Anonymous Referee #1

Comment:
The authors present a novel approach for model-measurement comparisons which introduces the use of several new metrics relevant to particle formation and growth. Using a global aerosol microphysics model and measurement data from 5 locations, the authors evaluate their new metrics, concluding that global models are appropriate tools for assessing the contribution of processes like nucleation to particle size and number. The paper is well written, interesting and certainly within the scope of ACP; I would recommend publication, following clarification on the below, minor issues.

Comment:
p8336, line 5: Rephrase "With increasing aerosols: : :" to be more specific?

**Response:**
**Changed so that revised manuscript now reads "With increasing aerosol number concentrations…"**

Comment:
p8338, line 19: since you go on to use this tuning factor for your ternary nucleation simulations some discussion of the value used might be helpful to the reader

**Response:**
**The manuscript now reads: "A modified version with a globally constant nucleation rate tuning factor of $10^{-5}$ has been incorporated into a regional aerosol model and shows reasonable agreement (Jung et al., 2010) with observations. The following results for ternary nucleation presented in this work include the $10^{-5}$ tuning factor, which was chosen on the basis of improvement of nucleation rate and aerosol number concentration predictions."**

Comment:
p8339, line 17: ": : :: : :.self-coagulation is much smaller than: : :" rephrase to clarify what you mean by smaller than, i.e. less important?

**Response:**
**We mean that the rate of coagulation growth is much slower than condensation growth. We agree that the wording is a bit confusing and the manuscript has been updated:**

**"…although this rate of coagulation growth is much slower than condensation growth and can generally be ignored."**

Comment:

The model has relatively coarse horizontal resolution, do you linearly interpolate acrossa djacent grid-cells to the observation locations? Perhaps add this at p8346, line 15-18 if so

**Response:**
**No, we do not do any linear interpretation. The values in the model are simply the corresponding grid boxes that contain the latitude/longitude points of the 5 sites. The coarse resolution probably limits the accuracy of predictions at the more urban sites compared to more remote locations (Hyytiala). Added the following: "We do not interpolate number size distribution values within the given grid cell. The horizontal resolution used here is a challenge for comparisons with observations, and results will generally show better comparisons to rural observations. This should be addressed in the future with improved model resolution, but additional long-term data sets of aerosol size distributions in rural locations are needed."**

Comment:
p8353, lines 20-28: both nucleation mechanisms seem to under-predict the number of event days during April at Hyytiala, any ideas why?

**Response:**
**This is likely due to an underprediction in sulfuric acid concentrations , consistent with the CDF in Figure 10. Added the following: "The model underpredicts nucleation days for both ternary and activation in April at Hyytiälä, most likely due to sulfuric acid concentration underpredictions (see Fig. 10)".**

Comment:
p8354, line 3: it looks like the near exact prediction in April and May at Pittsburgh is only with the ternary mechanism?

**Response**:
**That is correct. It is now mentioned in the revised manuscript:**

**"…with near exact prediction by the ternary simulation in April and May at Pittsburgh."**

Comment:
p8356, line 23-25: true, but the activation median is even closer (than ternary) to the observed median?

**Response:**
**Also correct. We note that the median isn't the only way to determine the accuracy of our predictions. The mean value is better with ternary, as is the LMNB (if only slightly). We have updated the manuscript so as not to imply ternary is necessarily better than activation.**

**"Generally, sulfuric acid is accurately predicted in the ternary and activation model as the median values agree within about 40%."**

Comment:
p8357, line 9: ": : :: : :small positive bias: : :: : :", based on what? The median and mean CoagS values are lower than the observed value

**Response:**
**This appears to be an oversight on our part. Indeed, the mean and median values are lower in the model than observed, and the LMNB values are negative. The manuscript now reads small <u>negative</u> bias.**

Comment:
p8358, line 11-12: Would be useful to give the relative growth rate numbers somewhere in this paragraph; from the plot it's difficult to compare the relative contributions of organics and sulfuric acid in the model v. measurements, since the total growth rate is different

**Response:**
**Yes, we agree. They are now mentioned in the paragraph.**

**"growth (measurements at HYY: 2.5 nm h$^{-1}$ organic, 0.4 nm h$^{-1}$ sulfate, model at HYY: 1.6, 0.1, respectively)"**

Comment:
You say here that this approach ": : :.should help isolate individual processes biasing model predictions." Have you been able to do this in this study? It would be useful here to give an example of how your approach has allowed this, if it has, I'm not sure...

