Reply to Anonymous Referee #1

Received and published: 18 September 2018

**This paper is an improvement of the Dal Maso et al. (2005) classification of New Particle Formation (NPF) days. That classification was limited to particles about 3 nm in mobility diameter and gave three classes: events, non-events and undefined days. In this paper a new proposal based on the ions participating in the nucleation is given and four classes can be obtained, excluding the undefined days. Other improvements have been incorporated, like the identification of regional and transported events. It has been applied to a large database and compared with the traditional manual procedure obtaining good results. My main concern is that it uses NAIS data, an instrument not very spread in the comC1 munity. This could limit the application of this new methodology. Anyway, this new method to classify NPF could be very useful in the near future. The results are discussed in an appropriate and balance way and the paper is well-structured. It is a significant work, concise and clearly written. I recommend publication in ACP and include some comments.**

We thank Referee #1 for their helpful suggestions. We replied to the comments below. The bold text refers to the referee's comments, and the text in italics are additions to the manuscript. The line numbers mentioned in the text below refer to the ACPD version of the manuscript.

**<u>Specific comments:</u>**

1. **Abstract: please, include the instrument used to obtain the data**

Instrument is added to Abstract as per suggestion from the reviewer.

2. **Line 69: "The station has accumulated 22 years. . ." Although the station is 22 years old, the dataset used in this paper is shorter, 11 years, please, indicate this here.**

We added the following to line 71:

*This study analyzes 10 years of data collected between 2006 and 2016.*

3. **Lines 116-117: "To decide whether the particle growth is observed, particle concentrations in the size range of 7 – 25 nm are examined. These particles represent the growth phase of freshly-formed clusters." Is not there any other possibility? For example, could they come from bigger particles that have suffered shrinkage? It has been observed in some sites particles below 20 nm after shrinkage.**

We thank the reviewer for interesting discussion and potential improvements of our paper. In fact, analyzing 20 years of data from Hyytiälä (Nieminen et al., 2014; Dada et al., 2017) we have not observed shrinkage of NPF in Hyytiälä. It is however a characteristic of certain events which are observed in rather urban environments (Yao et al., 2010; Alonso-Blanco et al., 2015; Salma et al., 2016).

Considering other locations for which this automated method would be applied, and for which shrinkage is observed, we can consider that the method looks for a growing mode (appearance of a peak in 7 – 25 nm) that occurs within 8 hours from the start time of the nucleating mode (Ions 2 – 4 nm). According to previous studies reporting particle shrinkage, it takes longer than 8 hours for the particles to first grow and then shrink back to the 7–25 nm size range, so the automated method would exclude most of these cases. Also, in general, shrunk particles should not be a problem in our case as those only originate due to NPF, so we can consider then part of the growth process and still consider the whole event as NPF. Thus it should not affect our criterion.

In another situation, if there are particles unrelated to NPF (with or without shrinkage), they are too small to have originated far away, so that the main source could be related to traffic (Rönkkö et al., 2017). In the case of Hyytiälä, this source is eliminated due to very small contribution from traffic due to the semi-remote location of our measurement site. In other locations, the influence of traffic is minimized as well, since usually the rush hour occurs before the peak of NPF in the small sizes, so traffic-related are not included in the detection criteria. However, in the case an unidentified plume of pollutants (in the size range of 7 – 25 nm) occurs simultaneously or directly after the NPF peak, the automated method malfunctions. The latter type of misclassification is the major in the discrepancy between manual classification and our automated method. Accordingly, we added a clarification to the reader in the main text describing the reason behind failed statistics to section 3.5:

*Our automated method fails sometimes as the result of the simultaneous appearance of an ion burst and a pollution plume. While the misjudgment of these days as regional events is largely minimized by correcting for the background concentrations of 7–25 nm particles, erroneous classification is still possible in some cases.*

4. **Point 2.4: In order to evaluate the improvement reached with this automated method, how long time do you need to classify a year using your method? And using the manual one?**

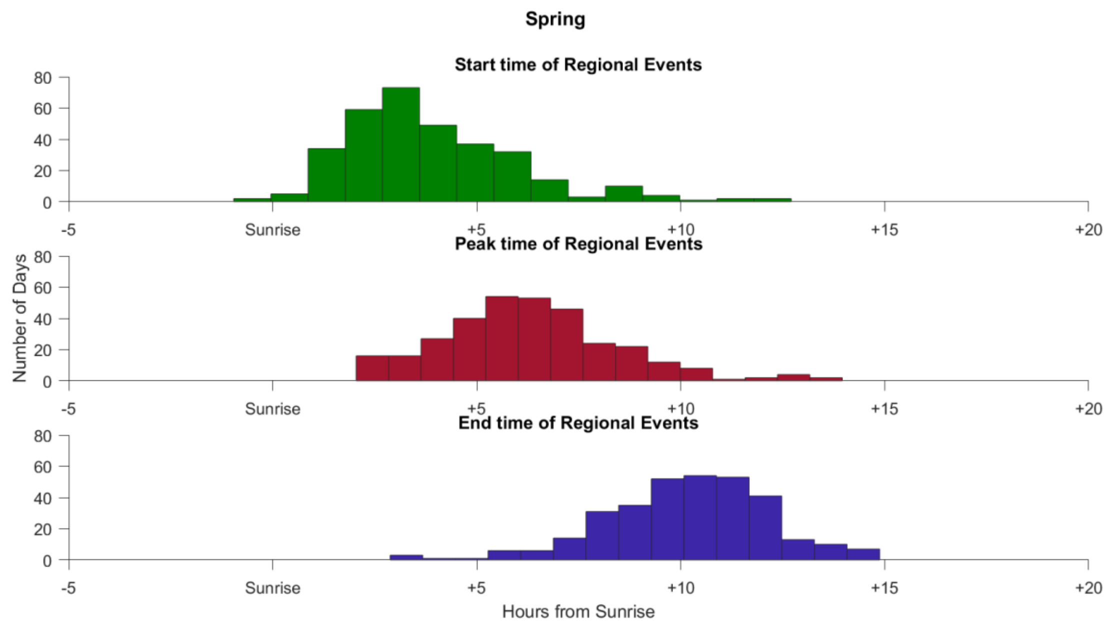
We thank the reviewer for his suggestion to stress the importance of the automated method. Accordingly we added the following sentence to Point 2.4, Line 136.

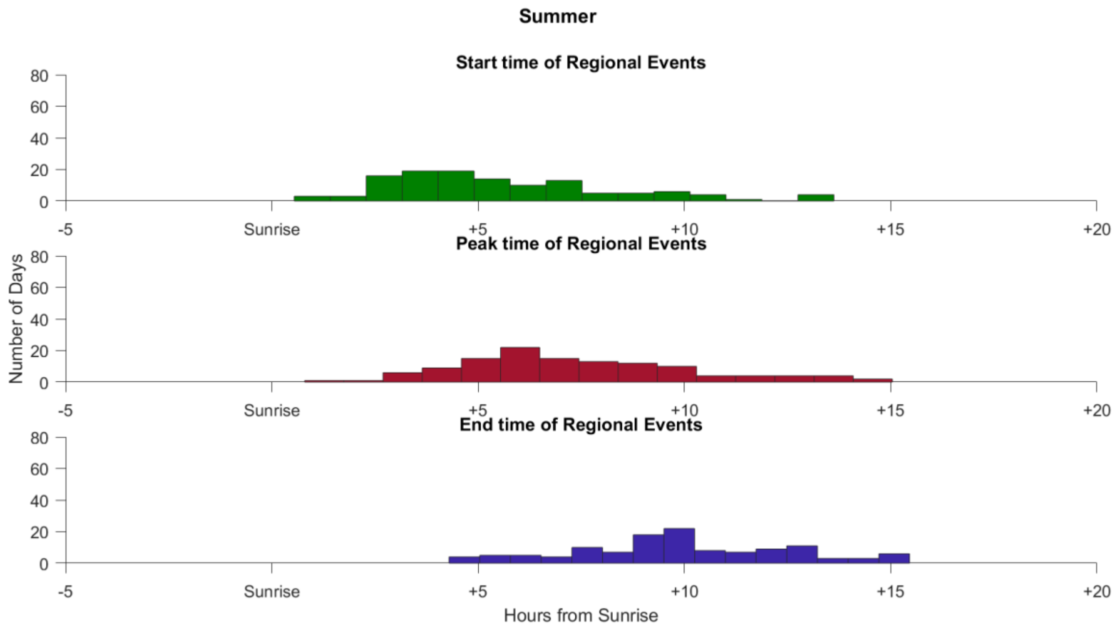
*Once the ion and particle data have been smoothed and precipitation time stamps eliminated, classification of event takes place within a couple of minutes with a click of a button using the new automated method. This can be compared to the manual method which, for classifying one year of measurement data, would require several hours of work ant at least two people to work with it.*

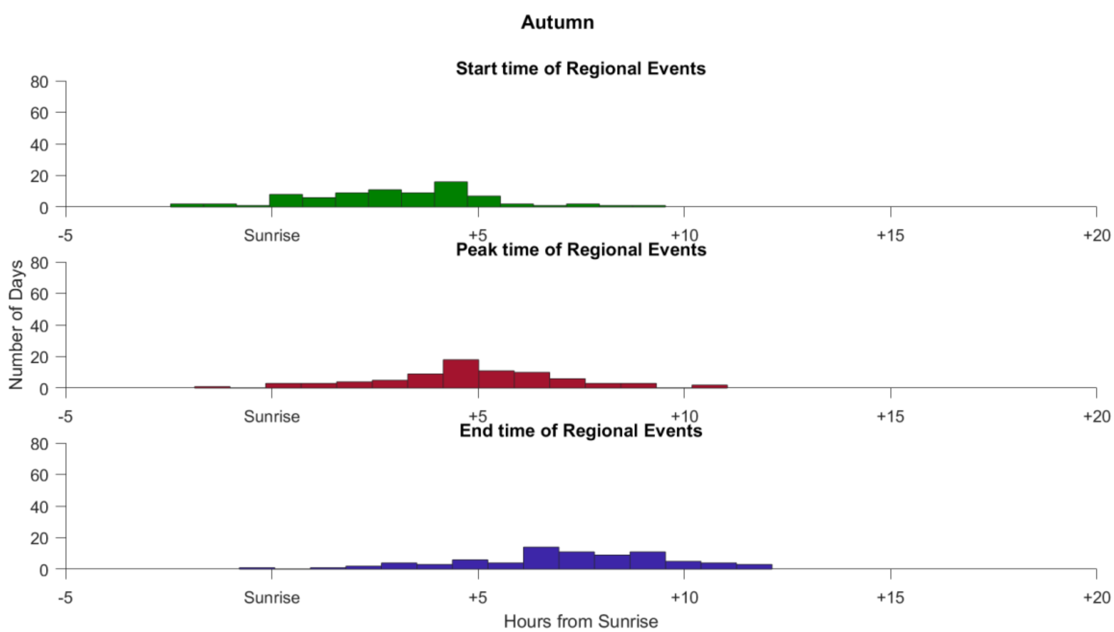
5. **Lines 219-222: "The peak times of the events had the most frequent occurrence at 5 to 6 hours after sunrise, which is between 10:30 and 11:30 local time, complementing our previous assumption that NPF peaks before noon. Finally, the ending times of the events had the most frequent occurrence at 10 hours after sunrise." Do these times depend on the season? Is this possible dependency seen in the data spread shown in the figure 6?**

We thank the reviewer for his suggestion. Indeed there is a large seasonal variability affecting the start, peak and end times of regional events. Below we present the figures separated by season. Since the major fraction of the regional events occurs during spring, the combined figure presented previously in the main text reflects mostly the spring start, peak and end times. Accordingly, we replace Figure 6 in the text with the one of the spring only as it follows nicely from Figure 5. The text of section 3.4 is modified as follows.

*During spring, when most of the NPF events occur, our results (Figure 6) show that indeed RE occur after sunrise and prior to noon, with the maximum number of days occurring between the sunrise and 5 hours past sunrise. The peak times of the events had the most frequent occurrence at 5 to 6 hours after the sunrise, which is between 10:30 and 11:30 local time, complementing our previous assumption that NPF peaks before noon. Finally, the ending times of the events had the most frequent occurrence at 9 to 11 hours after sunrise. During summer the events tend to start, peak and end later than in spring, and they show lower variability in comparison to spring. This observation could be attributed to longer daylight hours and less clouds. Whereas in autumn, the events, start, peak and end earlier than in spring. Exceptionally, during winter, ion concentrations might be affected by the accumulation of snow on or around the inlets. Overall, the variability of the event start, peak and end times can be affected by the solar cycle, degree of cloudiness and seasonality.*




Summer
Start time of Regional Events
Peak time of Regional Events
End time of Regional Events
Number of Days
80
60
40
20
0
-5
Sunrise
+5
+10
+15
+20
Hours from Sunrise


Autumn
Start time of Regional Events
Peak time of Regional Events
End time of Regional Events
Number of Days
80
60
40
20
0
-5
Sunrise
+5
+10
+15
+20
Hours from Sunrise

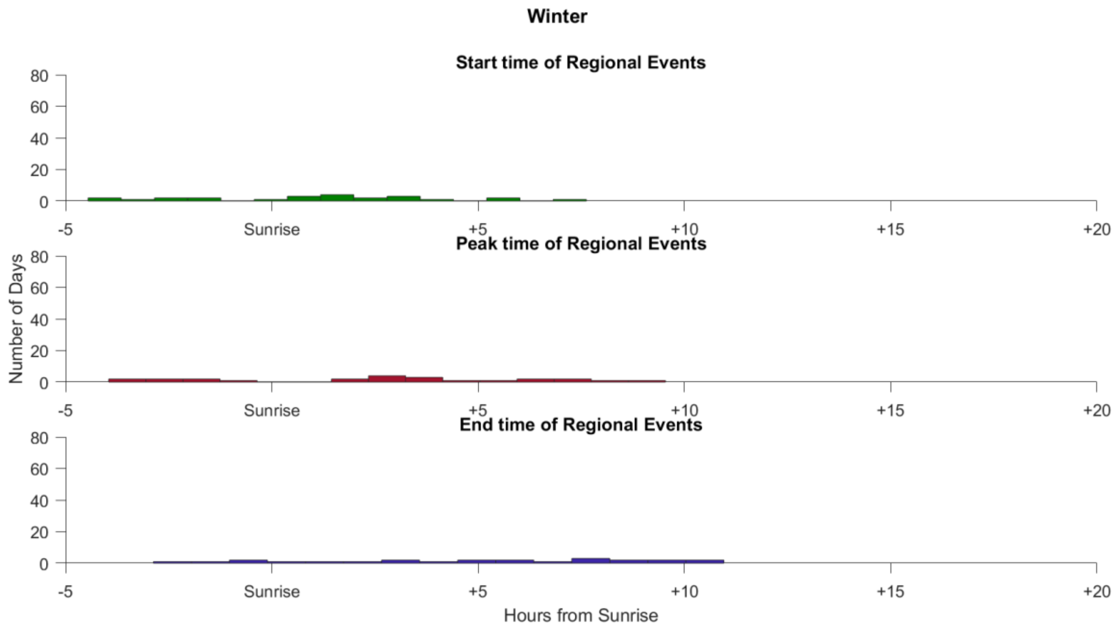



Figure S1. Frequency of days at which regional events start, peak and end past sunrise. For example, most events in Spring start within 5 hours from sunrise.

