MS No.: acp-2020-555
Title: The Effect of Meteorological Conditions and Atmospheric Composition in the Occurrence and Development of New Particle Formation (NPF) Events in Europe
Author(s): Dimitrios Bousiotis et al.

RESPONSE TO REVIEWERS

Referee #1

General comments
The current version of the manuscript has improved. Readability is better. The authors followed most of the suggestions. However, there are still few minor aspects I encourage to improve.

The readability of the abstract and introduction could still be polished.
RESPONSE: We have done our best to improve these sections.

I am not convinced by the use of term "NPF probability". You are probably aware that some readers might only (or first) check the abstract and conclusion. Using "NPF probability" term may be confusing. If you still think this is something you would like to keep, I suggest be more specific what it actually represents when discussing it in conclusion.
RESPONSE: As the term "NPF probability" seems to cause a lot of misunderstandings it was replaced by the more conventional "NPF frequency" throughout the manuscript. Small changes were also made in the Methods section to accommodate the change (line 238).

Please specify how the NPF frequency was calculated, and check values presented in Table 2. Which data is considered for this calculation? I think the answer was given in author's response (AR) file but I missed mentioning of that in the main manuscript. Further, some comments are addressed in AR but these clarifications do not always appear in text. Please double check.
RESPONSE: Due to the aforementioned change (removal of the "NPF probability" term), the term "NPF frequency" is now explained in the Methods section with the inclusion of the "full dataset" as an option of the possible groups for which it is calculated (as the term "NPF frequency" was previously used when full datasets were considered). Additionally, some information about the trends and results included in the AR (pointing to previous work) were also added in the manuscript as they can help in explaining some trends (please see the response on a later comment as well).

How sure are you that NPF occurred at UK sites if you only have SMPS data available >16 nm for these locations? Please comment in the manuscript. It is also one of the limitations of the study that you could mention. What was your approach for analysis of these sites?
RESPONSE: For the extraction of the NPF events in the UK sites additional criteria were set (including the variation of particle number concentration data from 7 nm, the variations of pollutant concentrations, the condensation sink, the effect of the nearby Heathrow Airport, etc.). The text from the Methods section in the paper in which these results were first presented reads:
"At this point it should be mentioned that due to the particle size range available, NPF events in which new formed particles failed to grow beyond 16.6 nm (if any) could not be identified. Bursts of new particles in the size range < 16.6 nm that were identified using the CPC data but did not appear in the SMPS dataset were ignored as their development was unknown. This type of development was rare and mainly found at the rural background site, occurring on a few days per year mainly in summer. Its main feature was the short duration of the bursts compared to event days. In the urban sites, this type of development was almost non-existent. High time resolution data for gaseous pollutants and aerosol constituents was used to identify pollution events affecting particle concentrations and these were removed from the data analysis. This analysis took account of the fact that nanoparticle emissions from Heathrow Airport affect size distributions at London sites (Harrison et al., 2018), and such primary emission influences were not included as NPF events."

As this is a lengthy clarification, a note was added in the Methods section in the present study, mentioning the limitation and referencing the work where this is further explained:
"As the available SMPS datasets for the sites in the U.K. are for particles of diameter greater than 16 nm, additional criteria were set to ensure the correct extraction of NPF events including the variations of the particle number concentrations from a Condensation Particle Counter (CPC – measuring from 7nm), as well as of the concentrations of gaseous pollutants and aerosol constituents (please refer to the Methods section in Bousiotis et al., 2019)." (line 186)

I have noticed that frequently two other papers published/submitted by Bousiotis et al (2019, 2020) are mentioned in AR. It would be good if the authors make sure that these are also mentioned in relevant places throughout the manuscript (mentioning the issue raised by reviewers e.g.: "xx was explored in Bousiotis et al. xxxx and is not the focus in this paper" or ""xx can be found in…").
Also make sure that crucial information on the study is provided in the current manuscript without the need to often look into two other papers to get a complete picture.
RESPONSE: We thank the reviewer for this suggestion. Some references were added in the manuscript that point to results from the previous studies. Also, some crucial information found in these works were also added in several points in the manuscript, to clarify the points made (results about the seasonality of the GR and J, the variation of the temperature, CS etc.).

