Response to Anonymous Referee #2

This paper examined the dependence of global SOA production from clouds on six variables including some focused on the meteorology (liquid water content, temperature) and some related to chemistry (parent hydrocarbon reaction, oxidants). The production rate of SOA in clouds was relatively well represented as a function of liquid water content (LWC) and total carbon loss (TCloss). The paper would be stronger with more details on the derivation of the parameterization (statistical thresholds used to include or eliminate variables, analysis of covariance between variables, independent verification, etc). Furthermore, it is unclear if this parameterization would remain valid in another model or different configuration of current model.

We thank the reviewer's thoughtful and helpful comments to clarify and improve this paper. In the revised manuscript, we have addressed these issues in detail. Please see our response to each comment below:

General comments:
1. Units. Values should always be presented with their associated units. For example, Table 2 shows partial regression coefficients, but there are no units. Please indicate unitless if appropriate. In addition, parameters alpha, beta, and gamma are often discussed without units. Given the units of alpha (which seem to involve m□1:8 and s□0:6 if gamma is dimensionless) could alpha be tied to any physical process?

We agree with the reviewer that alpha, beta and gamma should have units. In this case, the units for alpha, beta and gamma should be $(kg \cdot m^{-3} \cdot sec^{-1})^{1-gamma}$, $kg \cdot m^{-3} \cdot sec^{-1}$ and unitless, respectively. We have incorporated these units into our revised manuscript. For example, now table 2 is:

Table 2**.** Multiple linear regression between log(P_{SOAcld}, unit: $kg \cdot m^{-3} \cdot sec^{-1}$) and the logarithm of six factors

	LWC (kg_{water}/Kg_{air})	**TC_{loss}** ($kg \cdot m^{-3} \cdot sec^{-1}$)	**Temperature** (K)	**O_3** (mol/mol_{air})	**OH** (mol/mol_{air})	**VOC/NO_x** (mol/mol)
b*	1.0409 ±0.0002	0.3416 ±0.0003	14.77 ±0.01	-0.207 ±0.002	-0.0106 ±0.0005	0.0397 ±0.0004
B**	0.1704	0.0957	0.0571	-0.0052	-0.0015	0.0085

b***	1.0515 ±0.0002	0.4175 ±0.0002
B***	0.1721	0.1170

Alpha determines the overall SOA production efficiency when cloud water coincides with oxidation of hydrocarbon precursors and thereby controls the total amount of global in-cloud SOA production. It should be tied to the whole multiphase chemistry processes including gas-phase production of water-soluble gases (WSG), transport of WSG and oxidants from gas to cloud droplet surface, all liquid-phase reactions, and the formation of SOA inside the cloud droplet. In the revised manuscript, we have explicitly explained the meaning of each parameter in Section 4. Now we have:

"*Eq. (2) also indicates that the production of* SOA_{cld} *is simultaneously determined by LWC and* TC_{loss}*, which, respectively, represent the contribution from liquid-phase processes and gas-phase processes. The parameter α denotes the annual average intensity of chemical reactions in these processes. It determines the overall production efficiency of* SOA_{cld} *and is tied to whole multiphase processes to form SOA. The parameter β probably represents the influence of some physical processes, such as the transport of* SOA_{cld} *precursors from nearby regions. It may also represent the contribution from acetic acid, particularly over the remote locations. Besides α and β, we find exponents of LWC and* TC_{loss} *are approximately 1 and 0.4, respectively. This indicates that* P_{SOAcld} *responses linearly to the spatiotemporal variability of cloud liquid water, but nonlinearly (or concavely) to the gas-phase oxidation of hydrocarbons. The γ value reflects the complexity and combined effects of multiphase processes from the oxidation of VOCs to* SOA_{cld} *formation, including production of water-soluble species in the gas phase, the gas-liquid transport of oxidation products, and subsequent chemical reactions inside cloud droplets. Since γ is less than 1 (different to Ervens et al., 2008, where* SOA_{cld} *yield depends linearly on initial isoprene concentrations, see Table S2 in the supplementary material for major differences between this study and Ervens et al., 2008), it may also suggest the competition on oxidants between the gas-phase and liquid-phase chemistry to oxidize organic compounds as* TC_{loss} *increases.*"

2. Derivation of parameterization. What is the take away message from table 1? Use of the log based values points to all parameters being relevant. How did you go from Table 2 with coefficients from the multiple linear regression based on logarithms to a function of the form of equation 2 with pre-exponential as well as exponential coefficients? Are the coefficients in equation 2 also in Table 2?

