

Satellite studies like Veefkind et al. (2011) have found significant correlations between Aerosol Optical Thickness (AOT) and trace gas columns from the Ozone Monitoring Instrument (OMI) for various regions of the world. They show that enhanced aerosol and trace gas concentrations originate mostly from combustion, biomass burning and organic compounds released from the biosphere.

New aerosol particles are formed in the atmosphere constantly and all over the world (Kulmala et al., 2004). A large fraction of the newly formed particles are scavenged from the air by deposition or coagulation to larger particles but the remaining particles are able to grow into sizes where they start to have effects on the climate by forming CCN, which in turn affect cloud properties. The process where molecules in the air become CCN is presented in Figure 2 and will be described in more detail in following sections. The figure also represents the topics of this thesis, aerosol growth from nucleation to CCN. After reaching the CCN size some particles may form cloud droplets, and some may not, depending such properties as particle composition and ambient water supersaturation. Ultimately particles will be removed from the atmosphere by coagulation as well as wet and dry deposition processes.

In wet deposition aerosol particles are scavenged by precipitation, falling snow or collisions with cloud droplets. The particles are removed from the air only if the raindrops come all the way down to surfaces on the ground. If the droplets are evaporated the particles will be re-released to the air, possibly after transforming to some other form. Dry deposition takes place close to the surface, where the particles are deposited on the ground or vegetation. The main causes for dry deposition are gravitational sedimentation, interception in which particles arrive too close to obstacle, impaction caused by inertia in which the particle hits an obstacle as it is not able to follow turning streamlines, and Brownian motion induced diffusion, which is a process where random movement of aerosol particles causes diffusion to surfaces (Seinfeld and Pandis, 2006). The deposition efficiency depends on the particle size. The largest particles are removed by gravitational settling and the smallest by Brownian diffusion. Wet deposition is a more effective aerosol particle removal than dry deposition but in the viewpoint of this thesis dry deposition is more interesting since the new particle formation is observed only in dry weather conditions.

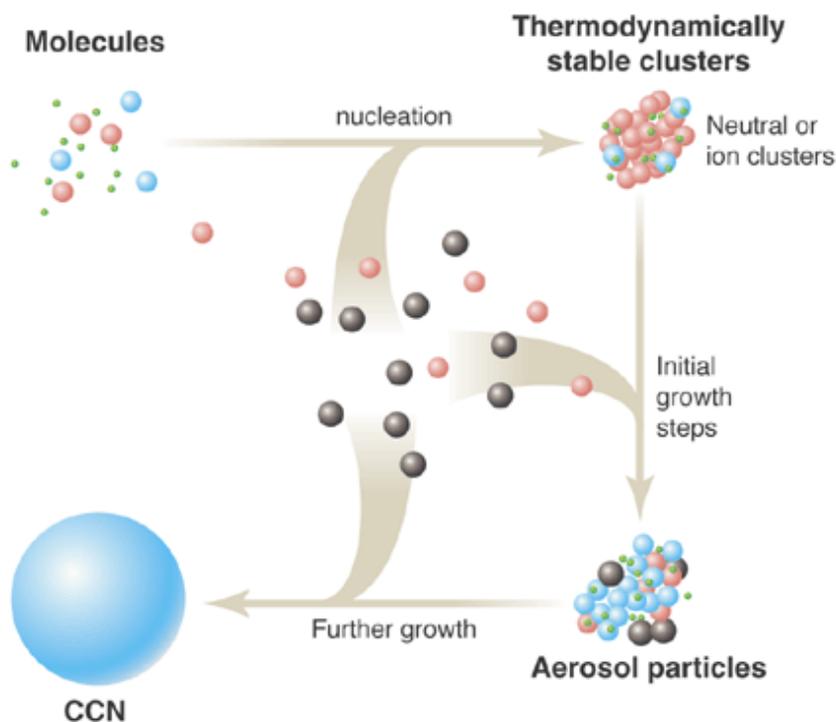


Figure 2. How particles form and grow [Reprinted from Kulmala (2004) with a small adjustment. Reproduced by permission of The American Association for the Advancement of Science (AAAS)]

2.1 New particle formation in the atmosphere

New particles are formed in the atmosphere when supersaturated vapours nucleate into liquid or solid particles and grow into detectable sizes. In the classical nucleation theory the molecules and atoms of the vapour form critical clusters which are able to grow into detectable sizes. Cluster being critical means that the cluster is stable and grows spontaneously when conditions remain constant. Atmospheric aerosol formation is initiated by photochemical reactions in the gas phase, in particular the formation of sulphuric acid (**Paper II; Paper IV**) and other vapours of very low volatility, such as multifunctional organic compounds (Hamed et al., 2007; Laaksonen et al., 2008b, Metzger et al, 2010) and iodine oxides (O'Dowd, 2002).

Aerosol formation dynamics are governed by complicated processes which are not completely understood. Even if the processes were fully known it would be impossible to provide an exact formulation of the formation process in aerosol models that are used in global scale applications

since the exact presentation of the processes would be computationally too expensive. Thus the processes need to be parameterized. Kerminen and Kulmala (2002) introduced a formula which calculates rate J of newly formed particles during new particle formation events. New nuclei of diameter d_{nuc} are produced at rate $J(t)$ (nuclei $\text{cm}^{-3} \text{s}^{-1}$) and at the same time their number concentration is decreased because of different removal mechanisms. The growth of the new nuclei is driven by condensation and self-coagulation, from which the latter is also important removal mechanisms for the nuclei. The formula is valid when it can be assumed that the only important sink for the nuclei as they grow is their coagulation to larger pre-existing particles, the nuclei grow by condensation at a constant rate, and the background aerosol size distribution remains unchanged during growth. Lehtinen et al. (2007) updated the formula for formation rate of particles with diameter d_p and currently it is given by,

$$J_{d_p} = J_{nuc} \exp\left(\frac{d_{nuc}}{m+1} \left[1 - \left(\frac{d_p}{d_{nuc}}\right)^{m+1}\right] \frac{\text{CoagS}(d_{nuc})}{GR}\right). \quad (1)$$

Where J_{nuc} is the nucleation rate, d_{nuc} is the initial size of nucleated clusters, $\text{CoagS}(d_{nuc})$ is their coagulation sink, i.e. the rate at which they coagulate with pre-existing aerosol particles, GR is their growth rate, and m ($\sim 1.5-2$) is a constant that depends on the shape of the particle number size distribution (Kerminen et al., 2010). This means that for solving the formation rate of particles of diameter d_p , one needs to estimate first the nucleation rate and the particle number size distribution, which is obtained from measurements. The relation between the nucleation rate and the identity and concentrations of nucleating vapours is one of the most important problems in this regard. For example, the functional dependence of the nucleation rate on the gaseous sulphuric acid concentration in the atmosphere appeared very different from that seen in laboratory experiments, and existing classical nucleation theories could not reconcile both field or laboratory data (Curtius, 2006; Sipilä et al, 2010).

Several different ways to estimate the nucleation rate have been suggested. It has been shown that simple power-law dependence on sulphuric acid concentration predicts atmospheric nucleation well (Weber et al., 1996; Sihto et al., 2006; Riipinen et al., 2007; Kuang et al., 2008; Paasonen et al., 2010). Kerminen et al. (2010) listed the most common sulphuric acid based formulations used for formation rate of 2 nm particles and new formulations introduced by Paasonen et al. (2010) which include the nucleating organic species:

$$J_2 = A[H_2SO_4] \quad (2)$$

$$J_2 = K[H_2SO_4]^2 \quad (3)$$

$$J_2 = A_{org}[NucOrg] \quad (4)$$

$$J_2 = K_{org}[NucOrg]^2 \quad (5)$$

$$J_2 = A_{s1}[H_2SO_4] + A_{s2}[NucOrg] \quad (6)$$

$$J_2 = K_{het}[H_2SO_4] \times [NucOrg] \quad (7)$$

$$J_2 = K_{SA1}[H_2SO_4]^2 + K_{SA2}[H_2SO_4] \times [NucOrg] \quad (8)$$

$$J_2 = K_{S1}[H_2SO_4]^2 + K_{S2}[H_2SO_4] \times [NucOrg] + K_{S3}[NucOrg]^2 \quad (9)$$

Here [NucOrg] refers to the concentration of currently unidentified organic vapour(s) participating in nucleation, and A_i and K_i are the first and second order nucleation coefficients, respectively. Paasonen et al. (2010) investigated how the suggested formulas predicted the measured formation rate in four different locations and found out that equations (8) and (9) gave the best prediction for nucleation rate in all sites. However, a problem in the estimation of the formation rate arises when [NucOrg] is taken as a predictor term. The organic vapour concentration is not a measured value, since NucOrg is itself unidentified, but it is usually derived from the closure of 2–4 nm particle growth rates (Paasonen et al., 2010), which is a significant source of uncertainty in the estimation procedure.

2.2 The role of sulphuric acid in new particle formation

As the gas phase sulphuric acid concentrations are difficult to measure the major problem in estimating the formation rate J is that the model needs to have some way of estimating the sulphuric acid concentration (Kerminen et al., 2010). The estimates for sulphuric acid concentration are based on the currently accepted mechanism of atmospheric SO_2 oxidation (Finlayson-Pitts and Pitts Jr., 2000):



As reactions R1-3 show, the production of sulphuric acid is defined by [OH] and [SO₂]. It is mainly removed by condensation to existing particles in the atmosphere.

Boy et al. (2005) introduced a zero-dimensional chemistry model to calculate sulphuric acid and OH concentrations. The model was described and successfully verified against measured

sulphuric acid data in Hyytiälä. Petäjä et al. (2009) derived three proximity measures ("proxies") for the sulphuric acid concentrations by using EUCAARI (European Integrated project on Aerosol Cloud Climate and Air Quality Interactions) 2007 campaign data and found that measured concentrations correlated well with proxies derived as well as with detailed pseudo-steady state chemical model results. However, the authors recognized that the proxies might be site-specific and should be verified against measurements prior to utilization in other environments. In **Paper IV** proxies for sulphuric acid concentration were introduced and tested in six different environments. It was shown that the sulphuric acid concentration can be predicted with global solar radiation, SO₂ concentration, total particle surface area and relative humidity.

When the nucleation rates, aerosol particle concentrations and sulphuric acid concentrations are estimated with semi-empirical parameterizations like in Laaksonen et al. (2008a), Paasonen et al. (2010), **Paper III** and **Paper IV**, there are some issues to take into account. First, the parameterizations are only simplifications of the real atmospheric processes. The chemistry behind the modelled processes is very complicated and all affecting factors cannot be taken fully into account with current computing resources. For example, reactions in which SO₂ is oxidised into sulphuric acid are known but the involved species interact also with several other compounds in atmospheric conditions, thus making it difficult to estimate how much sulphuric acid is actually formed (Finlayson-Pitts and Pitts Jr., 2000; Laaksonen et al., 2008a; **Paper IV**). Second, the measurements, which are the basis of the parameterizations, contain significant uncertainties. For example, the uncertainty in sulphuric acid measurements, which is caused by differences in measurement procedures which in the worst case might lead up to 50% differences between the instruments used in different sites (Paasonen et al., 2010; **Paper IV**). In addition, SO₂ measurement instruments can have some 50 to 100 ppt uncertainty, which can easily make 50% uncertainty at the low concentration levels (Christian Plaß Dülmer, personal communication, 2011). Possible error sources in measurements are for example inlet lines, often related to RH, and occasional drift of the instruments, which are hard to identify since usually only one instrument is operated at the time. Formation rates of small particles (J) may be up to a factor two lower or higher than the estimated ones due to uncertainties in measurements and data analysis (Manninen et al., 2010). A significant source of uncertainty in estimating the magnitude of J -values is that the measured sample air is not the same all the time but the air masses passing

by the measurement sites are in constant movement and the air mass in large area is not necessarily homogenous. Since the statistical models are based on the measured data, all the uncertainties in measurement may also cause additional uncertainty into the models. Thus it is necessary to acknowledge the uncertainties and try to take them account in the analysis of the data.

2.3 Observed formation of new aerosol particles

Figure 3 illustrates an intense new particle formation event day in San Pietro Capofiume, Italy. A day is considered as a NPF event day when sudden appearance of nucleation mode particles is followed by strong particle growth towards bigger sizes. The figure shows that the particle formation and initial growth process, taking place below 3 nm, starts just after 10 am and the particles start to grow rapidly. After noon the particles have reached the potential size for CCN, which the authors assume to be 50 nm (Laaksonen et al, 2005, **Paper III**).

