This paper describes a new option for WRF/Chem to allow the simulation of aerosol direct and indirect effects (based on the existing aerosol microphysics scheme (including a VBS treatment of SOA)) and their feedbacks on clouds and meteorology. The paper is interesting and certainly within the scope of GMD, I would recommend publication after clarification on the below issues.

Most of my comments relate to adding extra details/clarification that I think are probably necessary given that this is GMD.

We thank the reviewer for reading the manuscript with much attention and his/her very useful comments and suggestions that helped us to improve the clarity of the paper. The manuscript has been revised following the reviewer suggestions as described below.

My general thoughts are that it is not entirely clear whether including the new VBS treatment of SOA represents an improvement here or not, since the comparison is made to an equivalent simulation without SOA (not the previous version of WRF/Chem). This is something that could be addressed in the Conclusions and Abstract.

The difference between the new chemistry option and old mechanism is the SOA treatment. In the old mechanism, the SOA production is based on SORGAM model that, as explained in the Introduction, produces very little SOA mass concentrations (e.g. Tuccella et al., 2012). Therefore, we may assume that simulations with SORGAM and without SOA (in VBS option) are roughly equivalent. The advantage of this assumption is that the model is forced with the same initial and boundary meteorological conditions of the CTRL simulation. The use of SORGAM would require running the model on all three domains, leading to different results on domain 2. This would introduce dependencies on the input data for domain 3 making the comparison not directly comparable to the CTRL run. These observations have been added to first paragraph of the Section 5.

Being one day of simulation too short to establish if the introduction of SOA really improves the prediction of cloud optical and microphysical properties, we have performed the sensitivity tests also during different days. We have chosen three periods of three days each (with cloudy pixel coverage in MODIS data within the domain 3 larger than 60%) during the scavenging period (17-19 May) and six day during the long range transport period (25-27 and 28-29 May). Moreover, following one suggestion of reviewer 1, we have extended the analysis to cloud top effective radius and to liquid water clouds, being the microphysics scheme aerosol-aware only for liquid clouds. This analysis has shown enhanced skills of the model in reproducing clouds when

SOA are included in the run. As suggested, these results have been included in the Abstract and Conclusions.

Specific Comments:

p794, line 18: Specify here that you are talking about the change in aerosol due to anthropogenic emissions, since pre-industrial *times* >> ***The sentence has been rephrased "The change in global median radiative forcing associated to anthropogenic aerosol particles since pre industrial time is highly uncertain and it is estimated to be -0.9 W/m^2 within a range from -1.9 W/m^2 to -0.1 W/m^2 (Boucher et al., 2013)".***

p795, line 22: the range given by Scott et al. (2014) was +0.01 W/m2 to -0.77 W/m2, when accounting for the potential contribution of biogenics to new particle formation. >> ***The values have been corrected.***

p796, line 22-23: given that this is a new aspect that you are examining I think this requires slightly more detail than one sentence. The reader can look up Ahmadov et al 2012 but it would be good to summarise briefly here what this approach involves. >> ***We agree with the reviewer 2 that a better description of VBS could be useful. Therefore we have included a more detailed description of the VBS approach in Section 2.1.***

p799, line 15-17: at the moment this reads slightly like you have the PM2.5 and PM10 emissions and then add extra EC and OC? Might be worth rephrasing to clarify here, is the PM2.5 / PM10 disaggregated into different components? >> ***EC and OC emissions are from another dataset not included in the standard TNO inventory. This point has been clarified: "EC and primary OC emissions are taken from a specific TNO database that is part of the EUCAARI project (Kulmala et al., 2011). These EC and OC emissions are size resolved, they are separated for particles with diameter less than 1 µm, particles with diameter in the range of 1–2.5 µm and 2.5–10 µm".***

p799, line 25: which biogenic emissions do you include? Monoterpenes? Isoprene? Do the NMVOCs contribute anthropogenic SOA? >> ***SOA are formed by oxidation of anthropogenic (alkanes, alkenes, aromatics and cresols) and biogenic (isoprene, monoterpernes and sesquiterpenes) NMVOCs. This point has been included in the VBS description.***

p799, line 26: where do the dust and sea-salt fluxes come from? >> ***With the term "flux" we mean the sea salt and dust emissions from seawater and soil. The sentence has been rephrased: "Dust and sea salt emissions from soil and seawater are calculated on line in the simulations."***

p809, line 24-26: could you expand on this? Or at least clarify the implication, which I think is that if POM is underestimated, the amount of SOA being formed would also be underestimated (according to the VBS approach) >> ***The point has been clarified as suggested "Indeed, the partition between OCV and SOA used in VBS approach depends on the total OM (Equation 1 of Ahmadov et al., 2012), thus if POM is underpredicted the resulting SOA could be underestimated".***

p810, lines 11-15: in this section you refer to some of the specific flights, but other than the date/time of the flights given in Table S1, we don't have any other information about where these flights went; an additional figure to show the flight tracks would be useful, if possible. >> ***We agree that the lack of a figure showing the aircraft flight tracks was missing in the paper, therefore we have included a plot with aircraft tracks in Figure 1. The new figure has been discussed in Section 3.2.***

