Anonymous Referee #1

The manuscript describes and evaluates a module for simulating organic aerosol, its volatility, and the oxidation state in a global model framework. The module is one of the most advanced available for global scale atmospheric models and thus the manuscript is well within the scope of Geoscientific Model Development. The manuscript is for the most part well written. However, some parts which are detailed below need clarification and correcting. I can recommend publishing the manuscript after the following issues have been considered:

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

1/ The treatment of POA is unclear to me. Do you assume some volatility distribution for it? According to Page 13 this seems to be the case as it is said that it evaporates due to dilution (Line 340). What O:C ratio or O:C ratio distribution is assumed for POA?

Similar to ORACLE v1.0, POA in ORACLE-2D is treated as semi-volatile and chemically reactive. Specific information regarding the assumed volatility distribution and O:C of freshly emitted POA for the presented model application is given in section 3.2. POA from anthropogenic fuel combustion and open biomass burning are divided based on their volatility into LVOCs (with C^* at 298 K equal to $10^{-2}\,\mu g\,m^{-3}$) SVOCs (with C^* at 298 K equal to 10^0 and $10^2\,\mu g\,m^{-3}$) and IVOCs (with C^* at 298 K equal to 10^4 and $10^6\,\mu g\,m^{-3}$). Their corresponding emissions are estimated using the emission factors of Tsimpidi et al. (2016). Freshly emitted POA from anthropogenic and open biomass burning sources are assigned an initial O:C of 0.1 (Figure 1a) and 0.2 (Figure 1b), respectively (Donahue et al., 2011).

2/ What is assumed for the size distribution of OA which partitions from gas to particles? Does it follow the modal approach presented in Tsimpidi et al., 2014?

Yes, the modal approach is used. The change in aerosol mass of each size mode after the phase partitioning is determined by using a weighting factor as described in Tsimpidi et al. (2014). This information has been added in section 2.2.1

3/ Page 9: All the abbreviations of organic aerosol species are not explained in this paper, i.e. aSOA-v, bSOA-v, SOA-sv, SOA-iv

Thank you for pointing this out. The same naming convention as in Tsimpidi et al.(2014) has been used. Nevertheless, for clarity, we have added an explanation of the abbreviations in the revised text.

4/ Page 10, Section 3.5: After reading this section it is still not obvious to me how gas-particle partitioning is actually calculated. Does $C_{a,i}$ in Equations (11) and (12) correspond to

$SOA_{i,j}(t + \Delta t)$ in Equation (14)? If so, is it also correct that Equation (11) is first solved for one

volatility bin and then the total concentration in that bin is divided among the O:C bins using Equations (13) and (14) based on the O:C distribution on the previous time step? Do you use mole fraction from the previous time step for all the volatility bins? This procedure should be explained in more detail.

Equation 11 and 12 is a set of n nonlinear equations, where n is the number of volatility bins. Once this set of equations is solved, $C_{a,i}$ and $(C_{t,i} - C_{a,i})$ for each volatility bin i correspond to $\sum_{j=1}^{m} SOG_{i,j}(t+\Delta t)$ and $\sum_{j=1}^{m} SOA_{i,j}(t+\Delta t)$ in eqs. 13 and 14, where m is the number of O:C bins, and are used to calculate the $SOA_{i,j}(t+\Delta t)$ and $SOG_{i,j}(t+\Delta t)$ based on the O:C distribution of the previous time step. This has been made clearer in the revised manuscript.

5/ In Equation (11) it is assumed that the saturation ratio at the particle surface is dependent on the mole fraction of compounds in the organic phase. This is a common practice when solving SOA partitioning, but how good is this assumption in the atmosphere? In many cases, especially at the top of the boundary layer, humidity can be so high that the majority of moles in the aerosol is water decreasing significantly the mole fraction of individual SOA species compared to when only organic phase is taken into account. Thus SOA formation is underestimated. Can you comment on this?

We thank the reviewer for pointing this out. Indeed, water uptake by the organic-phase under high relative humidity conditions can increase the aerosol mass and decrease the mole fraction of individual SOA species which can encourage additional condensation of SOA (Jathar et al., 2016). However, this is not taken into account by ORACLE and can result in an underestimation of SOA in areas with high humidity and significant SOA concentrations. This discussion has been added to the text.

6/ The model is evaluated with observations only at the ground level. However, accounting for the volatility has implications on the vertical profiles of OA and as there are aircraft campaigns where OA has been measured with aerosol mass spectrometers, it would be interesting how well the model compares with the observations in the vertical.

This is a very interesting suggestion. We believe that such a comparison can provide significant insights on the vertical profile of OA and its properties (e.g., hygrsocopicity) due to the chemical aging and changes in its volatility and oxidation state. Therefore, we believe that such an analysis will be more suitable in a separate manuscript since the main scope of the current work is to present the new submodel, its capabilities and its technical implementation.

7/ Can you give some estimate on the increase in the computation time when moving from the 1-D ORACLE to 2-D ORACLE?

Similar to ORACLE, 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 chemical resolution employed here includes 164 organic aerosol surrogate compounds, compared to 34 OA compounds in the original ORACLE paper (Tsimpidi et al., 2014), which result in a 16% increase of the EMAC computational burden. This is discussed in section 3.1

Technical comments and typos:

1/ Page 2, Line 36: space missing before the word "Organic"

Corrected.

2/ Page 3, Line 44: pas-aerosol → gas-aerosol

Corrected.

3/ Page 10, Line 270: equation → equations

Corrected.

4/ Figure 1a: Compustion → Combustion

Corrected.

5/ Figure 8 caption: To be consistent with the rest of the manuscript, please change SV-OOA → SVOOA and LV-OOA → LVOOA

Done.

References

Donahue, N. M., Epstein, S. A., Pandis, S. N., and Robinson, A. L.: A two-dimensional volatility basis set: 1. organic-aerosol mixing thermodynamics, Atmos. Chem. and Phys., 11, 3303-3318, 2011.

Jathar, S. H., Mahmud, A., Barsanti, K. C., Asher, W. E., Pankow, J. F., and Kleeman, M. J.: Water uptake by organic aerosol and its influence on gas/particle partitioning of secondary organic aerosol in the United States, Atmospheric Environment, 129, 142-154, 10.1016/j.atmosenv.2016.01.001, 2016.

Tsimpidi, A. P., Karydis, V. A., Pozzer, A., Pandis, S. N., and Lelieveld, J.: ORACLE (v1.0): module to simulate the organic aerosol composition and evolution in the atmosphere, Geoscientific Model Development, 7, 3153-3172, 10.5194/gmd-7-3153-2014, 2014.

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.

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., HO_2, O_3) 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}, 10^0, 10^2, 10^4, and 10^6 µg 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 is 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 (~ 4 μg m^{-3}) and remain high (above 2 μg 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 concetrations 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 μg 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.

1 **ORACLE 2-D (v2.0): An efficient module to compute the volatility and oxygen content of**
2 **organic aerosol with a global chemistry – climate model**

4 **Alexandra P. Tsimpidi[1], Vlassis A. Karydis[1], Andrea Pozzer[1], Spyros N. Pandis[2, 3] and Jos**
5 **Lelieveld[1, 4]**
6
7 [1] Max Planck Institute for Chemistry, Mainz, Germany
8 [2] Department of Chemical Engineering, University of Patras, Patras, Greece
9 [3] Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA, USA
10 [4] Energy, Environment and Water Research Center, Cyprus Institute, Nicosia, Cyprus
11 *Corresponding author e-mail: ~~a.tsimpidi@mpic.de~~a.tsimpidi@mpic.de

Abstract

A new module, ORACLE 2-D, simulating the organic aerosol formation and evolution in the atmosphere has been developed and evaluated. The module calculates the concentrations of surrogate organic species in two-dimensional space defined by volatility and oxygen-to-carbon ratio. It is implemented into the EMAC global chemistry – climate model, and a comprehensive evaluation of its performance is conducted using an Aerosol Mass Spectrometer (AMS) factor analysis dataset derived from almost all major field campaigns that took place globally during the period 2001-2010. ORACLE 2-D uses a simple photochemical aging scheme that simulates efficiently the net effects of fragmentation and functionalization of the organic compounds. The module predicts not only the mass concentration of organic aerosol (OA) components, but also their oxidation state (in terms of O:C), which allows their classification into primary OA (POA, chemically unprocessed), fresh secondary OA (SOA, low oxygen content) and aged SOA (highly oxygenated). The explicit simulation of chemical OA conversion from freshly emitted compounds to a highly oxygenated state during photochemical aging enables the tracking of hygroscopicity changes of OA that result from these reactions. ORACLE 2-D can thus compute the ability of OA particles to act as cloud condensation nuclei, and serves a tool to quantify the climatic impact of OA.

1. Introduction

Atmospheric aerosols adversely affect human health and play a significant role in climate change on regional and global scales. Depending on their composition, aerosols affect the energy budget of the Earth's atmosphere by scattering and absorbing solar radiation (direct effect) and by influencing the reflective properties of clouds, their lifetime, and precipitation formation (indirect effects). In addition, climate change can play vital and complex role in the formation and removal of atmospheric particles ~~(Trail et al., 2013;Trail et al., 2014)~~ .(Trail et al., 2013;Trail et al., 2014). Organic aerosol (OA) is an important constituent of atmospheric particles contributing 20-90% to the total submicron particulate mass, depending on the region ~~(Zhang et al., 2007)~~.

