Answers to Anonymous Referee #2

We thank the referee for the 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.

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.