

Anonymous Referee #2

The manuscript by Tsimpidi et al. presents the extension of the ORACLE module in the EMAC global model to include the 2-D VBS for the calculation of organic aerosol evolution in the atmosphere. It is an improvement over the previous ORACLE module, developed by the same group, which included the 1-D VBS. The paper is very well written and contains a lot of information that are useful both to model developers that want to reproduce the results but also to the audience interested in the science output of the new model. I especially want to applaud the authors for providing tables with all model results, instead of just scatterplots, which makes the comparison of their work against future studies trivial. I also liked Figure 7 a lot; I found it very informative and interesting. I am listing a number of major and minor points below, but I do expect that after sufficient revision the manuscript will be eventually published in GMD.

We would like to thank the referee for his/her thoughtful review. Please see below our point by point response to reviewer's comments.

Major points

1/ Lines 79-80: While it is great that this capability exists, it has not been detailed in the manuscript, and it is expected for a model development paper. How easy is it to change the mechanism? The technical details discussed later (e.g. mass yield distribution of emissions in C and O:C bins, photochemical aging rate and products) will be adjusted automatically using some kind of interpolation or other internal model assumption, or the user needs to start from the beginning and introduce new assumptions for any change in the complexity? Has any other complexity been tested in the model, and if yes, how does it perform against the one presented? Is the one presented the optimal one after testing several configurations, or it is the first guess?*

ORACLE 2-D has a flexible interface in which the user can choose the resolution (number of bins used in each dimension) of the 2-D VBS space through a namelist file depending on the desired application and scientific goals. The namelist file mainly includes the variables that control the desired chemical resolution of OA (number of volatility/O:C bins, number of size modes, saturation concentrations, molecular weights, etc.) and the desired emission factors for the distribution of POA emissions into LVOC, SVOC and IVOC volatility bins (and the corresponding O:C bins). Changes to the gas-phase chemistry (e.g., photochemical reaction rate constants) need to be made in the MECCA submodel by modifying the ORACLE replacement file, which will automatically update the gas phase chemistry used by EMAC. Then, the interface layer of ORACLE-2D reads the namelist variables and automatically i) defines the new tracers for organic compounds, ii) performs the coupling with the emission modules and with the gas-phase chemistry module (MECCA), and iii) calls the core layer of the ORACLE-2D module. The core layer calculates the bulk equilibrium gas and aerosol concentrations and distributes the change in the bulk aerosol concentration into size modes.

In a recent study (Tsimpidi et al., 2017) we have analyzed the sensitivity of model results (and its performance compared to observations) to several parameters that control primary emissions, photochemical aging, and the scavenging efficiency of organic vapors. This analysis has been taken into account to choose the appropriate model parameters for the present application. The assumed O:C distribution of the freshly emitted LVOCs, SVOCs, IVOCs, and the first generation VOC oxidation products has not been tested within the ORACLE-2D module and relies on the finding of Donahue et al. (2011) and Murphy et al. (2011).

We have update the discussion in section 3.1 to include more information regarding the user's capability to control effectively the complexity of ORACLE-2D through a namelist file.

2/ The fragmentation discussion has a number of issues, which, in my opinion, will eventually lead to overestimation of OA in the model. First of all, the manuscript (line 233-234) states that volatility increases with fragmentation, which is not shown in figure 1. The net average decrease of volatility does not justify the elimination of the higher volatility product by simply adjusting the volatility of the low volatility product to higher values. This might work on field campaigns where over the measuring site the system remains open in terms of mass balance, but in a global model where mass conservation is paramount, a fraction of OA must leave the VBS framework as volatile compounds (e.g. CO, CO₂, HCHO, acetone). If not, the carbon that is not lost, will eventually oxidize and form OA later, contributing to the total OA mass and leading to OA overestimation. This carbon that was not lost will also affect the O:C ratio of the resulted OA, by probably resulting in a net (small) decrease of O:C at areas where fragmentation is important. Overestimating OA because of the additional fragmented carbon taken into account is also important for the partitioning of the semi-volatile OA: when more OA is present in the aerosol phase, partitioning of semi-volatile compounds will move towards the particulate phase, further adding to the OA overestimation.