Primary OA (POA) has been traditionally treated as non-volatile and inert in global scale chemistry climate models (CCMs). Robinson et al. ~~(2007)~~ demonstrated that OA emissions are semi-volatile and most of the emitted OA moves to the gas phase after emission due to dilution and evaporation. On the other hand, all organic vapors are subject to photochemical reactions with

43 OH in the gas phase forming organic products with lower volatility that can recondense to the

44 particulate phase as secondary organic aerosol (SOA). To describe the OA ~~pas~~gas-aerosol

45 partitioning, Donahue et al. ~~(2006)~~ developed the volatility basis-set (VBS) framework, where OA

46 is assumed to be semi-volatile and photochemically reactive, and is distributed in logarithmically

47 spaced volatility bins. With this innovative approach, the semi-volatile primary emissions, the

48 chemical aging, and the SOA formation were unified within a common framework that is ideally

49 suited for regional and global chemical modeling. Since 2006, many regional ~~(Lane et al.,~~

50 ~~2008;Murphy and Pandis, 2009;Tsimpidi et al., 2010;Tsimpidi et al., 2011;Ahmadov et al.,~~

51 ~~2012;Athanasopoulou et al., 2013;Koo et al., 2014;Fountoukis et al., 2014;Ciarelli et al., 2017;Gao~~

52 ~~et al., 2017) and global~~ (Pye and Seinfeld, 2010;Jathar et al., 2011;Jo et al., 2013;Tsimpidi et al.,

53 2014;Hodzic et al., 2016;Lane et al., 2008;Murphy and Pandis, 2009;Tsimpidi et al.,

54 2010;Tsimpidi et al., 2011;Ahmadov et al., 2012;Athanasopoulou et al., 2013;Koo et al.,

55 2014;Fountoukis et al., 2014;Ciarelli et al., 2017;Gao et al., 2017) and global (Pye and Seinfeld,

56 2010;Jathar et al., 2011;Jo et al., 2013;Tsimpidi et al., 2014;Hodzic et al., 2016) modeling studies

57 have used the VBS to account for the semi-volatile nature and chemical aging of organic

58 compounds, demonstrating improvements in reproducing the OA budget and its chemical

59 resolution.

60 The chemical aging of OA results in significant changes of its physical and chemical properties

61 due to the addition of oxygen atoms from the reaction with OH. This increase in OA oxygen

62 content is important for its impact on climate through changes in cloud condensation nuclei (CCN)

63 and ice nuclei (IN) activity. In fact, the oxygen content, expressed by the ratio of oxygen to carbon

64 atoms (O:C), influences the OA hygroscopic growth ~~(Chang et al., 2010;Lambe et al., 2011)~~ which

65 affects CCN activity. In addition, the phase-state changes of SOA during its atmospheric lifetime,

66 which can impact the IN activity, is also influenced by the O:C of the OA ~~(Shiraiwa et al., 2017)~~.

67 Donahue et al. ~~(2011)~~ extended the original one-dimensional VBS framework to two dimensions

68 (2-D VBS) tracking not only the saturation concentration but also the oxygen content of OA during

69 atmospheric transport. This approach further improved the description of the atmospheric

70 evolution of OA and its precursor gases that become increasingly more oxidized, less volatile, and

71 more hygroscopic during their atmospheric aging. However, the large number of additional

72 surrogate organic compounds required by the 2-D VBS framework has hindered implementation

73 in three-dimensional atmospheric models ~~(Napier et al., 2014).~~(Napier et al., 2014). Therefore, the

74 2-D VBS approach has been mostly adopted in box and ~~1D~~1-D Lagrangian models ~~(Murphy et al.,~~
75 ~~2011;Murphy et al., 2012;Chacon-Madrid et al., 2013;Zhao et al., 2015;Paciga et al., 2016)~~. Koo
76 et al. ~~(2014)~~ introduced a hybrid VBS approach for use in three dimensional chemical transport
77 models (CTMs) which combines the simplicity of the VBS with the ability to track the evolution
78 of OA in the 2-D space of volatility and oxygen content.

79   In this work, the computationally efficient module for the description of OA composition and
80 evolution in the atmosphere ~~(ORACLE; Tsimpidi et al., 2014)~~ has been extended to allow for the
81 first time in a global CCM the description of both the volatility and the oxygen content of OA
82 based on the 2-D VBS approach. Similar to ORACLE v1.0, the interface of the new version allows
83 the user to have full control of the complexity of the OA scheme by adjusting the number of species
84 and reactions (i.e., number of compounds, volatility bins, O:C bins) to optimize the computational
85 cost according to the application and the desired chemical resolution. The updated ORACLE
86 module can provide valuable information about the physicochemical evolution of OA during its
87 atmospheric lifetime in support of modeling studies and help quantify the climatic impact of OA.
88

89 **2. Model description**
90 **2.1 EMAC model**
91   The ECHAM/MESSy Atmospheric Chemistry (EMAC) model is a numerical chemistry
92 and climate simulation system that includes sub-models describing lower and middle atmospheric
93 processes and their interaction with oceans, land, and human influences. EMAC consists of the
94 Modular Earth Submodel System and an advanced version of the 5th generation of the European
95 Centre Hamburg (ECHAM) general circulation model. ECHAM5 ~~(Roeckner et al., 2006) serves~~
96 ~~as the atmospheric dynamic core that simulates the atmospheric flow and is integrated in the base~~
97 ~~model layer of MESSy. The interface structure of MESSy links the base model with several~~
98 ~~atmospheric submodels that simulate online the gas phase chemistry (MECCA; Sander et al.,~~
99 ~~2011), inorganic aerosol microphysics and dynamics (GMXe; Pringle et al., 2010), organic aerosol~~
100 ~~formation and growth (ORACLE; Tsimpidi et al., 2014), emissions (ONLEM and OFFLEM;~~
101 ~~Kerkweg et al., 2006b), dry deposition and sedimentation (DRYDEP and SEDI; Kerkweg et al.,~~
102 ~~2006a), cloud scavenging (SCAV; Tost et al., 2006), cloud microphysics (CLOUD; Jöckel et al.,~~
103 ~~2006), and aerosol optical properties (AEROPT; Lauer et al., 2007). EMAC has been extensively~~
104 ~~described and evaluated against ground-based and satellite observations~~ (Pozzer et al.,

105   2012;Tsimpidi et al., 2014;Tsimpidi et al., 2016, 2017;Karydis et al., 2016;Karydis et al.,

106   2017;Roeckner et al., 2006) serves as the atmospheric dynamic core that simulates the atmospheric

107   flow and is integrated in the base model layer of MESSy. The interface structure of MESSy links

108   the base model with several atmospheric submodels that simulate online the gas-phase chemistry

109   (MECCA; Sander et al., 2011), inorganic aerosol microphysics and dynamics (GMXe; Pringle et

110   al., 2010), organic aerosol formation and growth (ORACLE; Tsimpidi et al., 2014), emissions

111   (ONLEM and OFFLEM; Kerkweg et al., 2006b), dry deposition and sedimentation (DRYDEP and

112   SEDI; Kerkweg et al., 2006a), cloud scavenging (SCAV; Tost et al., 2006), cloud microphysics

113   (CLOUD; Jöckel et al., 2006;Bacer et al., 2018), and aerosol optical properties (AEROPT; Lauer

114   et al., 2007). EMAC has been extensively described and evaluated against ground-based and

115   satellite observations (Pozzer et al., 2012;Tsimpidi et al., 2014;Tsimpidi et al., 2016, 2017;Karydis

116   et al., 2016;Karydis et al., 2017). In this study, the applied spectral resolution of the EMAC model

117   is T63L31, corresponding to a horizontal grid resolution of 1.875°x1.875° and 31 vertical layers

118   extending to 18 km altitude (10hPa). EMAC is applied for 11 years covering the period 2000-2010

119   and the first year is used as spin-up.

120

121   **2.2 ORACLE module**

122   **2.2.1 Module description**

123   ORACLE is a computationally efficient module for the description of organic aerosol

124   composition and evolution in the atmosphere (Tsimpidi et al., 2014) that has been incorporated into

125   the EMAC model. The original version of ORACLE (V1.0) simulated the volatility distribution of

126   a wide variety of semi-volatile organic surrogate compounds using bins of logarithmically spaced

127   effective saturation concentrations. Organic emissions from multiple anthropogenic and natural

128   sources are taken into account using distinct surrogate species for each source category. These

129   surrogates can be subdivided into groups of organic compounds based on their volatility: Low

130   volatility organic compounds (LVOCs, $C* \leq 10^{-1}$ μg m^{-3}), semi-volatile organic compounds,

131   (SVOCs, $10^{-1} < C* \leq 10^2$ μg m^{-3}), intermediate volatility organic compounds (IVOCs, $10^2 < C* \leq$

132   10^6 μg m^{-3}) and volatile organic compounds (VOCs, $C* > 10^6$ μg m^{-3}). These organic compounds

133   are allowed to partition between the gas and aerosol phases resulting in the formation of OA. The

134   change in aerosol mass of each size mode after the phase partitioning is determined by using a

135   weighting factor as described in Tsimpidi et al. (2014). Gas-phase photochemical reactions that

modify the volatility of the organics are taken into account and the oxidation products of each group of precursors are simulated separately in the module to keep track of their origin. The model results for the different organic components in the particulate phase were compared with factor analysis results derived from a comprehensive dataset of aerosol mass spectrometer (AMS) measurements from multiple field campaigns across the Northern Hemisphere. The resulting good agreement between campaign average concentrations and model predictions supports the capability of the model to capture the spatial and temporal characteristics of OA levels. Tsimpidi et al. (2017) conducted an extensive sensitivity analysis of the EMAC OA predictions to uncertain parameters in the ORACLE module. One of the major conclusions of their analysis was that the model performance can be improved by assuming that freshly emitted organic compounds are relatively hydrophobic and become increasingly hygroscopic due to oxidation. As a first step to achieve this goal, the ORACLE module has been further developed here to account for the oxidation state of the organic surrogate compounds. The new version is called ORACLE 2-D.