6. **Lines 230-231: "For example, 65% of the originally-classified event days . . . were found to be RE, 10% were TE and 14% were IB". I don't understand how an event day obtained by the traditional method (Dal Maso et al. 2005) can be an ion burst event with C2 the new methodology. These IB events are characterized by not particle above 7 nm, so they should be considered non-event day by Dal Maso. Is there any reason for this?**

The reviewer is right, during an event day there must be a presence of nucleation mode (< 25 nm) particles for several hours. This mode is identified by an appearance of a peak after subtracting a morning background. In certain cases, when there is a high background during the morning hours, our method does not distinguish a growing mode in the size range 7-25 nm. Therefore, using our automated method, instead of identifying a day as RE, it ends up being classified as IB.

**<u>Minor comments:</u>**

**Check how the references are introduced in the text, sometimes not clear enough. Examples can be found in lines 35 and 48.**

Thanks to the reviewer, we modified the references in the text to a clearer format.

**Line 40: Merikanto et al., 2009 is not included in the reference list**

**Line 41: Idem for Salma et al., 2016**

**Line 111: Idem for Rose et a., 2018**

**Line 118: Idem for Yli-Juuti et al., 2011**

We revised all references used in the manuscript upon request from the reviewer.

**Lines 81-82: It is "the mobility distributions of charged and neutral aerosol particles and clusters in the size range of 0.8–47 nm and 2–42 nm, respectively, were measured with a Neutral cluster and Air Ion Spectrometer". I think it should be "the mobility distributions of neutral and charged aerosol particles and clusters. . ." to correlate with instrument list.**

We modified lines 81-82 based on the reviewer's suggestion to:

*For our proposed automated classification method, the mobility distributions of neutral and charged aerosol particles and clusters in the size ranges of 2–42 nm and 0.8–47 nm, respectively, were measured with a Neutral cluster and Air Ion Spectrometer (NAIS, Airel Ltd., Estonia, (Manninen et al., 2016; Manninen et al., 2009; Mirme and Mirme, 2013) between 2006 and 2016.*

**Line 165: "For 10 years of data (2006 – 2016). . ." Previously it has been said from 2006-2015, I think the right period is until 2016, please, unify the dates.**

Agreed, we are sorry for the typo.

**Lines 243-244: "Also, the growth can be interrupted by a sudden appearance of a cloud (Baranizadeh et al., 2014;Dada et al., 2017)." This idea already appears two lines before. Remove one of the sentences**

We thank the reviewer, and removed the sentence on Lines 243 – 244 and added the citation of Baranizadeh et al., 2014 to the previous sentence. Paragraph becomes as follows:

*The interruption mechanisms may include appearance of clouds (Baranizadeh et al., 2014; Dada et al., 2017), resulting in decreased radiation essential for particle formation and growth (Jokinen et al., 2017), or a change in the origin of arriving air masses from a clean to a rather polluted sector (Sogacheva et al., 2005).*

**Line 302: the DeSerio (2008) reference is incomplete**

Since the DeSerio(2008) refers to a report published online, we replaced the reference with a reference to a book which introduces the definition of Savitzky-Golay smoothing filters (Orfanidis, 1995).

*Orfanidis, Sophocles J. Introduction to signal processing. Prentice-Hall, Inc., 1995.*

References

Alonso-Blanco, E., Gómez-Moreno, F., Núñez, L., Pujadas, M., Cusack, M., Artíñano, B. J. A. C., and Discussions, P.: Towards a first classification of aerosol shrinkage events, 25231-25267, 2015.

Dada, L., Paasonen, P., Nieminen, T., Buenrostro Mazon, S., Kontkanen, J., Peräkylä, O., Lehtipalo, K., Hussein, T., Petäjä, T., Kerminen, V. M., Bäck, J., and Kulmala, M.: Long-term analysis of clear-sky new particle formation events and nonevents in Hyytiälä, Atmos. Chem. Phys., 17, 6227-6241, 10.5194/acp-17-6227-2017, 2017.

Manninen, H. E., Petäjä, T., Asmi, E., Riipinen, I., Nieminen, T., Mikkilä, J., Hõrrak, U., Mirme, A., Mirme, S., and Laakso, L.: Long-term field measurements of charged and neutral clusters using Neutral cluster and Air Ion Spectrometer (NAIS), Boreal Environ. Res, 14, 591-605, 2009.

Manninen, H. E., Mirme, S., Mirme, A., Petäjä, T., and Kulmala, M.: How to reliably detect molecular clusters and nucleation mode particles with Neutral cluster and Air Ion Spectrometer (NAIS), Atmospheric Measurement Techniques, 2016.

Mirme, S., and Mirme, A.: The mathematical principles and design of the NAIS–a spectrometer for the measurement of cluster ion and nanometer aerosol size distributions, Atmospheric Measurement Techniques, 6, 1061-1071, 2013.

Nieminen, T., Asmi, A., Dal Maso, M., Aalto, P. P., Keronen, P., Petäjä, T., Kulmala, M., and Kerminen, V.-M.: Trends in atmospheric new-particle formation: 16 years of observations in a boreal-forest environment, Boreal Env. Res., 19, 2014.

Orfanidis, S. J.: Introduction to signal processing, Prentice-Hall, Inc., 1995.

Rönkkö, T., Kuuluvainen, H., Karjalainen, P., Keskinen, J., Hillamo, R., Niemi, J. V., Pirjola, L., Timonen, H. J., Saarikoski, S., and Saukko, E. J. P. o. t. N. A. o. S.: Traffic is a major source of atmospheric nanocluster aerosol, 114, 7549-7554, 2017.

Salma, I., Németh, Z., Weidinger, T., Kovács, B., Kristóf, G. J. A. C., and Physics: Measurement, growth types and shrinkage of newly formed aerosol particles at an urban research platform, 16, 2016.

Yao, X., Choi, M., Lau, N., Lau, A. P., Chan, C. K., Fang, M. J. A. s., and technology: Growth and shrinkage of new particles in the atmosphere in Hong Kong, 44, 639-650, 2010.

Anonymous Referee #2

Received and published: 8 October 2018

**New particle formation has been demonstrated to play important roles in air quality and climate change. It's essential to classify the new particle formation events and non-events days accurately that can reduce the uncertainty when evaluating the contribution of NPF to aerosol and CCN budget. Previous methods were kind of subjective, and resulted in a poor comparability. This study present an automated method, which is more objective, to classify days into four categories including NPF events, non-events and two classes in between. This automated method was applied in a 10-year NAIS dataset at SMEAR II station. The classification using this methods almost matched the original method, but provided more reliable categories. Therefore, this automated method has the potential to be promoted widely. The manuscript is overall well written C1 ACPD Interactive comment Printer-friendly version Discussion paper and documented. The topic fits well in the scope of ACP. I recommend this manuscript can be published after some revisions.**

We thank Referee #2 for their helpful suggestions. We replied to the comments below. The bold text refers to the referee's comments, and the text in italics are additions to the manuscript. The line numbers mentioned in the text below refer to the ACPD version of the manuscript.

1. **A NAIS is needed to use this "new" method, which is not easy to be promoted. Can it be used with a SMPS or a DMPS? Hyde have SMPS/DMPS dataset, did the author compare the results that using a NAIS with a SMPS/DMPS? Are they identical?**

Our new method uses NAIS in ion mode to account for the process that occurs below 3 nm and represents the early steps of new particle formations which cannot measured by a typical DMPS system (Aalto et al. 2001). While the NAIS is not yet a wide spread instrument, it is important for this study since the ions play a very important role during the early stages of NPF. At these small sizes, the NAIS provides accurate ion concentrations in comparison to DMPS systems which are rather uncertain as seen in the figure below from Wiedensohler et al. (2012). We conclude that, unfortunately, our method is not applicable for data obtained merely from SMPS/DMPS measurements.

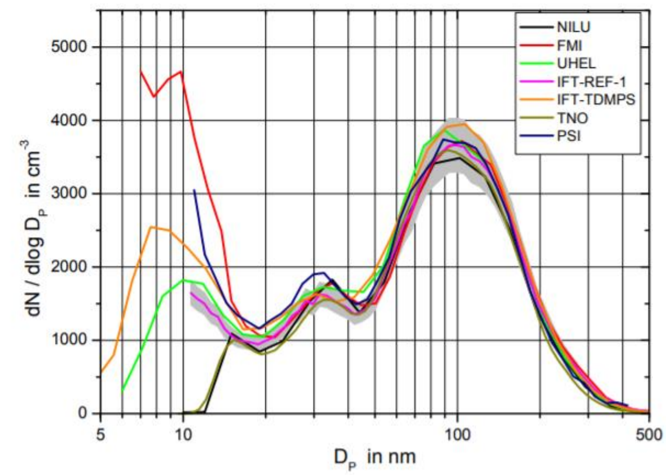



**Fig. 5.** Comparison of the ambient particle number size distributions of the different mobility particle size spectrometers against the reference mobility particle size spectrometer (IFT-REF-1). The shaded area marks the ±10 % range around IFT-REF-1. Internal particle losses in all instruments were corrected using the method of equivalent pipe length.

*Figure S 1 Comparison of particle number size distributions of different mobility particle spectrometers by Wiedensohler et al. 2012.*

2. **Line 154-156: definition of region events is "initiated over a large area including the measurement location and the particles continue to grow to bigger sizes". Since the SMEAR II station is a surface measurement site, how did the author make sure the identified "regional event" occurred over a large area?**

As discussed in a very recent review article (Kerminen et al., 2018), surface measurements at a fixed location can in principle distinguish whether the observed NPF events are regional rather than local in their character. Our classification criteria were designed so that what we call "regional events" would really represent large-scale production of new aerosol particles. The spatial extent of apparently "regional NPF" has been studied in more than 10 individual studies conducted in Europe, North America and China. The general conclusion from these studies is that the spatial extend of regional NPF is typically a few hundreds of km, possibly exceeding 1000 km in some cases (see section 3.1.3 in Kerminen et al., 2018).

3. **Transport events, is there any more evidence to support the definition? Any other possibility that other sources but not NPF contribute to the 7-25 nm particle?**

Our assumption regarding the transported events comes from the observation of downward flux of particles. One example is shown below by Leino et al. (2018). Interestingly, this is one of many similar events for which we are dedicating a complete separate study including more flight measurements above Hyytiälä (Lampilahti et al. 2019, In Prep).

However, there might for sure be contribution of a different source, for example the time period between midnight and 10 am in the figure below, and which is most possibly attributed to a pollution plume carried to our measurement location and which interferes with our automated method.