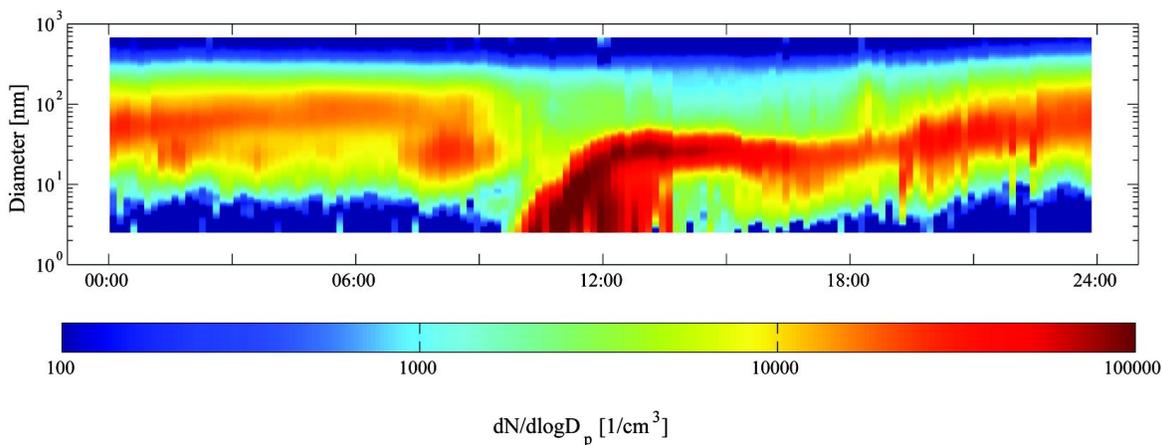


Figure 3. New particle formation event as measured with DMPS system on 25.3.2002 at the SPC station, Italy. The colour scale represents the number concentration of the particles. [Reprinted from Hamed et al., (2007). Reproduced by permission from the author]

NPF event days are classified into different groups with visual analysis from plots like that shown in Figure 3 (Mäkelä et al., 2000; Dal Maso et al., 2005). The categories depend on the intensity of the NPF event and consist of intensity groups 1, 2 and 3 for NPF events, referring to strong, intermediate, and weak events, respectively. If no NPF is observed, the day is classified

as a non-event day. A large number of days do not usually fulfil the criteria to be classified either as clear event or NE day and these are considered as unclassified days. The use of this classification method for our measurement data is introduced in Hamed et al. (2007). The results of the visual analysis are used in **Paper I** where discriminant analysis method was introduced for predicting the occurrence of the NPF event days.

Several different nucleation mechanisms have been proposed to act in the atmosphere. One of the most studied nucleation mechanisms is homogenous binary nucleation (Kulmala and Laaksonen, 1990; Vehkamäki et al., 2002). The binary nucleation theory of sulphuric acid and water is not sufficient to predict observed nucleation rates in boundary layer conditions (Kirkby et al., 2011), which means that some other compounds have to participate in atmospheric nucleation processes, for example ammonia, different aminium salts and other organic compounds (Ball et al., 1999; Hamed et al., 2007; Laaksonen et al., 2008b; **Paper II**), or even be the driving force for nucleation (Smith et al., 2010). One contributor in new particle formation is proposed to be ternary nucleation, which is driven by water, sulphuric acid and ammonia (Korhonen et al., 1999; Anttila et al., 2005; Merikanto et al., 2007). However, the ternary nucleation theory has not generally succeeded in reproducing observed particle formation rates. Recent studies (e.g. Kurtén et al., 2008) have shown that amines are able to enhance the addition of sulphuric acid to a cluster more efficiently than ammonia. In addition, Riipinen et al. (2007) found no significant correlation between ammonia and new particle formation and several studies have pointed out that ternary nucleation theory predicts 4-10 H₂SO₄ molecules in the critical cluster while according to observations the correct number is 1-2 (Weber et al., 1996; Kulmala et al., 2006; Laaksonen et al., 2008a; Kuang et al., 2008). Atmospheric ions have proposed to have an important role in nucleation (Arnold, 1980; Eisele et al., 2006). A modelling study by Yu et al. (2008) even proposed that ion induced nucleation is the main source of newly formed particles in the atmosphere. However, studies made e.g. by Manninen et al. (2010 and references therein) reported that the contribution of ions varies from 1% to 27% of total new particle formation, depending on the place and time of the measurements.

2.4 Growth of particles to Cloud Condensation Nuclei (CCN)

Dynamics of freshly-nucleated particles depends in a complicated way on the interplay between their formation rate, their condensation growth and their loss by coagulation and deposition to surfaces (Kerminen et al., 2004). When the newly formed particles grow to sizes larger than 50-100nm in diameter they start to become climatically important. Particles in this size range and larger can scatter sunlight back to space, thus having effect on the albedo of the Earth, and even more importantly 50 nm is considered to be the minimum potential size for CCN (Laaksonen et al., 2005; **Paper III**).

The newly formed particles grow mainly due to condensation of low-volatility vapours onto particle surfaces, nuclei self-coagulation, and possibly by heterogeneous reactions taking place in or on particles (Stolzenburg et al., 2005). Several methods relying on particle number size distribution measurements have been developed for determining nuclei growth rates from atmospheric observations (Leppä et al., 2011). If information about the time evolution of the mean diameter D_p of the new particle (nucleation) mode is available, the growth rate of particles can be calculated with formula given by (Seinfeld and Pandis, 2006)

$$GR = \frac{dD_p}{dt} \quad (10)$$

If suitable data on particle number size distributions to calculate growth rate are not available, GR at the size D can be estimated with information on the time evolution of nucleating precursors and the total nuclei number concentration (Kulmala et al., 2004)

$$GR = \frac{D - D_{crit}}{t_{prod} - t_{nuc}} \quad (11)$$

where D_{crit} is the critical size of nucleated particles, t_{prod} is the time at which particle formation is observed to begin, and t_{nuc} is the time nucleation is assumed to have begun. The GR is dependent on the vapour pressure of the condensable vapour. The quantity D_{crit} is obtained from a nucleation theory. t_{nuc} is typically set equal to the time at which the concentration of the assumed nucleating precursors, such as sulphuric acid, is observed to increase.

Condensational growth of newly formed particles competes with losses due to coagulation and deposition. The more rapid the condensation growth is the larger is the fraction of nucleated particles that will survive (Kulmala, 2003). The term describing how rapidly molecules will

condense onto pre-existing aerosols is condensation sink (CS), which is given by (Pirjola et al., 1998; Dal Maso et al., 2002),

$$CS = 2\pi D_d \int_0^{\infty} D_p \beta_m(D_p) n(D_p) dD_p = 2\pi D \sum_i \beta_i D_{pi} N_i \quad (12)$$

where D_{pi} is a measure of the diameter of the particle in the size class i and N_i is the particle number concentration in the respective size class. D_d is the diffusion coefficient of the condensing vapour, and β_m the correction factor for the transition and free molecular regimes (Fuchs and Sutugin, 1970).

After nucleating at diameters of about 1 nm in the atmosphere, a newly formed particle may grow by above mentioned mechanisms, and eventually reach particle sizes where it may act as a cloud condensation nucleus. A large fraction of the CCN sized particles are from primary sources such as sea salt particles originating from sea spray, or sulphate or carbonaceous particles formed from anthropogenic emissions (e.g. Merikanto et al., 2009; Korhonen et al., 2010; **Paper II**). Furthermore, atmospheric measurements and modelling studies have shown that newly formed particles can affect concentrations of CCN (Ghan et al., 2001; Lihavainen et al., 2003; Kerminen et al., 2005; Spracklen et al., 2006, 2008; **Paper II**) and the results are confirmed by theoretical frameworks (Pierce and Adams, 2007; Kuang et al., 2008). Lihavainen et al. (2003) concluded that in environments like Pallas in Northern Finland, where new particle formation events are not very common and the air is relatively clean, atmospheric new particle formation is likely to contribute only a minor amount to the overall CCN budget. Laaksonen et al. (2005) suggested that new particle formation might be an important source of CCN even in polluted environments where strong primary particle emissions are assumed to govern CCN formation. In SPC, which was the measurement site used in the study, NPF events are quite frequent and intense and the growth from nucleation to CCN-size may take only few hours if enough condensable vapours are available (Hamed et al., 2007). In **Paper II** in turn it was seen that while anthropogenic pollution levels in Melpitz, Germany decreased, growth rate of particles increased, which increased CCN production. This may be due to the fact that the concentration of secondary organic aerosol (SOA) did not decrease thus larger amount of condensable material was available for smaller number of particles. In addition, especially summers in the period 2003-2006 were warmer than in 1996-1997, which causes more biogenic VOC emissions, which

in turn lead to larger amount of condensable organic matter. Merikanto et al. (2009) suggested that 45% of global low-level cloud CCN (at 0.2% supersaturation) are secondary aerosol derived from nucleation and even over the most polluted continental regions with high primary emissions, such as in South-East Asia, nucleation accounts for 19 % of CCN.

2.5 Modelling the new particle formation and growth processes

The climate and atmosphere in general are investigated with global scale computer models. Most global models require significant computational resources and all of the complicated processes involved in atmospheric chemistry cannot be fully described in the large scale models. Taking account of aerosol processes in the models is a compromise between the accurate description of microphysical processes and computational efficiency, and often requires simplified parameterisations of these processes (Dusek et al., 2006; Kokkola et al., 2009 and references therein). The computational time can be significantly reduced with simplified parameterisations of the size distribution evolution and chemical composition of aerosol populations (Kokkola et al., 2009). Usually, when the aerosol particle size distribution is modelled, either modal approach (Whitby and McMurry, 1997) or sectional approach, employing size bins to the data (Gelbard et al., 1980), is used. However, Korhonen et al. (2003) found out that these methods are not very well suited for predicting new particle formation and growth to CCN sizes. A substantial part of the work presented in this thesis has focused on finding parameterizations that can be used in predicting new particle formation events (**Paper I**), particle concentration at a certain size (**Paper III**) and the concentration of sulphuric acid (**Paper IV**), the latter of which is the most important chemical compound involved in nucleation of aerosol particles in the atmosphere. With the estimation method of probability of a NPF event day we do not need visual analysis for occurrence of NPF events but the calculated probability can be used instead. By parameterizing the concentration of certain sized particles we can eliminate computationally expensive calculations of particles of smaller than the selected size, which was 50 nm in diameter in **Paper III**. Sulphuric acid parameterization requires only meteorological and trace gas parameters present in the models and thus detailed modelling of chemistry underlying the formation of sulphuric acid is not needed.

3 Methods

A variety of measurements conducted in several locations, techniques and data analysis methods were used in the work done for this thesis. The following sections introduce the experimental methodology presented in the articles.

3.1 Measurements

Data for studies included in this thesis are collected from different locations. Using data from varying environments makes it possible to generalize the results of the statistical analyses for areas outside of the measurement sites. For example, in **Paper I** discriminant analysis was used to predict NPF event days in SPC and the parameterization was updated in **Paper III** in order to introduce more generalized prediction method. In **Paper III** parameterization for concentration of 50nm particles was derived using three different datasets, obtained from differing environmental conditions. The parameterization was tested with Global Model of Aerosol Processes (GLOMAP) against different datasets collected several locations around the Europe during EUCAARI campaign. It was shown that the parameterization works adequately in different parts of continental Europe.

The measurement sites that are the focus of this work are Hyttiälä, Finland; Melpitz and Hohenpeissenberg in Germany; San Pietro Capofiume (SPC) in Po Valley, Italy; and Atlanta, Georgia and Niwot Ridge in Boulder, Colorado, USA. Locations of the sites are shown in Figure 4. The sites can be classified by the air pollution level at the measurement sites as follows: Hyttiälä and Niwot ridge are clean air stations, Melpitz and Hohenpeissenberg are semi-polluted and SPC and Atlanta are highly polluted environments.