This is a valid point. The sentence in lines 233-234 states that organic gases in the atmosphere can get oxidized and change their volatility due to functionalization (reducing the volatility) or fragmentation (increasing the volatility). This is not explicitly described by ORACLE-2D. Instead, we use a simple aging scheme to simulate the net effect of both fragmentation and functionalization. We do agree with the reviewer that assuming volatility transformations only into one direction (functionalization) and neglecting fragmentation can lead to increasingly higher OA concentrations in lower volatility bins. This may result in an overestimation of OA at long aging time scales, up to thousands of kilometers downwind of the source regions. However, we do believe that this overestimation can be partially offset by the conservative aging scheme used in ORACLE-2D, compared to the relatively aggressive functionalization scheme proposed by Murphy et al. (2012). Furthermore, the deposition of organic compounds during their long-range transport can also limit the diffusion of organic mass (and its accumulation) to the lower volatility bins. A relevant discussion has been added to the text.

3/ The addition of oxygen to the molecules is a statistical approach that lacks chemical intuition. Every OH oxidation step, especially at the first oxidation stages, should add one O atom. If two or three are to be added, this is most likely because of isomerization reactions, which requires a long-chain VOC to be the compound being oxidized, and there is no reason for the reaction rate to be faster (as described in section 3.4). Alternatively, more than one O atoms can be added if the reactant is an unsaturated VOC, but then the reaction rate can be an order of magnitude faster, at least. I understand that for simplicity certain choices have been made, but since the model is aware of the number of carbon atoms, I believe that this information can be used in that respect.

Our goal is the development of an efficient submodel for the description of OA suitable for mid-term and long-term simulations with global chemistry climate models. Therefore, we need to use a simple unified aging scheme for the description of all OA components in order to limit the computational cost due to the increase in the complexity of the gas phase chemistry used by the parent model. Here we use the reaction of the organic compounds with OH as a proxy for describing their photochemical aging. However, organic compounds can react with other oxidants as well (e.g., HO₂, O₃) and gain more than one oxygen atoms. By taking into account this probability, and combined with the assumed aging reaction rate with OH, the adopted photochemical aging parametrization has an encouraging performance against measurements here and in previous studies (Murphy et al., 2011).

4/ The temperature dependence on partitioning is completely absent in the manuscript. Its treatment though can have a huge impact on results, and should not only be presented, but also discussed in detail. Assuming that an enthalpy of vaporization is used in the model, is that constant across the two dimensions of the VBS framework, or it is changing (and how and why)? A high enthalpy of vaporization will make the volatility bins with low C very non-volatile when temperature drops even slightly, completely shutting down gas-phase oxidation, due to the absence of gas-phase material, affecting all results presented in the manuscript.*

The temperature dependence of saturation concentrations is described by the Clausius–Clapeyron equation:

$$c_i^* = c_{i,0}^* \frac{T_0}{T} \exp \left[\frac{\Delta H}{R} \left(\frac{1}{T_0} - \frac{1}{T} \right) \right],$$

where c_i^* and $c_{i,0}^*$ are the saturation concentrations at temperature T and T_0 , respectively, R is the gas constant, and ΔH is the enthalpy of vaporization. Here, an effective ΔH of 30 kJ mol⁻¹ is used for all aSOA-v and bSOA-v species based on data for α -pinene (Pathak et al., 2007). A ΔH of 112, 100, 88, 76, and 64 kJ mol⁻¹ is used for the 10⁻², 10⁰, 10², 10⁴, and 10⁶ $\mu\text{g m}^{-3}$ volatility bins, respectively, of all organic compounds from fuel combustion and biomass burning sources based on data for large saturated species commonly found in primary emissions (Donahue et al., 2006). This information has been added to the text.

Minor points

1/ Bottom of page 3: It would be useful to put ORACLE v1.0 in perspective there as well.

This paragraph presents previous studies and models used the 2D-VBS framework, having the ability to track the evolution of OA in the 2-D space of volatility and oxygen content. ORACLE v1.0 accounts only for changes in the volatility of OA, during its atmospheric aging, and not in the oxygen content.

2/ Lines 140-143: Although already published, since it is highly relevant here, a sentence or two about the distribution in volatility bins would be useful.