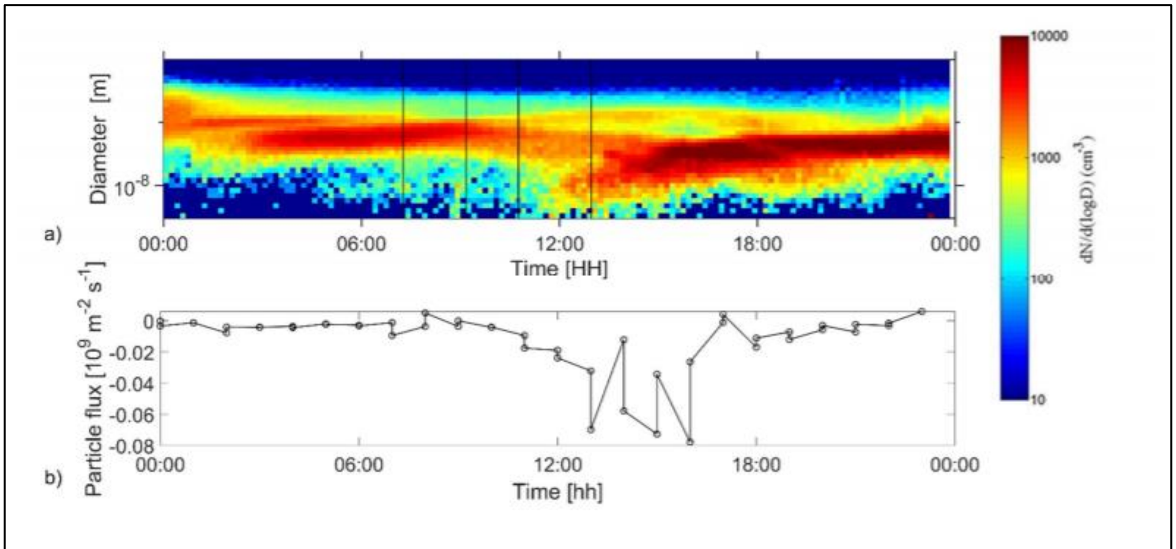


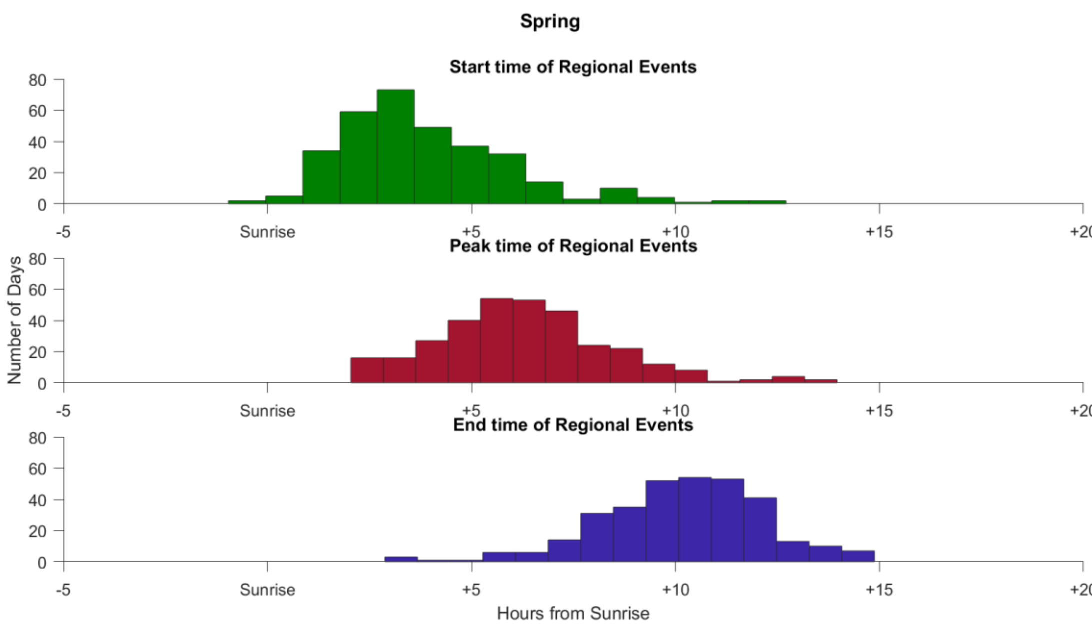

Figure 5. New particle formation event at SMEAR II station in Hyytiälä on 13th August 2015. Panel a) shows the number size distribution measured by Differential Mobility Particle Sizer at ground level inside the forest canopy. Start and end times of two measurement flights were marked by vertical lines in figure. Panel b shows the particle flux measured at 23 m above ground level at the station. Negative particle flux indicates particles flux downwards.

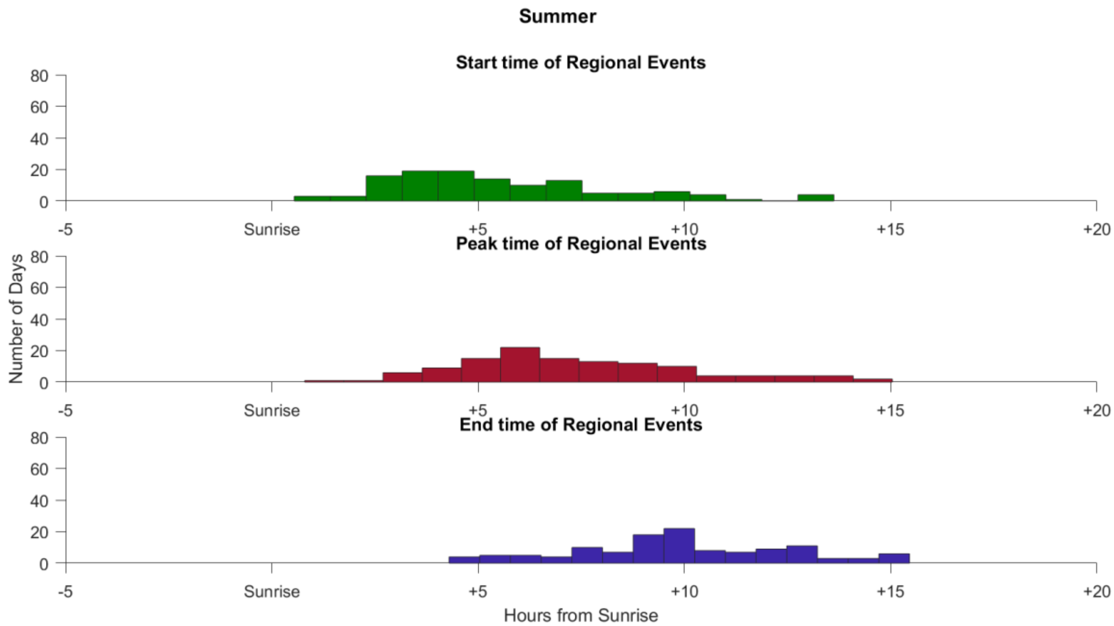
*Figure S 2 New particle formation event at SMEAR II station in Hyytiälä on 13th 550 August 2015. From Leino et al. 2018, ACPD.*

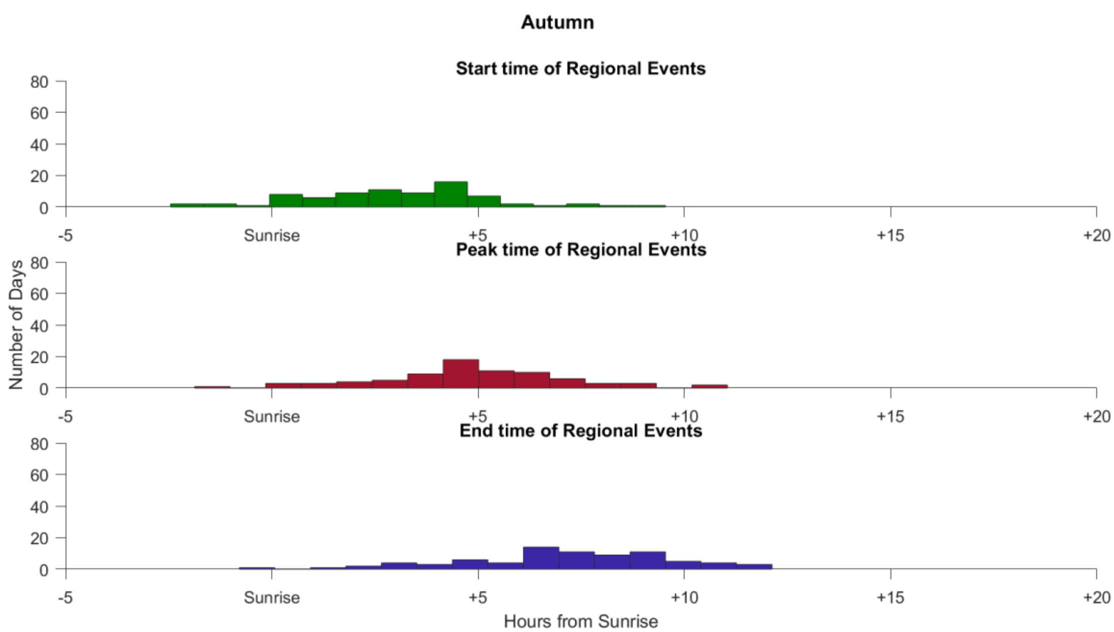
4. **Nighttime events: there are some regional events those were started and stopped before sunrise (Fig. 6)? Is it mean they are typical nighttime NPF events? Did they have the "banana" shape? If not, it means these events were not class A event, but still be defined as regional events (see comment 2)?**

We thank the reviewer for mentioning this point, we agree that there are a couple of days that have an untypical behavior in comparison to the majority of the days studied. Indeed, we do not observe typical 'bananas' that start during nighttime in Hyytiälä (Buenrostro Mazon et al., 2016; Rose et al., 2018). Our automated method matches the manual classification up to 94% and fails mostly during winter, as shown in Figure S3. The ion concentrations might be disrupted during winter due to snow accumulation. Based on the reviewer's comments we re-considered distributing the start, peak and end times analysis over the seasons, and replaced Figure 6 in the text by the similar one from only spring which is representative of the whole data set and still follows nicely from Figure 5. The results show similarity between all days and spring since the majority of RE occur during spring. Also, the redistribution of the plot into seasons confirms that these events that start and end before sunrise, are indeed in winter. The text in section 3.4 is modified as follows.

*During spring, when most of the NPF events occur, our results (Figure 6) show that indeed RE occur after sunrise and prior to noon, with the maximum number of days occurring between the sunrise and 5 hours past sunrise. The peak times of the events had the most frequent occurrence at 5 to 6 hours after the sunrise, which is between 10:30 and 11:30 local time, complementing our previous assumption that NPF peaks before noon. Finally, the ending times of the events had the most frequent occurrence at 9 to 11 hours after sunrise. During summer the events tend to start, peak and end later than in spring, and they show lower variability in comparison to spring. This observation could be attributed to longer daylight hours and less clouds. Whereas in autumn, the events, start, peak and end earlier than in spring. Exceptionally, during winter, ion concentrations might be affected by the accumulation of snow on or around the inlets. Overall, the variability of the event start, peak and end times can be affected by the solar cycle, degree of cloudiness and seasonality.*




Summer
Start time of Regional Events
Peak time of Regional Events
End time of Regional Events
Number of Days
0
20
40
60
80
-5
Sunrise
+5
+10
+15
+20
Hours from Sunrise


Autumn
Start time of Regional Events
Peak time of Regional Events
End time of Regional Events
Number of Days
0
20
40
60
80
-5
Sunrise
+5
+10
+15
+20
Hours from Sunrise

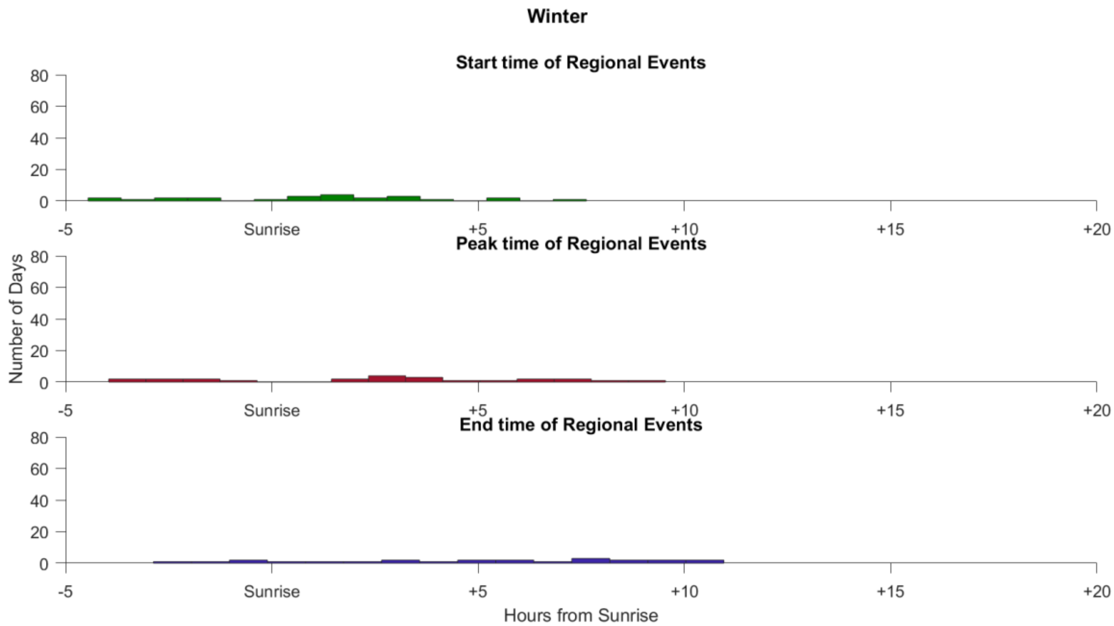
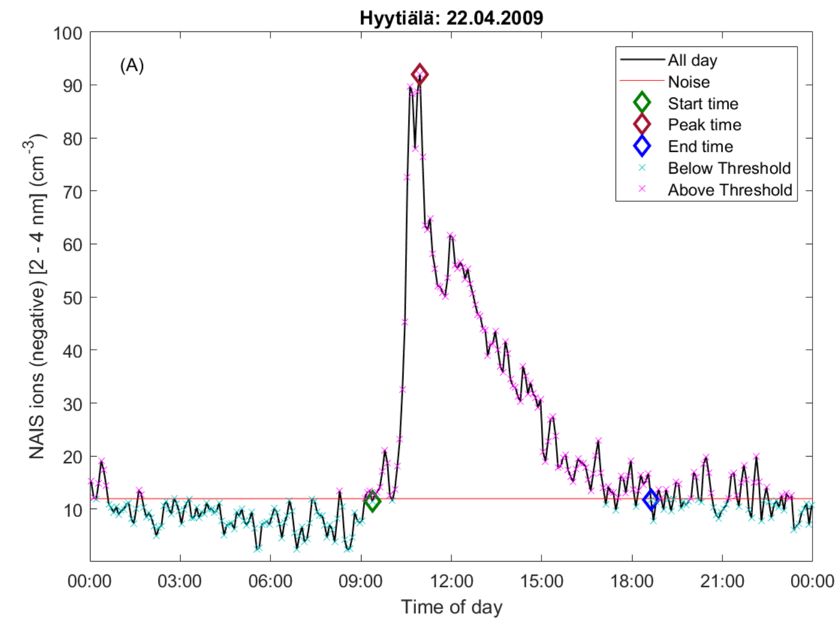
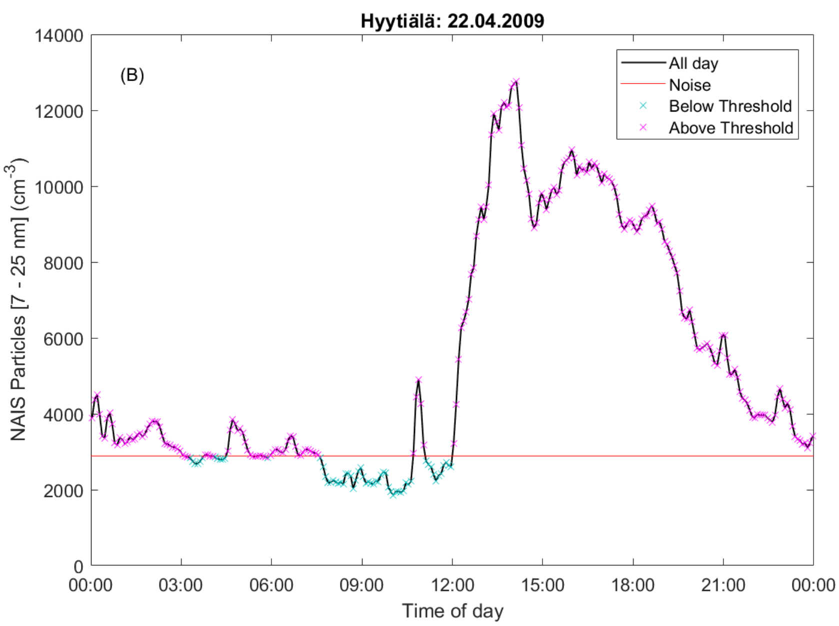



