Author's response to the referee comments

We thank the both referees for their comments and suggestions regarding our manuscript. Below we provide our answers (shown in italics) to these comments, and if changes were made to the manuscript also the modified text is given, with page and line numbers referring to the ACPD version of the paper. The revised manuscript is at the end of this documents, showing all the changes made to the ACPD version of the paper.

Answers to Anonymous Referee #1

The manuscript by Nieminen et al. reports on the predictability of new particle formation (NPF) events. Based on weather forecast (relevant parameters are the occurrence of clouds and rain), PM10 concentration (as a proxy for the condensation/coagulation sink), SO2 concentration (precursor for H2SO4) and the air mass history a day is classified either as NPF day, no NPF day or day with a weak possibility of NPF. The algorithm for the decision making is based on long-term measurements at the Hyytiälä/Finland field station. NPF event predictions were made during the PEGASOS (Pan-European Gas-Aerosol-Climate Interaction Study) campaign in central Finland during May and June 2013. The parameters used for the NPF forecast are from the Finnish Meteorological Institute (weather forecastas well as SO2 and PM10 from the SILAM (System for integrated modelling of atmospheric composition) air quality model). In addition, HYSPLIT trajectories are used for evaluating the air mass history. NPF forecasts were made three days and one day in advance in order to decide whether Zeppelin flights for NPF measurement should be launched or not. Comparison between predicted and observed NPF events is presented to evaluate the predictability of the method. The authors report that out of 11 NPF events 10 were accurately predicted. The manuscript is very clearly written and shows relevant data. Therefore, I recommend publication of the manuscript after some rather minor points have been addressed.

General remarks:
(1) The authors report that 10 out of 11 NPF event days were correctly predicted. However, Fig. 4 (colored bars) reveals that out of 11 observed NPF events "only" 6 were predicted as NPF event days and 4 were predicted to be undefined days with the possibility of NPF. In addition, 19 days classified as undefined days with the possibility of NPF occurred but only 10 of them were correctly predicted as undefined days and 7 were forecast to be non-NPF days (2 were predicted to be NPF days). The following table gives an overview on the statistics:

	observed (# of days)	"NPF" predicted (# of days)	"undefined" predicted (# of days)	"no-NPF" predicted (# of days)
NPF	**11**	6	4	1
undefined	**19**	2	10	7
no-NPF	**10**	0	2	8
sum	**40**	8	16	15

We did not predict "undefined" days but rather the occurance of NPF on the next day to be either (1) certain (NPF day), (2) possible (weak NPF/possibility of NPF), or (3) not possible (non-NPF day). The term "undefined day" is used for the classification of days based on observed particle size distribution evolution. In this regard, those 4 NPF days for which a possibility of NPF was forecasted can be considered succesfully forecasted. Out of the 19 days classified as undefined based on DMPS data, those days when NPF or possibility of NPF was forecasted, can also be considered to be succesfully forecasted.
We clarified the text on page 2469 lines 7–9 by adding a sentence: "Six of these days were also forecasted to be NPF days, and four to have a possibility of NPF to occur."

In this regard, could the authors please specify what they ultimately decided in terms of the Zeppelin flights? Was a flight scheduled whenever NPF or an undefined event was predicted? Or was the Zeppelin only launched if a clear NPF day was predicted? As the Zeppelin flights

were the motivation to forecast NPF events a bit more information on the actual decisions would be interesting. Maybe the authors could include a third row in Fig. 4 which indicates the Zeppelin flights.

The Zeppelin flights were also used to study other atmospheric phenomena than NPF. Therefore, the number of flight hours available for NPF studies was limited. Also other factors such as too high wind speeds or technical problems sometimes prevented the Zeppelin from flying, although NPF would have been forecasted to occur. As the focus of this technical note is to describe the NPF forecast method, we decided not add more details about the Zeppelin flights into the manuscript.

(2) It is not clear how the HYSPLIT trajectories were used. It is mentioned in section 2.2 (page 2463, line 21) that the trajectories were calculated 96 hours backwards in time but this would not allow making a prediction for the next three days.

The trajectories of the airmasses arriving at each hour to Hyytiälä were followed 96 hours prior to their arrival. Input meteorological data for the back-trajectory calculations was available for the next 8 days from GFS global weather forecasts. This way we were able to calculate the 96 hour forecast back-trajectories arriving at each hour to Hyytiälä during the next 3 days.

We modified the text in the beginning of chapter 2.2 to clarify this (page 2463, starting from line 16): "As input meteorological data for the model, we used the US National Weather Service's Global Forecasting System (GFS) weather forecast data which extends 192 hours forwards in time. The horizontal location accuracy of the air mass trajectory calculations using HYSPLIT has been estimated to be on the order of 10–30% of the total distance the air parcel has travelled (Stunder, 1996; Stohl, 1998; Draxler and Hess, 1998, 2010). We considered trajectories arriving each hour to Hyytiälä at 250-m height above ground calculated 96 hours backwards in time."

(3) It is mentioned that predictions were made both 3 days and 1 day in advance. How good are the 3 day predictions in comparison to the 1 day predictions?

We have not done comparisons between the accuracy of NPF forecasts done for 3 days in advance and the final NPF forecast for the next day. Typically there were some changes in the forecasts, especially if the general weather situation was changing rapidly (cloudiness and occurrence of rain). In order to clarify that the NPF forecast results presented in the manuscript are from the one day predictions, we added on page 2462 to line 17 the sentence: "All the NPF forecast results presented in this work refer to the final NPF forecasts, i.e. forecasts for the next day."

Other comments:

page 2461, line 10: please spell out "NT"
We added to the revised manuscript the definition "Neue Technologie" for the term "NT".

page 2461, line 12: "central"
When referring to a geographical area, "central" should be written with a capital letter as is done in the manuscript.

page 2462, line 3: "Lappalainen et al., 2009"

Corrected.

page 2463, line 5: What values for SO2 and PM10 were actually used for the forecast if hourly values are available? Was a forecast made for every hour of the day and then NPF was predicted if the algorithm yielded a positive outcome for just one specific hour, or was an average created somehow?
We paid special attention on the values forecasted by the SILAM model and weather forecasts for the early morning until around noon, as this is the time when NPF typically starts at the Hyytiälä station (see e.g. Dal Maso et al., 2005). We added the following explanation into the sentence on page 2464 lines 11–13: "Thus, our main criteria in forecasting NPF to occur was clear sky conditions, low condensation sink (in practise low PM10 concentration, which was obtained from SILAM) and low relative humidity in the early morning to noon-time, as this is the time when regional NPF events start in Hyytiälä (Kulmala et al., 2013)."

page 2463, line 24: insert blank before "but"
Corrected.

page 2465, line 4: delete the word "only"
We will keep the word "only" here, as we want to emphasize that NP1 is connected with only sulphuric acid proxy concentration while NP2 takes into account both sulphuric acid and oxidized organics proxies.

page 2466, line 20: after "2013": insert the total number of days (40?) of the campaign as the number of days for certain events is mentioned below
We added the total number of days the campaign lasted, and modified the first sentence of Section 3.1 on page 2466 lines 19–20 to: "The PEGASOS-Zeppelin Northern mission was a 40-day-long measurement campaign between 3 May and 11 June 2013."

page 2467, line 3: maybe better to write "... a longer period occurred during which transported polluted continental air dominated."
We agree with the referee that this sentence could be more clear, and modified it into: "At the end of May, a longer period occurred during which more polluted continental air was transported from Central Europe to Hyytiälä."

page 2467, line 12: remove the word "the"
Done.

page 2468, line 6: remove the word "the" before "Hyytiälä"
Done.

page 2468, line 7: remove the word "the" before "continental"
Done.

page 2468, line 9: remove the word "the" before "central"
Done.

page 2468, line 19: add the word "the" before "beginning"
Done.

page 2469, line 3: please spell out "DMPS" once

We added the full name of the instrument "Differential Mobility Particle Sizer" to this sentence.

page 2469, line 10: add the word "a" before "non-NPF"
Done.

page 2469, line 15: It is mentioned that on 10 days no NPF occurred but only one day was forecast to be an undefined day (28th May). However, there is another day with the same characteristics (17th of May, see Fig. 4).
The referee is correct, both 17th and 28th May were forecasted to have a possibility of NPF but were non-NPF days according to observations. We modified the sentence on page 2469 lines 13–15 to: "On 10 days of the campaign there was no particle formation occurring in Hyytiälä, and these were also forecasted to be non-NPF days, except two days (17th and 28th May) for which a possibility of NPF event was forecasted."

page 2469, line 23: "24" instead of "23"? please check
The referee is right, 24 is the correct number of days forecasted as either NPF days of possible NPF days. We corrected this in the revised manuscript.

page 2475, table 1: check the unit of the absolute humidity (ppth); parts per thousand should rather be abbreviated as "‰" or be spelled out
We spelled out the ppth notation as "part-per-thousand" in Table 1. We also changed the term "absolute humidity" to "H20" in order to avoid possible confusions. The ranges of H2O, SO2 and O3 on NPF and non-NPF days are all given as volume mixing ratios in Table 1.

page 2477, table 3: Could the authors please provide a short summary of the classification (class I and class II) in the manuscript text; a few explanatory sentences are probably sufficient.
We added a short explanation of the NPF classfication principles to page 2469, starting on line 5: "On NPF event days a new mode of particles smaller than 25 nm is observed and these particles can be observed growing to larger sizes during several hours. NPF event days are further classified according to the possibility to reliably derive particle formation and growth rates (Class I) or not (Class II). The days when no new sub-25 nm particles were appearing were classified as non-NPF days. Undefined days are those days for which it was not possible to unambiguously determine whether NPF occurred or not."