2.2.2. Emission inventory of OA precursors

The emissions of biogenic VOCs (i.e., isoprene and monoterpenes) are calculated online by EMAC with the ONLEM submodel. ~~The open biomass burning emissions of LVOCs, SVOCs, and IVOCs from savanna and forest fires are based on the Global Fire Emissions Database (GFED version 3.1). The emissions of anthropogenic LVOCs, SVOCs, IVOCs, and VOCs (i.e., aromatics, alkanes, olefins) from fossil and biofuel combustion are derived from the CMIP5 emission inventory for the RCP4.5 scenario. More details about the organic compound emissions used here can be found in Tsimpidi et al. (2014~~The open biomass burning emissions of organic compounds from savanna and forest fires are based on the Global Fire Emissions Database (GFED version 3.1; van der Werf et al., 2010) and are distributed into LVOCs (20%), SVOCs (40%), and IVOCs (40%) based on the emission factors proposed by May et al. (2013). The emissions of anthropogenic LVOCs, SVOCs, IVOCs, and VOCs (i.e., aromatics, alkanes, olefins) from fossil and biofuel combustion are derived from the CMIP5 emission inventory for the RCP4.5 scenario (Thomson et al., 2011). The volatility distribution of anthropogenic emissions to LVOCs (8%), SVOCs (72%), and IVOCs (170%) is based on the findings of (Robinson et al., 2007) and includes increased factors (the sum of the emission factors is 250%) to account for the missing IVOC

166 emissions from the traditional inventories. More details about the organic compound emissions
167 used here can be found in Tsimpidi et al. (2016).
168

169 **3 ORACLE 2-D description**
170 **3.1 Module overview**
171   The original ORACLE v1.0 (called hereafter ORACLE) uses saturation concentration bins
172 to describe the volatility distribution of the major OA components. In this work, ORACLE is
173 extended to also resolve the oxygen content of OA expressed by the O:C ratio. (ORACLE v2.0;
174 called hereafter ORACLE 2-D). The volatility dimension is discretized in up to 10 logarithmically-
175 spaced volatility bins separating the organic compounds in low volatility (LVOCs, expressed by
176 the volatility bins of 10^{-3}, 10^{-2}, and 10^{-1} μg m^{-3}), semi-volatile (SVOCs, $C*$ equal to 10^0, 10^1, and
177 10^2 μg m^{-3}), and those of intermediate volatility (IVOCs, $C*$ equal to 10^3, 10^4, 10^5 and 10^6 μg m^{-3}
178 $^{-3}$). Extremely low volatility organic compounds (ELVOCs, $C^* \leq 10^{-4}$) can be formed by the
179 ozonolysis of monoterpenes and sesquiterpenes (Liggio et al., 2010;Sun et al., 2009) playing an
180 important role for the formation and growth of new particles created in situ in the atmosphere by
181 nucleation (Ehn et al., 2014). The production of ELVOCs from biogenic VOCs is not currently
182 included in ORACLE since the simulation of new particle formation is outside the scope of the
183 current article, and part of work in progress. In addition, the oxygen content dimension is
184 discretized in up to 20 linearly spaced O:C bins subdividing the organic compounds into fresh
185 emissions (expressed by the O:C bins of 0.1 and 0.2), less oxygenated organic compounds (O:C
186 equal to 0.3, 0.4, 0.5, and 0.6), moderately oxygenated organic compounds (O:C equal to 0.7, 0.8,
187 0.9, and 1.0), and highly oxygenated species with O:C>1. The first bin includes also the
188 hydrocarbons with zero oxygen. The O:C range can be extended up to 2 (for CO_2). However, there
189 are only a few atmospheric organic compounds with O:C higher than unity and the observed O:C
190 of ambient OA rarely exceeds 1.1 (Ng et al., 2010; Kroll et al., 2011). The ability of ORACLE
191 v2.0 (called hereafter ORACLE 2-D) to simulate degree of oxidation of OA allows the simulation
192 of its hygroscopicity by using proposed parameterizations that link the hygroscopicity parameter
193 kappa with the O:C of OA (Chang et al., 2010;Lambe et al., 2011;Kuwata and Lee, 2017).(Chang
194 et al., 2010;Lambe et al., 2011;Kuwata and Lee, 2017). In the current application, the

195 hygroscopicity of each OA compound is represented by a linear function of the form $\kappa_{org}=0.18$

196 (O:C) + 0.03 ~~(Lambe et al., 2011).~~ (Lambe et al., 2011).

197 ORACLE 2-D has a flexible interface in which the user can choose the resolution (number

198 of bins used in each dimension) of the 2-D VBS space through a namelist file depending on the

199 desired application and scientific goals. ~~In the rest of this work, we employ ORACLE 2-D~~The

200 namelist file includes the variables that control the desired chemical resolution of OA (number of

201 volatility/O:C bins, number of size modes, saturation concentrations, molecular weights, etc.) and

202 the desired emission factors for the distribution of POA emissions into LVOC, SVOC and IVOC

203 volatility bins (and the corresponding O:C bins). Changes to the gas-phase chemistry (e.g.,

204 photochemical reaction rate constants) need to be made in the MECCA submodel by modifying

205 the ORACLE replacement file, which will automatically update the gas phase chemistry used by

206 EMAC. Then, the interface layer of ORACLE 2-D reads the namelist variables and automatically

207 i) defines the new tracers for organic compounds, ii) performs the coupling with the emission

208 modules and with the gas-phase chemistry module (MECCA), and iii) calls the core layer of the

209 ORACLE 2-D module. The core layer calculates the bulk equilibrium gas and aerosol

210 concentrations and distributes the change in the bulk aerosol concentration into size modes.

211 In this work, we employ ORACLE 2-D based on the sensitivity analysis results of Tsimpidi

212 et al. (2017) and using the 2-D VBS at a resolution suitable for medium-term simulations with

213 global chemistry-climate models. This chemical resolution includes ~~164~~150 organic aerosol

214 surrogate compounds, compared to 34 OA compounds in ORACLE, which result in a 16%

215 increase of the ~~EMAC computational burden.~~ overall EMAC computational burden. The

216 performance reduction comes mainly from the gas phase chemistry and secondarily from the

217 atmospheric transport of the additional tracers, since ORACLE 2-D uses an efficient method to

218 minimize the computational cost of the phase partitioning calculations as described in section 3.4.

219

220 ~~3.3 Theory: constructing~~3.2 Constructing the two dimensional grid

221 For the current application, ORACLE 2-D distributes the OA surrogate species into

222 logarithmically spaced volatility bins with C^* varying from 10^{-2} to 10^6 µg m^{-3} and linearly spaced

223 oxygen content bins with O:C varying from 0.1 to 1.2 (Figure 1). Each of the OA surrogate species

224 in this C^*-O:C 2-D space is characterized by a representative number of carbon atoms ~~(n_C~~ per

molecule (n_C) and a molecular weight (*MW*). Donahue et al. ~~(2011)~~ used structure activity

relationships ~~(Pankow and Asher, 2008)~~ to express the C^* as a function of n_c and the number of

oxygens per molecule (n_o):

$$\log_{10} C^* = 0.475(25 - n_C) - 2.3n_O + 0.6\frac{n_C n_O}{n_C + n_O} \quad (1)$$

Field Code Changed

$$\log_{10} C^* = 0.475(25 - n_C) - 2.3n_O + 0.6\frac{n_C n_O}{n_C + n_O} \quad (1)$$

Given that n_o is a function of O:C and n_c:

$$n_O = n_C(\text{O:C}) \quad (2)$$

$$n_O = n_C(\text{O:C}) \quad (2)$$

Field Code Changed

n_c can be expressed as a function of C^* and O:C:

$$n_C = \frac{11.875 - \log_{10} C^*}{0.475 + 2.3(\text{O:C}) - 0.6\dfrac{(\text{O:C})}{1 + (\text{O:C})}} \quad (3)$$

Field Code Changed

$$n_C = \frac{11.875 - \log_{10} C^*}{0.475 + 2.3(\text{O:C}) - 0.6\dfrac{(\text{O:C})}{1 + (\text{O:C})}}$$

Assuming that the organic compounds consist entirely of carbon, oxygen, and hydrogen atoms

(i.e., ignoring nitrogen and sulfur), the *MW* of each surrogate species is a function of n_C, n_O, and

Field Code Changed
Field Code Changed

the number of ~~carbon (n_C), oxygen (n_O), and~~ hydrogen (n_H) atoms per molecule (n_H) and

Field Code Changed

therefore it can be calculated as:

$$MW = n_H + 16n_O + 12n_C \quad (4)$$

247

$$MW = n_H + 16n_O + 12n_C \quad (4)$$

Field Code Changed

where,

~~$n_H = n_C(\text{H:C}) \quad (5)$~~

$$n_H = n_C(\text{H:C}) \quad (5)$$

Field Code Changed

H:C is the atomic ratio of hydrogen to carbon approximated by Heald et al. (2010) as:

~~$\text{H:C} = 2 - (\text{O:C}) \quad (6)$~~

$$\text{H:C} = 2 - (\text{O:C}) \quad (6)$$

Field Code Changed

Combining Equations 4, 2, and 5, we get

~~$MW = (15(\text{O:C}) + 14)n_C \quad (7)$~~

$$MW = (15(\text{O:C}) + 14)n_C \quad (7)$$

Field Code Changed

Given that n_c decreases as C^* increases (Eq. 3), the MW calculated by Eq. 7 is consistent with the molecular corridor approach ~~(Shiraiwa et al., 2014)~~ which suggests a tight inverse correlation between volatility and molar mass constrained by boundary lines of low and high O:C ratios.