Table 3: please indicate in the caption what values "in bold" indicate
RESPONSE: A clarification for what the values in bold indicate was added in the captions of Tables 3, 4 and S1.

Figure 7a, 7b: make one *caption* for figure 7(a and b) and only indicate on corresponding plots "a" and "b".
RESPONSE: The figure was updated with a single caption. Also, the plot 4b was removed from the Figure Legends table as it was moved to the SI

Referee #2

This work compiles already published results from 16 sites located in six European cities. Within this huge task, the effort is to identify relationships between meteo variables, gas phase composition and aerosol organic content with key NPF variables such as NPF frequency (the authors name it NPF probability), growth rate and formation rate. Several findings within this work justify publication; two most striking is the nonlinear relation of temperature with NPF probability and the fact that increased solar irradiance reduces the probability for NPF.

Some issues require attention though.

Starting from most important to least severe

The authors have chosen to use only Ia events and as a result the probability shown in Table 2 is several factors smaller than those reported in literature for the same sites. However, from a brief search, the difference can be up to a factor of six. It is well understood why the authors made such a choice as formation and growth rates can be calculated reliably from these type of events only. How does this choice reflect to the results shown? I strongly support that for one site (e.g. Hyytiälä) an intercomparison is carried out to indicate to the reader the tentative differences. My recommendation is to do so only for NPF probability. This is critical as most studies in the end

classify NPF as events, undefined and non-events lumping Ia, Ib and II classes into one. The authors should add to the caption of Table 1 the fact that only Ia events are considered.
RESPONSE: According to the results from the analysis of NPF events at the sites of the study it was found that the NPF events that did not meet the criteria for class Ia were up to double the number of those that are characterised as class Ia. Thus, in the methods section the following text was added:

"As only class Ia events were only considered, it is expected that the frequency of the events calculated should be lower than that expected if all types of events were included. This could result in values up to one third of those anticipated if all classes of events were considered. For the extent of this variation please refer to Bousiotis et al., (2019; 2020) in which there is an extended analysis of the NPF events for each site, including the special cases of NPF events that do not comply with the criteria set for class Ia." (line 316).

Additionally, the text was updated for the Table mentioning that the statistics refer to class Ia NPF events. (We assume that the caption of Table 2 is the one that needs updating).

How coarse is the time resolution of OC and sulfate measurements examined in this study? The first impression for the former is that they are derived from the thermal optical method. For the latter AMS is mentioned. However, AMS is typically used for short-term campaigns and this is a multi-year study. Do the authors mean ACSM? If the resolution of these measurements is coarser than 1 h, are they reliable to be used in NPF studies? The authors should clarify the time resolution of these measurements, both OC and SO4, in the manuscript and discuss any complications. If the authors indeed use AMS measurements what fraction of the time period discussed do they cover? Also, it would be worthwhile to mention the publications that refer to these measurements.
RESPONSE: ACSM data was used (updated in the text). The measurement resolution for all the sites for which such data was used is 1 hour. Data with 3-hour resolution or more was available but was not used as it would bias the results. The note "For all the sites, the data used in the present study are of either 1-hour resolution or less. Data with coarser resolution were omitted for reliability." was added in the Methods section for clarification (line 160). References for publications that reported the measurements of this study were added in the Site Description section.

In each pair of variables (e.g. NPF probability and RH; growth rate and temperature) presented in this work, there seems to be a norm and one (or more) site that is an exception. Since the authors cannot fully explain why (and this is perfectly understood), it is worthwhile to mention that in the abstract or the conclusions or in both.
RESPONSE: The note "though exceptions were found among the sites for all the variables studied" was added in the abstract (line 57). The note "in the majority of the sites (though exceptions were found as well, mostly in the southern sites)," was added in the Conclusions section to point out the exceptions found (line 863).