**Response:**
**We believe it has. For instance, we can say that for the most part, growth rates were not underpredicted in the model (despite lower amounts of SOA), so condensational growth in the model is likely not biased low. Similar things can be said for the coagulation process (comparing the coagulation sink).**

**Additionally, looking at Fig. 12, the model bias in the survival probability (SP100 especially, but also SP50) seems to be mainly a result of the low bias in the 3 nm formation rate (J3) and not the growth rate (GR). Growth rates have a LMNB of around -0.1 to +0.1, whereas J3 are worse.**

**We have added a few sentences to the conclusions and a sentence to the abstract to help summarize the overall take home message:**

**Conclusions:**

**"The model-measurement comparisons have also helped us identify the processes that lead to biased model predictions. For example, growth rates are predicted well across the sites (Fig. 12), whereas nucleation rates are more biased, causing most of the biases in CCN formation. Despite the use of a lower bound SOA source for organic growth, it is likely the nucleation rates and not the growth rates that cause a slight underprediction in quantities such as survival probability and CCN formation rates."**

**Abstract:**
**"...mostly due to bias in the nucleation rate predictions...**

Comment:
p8384, Figure 10: Panel A is also a CDF?

**Response:**
**Corrected in the revised manuscript.**

Comment:
p8386, Figure 12: You could perhaps add additional lines at (or shading between) +0.3 and -0.3 to emphasise the region that indicates a factor of 2 difference to aid the reader. You could move the explanation that is given in the text (p8359, lines 9-11) to the figure caption, on its own this figure is difficult to interpret.

**Response: We've added lines for +0.3 and -0.3, and also added to the caption to make things more clear.**

Technical Comments:

p8344, line 25: Is the Riipinen 2010 reference correct? should it be 2011?

p8345, line 10: Is the Spracklen 2009 reference correct? should it be 2010?

p8353, line 14-16: this is just the ternary model?

p8354, line 12: insert "," between Hyytiala and Pittsburgh

p8355, line 24: I think these are the mean values? rather than the medians

p8361, line 2: replace "they" with "it"

A few acronyms are used, e.g. CS for condensation sink, SP for survival probability: would be better to define the terms and then use consistently

**Response:**
**These have all been taken care of. Thanks.  The numbers above the bars in figure 3 are for ternary and activation nucleation simulations.**

Anonymous referee 2:

Comment:
This paper represents an interesting model vs. measurement comparison related to cloud condensation nuclei formation due to atmospheric nucleation. The paper is suitable for publication in ACP after the authors have addressed the few issues raised below.


Comment:
The authors have chosen to include a long, review-type introduction in their paper. I am fine with such an approach, but with the reservation that it should be made more carefully than done here. Firstly, I do not think that the chosen literature reflects the current understanding on atmospheric nucleation mechanisms, nuclei fate (growth vs. scavenging) and resulting CCN production to the extent that would be desirable for this kind of an introduction.

**Response: We have searched for more literature and mentioned and added several more references. For nucleation mechanisms, we add a citation to Chen et al. (2012), Zhao et al (2011), and Kulmala et al. (2013). For nuclei fate and CCN production, we cite and mention Lee et al. (2013) and Kerminen et al. (2012). Added to text:**

**"Recently, Zhao et al. (2011) observed neutral clusters of sulfuric acid and amines plus ammonia in atmospheric measurements for the first time, and Chen et al. (2012) proposed an acid-base nucleation mechanism involving these chemical species which achieved nucleation rate closure to within a factor of 10. Additionally, (Kulmala et al., 2013) have observed atmospheric nanoparticles and clusters as small as 1 nm in diameter and concluded that these particles are most likely comprised of sulfuric acid, strong bases, and organic vapors."**

**"Recently, (Lee et al., 2013) compiled 28 model parameters covering all important aerosol processes and ran Monte Carlo simulations to determine the uncertainty in CCN concentrations due to each parameter. They found that although roughly 40% of CCN are attributed to nucleation, CCN are generally insensitive to the nucleation rates across a wide range of boundary layer and free tropospheric nucleation assumptions."**

Comment:
Secondly, the text should not contain loose statements, or statements that may not be right. For example, is there a firm basis to claim that nucleated cluster are stable in the atmosphere and that their initial size is typically 1 nm (page 8336, lines 13-14)?