Organic compound emissions from anthropogenic fuel combustion and open biomass burning include LVOCs (with C^* at 298 K equal to 10^{-2} µg m^{-3}) SVOCs (with C^* at 298 K equal to 10^0 and 10^2 µg m^{-3}) and IVOCs (with C^* at 298 K equal to 10^4 and 10^6 µg m^{-3}). Their corresponding emissions are estimated using the emission factors of Tsimpidi et al. ~~(2016).~~(2016). Freshly emitted LVOCs, SVOCs, and IVOCs from anthropogenic and open biomass burning sources are assigned an initial O:C of 0.1 (Figure 1a) and 0.2 (Figure 1b), respectively (Donahue

275 et al., 2011). We distinguish anthropogenic and biogenic VOCs, and their first generation oxidation

276 products are distributed in four volatility bins (with C^* at 298 K equal to 10^0, 10^1, 10^2, and 10^3 μg

277 m^{-3}) by using the aerosol mass yields by Tsimpidi et al. (2014). The O:C distributions of the first

278 generation VOC oxidation products are given by Murphy et al. (2011) and vary with volatility

279 (Figures 1c and 1d).

280

3.43 New photochemical aging scheme

282   Similar to ORACLE, multiple generations of homogeneous gas-phase reactions with OH

283 are simulated for all OA compounds treated by ORACLE 2-D. After each oxidation step, oxygen

284 atoms are added to the reacting organic gas resulting in an increase of oxygen content and a change

285 of volatility due to functionalization (reducing the volatility) or fragmentation (increasing the

286 volatility). To minimize the computational cost and at the same time simulate the net effect of both

287 fragmentation and functionalization, ORACLE 2-D uses an approach similar to the aging scheme

288 proposed by Murphy et al. (2011). ORACLE 2-D assumes a net average decrease in volatility of

289 aSOA-v (SOA from the oxidation of anthropogenic VOCs) reacting with OH by a factor of 10 and

290 the addition of one or two oxygen atoms with the same probability of 50% (Figure 1c). The same

291 addition of oxygen atoms is assumed for bSOA-v.Assuming volatility transformations only into

292 one direction (functionalization) and neglecting fragmentation can lead to increasingly higher OA

293 concentrations in lower volatility bins. This may result in an overestimation of OA at long aging

294 time scales up to thousands of kilometers downwind of the source regions. However, this

295 overestimation can be partially offset by the conservative aging scheme used in ORACLE 2-D,

296 compared to the relatively aggressive functionalization scheme proposed by (Murphy et al., 2012).

297 The same addition of oxygen atoms (one or two) is assumed for bSOA-v (SOA from the oxidation

298 of biogenic VOCs), however, its volatility remains unchanged due to a balancing of fragmentation

299 and functionalization effects (Murphy et al., 2012) (Figure 1d). The OH reaction rate constant for

300 both aSOA-v and bSOA-v is 1×10^{-11} cm^3 molecule^{-1} s^{-1} (Tsimpidi et al., 2010). The reaction of

301 SOA-sv and SOA-iv(Donahue et al., 2006;Tsimpidi et al., 2010). The reaction of SOA-sv (SOA

302 from the oxidation of SVOCs) and SOA-iv (SOA from the oxidation of IVOCs) with OH with a

303 rate constant of 2×10^{-11} cm^3 molecule^{-1} s^{-1} (Tsimpidi et al., 2014) results in the addition of two or

304 three oxygen atoms (with equal probability) and the reduction of their volatilities by a factor of

305 100 (Figure 1a, 1b). The number of added oxygen atoms (n_{O+}) due to the reaction with OH is

306 then expressed as an increase of O:C in the 2-D space. Each of the OA surrogate compounds in

307 the 2-D space described in Section 3.3 has a representative number of carbon per molecule (n_C).

308 n_C). Assuming that carbon is conserved during the reaction with OH, the O:C of the product is

309 calculated as follows:

310

311
$$(O:C)_{product} = (O:C)_{reactant} + \frac{n_{O+}}{n_C} \qquad (8)$$

312

313 If the $(O:C)_{product}$ has more than one decimal places, then is distributed between the two

314 adjacent O:C bins of the 2-D space by using linear interpolation. Finally, since carbon is conserved,

315 the increase in organic mass (OM) due to the added oxygen after each oxidation reaction is

316 calculated as:

317

318
$$\frac{OM_{product}}{OM_{reactant}} = \frac{\left(\dfrac{OM}{OC}\right)_{product}}{\left(\dfrac{OM}{OC}\right)_{reactant}} \qquad (9)$$

319

320 where, following Murphy et al. (2011),

321

322
$$\left(\frac{OM}{OC}\right) = 1 + \frac{16}{12}(O:C) + \frac{1}{12}(H:C) \qquad (10)$$

323

324 **3.4 Phase partitioning calculations**

325 ORACLE 2-D uses the core layer of the ORACLE module to calculate the partitioning of

326 organic compounds between the gas and particle phases by assuming bulk equilibrium (Tsimpidi

327 et al., 2014). However, the computational time required for the phase partitioning increases super-

328 linearly with the number of species. As the condensation/evaporation of organic compounds

depends only on their saturation concentration and not their O:C ratio, only one equilibrium calculation is performed per volatility bin of each category (i.e., fuel combustion, biomass burning, anthropogenic VOC products, and biogenic VOC products). This approach reduces significantly the number of ~~equation~~equations to be solved and the corresponding computational cost of the phase partitioning calculations.

The ORACLE core layer calculates the aerosol composition at equilibrium by solving the following set of n nonlinear equations:

$$C_{a,i} = C_{t,i} \quad x_i C_i^* \text{ for } i = 1, n \qquad (11)$$

$$C_{a,i} = C_{t,i} - x_i C_i^* \text{ for } i = 1, n \qquad (12)(11)$$

$$x_i = \frac{C_{a,i}/M_i}{\sum_{i=1}^{n} C_{a,i}/M_i} \qquad (12) \qquad\qquad x_i = \frac{C_{a,i}/M_i}{\sum_{i=1}^{n} C_{a,i}/M_i}$$

where $C_{t,i}$ and $C_{a,i}$ are the total and aerosol-phase concentrations of product i in µg m^{-3}, respectively, C_i^* is the effective saturation concentration of species i, x_i is the mole fraction of product i in the absorbing organic phase, and M_i is the molecular weight of product i which corresponds to the weighted average molecular weight of the species with the same saturation concentration.

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

$$c_i^* = c_{i,0}^* \frac{T_0}{T} \exp\left[\frac{\Delta H}{R}\left(\frac{1}{T_0} - \frac{1}{T}\right)\right] \qquad (13)$$

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}, 10^0, 10^2, 10^4, and 10^6 µg m^{-3} volatility bins,

355 respectively, of all organic compounds from fuel combustion and biomass burning sources based

356 on data for large saturated species commonly found in primary emissions (Donahue et al., 2006).

357 Water uptake by the particulate organic phase under high relative humidity conditions can

358 increase the aerosol mass and decrease the mole fraction of individual organic products which may

359 allow additional condensation of SOA (Jathar et al., 2016). However, Eq. 12 calculates the mole

360 fraction of organic products only in the absorbing organic phase without taking into account the

361 presence of water in aerosol phase. This simplification can result in an underestimation of SOA in

362 areas with high humidity and significant SOA concentrations.

363 Assuming that the distribution of the species of the same volatility bin in the different O:C bins

364 does not change due to the condensation or evaporation, the gas and aerosol concentrations of each

365 compound in the 2-D space after the phase partitioning calculations are given by:

366

367
$$SOG_{i,j}(t+\Delta t) = \left(\sum_{j=1}^{m} SOG_{i,j}(t+\Delta t) \right) \frac{SOG_{i,j}(t) + SOA_{i,j}(t)}{\sum_{j=1}^{m} SOG_{i,j}(t) + \sum_{j=1}^{m} SOA_{i,j}(t)} \quad (13)$$

368
$$SOG_{i,j}(t+\Delta t) = \left(\sum_{j=1}^{m} SOG_{i,j}(t+\Delta t) \right) \frac{SOG_{i,j}(t) + SOA_{i,j}(t)}{\sum_{j=1}^{m} SOG_{i,j}(t) + \sum_{j=1}^{m} SOA_{i,j}(t)}$$

Field Code Changed

369

370 and

371

372
$$SOA_{i,j}(t+\Delta t) = \left(\sum_{j=1}^{m} SOA_{i,j}(t+\Delta t) \right) \frac{SOG_{i,j}(t) + SOA_{i,j}(t)}{\sum_{j=1}^{m} SOG_{i,j}(t) + \sum_{j=1}^{m} SOA_{i,j}(t)} \quad (14)$$

373
$$SOA_{i,j}(t+\Delta t) = \left(\sum_{j=1}^{m} SOA_{i,j}(t+\Delta t) \right) \frac{SOG_{i,j}(t) + SOA_{i,j}(t)}{\sum_{j=1}^{m} SOG_{i,j}(t) + \sum_{j=1}^{m} SOA_{i,j}(t)}$$

Field Code Changed

374

375 where, SOG and SOA are the gas and aerosol concentrations of each compound, respectively, i and

376 j are the volatility and O:C binsbin index in the 2-D space, respectively, m is the total number of

377 O:C bins, and t and $t+\Delta t$ are the times before and after the phase partitioning calculations,

378   respectively. $\sum_{j=1}^{m} SOG_{i,j}(t+\Delta t)$ and $\sum_{j=1}^{m} SOA_{i,j}(t+\Delta t)$ for each volatility bin i in eqs. 13 and 14

Field Code Changed

Field Code Changed

379   correspond to $C_{a,i}$ and $(C_{t,i} - C_{a,i})$ at equilibrium calculated from eqs. 11 and 12.