Figure S3. Frequency of days at which regional events start, peak and end past sunrise. For example, most events in Spring start within 5 hours from sunrise.

5. **Figure 2: it's better to give an example to show the variation of 2-4 nm particles and 7-25 nm particles in one event.**

We agree with the reviewer that showing an example of the variation of ions and particles on the same day improves the quality of our method, accordingly we changed the figures to the same day as shown below.


Hyytiälä: 22.04.2009
(A)
All day
Noise
Start time
Peak time
End time
Below Threshold
Above Threshold
NAIS ions (negative) [2 - 4 nm] (cm-3)
Time of day


Hyytiälä: 22.04.2009
(B)
All day
Noise
Below Threshold
Above Threshold
NAIS Particles [7 - 25 nm] (cm-3)
Time of day

**References**


Buenrostro Mazon, S., Kontkanen, J., Manninen, H. E., Nieminen, T., Kerminen, V.-M., and Kulmala, M.: A long-term comparison of nighttime cluster events and daytime ion formation in a boreal forest, Boreal Env. Res., 21, 242-261, 2016.

Carnerero, C., Pérez, N., Reche, C., Ealo, M., Titos, G., Lee, H.-K., Eun, H.-R., Park, Y.-H., Dada, L., and Paasonen, P. J. A. C. P. D., https://doi. org/10./acp--173, in review: Vertical and horizontal distribution of regional new particle formation events in Madrid, 2018.

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

Kerminen, V.-M., Chen, X., Vakkari, V., Petäjä, T., Kulmala, M., and Bianchi, F. J. E. R. L.: Atmospheric new particle formation and growth: review of field observations, 13, 103003, 2018.

Leino, K., Lampilahti, J., Poutanen, P., Väänänen, R., Manninen, A., Mazon, S. B., Dada, L., Nikandrova, A., Wimmer, D., Aalto, P. P., Ahonen, L. R., Enroth, J., Kangasluoma, J., Keronen, P., Korhonen, F., Laakso, H., Matilainen, T., Siivola, E., Manninen, H. E., Lehtipalo, K., Kerminen, V.-M., Petäjä, T., and Kulmala, M.: Vertical profiles of sub-3 nm particles over the boreal forest Atmos. Chem. Phys. Discuss. (Submitted), 2018.

Németh, Z., and Salma, I.: Spatial extension of nucleating air masses in the Carpathian Basin, J Atmospheric Chemistry

Physics, 14, 8841-8848, 2014.

Rose, C., Zha, Q., Dada, L., Yan, C., Lehtipalo, K., Junninen, H., Mazon, S. B., Jokinen, T., Sarnela, N., Sipilä, M., Petäjä, T., Kerminen, V.-M., Bianchi, F., and Kulmala, M.: Observations of biogenic ion-induced cluster formation in the atmosphere, Science Advances, 4, 10.1126/sciadv.aar5218, 2018.

Salma, I., Németh, Z., Kerminen, V.-M., Aalto, P., Nieminen, T., Weidinger, T., Molnár, Á., Imre, K., Kulmala, M. J. A. C., and Physics: Regional effect on urban atmospheric nucleation, 16, 8715-8728, 2016.

Vana, M., Komsaare, K., Hõrrak, U., Mirme, S., Nieminen, T., Kontkanen, J., Manninen, H. E., Petäjä, T., Noe, S. M., and Kulmala, M. J. B. E. R.: Characteristics of new-particle formation at three SMEAR stations, 2016.

Wiedensohler, A., Birmili, W., Nowak, A., Sonntag, A., Weinhold, K., Merkel, M., Wehner, B., Tuch, T., Pfeifer, S., and Fiebig, M. J. A. M. T.: Mobility particle size spectrometers: harmonization of technical standards and data structure to facilitate high quality long-term observations of atmospheric particle number size distributions, 5, 657-685, 2012.

# Refined classification and characterization of atmospheric new particle formation events using air ions

Lubna Dada[1], Robert Chellapermal[1], Stephany Buenrostro Mazon[1], Pauli Paasonen[1], Janne Lampilahti[1], Hanna E. Manninen[1,2], Heikki Junninen[1,3], Tuukka Petäjä[1,4], Veli-Matti Kerminen[1], and Markku Kulmala[1,4,5]

[1] Institute for Atmospheric and Earth System Research, University of Helsinki, Helsinki, Finland
[2] Experimental Physics Department, CERN, 1211 Geneva, Switzerland
[3]Institute of Physics, University of Tartu, Ülikooli 18, EE-50090 Tartu, Estonia
[4]Aerosol and Haze Laboratory, Beijing Advanced Innovation Center for Soft Matter Science and Engineering, Beijing University of Chemical Technology, Beijing, China
[5]Joint International Research Laboratory of Atmospheric and Earth System Sciences, Nanjing University, Nanjing, China
*Correspondence to*: Lubna Dada (lubna.dada@helsinki.fi)

**Abstract.** Atmospheric new particle formation (NPF) is a world-wide observed phenomenon that affects the human health and the global climate. With the growing network of global atmospheric measurement stations, efforts towards investigating NPF have increased. In this study, we present an automated method to classify days into four categories including NPF events, non-events and two classes in between, which then ensures the reproducibility and minimizes the man-hours spent on manual classification. We applied our automated method to 10 years of data collected at the SMEAR II measurement station in Hyytiälä, southern Finland using a Neutral and Air Ion Spectrometer (NAIS). In contrast to the traditionally-applied classification methods which categorize days into events, non-events and ambiguous days as undefined days, our method is able to classify the undefined days as it accesses the initial steps of NPF at sub-3 nm sizes. Our results show that on ~24% of the days in Hyytiälä, a regional NPF event occurred and was characterized by a 'nice weather' and favorable conditions such as a clear sky and low condensation sink. Another class found in Hyytiälä is the transported event class, which seems to be NPF carried horizontally or vertically to our measurement location and it occurred on 17% of the total studied days. Additionally, we found that an ion burst, where the ions apparently fail to grow to larger sizes, occurred on 18% of the days in Hyytiälä. The transported events and ion bursts were characterized by less favorable ambient conditions than regional NPF events, and thus experienced interrupted particle formation or growth. Non-events occurred on 41 % of the days and were characterized by a complete cloud cover and high relative humidity. Moreover, for the regional NPF events occurring at the measurement site, the method identifies the start time, peak time and end time, which helps us focus on variables within an exact time window to better understand NPF in a process level. Our automated method can be modified to work in other measurement locations where NPF is observed.

Keywords: NPF events, air ions, intermediate ions, boreal forest

## 1 Introduction

New particle formation (NPF) is an atmospheric phenomenon that results in a big addition to aerosol load in the global troposphere (Spracklen et al., 2010; Kerminen et al., 2018). NPF is observed frequently in different environments around the globe, ranging from pristine locations (Siberia –Kulmala et al. (2011); Asmi et al. (2016)), to boreal forests (Hyytiälä -Kulmala et al. (2013); Nieminen et al. (2014)), tropical forests (Amazon - Artaxo et al. (2013); Wimmer et al. (2018))), mountain tops (Jungfraujoch–Bianchi et al. (2016)), semi-polluted cities (European cities -Manninen et al. (2010) ) and even heavily polluted mega cities (China -Kulmala et al. (2016); (2017); Wang et al. (2017). The freshly formed particles that grow to larger sizes contribute largely to the cloud condensation nuclei load in the atmosphere (Merikanto et al., 2009; Kerminen et al., 2012; Salma et al., 2016) and thus indirectly affect the climate (IPCC, 2013).

In order to comprehend the phenomenon of NPF in a specific location, we first need to understand its frequency and characteristics as well as particle formation and growth rates associated with it. With the growing number of global stations (Kulmala, 2018), an automatic method is needed to classify the days into events and non-events. In addition to minimizing the effort of manual event classification, an automated method tends also to reduce any human error. In this study, we present an automated method which classifies days into four classes according to the observed characteristics of 2-4 nm sized air ions and 7-25 nm sized particles. The original classification method of days as events, non-events and undefined days was proposed by Dal Maso et al. (2005), and later modified by Kulmala et al. (2012), and is based on particle measurements starting from about 3 nm in particle mobility diameter, thus missing the initial steps of NPF. With the increased development of instrumentation, we are able to access sub-3 nm clusters and refine our classification method to account for the very initial steps of NPF. The classification proposed here divides days into regional events, transported events, ion bursts and non-events, thus excluding any 'undefined' days, which minimizes the number of days usually excluded from further data analysis. Furthermore, our automated method identifies the start, peak and end time of daytime regional events or ion bursts. By identifying the start and end times, we are able to concentrate on the conditions present during the actual NPF time window.

Our study focuses on the NPF occurring in Hyytiälä, a boreal forest site in southern Finland where the SMEAR II (Station for Measuring Forest Ecosystem-Atmosphere Relations) measurement station is located (Hari and Kulmala, 2005). The dataset collected at the station sums up more than 22 years of particle, meteorological and gas data, making extensive analyses of NPF and related parameters possible. Besides studying NPF occurrence in Hyytiälä, our method can be applied to other locations where NPF is observed, enabling scientists studying particle formation to focus on specific time windows by which active NPF occurs. Our specific aims in this study are i) to automatically classify days in Hyytiälä according to their initial NPF steps, ii) to minimize the number of undefined days by refining the classification, iii) to investigate different characteristics of classified days, iv) to identify the start, peak and end times of regional events and, thereby, v) to create a time series which allows us to focus on the exact time period during which a regional new particle formation event has occurred.

## 2 Materials and Methods

### *2.1 Measurement location*

The main results of our study are based on the measurements collected at the SMEAR II~~)~~ station located in the boreal forest site in Hyytiälä, Southern Finland (61°51'N, 24°17'E, 181 m a.s.l). The station has accumulated 22 years of comprehensive measurements including particle, radiation, gas, meteorological and complementary data. This study analyzes 10 years of data collected between 2006 and 2016. The location is considered a semi-clean boreal forest environment as it is far from anthropogenic pollutants (Asmi et al., 2011) and thus represents the northern-hemisphere boreal forests. A more detailed description of the site and the ongoing measurements can be found in Hari and Kulmala (2005) and Nieminen et al. (2014).

### *2.2 Instrumentation*

The traditional classification of days as NPF events and non-events follows the method proposed by Dal Maso et al. (2005); Kulmala et al. (2012). For this classification method, the particle number-size distributions measured with a twin-DMPS (Differential Mobility Particle Sizer) system (Aalto et al., 2001), were used. The twin DMPS system measured the aerosol number-size distribution over the size range 3-500 nm until 2004 and over the size range 3-1000 nm from 2005 onwards. The DMPS measurements are also used to calculate the condensation sink (CS) which is the rate at which non-volatile vapors condense onto a pre-existing particles (Kulmala et al., 2012).

For our proposed automated classification method, the mobility distributions of ~~charged and neutral~~neutral and charged aerosol particles and clusters in the size ranges of ~~0.8–47 nm and~~ 2–42 nm and 0.8–47 nm, respectively, were measured with a Neutral cluster and Air Ion Spectrometer (NAIS, Airel Ltd., Estonia, (Manninen et al., 2009; Mirme and Mirme, 2013; Manninen et al., 2016))) between 2006 and 2016~~5~~. No measurements using the NAIS were made during year 2008 when the instrument was used for an intensive campaign. Particle and air ion data are available in two-minute time steps.