The take away message from Table 1 is that individually each variable could be non-linearly correlated with the P_{SOAcld}, and should be transformed in order to better represent the association. Hence, in Table 2, we use the transformed variables to explore their combined effects on P_{SOAcld}. There are about two steps to derive Eq 2. We firstly determine the linear and nonlinear relationship between P_{SOAcld} and each variable based on a transformed multiple regression model. After a series of statistical tests by including/excluding each variable in light of the standardized regression coefficient in Table 2, LWC and TC_{loss} are chosen and their exponents (i.e., 1 and 0.4) in Eq 2 are determined based on values in Table 2. After transformed back from the pre-exponential coefficients, we conduct an additional linear regression to determine the parameters alpha and beta, which are not listed in Table 2. To make this clear, we add additional notes under Table 2:

"**** Results are based only on the most effective factors (i.e., LWC and TC_{loss}). Now the regression model is $log(P_{SOAcld}) \sim b1 \cdot log(LWC) + b2 \cdot log(TC_{loss})$. Please note, after transformed back from the pre-exponential coefficients, parameters in Eq. (2), namely α and β, are determined by an additional linear regression analysis, which are not shown here.*"

3. Applicability to other models/configurations. Section 2.2 indicates model outputs are archived every 3 hours. I assume these outputs are used both to develop the parameterization (equation 2) as well as compare the parameterization to the base model. Given how nonlinear your parameterization is, would you expect equation 2, with your fitted parameters, to still hold if implemented on a one hour timestep within your model (a common global chemical timestep)? Changes in horizontal dimensions might also cause the parameterization to perform poorly given how well models do or do not represent the spatial extent of clouds.

This is a very good comment. The parameters derived in Eq 2 are indeed based on the 3-hour averaged model output. By using the data with a higher archive frequency, we expect that individual parameters (e.g., alpha and beta) in Eq 2 should be slightly changed. But the structure of Eq 2 (i.e., the linear and nonlinear relationship to LWC and TC_{loss}) should be held as it physically depicts the essential aspects of multiphase chemistry to form SOA_{cld}. In the revised manuscript, we have implemented our parameterization into AM3 which has a 30min time step. We found by using the original parameters, the global production of SOA_{cld} is slightly lower than the 3-hour predicted results (~10%). We believe this is due to that by averaging each variable over a 3-hour period it increases the coincidence between LWC and TC_{loss}. In the revised manuscript, we give two sets of parameters, namely for the 3-hour and 30min time resolution, respectively. The latter is obtained by conducting a variety of sensitivity tests (with different sets of alpha and beta) in AM3. For horizontal dimensions, we do not expect significant changes of our parameterization since all variables are volume averaged. We agree with the reviewer that the parameterized P_{SOAcld} does depend on how well a model represents the spatial extent of clouds. If a model predicts more clouds when switching to a finer resolution, application of our parameterization would definitely have more SOA production (same results would be expected if detailed cloud chemistry is used). To clarify this, we add the following in Section 5.1:

"*Similar to other* SOA_{cld} *parameterizations, implementation of Eq (2) to global models with different time resolution maybe require slight change in parameter α (or β), to account for the difference in coincidence between LWC and* TC_{loss}*. However, the spatiotemporal pattern of* P_{SOAcld} *should be held according to Figure 5.*"

4. Evaluation. The parameterization seems to be evaluated with the data used to create it. The authors could have evaluated it using some sort of independent verification in which a subset of the model outputs were withheld from the parameterization and used only for purposes of evaluation. Also, if equation (2) is meant to be implemented inside a model to parameterize SOA production, a good test would be to put it inside the model and compare that to the base model. Another useful option would be to add observations to determine if places where the parameterization performs poorly may or may not be correlated with poor performance of the process based model.

Very good suggestion! In the revised manuscript, we have implemented Eq 2 into AM3 to testify the performance of our parameterization and compared the parameterized SOA production to the processed-based result. In addition, the evaluations on AM3 clouds have been addressed independently by Donner et al., 2011. Moreover, we substantially updated the text in Sections 4-6 as well as figures 3-5 in the revised manuscript. The new results in general look similar to the old ones despite a few slight changes (please see Figures 3-5 and Figure S2 in the supplementary material).

Specific comments:

1. *Page 26937, line 18-19, what fraction of the spatial and temporal outputs are covered by the threshold criteria?*

In the revised manuscript, we add this information near Page 26937, line 18-19 in the original manuscript.

"*The results are shown in Table 1 (data are paired in space (horizontal and vertical) and time (daily average)) and are restricted to the following conditions: pressure level below 200 hPa, cloud volume fraction >* 10^{-3}*, LWC >* 10^{-12} *kg(water)/kg(air), and* $TC_{loss} > 10^{-22}$ $kgC \cdot m^{-3} \cdot s^{-1}$ *(approximately 61% of the spatial and temporal outputs are covered by the above threshold criteria, which accounts for more than 95% of global total* SOA_{cld} *production in the process-based simulation).* "