There is little relationship between RH and CS at most sites. Is this because CS was based on dried measurements and was not corrected for hygroscopic growth? This would be understood since chemical composition was lacking on most sites. Please discuss if CS was corrected for hygroscopic growth and how that affects the results presented.
RESPONSE: A note has been added in the methods that CS was not corrected for hygroscopic growth as well as for the effect this has on the results presented.

ANOVA is only valid for normally distributed populations. Have the authors tested for normality? The F-test is typically used. How did the authors treat skewed distributions? Please discuss.
RESPONSE: The Shapiro-Wilk test was used to assess the normality and the vast majority of the variables were found to have $p > 0.05$ and thus were considered as normal. This is probably due to

the removal of the extreme values (for the calculations, 90% of each dataset was kept removing the extremely high and/or low values and the possible outliers included in them). While this was not done to promote the normality of the populations but to reduce the bias from extreme values, it indirectly assisted in making the distributions normal. For the few remaining (e.g. the growth rates associated with SO_2 concentrations for UKRO) for which normality was not present, the square root of the values of the variable were considered to achieve normality and proceed to the ANOVA test. This clarification was added in the Methods section (line 244).

In the supplement, several relationships are clearly non-linear, such as the temperature-NPF probability for a few sites, but the authors insist to use a linear fit (probably for consistency). May I ask the authors to note on the supplemental graphs in which cases linearity is not followed. The authors are better aware of the statistical significance of the related graphs than the reader is. In the case of Denmark Rural (S2b), it is not evident whether the deviation from linearity is statistically significant or not.
RESPONSE: As it is expected that most readers will not read the SI, such deviations are discussed in the text (one example is the case of the Danish rural site mentioned which is discussed in the text). A linear relationship is not always the best to describe the relationships found, but indeed was chosen for consistency (now discussed in the text – line 282). A metric for the consistency of the linearity can be given by the R^2 and the p-values (e.g. when R^2 is low then the linearity is not consistent, at least statistically). Apart from that, it is unknown even to the authors whether a trend that starts at the extreme values of a variable (e.g. the decline found in the Danish rural site with the NPF frequency at high temperatures) would consistently continue if the temperature increases further or it is an artefact, and thus it was decided not to be further discussed in detail, apart from the mentions made in the text (as only speculation can be made).

Please define in the methods section what is weak, strong and very strong correlation in this work. It will assist the reader further.
RESPONSE: As there is no specific mention for the relationships in the Methods section, the definition of weak, strong and very strong correlations is mentioned in the first reference to the coefficient of determination (line 334).

The effect of SO2 on NPF that the authors are discussing in Section 3.2.1 has been presented before. Please check the references below. These works relate particle acidity to NPF. Experimentally it has also been verified at the site named GRERU in this work.
RESPONSE: The works suggested are mentioned and referenced in the SO_2 section (line 548).

Line 247. “The remaining data” is better use of English than the “data left”
RESPONSE: Text changed to “remaining data” (now line 278).

If the authors prefer the term NPF probability it is fine. But please use it throughout the manuscript. The caption in Table 2 is a bit confusing.
RESPONSE: The “NPF probability” was addressed in an earlier comment. The caption in Table 2 was updated.

References
Jung, J., P. J. Adams, and S. N. Pandis (2006), Simulating the size distribution and chemical composition of ultrafine particles during nucleation events, Atmos. Environ., 40, 2248–2259, doi:10.1016/j.atmosenv.2005. 09.082.

Jung, J. G., S. N. Pandis, and P. J. Adams (2008), Evaluation of nucleation theories in a sulfur-rich environment, Aerosol Sci. Technol., 42, 495–504, doi:10.1080/02786820802187085.