The air temperature and the relative humidity are measured with 4-wired PT-100 sensors and relative humidity sensors (Rotronic Hygromet MP102H with Hygroclip HC2-S3, Rotronic AG, Bassersdorf, Switzerland) on a mast at a height level of 16.8 m, respectively. The temperature and relative humidity data are provided as 30-minute averages. Solar radiation in the wavelengths of global radiation (0.30-4.8 µm) is monitored using pyranometers (SL 501A UVB, Solar Light, Philadelphia, PA, USA; Reeman TP 3, Astrodata, Tõravere, Tartumaa, Estonia until June 2008, and Middleton Solar SK08, Middleton Solar, Yarraville, Australia since June 2008) above the forest at 18 m. We used global radiation data for calculating the cloudiness parameter ($P$), which is the ratio of global radiation to theoretical maximum radiation arriving at Hyytiälä, by following the method proposed by Dada et al. (2017). Values of $P \leq 0.3$ represent a complete cloud cover while values of $P \geq 0.8$ can be considered to represent clear-sky conditions.

### *2.3 Event classification decision tree*

Based on the concentrations of 2 – 4 nm ions, we are able to detect the initial steps of cluster formation (see Leino et al. (2016)),which would not be possible using the DMPS system alone and the traditional classification. This small size window available from the NAIS operating in ion mode gives an additional opportunity to investigate sub-3 nm clusters. Accordingly, we are able to estimate whether a regional NPF event occurred within the air mass in which the observations were made, or elsewhere and then carried to our measurement location. Similarly, undefined days are identified based on their sub-3 nm characteristics. We present in Figure 1 our refined classification decision tree and apply it to Hyytiälä data in this study. In order to attain this classification, we rely on the initial steps of cluster formation and their further growth, which we monitor using an automatic method. Since in our study we are interested in daytime NPF, we chose the time window between 06:00 and 19:00 when monitoring aerosol number concentrations. However, the automated method can be tweaked to include evening or night time event classification in places where these event types are present.

Our decision tree (Figure 1) first examines 2–4 nm ion concentrations representing the initial step of new particle formation. A notable increase in their concentration is interpreted as ion clustering on site. To be accounted as an increase, the number concentration of ions after 06:00 must increase above a relative threshold and persist for more than 1 hour. This threshold is calculated from ion concentration averaged over the time period 00:00–04:00 multiplied by a scaling factor (Figure 2A); we chose this time window as background as it is outside the time window when night time ion clusters are observed (Buenrostro Mazon et al., 2016; Rose et al., 2018). To be accounted as a notable increase past the threshold value, a concentration of 20 ions/cm$^3$ should be reached and should last for at least 1 hour. We chose the aforementioned value as it has been found to be an indicator for NPF in Hyytiälä (Leino et al., 2016). If this criterion is met, these ions are expected to either grow into bigger sizes and lead to regional NPF events (RE), or fail to grow further, in ~~which~~ this case the events are identified as ion bursts (IB) that do not form new particles.

To decide whether the particle growth is observed, particle concentrations in the size range of 7 – 25 nm are examined. These particles represent the growth phase of freshly-formed clusters. Since in Hyytiälä growth rates of 4 – 7 nm particles is reported to lie between 0.8 and 17 nm/h (Average 3.8 nm/h ) (Yli-Juuti et al., 2011), we considered a time delay of 1 to 8 hours between the initial increase of ion (2 – 4 nm) concentrations and particle (7 – 25 nm) concentrations. To be considered as an increase, the particle number concentration should exceed a relative threshold which in this case is the number concentration averaged over the time period of 03:00–05:00 (Figure 2B). We determined the background time window by comparing the automatic method to a manual classification that we performed for the years 2013-2014 from our data set. The increase in concentration should last for ~1.5 hr (100 minutes) and reach a peak of at least 3000 particles/cm$^3$. On one hand, if both 2 - 4 nm ions and 7 – 25 nm particles are present, the time period is considered as a regional event (RE). On the other hand, if the 2 - 4 nm ions are present but they do not grow to form 7 – 25 nm particles, the time period is classified as an ion burst (IB). Moreover, if 2 – 4 nm ions are not present, but we observe an increase in the particles, this leads to the assumption that the NPF event did not occur at the measurement location but was carried horizontally or vertically to our site (Leino et al., 2018). The latter has been previously described as a tail event (Buenrostro Mazon et al., 2009) or a transported event (TE). However, if neither criterion is met, which means that neither 2 – 4 nm ions nor 7 – 25 nm particles are present in sufficient concentrations, the time period is then classified as a non-event (NE).

### *2.4 Description of the automated method*

Our automatic method selects the start time, peak time and end time of negative NAIS ions in the size range 2 – 4 nm. The growth to an event is confirmed by an accompanying peak in the 7 - 25 nm particles measured by the NAIS. The outcome of the automatic method is the classification of days into the four classes, as well as a time series that identifies the time period of regional events and ion bursts in Hyytiälä (Pathways RE and IB in Figure 1). Once the ion and particle data are smoothed and the precipitation time stamps are eliminated, using the new automated method, the classified time series is generated within couple of minutes with a click of a button, in comparison to the manual method which could use several hours and at least 2 people in order to classify one year of data.

First, to investigate the appearance of 2 – 4 nm ions, the precipitation time stamps are excluded from our analysis as they interfere with the ion data (Leino et al., 2016), resulting in misinterpretations. After that, the ion concentrations are smoothed using Savitsky-Golayfilter (Orfanidis, 1995). We then search for an increase in the ion concentration that lasts for 12 consecutive points (5 minutes each) above a threshold value and reaches values greater than 20 cm$^{-3}$ (Leino et al., 2016). A maximum of 3 drops below the threshold value are allowed (Figure 2A). Finally, the method looks for a peak in the 7 - 25 nm particle concentration to identify the appearance of a growth phase (Figure 2B). The peak requires 15 consecutive points (5 minutes each) having concentrations larger than the threshold value and reaching a value larger than 3000 cm$^{-3}$. Also, a maximum of 3 drops below the threshold value are allowed. Accordingly, each time stamp is classified.

### *2.5 Start time, peak time and end time determination*

The start time, peak times and end times for regional events and ion bursts are defined based on the 2 – 4 nm ion concentration as follows: i) The start time is the first crossing of the threshold line which lasts for more than 12 consecutive points, ii) the peak time is when the concentration reaches the maximum and iii) the end time is the first trough after crossing the threshold line into lower concentrations which remains below the threshold for more than 3 consecutive points. An example day is demonstrated in Figure 2A. The threshold is taken as the 2 – 4 nm ion concentration averaged over the time period 00:00–04:00 multiplied by a scaling factor of 7. Our scaling factor was determined after we did a comparison with the manual classification of the data for the years 2013-2014.

## 3 Results and Discussion

### *3.1 Event Classification*

Our classification categorizes the days in Hyytiälä into four different categories following the pathway chart in Figure 1. Type RE, or regional NPF events, are those which are initiated over a large area including the measurement location and the particles continue to grow to bigger sizes. The type TE, or transported events (also known as tail events by Buenrostro Mazon et al. (2009)), are events whose beginning is not detected as it does not occur at the immediate vicinity of our measurement site. Such events could be attributed to events that were initiated outside our measurement site and transported to Hyytiälä (Leino et al. 2018). The aforementioned hypotheses could explain the observation that TE typically occur at around midday or later in the afternoon, while RE tend to occur concurrent with sunrise. The type IB, or ion bursts, are attempts of NPF, during which clusters form in Hyytiälä, however, they do not grow beyond a few nanometers in diameter. Changes in atmospheric conditions that could cause the limited, or interrupted, growth of the clusters are assessed in more detail in section 3.3. Finally, non-events (NE) are days for which we do not observe a forming mode of 2 – 4 nm ions nor a growing mode of 7 - 25 particles.

### *3.2 Frequency of Events*

For 10 years of data (2006 – 2016), excluding the days with missing NAIS data when the instrument was under maintenance or on campaigns, we classified a total of 2134 days. Using our refined classification method, we were able to classify the days into 4 categories as follows (Figure 3): 551 RE (24%), 410 TE (18%), 415 IB (18%) and 938 NE (40%). This refined classification is able to classify all days into categories and thus eliminate the undefined days that usually constitute around 40% of all the days in our location (Dal Maso et al., 2005;Buenrostro Mazon et al., 2009).

Moreover, we studied the inter-annual variation of each of the classes (Figure 4A). In general, RE constitute 20-30% of the total classified days. In 2006, the measurement started in September, which explains a lower fraction of RE. The gap in the analysis in 2008 is explained by a campaign during which the NAIS data is not available (Manninen et al., 2010). The data in 2009 includes data from spring only, which explains the high frequency of RE in 2009. While we can observe changes in the frequency of RE between the years, no clear trend exists. The annual variation of TE follows that of RE, also with no specific trend over the years. The type IB appears to have an almost constant fraction over the years. Finally, NE constitute between 40 and 50% of the days, except in 2009 which has the bias for spring favoring RE.

The monthly variation of RE follows the typical yearly cycle of NPF, with a peak in spring, followed by a smaller peak in autumn (Dal Maso et al., 2005; Nieminen et al., 2014; Dada et al., 2017). Interestingly, the refined classification shows that the events occurring in spring are mostly RE while those in autumn are dominated by TE. Additionally, RE rarely occur in winter, appearing on less than 5% of the days. IB have a steady 10-20% occurrence during the year. Finally, NE occur on 60 to 70% of winter days and less than 30% during spring. Interestingly, while previously it was understood that summer is dominated by NE (Nieminen et al., 2014;Dada et al., 2017), the refined classification shows that both TE and IB are frequent during summer, complementing observations by Buenrostro Mazon et al. (2009~~11~~) who reported 'failed events' during summer.

### *3.3 Characteristics of RE, TE, IB and NE*

For a regional event to take place, favorable conditions need to be present. These include a low condensation sink, low relative humidity, moderate temperature and plenty of radiation available during a clear sky (Dada et al., 2017; Hyvönen et al., 2005; Nieminen et al., 2014; Nieminen et al., 2015). In Figure 5, we present the characteristics of each type of event classified in terms of Condensation Sink (CS), relative humidity (RH), Temperature ($T$) and Cloudiness ($P$). The data in the plots represent half-hour averages of each variable between 7:00 and 12:00 during spring (March – May). We chose this season in order to capture the maximum NPF events and this time window in order to be consistent between all four studied classes. As expected, the median CS observed on RE was 1.7 x $10^{-3}$ $s^{-1}$ which is a factor of 2 lower than CS observed on TE days or on NE days (3 x $10^{-3}$ $s^{-1}$). To our understanding, high CS inhibits NPF, so that its higher values during the days classified as TE forbid the initial formation of particles at the measurement site. IB, on the other hand, are potential regional events whose growth has been interrupted. Since the median CS during IB was not high (2.5 x $10^{-3}$ $s^{-1}$), it does not explain the discontinuous growth of the clusters during these events. We proceed to study the effect of $T$ on the occurrence of each class of events. Since the data in Figure 5 are measurements during spring, the median value of temperature (2-7 ºC) was rather similar on all days and no specific trend or exception could be found.

In addition to CS and *T*, RH and cloudiness (*P*) play an important role in the occurrence of NPF (Dada et al., 2017; Hamed et al., 2011). A regional NPF event is more likely to occur on a clear-sky day rather than on a cloudy day. This conclusion is demonstrated nicely in Figure 5 which shows that the median value of *P* was close to 0.8 on the RE days and closer to 0.3 on NE day. TE usually took place when the conditions within the boundary layer were not favorable for a regional NPF to occur. However, the particle growth was much less sensitive to environmental conditions: a particle growth was often observed during all times of day and in every season, also on days (and nights) when NPF did not take place (Paasonen et al., 2018). Combined with a higher CS, the value of *P* was much lower on TE days than on RE days, describing a semi-cloudy day unfavorable for NPF to occur within the boundary layer, which could result in the occurrence of a TE in locations where the conditions are conducive enough to NPF. It is, however, important to mention that it is possible to have a regional NPF episode taking place simultaneously with a transported one, and when the latter is transported it gets mixed with the regional NPF so that this situation will be classified as a RE. Finally, since ion bursts are attempts of an event but do not grow, an interrupted clear sky could explain this phenomenon: for instance a sudden appearance of a cloud would result in the interruption of NPF (Baranizadeh et al., 2014), which then remains as an ion burst only. Finally, the RH, which in general correlates with cloudiness, showed a nice pattern between the event classes: RH was the lowest for RE and the highest for NE, and it fairly reflects cloudiness.

### *3.4 Start times, peak time and end time of RE*

Our method makes it possible to detect the start, peak and end times of every regional event classified during our study period. Although several previous studies state that ~~that~~ the occurrence of NPF starts with sunrise and peaks around midday, very few investigations have considered occurrence times accurately. We derived the start, peak and end times from 2 - 4 nm ions automatically, as mentioned in sections 2.4 and 2.5. During spring, when most of the NPF events occur, our results (Figure 6) show that indeed RE occur after sunrise and prior to noon, with the maximum number of days occurring between the sunrise and 5 hours past sunrise. The peak times of the events had the most frequent occurrence at 5 to 6 hours after the sunrise, which is between 10:30 and 11:30 local time, complementing our previous assumption that NPF peaks before noon. Finally, the ending times of the events had the most frequent occurrence at 9 to 11 hours after sunrise. During summer the events tend to start, peak and end later than in spring, and they show lower variability in comparison to spring. This observation could be attributed to longer daylight hours and less clouds. Whereas in autumn, the events, start, peak and end earlier than in spring. Exceptionally, during winter, ion concentrations might be affected by the accumulation of snow on or around the inlets. Overall, the variability of the event start, peak and end times can be affected by the solar cycle, degree of cloudiness and seasonality. The importance of the identification of the exact start and end times of the process helps to increase our understanding on the processes governing the NPF phenomenon. More specifically, they allow forming a time series where NPF is separated from non-event times, making it possible to compare the parameters responsible for the NPF process within appropriate time frames.