380

381

382   **4. Model results**

383   **4.1 Total OA and O:C**

384   Simulated total OA concentrations are high over regions affected by anthropogenic fuel

385   combustion and open biomass burning (Figure 22a). The highest annual average OA

386   concentrations at the surface are predicted over the densely populated areas of Eastern China,

387   Northern India and Bangladesh (10-28 μg m^{-3}), as well as over the tropical forest in the Congo

388   Basin (12-22 μg m^{-3}). Considerably high OA concentrations are also predicted over the tropical

389   forests of Southeast Asia and the Amazon (6-17 μg m^{-3}). Strong fossil fuel combustion sources

390   over the Arabian Peninsula result in annual average OA concentrations of 6-10 μg m^{-3}. The

391   calculated zonal average concentrations of total OA at surface peak over the mid-latitudes of the

392   Northern Hemisphere and the tropics (~ 4 μg m^{-3}) remain relatively high (>2 μg m^{-3}) up to 750

393   hPa altitude (Figure 2b). Over Europe, predicted OA annual average concentrations at the ground

394   level are in the 3-7 μg m^{-3} range. Over North America, the highest OA concentration is simulated

395   over Southern California, and also over the Mexico City metropolitan area, and the Southeastern

396   USA (2-8 μg m^{-3}). OA concentrations in the 0.5-1 μg m^{-3} range are predicted over parts of the

397   oceans due to the long-range transport of OA from the adjacent continental sources. Non-

398   negligible concentrations of OA can be found over the Greenland and Antarctica due to the long-

399   range transport of organic compounds and their condensation in the particulate phase under very

400   low temperatures. The calculated total tropospheric OA burden is 3.3 Tg, which is higher

401   compared to the calculated tropospheric burden of ORACLE v1.0 (2 Tg; Tsimpidi et al., 2016) but

402   is within the range of OA tropospheric burdens (0.7 to 3.8 Tg) from 31 global CTMs reported by

403   Tsigaridis et al. (2014).

404   Figure 3a depicts the annual average simulated O:C ratio of total OA at the surface. Lower

405   values of O:C are predicted close to OA sources, i.e., over the industrialized areas of the Northern

406   Hemisphere, and over the tropical and boreal forests. The lowest average values occur over the

407   boreal forests (as low as 0.3) due to the limited photochemical activity over these regions, in

408   contrast to the tropical forests, where O:C is around 0.5. Over the densely populated areas of Asia,

409   Europe and North America, the O:C is about 0.3-0.4, i.e., close to the anthropogenic sources. O:C

410   levels increase rapidly (in excess of 0.6) downwind of the sources due to photochemical aging of

411   the transported OA. The highest O:C values are calculated over the Sahara Desert and the remote

412   Atlantic and Pacific Oceans (0.8-1), however, over these regions, OA concentrations are low

413   (Figure 2). O:C ratios increase significantly with altitude according to the model (Figure 3b) since

414   organic vapors transported vertically react with OH, forming products with higher oxygen content.

415   Over the mid-latitudes of the Northern Hemisphere, the average O:C ratio near the surface is 0.55,

416   and increases as the air masses are transported aloft by approximately 0.05 for every 100 hPa

417   (Figure 3b).

418

419   **4.2 POA and SOA**

420   Figure 4a depicts the decadal average simulated POA concentrations at the surface. POA

421   concentrations are high over the densely populated areas of the Northern Hemisphere due to strong

422   fuel combustion emissions from the industrial, energy, residential and transport sectors. The

423   highest concentrations are calculated over Eastern China (3-13 μg m^{-3}), Bangladesh (2-8 μg m^{-3}),

424   and Eastern Europe (1-3 μg m^{-3}). Open biomass burning emissions from forest, woodland, peatland

425   and savannah fires result in high POA concentrations over the tropics (3-8 μg m^{-3} over the Congo

426   Basin) and the boreal forests (1-6 μg m^{-3} over Russia). However, a large fraction of POA

427   evaporates due to dilution as the air masses travel downwind from the sources, resulting in

428   significant reduction of the concentration (Figure 4a). Then, the material that is transferred to the

429   gas phase can be oxidized and recondense to the aerosol phase forming SOA that persists even far

430   from the sources (Figure 4b). This results in a relatively homogeneous regional distribution of

431   SOA with a continental background of 2 μg m^{-3} and high concentrations even far downwind from

432   anthropogenic (e.g., 7-22 μg m^{-3} over South and Eastern Asia) and open biomass burning (e.g., 9-

433   17 μg m^{-3} over Central Africa) sources. Lower concentrations are predicted over the boreal forests

434   (~1 μg m^{-3}) due to minor photochemical activity.

435   The ~~calculated~~ tropospheric burden of POA and SOA calculated by ORACLE 2-D is 0.25

436   Tg and 3.05 Tg, respectively. While the value of POA is very similar to the corresponding

437   tropospheric burden of ORACLE ~~v1.0~~ (0.24 Tg), the tropospheric burden of SOA has increased

438 significantly (1.74 Tg in ORACLE ~~v1.0~~). The tropospheric burden of POA calculated by ORACLE
439 2-D is much lower than most global CTMs in the AEROCOM intercomparison study (mean value
440 of 0.85 Tg) ~~(Tsigaridis et al., 2014).~~(Tsigaridis et al., 2014). This difference is due to the
441 evaporation of POA and its conversion into SOA in ORACLE 2-D, given that this dynamic
442 behavior of POA is not taken into account by most ~~of the~~ global CTMs. ~~On the other~~
443 ~~hand~~Consequently, the ORACLE 2-D calculated SOA tropospheric burden is higher than most
444 CTMs from the AEROCOM intercomparison study (mean value of 1 Tg) (Tsigaridis et al., 2014)
445 due to its stronger chemical production. As indicated by Tsimpidi et al. (2016), the POA burden is
446 underestimated by our model, especially in the urban environment in the winter~~, because our~~
447 ~~emission inventory does not yet account for residential wood burning (Denier van der Gon et al.,~~
448 ~~2015).~~. This underestimation was partially attributed to missing residential wood burning
449 emissions in our inventory (van der Gon et al., 2015).
450
451 **4.3 Fresh and ~~Aged~~aged SOA**
452   The major advantage, at least in this initial phase, of extending the ORACLE module to
453 describe the oxygen content of OA is the model's ability to quantify the degree of photochemical
454 processing of the OA. The model can distinguish the fresh SOA that is relatively less oxygenated
455 and the highly aged and oxygenated SOA. As a first approximation, OA compounds with O:C ≤
456 0.6 are considered "fresh SOA" and OA compounds with O:C > 0.6 are considered as "aged SOA".
457   Figure 5 depicts the average concentrations of fresh and aged SOA and their fractional
458 contributions to total SOA at the surface. Fresh SOA exceeds aged SOA close to the source areas
459 (Figures 5c). On the other hand, the spatial distribution of aged SOA extends further from the
460 sources with relatively high concentrations even over remote continental (e.g., Sahara) and oceanic
461 (e.g., North Pacific) regions (Figure 5b). Aged SOA maximizes over north India since southerly
462 monsoon winds, prevailing during summer, favor the transport of aged organic compounds to the
463 north where they can accumulate at the foothills of the Himalaya mountain range before being
464 removed by convection and precipitation. The fraction of aged SOA is higher over the
465 industrialized regions (0.4-0.7), where IVOCs comprise 70% of total fuel combustion emissions,
466 and lower over the tropics (0.1-0.5), where IVOCs represent 40% of total open biomass burning
467 emissions (Figure 5d). Overall, the tropospheric burden of fresh SOA is 1.26 Tg and of aged SOA
468 1.79 Tg.

469   The simulated vertical profiles of fresh and aged SOA are quite different (Figure 6). Both
470 fresh and aged SOA concentrations are high near the surface with zonal averages of 0.9-1.7 μg m⁻
471 ³ over the northern mid-latitudes and the tropics. However, at higher altitudes the oxidation of
472 fresh SOA continues, which leads to the transformation into aged SOA. Therefore, above 960 hPa
473 altitude the zonal average concentrations of fresh SOA decrease gradually (below 1 μg m^{-3}) and
474 that of aged SOA increase, exceeding in some cases 1.5 μg m^{-3} at 850 hPa altitude. Further aloft
475 in the atmosphere (above 700 hPa altitude) the levels of both fresh and aged SOA are reduced
476 significantly due to dilution and removal.

477

478 **4.3 2-D space distribution**

479 In this section, we present a new feature that comes along with the upgraded ORACLE 2-D.
480 Since ORACLE 2-D explicitly describes the concentration of organic surrogate species in two
481 dimensional space, defined by their volatility and O:C ratio, it can also provide as output ~~their~~its
482 distribution in this 2-D space. As an example, Figure 7 depicts the 2-D distribution of the average
483 total OA concentrations during the years 2001-2010 over central Europe and the Amazon basin.
484 The first left column represents all the species with $C^* \leq 10^{-2}$ μg m^{-3}, while the top line represents
485 all the species with O:C $\geq$ 1.2.

486   Over Europe, approximately 50% of OA has $C^* = 1$ μg m^{-3}. About 10% of the OA is emitted
487 directly in this volatility bin as POA with O:C = 0.1, while the rest, 40%, is SOA with higher O:C
488 from the aging of more volatile compounds. The volatility bin with $C^* = 10^{-2}$ μg m^{-3} is also
489 important, containing 18% of the total OA. On the other hand, volatility bins with $C^* > 10^2$ contain
490 less than 1% of total OA since organic species in these high volatility bins exist mainly in the gas
491 phase. Furthermore, about 30% of the total OA over Europe has O:C$\leq$0.2, mainly from direct
492 emissions of POA from fuel combustion sources with O:C=0.1 (20% of total OA). Organic species
493 with O:C = 0.4 are also important, representing 15% of the total OA. Overall, 40% of OA over
494 Europe has 0.3$\leq$O:C$\leq$ 0.6 and 30% is highly oxidized with O:C$\geq$0.7.