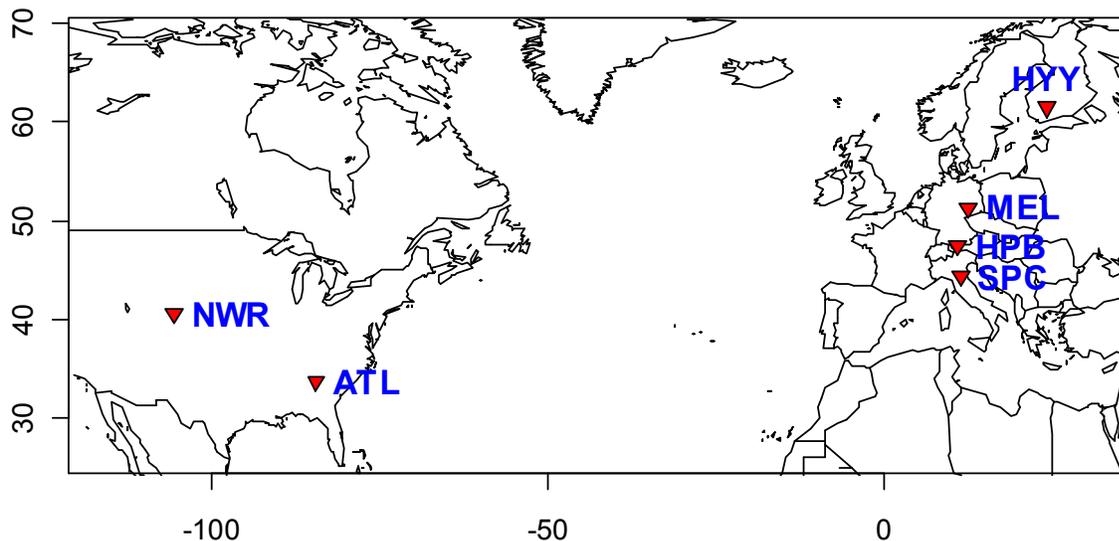


Figure 4. Locations of the six observation sites.

Hyttiälä is located in Southern Finland, in a rural environment surrounded by large areas of pine forest. The measurements are obtained at the SMEAR II (Station for measuring Forest Ecosystem-Atmosphere Relations), more information about the continuous measurements and the infrastructure can be found in Hari and Kulmala (2005). Anthropogenic pollution levels in Hyttiälä are low but surrounding boreal forest environment produces large amounts of organic matter which participates in new particle formation thus giving the site its peculiar attributes. In addition, Hyttiälä is a measurement site where frequent measurement campaigns take place, particularly QUEST (Boy et al., 2005) and EUCAARI (Petäjä et al., 2009; Kulmala et al., 2009) campaigns, which were the sources of sulphuric acid measurement data used in **Paper IV**.

Melpitz is a rural atmospheric research site in eastern Germany, operated by the Leibniz-Institute for Tropospheric Research (IfT). The site is situated on flat meadow grasslands surrounded by agricultural pastures and forests. Even though Melpitz is a rural observation site, the levels of anthropogenic pollution such as sulphur dioxide are higher than, for instance, Hyttiälä or at the Hohenpeissenberg site (Jaatinen et al., 2009; Hamed et al., 2010). For more information on atmospheric measurements at Melpitz including the climatology of particle and trace gas concentrations, see Engler et al. (2007), Birmili et al. (2008), Spindler et al. (2010) and **Papers II-IV**.

The Meteorological Observatory Hohenpeissenberg (HPB) is a GAW (Global Atmosphere Watch) site operated by the German Weather Service (DWD). It is situated in rural southern Germany about 30 km north of the Alpine mountain ridge. The observatory stands on top of Hohenpeissenberg Mountain at an altitude of 985 m a.s.l. and about 300m above the surrounding countryside. At night, HPB usually resides above the nocturnal surface layer inversion. In winter time, HPB may even reside above the daytime boundary layer. The surroundings of the mountain are mainly meadows and forests. For more information on previous aerosol and trace gas measurements at HPB, see Birmili et al. (2003), Paasonen et al. (2009) and **Papers III-IV**.

San Pietro Capofiume is located in Northern Italy, in a flat rural area in the eastern part of the Po Valley (Hamed et al., 2007; **Papers I, III, IV**). The Po Valley is the largest industrial, commercial and agricultural area in Italy, with a high population density and substantial anthropogenic gaseous and particulate emissions from diffuse sources such as industry, domestic heating and traffic. The station itself is in a sparsely inhabited area open to Adriatic Sea to the east side, but enclosed by densely populated areas on its southern, western and northern sides. The distance to the closest cities, Bologna and Ferrara, is about 40 km. There are power plants and industrial areas along the Po River and close to the harbours of Venice and Ravenna. In addition, surrounding mountains protect the valley from winds that might otherwise disperse the pollution plumes.

Niwot Ridge (NWR) is a forested station located on an east–west oriented ridge in the Front Range of the Rocky Mountains approximately 35 km west of Boulder, Colorado, USA, with the entire study site lying above 3000 m elevation. The site sits in a broad saddle bounded by low rounded hills and is flanked by an alpine tundra ecosystem. Winds are typically westerly at night (downslope drainage) bringing relatively clean air from the continental divide; whereas daytime heating creates easterly (upslope) flow, bringing somewhat more polluted air from the Denver-Boulder metropolitan area (Boy et al., 2008; **Paper IV**).

Atlanta, Georgia, USA, is an urban site where high relative humidity (RH) in the morning may influence sulphuric acid measurements (Fred Eisele, personal communication, 2010). Measurements were made during the 2002 Aerosol Nucleation and Real-time Characterization Experiment (ANARChE) at Jefferson Street Station (JST), which is located about 4 km

northwest of downtown Atlanta and about 9 km southeast of a coal-fired power station, the latter providing a rich source of H_2SO_4 (McMurry et al., 2005; **Paper IV**).

Some key variables for the different sites during the measurements introduced in **Paper IV** are compared in Table 1. Atlanta is the most polluted site, with the highest SO_2 concentrations (median 1.54 ppb) and condensation sink (CS, median $1.51 \times 10^{-2} \text{ s}^{-1}$), whereas the cleanest sites are NWR and Hyytiälä where median values of both pollution markers are almost one order of magnitude lower than in Atlanta. An example of uncertainty in the measurements was given in Melpitz, where SO_2 measurements between 8.5.–27.5.2008 had to be removed from the analysis due to instrumental failure. The highest magnitudes of global radiation were measured in SPC, with a median value of 376 W m^{-2} , whereas Hyytiälä 2003 campaign was early in the spring so the median radiation was only 90 W m^{-2} . RH was highest in Melpitz (median 74%) and lowest in NWR (median 52%). Ozone concentrations were highest in Niwot Ridge (median 56.3 ppb) and lowest in Atlanta (median 30.8 ppb). Sulphuric acid concentrations were highest in Melpitz (median $2.94 \times 10^6 \text{ molec cm}^{-3}$) and lowest in Hyytiälä 2007 (median $1.86 \times 10^5 \text{ molec cm}^{-3}$). The uncertainty in sulphuric acid measurements, caused by differences in measurement procedures which in the worst case might lead up to 50% differences between the instruments used in different sites (Paasonen et al., 2010), has to be taken into account when comparing the sulphuric acid concentrations. Note that standard deviations of $[\text{H}_2\text{SO}_4]$ and $[\text{SO}_2]$ are quite large. This is due to diurnal variation of $[\text{H}_2\text{SO}_4]$ and occasional pollution events, either from local sources or from long range transport, which cause high peaks in $[\text{SO}_2]$.

Table 1. Mean, median and standard deviation (sd) of some key variables measured in different locations during campaigns introduced in **Paper IV**. Detection limits for $[\text{H}_2\text{SO}_4]$ and $[\text{SO}_2]$ measurements are 10^4 molec cm^{-3} and 0.1 ppb, respectively. In global radiation entries, only times when it was not completely dark are counted.

		SPC	Melpitz	Hyytiälä 2003	Hyytiälä 2007	NWR	Atlanta	HPB
$[\text{H}_2\text{SO}_4]$ 10^6 molec cm^{-3} ($>10^4$)	mean	5.40	6.43	1.42	0.43	1.83	12.9	1.70
	median	2.40	2.94	0.57	0.18	1.40	2.85	0.61
	sd	6.50	8.08	1.99	0.67	1.94	31.2	2.85
$[\text{SO}_2]$ ppb (>0.1)	mean	0.35	1.00	0.51	0.32	0.23	3.40	0.40
	median	0.24	0.59	0.34	0.21	0.18	1.54	0.26
	sd	0.28	0.98	0.50	0.32	0.17	7.16	0.46
CS 10^{-3} s^{-1}	mean	6.55	13.3	1.99	4.15	4.21	15.9	4.87
	median	6.33	12.0	1.42	3.41	3.90	15.1	4.27
	sd	2.61	5.57	1.48	3.33	1.90	6.37	6.79
Radiation W m^{-2} (>0)	mean	405	347	158	221	333	242	277
	median	376	296	90	157	207	89	187
	sd	312	280	174	213	327	288	261
RH %	mean	67	73	67	60	56	66	75
	median	67	74	66	60	52	67	79
	sd	17	21	17	20	25	19	29
$[\text{O}_3]$ ppb	mean	37.8	35.8	39.2	35.2	56.6	36.7	41.9
	median	38.0	37.0	39.0	35.6	56.3	30.8	40.8
	sd	21.9	17.9	5.9	7.9	10.1	29.6	13.3

Although the different measurement instruments are not of specific interest in this study, one has to know something about the measurements in order to take into account the uncertainties caused by the instrumentation. The most important instrument for all of the **Papers I-IV** was the Differential Mobility Particle Sizer (DMPS), which is used for atmospheric aerosol size distribution measurements (see e.g. Wiedensohler et al., 1994). The DMPS consists of a Differential Mobility Analyzer (DMA) and a Condensation Particle Counter (CPC), which is a widely-used instrument for detecting the number concentration of aerosol particles too small to be observed with optical techniques (Stolzenburg and McMurry, 1991; McMurry, 2000).

The second most important instrument which needs to be introduced was Chemical Ionization Mass Spectrometry (CIMS), which is used to measure sulphuric acid concentrations in the atmosphere (**Paper IV**). The first measurements of atmospheric gas-phase sulphuric acid have been made on stratospheric balloons and research air craft by MPIK-Heidelberg, using PACIMS

(Passive Chemical Ionization Mass Spectrometry), a novel method developed by the MPIK-group (Arnold and Fabian, 1980). Later, measurements in lower troposphere air have been made using active CIMS (Eisele and Tanner, 1993; Berresheim et al., 2000). The challenges associated with sulphuric acid measurements combined with subtle differences between CIMS instruments have resulted in variations in the measurement results. Paasonen et al. (2010) noted that there might in the worst case be up to 50% differences between the instruments, which has been taken account in analyzing the data in **Paper IV**.

The measurements with DMPS and CIMS systems were supported by measurements of various meteorological and gas phase variables. More details on the instrumentation used in these measurements are described in Hamed et al. (2007) and references therein. In addition, as already mentioned in Section 2.1.1, significant uncertainties are also known to exist in SO₂ measurements when the concentrations are low, which was also taken account in all papers.

3.2 Statistical methodology

Various statistical analysis methods were used to analyze the datasets included in this thesis. Descriptive statistics methods were the same in every paper: average values and variances describing the measured variables, correlations to find the interdependencies of the different variables, simple linear regression and variance analyses to test dependencies between different phenomena. Also, figures such as scatter plots, histograms and box plots for visual analysis of the data were shown. In **Papers I, III and IV** more advanced multivariate methodologies were used and they are described in more detail in the following sections. One of the greatest advantages of multivariate models is that the interpretation of estimates of single parameters is easy and the results are more valid than in single variable analysis.

3.2.1 Discriminant analysis

Discriminant analysis (DA) is a multivariate statistical analysis method, which is commonly used to build a predictive or descriptive model of group discrimination based on observed predictor variables and to classify observations into different groups. DA was used in **Paper I** for finding factors affecting the occurrence of new particle formation and to build a model that can be used

in the prediction of events. Analysis results of discriminant analysis for **Paper I** were also used while interpreting results of Hamed et al. (2007).

The purpose of the classical, parametric form of discriminant analysis (Fisher, 1936) is to express one dependent variable as a linear combination of other features or measurements. DA is closely related to variance analysis and regression analysis, which also attempts to express one dependent variable as a linear combination of other measurements with the only exception that in DA the dependent variable is a categorical and not continuous numerical variable as in the other two methods. A categorical variable can be for example NPF event occurrence, such as in **Paper I**, where the measurement days we categorised as NPF event days, nonevent days and unclassified days. When the distribution in each group cannot be assumed multivariate normal, which is the case with atmospheric aerosol measurements, non-parametric discriminant methods must be used. Non-parametric methods are based on group-specific probability densities and they are used to produce a classification criterion based on those probabilities. In addition, the non-parametric method is more robust for multicollinearity (i.e. when some variables measure partly the same effect thus correlating strongly between each other), which might occur in the analysis of atmospheric data. Multicollinearity may cause significant errors in the estimation of the coefficients of the more sensitive statistical models, such as parametric form of the discriminant analysis. In **Paper I** we used non-parametric Epanechnikov kernel method (Epanechnikov, 1969) to find factors which classify the days as new particle formation event days or nonevent days.