### *3.5 Comparison to previous classification*

In order to estimate the goodness of our automatic method, it is crucial to compare our results with the previous classifications (Dal Maso et al., 2005;Kulmala et al., 2012). Although such a comparison is not straightforward, we show one version of such a comparison in Figure 7. On the x-axis, the original classified days are shown, and the refined classes are shown on the y-axis as a fraction of each original class. For example, 65% of the originally-classified event days (event days make 25% of the total days in Hyytiälä according to the original classification) were found to be RE, 10% were TE and 14% were IB. The remaining 11% were considered as misclassified or bad data (by manual classification) and were excluded from the plot. In total, our automatic method was able to classify 89% of the original NPF events into some of the new event classes (RE, TE or IB). The original non-events (which made 40% of the total days) were split between the TE (20%), IB (19%) and NE (53%). The remaining 8% were bad data according to the manual classification and were excluded from the plot.

Finally, undefined days, which according to the traditional classification were 35% of the total days, were split between all the classes. Our results show that 17% of those were RE, 21% were TE, 19% were IB and 42% were non-events. Those days were usually excluded from further analysis because they did not belong to a defined class according to the original classification method. Previous extensive studies of undefined days in Hyytiälä by Buenrostro Mazon et al. (2009) showed that a fraction of undefined days resembles interrupted events which, in our case, were 83% of the days (TE, IB or NE), and which all in all were related to unfavorable conditions for regional NPF. The interruption mechanisms may include appearance of clouds (Baranizadeh et al., 2014; Dada et al., 2017), resulting in decreased radiation essential for particle formation and growth (Jokinen et al., 2017) ), or a change in the origin of arriving air masses from a clean to a rather polluted sector. ~~. Others include a change of arriving air masses from a clean to a rather polluted sector~~(Sogacheva et al., 2005). ~~Also, the growth can be interrupted by a sudden appearance of a cloud (Baranizadeh et al., 2014;Dada et al., 2017).~~ Our automated method fails sometimes as the result of the simultaneous appearance of an ion burst and a pollution plume. While the misjudgment of these days as regional events is largely minimized by correcting for the background concentrations of 7–25 nm particles, erroneous classification is still possible in some cases.

## 4 Conclusions

Using 10 years of measurement using the NAIS at SMEAR II station, we were able to create an automated method to classify days into 4 classes based on their ion (2 – 4 nm) and particle (7 - 25 nm) number concentrations, including regional events, transported events, ion bursts and non-events. Our method minimizes the efforts used in manual day-by-day classification as well as the errors due to human bias. In addition, our method allows for the complete classification (sub-3 nm) of all days, i.e. reduces the number of previously known 'undefined days', which have always been excluded from previous analyses.

Our results show that on ~ 40% of the days during spring in Hyytiälä, a regional NPF event occurs and is characterized by a set of favorable conditions, such as a clear sky, low condensation sink, medium temperature and low relative humidity. On the contrary, NE were ~25 % of the days and were characterized by a complete cloud cover, high RH and high CS. Interestingly, TE and IB fall in the category between RE and NE in this respect. While IB are interrupted growth of initially started RE due to a probable change to polluted air mass or an appearance of a cloud, TE occurred on days when there was little chance for the cluster to form within our measurement location but still they had a chance to grow if reaching our site. Both IB and TE were characterized by intermediate values of CS, RH and *P* compared with RE and NE. Moreover, using the new method we are able to identify the start time, peak time and end time of events occurring in Hyytiälä. Our results show that most RE started within 5 hours from the sunrise, peaked before noon, and ended 10 hours after sunrise. Finally, with small changes the classification method can be applied to other places around the globe where NPF takes place providing deeper understanding yet less effort for atmospheric scientists.

*Acknowledgements:* Lubna Dada acknowledges the doctoral programme in Atmospheric Sciences (ATM-DP, University of Helsinki) for financial support. This project has received funding from the European Union's Horizon 2020 research (ERA-PLANET (689443)) and innovation programme under grant agreement No 654109. This work was supported by the European Commission via projects ACTRIS2, European Research Council via ATM-GTP (742206), and Academy of Finland Centre of Excellence in Atmospheric Sciences (grant number: 272041).

*Data availability:* Data measured at the SMEAR II station are available on the webpage: http://avaa.tdata.fi/web/smart/. The classification, start times, peak times and end times are available from Lubna Dada (lubna.dada@helsinki.fi) upon request.

**References**

Aalto, P., Hämeri, K., Becker, E., Weber, R., Salm, J., Mäkelä, J. M., Hoell, C., O'dowd, C. D., 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 10.1034/j.1600-0889.2001.530403.x, 2001.

Artaxo, P., Rizzo, L. V., Brito, J. F., Barbosa, H. M., Arana, A., Sena, E. T., Cirino, G. G., Bastos, W., Martin, S. T., and Andreae, M. O.: Atmospheric aerosols in Amazonia and land use change: from natural biogenic to biomass burning conditions, Faraday discussions, 165, 203-235, 2013.

Asmi, A., Wiedensohler, A., Laj, P., Fjaeraa, A.-M., Sellegri, K., Birmili, W., Weingartner, E., Baltensperger, U., Zdimal, V., and Zikova, N.: Number size distributions and seasonality of submicron particles in Europe 2008–2009, Atmos. Chem. Phys., 11, 5505-5538, 10.5194/acp-11-5505-2011, 2011.

Asmi, E., Kondratyev, V., Brus, D., Laurila, T., Lihavainen, H., Backman, J., Vakkari, V., Aurela, M., Hatakka, J., Viisanen, Y., Uttal, T., Ivakhov, V., and Makshtas, A.: Aerosol size distribution seasonal characteristics measured in Tiksi, Russian Arctic, Atmos. Chem. Phys., 16, 1271-1287, 10.5194/acp-16-1271-2016, 2016.

Baranizadeh, E., Arola, A., Hamed, A., Nieminen, T., Mikkonen, S., Virtanen, A., Kulmala, M., Lehtinen, K., and Laaksonen, A.: The effect of cloudiness on new-particle formation: investigation of radiation levels, Boreal Env. Res., 19, 343-354, 2014.

Bianchi, F., Tröstl, J., Junninen, H., Frege, C., Henne, S., Hoyle, C., Molteni, U., Herrmann, E., Adamov, A., Bukowiecki, N., Chen, X., Duplissy, J., Gysel, M., Hutterli, M., Kangasluoma, J., Kontkanen, J., Kürten, A., Manninen, H. E., Münch, S., Peräkylä, O., Petäjä, T., Rondo, L., Williamson, C., Weingartner, E., Curtius, J., Worsnop, D. R., Kulmala, M., Dommen, J., and Baltensperger, U.: New particle formation in the free troposphere: A question of chemistry and timing, Science, 352, 1109-1112, 10.1126/science.aad5456, 2016.

Buenrostro Mazon, S., Riipinen, I., Schultz, D., Valtanen, M., Maso, M. D., Sogacheva, L., Junninen, H., Nieminen, T., Kerminen, V.-M., and Kulmala, M.: Classifying previously undefined days from eleven years of aerosol-particle-size distribution data from the SMEAR II station, Hyytiälä, Finland, Atmos. Chem. Phys., 9, 667-676, 10.5194/acp-9-667-2009, 2009.

Buenrostro Mazon, S., Kontkanen, J., Manninen, H. E., Nieminen, T., Kerminen, V.-M., and Kulmala, M.: A long-term comparison of nighttime cluster events and daytime ion formation in a boreal forest, Boreal Env. Res., 21, 242-261, 2016.

Dada, L., Paasonen, P., Nieminen, T., Buenrostro Mazon, S., Kontkanen, J., Peräkylä, O., Lehtipalo, K., Hussein, T., Petäjä, T., Kerminen, V. M., Bäck, J., and Kulmala, M.: Long-term analysis of clear-sky new particle formation events and nonevents in Hyytiälä, Atmos. Chem. Phys., 17, 6227-6241, 10.5194/acp-17-6227-2017, 2017.

Dal Maso, M., Kulmala, M., Riipinen, I., Wagner, R., Hussein, T., Aalto, P. P., and Lehtinen, K. E.: Formation and growth of fresh atmospheric aerosols: eight years of aerosol size distribution data from SMEAR II, Hyytiala, Finland, Boreal Env. Res., 10, 323, 2005.

Hamed, A., Korhonen, H., Sihto, S. L., Joutsensaari, J., Järvinen, H., Petäjä, T., Arnold, F., Nieminen, T., Kulmala, M., and Smith, J. N.: The role of relative humidity in continental new particle formation, Journal of Geophysical Research: Atmospheres, 116, 2011.

Hari, P., and Kulmala, M.: Station for measuring ecosystem-atmosphere relations, Boreal Env. Res., 10, 315-322, 2005.

Hyvönen, S., Junninen, H., Laakso, L., Maso, M. D., 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, 10.5194/acp-5-3345-2005, 2005.

IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp., 2013.

Jokinen, T., Kontkanen, J., Lehtipalo, K., Manninen, H. E., Aalto, J., Porcar-Castell, A., Garmash, O., Nieminen, T., Ehn, M., and Kangasluoma, J.: Solar eclipse demonstrating the importance of photochemistry in new particle formation, Scientific Reports, 7, 45707, 2017.

Kerminen, V.-M., Paramonov, M., Anttila, T., Riipinen, I., Fountoukis, C., Korhonen, H., Asmi, E., Laakso, L., Lihavainen, H., Swietlicki, E., Svenningsson, B., Asmi, A., Pandis, S. N., Kulmala, M., and Petäjä, T.: Cloud condensation nuclei production associated with atmospheric nucleation: a synthesis based on existing literature and new results, Atmos. Chem. Phys., 12, 12037-12059, 10.5194/acp-12-12037-2012, 2012.

Kerminen, V.-M., Chen, X., Vakkari, V., Petäjä, T., Kulmala, M., and Bianchi, F. J. E. R. L.: Atmospheric new particle formation and growth: review of field observations, 13, 103003, 2018.

Kulmala, M., Alekseychik, P., Paramonov, M., Laurila, T., Asmi, E., Arneth, A., Zilitinkevich, S., and Kerminen, V.-M.: On measurements of aerosol particles and greenhouse gases in Siberia and future research needs, Boreal Environment Research, 16, 2011.

Kulmala, M., Petäjä, T., Nieminen, T., Sipilä, M., Manninen, H. E., Lehtipalo, K., Dal Maso, M., Aalto, P. P., Junninen, H., and Paasonen, P.: Measurement of the nucleation of atmospheric aerosol particles, Nature protocols, 7, 1651-1667, 10.1038/nprot.2012.091, 2012.

Kulmala, M., Kontkanen, J., Junninen, H., Lehtipalo, K., Manninen, H. E., Nieminen, T., Petäjä, T., Sipilä, M., Schobesberger, S., Rantala, P., Franchin, A., Jokinen, T., Järvinen, E., Äijälä, M., Kangasluoma, J., Hakala, J., Aalto, P., Paasonen, P., Mikkilä, J., Vanhanen, J., Aalto, J., Hakola, H., Makkonen, U., Ruuskanen, T., Mauldin, R. r., Duplissy, J., Vehkamäki, H., Bäck, J., Kortelainen, A., Riipinen, I., Kurtén, T., Johnston, M., Smith, J., Ehn, M., Mentel, T., Lehtinen, K., Laaksonen, A., Kerminen, V., and Worsnop, D.: Direct observations of atmospheric aerosol nucleation, Science, 339, 943-946, 10.1126/science.1227385, 2013.

Kulmala, M., Petäjä, T., Kerminen, V.-M., Kujansuu, J., Ruuskanen, T., Ding, A., Nie, W., Hu, M., Wang, Z., Wu, Z., Wang, L., and Worsnop, D. R.: On secondary new particle formation in China, Frontiers of Environmental Science & Engineering, 10, 8, 10.1007/s11783-016-0850-1, 2016.

Kulmala, M., Kerminen, V.-M., Petäjä, T., Ding, A., and Wang, L.: Atmospheric gas-to-particle conversion: why NPF events are observed in megacities?, Faraday discussions, 200, 271-288, 2017.

Kulmala, M.: Build a global Earth observatory, Nature, 553, 21-23, 2018.

Leino, K., Nieminen, T., Manninen, H. E., Petäjä, T., Kerminen, V.-M., and Kulmala, M.: Intermediate ions as a strong indicator of new particle formation bursts in a boreal forest, 2016.

Leino, K., Lampilahti, J., Poutanen, P., Väänänen, R., Manninen, A., Mazon, S. B., Dada, L., Nikandrova, A., Wimmer, D., Aalto, P. P., Ahonen, L. R., Enroth, J., Kangasluoma, J., Keronen, P., Korhonen, F., Laakso, H., Matilainen, T., Siivola, E., Manninen, H. E., Lehtipalo, K., Kerminen, V.-M., Petäjä, T., and Kulmala, M.: Vertical profiles of sub-3 nm particles over the boreal forest Atmos. Chem. Phys. Discuss. (Submitted), 2018.

Manninen, H. E., Petäjä, T., Asmi, E., Riipinen, I., Nieminen, T., Mikkilä, J., Hõrrak, U., Mirme, A., Mirme, S., and Laakso, L.: Long-term field measurements of charged and neutral clusters using Neutral cluster and Air Ion Spectrometer (NAIS), Boreal Environ. Res, 14, 591-605, 2009.