**Response: We have changed the text to say to avoid specifying a hard-and-fast size definition. The manuscript now reads: "Nucleated particles have initial sizes on the order of a few nanometers or less, which is much smaller than typical primary emission size ranges (Kulmala et al., 2000, 2004; Mäkelä et al.,1997.; Vehkamaki et al., 2004, Kulmala et al. 2013). "**

Comment:
Third, I have a hard time of seeing how large nucleation rates and nucleation probabilities as low as 10 to the power -8 would be related to each other as stated on page 8337, lines 9-13. Think an extreme case where nucleated particles grow solely by their self-coagulation, not by condensation at all. It takes a thousand 1 nm particles to make one 10 nm particle, so the survival probability against self-coagulation would be 0.001 at 10 nm and 10 to the power -6 at 100 nm. Lower survival probabilities are possible only if larger pre-existing particles are the main sink for growing nuclei, but if this is the case then the survival probability would be independent of the nucleation rate (as nuclei themselves do not contribute to the sink).

**Response: These values came from Pierce and Adams (2009b**). A detailed explanation of those results here would be tangential, so we simply insert a very brief explanation.
**" The reason for these low survival probabilities is a nucleation-CCN feedback mechanism in which higher nucleation rates lead to faster coagulation removal rates and a higher condensation sink, which in turn lowers survival probability and dampens CCN formation."**

**Pierce, J. R. and Adams, P. J.: Uncertainty in global CCN concentrations from uncertain aerosol nucleation and primary emission rates, Atmospheric Chemistry and Physics, 9(4), 1339–1356 [online] Available from: //000263642000018, 2009b.**

We have confidence in the model result that faster nucleation rates cause a strong decrease in survival probability such that CCN concentrations are fairly insensitive to changing nucleation rates. However, we are aware that these model results are counter-intuitive to many researchers. In fact, we are close to submitting a manuscript entirely devoted to answering this question in nitty gritty detail, so we ask the reviewer's patience in this regard.

Briefly, the reviewer is right that coagulation with pre-existing larger particles is a dominant loss mechanism controlling the survival probability of nucleated particles. Faster nucleation rates do increase CCN concentrations somewhat, but the CCN mode is centered at smaller sizes and has a greater surface area. This has two effects. First, condensational growth rates are slowed. Second, pre-existing particles become more efficient in scavenging nucleated particles. We will show in our upcoming paper that these two effects explain quantitatively the decreased survival probabilities predicted by our model. It is also worth noting that the Pierce and

Adams paper used the fast nucleation rates with an untuned ternary nucleation parameterizations, so the results are an extreme case.

Comment:
Forth, is there some specific reason to select these two sites (Hyytiala in Finland and South Africa) when discussing observed particle growth rates? Certainly, both clearly smaller and larger growth rates have been observed in other locations.

**Response: The South Africa site is among the fastest with median growth of 9 nm hr$^{-1}$. But we agree that the choice of these two sites is somewhat arbitrary. The text now reads:**

**"Diameter growth rates from 3 to 25 nm during nucleation events in 2007 at Hyytiälä, Finland have a median value around 2 or 3 nm hr$^{-1}$, although median rates much faster and slower have been reported elsewhere."**

Comment:
Finally, what is meant by CCN sensitivity to nucleation and can it be measured using percentages? Normally, sensitivity means how much a change in one quantity affect a change in another quantity.

