We would like to thank the referees for the effort to critically review this manuscript which has lead to its substantial improvement. We also gratefully acknowledge the editor and the Copernicus editorial team for their work.

Our point-by-point replies to the referees` comments are given in the following. For completeness, original referee comments are cited (font style: italic) and addressed in sequence (regular font style), beginning with a "→" symbol.

Anonymous Referee #1

General Comments:

I find it important to emphasize what is new/different compared to the paper by Borrmann et al., ACP 10, 5573-5592, 2010, who used the "same" data (cf. e.g., Fig. 1 in this study and Fig, 6 in Borrmann et al.) from the same campaigns and came to similar conclusions e.g., the broad particle concentration maximum in the 340 K to 390 K altitude range and its role in maintaining the stratospheric aerosol layer. There are new/different data/findings in the manuscript, but this should be emphasized.

→ While the data used in our paper indeed largely overlaps with the data presented by Borrmann et al (Atmos. Chem. Phys., 10, 1–19, 2010), the results discussed in Weigel et al. have a distinct focus on certain aspects of new particle formation as compared to the more general perspective provided by Borrmann et al.. Weigel et al. aim at a detailed discussion of NPF events including detailed comparison with modeling results. The conditions for NPF occurrence, air mass history, as well as event strength and duration are studied.

To emphasize the new/different findings compared to Borrmann et al. (2010), the following text has been added in the revised manuscript version (Section 1, last paragraph):
"With this study the location, the occurrence abundance, the strength and duration of observed NPF events are analyzed. This work refines the general findings published by Borrmann et al. (2010), which are largely based on the same data sets as the one used here."

Furthermore we added (Section 6, end of the first paragraph):
"This indicates - refining the conclusions of Borrmann et al. (2010) - that NPF is strongest above the convective outflow regions of the upper troposphere and in the lowest part of the TTL."

Generally, in some parts the paper is too lengthy. For instance the "Introduction" should be shortened (this is not a review paper on the source of stratospheric aerosol particles), rearranged, and focused on the for this study essential issues.

→ We tried to be more concise in the revised version and shortened the paper at several places, e.g. where redundant information was given. We shortened the introduction by nearly 20%. Nevertheless we wanted to provide a fairly broad introduction into the subject and to acknowledge important results from previous studies.

For instance, on page 9252 the discussion on SO2 as particle precursor gas is started but not completed and is restarted again on page 9254.

→ We confined the discussion of SO_2 in the introduction to one instance (section 1, paragraph 4)

Furthermore the purpose of the microphysical box modeling is not clear to me. When you look into the literature you will find that parametrizations of particle formation vary over a wide range in the relevant parameter space (cf. e.g. Zhang et al. JGR, 2010). Hence, as long as there is no progress on understanding the relevant mechanisms, one will always find a nucleation parametrization which reproduces the observed concentrations. But this is no poof. A similar approach as chosen by the authors here, trajectories and a microphysical box model to explain observations, was already made by others, e.g., Kanawade and Tripathi, JGR, 2006 (by the way, this reference is missing in the manuscript). So please convince me and make it clear to the reader why the MAIA simulations are helpful and what do we learn from it what we didnt already know.

- ➔ The manuscript text has been rephrased (p. 9256, sentence starting at line 18): "Studies with an aerosol nucleation model are presented for investigating the strength of NPF under various environmental conditions, and to investigate the simulations sensitivity concerning used parameters and assumptions."

  The suggested publication (Kanawade and Tripathi, JGR, 2006) is considered in the revised manuscript version.

  MAIA does not use just yet another parameterization of aerosol nucleation; instead, it describes the aerosol nucleation process based on thermochemical data for neutral and charged nucleation of sulfuric acid and water that were determined in laboratory measurements (see Kazil et al., 2010, and references therein). To our knowledge, there is no other aerosol nucleation mechanism for which laboratory thermochemical data exist and are used for atmospheric modeling. The nucleation scheme of the MAIA model is based on the PARNUC model (Kazil et al., 2007). The predictions for H2SO4/H2O ion-induced nucleation by the PARNUC model have just been confirmed in a very convincing way by the CLOUD experiment at CERN (Kirkby et al., Nature, 2011).

  We do agree that the MAIA simulations do not conclusively identify the nucleation mechanism that is responsible for the observed NPF, mainly because of the large uncertainty in the conditions at the start and along the trajectories. The main significance of the simulations is, however, the following: The FLEXPART model has determined that trajectories from regions with elevated SO_2 emissions lead to locations with observed NPF events; the MAIA simulations support a causal link between the SO_2 and the NPF events, even with substantial removal of SO_2 by deep convection before the start of the simulated trajectories. MAIA confirms the causal connection between SO_2 and the new particle formation and shows that the nucleation mode particle concentration lay roughly in the expected range.

On the other hand, the paper addresses one important question in UT/LS aerosol science: Does particle formation take place inside cirrus clouds and if yes, under which conditions? The authors should strengthen this part (maybe there are even more cases available for analysis).

- ➔ The authors agree with Referee 1 and improved the manuscript concerning this issue. In fact, a new figure (Figure 12) and the following text is included in the manuscript (p9285 before the sentence starting at line 19):

  "This upper limit for the cloud particle concentration is supported by the comparison of N_{NM} of all NPF events with $N_{>2\mu m}$ from the measurements over West Africa (SCOUT-AMMA, 2006) and Australia (SCOUT-O_3, 2005) when simultaneously cloud particles where observed (**Fehler! Verweisquelle konnte nicht gefunden werden.**). Additionally, **Fehler! Verweisquelle konnte**

nicht gefunden werden. shows that the coexistence of nucleation mode particles with cloud elements is observed frequently."

Likewise, the information that particle formation takes place mainly in the lower part of the TTL is important the question is why there and not above.

➔ The question why this NPF takes place preferably in the lowermost TTL is addressed by the suggestion that the lowermost TTL is also the level where the convective outflow mainly occurs. Thus, it seems plausible that NPF is favored to occur in the lower TTL due to large quantities of NPF precursor gases which are rapidly transported upwards by convection and are then present in the vicinity of convective outflow. Additionally, the aerosol surface area is lowest in the outflow, as most of the particles were scavenged during the uplift inside clouds. Furthermore the ambient air temperatures are low so that NPF precursor gases are more readily saturated or supersaturated.