Estimation of discriminant function is usually based on Mahalanobis squared distance (Mahalanobis, 1936) of the observation x from the mean \bar{x}_i of the sample i , given by (Morrison, 2005)

$$D_i^2 = (x - \bar{x}_i)'S^{-1}(x - \bar{x}_i) \quad (13)$$

The Mahalanobis distance is used to define probability that the observation belongs to each data group. In the most simple case observation x is assigned to group i if $D_i^2 = \min\{D_1^2, \dots, D_k^2\}$ where k is the number of groups.

In kernel estimation methods, like the one used in Paper I, the classification of observations x is based on the estimated group specific densities of the data. Observation x is classified into group

u if the conditional probability $p(u|x)=\max p(u|x) = \max\{p(j|x)\}$, where $j=1,\dots,k$. More details on the analysis can be found in **Paper I**.

In **Paper I** discriminant analysis was used to predict NPF event days in SPC. The parameters needed for the prediction were solar radiation, ozone concentration and relative humidity. The same method was used in **Paper III** in order to find estimates for probabilities of the occurrence of NPF event days in three different datasets. It appeared that when the method was applied to data from Melpitz and Hohenpeissenberg in **Paper III** the parameters used in SPC were inadequate to represent the different conditions in the other two locations. That study demonstrated that concentrations of SO_2 and NO_2 and air temperature had to be taken into account in the analysis in order to get a more general parameterization for the prediction method. The best predictor models for each site are shown in Table 2 with prediction errors. Finally, we combined all the data from different sites in order to find the best discriminates for the NPF events for the full dataset. We found out that the best predictor sets for the individual sites are subgroups of the best predictors of the combined data, i.e., in specific sites it is possible to get equally good classification with fewer predictors. The differences between the prediction errors of the different datasets are relatively small. However, for more general analysis the parameter set defined from the combined data should be used.

Table 2. Results from Discriminant analysis as found in **Paper I** and **Paper III**

Station	Predictors	Classification error	Missed events	predicting non-event to event
SPC	RH, O_3 , Radiation	13.4 %	1.9 %	4.9 %
Melpitz	RH, O_3 , SO_2 , Radiation, Temperature	6.7 %	7.4 %	0.7 %
Hohenpeissenberg	RH, O_3 , SO_2 , Radiation, NO_2	6.1 %	11.7 %	2.7%
Combined data	RH, O_3 , SO_2 , NO_2 , Radiation, Temperature	7.7 %	8.4 %	1.5 %

3.2.2 Mixed Models

While constructing parameterization for 50 nm sized particles in **Paper III**, the complexity of the physical and chemical processes affecting the data was taken account with a mixed model structure. Mixed model is a generic term for a model family that combines several different kinds

of models used in multivariate analysis when the data are not fulfilling the standard independency and homogeneity assumptions. These models are usually generalizations of different linear or nonlinear regression models, variance analysis models or repeated measures analysis models. A mixed effects model is a combination of fixed and random effects models. Matrix formulations of the models are given in Table 3.

Table 3. Matrix type formulations of different models.

	Linear	Generalized linear	nonlinear
Fixed effects model	$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$	$g(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$	$\mathbf{y} = f(\mathbf{X}\boldsymbol{\beta}) + \boldsymbol{\varepsilon}$
Random effects model	$\mathbf{y} = \boldsymbol{\mu} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon}$	$g(\mathbf{y}) = \boldsymbol{\mu} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon}$	$\mathbf{y} = \boldsymbol{\mu} + f(\mathbf{Z}\mathbf{u}) + \boldsymbol{\varepsilon}$
Mixed model	$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon}$	$g(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon}$	$\mathbf{y} = f(\mathbf{X}\boldsymbol{\beta}, \mathbf{Z}\mathbf{u}) + \boldsymbol{\varepsilon}$

In Fixed effects models \mathbf{y} is the vector of measurements of the studied variable, \mathbf{X} is a known model matrix containing the observations of predictor variables x_i , $\boldsymbol{\beta}$ is a vector of parameter estimates b_i estimated from the data and $\boldsymbol{\varepsilon}$, is the vector of the residuals of the model. In the simplest case of a Fixed effects model, y is a linear function of predictors x_i . In generalized linear model $g(\bullet)$ refers to so-called link function, which relates the linear predictors x_i to the expected value of \mathbf{y} (McCullagh and Nelder, 1989). The link function is needed when the dependent variable is not normally distributed in nature, as the atmospheric measurement variables seldom are (Schultz et al., 2007; **Paper III**). The most typical link functions are logarithm, logit, reciprocal and various non-canonical links. If the dependency of the predictor variables and the study variable cannot be normalized with a simple link function, the analysis is made with nonlinear model, where f is commonly some p-dimensional function or exponent-based function or combination of the two. More details on nonlinear analysis are given in section 3.2.3.

In a Random effects model $\boldsymbol{\mu}$ is the mean value of \mathbf{y} , \mathbf{Z} is a model matrix such as \mathbf{X} and \mathbf{u} is the vector of random effects that occur in the data vector \mathbf{y} . Vector \mathbf{u} defined by $\mathbf{u} \sim (\mathbf{0}, \mathbf{D})$, meaning that expected value $E[\mathbf{u}] = \mathbf{0}$ and variance $\text{var}(\mathbf{u}) = \mathbf{D}$. This means that in Random effects model \mathbf{y} varies around $\boldsymbol{\mu}$ with variance defined by \mathbf{u} . Notations related to the Random and Mixed effects models are based on McCulloch and Searle (2001).

A Mixed model is a combination of fixed and random effects. In linear models, fixed effects are used for modelling the mean of \mathbf{y} and random effects are determining the variance-covariance structure of the data. Successful modelling of variances and covariances of the observations

provides valid statistical inference for the fixed effects β of the mixed model. In contrast to general linear models, the error terms ε can be correlated, the covariance matrix of the residuals is denoted by \mathbf{R} , and the correlation structure can be defined to fit the data in order to reduce bias in the estimates of β and \mathbf{u} . In mixed models \mathbf{u} is a vector with a q -dimensional normal distribution with zero expectation and covariance matrix denoted by \mathbf{G} . From this it follows that the distribution of observations y can be postulated as a normal distribution with the expectation of $\mathbf{X}\beta$ and covariance matrix \mathbf{V} , which is given by $\mathbf{V} = \mathbf{ZGZ} + \mathbf{R}$. Details on the estimation of the parameters of the mixed models can be found in Patterson & Thompson (1971), Lindstrom & Bates (1988) and in McCulloch & Searle (2001).

One of the greatest advantages of a multivariate model is that when all parameters are in the same model the interpretation of estimates of single parameters is easy and the results are more valid than in single variable analysis. For example, bias caused by yearly variation can be cleaned from the other variables in the model. A common way to make the analysis is to use a stepwise manner, i.e. the parameters are added and removed to and from the model according to their statistical significance and the total increase or decrease of some measure of explanation capability of the model (e.g. log-likelihood or coefficient of determination) while all the other parameters are kept constant.

In **Paper III** a predictive model for number concentration of 50 nm particles in diameter was constructed with mixed model structure. We first performed tests for the data from each of the three measurement sites separately, but as the parameters for the three datasets were very close to each other, we could merge the data and find the best predictors for the combined dataset. The best predictive model was attained with a combination of relative humidity, new particle formation event probability, temperature, condensation sink and concentrations of SO_2 , NO_2 and ozone. The seasonal variation was also taken into account in the mixed model structure. The final model, giving the best prediction ability in the combined data can be written as

$$\ln(dN / d \log Dp|_{50}) = (\beta_0 + u_m) + (\beta_1 + v_{1m})RH + (\beta_2 + v_{2m})[SO_2] + (\beta_3 + v_{3m})[NO_2] \\ + (\beta_4 + v_{4h})[O_3] + \beta_5 \cdot PrNE + \beta_6 \cdot Temp + v_5 \cdot CS^{-1} + \varepsilon_{jt} \quad (14)$$

where β_0 is the fixed intercept term, β_i are the fixed slopes, u_m are month-specific random intercepts and $v_1 - v_4$ are the random month-specific (m) or hour-specific (h) slopes. Here m lies

in range [1-12] and h in range [0-23]. v_5 is location specific random effect for taking account the condensation sink in the estimation of other parameters. The other variables used in the model are: relative humidity (RH, %), $[\text{SO}_2]$ ($\mu\text{g}/\text{m}^3$), $[\text{NO}_2]$ ($\mu\text{g}/\text{m}^3$), $[\text{O}_3]$ ($\mu\text{g}/\text{m}^3$), the probability that the day is not a NPF event day (PrNE), and air temperature (K).

The parameterization derived in **Paper III** with mixed models was validated with GLOMAP model. Figure 5 compares the model-predicted potential CCN concentration ($D_p > 50$ nm) against measured April mean data at Melpitz, SPC, Hohenpeissenberg, and 15 other sites. For the three sites analysed in **Paper III**, we presented averages of several years of measurement data (i.e., multi-annual averages), since the analysed measurement periods at the three sites did not overlap. The data for the other 15 sites is from the European Integrated project on Aerosol Cloud Climate and Air Quality Interactions (EUCAARI) during April 2008 and/or 2009, and were chosen because of their comprehensiveness. Note that since the model data is only available for April 2000 (due to computational expense of a global model) one can expect only a rough agreement between the model and measurements. However, we wanted to include the EUCAARI data in order to demonstrate that the parameterization derived with a mixed model gives reasonable results also outside the geographical domain for which it was derived.

While this preliminary test of the parameterization against observations is incomplete in the way described above, it does indicate that the derived parameterization has potential to describe CCN formation at very different environments from Arctic to polluted continental areas. These results give confidence to apply the mixed model framework to measurements from other sites in order to further improve the derived parameterization.

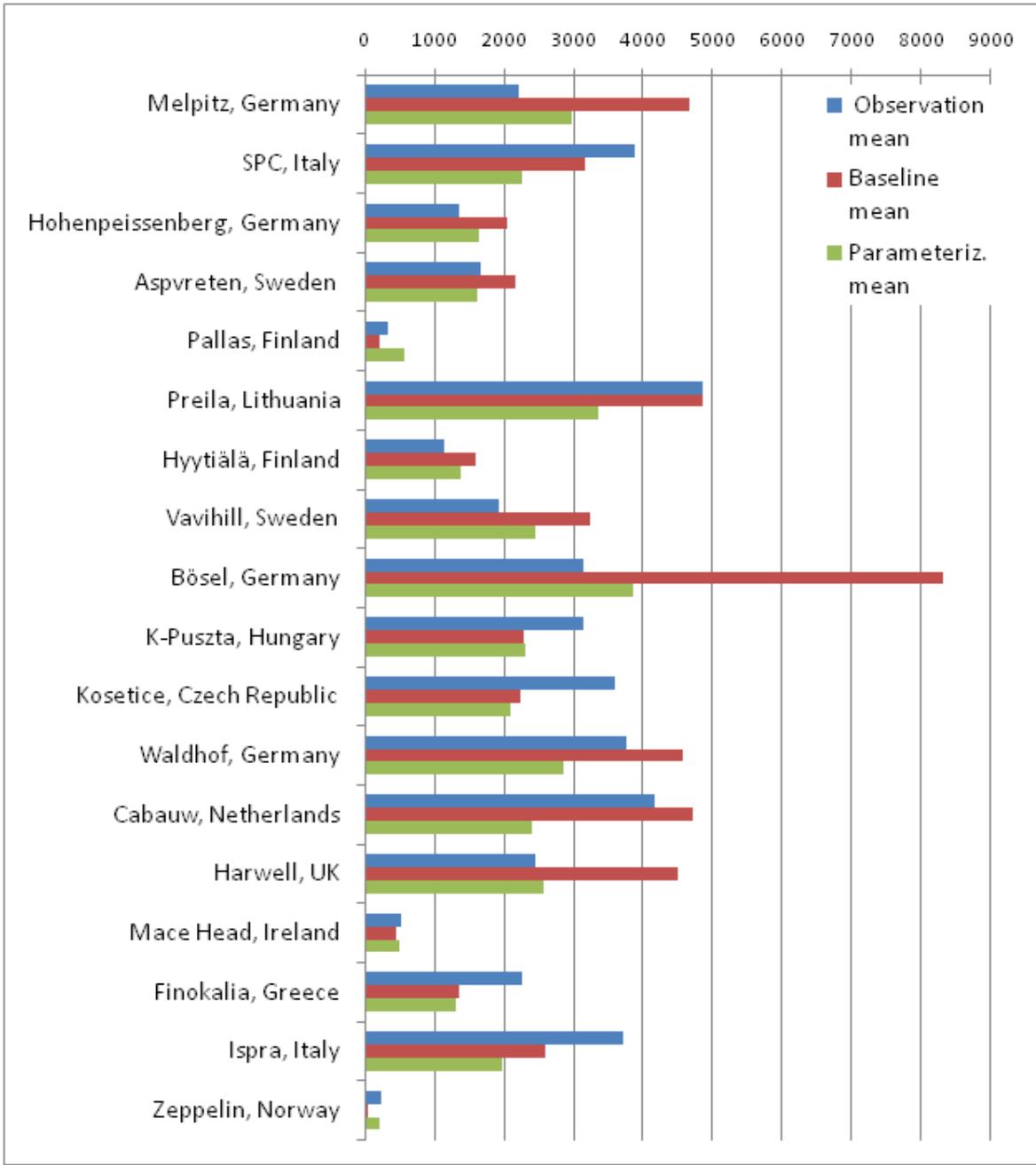


Figure 5. Mean observed and predicted potential CCN ($D_p > 50$ nm) concentrations at the 3 measurement sites (**Paper III**) and at 15 EUCAARI sites. Units are cm^{-3} . "Baseline mean" is the result of the standard set-up of the GLOMAP and "Parameteriz. Mean" is the result gained with the parameterization derived in **Paper III**.