Answers to Anonymous Referee #2

The Authors present the application of a methodology for new-particle formation events optimized for a site in the European boreal forest and relying on the previous work on stochastic methods presented in Hyvönen et al. (ACP 2005). The usefulness of these methods for field campaign planning is clearly stated in the abstract of the paper. Stochastic methods represent a suitable alternative to chemical models incorporating a more mechanistic representation of new-particle formation. However, they must be optimized for the conditions encountered at the specific sites. In other words, the protocols for new particle formation (NPF) prediction presented in this study are just suitable for Hyytiälä and cannot be extrapolated to other environments. Therefore, the results presented in this paper must be treated mainly as an example of the suitability of stochastic models for NPF forecasts. The approach can be attractive for its simplicity (look at the decision tree in Figure 1), but it is actually based on an in-depth data mining work necessary to extract key predictors (Hyvönen et al., ACP 2005). The work presented here shows that, as the key parameters for Hyytiälä can be estimated by weather and chemical weather forecast models, the NPF occurrence can be predicted three days in advance. In the conclusions, the Authors seem to suggest that a similar methodology can be developed for other environments, provided that a sufficiently long record of measurements of the possible key predictors is available. However, this cannot be known with certainty, because it is possible that at other sites none of the simple physical and chemical parameters that are normally measured even at a well-equipped observatory can result to be optimal predictors. I would encourage the Authors to provide a more clear discussion of the applicability of their method outside the boreal forest.

We agree with the referee that the method presented in this work relies on knowing the typical range of atmospheric conditions relevant to nucleation (e.g. SO2 concentration, solar radiation, condensation sink) on NPF and non-NPF days at the location for which the NPF forecasts are done. For example, Jaatinen et al. (Boreal Env. Res. 14, 481–498, 2009) compared several prediction methods of NPF occurrence at three different sites in Finland and Central Europe, and their conclusion was that no single prediction method worked well at all the sites.

We added to the end of the abstract (page 2460, line 21): "To our knowledge, no similar forecasts of NPF occurance have been developed for other sites. This method of forecasting NPF occurance could be applied also at other locations provided that long-term observations of conditions favouring particle formation are available."

I have some major comments also on the methodology:

a) The work of Hyvönen et al. (ACP 2005) provides a detailed analysis of the best NPF predictors in Hyytiälä concluding that "This resulted in two key parameters, relative humidity and preexisting aerosol particle surface (condensation sink), capable in explaining 88% of the nucleation events. The inclusion of any further parameters did not improve the results notably". Instead here, other variables are taken into account (radiation, air mass origin), while RH disappears from the decision tree (Fig. 1). Why? Is the information on RH implicit in the "cloudy, rainy conditions"? But why not using directly RH instead of parameters difficult to quantify (cloudiness)?

The referee is right that typically cloudy and rainy days have high RH and warm sunny days lower RH. We decided to use cloudiness as one parameter in the forecasts, because the amount of cloudiness (clear sky, isolated clouds, broken clouds) has an influence not only on

the occurrence probability of NPF, but also on what kind of NPF event occurs (Sogacheva et al., 2008). Also, parameters that give best predictions of NPF events are not necessarily the same as those that separate NPF event and non-event days based on measurements best. In forecasting the NPF events we need to use forecasts of the other parameters, which all have uncertainties. Especially the predictions of CS have significant uncertainties. Therefore, even though based on measured RH and CS it is possible to distinguish between NPF and non-NPF days well, NFP predictions based on just these two parameters would have rather large uncertainty. Using several parameters enhances the predictability of the NPF events even if based on measurements the additional parameters would not enhance the separation efficiency of event and non-event days.

b) If the back-trajectories map the more or less anthropogenic impact on air mass composition, why using them as a separate variable with respect to CS and SO2?

Dal Maso et al. (2007) found that the airmasses arriving to Hyytiälä from the directions of north-west to north clearly favor NPF occurrence in Hyytiälä. On one hand this can be related to the less polluted airmasses arriving from this direction, favoring NPF due to typically low coagulation sink for newly formed particles. On the other hand, NPF is also promoted in the airmasses arriving from this direction due to various BVOCs emitted from the Scandinavian forests. Hence, the airmass source areas do not tell only about the anthropogenic impact as the referee points out, but also about the biogenic impact.

c) For the decision making flow chart in Figure 1, you set thresholds (6.6 μg/m3 of PM, 0.23 ppb of SO2, etc.). How were they chosen?

The threshold values are based on the range of conditions observed on NPF and non-NPF days in Hyytiälä presented in Table 1. We added to page 2465 on line 2 the following sentence: "The threshold values for SO2 and PM10 shown in the flowchart are based on the observed range of these variables on NPF and non-NPF days (Table 1)."

Some more major comments about results and conclusions:

a) Figure 1 shows a classification of forecasted events into three categories only. Why Table 3 shows multiple ways to describe the undefined events? Were there different types of undefined events?

The flowchart of Figure 1 shows three forecast categories: (1) NPF, (2) weak or possible NPF, or non-continuous growth of the particles, (3) no NPF. Categories (1) and (3) are the same in Table 3, but for category (2) there are given finer details of the classification, which was used during the campaign. In order to avoid confusion, we decided to change in the revised manuscript the terminology used in Table 3 for the category (2) forecasts to be exactly the same as in Figure 1 flowchart, i.e. either "weak NPF", "possible NPF" or "non-continuous growth of the particles".

b) When reporting the scores of the model, please use clear indexes for the missed, false and total misclassified events, such as in Hyvönen et al. (ACP 2005). Provide these score indexes for NPF and for non-NPF events.

In the following table, green shows days when the NPF forecast agrees with the observations, red shows days when there's disagreement (either missed or false forecast of NPF or non-NPF day):

	"NPF" forecasted (8 days)	"Weak NPF / Possibility of NPF / No continuous growth" forecasted (16 days)	"No-NPF" forecasted (16 days)
NPF day observed (11)	6	4	1
Undefined day observed (19)	2	10	7
Non-NPF day observed (10)	0	2	8

During the 40-day campaign, there were 11 NPF days and 10 non-NPF days according to DMPS data. The remaining 19 days were undefined, and are left out when calculating the score indexes, similarly as in Hyvönen et al. (2005). Out of these 21 days our forecasts had

- *2 false NPF event days (non-event day forecasted to be either event or to have a possibilty for event) → false NPF events 2/21 = 10%,*
- *1 NPF event day forecasted to be non-event day → missed NPF events days 1/21 = 5%.*

The total error of the NPF forecasts (false events + missed events) during classified days of the 40-day campaign was (2+1)/21 = 14%.

We add the above table to the revised manuscript as Table 4, and the following text to the revised manuscript at the end of Section 3.2 (page 2469, line 17): "Comparison of the event classification and the event forecasts is shown in Table 4. We follow the method of Hyvönen et al. (2005) for calculating the score indexes for the performance of the event forecasts on the 21 days classified as either NPF or non-NPF days (undefined days are removed from this comparison). Out of these 21 days our forecasts had two false NPF event days (non-event day forecasted to be either event or to have a possibilty for event) giving a 10% false-event fraction, and one NPF event day forecasted to be a non-event day giving a 5% missed-event fraction. The total error of the NPF forecasts (false and missed events) during the 21 classified days of the 40-day campaign was (2+1)/21 = 14%, which is comparable to the performance of the classification methods presented in the study by Hyvönen et al. (2005)."

c) The Authors' conclusions about the usefulness of the nucleation parameters NP1 and NP2 for NPF forecasting are unclear.

We added the following text (page 2469, line 2) discussing the usefulness of the nucleation parameters: "The nucleation parameters NP1 and NP2 have a clear connection to the NPF: they represent the ratios between the source and sink terms for the newly formed particles. However, the numerical values for NP1 and NP2 and especially their uncertainty depend greatly on the weather forecast and air-quality forecast data taken from the SILAM model. As it is out of the scope of this work to evaluate the accuracy of the SILAM predictions for the various parameters used, the values of NP1 and NP2 should be regarded as qualitative. When comparing the different days during the campaign, they did however provide useful information to support the NPF forecasting."

Finally, some specific comments:

a) Please, add some details on the SILAM model (resolution etc.)
We added the following details on page 2462, starting from line 25: "Input information for SILAM includes anthropogenic emission from the TNO-MACC data set, IS4FIRES information on wild-land fires, as well as emission calculations for sea salt, pollen, wind-bloan dust, and natural volatile organic compounds. The weather forecast input data is obtained from the FMI HIRLAM model. The horizontal resolution of SILAM in the Scandinavian area is 6–7 km."

b) Why using 96 h back-trajectories instead of shorter/longer ones?
96 h provide a good overview of the source areas of the airmasses arring at Hyytiälä, and are still reasonably accurate. Also, this is the same length for the trajectories as was used by Dal Maso et al. (2007) in their study which detemined that airmass arrival directions from southwest to north are favourable for NPF occurrence in Hyytiälä. As this information was used in our NPF forecasts, we wanted to keep the same calculation length for the forecast airmass back-trajectories.

1 **New particle formation event forecasts during PEGASOS–Zeppelin**

2 **Northern mission 2013 in Hyytiälä, Finland**

3 T. Nieminen[1,2], T. Yli-Juuti[1,3,4], H. E. Manninen[1], T. Petäjä[1], V.-M. Kerminen[1], and M. Kulmala[1]

4 [1] Department of Physics, University of Helsinki, P.O. Box 64, FI-00014, University of Helsinki,

5 Finland

6 [2] Helsinki Institute of Physics, P.O. Box 64, FI-00014, University of Helsinki, Finland

7 [3] Multiphase Chemistry Department, Max Planck Institute for Chemistry, Hahn-Meitner Weg 1, 55128

8 Mainz, Germany

9 [4] Department of Applied Physics, University of Eastern Finland, P.O. Box 1627, FI-70211 Kuopio,

10 Finland

11 **Abstract**

12 New particle formation (NPF) occurs frequently in the global atmosphere. During recent years detailed

13 laboratory experiments together with intensive field observations in different locations have provided

14 insights into the vapours responsible for the initial formation of particles and their subsequent growth.

15 In this regard, the importance of sulphuric acid, stabilizing bases such as ammonia and amines as well

16 as extremely low volatile organics have been proposed. The instrumentation to observe freshly formed

17 aerosol particles has developed to a stage where the instruments can be implemented as part of

18 airborne platforms, such as aircrafts or a Zeppelin-type airship. Flight measurements are technically

19 more demanding and require a greater detail of planning than field studies at the ground level. The

20 high cost of flight hours, limited time available during a single research flight for the measurements

21 and different instrument payloads in Zeppelin airship for various flight missions demanded an analysis

22 tool that would forecast whether or not there is a good chance for a NPF event. Here we present a

23 generalized methodology to derive NPF event probability that was used to optimise flying hours

24 during Zeppelin campaign in Finland during the PEGASOS-Zeppelin Northern mission in 2013.