p811, lines 18-19: how do you actually calculate CCN from the model? Sorry if I've missed this somewhere, Figure 8 just says calculated at 0.2%. Later on you refer to calculating CCN using the particles in the accumulation mode, and mention the hygroscopicity of aerosol, so some clarification on how this calculation is done in WRF/Chem would be beneficial. >> ***WRF/Chem calculates the number and mass of aerosol activated as cloud droplet, and the spectrum of CCN, i.e. the aerosol activated at some specific supersaturation thresholds. We agree that calculation of CCN is not well described in the text, therefore we have included some clarifications in Section 2.1: "The activation of aerosols is based on a maximum supersaturation determined from a gaussian spectrum of updraft velocities and bulk hygroscopicity of each lognormal mode. Bulk hygroscopicity is based on the volume weighted average of the hygroscopicity of each aerosol component. In addition to the activated aerosols at environmental conditions, the CCN spectrum is also determined, i.e. the aerosol particles acting as CCN at some given maximum supersaturation (0.02, 0.05, 0.1, 0.2, 0.5, and 1%) are computed." Moreover it has been also specified in the text that model/observation comparison is done with CCN at 0.2% of supersaturation.***

p811, lines 22-24: this seems like a point that is relevant to your observed v. simulated CN comparisons? if you are missing everything below 15 nm in the observations surely this would account for some of your over prediction, unless you are calculating your modelled CN above 15 nm? Then you say "Therefore the so calculated CCN efficiency is underestimated." But if you are talking about the observations, I would have thought the CCN efficiency is overestimated if your CN value is lower than it should be? Or do you mean relative to the modelled CCN efficiency? Either way this paragraph requires some clarification. >> ***In the comparison between predicted/observed CN we used the measurement of CPC 3010 instrument that measures the number of particles larger than 15 nm because the model does not take into account the ultrafine particle mode, indeed the modelled CN is given by the sum of particles of Aitken, accumulation and coarse modes. Considering that during the campaign Aitken mode had the mean diameter of 30-60 nm (please refer to Crumeyrolle et al., ACP, 2013), we may say that we are not missing anything in the observations below 15 nm.***

When we say "Therefore the so calculated CCN efficiency is underestimated" we are referring to both observed and modelled CCN, because this should be calculated by using the particles larger than minimum activation diameter. Therefore, the model CCN efficiency is calculated with particles of the accumulation mode (the most favored particles to act as CCN) and it is qualitatively compared to the observed CCN efficiency reported in Crumeyrolle et al., ACP, 2013 that has been computed by using the particles larger than 100 nm. This helps to better characterize the relationship between CCN and the corresponding aerosol populations.

For more clarity, the last paragraph of Section 4.4 has been rephrased: "The calculated observed and modelled CCN efficiencies could be underestimated. In general, the CCN efficiency should be computed with the aerosol population with size larger than the minimum activation diameter (Asmi et al., 2012). The latter depends on the aerosol type and ranges from about 50 to 125 nm. We calculated the observed CCN/CN ratio with the measurements of CPC 3010 which gives the total number of particles larger than 15 nm, and modelled CCN fraction is calculated with total particle number given by the sum of the three modes of the lognormal distribution (Aitken, accumulation and coarse). In order to better characterize the relationship between CCN and corresponding aerosol population in the model, predicted CCN efficiency was also calculated with particles of the accumulation and coarse modes (the most favored particles to act as CCN) and it was qualitatively compared to observed efficiency during the IMPACT campaign computed with particles larger than 100 nm. Observed values of CCN efficiency are in the range of 0.28-0.4 and 0.38-0.6 in the PBL and FT (Crumeyrolle et al., 2013), respectively. The simulated CCN fraction calculated with the particles of the accumulation and coarse modes, is always

underestimated with respect to the observations, and it is in the range of 0.17-0.3 in PBL and 0.23-0.36 in FT. The model deficiency in simulating the CCN/CN ratio could be attributable to the uncertainties in geometrical diameter and bulk hygroscopicity of the lognormal modes, and updraft velocity that lead to error in the prediction of minimum activation diameter of each mode".

p812, line 13-14: this is tricky to attribute to anything because we don't have any spatial information on where the number (and mass which is probably more important for AOT at 550 nm?) is over predicted, whereas the AOT is over predicted predominantly in the east of the domain >> ***Unfortunately MODIS data have a good coverage of AOT measurements only for 14 May. This effectively does not allow us to have a general overview of model skill in predicting AOT. Moreover as it is possible to note looking at Figure 1, the aircraft observations of 14 May are over the area to the South-Est of Cabauw that is not covered by satellite data. Therefore we have removed this sentence and we have highlighted that it is not easy characterize the bias of the model in simulating AOT within this study.***