The assumed distribution of the primary organic compounds to LVOCs, SVOCs, and IVOCs from open biomass burning and anthropogenic emissions based on their volatility has been added to the text.

3/ Lines 175-177: How many of those species are advected? Increasing the number of advected tracers by 130 and only getting a 16% slowdown of the model sounds too good to be true. Is this 16% degradation of performance a metric for EMAC (the CTM model) or ECHAM-MESSy (the CCM)? What is the total number of advected tracers, for comparison?

The degradation of performance refers to the EMAC chemistry climate model that consists of the MESSy submodel system and the ECHAM general circulation model. ECHAM simulates the atmospheric flow and is integrated in the base model layer of MESSy, which uses different modules to simulate the atmospheric chemistry and transport. In EMAC, the simulation of atmospheric transport consumes only a small portion of the computational time (~10%), since the equations used for the advection comprise a system of non-coupled differential equations, and the computational time increases approximately linearly with the number of species. The main computational burden originates from the gas-phase chemistry (MECCA submodel, ~ 40%), aerosol dynamics (~20%, which includes inorganic aerosol thermodynamics and phase partitioning of organic aerosols). All these processes include the solution of numerically stiff differential equations, and the computational time required for these processes increases super-linearly with the number of species. However, ORACLE-2D uses an efficient method to minimize the computational cost of the phase partitioning calculations by performing only one equilibrium calculation per volatility bin since the condensation/evaporation of organic compounds depends only on their saturation concentration and not their O:C ratio. Therefore, the degradation of performance comes mainly from the gas phase chemistry and secondarily from the atmospheric transport of the additional tracers (in total: 524 in ORACLE-2D and 292 in ORACLE).

4/ Line 320 and figure 2: Boreal forest emissions maximize during summer. In addition, although the net photochemical activity is probably less in absolute values there than in tropical forests, the

length of the day is much longer, and the net photochemical activity there is not negligible. I would expect to see more OA over boreal forests, which are not present there. On the other hand, there are local maxima over very cold and remote regions with high altitude (Greenland, Antarctica) which are probably unrealistic. Is this a combined effect of long-range transport and temperature dependence?

We would like to clarify that the presented values in line 320 correspond to the O:C ratio and not to the concentration of OA. We agree with the reviewer that despite the extended photochemical activity in boreal forests during summer, the O:C in the region is moderate (~0.3) and lower than the O:C calculated over the tropical forests (~0.5). However, this value corresponds to the annual average O:C of the region, and therefore summertime O:C values are averaged with the very low O:C values during the boreal winter, where the photochemical activity is very limited. In addition, the relatively low temperatures over the boreal forests favor the partitioning of SVOCs into the particulate phase forming POA and reducing the overall OA ratio of O:C. High values of O:C (and non-negligible concentrations of OA) can be found over the Greenland and Antarctica due to the long range transport of highly oxidized organic compounds and their condensation in the particulate phase under very low temperatures.

5/ Line 359: Please soften this statement. The way it is currently written implies that this is the one and only reason of underestimation.

The sentence has been rewritten to clarify that the underestimation of OA was only partially attributed to the missing residential wood burning emissions in our inventory.

6/ Line 413-415: Why OA are less oxidized in the tropics, where photochemical activity (thus aging) is higher?

The multi-generational oxidation of IVOCs can increase the overall oxidation state of OA. OA are less oxidized in the tropics because IVOCs comprise 40% of total open biomass burning emissions compared to 70% of total fuel combustion emissions (that dominate the organic compound emissions in the northern industrialized regions).

7/ Line 449: I find the statement that POA is simulated well overly optimistic. The plots in figure 8 don't resemble a straight line, and roughly a third of the data points are under- or over-estimating measurements by a factor of 3 or more. In addition, the data points are not representative for the global atmosphere, since major OA-rich regions are absent, due to lack of measurements there. I have no doubt that the authors did their best to compare the model with whatever relevant data is available, but the fact that the mean bias is small does not mean that very large differences exist when comparing individual stations.

We would like to clarify the model reproduces the observed campaign average POA concentrations within a factor of two in 40% of the cases over urban downwind and rural locations. Despite the fact that such discrepancies are expected by global models, we do agree with the reviewer that the small mean bias does not necessarily indicate a good model performance over individual measuring stations. The sentence has been revised accordingly.