25 Based on the existing knowledge we derived a method for estimating the nucleation probability that

26 utilizes forecast air mass trajectories, weather forecasts and air quality model predictions. With the

27 forecast tool we were able to predict the occurrence of NPF events for the next day with more than

28 90% success rate (10 out of 11 NPF event days correctly predicted). To our knowledge, no similar

29 forecasts of NPF occurance have been developed for other sites. This method of forecasting NPF

30 occurance could be applied also at other locations, provided that long-term observations of conditions

31 favouring particle formation are available.

1. Introduction

Formation and growth of secondary aerosol particles has been observed in numerous locations and in different environments in the planetary boundary layer (for an overview see e.g. Kulmala et al., 2004; Kulmala and Kerminen, 2008). Numerous investigations have attempted to connect new particle formation (NPF) to atmospheric trace gas concentrations, atmospheric chemistry and meteorological processes (e.g. Weber et al., 1995; Riipinen et al., 2007; Paasonen et al., 2010). Most of the NPF observations are based on stationary ground-level measurements during which the sampled air masses and prevailing meteorological conditions are continuously changing. Typically the growth of the newly formed particles can be followed for several hours from these fixed points measurements, indicating that NPF usually occurs over large areas (Dal Maso et al., 2007; Hussein et al., 2009). In order to obtain more information on the spatial extent of NPF events both in the vertical and horizontal directions, measurements using aircrafts are needed. As part of the 4 year-long EU funded PEGASOS (Pan-European Gas-Aerosol-Climate Interaction Study) project, a Zeppelin NT (Neue Technologie) airship was performing atmospheric aerosol, trace gas and photochemistry measurement flights in Central Finland during May–June 2013. In order to most efficiently utilize the flight hours of the airship, it was necessary to prepare forecasts on the probability of NPF events in the coming days.

Most of the Zeppelin measurement flights during the campaign were directed to the vicinity of the University of Helsinki SMEAR II measurement station in Hyytiälä (Hari and Kulmala, 2005). Measurements of aerosol number size-distributions, trace gas concentrations and basic meteorological quantities were started at the SMEAR II station in January 1996. These long time-series records have been used to extensively characterize the conditions in which NPF occurs (or does not occur) in this boreal forest environment, based on both the local atmospheric conditions as well as the synoptic situation and air mass origins and transport route to the station (Boy and Kulmala, 2002; Boy et al., 2003; Lyubovtseva et al., 2005; Dal Maso et al., 2007; Sogacheva et al., 2008, Nieminen et al., 2014).

Field observations, laboratory experiments and theoretical considerations have shown that sulphuric acid is one of the key components in atmospheric NPF events, but in addition also trace amounts of other vapors such as ammonia, amines or oxidized organics are needed (e.g. Kulmala et al., 2013). Particularly the contribution of extreme low volatile organics seems to be crucial in the boreal forest environment (Kulmala et al., 1998; Yli-Juuti et al., 2011; Ehn et al., 2014). Proxies for the concentrations of these trace gases or their precursors have been developed based on campaign-wise measurements (Petäjä et al., 2009; Lappalainen et al., 2009). Based on the concentrations and emissions of these trace gases, several parametrizations have been developed to describe the occurance

64 and intensity of NPF (e.g. Buzorius et al., 2003; Bonn et al., 2008; Paasonen et al., 2010; Häkkinen et
65 al., 2013).

66 In this work, we describe forecasts for the occurrence of NPF at the SMEAR II station. The forecasts
67 are based on the above-mentioned long-time series observations of the typical conditions during NFP
68 days and non-NPF days, the air mass origins as well as weather and air-quality forecasts.

69 ## 2. Materials and methods

70 The main objective of the NPF forecasts was to predict whether during the next three days NPF events
71 are likely to occur at the SMEAR II station area. A time period of three days was chosen in order to
72 have long enough time for preparing the measurement instruments needed on different flights while
73 still maintaining reliability of the input data used in making the NPF forecasts. The final NPF forecast
74 was always given for the next day, as the Zeppelin measurement flights were typically planned one
75 day in advance. All the NPF forecast results presented in this work refer to the final NPF forecasts, i.e.
76 forecasts for the next day.

77 2.1 Predictions for trace gas concentrations, particulate matter and meteorology

78 Forecasts for concentrations of trace gases SO_2, O_3, NO_x, CO and OH as well as particulate matter
79 (PM_{10}, comprising the total mass concentration of particles smaller than 10 μm in diameter) and
80 relative humidity were obtained from the Finnish Meteorological Institute's SILAM (System for
81 Integrated modeLling of Atmospheric coMposition) air quality model (Sofiev et al., 2006). This model
82 provides predictions for the above mentioned variables for the next 5 days at several heights above the
83 ground. Input information for SILAM includes anthropogenic emissions from TNO-MACC data set,
84 IS4FIRES information on wild-land fires, as well as emission calculations for sea salt, pollen, wind-
85 bloan dust, and natural volatile organic compounds. The weather forecast input data are obtained from
86 the FMI HIRLAM model. The horizontal resolution of SILAM in the Scandinavian area is 6–7 km. All
87 SILAM forecast data are freely accessible via internet (http://silam.fmi.fi/), and the forecast for the
88 Northern Europe area is updated once per day. For the purposes of the current NPF event forecasts, we
89 used predictions for the ground level (15 m above ground) during next 3 days from the model grid
90 point nearest to Hyytiälä SMEAR II station with the time resolution of 1 hour.

91 As supporting data, we also used several "traditional" weather forecasts available on the internet
92 (including forecasts by Finnish Meteorological Institute, Foreca and Norwegian Meteorological
93 Institute), mainly to evaluate the probabilities of cloudiness and rain. During the campaign time, the

94 weather was rather variable and the forecasts were changing rapidly (even several times a day) from
95 clear skies to partly cloudy and possibly rainy. All these conditions are known to affect directly the
96 probability of NPF.

97 ## 2.2 Air mass back-trajectories

98 Air mass arrival directions and source areas were forecasted for 96 hours prior to the arrival of air at
99 Hyytiälä using the HYSPLIT single particle Lagrangian transport model developed by NOAA and
100 freely available on the internet (http://www.arl.noaa.gov/HYSPLIT.php). As input meteorological data
101 for the model, we used the US National Weather Service's Global Forecasting System (GFS) weather
102 forecast data which extends 192 hours forwards in time. The horizontal location accuracy of the air
103 mass trajectory calculations using HYSPLIT has been estimated to be on the order of 10–30% of the
104 total distance the air parcel has travelled (Stunder, 1996; Stohl, 1998; Draxler and Hess, 1998, 2010).
105 We considered trajectories arriving each hour to Hyytiälä at 250-m height above ground calculated
106 96 hours backwards in time. Typically air masses travelled less than 1000 km during this time,
107 meaning that the air mass source area predictions were accurate within 100–300 km or less. Also,
108 since we did not consider just individual air mass back-trajectories but rather took into account all the
109 air masses that were to arrive during the morning and early afternoon (which is the typical time of
110 NPF occurance in Hyytiälä), the effect of uncertainties in the position of individual trajectories was
111 diminished.

112 ## 2.3 NPF event forecasts and nucleation probability parameters

113 Typical conditions on NPF and non-NPF days in Hyytiälä are shown in Table 1 for May and June
114 during years 1996–2012. In a data-mining study of the SMEAR II station long-time series records of
115 aerosol size distributions and meteorological parameters, Hyvönen et al. (2005) found that the
116 condensation sink (describing the pre-existing aerosol surface area) and relative humidity were the two
117 parameters most effectively separating NPF days from non-NPF days. Particle formation was occuring
118 only on days with a low CS and low RH. On the other hand, photochemical production of vapors
119 participating in nucleation and growth, namely sulphuric acid and oxidation products of organics, is
120 more efficient in clear-sky conditions with high UV radiation intensity compared to cloudy conditions.
121 Thus, our main criteria in forecasting NPF to occur was clear sky conditions, low condensation sink
122 (in practise low PM_{10} concentration, which was obtained from SILAM) and low relative humidity. in
123 the early morning to noon-time, as this is the time when regional NPF events start in Hyytiälä
124 (Kulmala et al., 2013). Note that typically in spring and summer time, days with low relative humidity
125 are typically also warm and sunny, so these conditions are not necessarily independent of each other.
126 However, the difference between NPF days and non-NPF days is also seen in the absolute humidity

127   (water vapor concentration, see Table 1).

128   The air mass source area and transport route to Hyytiälä were considered when making the NPF
129   forecasts. In the long-time series analysis by Dal Maso et al. (2007), the occurrence of NPF in
130   Hyytiälä was observed to be highly favorable in air masses originating from the Arctic Ocean and
131   Northern Atlantic, and on the other hand suppressed in southern air masses. This is typically connected
132   to clean air arriving from the west and more polluted air originating from Central and Eastern Europe,
133   directly influencing the sink for newly formed particles. However, in air masses originating from south
134   and south-east to Hyytiälä, SO$_2$ concentrations are typically higher than in westerly air masses, which
135   would favor NPF due to a higher production rate of sulphuric acid (Riuttanen et al., 2013). Table 2
136   summarizes the criteria used for making the NPF forecasts. The flowchart representing the main
137   decision making process for the NPF forecasts is shown in Figure 1. The threshold values for SO2 and
138   PM10 shown in the flowchart are based on the observed range of these variables on NPF and non-NPF
139   days (Table 1).

140   We also developed several "nucleation parameters" to forecast the intensity of NPF. The parameters
141   that worked best were either related to only the proxy concentration of sulphuric acid, or to proxies for
142   both sulphuric acid and oxidation products of volatile organic compounds (such as monoterpenes).
143   Paasonen et al. (2010) studied several different parameterizations for the formation rate of 2 nm
144   particles, and found that at the Hyytiälä site nucleation rate could be mainly explained by the sulphuric
145   acid concentration to the power of one or two.