These statements are included in the revised manuscript version (section 6 and section 4.1 "Relevance for the TTL"))

Specific Comments:

- p. 9251 l. 9, The term "ultrafine particles" is nowadays used for particles smaller than about 100 nm. The particle size range the authors refer to, 6-15 nm, is better described by "nucleation mode particles". Please change accordingly in the whole manuscript.

➔ Throughout the manuscript the term "ultrafine" is replaced by "nucleation mode" whenever the text refers to properties of particles in the size range of 6-15nm. Accordingly, the expression N_{uf} is replaced by N_{NM}

- p. 9252 l. 17, For the ultrafine particles (< 100 nm) sedimentation is not important, even at the low pressures at the tropical tropopause. Please remove "sedimentation" from the list of microphysical processes.

➔ "sedimentation" is removed

- p. 9252 l. 24, Remove "lower".

➔ "lower" is removed

- p. 9252 l. 26, The authors describe the stratospheric aerosol life cycle, but if you do so, please include all important processes, e.g., the sink of stratospheric aerosol particles is the transport into the troposphere, either by sedimentation, or tropopause folds, or downward transport in the polar vortex.

➔ The section has been modified, and now mentioned processes are considered.

- p. 9253 l. 29, The non-volatile components, please specify which inorganic salts are meant. And concerning the mineral dust, I doubt that it reaches altitudes of 11 km or above. I the literature one can find individual cases where this was the case, but mineral dust is likely not a component of the majority of aerosol particles in the UT/LS. If the authors are aware of references with opposite statements, please provide them to the reader.

➔ Literature references concerning the variety of non-volatile particles are given in the manuscript just one sentence before where we introduce non-volatile particles in general. We thought this the best place to provide the references. Indeed, the variety of non-volatile species that were found so far is large. But at this point of the manuscript there is no statement about abundances (neither absolute nor relative) of these species and we did not intend to make the reader believe that the mineral dust particles, for example, is a component of the majority of the aerosol.

Conceivable inorganic salts are, for example, ammonium sulfate and ammonium bisulfate

particles. Note that already very small fractions of e.g. inorganic salt are sufficient that the particle will not evaporate entirely during heating and the particle will be counted as non-volatile. A residuum of only 10 nm after heating will be sufficient that the particle is counted as non-volatile.

- p. 9254 l. 20, Barth et al., 2001 is only one reference. There are other references which state that only a smaller fraction of the SO2 survives the transport inside deep convective clouds, and not the "major fraction" (e.g., 30% given by Ekman et al., 2006, J. Atmos. Sci.). In my point of view the fraction of SO2 reaching the upper troposphere is one of the important open questions with respect to UT/LS particle formation.

➔ The reviewer is correct; one cannot state that generally the majority of the original SO_2 reaches the outflow of deep convection. We have corrected this in the paper, and added the reference to Ekman et al. (2006). However, this more detailed discussion supports our line of thought on transport of SO_2 in deep convection and subsequent aerosol nucleation in the convective outflows: Ekman et al. (2006) find that "Polluted air is transported from the boundary layer during the convective event, and relatively high concentrations of chemical compounds such as CO and SO_2 can be found at the top of the cloud." Further, in the same chapter, Ekman et al. (2006) write "The concentration of SO_2 in the anvil is relatively high and a substantial amount of new particles are formed (up to 1–3 x 10^4 cm^3). This number is in agreement both with observations downwind of convective clouds (Lee et al. 2004; Twohy et al. 2002; Clarke 1993) and with previous 2D simulations by Zhang et al. (1998)." The SO_2 in the simulations of Ekman et al. (2006) reaches 62 pptm (29 pptv) in the cloud top region, which is close to the 50 pptv of SO_2 used in our simulations.

Furthermore a publication by Crutzen and Lawrence, 2000 (The Impact of Precipitation Scavenging on the Transport of Trace Gases: A 3-Dimensional Model Sensitivity Study, Journal of Atmospheric Chemistry 37: 81–112, 2000) shows the following:

The transfer efficiency of SO_2 through a cloud depends strongly on the clouds liquid phase acidity. On a global average of cloud acidity the authors indicate that the uplift of SO_2 (Henry's law constant of about $5 \cdot 10^3$ mol l^{-1} atm^{-1}) into the upper troposphere is controlled by precipitation scavenging (SO_2 removal by 10-60%) or conversion of SO_2 by in-cloud oxidation with H_2O_2 and O_3. Under certain conditions, when the cloud droplet uptake of soluble acids e.g. HNO_3, H_2SO_4) is not balanced by the uptake of alkaline compounds, SO_2 will remain largely in the gas phase, allowing significant upwards transport of SO_2 to the upper troposphere (Crutzen and Lawrence, 2000).

Hence in the revised manuscript , the corresponding section has been modified and further references as well as previous argumentation were added.

- p. 9261 l. 13, Where there no trace gas measurements on the Falcon or are they not used?

➔ There were trace gas measurements on board the DLR Falcon. But as the Falcon has a maximum ceiling of 12km and as the Falcon generally followed other flight tracks or even made flights on other days than the Geophysica, the information from the trace gas measurement performed on board the Falcon are limited for interpreting the measurements of the Geophysica instrumentation.

The nucleation mode particle measurements on board the Falcon were used as additional information concerning the vertical distribution of potentially "preferred" NPF levels

But a detailed discussion of the lower level Falcon NPF observations is beyond the scope of this manuscript.

- p. 9262 l. 10, Where are the 31 additional level for simulating the SCOUT-AMMA flight located, were they just added above or was a higher vertical resolution used? Why are the numbers of vertical levels different for the two campaigns?

➔ We used meteorological data from the operational version of the ECMWF model for driving FLEXPART. ECMWF has increased the vertical resolution of its model between the two campaigns and, therefore, a different number of vertical model levels was used. The vertical resolution increased everywhere but particularly near the surface and in the stratosphere.

A clarifying statement has been added: "Generally 60 vertical levels were considered with additional 31 for the SCOUT-AMMA flights (due to ECMWF data which were available with increased vertical resolution for the SCOUT-AMMA campaign)."