**Response:**
**By sensitivity, we mean the percent increase in CCN concentrations from a particular increase in the nucleation rate. For example, if we change the nucleation rate in the model from binary to ternary (~4 orders of magnitude), by what percent does CCN increase? Most global nucleation-CCN modelers have chosen to present CCN sensitivity to nucleation in this manner. So it is not a sensitivity in the formal sense, dCCN/dJ (or d(percent CCN)/d(percent J), although it is qualitatively similar.**

**We have added this brief phrase where sensitivity is first mentioned to help clarify:**
**"(percent changes in CCN concentrations due to changes in the nucleation rate)"**

Comment:
The survival probability requires some clarifications. First, the definition of this quantity in section 2.5.4 includes coagulation losses only. The authors should tell the readers that this is not necessarily the real survival property of the nucleated particles because they may be removed by other ways before reaching CCN sizes. For example, in many cases removal by wet scavenging is much more efficient than removal by coagulation.

**Response: Typically in the model, wet deposition lifetimes are on the order of several days to one week, whereas coagulation loss lifetimes are on the order of hours (see Fig 10c). Additionally, we are concerned here specifically with nucleation and growth days, which tend to be photochemically active, cloud and rain-free days. If nucleated particles were later being removed by wet scavenging, we would see that in the banana plots. A growing nucleation mode would disappear or at least lose a lot of number concentration. In this case, the day wouldn't be classified as a nucleation event. Added to manuscript:**

**"We do not consider the loss due to wet deposition, which typically occurs at much slower timescales than coagulation (days versus hours). Also, this comparison focuses on nucleation and growth days, which tend to be free of clouds by their nature. Additionally, if wet scavenging did happen, it should be noticeable in the size distribution evolution plots."**

Comment:
Second, how the survival probability is determined in practice from measurements and model simulations? By determining the relevant time scales from observations and simulations and then using equation 6, or by trying to determine this quantity somehow more directly from observations and simulations?

Response:
**We have clarified the calculation with the following addition to the manuscript:**

**When calculating survival probabilities from observations, using Eq. 6, we take the coagulational and condensational timescales from Eqs. 4 and 5, in which the growth rate is inferred from observations, the size-resolved number concentrations are also from observations, and coagulation coefficients are calculated from theory. Therefore, although some theoretical calculations are required, the key inputs related to condensational growth and coagulational scavenging, are based on observations. An analogous procedure is used for calculating survival probabilities from model nucleation and growth events.**

What are the related uncertainties in calculated survival probabilities and do these uncertainties affect the model-simulation comparisons in Figures 5-9?

**Response: The calculation procedure is the same for the model output and the measurements. Thus, if there are uncertainties, figures 5-9, the model-measurement comparisons, are subject to the same uncertainties and should be valid. That being said, any uncertainty would from two places: 1) the model itself, 2) the instruments themselves, and 3) the assumptions made in the survival and growth calculation. The first is the overall goal of the paper and we assume that the second is small in this context, so we will focus on (3). Nucleation rates assume a 3-25 nm mode, which may overlap slightly with**

primary sources especially at the urban sites, biasing our nucleation rates high. Growth rates have a similar uncertainty. Although a distinct and noticeable mode is usually observed in both the model and the measurements, there is a possibility that primary particles may contaminate the growth rate. Finally, the survival probability calculation assumes that the entire range of growth (3-100 nm) can be modelled using two growth rates (3-25 and 25-100). Compared to using an instantaneous growth rate, this may bias survival probabilities higher than expected.

The following discussion has been added to the manuscript:
"There are also uncertainties in the calculations of the modeled and measured growth rate, nucleation rates, survival probabilities, etc. One uncertainty is that primary particles may make some contribution to N3-25, therefore contributing to the apparent nucleation rate. A second uncertainty is our use of two growth rates for each nucleation event. However, since we apply the same calculations under the same assumptions for the model output and the observations, our model-measurement comparison results should be largely unaffected."