146   The simplest nucleation parameter is described by the following equation:

147
$$\mathrm{NP}_1 \;=\; \frac{[\mathrm{SO_2}]\cdot[\mathrm{OH}]}{\mathrm{PM_{10}\cdot RH}} \tag{1}$$

148   where the sulphur dioxide concentration (SO$_2$), hydroxyl radical concentration (OH), particulate mass
149   concentration (PM$_{10}$) and relative humidity (RH) are taken from the SILAM air quality forecasts for
150   the grid point closest to Hyytiälä. From SILAM forecast, the particulate mass concentration is
151   available. In Hyytiälä, the PM$_{10}$ concentrations correlate well with the condensation sink CS which
152   describes the total sink of the newly formed particles due to the pre-existing aerosol population. The
153   PM$_{10}$ concentrations (in units µg m^{-3}) can be scaled to CS (in units s^{-1}) using the linear relationship CS
154    $= 4.59\cdot10^{-4}\cdot$PM$_{10}$ (linear regression based on measurement data from Hyytiälä in 1996–2012 with
155   correlation coefficient r=0.81). The relative humidity is included as RH^{-1} in Equation 1 in order to take
156   into account the observed anti-correlation between the relative humidity and particle formation

157   intensity, mainly due to the fact that the highest sulphuric acid concentrations are limited to times of

158   low ambient relative humidity (Hamed et al., 2011).

159   A nucleation parameter taking into account the oxidation products of monoterpenes, in addition to

160   sulphuric acid, is described by the following equation:

161
$$\mathrm{NP_2} \;=\; \frac{[\mathrm{SO_2}]\cdot[\mathrm{OH}]}{\mathrm{PM_{10}}\cdot\mathrm{RH}}\cdot\frac{\exp(aT)\cdot\big(k_{\mathrm{OH}}[\mathrm{OH}]+k_{\mathrm{O3}}[\mathrm{O_3}]\big)}{\mathrm{BLH}\cdot\mathrm{PM_{10}}} \tag{2}$$

162   Here, the concentrations of sulphur dioxide SO_2, hydroxyl radicals OH and ozone O_3 (in units of cm^{-3}),

163   particulate mass PM_{10} (in units μg m^{-3}), as well as relative humidity RH (in percentages) and

164   temperature T (in units ºC) were obtained from SILAM forecasts. The concentrations of monoterpenes

165   were predicted based on the ambient temperature, as their concentrations have been shown to follow

166   an exponential temperature dependence in Hyytiälä with the scaling coefficient a=0.078 $ºC^{-1}$

167   (Lappalainen et al., 2009). The OH and O_3 concentrations were used to calculate the proxy

168   concentrations of the monoterpene oxidation products, and the reaction coefficients k_{OH}=7.5×10^{-11}

169   cm^3s^{-1} and k_{O3}=1.4×10^{-17} cm^3s^{-1} are the averages of the reaction coefficients for individual

170   monoterpene species weighted according to their typical concentrations observed in Hyytiälä (Hakola

171   et al., 2003; Yli-Juuti et al., 2011). The modelled boundary layer height BLH is included in Equation 2

172   to take into account the dilution of monoterpene emissions into the developing boundary layer.

173   **3. Results**

174   3.1 Overview of the conditions during the campaign

175   The PEGASOS-Zeppelin Northern mission ~~took place~~was a 40-day-long measurement campaign

176   between 3 May and 11 June 2013. An overview of the meteorological conditions as well as trace gas

177   and particle concentrations observed at the SMEAR II station during the campaign is shown in Figure

178   2. Most of the days were sunny with either clear or partly clear skies. Rain occurred on 13 days during

179   the campaign. The air was rather clean from anthropogenic pollution, especially in the first and last

180   week of the campaign. Occasionally, there were pollution episodes seen e.g from a ten-fold rise of the

181   SO_2 concentration from its typical level of about 0.1 ppb. At the end of May, a longer period ~~of~~

182   ~~transport of polluted continental air occurred.~~occurred during which more polluted continental air was

183   transported from Central Europe to Hyytiälä.

184   Figure 3 shows the arrival routes of air masses to Hyytiälä during the period of our measurement

185   campaign. These trajectories were calculated for the 250-m arrival height above ground, and 96 hours

186   backwards in time. From the beginning of the campaign until middle of May, approximately 17th May,

187   the air masses originated mainly from over the Atlantic, and arrived at Hyytiälä either directly from

188   the west over Scandinavia or from south-west making a turn over the Baltic sea. Air in Hyytiälä was

189   relatively clean during this time, characterized by low particulate mass and trace gas concentrations.

190   Especially SO_2 had very low concentrations during this time, with the exception of one pollution-

191   related peak on 9 May. After the mid-May, air masses turned to arrive mainly from east at Hyytiälä,

192   originating either from over the Arctic Ocean or from the continental north-west Russia. During this

193   time until early June, the condensation sink and PM_{10} concentrations were higher than in early May,

194   indicating more polluted air. Also high concentration peaks in the trace gases SO_2 and CO were more

195   frequent during this time. During the last weeks of the campaign in the beginning of June, air masses

196   turned again to arrive at Hyytiälä from west over Scandinavia, resulting in cleaner air with low

197   particulate matter and trace gas concentrations.

198   3.2 Performance of the NPF forecasts and nucleation parameters

199   Figure 4 shows the particle number size-distributions along with the forecasted NPF occurrence and

200   the time-series of the nucleation parameters NP_1 and NP_2. In the beginning of the campaign, several

201   strong NPF bursts occurred (high nucleation mode particle concentrations on 3, 6 and 8 of May), and

202   our forecasts were able to capture these as well as the days with no new particle formation. Both of the

203   nucleation parameters peaked on these three NPF event days, and were clearly lower on the days

204   between NPF events, except NP_1 which had a relatively high value also on 4 May. During the

205   beginning of the campaign time, air masses originated mainly from over the Atlantic Ocean and

206   arrived at Hyytiälä after passing over Scandinavia. On some of these days, the air was remarkably

207   clean, characterized by very low SO_2 concentrations (below 0.1 ppb), resulting in low sulphuric acid

208   concentrations and weak or no NPF event on clear-sky conditions. The daytime peak value of 10^4 or

209   higher for nucleation parameter NP_1 was typically associated with the occurance of NPF.

210   After mid-May until early June, the air masses arrived at the Hyytiälä mainly from east, either

211   spending several days over the continental Russia or, in some cases, coming more directly from over

212   the Arctic Ocean via northwest Russia. The air mass circulation was driven by a persistent high-

213   pressure system residing over the cCentral Finland. This resulted in a rather unusual air mass transport

214   pattern to Hyytiälä, and also made the NPF forecasting more challenging. During this time, there were

215   situations when the polluted air masses resulted in a high condensation sink, preventing the occurrence

216   of NPF. Also the SILAM forecasts for the SO_2 and PM_{10} concentrations were less accurate during the

217   easterly air masses compared with air masses coming from west or south. This might be related to less

218   accurate emission data for these species over the Russian area.

219   The nucleation parameter NP_2 started to have high values more frequently after the middle of May.
220   One factor influencing this was the higher air temperatures during this time compared to the beginning
221   of the campaign, as the emissions of monoterpenes are highly influenced by the ambient temperature.
222   NPF events, however, were not as frequent during this time. On one hand, this period was influenced
223   by the more polluted air masses arriving at Hyytiälä from the east. On the other hand, during this time
224   there were quite many days (13 out of 22 days after 20th May) when a growing particle mode was
225   observed to appear in Hyytiälä starting from sizes above 10–20 nm. These types of NPF events are
226   typically observed during the summer time in Hyytiälä, and they might be connected to higher particle
227   growth rates during the summer, leading to observing the newly formed particles after they have
228   already grown for several hours (Buenrostro Mazon et al., 2009). When the nucleation parameter NP_2
229   had daytime maximum values higher than 0.02, NPF occurrence started to be more probable than the
230   day being a non-event day.
231   The nucleation parameters NP_1 and NP_2 have a clear connection to the NPF: they represent the ratios
232   between the source and sink terms for the newly formed particles. However, the numerical values for
233   NP_1 and NP_2 and especially their uncertainty depend greatly on the weather forecast and air-quality
234   forecast data taken from the SILAM model. As it is out of the scope of this work to evaluate the
235   accuracy of the SILAM predictions for the various parameters used, the values of NP_1 and NP_2
236   presented in this study should be regarded as qualitative. When comparing the different days during
237   the campaign, they did however provide useful information to support the NPF forecasting.

238   The particle number size distributions measured by the Differential Mobility Particle Sizer (DMPS)
239   during the whole campaign are shown in the upper panel of Figure 4. Using the criteria developed by
240   Dal Maso et al. (2005), each day was classified as either a NPF event, non-event or undefined day. On
241   NPF event days a new mode of particles smaller than 25 nm is observed and these particles can be
242   observed growing to larger sizes during several hours. NPF event days are further classified according
243   to the possibility to reliably derive particle formation and growth rates (Class I) or not (Class II). The
244   days when no new sub-25 nm particles were appearing were classified as non-NPF days. Undefined
245   days are those days for which it was not possible to unambiguously determine whether NPF occurred
246   or not. Table 3 shows the forecast and the corresponding event classification for each day. During the
247   40 day campaign, clear regional NPF events occurring during several hours were observed on 11 days
248   in Hyytiälä, and 10 of these days were also forecasted to be either NPF day or to have a possibility of
249   NPF to occur. Six of these days were also forecasted to be NPF days, and four to have a possibility of
250   NPF to occur. The NPF day which we forecasted to be a non-NPF day (9th June) was cloudy and had a
251   possibility of rain according to weather forecasts, and the airmasses were forecasted to originate from
252   west, which is not the direction from where airmasses typically arrive to Hyytiälä on NPF event days

16

253 (Dal Maso et al., 2007). On 10 days of the campaign there was no particle formation occurring in
254 Hyytiälä, and these were also forecasted to be non-NPF days, except ~~one day (28th May)~~two days
255 (17th and 28th May) for which a possible NPF event was forecasted. This was most probably caused
256 by the very low SO_2 concentration. On only one of the days forecasted to be non-NPF day, was there
257 appearance and growth of new nucleation mode particles.