- p. 9264 l. 29, The estimation of the pre-existing particle surface area, where does it come from, which numbers or references are used?

➔ In fact these numbers are guesses. There was no particular reference these values were based on. The idea was to cover a broad range of possible pre-existing particle surface area in order to have two extremes with initial steps to indicate a kind of progress of resulting newly formed particles with increasing/decreasing pre-existing particle surface area.

- p. 9265 l. 20, "The MAIA simulations of this work serve as sensitivity study" for what? Considering all the uncertainties likely only an estimate on which of the input parameters is the most critical one can be obtained, but probably not a quantitative result.

➔ The reviewer is correct, and the "sensitivity study" has been changed to "study" in the corresponding passages in the revised manuscript.

- p. 9268 l. 6, I dont see that the Falcon measurements exceed the Geophysica measurements. For an easier comparison, a log-scale on the x-axis of Fig 2.a would help.

➔ We do not agree with referee 1 at this point: One can clearly see in previous Figure 2a) (now Figure 3a) that between 340K and 350K the number concentration of nucleation mode particles from Falcon measurements (rose color) exceeds the values from the Geophysica measurements (orange and dark-violet, dotted).

The point is that a log scale would eliminate all data points where the number of nucleation mode particles is ≤ 0. These values are caused by circumstances that are naturally linked to instrumental measurements. Nevertheless, those data should be shown to illustrate that if NPF is observed than in quantities that are significantly larger than the instrumental uncertainties.

A misleading point with the figure was that the number of minor ticks should be changed. This is considered in the revised version.

- p. 9269 l. 1, Remove the sentence "It can be speculated …", it is of no help to understand the data.

➔ This sentence is removed.

- p. 9269 l. 11, What is the reason to provide the aircraft speed here. I my point of view the sentence can be removed. (And by the way, the True Air Speed, TAS, and not the ground speed would be relevant to derive spatial scales from measurement periods.)

➔ This sentence is removed.

- p. 9269 l. 16, Reading the description of Fig. 2 in the above paragraph is a little bit exhausting, wouldn't it be enough to describe the shape of the curves and to discuss the statistical numbers provided in Tab. 1?

- ➔ We have to admit that the description of previous figure 2 (now Figure 3) is detailed - according to the complexity of the figure. As the previous figure 2 (now Figure 3), in general, represents a comprehensive illustration of the observations, and all subfigures, in particular, highlight different aspects that have to be considered to make the conclusions that we make. We believe that the data have to be described and evaluated with care and in detail. This might lead to text sections that seem to be too descriptive but these are provided in this manuscript for the sake of avoiding misunderstanding.

- p. 9273 l. 22, As you are making case studies I would say that it is not totally impossible to distinguish between the different kinds of vertical transport. You could take respective weather maps and satellite pictures together with your FLEXTRA trajectory and see which kind of cloud (convective, frontal) is prevailing in the region with the strongest SO2 source contribution. This would be quite a simple exercise.

- ➔ We principally agree with the referee as a consultation of weather maps and satellite images would help to distinguish whether the uplift is driven by convection or frontal activity. In fact the text was phrased in a way that could have caused a misunderstanding. It is not important to distinguish whether or not the air was uplifted by convection or by slower, synoptic scale lifting. Important is, to quantify the uplift. Also weather maps and satellite images would not be able to give uplift rates of air parcels for respective cases. Thus the sentence is rephrased into: "Both lifting processes are considered within the FLEXPART simulations, nevertheless, it is impossible to quantify (i.e. by a lifting rate) respective contribution to the resulting uplift."

- p. 9273 l. 17, In order to estimate how reliable the SO2 loads to the air masses are (calculated with FLEXPART), please plot the whole flight track together with the respective SO2 loads in Fig. 4. In this way it would be possible to see, if the high SO2 loads at the time of particle formation are just coincidence.

- ➔ The previous figure 4 (now Figure 5) has been changed according to the referee's suggestion. Additionally the caption of previous figure 4 (now Figure 5) has been completed with the expressions within brackets: FLEXPART model simulation of SO_2 mixing ratio (colored) along the flight track (black line).

- p. 9276 l. 15, Where do the numbers for the particle surface area come from? Please give a reference or describe why you have chosen these values.

- ➔ As mentioned above these numbers are guesses. The idea was to cover a broad range of possible pre-existing particle surface area in order to have two extremes with initial steps to indicate a kind of progress of resulting newly formed particles with increasing/decreasing pre-existing particle surface area. Thus in section 4.4 of the revised manuscript following sentence is included to make this clearer: "…using values of 0.5, 1, 2, 4, and 8 $\mu m^2\ cm^{-3}$. The maximum and minimum values for the pre-existing surface area (i.e. 0.5 and 8 $\mu m^2\ cm^{-3}$) were chosen with the aim to cover a representative range that is realistic and that naturally prevails from boundary layer up to 13-15 km altitude. These aerosol surface area densities…"

- p. 9277 l. 7, The calculations of the new particle formation shown in Fig. 6 partly range back 6 days. I can not imagine that the respective air masses/trajectories didn't have any cloud contact within this period, even considering the high altitude of the events. Most of the trajectory path lies in the tropics!

And on p. 9280 you even state that "cloud processing along the trajectory" is likely. However, if a cloud contact occurred particle concentrations will be different (scavenging, particle formation). Hence how realistic/reliable is this modeling?

➔ In one of the four case studies in this manuscript, indeed a MAIA run covers a time span of six days, the other runs are shorter.
Where we wrote "...efficient cloud processing appears very likely if the SO_2 is to be reduced to values around 50 pptv." we were referring to cloud processing in deep convection that takes place before the start of the MAIA trajectories.
In section 5.4 we have provided a detailed list of the uncertainties that affect the MAIA simulations. However, even with these uncertainties, the main significance of the simulations can be upheld: The FLEXPART model has determined that trajectories from regions with elevated SO_2 emissions lead to locations with observed NPF events; the MAIA simulations support a causal link between the SO_2 and the NPF events, even with substantial removal of SO_2 by deep convection before the start of the simulated trajectories. Without the MAIA simulations, only speculations on a causal connection between the SO_2 and the new particle formation would be possible. We have addressed the referee's comment by including an additional point in the list in section 5.4:

"(6) a growing probability, with increasing time span covered by a MAIA run (i.e. up to 6.5 days for one of the case studies shown herein) that nucleation mode particle numbers at the end of the trajectories are unrealistically high, considering that in reality the formation of clouds (even at very high altitudes in the tropics) and interaction of those with the newly formed particles becomes likely with time."