To more quantitatively address the survival probability uncertainty, we show below a scatterplot of observed survival probability at Hyytiälä using two methods: our method identified in the paper, and the method described by Kuang et al. (2009) in which the ratio of N100 to N3 for a given growth trajectory is defined as the survival probability. Note that there are some differences, yet both methods appear to yield similar conclusions. The vast majority of the days have nearly identical survival probabilities with both methods. More days have larger SP in the Kuang et al. (2009) method than the Westervelt et al. (2013) method. We leave a more thorough analysis of survival probability intercomparisons to a future paper.
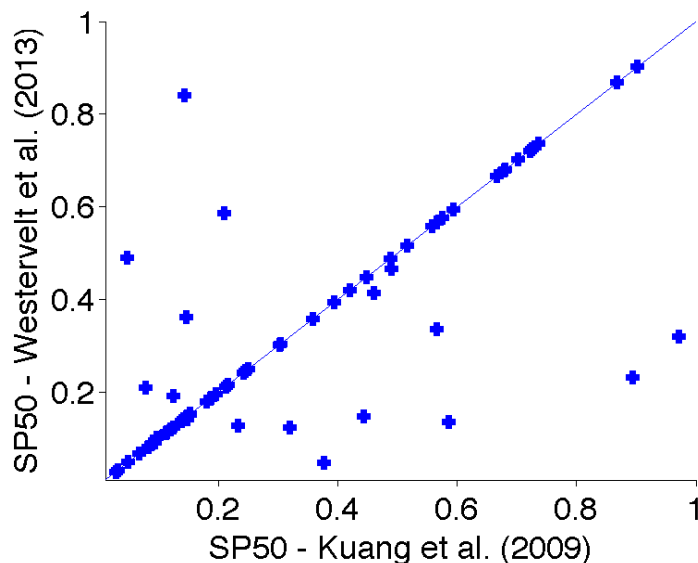
Figure 1: Survival probability to 50 nm in Hyytiälä observations using two different methods

Comment:
Thirdly, nucleated particles grow, on average, too slowly too reach 100 nm within one day at most of the sites? The authors mention this briefly, but do not discuss the consequences of this fact on the results and their interpretation.

**Response:  Yes, this is correct, growth to 100 nm is not very common. We have added significant discussion regarding the single-day assumption.:**

**"For the purposes of model-measurement comparison, we do not consider growth of particles beyond the nucleation day. There are a few reasons for this. First, air masses often shift over a given location after one day, making it difficult to make model-measurement comparisons after the first day. Secondly, if the model compares growth and loss processes well for the first day of nucleation, we can expect the model to perform similarly over subsequent days. Finally, we acknowledge that for the goal of defining the CCN budget contribution by nucleation, multi-day growth is more important. We leave that for future work and focus on model-measurement comparison here, where the use of single day estimates of nucleation metrics is justifiable. "**

Comment:
Figure 3 and related text: depending on the assumed nucleation mechanism and site, the model predicts whether nucleation takes place or not with 50 to 64% probability. Noting that purely random guess should produce 50% probability of being right, these number do not sound very large. The authors should open up this issue a bit further and not just state about the overall success of the model (page 8360, line 5).

**Response:**
**We agree. We have toned down the discussion in the results stressing that the model is able to get the overall number of events (1/4 to 1/3 of the year) with decent accuracy. We have completely revised the discussion of the model's (limited) ability to get individual days right to read as follows:**

**"It is encouraging that the model can predict accurately the number of nucleation days in one year, but the model shows only limited skill in correctly forecasting whether nucleation occurs or not on any given day. "**

and

**"Besides the accuracy of the nucleation parameterization itself, a host of related variables and processes have to be predicted correctly to get each day right: precipitation, cloud cover, and the emissions and transport of precursor**

species such as sulfur dioxide and ammonia, for ternary nucleation. Future work should investigate whether the limited skill in predicting day-to-day variability results primarily from weaknesses in the nucleation parameterizations themselves or from other processes controlling precursor concentrations."

Comment:
Most of the measurement sites used in the analysis are kind of urban locations, and no remote site is included. This should be mentioned somewhere in the text. Does this fact have any consequences on how the main conclusions can be generalized?