Manninen, H. E., Nieminen, T., Asmi, E., Gagné, S., Häkkinen, S., Lehtipalo, K., Aalto, P., Vana, M., Mirme, A., Mirme, S., Hõrrak, U., Plass-Dülmer, C., Stange, G., Kiss, G., Hoffer, A., Törő, N., Moerman, M., Henzing, B., de Leeuw, G., Brinkenberg, M., Kouvarakis, G. N., Bougiatioti, A., Mihalopoulos, N., O'Dowd, C., Ceburnis, D., Arneth, A., Svenningsson, B., Swietlicki, E., Tarozzi, L., Decesari, S., Facchini, M. C., 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., Petäjä, T., Kerminen, V. M., and Kulmala, M.: EUCAARI ion spectrometer measurements at 12 European sites – analysis of new particle formation events, Atmos. Chem. Phys., 10, 7907-7927, 10.5194/acp-10-7907-2010, 2010.

Manninen, H. E., Mirme, S., Mirme, A., Petäjä, T., and Kulmala, M.: How to reliably detect molecular clusters and nucleation mode particles with Neutral cluster and Air Ion Spectrometer (NAIS), Atmospheric Measurement Techniques, 9, 3577-3605, 10.5194/amt-9-3577-2016, 2016.

Merikanto, J., Spracklen, D., Mann, G., Pickering, S., and Carslaw, K.: Impact of nucleation on global CCN, Atmos. Chem. Phys., 9, 8601-8616, 10.5194/acp-9-8601-2009, 2009.

Mirme, S., and Mirme, A.: The mathematical principles and design of the NAIS–a spectrometer for the measurement of cluster ion and nanometer aerosol size distributions, Atmospheric Measurement Techniques, 6, 1061-1071, 2013.

Nieminen, T., Asmi, A., Dal Maso, M., Aalto, P. P., Keronen, P., Petäjä, T., Kulmala, M., and Kerminen, V.-M.: Trends in atmospheric new-particle formation: 16 years of observations in a boreal-forest environment, Boreal Env. Res., 19, 2014.

Nieminen, T., Yli-Juuti, T., Manninen, H., Petäjä, T., Kerminen, V.-M., and Kulmala, M.: Technical note: New particle formation event forecasts during PEGASOS–Zeppelin Northern mission 2013 in Hyytiälä, Finland, Atmos. Chem. Phys., 15, 12385-12396, 10.5194/acp-15-12385-2015, 2015.

Orfanidis, S. J.: Introduction to signal processing, Prentice-Hall, Inc., 1995.

Paasonen, P., Peltola, M., Kontkanen, J., Junninen, H., Kerminen, V.-M., and Kulmala, M.: Comprehensive analysis of particle growth rates from nucleation mode to cloud condensation nuclei in Boreal forest, Atmos. Chem. Phys. Discuss. (Submitted), 2018.

Rose, C., Zha, Q., Dada, L., Yan, C., Lehtipalo, K., Junninen, H., Mazon, S. B., Jokinen, T., Sarnela, N., Sipilä, M., Petäjä, T., Kerminen, V.-M., Bianchi, F., and Kulmala, M.: Observations of biogenic ion-induced cluster formation in the atmosphere, Science Advances, 4, 10.1126/sciadv.aar5218, 2018.

Salma, I., Németh, Z., Kerminen, V.-M., Aalto, P., Nieminen, T., Weidinger, T., Molnár, Á., Imre, K., and Kulmala, M.: Regional effect on urban atmospheric nucleation, Atmos. Chem. Phys., 16, 8715-8728, 10.5194/acp-16-8715-2016, 2016.

Sogacheva, L., Dal Maso, M., Kerminen, V.-M., and Kulmala, M.: Probability of nucleation events and aerosol particle concentration in different air mass types arriving at Hyytiälä, southern Finland, based on back trajectories analysis, Boreal Env. Res., 10, 2005.

Spracklen, D. V., Carslaw, K. S., Merikanto, J., Mann, G. W., Reddington, C. L., Pickering, S., Ogren, J. A., Andrews, E., Baltensperger, U., Weingartner, E., Boy, M., Kulmala, M., Laakso, L., Lihavainen, H., Kivekäs, N., Komppula, M., Mihalopoulos, N., Kouvarakis, G., Jennings, S. G., O'Dowd, C., Birmili, W., Wiedensohler, A., Weller, R., Gras, J., Laj, P., Sellegri, K., Bonn, B., Krejci, R., Laaksonen, A., Hamed, A., Minikin, A., Harrison, R. M., Talbot, R., and Sun, J.: Explaining global surface aerosol number concentrations in terms of primary emissions and particle formation, Atmos. Chem. Phys., 10, 4775-4793, 10.5194/acp-10-4775-2010, 2010.

Wang, Z., Wu, Z., Yue, D., Shang, D., Guo, S., Sun, J., Ding, A., Wang, L., Jiang, J., and Guo, H.: New particle formation in China: Current knowledge and further directions, Science of the Total Environment, 577, 258-266, 2017.

Wimmer, D., Buenrostro Mazon, S., Manninen, H. E., Kangasluoma, J., Franchin, A., Nieminen, T., Backman, J., Wang, J., Kuang, C., Krejci, R., Brito, J., Goncalves Morais, F., Martin, S. T., Artaxo, P., Kulmala, M., Kerminen, V. M., and Petäjä, T.: Ground-based observation of clusters and nucleation-mode particles in the Amazon, Atmos. Chem. Phys., 18, 13245-13264, 10.5194/acp-18-13245-2018, 2018.

Yli-Juuti, T., Nieminen, T., Hirsikko, A., Aalto, P., Asmi, E., Hõrrak, U., Manninen, H., Patokoski, J., Maso, M. D., and Petäjä, T.: Growth rates of nucleation mode particles in Hyytiälä during 2003− 2009: variation with particle size, season, data analysis method and ambient conditions, Atmospheric Chemistry and Physics, 11, 12865-12886, 2011.

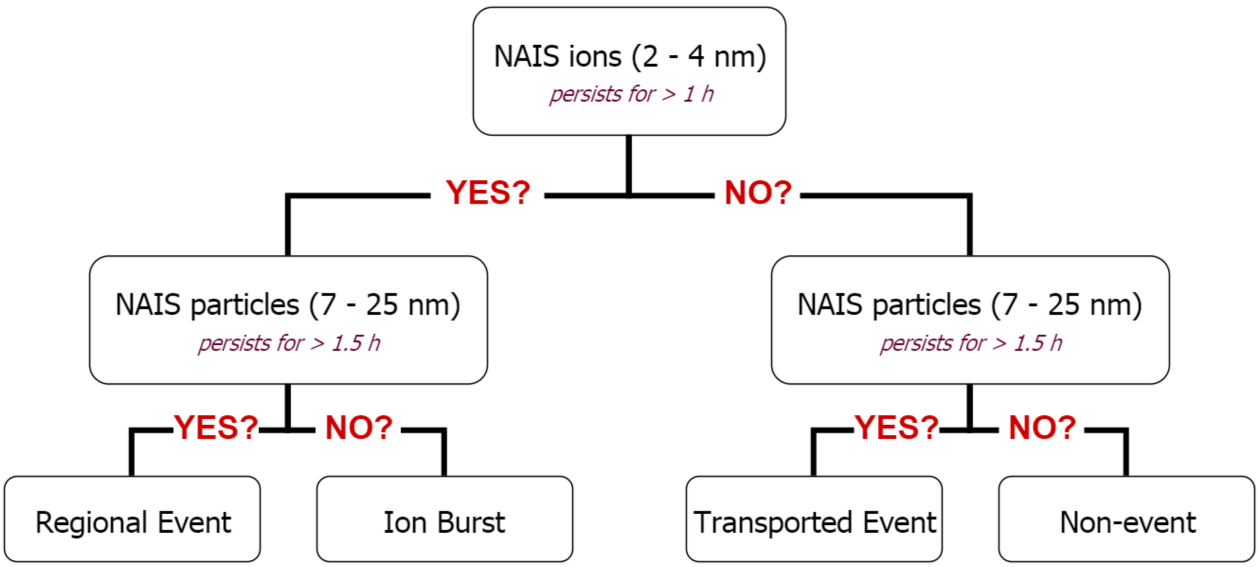



***Figure 1 A flow chart for the decision path during event classification in Hyytiälä using new classification method.***

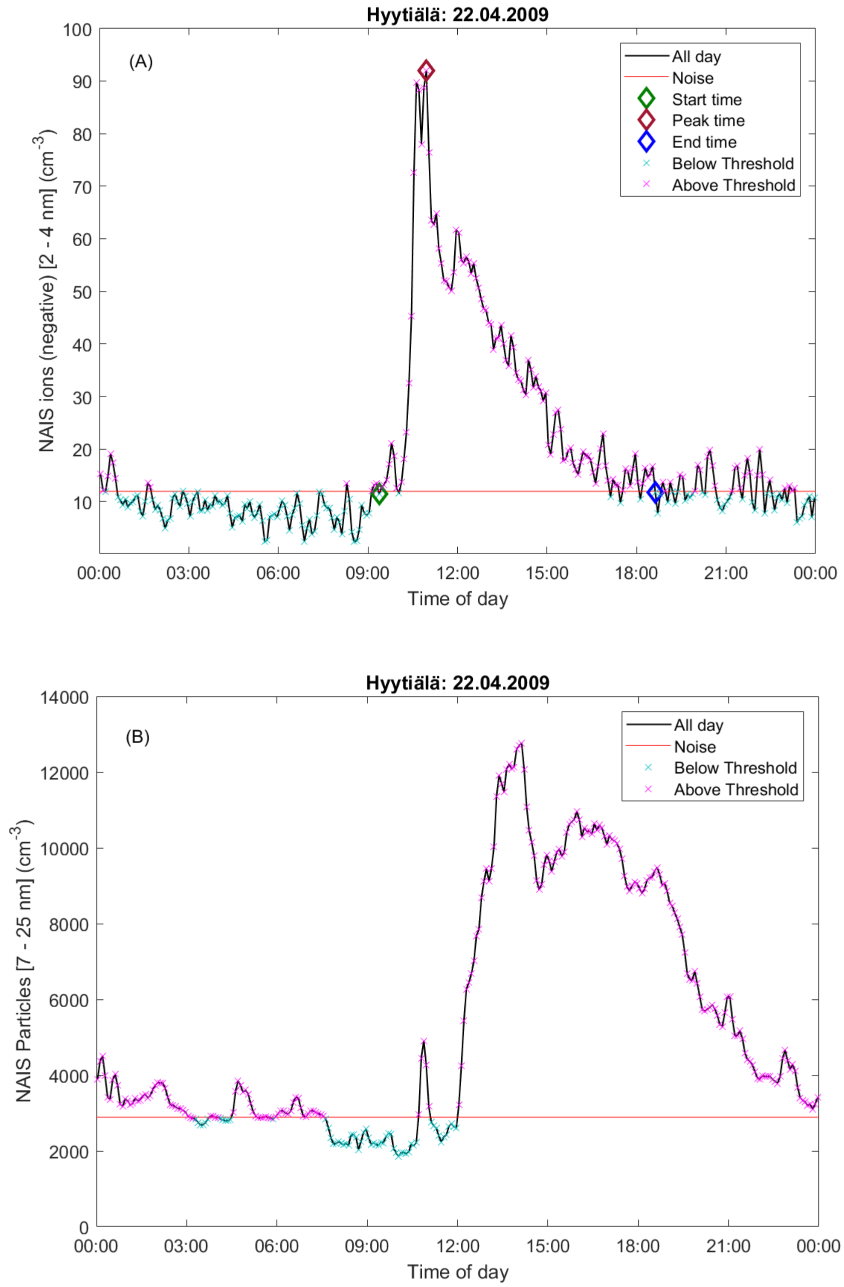



***Figure 2 Automatic method applied to (A) 2 – 4 nm ions (negative) example, ion concentration passed threshold and persisted > 1 hour and (B) 7 – 25 nm particles example, particle concentration passed threshold and persisted for > 1.5 hours.***

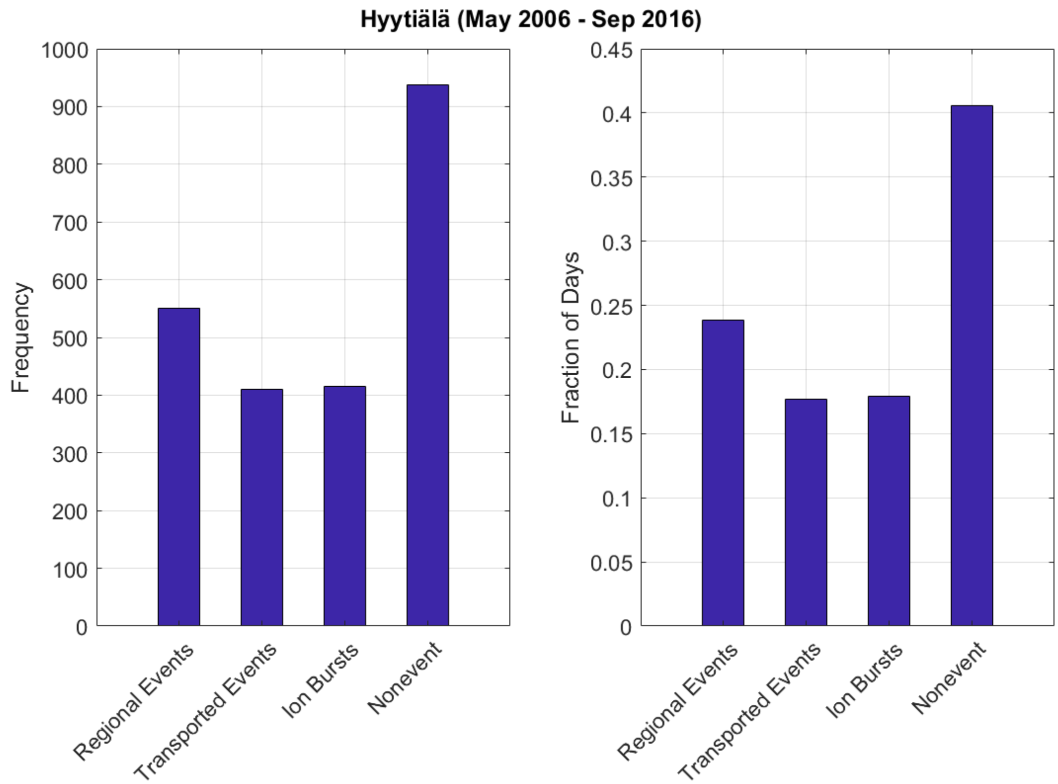



***Figure 3 Frequency and fraction of events, ions burst and non-events in Hyytiälä using the new classification method.***