Technical Comments:

- p. 9252 l. 18, The sentence starting with "Strong new…" prolongs over six lines. Please shorten or separate into two or three sentences.

➔ The sentence has been split.

- p. 9253 l. 25, Make the description of the aerosol particle composition a new paragraph.

➔ Done.

- p. 9257 l. 12, I'm not a native speaker, but "beside others" seems not to be the right expression to me.

➔ *"beside others" is replaced by "amongst others"*

- p. 9257 l. 14, Please use "LMS" for lowermost stratosphere, because the lower stratosphere (LS) is a different region.

➔ The acronym "LS" is removed at this point as this is the only point that the expression "lowermost stratosphere" is used in the context of the subtropics. Further in the text indeed the lower stratosphere (LS) is meant.

- p. 9257 l. 20, "describe" instead of "described"

➔ Done.

- p. 9257 l. 20, Remove "at".

➔ Done.

- p. 9257 l. 20, What is meant with "transformation of H2O"?

➔ The expression "transformation" is replaced by "conversion"

- p. 9257 l. 21, The "impact" of what "on the stratospheric water budget"?

➔ In the revised manuscript version the sentence is now: ”The studies of Schiller et al. (2009) describe the transport and conversion of H_2O (e.g. for the formation of H_2SO_4) in the TTL and the impact of this conversion on the stratospheric water budget.”

- p. 9258 l. 3, Remove “Fierli et al.” reference, it is already provided in the precious sentence.

➔ Done.

- p. 9258 l. 21, Exchange “has been” with “was”.

➔ Done.

- p. 9258 l. 27, Exchange “size” with “particle”.

➔ Done, here and elsewhere when “size diameter” was written.

- p. 9259 l. 2, Exchange “These” with “Generally, CPC”.

➔ Done.

- p. 9259 l. 3, Insert a comma before “which”.

➔ Done.

- p. 9260 l. 2, Is “automation” correct or shouldn’t it be “automatic”?

➔ The dictionary says that “automation” is a valid expression.

- p. 9260 l. 8, “tropics of the Americas” should rather be “tropics of Central and South America” as I don’t see any tropics in North America.

➔ Replaced according to the suggestion of the referee.

- p. 9260 l. 11, Remove the “of”s before the years.

➔ Done.

- p. 9260 l. 20, Remove “size”.

➔ Done.

- p. 9260 l. 21, Exchange “size” with “diameter”.

➔ Done.

- p. 9261 l. 2, Please specify what “in the vicinity of the aircraft” means, 10 m, 100 m, 1000 m?

➔ In the revised manuscript version the sentence is now: “Responding to the presence of optically active particles with sizes above 0.2 µm the MAS also provides measurements of aerosol backscatter and depolarization (at 532 nm and 1064 nm wavelengths) in the immediate vicinity (3-10 m) of the aircraft with 5 s time resolution and 5 % accuracy.”

- p. 9263 l. 22, Please insert “freshly formed” before “aerosol particles”.

➔ Done.

- p. 9264 l. 3, Remove “(decay probability)”, this was already introduced as synonym for stability the sentence before.

➔ Done.

- p. 9264 l. 5, The sentence starting with “The thermodynamical data…” prolongs over six lines. Please separate into two sentences.

➔ Done.

- p. 9264 l. 10, A trifle, but sometimes “see” and sometime ”cf.” is used to refer to a paper or figure. Please harmonize in the whole manuscript.

➔ harmonized towards using “cf.”, here and elsewhere throughout the manuscript.

- p. 9265 l. 8/9, Please use either “particle size range” or “diameter range” but not “diameter size range”.

➔ Done.

- p. 9266 l. 15, “data set (…) is a summarized result”, please change wording.

➔ Done.

- p. 9266 l. 20, Change to either "is between" or "ranges from… to".

➔ Done.

- p. 9267 l. 13, Change "applied on" to "applied to".

➔ Done.

- p 9267 l. 14, The sentence "Intentionally a conservative…", please rephrase this sentence.

➔ The sentence was rephrased into: "For both instruments intentionally a more rigorous criterion was adopted for our data analysis which excludes the statistical fluctuations of the counters. NPF criteria, e.g. as suggested by Lee et al. (2003), on our measurements data would lead to more frequent identification of NPF events with longer duration of each event."

- p 9267 l. 18, Add an "s" to "lead".

➔ Done.

- p 9268 l. 6, Replace "according" with "respective", similar p. 9269, l. 12.

➔ Done.

- p 9270 l. 2, Shouldn't it be "Fig. 2c and d"?

➔ Corrected.

- p 9271 l. 26, Replace "by" with "according to".

➔ Done.

- p 9273 l. 16, Remove the heading "Results of…", it has no numbering and is actually not needed.

➔ Done.

- p 9274 l. 9, The reference "Kazil and Lovejoy, 2007" is missing in the reference list. Or it must be "Kazil et al., 2007" or "Kazil and Lovejoy, 2004" here.

➔ It should be "Kazil and Lovejoy, 2004" and is corrected.

- p 9282 l. 21, Remove "(Nuf)", you use this abbreviation already before.

➔ Done.

- p 9283 l. 15, Please use past tense when describing measurements.

➔ Corrected.

- p 9283 l. 18, Change "in average" to "on average".

➔ Done.

- p 9283 l. 21, Must be "Table 3", do. p.9284 l. 10.

➔ Corrected in both cases.

- p 9284 l. 12, Remove "and".

➔ Done.

- p 9286 l. 2, Change "aircraft borne" to "aircraft-borne".

➔ Done here and elsewhere in the manuscript.

- p 9286 l. 7, Please separate the sentence, the second sentence starting with "One NPF event…".

➔ Done.