495   Over the Amazon, similar to Europe, approximately 50% of total OA has $C^* = 1$ μg m^{-3}
496 while bins with $C^* > 10^2$ contain only 2% of total OA. However, volatility bins with $10^2 \leq C^* \leq 10^3$
497 are more important compared to Europe containing about 30% of the total OA concentrations. This
498 discrepancy can be attributed to the high bSOA-v concentrations over the Amazon rainforest,
499 which remain within higher volatility bins during their photochemical aging compared to aSOA-

v. Furthermore, similar to Europe, about 30% of total OA over the Amazon has O:C ≤ 0.2, mainly affected by the direct aerosol emissions from biomass burning with O:C = 0.2 (~20% of total OA). OA species with 0.3≤O:C≤0.6 represent 55% of the total OA with mostly O:C = 0.4 (35% of total OA). On the other hand, only 15% of the total OA is highly oxidized with O:C ≥ 0.7. Comparing these results to Europe shows that OA over biomass burning areas (e.g., the Amazon) are estimated to be less oxidized than OA over anthropogenic areas (e.g., Europe). despite the higher photochemical activity over the tropics. This perhaps unexpected result can be explained by the higher fraction of IVOC emissions over the industrialized regions that can increase the overall oxidation state of OA through their multi-generational aging.

5. Model performance evaluation

The mean absolute gross error (MAGE), mean bias (MB), normalized mean error (NME), normalized mean bias (NMB), and the root mean square error (RMSE) are used to assess the model performance of the simulations of OA components (i.e., POA, fresh and aged SOA) and O:C.

$$MAGE = \frac{1}{N}\sum_{i=1}^{N}|P_i - O_i| \quad (15) \qquad MB = \frac{1}{N}\sum_{i=1}^{N}(P_i - O_i) \quad (16)$$

$$NME = \frac{\sum_{i=1}^{N}|P_i - O_i|}{\sum_{i=1}^{N}O_i} \quad (17) \qquad NMB = \frac{\sum_{i=1}^{N}(P_i - O_i)}{\sum_{i=1}^{N}O_i} \quad (18)$$

$$RMSE = \left[\frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)^2\right]^{\frac{1}{2}} \quad (19)$$

where O_i is the observed campaign average value of the ith OA component or O:C, P_i is the corresponding modelled value during the same period, and N is the total number of comparisons used for the evaluation. (Tsimpidi et al., 2016).

5.1 Organic aerosol concentrations

526   The simulated POA and SOA can be compared against AMS factor analysis results from
527   61 field campaigns performed during the period 2001–2010 over urban-downwind and rural
528   environments in the Northern Hemisphere (Figure 8). Information for each of these campaigns is
529   given in Tsimpidi et al. (2016). Factor-analysis techniques classify OA into hydrocarbon-like OA
530   (HOA), biomass burning OA (BBOA), and oxygenated OA (OOA). HOA is assumed to
531   correspond to POA from fossil fuel combustion, and BBOA corresponds to POA from biomass
532   burning (Crippa et al., 2014). Therefore, simulated POA is compared here against the sum of AMS
533   HOA and BBOA (Table 1). OOA corresponds to modelled SOA (Table 2) and can be further
534   classified into two subtypes, semi-volatile OOA (SVOOA) and low volatility OOA (LVOOA)
535   (Crippa et al., 2014). Recent studies have suggested that the main difference between these two
536   OOA types is often not so much their volatility, but mostly their oxygen content (Kostenidou et
537   al., 2016; Louvaris et al., 2017). These two OOA types represent distinct oxidation states with O:C
538   of 0.33-0.67 for SVOOA and 0.67-1.00 for LVOOA (Donahue et al., 2012). Therefore, SVOOA
539   can be compared at least as a zero-order approximation against the simulated fresh SOA (Table 3)
540   and LVOOA against the simulated aged SOA (Table 4).

541   The model reproduces the observed campaign average POA concentrations ~~well~~within a factor
542   of two in ~~most~~40% of the cases over ~~both the~~ urban downwind and rural locations (Table 1, Figure
543   8a). The average simulated POA concentration over the urban downwind regions is around 0.65
544   $\mu g\ m^{-3}$ and it decreases further after continued transport from the sources to 0.45 $\mu g\ m^{-3}$ over the
545   rural areas. Compared to AMS HOA concentrations, modelled POA is unbiased over rural
546   environments, however, it is underestimated downwind of urban areas~~, as indicated in Sect. 4.2~~
547   (MB=-0.17 $\mu g\ m^{-3}$). The model has the best performance during summer with MB=-0.05 $\mu g\ m^{-3}$
548   and RMSE = 0.40 $\mu g\ ——m^{-3}$. Compared to ORACLE (Tsimpidi et al., 2016), ORACLE 2-D
549   produces almost identical concentrations of POA, hence the model performance is unchanged.

550   Calculated SOA concentrations are higher than POA both downwind of major urban centers
551   (2.76 $\mu g\ m^{-3}$) and rural locations (2.53 $\mu g\ m^{-3}$). The model reproduces the observed campaign
552   average SOA concentrations within a factor of two in 65% of the cases over urban downwind and
553   rural locations. Simulated average SOA concentrations are slightly low-biased compared to AMS
554   OOA measurements downwind of urban areas (MB=-0.22 $\mu g\ m^{-3}$) and over rural areas (MB=-0.18
555   $\mu g\ m^{-3}$). While the model performs well during spring (NMB=~~10~~11%), summer (NMB=-10%),
556   and autumn (NMB=0%), it strongly underestimates SOA concentrations during winter (NMB=-

557 76%). This underestimation of SOA by the model is mostly due to missing LVOOA. For the 41

558 campaigns in which both OOA types were identified, EMAC reproduces the fresh SOA

559 concentrations well (compared to SVOOA) with a MB of 0.41 and 0.32 μg m^{-3} downwind of urban

560 and in rural locations, respectively (Table 4). Furthermore, except during winter, the model is also

561 able to capture the seasonal variations of the fresh SOA concentration (Table 4). However, the

562 model underpredicts the aged SOA concentrations (compared to AMS measured LVOOA) with a

563 MB of -0.95 and -0.70 μg m^{-3} over urban downwind and rural locations, respectively. This aged

564 SOA underestimation is also evident during the four seasons of the year and especially during

565 winter (NMB = -93%). This result may indicate that the model underestimates the atmospheric

566 aging rate of SOA or misses processes forming highly oxidized OA, e.g., highly oxidized ELVOCs

567 from the ozonolysis of terpenes (Ehn et al., 2014; Jokinen et al., 2016). ~~Another~~Therefore, another

568 useful feature of ORACLE 2-D is that detailed AMS measurements can be used to gain further

569 insights into what causes biases and errors in its OA predictions. Compared to ORACLE (Tsimpidi

570 et al., 2016), ORACLE 2-D tends to produce higher fresh SOA concentrations and slightly lower

571 aged SOA concentrations. However, this discrepancy may be partially due to the fact that in

572 ORACLE fresh SOA is assumed to correspond only to first generation oxidation products, while

573 in ORACLE 2-D fresh SOA is defined based on the O:C ratio, and includes higher generation

574 oxidation products as well. Overall, the simulated average total SOA concentrations in ORACLE

575 2-D (2.59 μg m^{-3}) are higher than ORACLE (1.91 μg m^{-3}) reducing the gap with the corresponding

576 AMS OOA (2.78 μg m^{-3}).

577

578 **5.2 O:C ratio**

579   The simulated O:C ratio of total OA and SOA is compared against observed O:C of OA

580 and OOA from 30 and ~~58~~57 field campaigns, respectively, performed during the period 2001–

581 2010 in the Northern Hemisphere (Tables 5, 6). Given that global models, including EMAC,

582 underestimate concentrations of POA and SOA over urban locations (Tsigaridis et al., 2014;

583 Tsimpidi et al., 2016), AMS data from these locations is not included for the statistical evaluation

584 of the calculated O:C (Figure 9, Tables 7, 8). In fact, the model tends to overestimate the O:C of

585 total OA compared to observations from urban locations (e.g., Mexico City, Barcelona, New York,

586 Riverside, Paris; Table 5). This overprediction can be attributed to the coarse spatial resolution

587 and the difficulty to represent freshly emitted (or formed) OA on a local scale. On the other hand,

588 the model performs remarkably well in reproducing the O:C ratio of both total OA (NMB=7%,
589 RMSE=0.13) and SOA (NMB=5%, RMSE=0.12) in non-urban areas.

590   The model tends to overestimate the O:C ratio of total OA compared to observations from field

591 campaigns close to the coasts (e.g., western coast of Chile, Mace Head). The model reproduces
592 very well the low O:C ratios of total OA during winter and the higher values during spring and
593 autumn (Table 7). Even if the model substantially underpredicts all types of OA during wintertime,
594 the calculated OA O:C is similar to observations because the relative contribution of each OA
595 compound is captured by the model. The slightly larger underprediction of aged SOA (compared
596 to fresh SOA and POA) during winter results in a small underprediction of total OA O:C (NMB =
597 -7%). This indicates that the missing OA during winter cannot be attributed only to missing POA
598 sources (e.g., residential biofuel use) but also to missing SOA formation pathways (e.g.,
599 multiphase oxidation). On the other hand, it overestimates the O:C of total OA during summer
600 (MB=0.14). However, there are only 4 fieldField campaigns conducted during the summer months
601 that provide O:C of total OA with surprisingly low O:C values for this time of the year (0.3-0.4)
602 which are not captured by the model. Overall, the O:C of total OA is slightly overestimated by the
603 model (MB=0.04).