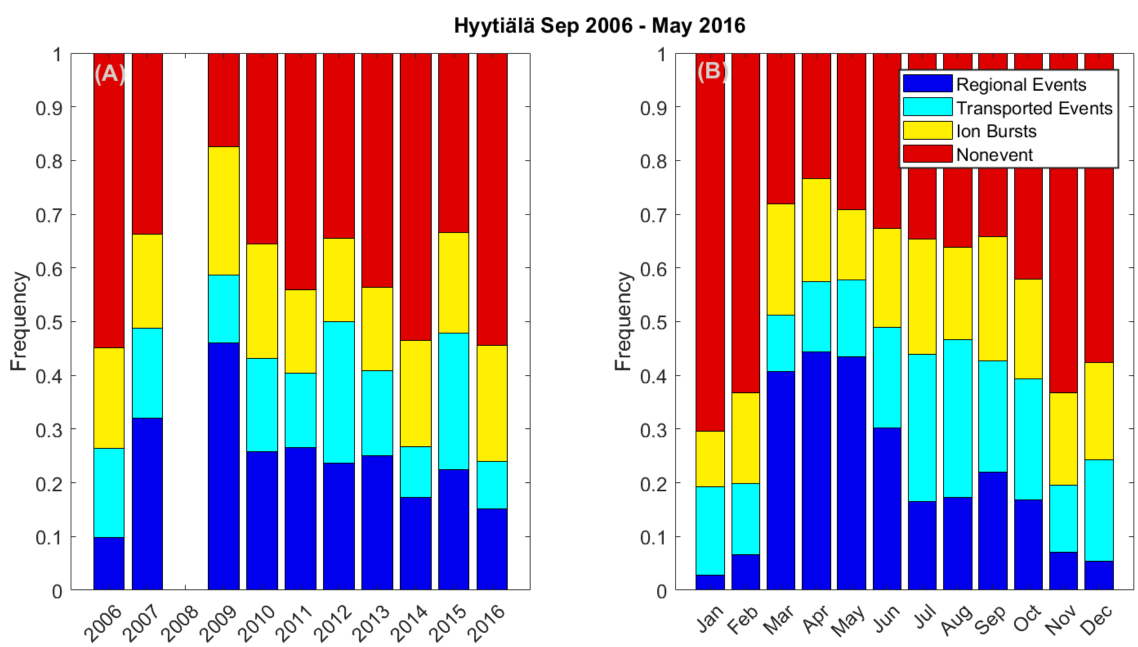



***Figure 4 (A) Yearly and (B) monthly fraction of days classified as Regional events (RE), Transported events (TE), Ion bursts (IB), and non-events(NE) using the new classification method. The data of year 2009 is bias to spring months, which could explain the much higher number of events. No data was available during 2008.***

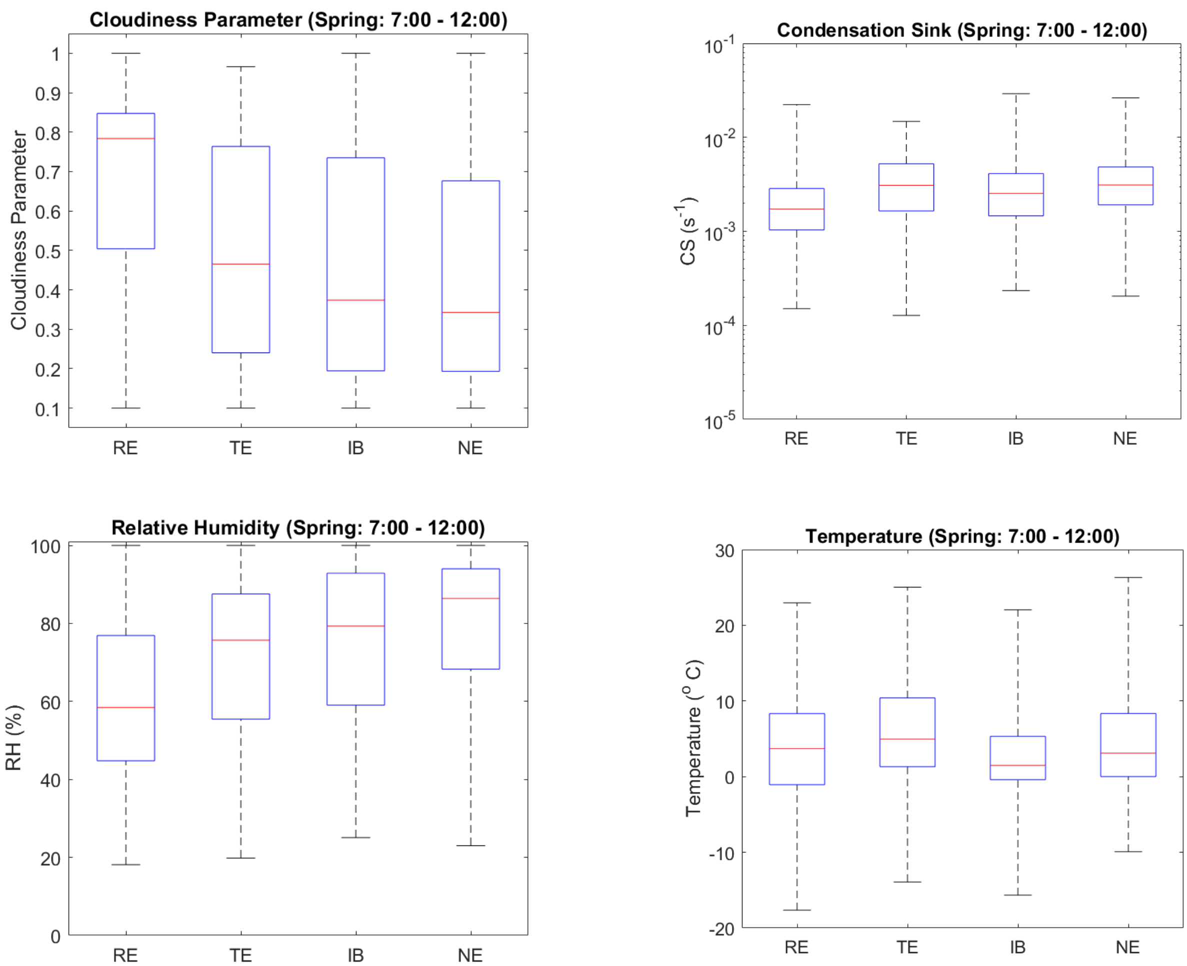



***Figure 5 (A) Cloudiness parameter, (B) condensation sink, (C) Relative humidity and (D) Temperature during different days classified with the new classification method for Spring (Mar-May) of 2006-2016 during maximum NPF window (7:00 – 12:00). The acronyms RE, TE, IB and NE stand for regional events, transported events, ions bursts and non-events, respectively. The red line represents the median of the data and the lower and upper edges of the box represent 25th and 75th percentiles of the data respectively. The lines extending from the central box represent the minimum and the maximum of the data inclusive.***

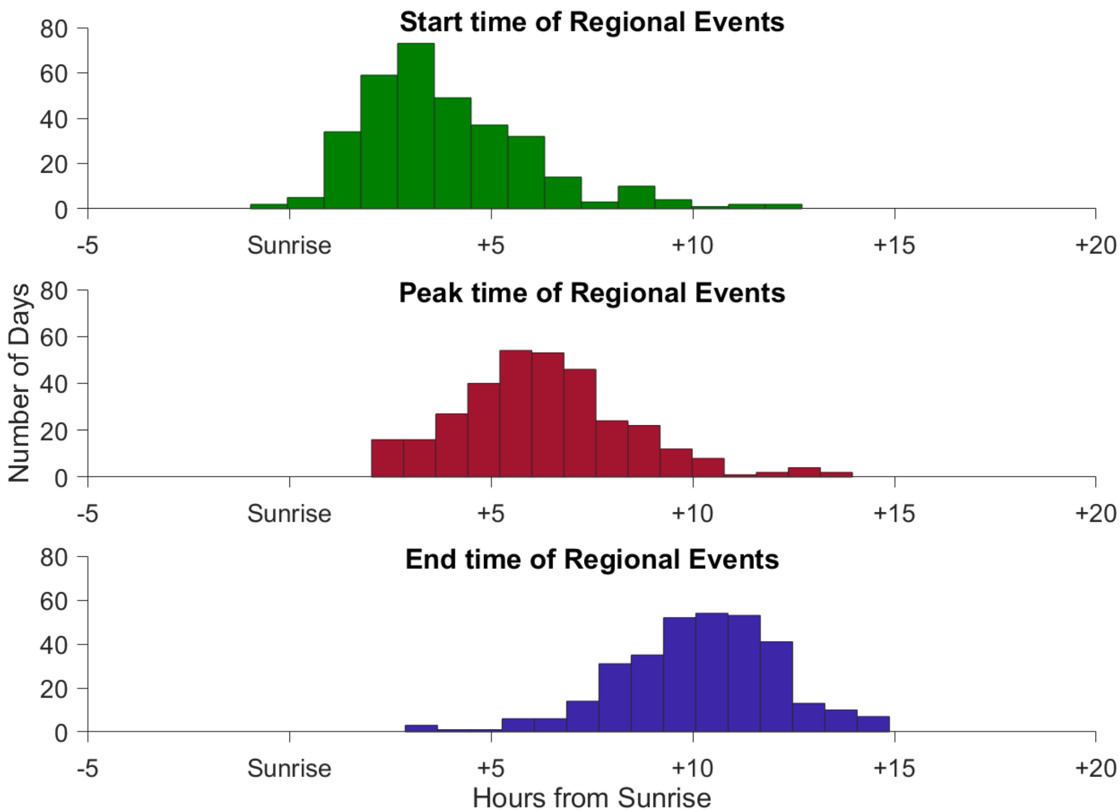



***Figure 6 Frequency of days during spring at which regional events start, peak and end past sunrise. For example, most events start within 3~~5~~ hours from sunrise.***

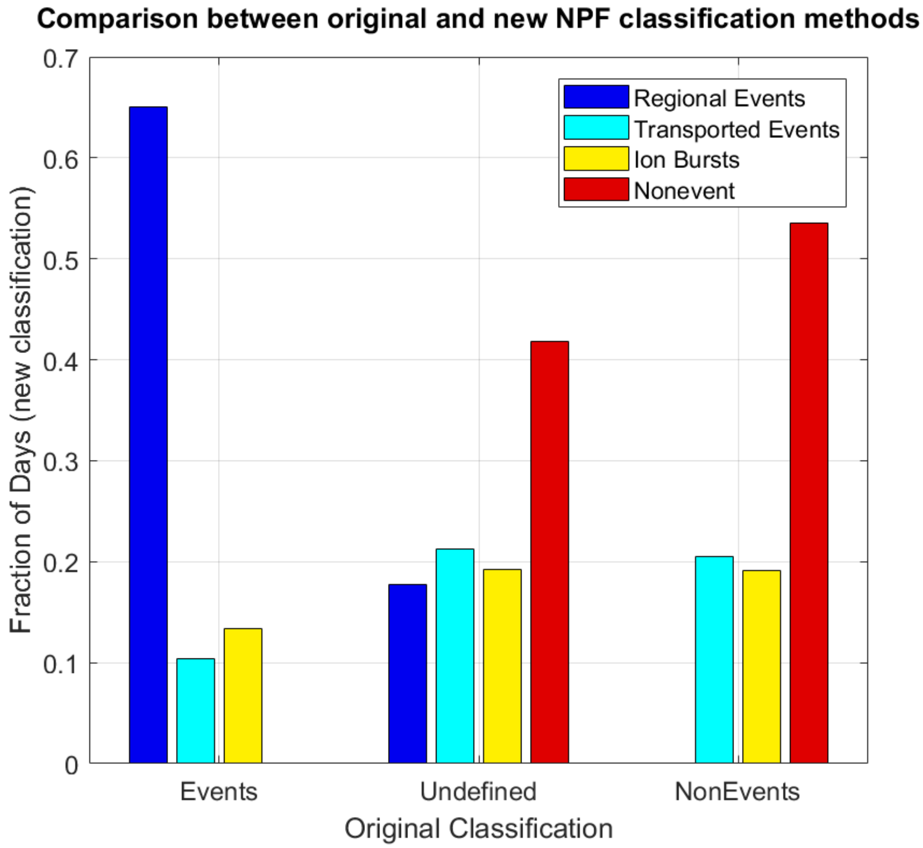



***Figure 7 Comparison between original and new NPF classification methods. The refined classification matches 94% with original event and non-event classification.***