- Tab 3: Please make the diameter ranges for the surface area a new line.

➔ This change would lead to misunderstanding. It has to be clear that the determined surface area is based on these diameter ranges. Making the diameter range a new line could be understood as if the aerosol volume backscatter coefficient and the aerosol volume depolarization are also valid for this diameter range only, which is not the case.

- Fig. 1: c) There are not "red circles" as mentioned in the figure caption.

➔ The figure caption (Fig 1) in the revised version of the manuscript was corrected. The expression in brackets was removed.

- Fig. 5: Caption: l. 3 Exchange "or" with "and".

➔ "or" was replaced by "and".

- Fig. 6: Please, make the figure larger.

➔ Actually, this is a problem with the online page layout. The figure is principally large enough but as the figure and the caption have to share one page the figure is shrinked.

- Fig. 6: Caption: l. 8 There are no "cyan circles" in the figure, but "green bars", please correct.

➔ The caption was rephrased.

Anonymous Referee #2

The one somewhat major comment has to do with Fig. 4 and the FLEXPART modeling. What is the grid scale for the FLEXPART model and how does it compare to the "spatial scale" in Fig. 4. Although the abscissa of Fig. 4 is time it represents a horizontal distance traveled by the a/c particularly for the South America flight. Since the air mass trajectories are so long for the NPF events, what leads to the high temporal/spatial variability of the so2 plumes?

➔ The FLEXPART calculations were not performed for a regular grid. Whenever the aircraft changed its position by more than 0.2 degrees, a backward simulation was initiated. Also, whenever it changed its altitude by 50 m below 300 m, 150 m below 1000 m, 200 m below 3000 m, or 400 m above, a new backward simulation was initiated. This somewhat varying resolution along the flight track is higher than the resolution of the ECMWF data. However, it is well known that Lagrangian models can resolve fine-scale structures, which can be below the grid scale of the driving meteorological data (here, 0.5 deg).

A more difficult and more important question is, what fraction of the time were so2 plumes observed corresponding to flight legs when there was no NPF observed? Would the so2 concentrations on other flight legs have a similar appearance as the ones associated with the two flight legs shown, even when no NPF events were observed? This question may exceed the work envisioned for this paper, but such a nice correspondence here, naturally raises the question as to how often such a correspondence is observed. NPF events were observed on half of the South American flights and all African flights, so if the modeling is not too hard, a statistical answer would strongly reinforce the importance of so2 to the NPF events observed.

➔ The previous figure 4 (now Fig 5) was complemented by a panel c) showing the relation between N_{NM} during NPF versus predicted SO2 along the Geophysica flight track and the figure caption was extended. Text was added (p9274,4): "In **Fehler! Verweisquelle konnte nicht gefunden werden.**c) the measured N_{NM} for all NPF cases are shown versus the predicted SO_2 mixing ratio along the flight track of the *Geophysica* for the TROCCINOX campaign. FLEXPART simulations for SCOUT-AMMA were made only for one selected flight. When NPF was observed while SO_2 was predicted to be lower than 0.1 ppbv this can still be sufficient for (moderate) NPF if other conditions promoting particle formation are favorable (low temperature, low condensation sink from preexisting particle surfaces and high relative humidity) (Kirkby et al., 2011). However, whenever FLEXPART predicted SO_2 to be elevated (> 0.1 ppbv) along the flight track, NPF was observed."

It would be nice to provide a rationale for the selection of case studies since there were many NPF events observed. Why for example was a ferry flight used for the South American case rather than a targeted scientific flight?

- ➔ A rationale for this selection is in fact given with (9271, 3-6): "The highest concentrations of nucleation mode particles N_{NM} were found to occur within considerable spatial range during two flights on 24 February 2005 (TROCCINOX, South America) and on 07 August 2006 (SCOUT-AMMA, West Africa). These two flights were therefore selected for detailed case studies, but also because of the strength of the observed NPF events. Moreover, the second flight allows for studying the presence of nucleation mode particles within clouds."
  The reason for the selection of the ferry flight is that here the most intense NPF event of largest spatial extent was observed throughout the mission in South America.
  None of the measurement flights were directly targeted to the observations of NPF events.

Detailed comments:

9255.3-5: It is not obvious how "mixing of two subsaturated air masses can result in supersaturation". If they are both subsaturated, mixing them will not change that.

- ➔ This sentence was rephrased into: "Isobaric mixing of two saturated (or slightly subsaturated) air masses can result in supersaturation (e.g. with respect to H_2SO_4) and enhance the NPF processes considerably (Khosrawi and Konopka, 2003)."

9257.20: … described the transport …

- ➔ Corrected.

9257.24 … predominantly originates from…?

- ➔ Corrected.

9266.20: 350-370 K is a little restrictive for the high concentration layers for SCOUT-AMMA, where the layers extend to 340 K.

- ➔ Sentence is rephrased into: " However, the comparison reveals one common feature: The altitude level which is characterized by highest particles number concentrations N_6 ranges from 350 K (340 K, for the SCOUT-AMMA observations) to 370 K, the bottom of the TTL."

*9267.6-20: Suggest replacing all of this text, which is quite confusing, with the following. For the COPAS measurements analysis a conservative definition of an NPF event is used, which includes the measurement uncertainty of 15% for each COPAS channel. Thus, a positive difference between N6 and N15 is only considered to be an NPF event if 0.85*N6 – 1.15*N15 > 100 particles per cm3. At background conditions, without NPF, the subtraction of number concentrations that are measured by two independently operated CPCs lead to positive and negative values of Nuf that statistically vary around zero. Intentionally a conservative approach was adopted here. Application of other NPF criteria, as suggested by Lee et al. (2003), on our measurements would lead to more frequent observation of NPF events with longer duration of each event.*

- ➔ This section was replaced in the revised version by: "For the COPAS measurement analysis a conservative definition of an NPF event is used, which includes the measurement uncertainty of 15% for each COPAS channel. Thus, a positive difference between N_6 and N_{15} is only considered to be an NPF event if $0.85 \cdot N_6 - 1.15 \cdot N_{15} > 0$ particles per cm^3 (if ≥ 100 particles per cm^3, the NPF event is considered to be of high intensity). At background conditions, without NPF, the subtraction of number concentrations that are measured by two independently operated CPCs leads to positive and negative values of N_{NM} that statistically vary around zero. A similar NPF criterion was applied to the measurement data from the DLR CPC system, assuming an instrumental uncertainty of 10 %. For both instruments intentionally

a more rigorous criterion was adopted for our data analysis which excludes the statistical fluctuations of the counters. NPF criteria, e.g. as suggested by Lee et al. (2003), on our measurements data would lead to more frequent identification of NPF events with longer duration of each event."