258 Comparison of the event classification and the event forecasts is shown in Table 4. We follow the
259 method of Hyvönen et al. (2005) for calculating the score indexes for the performance of the event
260 forecasts on the 21 days classified as either NPF or non-NPF days (undefined days are removed from
261 this comparison). Out of these 21 days our forecasts had two false NPF event days (non-event day
262 forecasted to be either event or to have a possibilty for event) giving a 10% false-event fraction, and
263 one NPF event day forecasted to be a non-event day giving a 5% missed-event fraction. The total error
264 of the NPF forecasts (false and missed events) during the 21 classified days of the 40-day campaign
265 was (2+1)/21 = 14%, which is comparable to the performance of the classification methods presented
266 in the study by Hyvönen et al. (2005).

4. Summary and conclusions

268 Here we present a way to forecast new particle formation events. Being able to make such forecast
269 accurately is very important, for example, when airborne measurements are performed. As a summary,
270 we made NPF forecast for 40 days. The forecasts were found to work reasonably well. Only one day
271 when nucleation was forecasted to occur was a non-nucleation event day. In total ~~23~~24 days were
272 predicted to be either NPF event days or probable NPF event days, and on 10 of them there was an
273 NPF event, 11 were undefined (when it could not be reliably determined whether NPF occurred or
274 not), and two were non-event days.

275 The main challenges in making the NPF forecasts were to obtain as reliable input data as possible
276 from SILAM, HYSPLIT and weather forecasts. The methods utilized here are most likely also
277 applicable to other locations where there is sufficiently long datasets available characterizing the
278 conditions favourable for the occurrence of particle formation.

Acknowledgements

References

288 **References**

289 Bonn, B., Boy, M., Kulmala, M., Groth, A., Trawny, K., Borchert, S., and Jacobi S.: A new
290 parametrization for ambient particle formation over coniferous forests and its potential implications
291 for the future. Atmosperic Chemistry and Physics 9, 8079–8090, 2008.

292 Boy, M., and Kulmala, M.:Nucleation events in the continental boundary layer: Influence of physical
293 and meteorological parameters. Atmospheric Chemistry and Physics 2, 1–16, 2002.

294 Boy, M., Rannik, Ü., Lehtinen, K. E. J., Tarvainen, V., Hakola, H., and Kulmala, M.: Nucleation
295 events in the continental boundary layer: Long-term statistical analyses of aerosol relevant
296 characteristics. Journal of Geophysical Research 108(D21), 4667, 2003.

297 Buenrostro Mazon, S., Riipinen, I., Schultz, D. M., Valtanen, M., Dal Maso, M., Sogacheva, L.,
298 Junninen, H., Nieminen, T., Kerminen, V.-M., and Kulmala, M.: Classifying previously undefined
299 days from eleven years of aerosol-particle-size distribution data from the SMEAR II station, Hyytiälä,
300 Finland. Atmospheric Chemistry and Physics 9, 667–676, 2009.

301 Buzorius, G., Rannik, Ü., Aalto, P., Dal Maso, M., Nilsson, E. D., Lehtinen, K. E. J., and Kulmala, M.:
302 On particle formation prediction in continental boreal forest using micrometeorological parameters.
303 Journal of Geophysical Research 108(D13), 4377, 2003.

304 Dal Maso, M., Kulmala, M., Riipinen, I., Wagner, R., Hussein, T., Aalto, P. P., and Lehtinen, K. E. J.:
305 Formation and growth of fresh atmospheric aerosols: eight years of aerosol size distribution data from
306 SMEAR II, Hyytiälä, Finland. Boreal Environment Research 10, 323–336, 2005.

307 Dal Maso, M., Sogacheva, L., Aalto, P. P., Riipinen, I., Komppula, M., Tunved, P., Korhonen, L., Suur-
308 Uski, V., Hirsikko, A., Kurtén, T., Kerminen, V.-M., Lihavainen, H., Viisanen, Y., Hansson, H.-C., and
309 Kulmala, M.: Aerosol size distribution measurements at four Nordic field stations: identification,
310 analysis and trajectory analysis of new particle formation bursts. Tellus 59B, 350–361, 2007.

311 Draxler, R. R., and Hess, G. D.: An overview of the HYSPLIT_4 modelling system for trajectories,
312 dispersion and deposition. Australian Meteorological Magazine 47, 295–308, 1998.

313 Draxler, R. R., and Hess, G. D.: Description of the HYSPLIT_4 modeling system. NOAA Technical
314 Memorandum ERL ARL-224. http://www.arl.noaa.gov/documents/reports/arl-224.pdf

315 Ehn, M., Thornton, J. A., Kleist, E., Sipilä, M., Junninen, H., Pullinen, I., Springer, M., Rubach, F.,
316 Tillmann, R., Lee, B., Lopez-Hilfiker, F., Andres, S., Acir, I.-H., Rissanen, M., Jokinen, T.,
317 Schobesberger, S., Kangasluoma, J., Kontkanen, J., Nieminen, T., Kurtén, T., Nielsen, L. B.,
318 Jørgensen, S., Kjaergaard, H. G., Canagaratna, M., Dal Maso, M., Berndt, T., Petäjä, T., Wahner, A.,
319 Kerminen, V.-M., Kulmala, M., Worsnop, D. R., Wildt, J., and Mentel, T. F.: A large source of low-
320 volatility secondary organic aerosol. Nature 506, 476–479, 2014.

321 Hakola, H., Tarvainen, V., Laurila, T., Hiltunen, V., Hellen, H., and Keronen, P.: Seasonal variation of
322 VOC concentrations above a boreal coniferous forest. Atmospheric Environment 37, 1623–1634,
323 2003.

324 Hari, P., and Kulmala, M.: Station for measuring ecosystem–atmosphere relations (SMEAR II). Boreal
325 Environment Research 10, 315–322, 2005.

326 Hamed, A., Korhonen, H., Sihto, S.-L., Joutsensaari, J., Järvinen, H., Petäjä, T., Arnold, F., Nieminen,
327 T., Kulmala, M., Smith, J. N., Lehtinen, K. E. J., and Laaksonen, A.: The role of relative humidity in
328 continental new particle formation. Journal of Geophysical Research 116, D03202, 2011.

329 Hussein, T., Junninen, H., Tunved, P., Kristensson, A., Dal Maso, M., Riipinen, I., Aalto, P. P.,
330 Hansson, H.-C., Swietlicki, E., and Kulmala, M.: Time span and spatial scale of regional new particle
331 formation events over Finland and Southern Sweden. Atmospheric Chemistry and Physics 9, 4699–
332 4716, 2009.

333 Hyvönen, S., Junninen, H., Laakso, L., Dal Maso, M., Grönholm, T., Bonn, B., Keronen, P., Aalto, P.,
334 Hiltunen, V., Pohja, T., Launiainen, S., Hari, P., Mannila, H., and Kulmala M.: A look at aerosol
335 formation using data mining techniques. Atmospheric Chemistry and Physics 5, 3345–3356, 2005.

336 Häkkinen, S. A. K., Manninen, H. E., Yli-Juuti, T., Merikanto, J., Kajos, M. K., Nieminen, T.,
337 D'Andrea, S. D., Asmi, A., Pierce, J. R., Kulmala, M., and Riipinen, I.: Semi-empirical
338 parameterization of size-dependent atmospheric nanoparticle growth in continental environments.
339 Atmospheric Chemistry and Physics 13, 7665–7682, 2013.

340 Kulmala, M., Toivonen, A., Mäkelä, J. M., and Laaksonen, A.: Analysis of the growth of nucleation
341 mode particles observed in Boreal forest. Tellus 50B, 449 462, 1998.

342 Kulmala, M., Vehkamäki, H., Petäjä, T., Dal Maso, M., Lauri, A., Kerminen, V.-M., Birmili, W., and
343 McMurry, P. H.: Formation and growth rates of ultrafine atmospheric particles: a review of
344 observations. Journal of Aerosol Science 35, 143–176, 2004.

345 Kulmala, M., and Kerminen, V.-M.: On the formation and growth of atmospheric nanoparticles.
346 Atmospheric Research 90, 132 – 150, 2008.

347 Kulmala, M., Kontkanen, J., Junninen, H., Lehtipalo, K., Manninen, H. E., Nieminen, T., Petäjä, T.,
348 Sipilä, M., Schobesberger, S., Rantala, P., Franchin, A., Jokinen, T., Järvinen, E., Äijälä, M.,
349 Kangasluoma, J., Hakala, J., Aalto, P. P., Paasonen, P., Mikkilä, J., Vanhanen, J., Aalto, J., Hakola, H.,
350 Makkonen, U., Ruuskanen, T., Mauldin, R. L., Duplissy, J., Vehkamäki, H., Bäck, J., Kortelainen, A.,
351 Riipinen, I., Kurtén, T., Johnston, M. V., Smith, J. N., Ehn, M., Mentel, T. F., Lehtinen, K. E. J.,
352 Laaksonen, A., Kerminen, V.-M., and Worsnop, D. R.: Direct observations of atmospheric aerosol
353 nucleation. Science 339, 943–946, 2013.

354 Lappalainen, H. K., Sevanto, S., Bäck, J., Ruuskanen, T. M., Kolari, P., Taipale, R., Rinne, J.,
355 Kulmala, M., and Hari, P.: Day-time concentrations of biogenic volatile organic compounds in a boreal
356 forest canopy and their relation to environmental and biological factors. Atmospheric Chemistry and
357 Physics 9, 5447–5459, 2009.

358 Lyubovtseva, Y. S., Sogacheva, L., Dal Maso, M., Bonn, B., Keronen, P., and Kulmala, M.: Seasonal
359 variations of trace gases, meteorological parameters, and formation of aerosols in boreal forests.
360 Boreal Environment Research 10, 493–510, 2005.