3.2.3 Nonlinear regression analysis

A nonlinear regression analysis was used in **Paper IV** in order to find applicable predictive model (proximate measure, "proxy") for atmospheric sulphuric acid concentration. The basic idea of nonlinear regression is to relate a response variable y to predictor variables x_i , similarly as in linear analysis, with the only difference that the prediction equation depends nonlinearly on one or more unknown parameters (Smyth, 2002). Nonlinear regression is usually needed when there are physical reasons for believing that the relationship between the response and the predictors follows a particular functional form. The general form of nonlinear regression model is given by

$$y_i = f(x_i, \beta) + \varepsilon_i \quad (15)$$

where y_i are the measured response observations, f is a known nonlinear function of the measured predictor variables x_i , β are the estimated parameters of the model, ε_i are the random residuals of the model which are usually assumed to be uncorrelated with mean zero and constant variance. The definition of nonlinearity relates to the unknown parameters and not to the relationship between the covariates and the response. For example polynomial regression models $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots$ are considered to be linear because the relationship between parameters β_j is linear and the model can be estimated by using linear regression methods. To quantify this, take derivatives of y with respect to the parameters β_0 , β_1 , and β_2 : $dy/d\beta_0 = 1$, $dy/d\beta_1 = x$, $dy/d\beta_2 = x^2$. Since none of these derivatives depends on a model parameter, the model is linear. For example the best proxy function found for sulphuric acid concentration presented in **Paper IV** was in the form

$$[H_2SO_4] = a \cdot k \cdot Radiation \cdot [SO_2]^c \cdot (CS \cdot RH)^d \quad (16)$$

where a , c and d are unknown parameters and thus the function fulfils the nonlinearity assumption.

Nonlinear regression models are fitted with a nonlinear least squares fitting procedure (Bates and Watts, 1988; R Development Core Team, 2010), which gives some of the same advantages (and disadvantages) that linear least squares regression has over other methods. A remarkable advantage is efficient use of data; nonlinear regression can produce good estimates of the unknown parameters in the model with relatively small data sets. One of the greatest advantages

of the nonlinear approach is that the non-equilibrium conditions of the modelled processes can be taken account in the analysis. In addition, the asymptotic behaviour of some processes and other features of physical processes can often be expressed more easily using nonlinear models than with simpler model types.

The major disadvantage of nonlinear least squares regression in comparison to simpler modelling techniques, like linear least squares, has been the computational cost of the iterative parameter estimation procedure. However, increasing computing power is reducing this disadvantage. With linear type functions the least squares estimates of the parameters can always be obtained analytically, while that is generally not the case with nonlinear models. The use of iterative procedures requires the user to provide starting values fairly close to the unknown parameters before the software can succeed in the optimization. The optimization procedure does not calculate the difference between local and global minimum values that defines the least squares estimates, so bad starting values may cause significant errors in the interpretation of the results. One of the disadvantages shared with the linear least squares procedure is a strong sensitivity to outliers. Presence of one or two outliers in the data can seriously affect the results of least squares analysis.

In **Paper IV** we found out that the connection between sulphuric acid concentration in the atmosphere and the variables used in our proxy was not linear. In order to find the optimal parameterization for the proxy, a nonlinear least squares fitting procedure was applied to the data, with various possible fit functions. The advantage of using the nonlinear approach is that the non-equilibrium conditions are taken into account by estimating individual power terms for proxy variables from the data. The proxy introduced in Eq. (16) was found to give the best approximate value for measured concentration as well in the campaign datasets used to construct the proxy and in separate, long term data recorded in Hohenpeissenberg. Figure 6 shows that the predictive ability of the derived proxy is good, correlation between observed and predicted values given by the proxy was 0.87, except in data points where *Radiation* or $[SO_2]$ values are small, i.e. the proxy value is small.

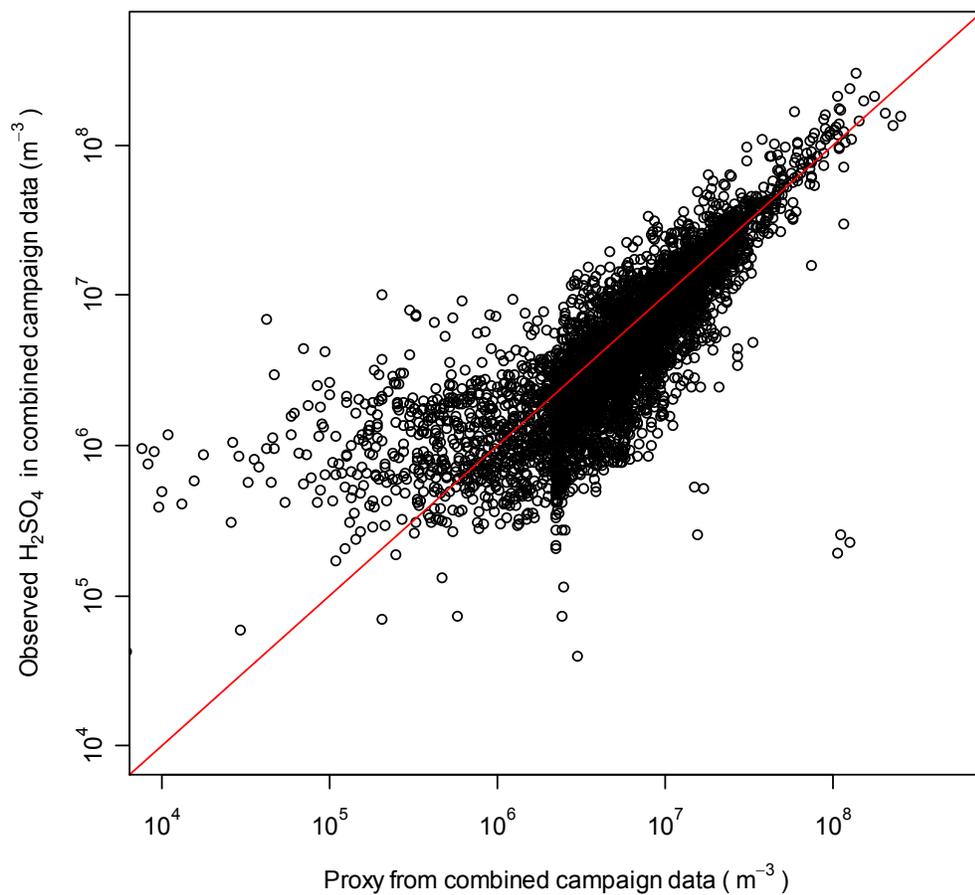


Figure 6. Observed $[H_2SO_4]$ in combined campaign data vs. predicted values given by Proxy in **Paper IV**. Diagonal line represents the one-to-one correspondence.

4 Review of papers and author's contribution

Paper I investigates the factors that affect the occurrence of new particle formation events in San Pietro Capofiume, Italy with more than three years continuous measurement data. The paper introduces a nonparametric discriminant analysis method, which is able to find applicable classification parameters from the measured trace gas and meteorological variables in order to predict the occurrence of a new particle formation event. The best classification result in the data from SPC was reached with the combination of relative humidity, ozone concentration and global radiation. RH appeared to have a preventing effect on the new particle formation whereas the effects of O₃ and radiation were found to have a positive effect. The concentrations of SO₂ and NO₂ also appeared to have a significant effect on the emergence of NPF events, but due to a great number of missing observations, we had to exclude them from the final analysis.

I performed the statistical analysis and wrote most of the paper.

Paper II reports the significant decrease in SO₂ concentration in Melpitz, Germany, between two measurement periods recorded in 1996-1997 and 2003-2006. The paper studies how the changed conditions affect new particle formation and the subsequent CCN production. Our analysis showed a statistically significant and concurrent drop in the frequency of new particle formation events between the two observation periods. Along with this, a decrease in the particle formation rates was observed. The results suggest a connection between the decreasing availability of anthropogenic SO₂ and the diminishing production of new particles. A simple proxy for sulphuric acid concentration was introduced, making use of the fact that H₂SO₄ is formed in the reaction between SO₂ and OH radicals, and the concentration of the latter can be correlated with intensity of solar radiation and that the condensation sink explains the losses of H₂SO₄. This gave a starting point for **Paper IV**.

I did the statistical calculations needed for the paper and participated in the writing.

Paper III introduced a parameterization for number concentration of 50 nm particles, N_{50} , in the atmosphere. Measurements of aerosol size distribution and different gas and meteorological parameters were made in three polluted sites in Central and Southern Europe: Po Valley, Italy, Melpitz and Hohenpeissenberg in Germany. The data were analysed to examine which of the

meteorological and trace gas variables affect the number concentration of Aitken (here $D_p = 50$ nm) mode particles, which can be considered the minimum potential CCN size. The variables were then used to construct a parameterization that can be used in global atmospheric models to predict the concentration of the smallest climatically active particles. The parameterization was tested with the GLOMAP model and the results indicate that the parameterization works well at least in boundary layer but it still does not solve the computational efficiency problems in free troposphere. The statistical parameterization also gives reasonable results outside the geographical domain for which it was derived. Equally, the use of the statistical model for N_{50} could bypass some of the current uncertainties in the theoretical description of the nucleation and growth process, particularly when predicting potential CCN concentration.

I did the statistical analysis, constructed the parameterization and was mostly responsible for the writing of the paper.

Paper IV presents a proximity measures ("proxies") for sulphuric acid concentration in the atmosphere. The study evaluates a body of experimental gas phase sulphuric acid concentrations, as measured by Chemical Ionization Mass Spectrometry (CIMS) during six intensive measurement campaigns and one long-term observational period. The objective of this study was to find a proxy for sulphuric acid that is valid in different atmospheric environments. Our most accurate and universal formulation of the sulphuric acid concentration proxy uses global solar radiation, SO_2 concentration, condensation sink and relative humidity as predictor variables. Interestingly, the role of the condensation sink in the proxy was only minor, since similarly accurate proxies could be constructed with global solar radiation and SO_2 concentration alone.

I constructed the proxies with statistical methods and wrote a major part of the paper.

5 Conclusions

In the work presented in this thesis we studied the use of different statistical methods in the analysis of aerosol measurement data. We first examined new particle formation, then the growth of newly formed and primary particles into size range where they become climatologically active, and finally the formation of sulphuric acid, which is one the most significant compounds relating related to particle formation and growth. The findings can be summarized and the main objectives in Section 1 were achieved as follows:

- 1) New particle formation is a frequent phenomenon in semi-polluted and polluted areas such as San Pietro Capofiume, Italy, Melpitz and Hohenpeissenber, Germany (**Papers I, II, III**). Conditions favouring new particle formation can be characterized with measured atmospheric variables.