Response:
The nearest urban locations to Hyytiälä are 50 km to the southwest (Tampere) and 100 km to the northeast (Jyvaskyla). These are not large cities and probably have limited influence on the Hyytiälä observations and the large model grid cell. San Pietro Capofiume is neither urban nor pristine; it can best be described as polluted continental background, although still challenging for the horizontal resolutions used here. We mentioned in section 2.3 the following: "
These locations span a range of conditions, making the set a good test for a global aerosol microphysics model. For instance, growth at Hyytiälä is dominated by organic condensation (Riipinen et al., 2011), whereas at Pittsburgh, sulfuric acid condensation is the leading mechanism for particle growth (Jung et al., 2010). Urban, polluted continental, and clean continental sites are represented, although the fact that 3 out of 5 sites are urban complicates comparison with a global model."

We have also pointed out that at urban locations (STL, ATL) our model comparisons seem to be the poorest. Finally we mention the need for more remote site measurements in the conclusions.

Comment:
Beginning of section 2.2.1: EC is not part of the organic aerosol.

Response: Change to carbonaceous aerosol

Comment:
Section 2.2.2: The authors should mention briefly that organic may influence not only the growth rate but also the nucleation rate.

Response: Fixed.

Comment:
page 8351, line7: Seinfeld and Pandis (2006), not 2005.

**Response: Fixed.**

Comment:
page 8357, line 7: what is meant by feedback between nucleation rate and condensation sink?

**Response: This was explained in an earlier comment. We are preparing another paper for submission very soon that will explore this feedback in greater detail.**

**References:**

Chen, M., Titcombe, M., Jiang, J., Jen, C., Kuang, C., Fischer, M. L., Eisele, F. L., Siepmann, J. I., Hanson, D. R., Zhao, J. and McMurry, P. H.: Acid-base chemical reaction model for nucleation rates in the polluted atmospheric boundary layer., Proceedings of the National Academy of Sciences of the United States of America, 109(46), 18713–8, doi:10.1073/pnas.1210285109, 2012.

Jung, J., Fountoukis, C., Adams, P. J. and Pandis, S. N.: Simulation of in situ ultrafine particle formation in the eastern United States using PMCAMx-UF, J. Geophys. Res., 115, D03203, doi:10.1029/2009JD012313, 2010.

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, Atmospheric Chemistry and Physics, 12(24), 12037–12059, doi:10.5194/acp-12-12037-2012, 2012.

Kuang, C., McMurry, P. H. and McCormick, A. V.: Determination of cloud condensation nuclei production from measured new particle formation events, Geophys. Res. Lett., 36, L09822,doi:10.1029/2009GL037584, 2009

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. P., Paasonen, P., Mikkilä, J., Vanhanen, J., Aalto, J., Hakola, H., Makkonen, U., Ruuskanen, T., Mauldin, R. L., Duplissy, J., Vehkamäki, H., Bäck, J., Kortelainen, A., Riipinen, I., Kurtén, T., Johnston, M. V, Smith, J. N., Ehn, M., Mentel, T. F., Lehtinen, K. E. J., Laaksonen, A., Kerminen, V.-M. and Worsnop, D. R.: Direct observations of atmospheric aerosol nucleation., Science (New York, N.Y.), 339(6122), 943–6, doi:10.1126/science.1227385, 2013.

Lee, L. a., Pringle, K. J., Reddington, C. L., Mann, G. W., Stier, P., Spracklen, D. V., Pierce, J. R. and Carslaw, K. S.: The magnitude and causes of uncertainty in global model

simulations of cloud condensation nuclei, Atmospheric Chemistry and Physics Discussions, 13(3), 6295–6378, doi:10.5194/acpd-13-6295-2013, 2013.

Pierce, J. R. and Adams, P. J.: Uncertainty in global CCN concentrations from uncertain aerosol nucleation and primary emission rates, Atmos. Chem. Phys., 9(4), 1339-1356, doi:10.5194/acp-9-1339-2009, 2009b.


Zhao, J., Smith, J. N., Eisele, F. L., Chen, M., Kuang, C. and McMurry, P. H.: Observation of neutral sulfuric acid-amine containing clusters in laboratory and ambient measurements, Atmospheric Chemistry and Physics, 11(21), 10823–10836, doi:10.5194/acp-11-10823-2011, 2011.