361 Nieminen, T., Asmi, A., Dal Maso, M., Aalto, P. P., Keronen, P., Petäjä, T., Kulmala, M., and
362 Kerminen, V.-M.: Trends in atmospheric new particle formation: 16 years of observations in boreal
363 forest environment. Boreal Environment Research 19 (suppl. B), 191–214, 2014.

364 Paasonen, P., Nieminen, T., Asmi, E., Manninen, H. E., Petäjä, T., Plass-Dülmer, C., Flentje, H.,
365 Birmili, W., Wiedensohler, A., Hõrrak, U., Metzger, A., Hamed, A., Laaksonen, A., Facchini, M. C.,
366 Kerminen, V.-M., and Kulmala, M.: On the roles of sulphuric acid and low-volatility organic vapours
367 in the initial steps of atmospheric new particle formation. Atmosperic Chemistry and Physics 10,
368 11223–11242, 2010.

369 Petäjä, T., Mauldin, R. L., Kosciuch, E., McGrath, J., Nieminen, T., Paasonen, P., Boy, M., Adamov,
370 A., Kotiaho, T., and Kulmala, M: Sulfuric acid and OH concentrations in a boreal forest site.

371   Atmospheric Chemistry and Physics 9, 7435–7448, 2009.

372   Riipinen, I., Sihto, S.-L., Kulmala, M., Arnold, F., Dal Maso, M., Birmili, W., Saarnio, K., Teinilä, K.,
373   Kerminen, V.-M., Laaksonen, A., and Lehtinen, K. E. J.: Connections between atmospheric sulphuric
374   acid and new particle formation during QUEST III–IV campaigns in Heidelberg and Hyytiälä.
375   Atmospheric Chemistry and Physics 7, 1899–1914, 2007.

376   Riuttanen, L., Hulkkonen, M., Dal Maso, M., Junninen, H., and Kulmala, M.: Trajectory analysis of
377   atmospheric transport of fine particles, SO_2, NO_x and O_3 to the SMEAR II station in Finland in 1996–
378   2008. Atmospheric Chemistry and Physics 13, 2153–2164, 2013.

379   Sofiev, M., Siljamo, P., Valkama, I., Ilvonen, M., and Kukkonen, J.: A dispersion modelling system
380   SILAM and its evaluation against ETEX data. Atmospheric Environment 40, 674–685, 2006.

381   Sogacheva, L., Saukkonen, L., Nilsson, E. D., Dal Maso, M., Schultz, D. M., De Leeuw, G., and
382   Kulmala, M.: New aerosol particle formation in different synoptic situations at Hyytiälä, Southern
383   Finland. Tellus 60B, 485–494, 2008.

384   Stohl, A.: Computation, accuracy and applications of trajectories–a review and bibliography.
385   Atmospheric Environment 32, 947–966, 1998.

386   Stunder, B. J. B.: An assessment of the quality of forecast trajectories. Journal of Applied Meteorology
387   35, 1319–1331, 1996.

388   Weber, R. J., McMurry, P. H., Eisele, F. L., and Tanner, D. J.: Measurement of expected nucleation
389   precursor species and 3 to 500 nm diameter particles at Mauna Loa Observatory, Hawaii. J. Atmos.
390   Sci. 52, 2242–2257, 1995.

391   Yli-Juuti, T., Nieminen, T., Hirsikko, A., Aalto, P. P., Asmi, E., Hõrrak, U., Manninen, H. E.,
392   Patokoski, J., Dal Maso, M., Petäjä, T., Rinne, J., Kulmala, M., and Riipinen, I.: Growth rates of
393   nucleation mode particles in Hyytiälä during 2003–2009: variation with particle size, season, data
394   analysis method and ambient conditions. Atmosperic Chemistry and Physics 11, 12865–12886, 2011.

Tables and Figures

Table 1. Conditions observed at Hyytiälä during NPF and non-NPF days in months May–June 1996–2012. For each variable the median value is given and the interquartile range (25th and 75th percentiles) is shown in brackets.

Parameter	NPF day	Non-NPF day
Temperature ($^{\circ}$C)	11 (8–14)	13 (9–17)
Global radiation (W m^{-2})	560 (430–610)	230 (120–530)
Relative humidity (%)	45 (39–55)	76 (59–91)
~~H$_2$O~~Absolute humidity (~~ppth~~parts-per-thousand)	6.5 (4.9–8.0)	10.2 (8.6–12.4)
SO$_2$ (ppb)	0.12 (0.04–0.23)	0.09 (0.04–0.19)
O$_3$ (ppb)	39 (35–44)	34 (28–41)
Condensation sink (10^{-3} s^{-1})	2.2 (1.4–3.3)	3.9 (2.7–6.2)
PM$_{10}$ (µg m^{-3})	3.7 (2.3–5.5)	6.6 (4.6–10.5)

399   **Table 2.** Criteria for the NPF forecasts (the source for each data is shown in parenthesis).

NPF forecast	Criteria
NPF day	• Sunny, clear skies (according to weather forecasts) • Low PM_{10} concentration, smaller than 3.7 µg m^{-3} (SILAM) • Low RH during the day, smaller than 45% (SILAM) • SO_2 concentration higher than 0.12 ppb (SILAM) • Airmasses originating from the Arctic Ocean or Northern Atlantic (HYSPLIT trajectories)
Weak NPF / Possibility of NPF / No continuous growth of nucleation mode particles	• Sunny day with some clouds, or partly cloudy (according to weather forecasts) • PM_{10} 3.7–6.6 µg m^{-3} and SO_2 > 0.12 ppb; OR PM_{10} > 6.6 µg m^{-3} and SO_2 > 0.23 ppb (SILAM) • RH during the day 45–76% (SILAM) • Airmasses not coming directly from the west–north-west, or passing over known areas of anthropogenic pollution (HYSPLIT trajectories)
No NPF	• Cloudy day, rain (according to weather forecasts) • High PM_{10}, higher than 6.6 µg m^{-3} (SILAM) • High RH, higher than 76% (SILAM) • SO_2 concentration smaller than 0.09 ppb (SILAM) • Airmasses originating from south (continental Europe) or east, or passing over known areas of anthropogenic pollution (HYSPLIT trajectories)

Table 3. NPF event forecasts (2nd column), and NPF event classification based on measured particle size distributions (3rd column) for each day of the campaign. Class I and II NPF events refer to the classification by Dal Maso et al. (2005). Remarks on the 4th column show the basis for the NPF event forecast.

| Date	Forecast	Classification	Remarks
3 May	Possibility of NPF~~NPF event very probable~~	NPF (class II)	Airmasses coming from north to north-west. Low SO_2 and PM_{10}. Sunny the whole morning until mid-afternoon. RH drops to 25% during the day.
4 May	No NPF	Undefined	Airmasses coming from south-west. Elevated SO_2 and PM_{10} concentrations. Cloudy day with small chance of occasional rain.
5 May	No NPF	No NPF	Airmasses coming from south-west over southern Sweden, turning to west late in evening. Low SO_2 and PM_{10} concentration. Cloudy day with light rain throughout the day until evening. In the evening partly cloudy.
6 May	NPF day	NPF (class I)	Airmasses originating from west and circulating over western Finland. Relatively clean air, SO_2 and PM_{10} concentrations low during morning and increasing towards evening. Sunny until midday, afternoon clouds possible. RH dropping to 40% during the morning.
7 May	~~NPF occuring, but n~~No continuous growth of nucleation mode particles	NPF (class II)	Airmasses coming from south-west over Denmark and southern Sweden. Elevated SO_2 and PM_{10} concentrations. Partly cloudy day. RH dropping to 55%.
8 May	NPF day	NPF (class I)	Westerly airmasses coming over Central Sweden. Low SO_2 and PM_{10} concentrations. Sunny day, warm temperatures (over +15°C). RH dropping to 35% during the morning.
9 May	No NPF	No NPF	Airmasses coming from southwest-south and circulating over southern Finland. Elevated SO_2 and PM_{10} concentrations. Cloudy and rainy day.
10 May	No NPF	No NPF	Airmasses coming from south over Baltic countries and southern Finland. Low SO_2 and somewhat elevated PM_{10} concentrations. Cloudy day with occasional light rain.
11 May	No NPF	Undefined	Airmasses coming from southwest over northern Germany, Denmark and southern Sweden. Low SO_2, somewhat elevated PM_{10} concentrations. Cloudy morning, partly cloudy in afternoon, small chance of rain. Clear skies in the evening.
12 May	~~Chance of w~~Weak NPF	Undefined	Airmasses from south-west over northern France, Denmark and southern Sweden. Elevated SO_2 and relatively low PM_{10} concentrations. Partly cloudy, warm (over +15°C), RH dropping to 45%.
13 May	No NPF	Undefined	Airmasses from south-west over England, Denmark and southern Sweden. SO_2 concentration somewhat elevated, high PM_{10} concentrations. Partly cloudy day, early morning sunny. Warm (over +15°C), RH dropping to 45%.
14 May	Possibility of NPF	Undefined	Airmasses from south-west over Denmark and southern Sweden. Somewhat elevated SO_2 and relatively low PM_{10} concentrations. Possibility for clear sky in the morning, more clouds in the afternoon. Warm day (over +15°C), RH dropping to 40%.