8/ Section 5.2: This is one of the most innovative parts of the manuscript, but stays in just describing the results. Expanding this discussion in implications would be valuable here. Examples include, but are not limited to, why the model performs better during winter, when the OA performance is the worst? What does that tell us for the missing source of OA? Also, how does the O:C ratio compare against OM/OC values used by other models? How can the improved O:C simulation with ORACLES inform other models without VBS capabilities to improve their simulations by probably using different OM/OC ratios?

We thank the reviewer for the suggestions. The model can reproduce well the O:C ratio of total OA only if it can capture well the relative contributions of POA, fresh SOA, and aged SOA to total OA mass. Therefore, even if the model severely underpredicts all types of OA during wintertime, the calculated OA O:C is similar to observations because the relative contribution of each OA compound is captured by the model. The slightly higher underprediction of aged SOA (compared to fresh SOA and POA) during winter results in a small underprediction of total OA O:C (NMB = -7%). This indicates that the missing OA during winter cannot be attributed only to missing POA sources (e.g., residential biofuel use) but also to missing SOA formation pathways (e.g., multiphase oxidation).

In addition, the total OA O:C ratios presented here can be used to calculate the OM/OC based on Eqs. 10 and 6 of the manuscript. Here, we found that the average calculated OM/OC is 1.8, compared to the observed 1.75. This is also in accordance with the OM/OC value of 1.84 reported by Canagaratna et al. (2015) obtained from a vast dataset of chamber and ambient OA measurements. The calculated OM/OC ranges from 1.6 during winter to 2 during autumn, while the observed O:C ranges from 1.6 during summer to 1.8 during spring and autumn.

We have expanded the discussion in section 5.2 accordingly.

9/ Line 501: I do not understand why caution the reader here that there are only 4 field campaigns during summer months, when there are also only 4 during winter and just 2 during autumn (Table 7).

This is a valid point, we have removed this information from the text.

10/ The summary lacks two important conclusions from this work: one, the fact that the model, same as probably all other global models, underestimates OA during winter, and two, the seasonality of O:C.

Thank you for pointing this out. We have added these two conclusions in the summary.

11/ Code data and availability, line 565: Although it is clear what steps need to be taken if a user wants to use ORACLES in EMAC, there is no mention on whether the ORACLES model is modular enough to be ported in other models, and whether its code is even available for a user (a user is allowed) to do such a thing.

ORACLE needs to be modified in order to get implemented into other CCMs. ORACLE consists of two layers: the submodel interface layer and the submodel core layer. The interface layer collects all relevant information/data from the parent CCM (i.e., emissions, temperature), transfer them via parameter lists to the core layer, call the core layer routines, and distribute the calculated results from the parameter lists back to the parent CCM. The coupling/feedback between different submodels are also managed within this layer. Since the interface layer performs the data exchange for the submodel, all the required modifications for implementing the submodel into other CCMs will have to be made solely in this layer letting the core layer of the model untouched. The complete ORACLE code can be obtained by applying for an EMAC license or upon request by emailing the first author. We have added the relevant information to the “code availability section”.

12/ A zonal mean plot for total OA would be a great supplement in Figure 2 and the discussion.

Following the reviewer’s suggestion, we have added the zonal plot for total OA and the corresponding discussion. In brief, the calculated zonal average concentrations of total OA at surface peak over the mid-latitudes of the Northern Hemisphere and the tropics ($\sim 4 \mu\text{g m}^{-3}$) and remain high (above $2 \mu\text{g m}^{-3}$) up to 750 hPa altitude.

13/ Figure 5b: Why aged SOA maximizes over India?

Aged SOA maximizes over north India since southerly monsoon winds, prevailing during summer, favor the transport of aged organic compounds to the north where they are trapped by the presence of the Himalaya mountain range and accumulate their concentrations. This information has been added in section 4.3.

Technical comments

1/ Please explain what aSOA-v, b-SOAv, SOAsv and SOAiv mean in section 3.4.

Done.

2/ Line 241: I believe the correct reference for the OH reaction rate is Donahue et al. (2006).