Occurrence of NPF events can be predicted with meteorological and trace gas variables by using a model defined with discriminant analysis (**Paper I**). In **Paper II** it was seen that while the SO₂ concentration reduced by 65% the intensity of new particle formation diminished in Melpitz, Germany. Frequency of NPF events was lowered by 45% and the drop in formation rate was 68% between the two observation periods 1996-1997 and 2003-2006. This result is not surprising, since the NPF is driven by sulphuric acid, which is an oxidation product of SO₂.

- 2) The concentration of 50 nm sized particles can be predicted with in-situ meteorological and gas phase parameters, by using a multivariate mixed effects model.

The developed parameterization was a combination of measured relative humidity, temperature, condensation sink and concentrations of SO₂, NO₂ and ozone and new particle formation event probability. The NPF event probability was calculated with the model introduced in **Paper I** and updated in **Paper III**. However, the parameterization is not dependent on the method for calculating the NPF event probability but the probability can also be estimated with some other applicable methods. The seasonal variation was also taken into account in the mixed model structure (**Paper III**). The statistical model needed to describe, amongst others, the factors affecting the growth of newly formed aerosol particles (below 10 nm) to 50 nm size, but also sources of direct particle emissions in that size range. The parameterization is able to predict the concentration of 50 nm particles in continental areas but it is probably not valid in other regions.

The model was proven to be valid also in mountain areas at the height of the top of the daytime boundary layer. The validity of the model in the higher altitudes needs to be quantified. Over the oceans totally different processes control the aerosol and CCN concentration and more work needs to be done before particle concentrations can be approximated all over the globe. Possibility to use satellite measurement in the analysis of the concentration of small particles has to be studied.

3) New particle formation was found to be important source for CCN formation. NPF probability was found to be significant predictor for the number concentration of CCN-sized particles (**Paper III**).

The role of SO₂ concentration in CCN formation was examined in **Paper II**. The results indicate that despite the fact that a decreasing SO₂ concentration has decreased new particle formation, the growth rate of the formed particles has increased by 22%. This in turn has increased the CCN production by tens of percent. Since the reducing SO₂ concentration also reduces sulphuric acid concentration, it can be concluded that the particle growth is also driven by some other compound than sulphuric acid, most probably low volatility organics.

4) Decreasing sulphur dioxide concentration was suggested to decrease the sulphuric acid concentration in **Paper II** and a universal proxy for sulphuric acid concentration was introduced in **Paper IV**.

The proxy was based on the fact that sulphuric acid is formed in reaction between SO₂ and OH radicals and it is removed from the air by condensation onto existing particles. Interestingly, the role of the condensation sink in the proxy was only minor, which lead to the finding that accurate proxies could be constructed with global solar radiation and SO₂ concentration alone. This could be attributed to SO₂ being an indicator for anthropogenic pollution, including particulate and gaseous emissions which represent sinks for the OH radical that, in turn, is needed for the formation of sulphuric acid. The derived proxy was found to give good predictions of the sulphuric acid concentration in daytime, when the amount of radiation is sufficient. However, the concentration of sulphuric acid will most probably need to be approximated on nighttimes also and a new proxy has to be developed for that purpose. The study sites were all in continental areas, so applicability of the proxy in marine areas is not known. Two of the study sites were in

mountain areas, so the proxy is proven to be valid in higher altitudes. The proxy is expected to be time invariant, so it can be extrapolated over time, which is important e.g. for the climate models.

6 References

- Aalto, P., Hämeri, K., Becker, E., Weber, R., Salm, J., Mäkelä, J. M., Hoell, C., O'dowd, C. D., Karlsson, H., Hansson, H.-C., Väkevä, M., Koponen, I. K., Buzorius, G. and Kulmala, M.: Physical characterization of aerosol particles during nucleation events. *Tellus B*, 53: 344–358. doi: 10.1034/j.1600-0889.2001.530403.x, 2001.
- Albrecht, B. A., Aerosols, cloud microphysics, and fractional cloudiness, *Science*, 245, 1227–1230, 1989.
- Anttila, T, Vehkamäki H, Napari I & Kulmala M: Effect of ammonium bisulphate formation on atmospheric water-sulphuric acid-ammonia nucleation. *Boreal Env.Res.* 10:511–523, 2005.
- Arnold, F.: Multi-ion complexes in the stratosphere – implications for trace gases and aerosol, *Nature* 284, 610–611, 1980.
- Arnold, F. and Fabian, R.: First measurements of gas phase sulfuric acid in the stratosphere. *Nature* 282, 55, 1980.
- Ball, S. M., Hanson, D. R., Eisele, F. L., and McMurry, P. H.: Laboratory studies of particle nucleation: Initial results for H₂SO₄, H₂O, and NH₃ vapors. *Journal of Geophysical Research-Atmospheres*, 104(D19):23709–23718, 1999.
- Bates, D. M. and Watts, D. G.: *Nonlinear Regression Analysis and Its Applications*, Wiley, New York, 1988.
- Berresheim, H., Elste, T., Plass-Dülmer, C., Eisele, F. L., and Tanner, D. J.: Chemical ionization mass spectrometer for longterm measurements of atmospheric OH and H₂SO₄, *Int. J. Mass. Spectrom.*, 202, 91–109, 2000.
- Birmili, W., Berresheim, H., Plass-Dülmer, C., Elste, T., Gilge, S., Wiedensohler, A., and Uhrner, U.: The Hohenpeissenberg aerosol formation experiment (HAFEX): A long-term study including size-resolved aerosol, H₂SO₄, OH, and monoterpenes measurements, *Atmos. Chem. Phys.*, 3, 361–376, 2003
- Birmili, W., Schepanski, K, Ansmann, A, Spindler, G, Tegen, I, Wehner, B., Nowak, A., Reimer, E., Mattis, I., Müller, K., Brüggemann, E., Gnauk, T., Herrmann, H., Wiedensohler, A., Althausen, D., Schladitz, A., Tuch, T., and Löschau, G.: A case of extreme particulate

- matter concentrations over Central Europe caused by dust emitted over the southern Ukraine, *Atmos. Chem. Phys.*, 8, 997-1016, doi:10.5194/acp-8-997-2008, 2008.
- Boy, M., Kulmala, M., Ruuskanen, T. M., Pihlatie, M., Reissell, A., Aalto, P. P., Keronen, P., Dal Maso, M., Hellen, H., Hakola, H., Jansson, R., Hanke, M., and Arnold, F.: Sulphuric acid closure and contribution to nucleation mode particle growth, *Atmos. Chem. Phys.*, 5, 863-878, doi:10.5194/acp-5-863-2005, 2005.
- Boy, M., Karl, T., Turnipseed, A., Mauldin, R. L., Kosciuch, E., Greenberg, J., Rathbone, J., Smith, J., Held, A., Barsanti, K., Wehner, B., Bauer, S., Wiedensohler, A., Bonn, B., Kulmala, M., and Guenther, A.: New particle formation in the Front Range of the Colorado Rocky Mountains, *Atmos. Chem. Phys.*, 8, 1577-1590, doi:10.5194/acp-8-1577-2008, 2008.
- Clifford, S., Low Choy, S., Hussein, T., Mengersen, K., Morawska, L.: Using the Generalised Additive Model to model the particle number count of ultrafine particles. *Atmospheric Environment*, 45, 5934-5945, doi:10.1016/j.atmosenv.2011.05.004, 2011.
- Curtius, J.: Nucleation of atmospheric aerosol particles, *C. R. Physique* 7, 1027–1045, 2006
- Dal Maso, M., M. Kulmala, K. E. J. Lehtinen, J. M. Mäkelä, P. Aalto, and C. D. O'Dowd: Condensation and coagulation sinks and formation of nucleation mode particles in coastal and boreal forest boundary layers, *J. Geophys. Res.*, 107(D19), 8097, doi:10.1029/2001JD001053, 2002.
- Dal Maso, M., Kulmala, M., Riipinen, I., Wagner, R., Hussein, T., Aalto, P. P., and Lehtinen, K. E. J.: Formation and growth of fresh atmospheric aerosols: eight years of aerosol size distribution data from SMEAR II, Hyytiälä, Finland, *Boreal Environ. Res.*, 10, 323–336, 2005.
- Dusek, U., Frank, G. P., Hildebrandt, L., Curtius, J., Schneider, J., Walter, S., Chand, D., Drewnick, F., Hings, S., Jung, D., Borrmann, S., and Andreae, M. O.: Size Matters More Than Chemistry for Cloud-Nucleating Ability of Aerosol Particles, *Science*, 312, 1375–1378, 2006, doi:10.1126/science.1125261.

- Eisele, F. and Tanner, D.: Measurement of the gas phase concentration of H₂SO₄ and methane sulfonic acid and estimates of H₂SO₄ production and loss in the atmosphere, *J. Geophys. Res.*, 98, 9001–9010, 1993.
- Eisele, F. L., Lovejoy, E. R. , Kosciuch, E. , Moore, K. F., Mauldin III, R. L. , Smith, J. N., McMurry, P. H., and Iida, K.: Negative atmospheric ions and their potential role in ion-induced nucleation, *J. Geophys. Res.*, 111, D04305, doi:10.1029/2005JD006568, 2006.
- Engler, C., Rose, D., Wehner, B., Wiedensohler, A., Brüggemann, E., Gnauk, T., Spindler, G., Tuch, T., and Birmili, W.: Size distributions of non-volatile particle residuals (D_p<800 nm) at a rural site in Germany and relation to air mass origin, *Atmos. Chem. Phys.*, 7, 5785–5802, doi:10.5194/acp-7-5785-2007, 2007.
- Epanechnikov, V. A.: Nonparametric estimation of a multidimensional probability density. *Theory Probab. Appl.*, 14, 153–158, 1969
- Finlayson-Pitts, B. J. and J. N. Pitts Jr.: *Chemistry of the Upper and Lower Atmosphere: Theory, Experiments, and Applications*, Academic Press, San Diego, CA, USA, 298–299, 2000.
- Fisher, R. A.: The Use of Multiple Measurements in Taxonomic Problems, *Annals of Eugenics*, 7: 179-188, 1936.
- Fuchs, N. A. and Sutugin, A. G.: *Highly dispersed aerosols*. Ann Arbor Science Publishers, Ann Arbor, London, 1970.
- Gelbard, F., Tambour, Y., and Seinfeld, J. H.: Sectional representations for simulating aerosol dynamics. *Journal of Colloid and Interface Science*, 76(2):541–556, 1980.
- Ghan, S., R. Easter, J. Hudson, and F. Breon: Evaluation of aerosol indirect radiative forcing in MIRAGE, *J. Geophys. Res.*, 106(D6), 5317– 5334, 2001.
- Hamed, A., Joutsensaari, J., Mikkonen, S., Sogacheva, L., Dal Maso, M., Kulmala, M., Cavalli, F., Fuzzi, S., Facchini, M. C., Decesari, S., Mircea, M., Lehtinen, K. E. J., Laaksonen, A.: Nucleation and growth of new particles in Po Valley, Italy *Atmos. Chem. Phys.*, 7, 355-376, 2007.
- Hamed, A., Birmili, W., Joutsensaari, J., Mikkonen, S., Asmi, A., Wehner, B., Spindler, G., Jaatinen, A., Wiedensohler, A., Korhonen, H., Lehtinen, K. E. J., and Laaksonen, A.:

- Changes in the production rate of secondary aerosol particles in Central Europe in view of decreasing SO₂ emissions between 1996 and 2006, *Atmos. Chem. Phys.*, 10, 1071-1091, doi:10.5194/acp-10-1071-2010, 2010.
- Hari, P. and Kulmala, M.: Station for measuring ecosystem-atmosphere relations (SMEAR II), *Boreal Environ. Res.*, 10, 315–322, 2005.
- Hyvönen, S., Junninen, H., Laakso, L., Dal Maso, M., Grönholm, T., Bonn, B., Keronen, P., Aalto, P., Hiltunen, V., Pohja, T., Launiainen, S., Hari, P., Mannila, H. and Kulmala, M.: A look at aerosol formation using data mining techniques, *Atmos. Chem. Phys.*, 5, 3345-3356, 2005.
- IPCC, 2007: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, UK and New York, NY, USA.
- Jaatinen, A., Hamed, A., Joutsensaari, J., Mikkonen, S., Birmili, W., Wehner, B., Spindler, G., Wiedensohler, A., Decesari, S., Mircea, M., Facchini, M. C., Junninen, H., Kulmala, M., Lehtinen, K. E. J. & Laaksonen, A.: A comparison of new particle formation events in the boundary layer at three different sites in Europe. *Boreal Environ. Res.*, 14: 481–498, 2009.
- Kennedy, I. M.: The health effects of combustion-generated aerosols. *Proceedings of the Combustion Institute*, 31 (2), Pages 2757-2770, doi:10.1016/j.proci.2006.08.116, 2007.
- Kerminen, V.-M. and Kulmala, M.: 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, 2002.
- Kerminen, V.-M., Lehtinen, K. J., Anttila, T., and Kulmala M.: Dynamics of atmospheric nucleation mode particles: a timescale analysis, *Tellus*, 56B, 135–146, 2004.
- Kerminen, V.-M., Lihavainen, H., Komppula, M., Viisanen, Y. & Kulmala, M. Direct observational evidence linking atmospheric aerosol formation and cloud droplet activation. *Geophys. Res. Lett.*, 32, L14803, doi:10.1029/2005GL023130, 2005.