604   Simulated O:C of SOA is relatively high (0.5-0.7) at most observational sites, which agrees
605 well with measurements. The model also captures the very high O:C values (larger than 0.7)
606 observed over very remote areas (e.g., Okinawa, Finokalia). On the other hand, low O:C ratios
607 (lower than 0.5), reported by a few field campaigns, are overestimated by the model (Duke forest,
608 Rhine Valley, Jiaxing). Remarkably, the model performance during winter is unbiased for the O:C
609 ratio of SOA (Table 8). This is in contrast to its inability to reproduce SOA concentrations during
610 winter (Tsimpidi et al., 2016). During the other seasons the model slightly overestimates the O:C
611 ratio of SOA with an MB ranging from 0.02 (during spring) to 0.07 (during summer and autumn).

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

618

619

6. Conclusions

The ORACLE module for the description of OA composition and evolution in the atmosphere (Tsimpidi et al., 2014) has been extended to simultaneously simulate the volatility and the oxygen content of OA that results from atmospheric aging. Similar to ORACLE v1.0, the new version is implemented in the EMAC CCM, and considers the formation of OA from the emissions and chemical aging of LVOCs, SVOCs and IVOCs from fossil fuel, biofuel and biomass burning related sources, as well as from the oxidation of anthropogenic and biogenic VOC precursors. The updated ORACLE module employs the 2-D VBS framework that uses logarithmically spaced effective saturation concentration bins to describe the volatility of organic compounds, and linearly spaced oxygen per carbon ratio bins to describe their oxygenation state.

The simulated concentrations of OA and its components (i.e., POA and SOA) are similar to ORACLE v1.0 with relatively high concentrations over the industrialized areas in the Northern Hemisphere and the biomass burning areas in the tropics. The tropospheric burdens of POA and SOA are calculated to be 0.25 and 3.05 Tg respectively, the latter being higher than with ORACLE v1.0. The new ORACLE 2-D module additionally allows calculation of the oxidation state of OA (in terms of O:C), and therefore, its classification into fresh SOA (with O:C lower than 0.6) and aged SOA (with O:C higher than 0.6). O:C is calculated to be relatively low close to source regions and at high latitudes. The predicted O:C ratio is as low as 0.3 over the boreal forests, 0.3-0.4 close to the anthropogenic sources in the Northern Hemisphere, 0.5 over the tropical forests and higher than 0.6 downwind of source areas and at altitudes aloft. Accordingly, fresh SOA concentrations are higher close to sources while aged SOA increases as the air masses are transported away from the sources and to higher altitudes. The estimated tropospheric burden of fresh SOA is 1.26 Tg and of aged SOA 1.79 Tg. The analysis of model results regarding the distribution of OA in the 2-D space of volatility and oxygen content showed that half of OA has $C^* = 1$ μg m^{-3} over both anthropogenic areas (e.g., Europe) and tropical forests (e.g., Amazon). Furthermore, over Europe OA compounds are more strongly oxidized than over the Amazon, and consist of 40% fresh OA and 30% aged SOA (compared to 55% and 15%, respectively, over the Amazon). The remainder 30% in both areas consists of POA or very low oxidized material with O:C≤0.2.

The simulated OA components (POA, fresh, aged and total SOA) have been compared with observed subtypes of OA (sum of HOA and BBOA, SVOOA, LVOOA and OOA) derived

from AMS factor analysis results from 61 field campaigns, performed during the period 2001–2010 over urban-downwind and rural environments in the Northern Hemisphere. The model reproduces the POA concentrations over both types of environment with low bias. While the model appears to capture the fresh SOA concentrations reasonably well, it underestimates the concentrations of aged SOA, resulting in an overall underprediction of total SOA. While this aged SOA underestimation is evident throughout the year, it is more pronounced during winter. The underestimation of aged SOA by the model emphasizes the need to better describe chemical aging processes and further explore the effect of ELVOCs on the formation of SOA. Furthermore, the model realistically reproduces the observed O:C ratio of total OA (NMB=7%) and OOA (NMB=5%) from 50 field campaigns over urban-downwind and rural areas. It adequately captures the low values during winter and the higher during spring and autumn, however, it overestimates the O:C of total OA during summer. In addition, the model slightly overestimates the O:C ratio of total OOA (NMB=5%), except during winter when modeled OOA O:C is unbiased.

ORACLE 2-D is a flexible module that efficiently describes the organic aerosol composition and evolution in the atmosphere by simulating the changes in OA volatility and oxygen content throughout its lifetime in the atmosphere. The ability of ORACLE 2-D to simulate the degree of OA oxidation can help determine the changing OA hygroscopicity during atmospheric aging. ORACLE 2-D can potentially provide valuable insights into the composition and reactivity of OA and the physicochemical evolution during atmospheric transport of OA, which can help reduce aerosol-related uncertainties that persist in global atmospheric chemistry and climate modeling.

Code and Data Availability

The complete ORACLE 2-D code can be obtained by applying for an EMAC license or upon request by emailing the first author. To use ORACLE 2-D as part of EMAC, please first apply for an ECHAM5 and a MESSy license. The GCM ECHAM5 has been developed at the Max Planck Institute for Meteorology in Hamburg (see: www.mpimet.mpg.de/en/wissenschaft/modelle/echam/echam5.html). The Modular Earth Submodel System (MESSy) is developed and applied by a consortium of institutions, initiated by

the Max Planck Institute for Chemistry. The usage of MESSy and access to the source code is licensed to all affiliates of institutions that are members of the MESSy Consortium. Institutions can be a member of the MESSy Consortium by signing the MESSy Memorandum of Understanding. More information can be found on the MESSy Consortium Website (www.messy-interface.org). The measurement data used for the evaluation of the model can be found in Tables 5 and 6 of this manuscript and Tables S6 and S7 in the supplement of Tsimpidi et al. (2016) (available online at ~~https://doi.org/10.5194/acp-16-8939-2016-supplement)~~.https://doi.org/10.5194/acp-16-8939-2016-supplement).

Acknowledgements

A.P. Tsimpidi acknowledges support from a DFG individual grand programme (project reference TS 335/2-1) and V.A. Karydis acknowledges support from a FP7 Marie Curie Career Integration Grant (project reference 618349).

References

Ahmadov, R., McKeen, S. A., Robinson, A. L., Bahreini, R., Middlebrook, A. M., de Gouw, J. A., Meagher, J., Hsie, E. Y., Edgerton, E., Shaw, S., and Trainer, M.: A volatility basis set model for summertime secondary organic aerosols over the eastern United States in 2006, J. Geophys. Res. Atmos., 117, 10.1029/2011jd016831, 2012.

Aiken, A. C., Decarlo, P. F., Kroll, J. H., Worsnop, D. R., Huffman, J. A., Docherty, K. S., Ulbrich, I. M., Mohr, C., Kimmel, J. R., Sueper, D., Sun, Y., Zhang, Q., Trimborn, A., Northway, M., Ziemann, P. J., Canagaratna, M. R., Onasch, T. B., Alfarra, M. R., Prevot, A. S. H., Dommen, J., Duplissy, J., Metzger, A., Baltensperger, U., and Jimenez, J. L.: O/C and OM/OC ratios of primary, secondary, and ambient organic aerosols with high-resolution time-of-flight aerosol mass spectrometry, Environmen. Sci. & Technol., 42, 4478-4485, 2008.

Allan, J. D., Williams, P. I., Morgan, W. T., Martin, C. L., Flynn, M. J., Lee, J., Nemitz, E., Phillips, G. J., Gallagher, M. W., and Coe, H.: Contributions from transport, solid fuel burning and cooking to primary organic aerosols in two UK cities, Atmospheric Chemistry and Physics, 10, 647-668, 10.5194/acp-10-647-2010, 2010.

Athanasopoulou, E., Vogel, H., Vogel, B., Tsimpidi, A. P., Pandis, S. N., Knote, C., and Fountoukis, C.: Modeling the meteorological and chemical effects of secondary organic aerosols during an EUCAARI campaign, Atmos. Chem. Phys., 13, 625-645, 2013.

Bacer, S., Sullivan, S. C., Karydis, V. A., Barahona, D., Krämer, M., Nenes, A., Tost, H., Tsimpidi, A. P., Lelieveld, J., and Pozzer, A.: Implementation of a comprehensive ice crystal formation parameterization for cirrus and mixed-phase clouds into the EMAC model (based on MESSy 2.53), Geosci. Model Dev. Discuss., 2018, 1-26, 10.5194/gmd-2018-62, 2018.