9267.11: I do not understand the point of the following sentence, "According to a measurement uncertainty of 10% an NPF criterion, equivalent to the one used for the COPAS data, was applied on the measurement data from the DLR CPC system." Was a different uncertainty applied to the DLR CPC system? What is meant with this sentence? Is this necessary to include?

➔ Indeed the DLR-CPCs are assumed by the operators to work with an uncertainty of 10% instead of the 15% uncertainty which is valid for the COPAS measurements. It is necessary to have this information to see that the criterion depends on the instrument uncertainty.

9268.6:… at corresponding altitude levels…?

➔ Corrected.

9269.12: Here and elsewhere (9276.8, 9284.5, …) this is not the proper use of according, in English. Replace it with corresponding.

➔ Done.

9269.6-16: The times of various NPF events and speeds of the various a/c are not all that interesting unless the picture is completed by providing the spatial scale of the events. This is what is interesting, not the time over which the events are observed in a speeding a/c. In fact, skip the detail about a/c speeds and times of observation, and just say that, based on the time over which the events were observed, and the relative a/c speeds, the spatial scale of the NPF events ranged from x – y km. Same for Table 1. Replace the last three columns with the distances representing the NPF events.

➔ We agree with the Referee in the point that comments concerning the spatial extension of NPF events (from NPF duration and air craft speed) in the text should be skipped.

Thus, in the revised version, the text on page 9269, lines 8-16, was replaced and the last sentence in the caption of table 1 was erased.

We do not agree with erasing the three last columns in Table 1. These are in fact the most accurate statements we can deliver about the "extension" of observed NPF events – on a scale of time! From the data in this table the reader gets information on the time span of NPF observations, on which level NPF occurs with largest "temporal extension", and on how much time the aircraft spent at this flight level.

To use the aircraft speed to determine a concrete spatial extension of NPF fields would be misleading as the real flight track (including turns, ascents or descents during NPF events) is not considered, here. Thus, this way calculated spatial extensions would likely be a strong overestimation of the true extension.

9271.13: … over Brazil…

➔ Corrected.

9273.20-23: Why mention 4-7 days? This represents less than 50% of the so2 in South America and well less than 50% in Africa. An equivalent or larger fraction of so2 comes from 15-20 days which implies synoptic scale lifting. This raises the question as to why the great temporal dependence of the so2, see the major question above?

➔ "The air masses were lifted to the measurement altitude within the last 4 to 7 days…." is mentioned to provide this information to the reader as the FLEXTRA result concerning the air mass trajectory is exactly the pathway along which the MAIA simulation is running. Thus,

MAIA can consider only what happens along this way over a certain time before the measurement happens (t=0).

9281.9-11: The standards for agreement seem to have relaxed somewhat here. The measurements fall within the four order of magnitude range covered by the model results, so yes it should be easy to find a combination of so2 concentrations and surface area which match the measurements; however, all would require low values of so2 and surface areas between 2 and 1 cm-3.

- ➔ This sentence was replace in the revised manuscript version by: "Thus, a low SO_2 mixing ratio (≤ 50 pptv) and only a narrow range of pre-existing aerosol surface area (between 1 -2 $\mu m^2\ cm^{-3}$) seem to be the preconditions for predicting the number concentration of nucleation mode particles with MAIA in the order of magnitude as observed by in-situ measurements."

9282.19-25: It is not clear where the CO mixing ratio numbers come from. For example in the CO range from 67-82, Nuf ranges from 4-4000 cm-3. But also in this Nuf range CO ranges from 67 to over 90. So where does 82 come from? I assume it is the highest CO value when Nuf < 100 cm-3, but even this isn't quite true as Nuf is < 100 for CO up to 85.

- ➔ The text was indeed more confusing than it should be. Thus criticized section was replaced by: "Generally, the occurrence of NPF events with N_{NM} of more than ~ 100 particles per cm^3 is accompanied by CO mixing ratio in excess of ~ 70ppbv (typical CO background levels are 60-70 ppbv in the pristine marine boundary layer). In particular, high number concentrations of nucleation mode particles (N_{NM} of more than 2000 particles per cm^3) occur when the analyzed air masses are CO enriched (> 85 ppbv of CO) at 350 K < Θ < 360 K. "

9283.24: Do not use a colon here. Make it two sentences to not confuse the reader.

- ➔ Corrected

9284.10: What is meant by, "(see … black curve in Fig. 10)? That curve is Nuf and not related to backscatter or depolarization.

- ➔ Corrected in the revised version.

9284.11-12: … up to and 4000 ???

- ➔ Corrected.

9286.9-11: Change to, "New particle formation, which was observed on half of all flights over South America and during each local flight over West Africa, was confined…",

- ➔ Changed according to referees suggestion

9286.14: Suggest, … were to large extent volatile…, but it should be quantified with a per cent or fraction.

- ➔ Corrected

9287.10-12: The conclusion from this statement, " Nucleation mode particles were detected not only in clear air but also within thin cirrus cloud layers, indicating that new particle formation occurred in both, clear air, as well as in clouds.", does not follow from the modeling shown here. The observation, as the modeling shows, is not made at the time of the new particle formation, but measures the residue of such an event. Thus, high concentrations of Nuf in a cloud does not imply NPF in the cloud, merely that the resulting Nuf were still present even in cloudy air. The authors do qualify this statement in the following text, but not before this bold statement is made. The qualifications should rule and the bold statement eliminated or strictly qualified.

- ➔ This text section was rephrased into: "Nucleation mode particles were detected not only in clear air but also within thin cirrus cloud layers. Mixing of cloud-free and cloudy air parcels, aerosol nucleation and cirrus formation occurring in sequence, and ice particles which have

sedimented into an air parcel where aerosol nucleation has occurred could explain the observation of nucleation mode aerosol inside cirrus clouds."