- Kerminen, V.-M., Petäjä, T., Manninen, H. E., Paasonen, P., Nieminen, T., Sipilä, M., Junninen, H., Ehn, M., Gagné, S., Laakso, L., Riipinen, I., Vehkamäki, H., Kurten, T., Ortega, I. K., Dal Maso, M., Brus, D., Hyvärinen, A., Lihavainen, H., Leppä, J., Lehtinen, K. E. J., Mirme, A., Mirme, S., Hörrak, U., Berndt, T., Stratmann, F., Birmili, W., Wiedensohler, A., Metzger, A., Dommen, J., Baltensperger, U., Kiendler-Scharr, A., Mentel, T. F., Wildt, J., Winkler, P. M., Wagner, P. E., Petzold, A., Minikin, A., Plass-Dülmer, C., Pöschl, U., Laaksonen, A., and Kulmala, M.: Atmospheric nucleation: highlights of the EUCAARI project and future directions, *Atmos. Chem. Phys.*, 10, 10829-10848, doi:10.5194/acp-10-10829-2010, 2010.
- Kirkby, J., J. Curtius, J. Almeida, E. Dunne, J. Duplissy, S. Ehrhart, A. Franchin, S. Gagné, L. Ickes, A. Kürten, A. Kupc, A. Metzger, F. Riccobono, L Rondo, S. Schobesberger, G. Tsagkogeorgas, D. Wimmer, A. Amorim, F. Bianchi, M. Breitenlechner, A. David, J. Dommen, A. Downard, M. Ehn, R. C. Flagan, S. Haider, A. Hansel, D. Hauser, W. Jud, H. Junninen, F. Kreissl, A. Kvashin, A. Laaksonen, K. Lehtipalo, J. Lima, E. R. Lovejoy, V. Makhmutov, S. Mathot, J. Mikkilä, P. Minginette, S. Mogo, T. Nieminen, A. Onnela, P. Pereira, T. Petäjä, R. Schnitzhofer, J. H. Seinfeld, M. Sipilä, Y. Stozhkov, F. Stratmann, A. Tomé, J. Vanhanen, Y. Viisanen, A. Vrtala, P. E. Wagner, H. Walther, E. Weingartner, H. Wex, P. M. Winkler, K. S. Carslaw, D. R. Worsnop, U. Baltensperger & M. Kulmala: Role of sulphuric acid, ammonia and galactic cosmic rays in atmospheric aerosol nucleation. *Nature*, 476, 429–433, doi:10.1038/nature10343, 2011.
- Kokkola, H., Hommel, R., Kazil, J., Niemeier, U., Partanen, A-I., Feichter, J., and Timmreck, C.: Aerosol microphysics modules in the framework of the ECHAM5 climate model – intercomparison under stratospheric conditions, *Geosci. Model Dev.*, 2, 97-112, 2009.
- Koponen, I.K., Virkkula, A., Hillamo, R., Kerminen, V.M. and Kulmala, M.: Number size distributions and concentrations of the continental summer aerosols in Queen Maud Land, Antarctica. *J. Geophys. Res.*, 108, 4587, 2003.
- Korhonen, P., Kulmala, M., Laaksonen, A., Viisanen, Y., McGraw, R., and Seinfeld, J.: Ternary nucleation of H₂SO₄, NH₃ and H₂O in the atmosphere, *J. Geophys. Res.*, 104, 26349–26353, 1999.

- Korhonen H., Napari I., Timmreck C., Vehkamäki H., Pirjola L., Lehtinen K.E.J., Lauri A., & Kulmala M.: Heterogeneous nucleation as a potential sulphate-coating mechanism of atmospheric mineral dust particles and implications of coated dust on new particle formation. *J. Geophys. Res.*, 108, D17, 4546, doi: 10.1029/2003JD003553, 2003.
- Korhonen, H., Carslaw, K. S., Forster, P. M, Mikkonen, S., Gordon, N. D., and Kokkola, H.: Aerosol climate feedback due to decadal increases in southern hemisphere wind speeds, *Geophys. Res. Lett.*, 37, L02805, doi:10.1029/2009GL041320, 2010.
- Kuang, C., P. H. McMurry, A. V. McCormick, and F. L. Eisele: Dependence of nucleation rates on sulfuric acid vapor concentration in diverse atmospheric locations, *J. Geophys. Res.*, 113, D10209, doi:10.1029/2007JD009253, 2008.
- Kuang, C., Riipinen, I., Sihto, S.-L., Kulmala, M., McCormick, A. V., and McMurry, P. H.: An improved criterion for new particle formation in diverse atmospheric environments, *Atmos. Chem. Phys.*, 10, 8469-8480, doi:10.5194/acp-10-8469-2010, 2010.
- Kulmala, M. and Laaksonen, A.: Binary nucleation of water-sulfuric acid system: Comparison of classical theories with different H₂SO₄ saturation vapor pressures. *J. Chem. Phys.*, 93:696-791, 1990.
- Kulmala, M., Pirjola, L and Mäkelä, J.M.: Stable sulphate clusters as a source of new atmospheric particles, *Nature* 404, 66-69, doi:10.1038/35003550, 2000.
- Kulmala, M.: How Particles Nucleate and Grow, *Science* 7: 302 (5647), 1000-1001., doi: 10.1126/science.1090848, 2003.
- Kulmala, M., Vehkamäki, H., Petäjä, T., Dal Maso, M., Lauri, A., Kerminen, V.-M., Birmili, W., and McMurry, P.: Formation and growth rates of ultrafine atmospheric particles: a review of observations, *J. Aerosol Sci.*, 35, 143–176, 2004.
- Kulmala, M., Lehtinen, K. E. J., and Laaksonen, A.: Cluster activation theory as an explanation of the linear dependence between formation rate of 3nm particles and sulphuric acid concentration, *Atmos. Chem. Phys.*, 6, 787-793, 2006.
- Kulmala, M., Riipinen, I., Sipila, M., Manninen, H. E., Petaja, T., Junninen, H., Dal Maso, M., Mordas, G., Mirme, A., Vana, M., Hirsikko, A., Laakso, L., Harrison, R. M., Hanson, I.,

- Leung, C., Lehtinen, K. E. J., and Kerminen, V. M.: Toward direct measurement of atmospheric nucleation, *Science*, 318, 89–92, 2007.
- Kulmala, M., Asmi, A., Lappalainen, H. K., Carslaw, K. S., Pöschl, U., Baltensperger, U., Hov, Ø., Brenquier, J.-L., Pandis, S. N., Facchini, M. C., Hansson, H.-C., Wiedensohler, A., and O'Dowd, C. D.: Introduction: European Integrated Project on Aerosol Cloud Climate and Air Quality interactions (EUCAARI) – integrating aerosol research from nano to global scales, *Atmos. Chem. Phys.*, 9, 2825-2841, doi:10.5194/acp-9-2825-2009, 2009.
- Kurtén, T., Loukonen, V., Vehkamäki, H. and Kulmala, M.: Amines are likely to enhance neutral and ion-induced sulfuric acid-water nucleation in the atmosphere more effectively than ammonia. *Atmos. Chem. Phys.*, 8, 4095–4103, 2008.
- Laaksonen, A., Hamed, A., Joutsensaari, J., Hiltunen, L., Cavalli, F., Junkermann, W., Asmi, A., Fuzzi, S., and Facchini, M. C.: Cloud condensation nucleus production from nucleation events at a highly polluted region, *Geophys. Res. Lett.*, 32, 1–4, 2005.
- Laaksonen, A., Kulmala, M., Berndt, T., Stratmann, F., Mikkonen, S., Ruuskanen, A., Lehtinen, K. E. J., Dal Maso, M., Aalto, P., Petäjä, T., Riipinen, I., Sihto, S.-L., Janson, R., Arnold, F., Hanke, M., Ücker, J., Umann, B., Sellegri, K., O'Dowd, C. D., and Viisanen, Y.: SO₂ oxidation products other than H₂SO₄ as a trigger of new particle formation – Part 2: Comparison of ambient and laboratory measurements, and atmospheric implications, *Atmos. Chem. Phys.*, 8, 7255-7264, 2008a.
- Laaksonen, A., Kulmala, M., O'Dowd, C. D., Joutsensaari, J., Vaattovaara, P., Mikkonen, S., Lehtinen, K. E. J., Sogacheva, L., Dal Maso, M., Aalto, P., Petäjä, T., Sogachev, A., Yoon, Y. J., Lihavainen, H., Nilsson, D., Facchini, M. C., Cavalli, F., Fuzzi, S., Hoffmann, T., Arnold, F., Hanke, M., Sellegri, K., Umann, B., Junkermann, W., Coe, H., Allan, J. D., Alfarra, M. R., Worsnop, D. R., Riekkola, M. -L., Hyötyläinen, T., and Viisanen, Y.: The role of VOC oxidation products in continental new particle formation, *Atmos. Chem. Phys.*, 8, 2657-2665, 2008b.

- Lehtinen, K. E. J., Dal Maso, M., Kulmala, M., and Kerminen, V.-M.: Estimating nucleation rates from apparent particle formation rates and vice versa: Revised formulation of the Kerminen- Kulmala equation, *J. Aerosol Sci.*, 38, 988–994, 2007.
- Leppä, J., Anttila, T., Kerminen, V.-M., Kulmala, M., and Lehtinen, K. E. J.: Atmospheric new particle formation: real and apparent growth of neutral and charged particles, *Atmos. Chem. Phys.*, 11, 4939-4955, doi:10.5194/acp-11-4939-2011, 2011.
- Lighty, J. S., Veranth, J.M., and Sarofim, A. F.: Combustion Aerosols: Factors Governing Their Size and Composition and Implications to Human Health. *J. Air & Waste Manage. Assoc.* 50:1565-1618, 2000.
- Lihavainen, H.; Kerminen, V.-M.; Komppula, M.; Hatakka, J.; Aaltonen, V.; Kulmala, M.; Viisanen, Y. Production “potential” cloud condensation nuclei associated with atmospheric new-particle formation in northern Finland, *J. Geophys. Res.*, 108 (D24) 4782, 2003.
- Lindstrom, M., & Bates, D.: Newton-Raphson and EM algorithms for linear mixed effects models for repeated measures data. *Journal of the American Statistical Association*, 83, 1014-1022, 1988.
- Mahalanobis, P.C.: On the generalized distance in statistics, *Proceedings of the National Institute of Sciences of India*, vol. 12, pp. 49-55. 1936.
- Manninen, H. E., Nieminen, T., Asmi, E., Gagne, S., Hakkinen, S., Lehtipalo, K., Aalto, P., Vana, M., Mirme, A., Mirme, S., Horrak, U., Plass-Dulmer, C., Stange, G., Kiss, G., Hoffer, A., Toeroe, N., Moerman, M., Henzing, B., de Leeuw, G., Brinkenberg, M., Kouvarakis, G., Bougiatioti, A., Mihalopoulos, N., O’Dowd, C., Ceburnis, D., Arneth, A., Svenningsson, B., Swietlicki, E., Tarozzi, L., Decesari, S., Facchini, M., Birmili, W., Sonntag, A., Wiedensohler, A., Boulon, J., Sellegri, K., Laj, P., Gysel, M., Bukowiecki, N., Weingartner, E., Wehrle, G., Laaksonen, A., Hamed, A., Joutsensaari, J., Petaja, T., Kerminen, V.-M., and Kulmala, M.: EUCAARI ion spectrometer measurements at 12 European sites - analysis of new particle formation events. *Atmospheric Chemistry and Physics*, 10(16):7907–7927, 2010.
- McCullach, P. and Nelder, J.A.: *Generalized Linear Models*, London: Chapman and Hall, 1989.