15 May	~~Weak p~~Possibility of NPF	NPF (class I)	Aimasses from south-west over England, Northern Germany and Southern Sweden. Somewhat elevated SO_2. Low PM_{10} concentrations. Partly cloudy early morning (possibility for clear sky), clear sky in the afternoon. Warm (over +15°C). RH dropping to 35% during the day.
16 May	NPF day	NPF (class II)	Airmasses from south-west over Northern Germany and Southern Sweden, circulating over Southern Finland. Low SO_2. Low PM_{10} concentrations. Clear sky in the morning, possibility of some clouds towards afternoon, warm (over +15°C). RH dropping to 35% during the day.
17 May	~~NPF occurring but n~~No continuous growth of ~~the~~ nucleation mode particles	No NPF	Airmasses from south-west over England, Germany, and Southern Sweden, towards afternoon circulating over Northeast Russia. Low SO_2. Slightly elevated PM_{10} concentrations. Partly cloudy, very warm (over +20°C). RH dropping to 45% during the day.
18 May	No NPF	Undefined	Airmasses circulating over Finland and Northwest Russia. High SO_2, elevated PM_{10} concentrations. Cloudy, some rain, warm (over +15°C). RH dropping to 45% during the day.
19 May	No NPF	No NPF	Airmasses coming from east. Low SO_2 and PM_{10} concentrations. Cloudy and some rain, very warm (over +20°C). RH dropping to 50% during the day.
20 May	No NPF	Undefined	Airmasses coming from east. Low SO_2 and PM_{10} concentrations. Cloudy and some rain in the morning, very warm (over +20°C). RH > 90% during the day.
21 May	Possibility of NPF~~, no continuous growth of nucleation mode particles~~	Undefined	Airmasses coming from northeast. Elevated SO_2, low PM_{10} concentrations. Partly cloudy until afternoon (no continuous growth of nucleation mode particles), possibly clear skies in the evening. Very warm (over +20°C). RH dropping to 55% during the day.
22 May	Possibility of ~~weak~~ NPF~~, but no continuous growth of nucleation mode~~	NPF (class II)	Airmasses coming from Arctic Ocean and circulating via northwest Russia. Elevated SO_2 and PM_{10} concentrations. Partly cloudy day, cloudier towards the afternoon (no continuous growth of nucleation mode) with a chance of rain. Warm (over +15°C). RH dropping to 45% during the day.
23 May	Possibility of ~~weak~~ NPF~~, but no continuous growth of nucleation mode~~	NPF (class II)	Airmasses coming from Arctic Ocean and circulating over northwest Russia. Low SO_2 and PM_{10} concentrations. Partly cloudy in the morning, cloudy skies towards the evening (no continuous growth of nucleation mode).
24 May	No NPF	No NPF	Airmasses coming from Arctic Ocean and circulating via northwest Russia. Low SO_2 and PM_{10} concentrations. Cloudy day, chance of rain throughout the day.
25 May	NPF day	NPF (class II)	Airmasses coming from south-east and circulating over Central Finland. Low SO_2 and somewhat elevated PM_{10} concentrations. Sunny day with few clouds, very warm (over +20°C), RH dropping to 35% during the morning.
26 May	NPF day	NPF	Airmasses coming from east and circulating via Northern Finland to Hyytiälä. Low SO_2 and PM_{10} concentrations. Sunny day, some clouds in the afternoon, very warm (over +20°C), RH dropping to 35% during the morning.
27 May	No NPF	Undefined	Airmasses coming from east. High SO_2 and elevated PM_{10} concentrations. Cloudy and a chance of rain in the

|			morning, partly cloudy in the afternoon.
28 May	~~NPF occuring, n~~No continuous growth of nucleation mode particles	No NPF	Airmasses coming from Arctic Ocean via northwest Russia to Hyytiälä. Low SO_2 and PM_{10} concentrations. Partly cloudy in the morning, clear skies towards afternoon. Very warm (over +20°C), RH dropping to 35% during the morning.
29 May	Possibility of NPF	Undefined	Airmasses coming from Arctic Ocean via northwest Russia to Hyytiälä. Low SO_2 and PM_{10} concentrations. Partly cloudy in the morning, clear skies towards the evening. Very warm (over +20°C).
30 May	Possibility of NPF	Undefined	Airmasses coming from Arctic Ocean via northwest Russia to Hyytiälä. Low SO_2 and low PM_{10} concentrations. Partly cloudy, with chances of rain in the evening. Very warm (over +20°C), RH dropping to 45% during the morning.
31 May	No NPF	Undefined	Airmasses coming from Arctic Ocean via northwest Russia to Hyytiälä. Low SO_2 and PM_{10} concentrations. Partly cloudy and chances of rainshowers during the day.
1 June	No NPF	No NPF	Airmasses coming from southeast. Low SO_2 and PM_{10} concentrations. Partly cloudy in the morning, more clouds towards afternoon, chances of light rain in the afternoon the day.
2 June	Possibility of NPF	Undefined	Airmasses coming from southeast. Elevated SO_2 and PM_{10} concentrations. Partly cloudy, very warm (over +20°C), RH dropping to 45% during the morning.
3 June	NPF day	Undefined	Airmasses coming from southeast. High SO_2 and elevated PM_{10} concentrations. Clear skies, some clouds in the afternoon, very warm (over +25°C), RH dropping to 35% during the morning.
4 June	Possibility of ~~weak~~ NPF	Undefined	Airmasses coming from southeast. Low SO_2 and PM_{10} concentrations. Partly cloudy morning, some rain in the afternoon, warm (over +20°C). RH dropping to 35% during the day.
5 June	Possibility of NPF	Undefined	Airmasses coming from southeast. Elevated SO_2 and low PM_{10} concentrations. Clear sky in the morning, partly cloudy starting from noon, possibility of rain in the evening, very warm (over +25°C). RH dropping to 25%.
6 June	Possibility of ~~weak~~ NPF	Undefined	Airmasses coming from east in the morning and turning to northwest in the afternoon. High SO_2 and elevated PM_{10} concentrations. Partly cloudy morning, rain and thunderstorms in the afternoon, warm (over +20°C).
7 June	No NPF	No NPF	Airmasses coming from northeast. Low SO_2 and elevated PM_{10} concentrations. Cloudy, rain and thunderstorms, moderate temperature (over +15°C).
8 June	NPF day	NPF (class II)	Airmasses coming from west, circulating over Sweden. Elevated SO_2 and low PM_{10} concentrations. Almost clear sky in the morning, partly cloudy in the afternoon, warm (over +20°C). RH dropping to 25% during the day.
9 June	No NPF	NPF (class II)	Airmasses coming from west. Low SO_2 and PM_{10} concentrations. Cloudy/partly cloudy, rain in the afternoon, moderate temperature (over +15°C). RH dropping to 45% during the day.

10 June	Possibility of ~~weak~~ NPF	Undefined	Airmasses coming from northwest. Low SO_2 and PM_{10} concentrations. Almost clear sky in the early morning, partly cloudy towards noon, chance of rain in the afternoon, moderate temperature (over +15°C).
11 June	No NPF	No NPF	Airmasses coming from northwest. Elevated SO_2 and PM_{10} concentrations. Partly cloudy, chance of rain, moderate temperature (over +15°C). RH dropping to 50% during the day.