9287.19-22: These statements are again not strictly supported by the modeling or observations. In all case studies the times of nucleation in the NPF events always occurs 1-2 days prior to the observation. This is not consistent with nucleation occurring in thunderstorm anvils which are only a few hours old. Thus this has to be qualified as a pathway far upstream of the observations for providing source material to the TTL.

- ➔ This statement was mitigated, nevertheless it should be mentioned as an option which our observations and modeling do (admittedly) not proof but they do not exclude it, either. Thus the sentence is rephrased into: "Considering that this in-cloud observation occurred in the convective outflow region and, thus, in proximity of a tropical convective cell (cf. Frey et al., 2011), in-cloud (cirrus) NPF in the tropical UT is conceivable, if sufficient amounts of NPF precursors are carried aloft by deep convection from the boundary layer. Despite significant removal of the gaseous precursors by scavenging during the upward transport, enough gaseous material may remain inside the Cb anvil and its outflow to enable NPF."

Fig1: How are the N4 particles shown? There are no symbols included in the N4 legend. What do the percentile gray scales refer to, the N6 or N4 particles? Where are the "red dots"

- ➔ Both N6 (Theta>350K) and N4 (Theta <350K) median are shown in a black line. The percentiles gray scales refer to N6 (Theta >350K) and N4 (Theta <350K), respectively. The "red circles" is in fact a residual and erased in the figure caption.

Fig2: The caption is confused. The description for b) is for c), and c) for d), and there is no description for b). Where is the black scatter-line plot?

- ➔ Former Fig 2 caption (now Fig 3) was reorganized and rephrased

Fig. 4: Include on the figure, with another scale, the particle number concentrations as in Fig. 3, so the correspondence between so2 and NPF can be seen clearly.

- ➔ For visualization in Fig 4 grey shaded areas are included which clearly indicate the periods when increased number concentrations of nucleation mode particles were observed. In the revised manuscript version the range of the time axis was expanded to the full scale (according to a comment of Referee 1) for respective flight such that SO2 data are shown along the complete flight track – the grey shaded areas to mark the NPF observation periods are still included.

Fig. 6: There are no cyan circles on the plot.

- ➔ Corrected.

Anonymous Referee #3

Major comments:

1. As the authors have acknowledged, one limitation of their modeling approach is that the mixing of air masses with surrounding air, which are well-known to occur significantly in the convective outflows, was not considered. One question is that how this limitation may affect the conclusion of the modeling results that "predicted concentrations of ultrafine particles in reasonable agreement with the in situ observations." As can be seen from Figs. 3 and 10, high concentrations of Nuf exit in limited areas and are essentially zero in many areas. If the MAIA model allows the dilution (say by a factor of 10 or more) of freshly nucleated particles (Nuf), the simulated Nuf at t=0 in Fig. 6 may be much smaller than the observed values.

- ➔ Expressions like" predicted concentrations of ultrafine particles in reasonable agreement with the in situ observations" are removed or significantly mitigated in the revised manuscript version. According to comments by Referee 1 and 2 a major problem arose from this phrasing which probably led to the assumption that model results and observations were aimed to be qualitatively compared. The MAIA simulations were aimed to investigate how sensitive the model result reacts on the changing input parameters.

2. Figure 6. Based on modeling results, most of Nuf at t=0 were formed 2-6 days ago. Can you still called these particles freshly nucleated? In discussing Fig. 2, the authors mentioned that many nucleation events were observed over periods of 15 min or less. If these Nuf particles were formed 2-6 days ago, why they didn't spread to more wide regions?

In a few days, these particles might have been transported far away from cloud outflow regions. Base on Figs. 3 and 10, high concentrations of Nuf are limited to the cloud outflow regions. Could these observations suggest the observed Nuf (Figs. 3, 6, and 10) were formed on the day of the measurements instead of being formed

2-6 days ago somewhere else)?

- ➔ The term "freshly nucleated" has been removed in the revised manuscript where it is not applicable.

  In fact, the nucleation mode particles did spread over distances of many tens of km. The GEOPHYSICA aircraft flies with approximately 600 km/h, hence in 15 minutes, it covers a horizontal range of about 150 km.

  The MAIA simulations take place on FLEXPART trajectories, which do not resolve convection in detail. It is therefore conceivable that the observed particles formed on trajectories that differ from those provided by FLEXPART and used by MAIA. However, we have no indication that this is the case. In addition, Figure 3 and 10 do not indicate in our opinion that the measurements were taken close to the cloud outflow region.

3. Figure 6. The influences of initial SO2 concentrations and surface areas on the predicted Nuf are presented. Both SO2 concentrations and surface areas affect new particle formation by controlling sulfuric acid vapor concentrations ([H2SO4]). [H2SO4] is a more direct and important parameter. In addition, it is hard to infer the nucleation rates from Fig.6. I would suggest the time series of modeled [H2SO4] along with nucleation rates be presented in an additional figure or in supplementary materials. Such additional information will be very useful to the interested readers.

- ➔ In the revise manuscript version the suggestion of this referee is realized by replacing previous figure 7 by a new one (now figure 8) showing the time series of the nucleation rate along the trajectory. The data points are colored regarding the respective contribution of neutral and ion induced nucleation (using the aspect of former figure 6, now figure 7). The H_2SO_4 production is implicitly shown in this new figure 8 again as the nucleation rate depends on the presence of H_2SO_4.

4. p9278, lines 14-15. Why neutral nucleation dominates at night? What is [H2SO4] at the night? How do you separate ion versus neutral nucleation? Provide time series of [H2SO4] and nucleation rates (both ion and neutral) will help to clarify the picture.

- ➔ Overall nucleation rates are extremely small at night, and the apparent dominance of neutral nucleation can therefore not be interpreted in a significant way. Neutral and charged nucleation rates are calculated in the MAIA simulation as formation rates of neutral and

charged supercritical particles. Clarification should be provided by the new figure 8 (cf. previous comment.)