- McCulloch, C.E. and Searle, S.R. *Generalized, Linear, and Mixed Models*, New York: Wiley, 2001.
- McMurry, P. H. : The history of Condensation Nucleus Counters, *Aerosol Sci. Technol.*, 33, 297–322, 2000.
- McMurry, P. H., Fink, M., Sakurai, H., Stolzenburg, M. R., Mauldin, R. L., Smith, J., Eisele, F., Moore, K., Sjostedt, S., Tanner, D., Huey, L. G., Nowak, J. B., Edgerton, E., and Voisin, D.: A criterion for new particle formation in the sulfur-rich Atlanta atmosphere, *J. Geophys. Res.*, 110, D22S02, doi:10.1029/2005JD005901, 2005.
- Merikanto, J., Napari, I., Vehkamäki, H., Anttila, T., and Kulmala, M.: New parameterization of sulfuric acid-ammonia-water ternary nucleation rates at tropospheric conditions, *J. Geophys. Res.* 112, D15207, doi:10.1029/2006JD007977, 2007.
- Merikanto, J., Spracklen, D. V., Mann, G. W., Pickering, S. J., and Carslaw, K. S.: Impact of nucleation on global CCN, *Atmos. Chem. Phys.*, 9, 8601-8616, doi:10.5194/acp-9-8601-2009, 2009.
- Metzger, A., Verheggen, B., Dommen, J., Duplissy, J., Prevot, A., Weingartner, E., Riipinen, I., Kulmala, M., Spracklen, D., Carslaw, K., and Baltensperger, U.: Evidence for the role of organics in aerosol particle formation under atmospheric conditions. *Proceedings of The National Academy of Sciences of The USA*, 107(15):6646–6651, 2010.
- Morrison, D. F.: *Multivariate Statistical Methods*, Belmont CA: Thomson/Brooks/Cole, 2005.
- Mäkelä, J., Dal Maso, M., Pirjola, L., Keronen, P., Laakso, L., Kulmala, M., and Laaksonen, A.: Characteristics of the atmospheric particle formation events observed at a boreal forest site in southern Finland, *Boreal Environ. Res.*, 5, 299–313, 2000.
- Mäkelä, J. M., Aalto, P., Jokinen, V., Pohja, T., Nissinen, A., Palmroth, S., Markkanen, T., Seitsonen, K., Lihavainen, H., and Kulmala, M.: Observations of ultrafine aerosol particle formation and growth in boreal forest, *Geophys. Res. Lett.*, 24, 1219–1222, 1997.
- Mönkkönen, P., Koponen, I.K., Lehtinen, K.E.J., Hämeri, K., Uma, R. and Kulmala, M.: Measurements in a highly polluted Asian mega city: observations of aerosol number size

- distribution, modal parameters and nucleation events. *Atmos. Chem. Phys.*, 5, 57–66, 2005.
- Nel, A.: Air Pollution-Related Illness: Effects of Particles. *Science* 6 May 2005: Vol. 308 no. 5723 pp. 804-806, doi: 10.1126/science.1108752, 2005.
- O'Dowd CD, Jimenez JL, Bahreini R, Flagan RC, Seinfeld JH, Hämeri K, Pirjola L, Kulmala M, Jennings SG, Hoffmann T.: Marine aerosol formation from biogenic iodine emissions. *Nature* 417, 632-636, doi:10.1038/nature00775, 2002.
- Paasonen, P., Sihto, S.-L., Nieminen, T., Vuollekoski, H., Riipinen, I., Plaß-Dülmer, C., Berresheim, H., Birmili, W. & Kulmala, M.: Connection between new particle formation and sulphuric acid at Hohenpeissenberg (Germany) including the influence of organic compounds. *Boreal Environ. Res.*, 14: 616–629, 2009.
- Paasonen, P., Nieminen, T., Asmi, E., Manninen, H. E., Petäjä, T., Plass-Dülmer, C., Flentje, H., Birmili, W., Wiedensohler, A., Hörrak, U., Metzger, A., Hamed, A., Laaksonen, A., Facchini, M. C., Kerminen, V.-M., and Kulmala, M.: On the roles of sulphuric acid and low-volatility organic vapours in the initial steps of atmospheric new particle formation, *Atmos. Chem. Phys.*, 10, 11223-11242, doi:10.5194/acp-10-11223-2010, 2010.
- Patterson, H.D.; Thompson, R: Recovery of inter-block information when block sizes are unequal. *Biometrika* 58 (3): 545–554. doi:10.1093/biomet/58.3.545, 1971.
- Petäjä, T., Mauldin, III, R. L., Kosciuch, E., McGrath, J., Nieminen, T., Paasonen, P., Boy, M., Adamov, A., Kotiaho, T., and Kulmala, M.: Sulfuric acid and OH concentrations in a boreal forest site, *Atmos. Chem. Phys.*, 9, 7435-7448, doi:10.5194/acp-9-7435-2009, 2009.
- Pierce, J. R., and P. J. Adams: Efficiency of cloud condensation nuclei formation from ultrafine particles, *Atmos. Chem. Phys.*, 7, 1367–1379, 2007.
- Pirjola, L., Laaksonen, A., Aalto, P., and Kulmala, M.: Sulfate aerosol formation in the Arctic boundary layer, *J. Geophys. Res.*, 103, 8309–8322, 1998.
- Pope, C. A. III and Dockery, D. W.: 2006 Critical Review - Health Effects of Fine Particulate Air Pollution: Lines that Connect, *J Air Waste Manage.*, 56, 709–742, 2006.

- R Development Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>, 2010.
- Ramanathan, V., P.J. Crutzen, J.T. Kiehl and D. Rosenfeld: Aerosols, Climate, and the Hydrological Cycle. *Science* 7, 294 (5549), 2119-2124, doi:10.1126/science.1064034, 2001.
- Riipinen, I., Sihto, S.-L., Kulmala, M., Arnold, F., Dal Maso, M., Birmili, W., Saarnio, K., Teinilä, K., Kerminen, V.-M., Laaksonen, A., and Lehtinen, K. E. J.: Connections between atmospheric sulphuric acid and new particle formation during QUEST III–IV campaigns in Heidelberg and Hyytiälä, *Atmos. Chem. Phys.*, 7, 1899–1914, doi:10.5194/acp-7-1899-2007, 2007.
- Schultz, D. M., S. Mikkonen, A. Laaksonen, and M. B. Richman, Weekly precipitation cycles? Lack of evidence from United States surface stations, *Geophys. Res. Lett.*, 34, L22815, doi:10.1029/2007GL031889, 2007.
- Seinfeld, J.H. and Pandis, S.N.: *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. John Wiley & Sons, Inc., New York, 2006.
- Seinfeld, J.: Atmospheric science: Black carbon and brown clouds, *Nature Geosci.*, 1, 15 – 16, doi:10.1038/ngeo.2007.62, 2008.
- Sihto, S.-L., Kulmala, M., Kerminen, V.-M., Dal Maso, M., Petäjä, T., Riipinen, I., Korhonen, H., Arnold, F., Janson, R., Boy, M., Laaksonen, A., and Lehtinen, K. E. J.: Atmospheric sulphuric acid and aerosol formation: implications from atmospheric measurements for nucleation and early growth mechanisms, *Atmos. Chem. Phys.*, 6, 4079-4091, doi:10.5194/acp-6-4079-2006, 2006.
- Sipilä, M., Berndt, T., Petäjä, T., Brus, D., Vanhanen, J., Stratmann, F., Patokoski, J., Mauldin III, R.L., Hyvärinen, A.-P., Lihavainen, H. and Kulmala, M.: The role of sulfuric acid in atmospheric nucleation, *Science*, 327, pp. 1243-1246, 2010.

- Smith, J.N., K.C. Barsanti, H.R. Friedli, M. Ehn, M. Kulmala and D.R. Collins, Observations of ammonium salts in atmospheric nanoparticles and possible climatic implications, *Proc. Natl. Acad. Sci. U.S.A.* 107, pp. 6634–6639, 2010.
- Smyth, G. K.: Nonlinear regression, in: *Encyclopedia of Environmetrics*, Volume 3, Editors A.H. El-Shaarawi and W.W. Piegorisch, John Wiley & Sons, Ltd, Chichester, 1405–1411, 2002.
- Spindler, G., E. Brüggemann, T. Gnauk, A. Grüner, K. Müller, H. Herrmann, A four-year size-segregated characterization study of particles PM₁₀, PM_{2.5} and PM₁ depending on air mass origin at Melpitz. *Atmos. Environ.*, 44, 164-173, 2010.
- Spracklen, D. V., Carslaw, K. S., Kulmala, M., Kerminen, V.-M., Mann, G. W., and Sihto, S.-L.: The contribution of boundary layer nucleation events to total particle concentrations on regional and global scales, *Atmos. Chem. Phys.*, 6, 5631-5648, doi:10.5194/acp-6-5631-2006, 2006.
- Spracklen, D. V., Carslaw, K. S., Kulmala, M., Kerminen, V.-M., Sihto, S.-L., Riipinen, I., Merikanto, J., Mann, G. W., Chipperfield, P. M., Wiedensohler, A., Birmili, W., and Lihavainen, H.: Contribution of particle formation to global cloud condensation nuclei concentrations, *Geophys. Res. Lett.*, 35, L06808, doi:10.1029/2007GL033038, 2008.
- Stolzenburg, M. and McMurry, P. H.: An ultrafine aerosol condensation nucleus counter, *Aerosol Sci. Technol.*, 14, 48–65, 1991.
- Stolzenburg, M. R., McMurry, P. H., Sakurai, H., Smith, J. N., Mauldin III, R. L., Eisele, F. L., and Clement, C. F.: Growth rates of freshly nucleated atmospheric particles in Atlanta, *J. Geophys. Res.*, 110, D22S05, doi:10.1029/2005JD005935, 2005.
- Twomey, S., The influence of pollution on the short wave albedo of clouds, *J. Atmos. Sci.*, 34, 1149–1152, 1977.
- Veefkind, J. P., Boersma, K. F., Wang, J., Kurosu, T. P., Krotkov, N., Chance, K., and Levelt, P. F.: Global satellite analysis of the relation between aerosols and short-lived trace gases, *Atmos. Chem. Phys.*, 11, 1255-1267, doi:10.5194/acp-11-1255-2011, 2011.
- Vehkamäki, H., Kulmala, M., Napari, I., Lehtinen, K. E. J., Timmreck, C., Noppel, M., and Laaksonen, A.: An improved parameterization for sulfuric acid/water nucleation rates for

- tropospheric and stratospheric conditions, *J. Geophys. Res.*, 107(D22), 4622, doi:10.1029/2002JD002184, 2002.
- Voutilainen A, Kolehmainen V, Kaipio JP: Statistical inversion of aerosol size measurement data, *Inv Probl Eng* 9:67-94, 2001.
- Weber, R., Marti, J., McMurry, P., Eisele, F., Tanner, D., and Jefferson, A.: Measured atmospheric new particle formation rates: Implications for nucleation mechanisms. *Chemical Engineering Communications*, 151:53–64, 1996.
- Wehner, B. and Wiedensohler, A.: Long term measurements of submicrometer urban aerosols: statistical analysis for correlations with meteorological conditions and trace gases, *Atmos. Chem. Phys.*, 3, 867-879, doi:10.5194/acp-3-867-2003, 2003.
- Whitby, E. R. and P. H. McMurry: Modal aerosol dynamics modelling. *Aerosol Sci. Technol.* 27:673-688, 1997.
- Wiedensohler, A., Aalto, P., Covert, D., Heintzenberg, J., and McMurry, P.: Intercomparison of four methods to determine size distributions of low- concentration ($<100 \text{ cm}^{-3}$), ultrafine aerosols ($3 \text{ nm} < dp < 10 \text{ nm}$) with illustrative data from arctic, *Aerosol Sci. Tech.*, 21, 95-109, 1994.
- World Health Organization: Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide - Global update 2005 - Summary of risk assessment. WHO World Health Organization ; 2006.
- Yu, F.: Updated H₂SO₂-H₂O binary homogeneous nucleation look-up tables, *J. Geophys. Res.*, 113, D24201, doi:10.1029/2008JD010527, 2008.
- Zhang, Renyi: Getting to the Critical Nucleus of Aerosol Formation. *Science* 11: 328 (5984), 1366-1367, 2010. (doi:10.1126/science.1189732)