|	"NPF" forecasted (8 days)	"Weak NPF / Possibility of NPF / No continuous growth" forecasted (16 days)	"No-NPF" forecasted (16 days)
NPF day observed (11 days)	6	4	1
Undefined day observed (19 days)	2	10	7
Non-NPF day observed (10 days)	0	2	8
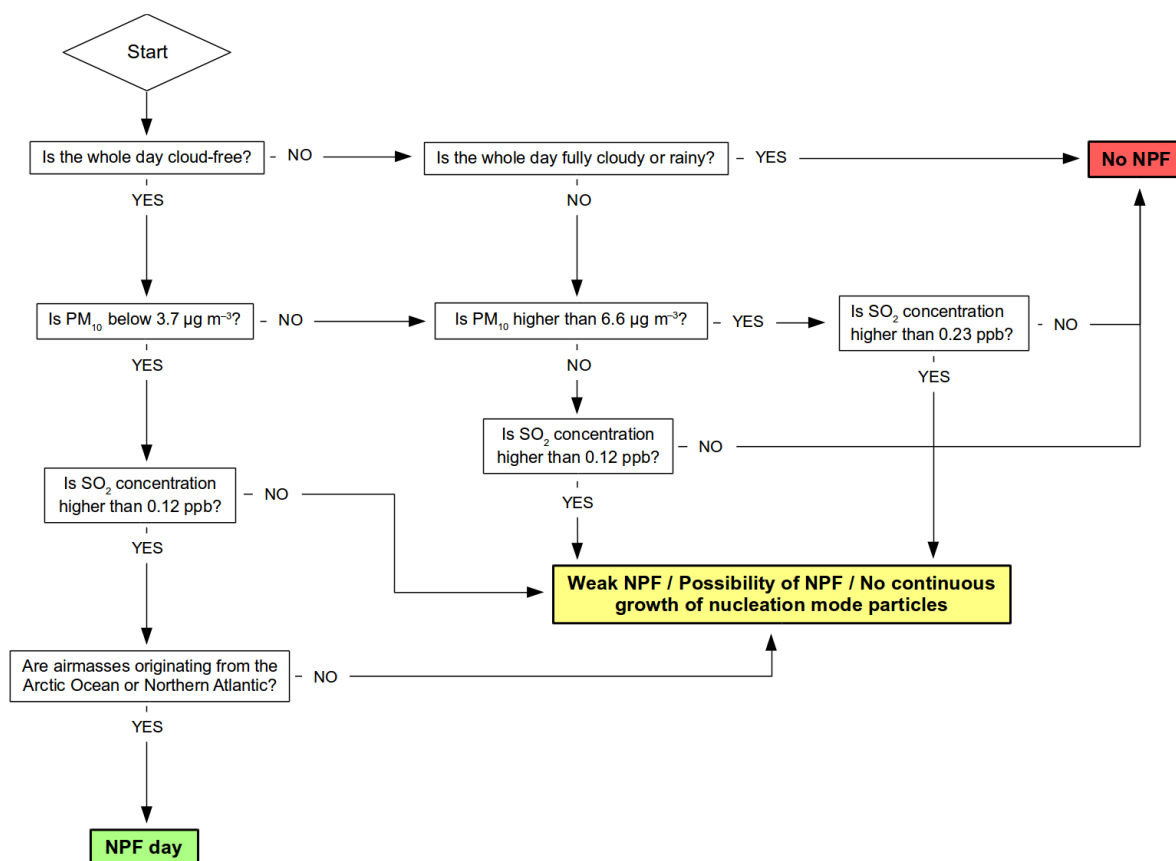
**Figure 1.** Flowchart of the decision making process for the NPF forecasts.
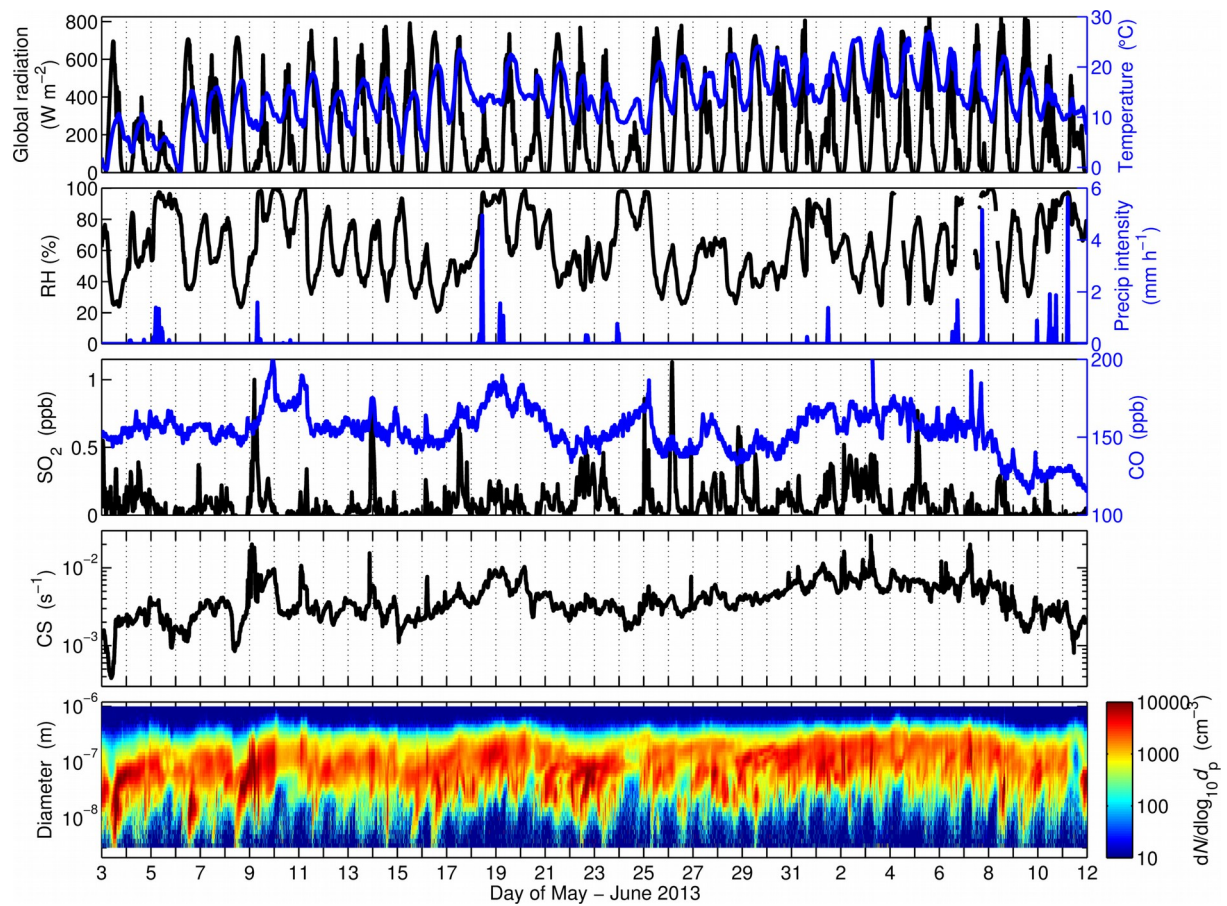
Figure 2. Overview of the meteorological parameters, trace gas concentrations and particle size distributions during the campaign 3.5.–11.6.2013.
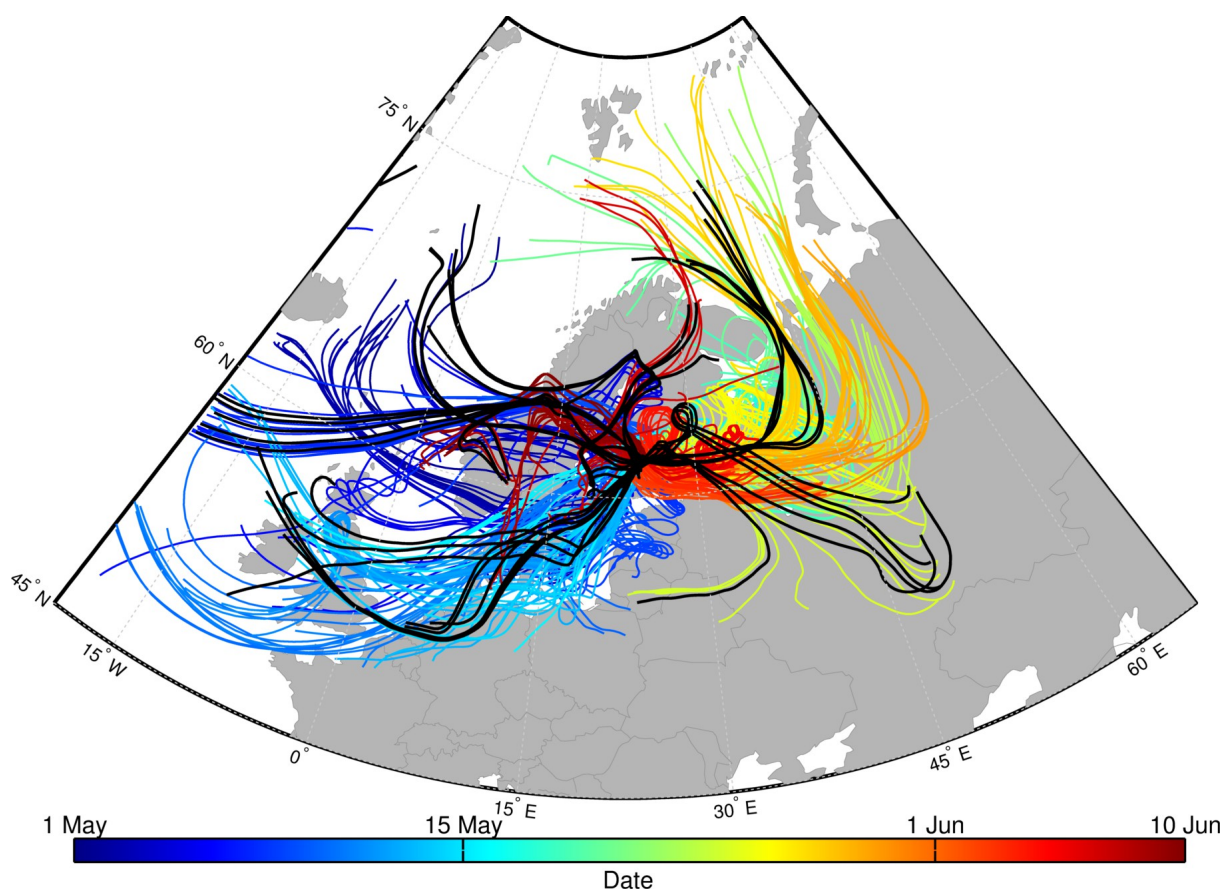
Figure 3. Airmass arrival trajectories to Hyytiälä 3.5.–11.6.2013 calculated using HYSPLIT model. The color indicates the arrival date and each trajectory represents airmass route during 96 hours before arrival. Airmass trajectories arriving on NPF days between 10–14 local time are marked with black lines.
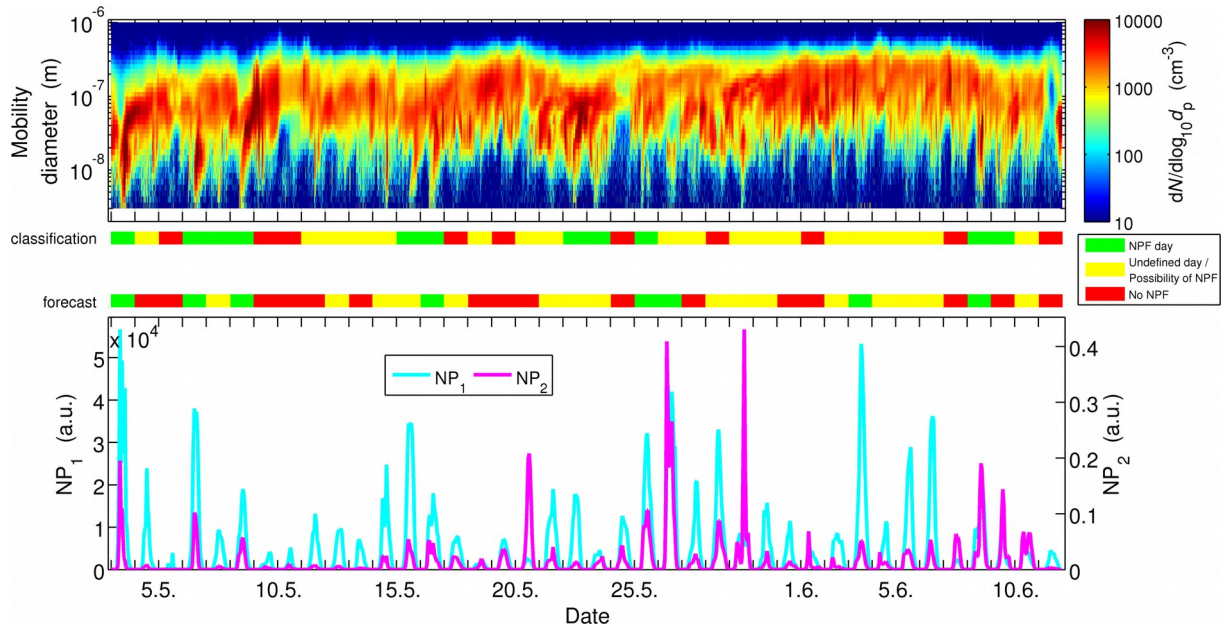
Figure 4. Particle number concentration size distributions (top panel), and nucleation parameters NP_1 and NP_2 (bottom panel) during the campaign time 3.5.–11.6.2013. The colorbars between the panels indicate the NPF forecast and classification: green for NPF days, yellow for weak or possible NPF days, and red for non-NPF days (upper colorbar shows the NPF event classification based on the DMPS data, and lower colorbar the forecast for each day).