5. p9281, line 23. How do you maintain the constant surface area densities? The nucleated particles coagulate and grow and will contribute to the surface area densities. Since MAIA is a size resolved and kinetic model, it will be useful if the authors can provide a figure showing the evolution of particle size distribution (dN/dlogDp versus time) for a representative case.

- ➔ The initial surface area densities are not held constant in the course of the MAIA simulations. The text has been corrected in the revised manuscript as follows: "These aerosol surface area densities change in the course of the runs due to growth and coagulation of the pre-existing aerosol particles."

6. Section 3.3. Please give a brief description of size ranges and number of bins used in the model.

- ➔ Citations to works which describe the MAIA model are given in the manuscript.

Minor comments:

1. p9252, lines 8-11. Not up to date and might be not true anymore. Should check and discuss more recent global NPF studies using various schemes for example: Yu, F., G. Luo , T. Bates , B. Anderson , A. Clarke , V. Kapustin , R. Yantosca , Y. Wang , S. Wu: Spatial distributions of particle number concentrations in the global troposphere: Simulations, observations, and implications for nucleation mechanisms, J. Geophys. Res., 115, D17205, doi:10.1029/2009JD013473, 2010.

- ➔ The section was rephrased in the revised manuscript version into: "However, an evaluation of the different mechanisms forming new particles by Lucas and Akimoto (2006) revealed that qualitatively, due to the large uncertainties of parameterizations, NPF was so far inadequately represented in global models. Recent studies (Yu et al., 2010; Zhang et al., 2010) suggest that uncertainties in the effectiveness of NPF mechanisms themselves may cause problems in simulating global particle number densities."

2. p9252, line 20. Why at ocean surface? DMS is emitted from ocean and then gradually oxidized to become SO2.

- ➔ Sentence was rephrased.

3. p9256, lines 9-11. Should provide some representative references for this statement.

- ➔ This sentence is kind of a conclusive statement concerning the previous text section which includes several references that provide different opinions concerning the contribution of ion-induced NPF. Thus we believe to have sufficiently demonstrated the controversial discussion on this issue. However, in the revised manuscript version this sentence and the following was removed.

4. p9279, last paragraph. The picture will be clearer if the [H2SO4] values were given.

- ➔ Most likely solved by the modifications that result from major comments 3 through 5 by referee 3.

5. p9286, line 4 and line 14. Not "recent new" or "freshly" based on Fig. 6 (formed 3-6 days ago).

- ➔ On the one hand, as one conclusion of comment 2 by referee 3, the location and proximity to the convective outflow suggests that these particles were "recently" or "freshly" formed. The fact that these nucleated particles, according to the MAIA modeling, might be processed over time span (like days) cannot exclude to call them "recently formed" in contradiction to particles from e.g. volcanic eruptions which have live times in the UT/LS of weeks and months.

Additionally added, replaced or moved text – independent on the referee's comments:

P9258, 4-12 is moved to section 2.1

p9259, 27 here ending sentence is completed by the add: "…and FCAS (Focused Cavity Aerosol Spectrometer; Jonsson et al., 1995)" Also the new reference is included in the revised manuscript version.

p9260, 4 following sentence is included: "FCAS analyses scattered light by aerosol particles revealing aerosol size distribution between 0.06 µm < d_p < 2 µm within 31 detector channels (with mean $\Delta d_p/d_p$ = 0.111). NMASS and FCAS data which were acquired in recognized aircraft plumes and in air with relative humidity with respect to ice exceeding 95 % were excluded."

P9257, 10 added text: "(2) during the SCOUT-O3 mission (November – December 2005, from Darwin, Australia, at 12.5°S, 130.8°E; cf. Vaughan et al., 2008; Brunner et al., 2009; http://www.atmos-chem-phys.net/special-issue124.html) and "

P9257, 21 added text: "An objective of the SCOUT-O_3 campaign was the characterization of the "Hector" storm system that develops north of Darwin, over the Tiwi Islands, nearly on a daily basis during the pre-monsoon season (Connolly et al., 2006). The top of the developing storm cell is capable of reaching altitudes of up to 20 km, which indicated that "Hector" could play an important role for the vertical transport of aerosol particles and trace gases into the TTL or into the lower stratosphere. Further information on the meteorological situation during SCOUT-O_3 are provided by Brunner et al. (2009)."

P9258, 11 sentence extended: "Details on the aircraft, the instrumentation, and the campaign deployments can be obtained for TROCCINOX from http://www.pa.op.dlr.de/troccinox/, for SCOUT-O_3 from http://www.ozone-sec.ch.cam.ac.uk/scout_o3/ and for SCOUT-AMMA from http://amma.igf.fuw.edu.pl/."

P9267, 21 a further section is in the revised version together with a new figure (now figure 2): "The two NPF criteria (Lee et al., 2003 and the one described here) are compared for 5530 size distributions measured with the NMASS-FCAS system on board the NASA WB-57F (**Fehler! Verweisquelle konnte nicht gefunden werden.**). The fraction of 16 % of the NMASS-FCAS size distributions met our NPF criterion and highest fractions were found between potential temperatures of 325 K and 375 K. A fraction of 48 % met the Lee et al. (2003) criterion and large fractions extended from 325 K to 395 K. Size distributions meeting the Lee et al. (2003) criterion but not ours had distribution functions that increased as particle size decreased from 10 nm to 4 nm in diameter. However, those NPF events meeting our criterion had a larger fraction of the particle in the 4 – 10 nm diameter range. This suggests that our criterion selected those events with particularly small particles while the Lee et al. (2003) criterion included events that had aged and lost some of the particle fraction in the 4 – 10 nm diameter range by coagulation. Thus, the conservative approach of the present work selects particularly those NPF events with the smallest, most recently nucleated particles, and the conservative criterion narrows down the location where these particles are found to the lower TTL region.

Expressions like "d_{p50}" and "d_p" are written in italic characters in the revised paper version.

The abstract was rearranged for the sake of being better structured.

Two more Co-authors were added:
C. Schiller, Institut für Energie- und Klimaforschung – Stratosphäre, Forschungszentrum Jülich, Germany and

T. Peter, Institute for Atmospheric and Climate Science, Eidgenössische Technische Hochschule (ETH), Zürich, Switzerland.