Donahue et al. (2006) has been added as a reference

3/ Section 5 introduction: These equations are textbook material, they should be moved in the supplement.

The equations have been removed from the text. We have add a reference to our previous ORACLE evaluation (Tsimpidi et al., 2016) where the same evaluation metrics are used and explained in detail.

4/ The second paragraph of section 5.2 sounds a lot like a repetition of the first one, with the addition of some numbers. Please consider revising the text.

The first paragraph in section 5.2 discusses in brief the performance of the model compared to measurements from urban locations. On the other hand, the second paragraph refers to field campaigns in non-urban areas and to the seasonal evaluation of the model.

5/ Figure 1: What happens with the empty bins, e.g. high volatility and high O:C? Are they being considered in the model, or there is a smart way implemented in the model to ignore them, which saves CPU time?

The empty bins are not considered by the model since the chemical mechanism does not produce any specie in this area (with high volatility and high O:C). The number of volatility bins and the O:C bins in each volatility is also controlled by the user from the namelist file.

6/ Figure 5: It appears some values in the remote oceans have been filtered out. What is the criterion for this? I don't think that there are numerical instabilities in the ratios as presented in the figure for when the concentrations are really low, so I don't understand why filter them out. If e.g. everything is red in Figure 5d over the south Pacific, it is a perfectly valid and expected result.

In figures 5c and 5d, the values have been filtered out when total OA is lower than $0.01 \mu\text{g m}^{-3}$. We have used this filter because the ratio of SOA to total OA in such areas is very high and attracts the attention of the reader, even though this result is of no account due to the very low concentrations of OA in these regions.

7/ Figure 8: the 1:2 and 2:1 lines don't meet the axes at the relevant tick marks. Maybe a plotting problem?

Corrected.

References

- Canagaratna, M. R., Jimenez, J. L., Kroll, J. H., Chen, Q., Kessler, S. H., Massoli, P., Ruiz, L. H., Fortner, E., Williams, L. R., Wilson, K. R., Surratt, J. D., Donahue, N. M., Jayne, J. T., and Worsnop, D. R.: Elemental ratio measurements of organic compounds using aerosol mass spectrometry: characterization, improved calibration, and implications, *Atmospheric Chemistry and Physics*, 15, 253-272, 10.5194/acp-15-253-2015, 2015.
- Donahue, N. M., Robinson, A. L., Stanier, C. O., and Pandis, S. N.: Coupled partitioning, dilution, and chemical aging of semivolatile organics, *Environ. Sci. Technol.*, 40, 2635-2643, 2006.
- Murphy, B. N., Donahue, N. M., Fountoukis, C., and Pandis, S. N.: Simulating the oxygen content of ambient organic aerosol with the 2D volatility basis set, *Atmos. Chem. Phys.*, 11, 7859-7873, 2011.
- Murphy, B. N., Donahue, N. M., Fountoukis, C., Dall'Osto, M., O'Dowd, C., Kiendler-Scharr, A., and Pandis, S. N.: Functionalization and fragmentation during ambient organic aerosol aging: application of the 2-D volatility basis set to field studies, *Atmospheric Chemistry and Physics*, 12, 10797-10816, 10.5194/acp-12-10797-2012, 2012.
- Pathak, R. K., Presto, A. A., Lane, T. E., Stanier, C. O., Donahue, N. M., and Pandis, S. N.: Ozonolysis of alpha-pinene: parameterization of secondary organic aerosol mass fraction, *Atmospheric Chemistry and Physics*, 7, 3811-3821, 2007.
- Tsimpidi, A. P., Karydis, V. A., Pandis, S. N., and Lelieveld, J.: Global combustion sources of organic aerosols: model comparison with 84 AMS factor-analysis data sets, *Atmos. Chem. Phys.*, 16, 8939-8962, 10.5194/acp-16-8939-2016, 2016.
- Tsimpidi, A. P., Karydis, V. A., Pandis, S. N., and Lelieveld, J.: Global-scale combustion sources of organic aerosols: sensitivity to formation and removal mechanisms, *Atmospheric Chemistry and Physics*, 17, 7345-7364, 10.5194/acp-17-7345-2017, 2017.