p812, line 14-17: so what is the implication of this? It's probably worth reiterating what your assumptions are regarding mixing state >> ***We think that this is a point to clarify. We have rephrased the end of the first paragraph of the Section 4.5: "In general, model intercomparisons revealed that a large part of the uncertainties in simulating the AOT arises from the assumption on the mixing state. For example, AOT computed with external mixing is larger by 30-35% of that calculated with internal mixing assumption (Curci et al., 2014b). For typical atmospheric particle sizes and in the visible wavelength range, the AOT is then expected to be lower under internal mixing assumption (that is the assumption done in this work). Moreover, a 10% error in predicting AOT may be attributable to the choice of species density, refractive index, and hygroscopic growth factor (Curci et al., 2014b)".***

p813, line 1-2: I think this sentence requires some rephrasing, shift in what sense? >> ***The sentence has been removed after new analysis (please, see the answer to major comments).***

p813, line 8+: it would be useful here to clarify what you are actually turning off, the biogenic emissions? just in this domain? (+ what are the implications of that) It would also be useful to have some domain wide summary statistics (like Table 3) for the NOSOA simulation. Also it would be good to be consistent with how you refer to these simulations i.e., CTRL (sometimes called

"reference run") and NOSOA (sometimes called "sensitivity run") >> ***The sensitivity tests have been conducted only in domain 3. The model array of the gas and aerosol concentrations has been set to zero for all anthropogenic and biogenic SOA. Therefore SOA concentrations are zero everywhere in the domain, and simulations are not affected by incoming SOA from boundaries. This point has been included in Section 5. Moreover, the domain summary statistics has been included for both liquid and all phase clouds in Tables 4 and 5.***

p814, line 17: I wouldn't use the word larger (this suggests something about the size of the CCN), do you mean higher concentrations? The increase in droplet concentration does not spatially overlap with the increase in CCN concentration, it would be worth some explanation of this, particularly since in the previous section you are suggesting that overestimating CCN would lead to overestimating CDN which would lead to overestimating CWP. >> ***The aim of maps showing $CCN_{0.2}$ and cloud droplet column was to show that CTRL simulations has more CCN and cloud droplets with respect NOSOA run. It is normal that the increase in droplet column does not overlap with the increase in CCN column, because in these maps CCN are calculated at 0.2% of supersaturation while cloud droplet are calculated from aerosol distribution at environmental supersaturation (please, see our comment about calculation of CCN). If we were supposed to see the overlap between CCN and droplet we should have plotted the columns of activated aerosol particles (cloud born aerosols). However, the sentence has been removed after new analysis.***

p814, line 24-26: Unless I've misunderstood, something is the wrong way around here, which simulation has 10% more optically thin clouds? Or is it just a coincidence that the difference between the amount of the total clouds accounted for by thin clouds in each sim is also 10%? **NOSOA** ***run is the simulation with more optically thin clouds. The sentence has been updated including also the results from other days of simulation.***

p815, lines 3-4: where does the smaller effective radius come from? The previous sentence made sense, perhaps remove (or clarify) this one >> ***The sentence has been removed after new analysis.***

p819, line 5: it would be good to keep the description consistent, if you're using VBS to simulate the partitioning of secondary organics into the particle phase, you're not excluding "SOA particles" as such, would be better to say "SOA is excluded" >> ***The sentence has been rephrased: "As test application of the new chemistry option, we performed a sensitivity simulation where SOA mass concentration is set to zero".***

p819, lines 20-21: these are not the same values as you give on p814, so one set needs updating. >> ***The paragraph has been modified after new analysis.***

Technical Comments:

P793, line 4: insert “the” before WRF/Chem? >> ***Done***

p797, line 7: insert “the” before “new chemistry package”? >> ***Done***

p798, line 17: remove “the” before Table 1 >> ***Done***

p801, line 9: replace “counters” with counter? >> ***Done***

p802, line 4-6: refer the reader to Table 2 here. >> ***Done***

p807, line 10-11: refer the reader to Figure 6 here >> ***Done***

p811, line 4: the end of this sentence doesn’t make sense, please rephrase >> ***The sentence has been rephrased “Using several nucleation parameterizations, Pierce and Adams (2009) showed that CCN on average varies by up to 12% within the PBL”.***

p811, line 6-7: again doesn’t quite make sense, please rephrase (perhaps replace “of predicted CN larger” with “between simulated and observed CN being larger”) >> ***We agree that sentence in this form is difficult to understand, it has been rephrased as suggested.***

p811, line 12: “Andrea” should be “Andreae” (it’s correct in the Reference list) >> ***Done***

P813, line 20: replace “ticker” with “thicker” >> ***Done***

p835, Table 2: I think it would be useful to specify in the caption what the abbreviations (WS, WD etc) mean so that the reader doesn’t need to look elsewhere. Also should the 6th entry in the WD section be WD200 (currently reads WS200)? >> ***The abbreviations have been specified in the caption of the Table 2. WS200 mean the observed/modelled wind speed at 200m. The abbreviations of the variables have been changed for example as follow: WS (m/s) at 10m, WS (m/s) at 20m, etc …***

p844 and p845, Figures 7 and 8: although these are the same as Figure 6 it would be useful to include in the caption what the different colours represent so the reader does not have to keep referring back to Figure 6. >> ***Done***

Supp Info: I think the second “Figure S1” should be called Figure S2. >> ***Done***