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

719 calibration, and implications, Atmospheric Chemistry and Physics, 15, 253-272, 10.5194/acp-15-253-2015,
720 2015.
721 Carbone, S., Aurela, M., Saarnio, K., Saarikoski, S., Timonen, H., Frey, A., Sueper, D., Ulbrich, I. M., Jimenez,
722 J. L., Kulmala, M., Worsnop, D. R., and Hillamo, R. E.: Wintertime Aerosol Chemistry in Sub-Arctic Urban
723 Air, Aerosol Science and Technology, 48, 313-323, 10.1080/02786826.2013.875115, 2014.
724 Chacon-Madrid, H. J., Henry, K. M., and Donahue, N. M.: Photo-oxidation of pinonaldehyde at low NOx:
725 from chemistry to organic aerosol formation, Atmospheric Chemistry and Physics, 13, 3227-3236,
726 10.5194/acp-13-3227-2013, 2013.
727 Chang, R. Y. W., Slowik, J. G., Shantz, N. C., Vlasenko, A., Liggio, J., Sjostedt, S. J., Leaitch, W. R., and Abbatt,
728 J. P. D.: The hygroscopicity parameter (kappa) of ambient organic aerosol at a field site subject to biogenic
729 and anthropogenic influences: relationship to degree of aerosol oxidation, Atmospheric Chemistry and
730 Physics, 10, 5047-5064, 10.5194/acp-10-5047-2010, 2010.
731 Ciarelli, G., El Haddad, I., Bruns, E., Aksoyoglu, S., Mohler, O., Baltensperger, U., and Prevot, A. S. H.:
732 Constraining a hybrid volatility basis-set model for aging of wood-burning emissions using smog chamber
733 experiments: a box-model study based on the VBS scheme of the CAMx model (v5.40), Geoscientific
734 Model Development, 10, 2303-2320, 10.5194/gmd-10-2303-2017, 2017.
735 Crippa, M., El Haddad, I., Slowik, J. G., DeCarlo, P. F., Mohr, C., Heringa, M. F., Chirico, R., Marchand, N.,
736 Sciare, J., Baltensperger, U., and Prevot, A. S. H.: Identification of marine and continental aerosol sources
737 in Paris using high resolution aerosol mass spectrometry, Journal of Geophysical Research-Atmospheres,
738 118, 1950-1963, 10.1002/jgrd.50151, 2013.
739 Docherty, K. S., Aiken, A. C., Huffman, J. A., Ulbrich, I. M., DeCarlo, P. F., Sueper, D., Worsnop, D. R., Snyder,
740 D. C., Peltier, R. E., Weber, R. J., Grover, B. D., Eatough, D. J., Williams, B. J., Goldstein, A. H., Ziemann, P.
741 J., and Jimenez, J. L.: The 2005 Study of Organic Aerosols at Riverside (SOAR-1): instrumental
742 intercomparisons and fine particle composition, Atmospheric Chemistry and Physics, 11, 12387-12420,
743 10.5194/acp-11-12387-2011, 2011.
744 Donahue, N. M., Robinson, A. L., Stanier, C. O., and Pandis, S. N.: Coupled partitioning, dilution, and
745 chemical aging of semivolatile organics, Environ. Sci. Technol., 40, 2635-2643, 2006.
746 Donahue, N. M., Epstein, S. A., Pandis, S. N., and Robinson, A. L.: A two-dimensional volatility basis set: 1.
747 organic-aerosol mixing thermodynamics, Atmos. Chem. and Phys., 11, 3303-3318, 2011.
748 Ehn, M., Thornton, J. A., Kleist, E., Sipila, M., Junninen, H., Pullinen, I., Springer, M., Rubach, F., Tillmann,
749 R., Lee, B., Lopez-Hilfiker, F., Andres, S., Acir, I.-H., Rissanen, M., Jokinen, T., Schobesberger, S.,
750 Kangasluoma, J., Kontkanen, J., Nieminen, T., Kurten, T., Nielsen, L. B., Jorgensen, S., Kjaergaard, H. G.,
751 Canagaratna, M., Maso, M. D., Berndt, T., Petaja, T., Wahner, A., Kerminen, V.-M., Kulmala, M., Worsnop,
752 D. R., Wildt, J., and Mentel, T. F.: A large source of low-volatility secondary organic aerosol, Nature, 506,
753 doi: 10.1038/nature13032, 2014.
754 Fountoukis, C., Megaritis, A. G., Skyllakou, K., Charalampidis, P. E., Pilinis, C., van der Gon, H., Crippa, M.,
755 Canonaco, F., Mohr, C., Prevot, A. S. H., Allan, J. D., Poulain, L., Petaja, T., Tiitta, P., Carbone, S., Kiendler-
756 Scharr, A., Nemitz, E., O'Dowd, C., Swietlicki, E., and Pandis, S. N.: Organic aerosol concentration and
757 composition over Europe: insights from comparison of regional model predictions with aerosol mass
758 spectrometer factor analysis, Atmospheric Chemistry and Physics, 14, 9061-9076, 10.5194/acp-14-9061-
759 2014, 2014.
760 Gao, C. Y., Tsigaridis, K., and Bauer, S. E.: MATRIX-VBS (v1.0): implementing an evolving organic aerosol
761 volatility in an aerosol microphysics model, Geoscientific Model Development, 10, 751-764, 10.5194/gmd-
762 10-751-2017, 2017.
763 Ge, X. L., Setyan, A., Sun, Y. L., and Zhang, Q.: Primary and secondary organic aerosols in Fresno, California
764 during wintertime: Results from high resolution aerosol mass spectrometry, Journal of Geophysical
765 Research-Atmospheres, 117, 10.1029/2012jd018026, 2012.

766 Gong, Z. H., Lan, Z. J., Xue, L., Zeng, L. W., He, L. Y., and Huang, X. F.: Characterization of submicron aerosols
767 in the urban outflow of the central Pearl River Delta region of China, Frontiers of Environmental Science
768 & Engineering, 6, 725-733, 10.1007/s11783-012-0441-8, 2012.
769 Guo, H. Y., Liu, J. M., Froyd, K. D., Roberts, J. M., Veres, P. R., Hayes, P. L., Jimenez, J. L., Nenes, A., and
770 Weber, R. J.: Fine particle pH and gas-particle phase partitioning of inorganic species in Pasadena,
771 California, during the 2010 CalNex campaign, Atmospheric Chemistry and Physics, 17, 5703-5719,
772 10.5194/acp-17-5703-2017, 2017.
773 Hawkins, L. N., Russell, L. M., Covert, D. S., Quinn, P. K., and Bates, T. S.: Carboxylic acids, sulfates, and
774 organosulfates in processed continental organic aerosol over the southeast Pacific Ocean during VOCALS-
775 REx 2008, Journal of Geophysical Research-Atmospheres, 115, 10.1029/2009jd013276, 2010.
776 Hayes, P. L., Ortega, A. M., Cubison, M. J., Froyd, K. D., Zhao, Y., Cliff, S. S., Hu, W. W., Toohey, D. W., Flynn,
777 J. H., Lefer, B. L., Grossberg, N., Alvarez, S., Rappenglueck, B., Taylor, J. W., Allan, J. D., Holloway, J. S.,
778 Gilman, J. B., Kuster, W. C., De Gouw, J. A., Massoli, P., Zhang, X., Liu, J., Weber, R. J., Corrigan, A. L.,
779 Russell, L. M., Isaacman, G., Worton, D. R., Kreisberg, N. M., Goldstein, A. H., Thalman, R., Waxman, E. M.,
780 Volkamer, R., Lin, Y. H., Surratt, J. D., Kleindienst, T. E., Offenberg, J. H., Dusanter, S., Griffith, S., Stevens,
781 P. S., Brioude, J., Angevine, W. M., and Jimenez, J. L.: Organic aerosol composition and sources in
782 Pasadena, California, during the 2010 CalNex campaign, Journal of Geophysical Research-Atmospheres,
783 118, 9233-9257, 10.1002/jgrd.50530, 2013.
784 He, L. Y., Huang, X. F., Xue, L., Hu, M., Lin, Y., Zheng, J., Zhang, R. Y., and Zhang, Y. H.: Submicron aerosol
785 analysis and organic source apportionment in an urban atmosphere in Pearl River Delta of China using
786 high-resolution aerosol mass spectrometry, Journal of Geophysical Research-Atmospheres, 116,
787 10.1029/2010jd014566, 2011.
788 Hodzic, A., Kasibhatla, P. S., Jo, D. S., Cappa, C. D., Jimenez, J. L., Madronich, S., and Park, R. J.: Rethinking
789 the global secondary organic aerosol (SOA) budget: stronger production, faster removal, shorter lifetime,
790 Atmos. Chem. Phys., 16, 7917-7941, 10.5194/acp-16-7917-2016, 2016.
791 Holzinger, R., Kasper-Giebl, A., Staudinger, M., Schauer, G., and Rockmann, T.: Analysis of the chemical
792 composition of organic aerosol at the Mt. Sonnblick observatory using a novel high mass resolution
793 thermal-desorption   proton-transfer-reaction   mass-spectrometer   (hr-TD-PTR-MS),   Atmospheric
794 Chemistry and Physics, 10, 10111-10128, 10.5194/acp-10-10111-2010, 2010.
795 Hu, W. W., Hu, M., Hu, W., Jimenez, J. L., Yuan, B., Chen, W. T., Wang, M., Wu, Y. S., Chen, C., Wang, Z. B.,
796 Peng, J. F., Zeng, L. M., and Shao, M.: Chemical composition, sources, and aging process of submicron
797 aerosols in Beijing: Contrast between summer and winter, Journal of Geophysical Research-Atmospheres,
798 121, 1955-1977, 10.1002/2015jd024020, 2016.
799 Huang, X. F., He, L. Y., Hu, M., Canagaratna, M. R., Sun, Y., Zhang, Q., Zhu, T., Xue, L., Zeng, L. W., Liu, X.
800 G., Zhang, Y. H., Jayne, J. T., Ng, N. L., and Worsnop, D. R.: Highly time-resolved chemical characterization
801 of atmospheric submicron particles during 2008 Beijing Olympic Games using an Aerodyne High-
802 Resolution Aerosol Mass Spectrometer, Atmospheric Chemistry and Physics, 10, 8933-8945, 10.5194/acp-
803 10-8933-2010, 2010.
804 Huang, X. F., He, L. Y., Hu, M., Canagaratna, M. R., Kroll, J. H., Ng, N. L., Zhang, Y. H., Lin, Y., Xue, L., Sun, T.
805 L., Liu, X. G., Shao, M., Jayne, J. T., and Worsnop, D. R.: Characterization of submicron aerosols at a rural
806 site in Pearl River Delta of China using an Aerodyne High-Resolution Aerosol Mass Spectrometer,
807 Atmospheric Chemistry and Physics, 11, 1865-1877, 10.5194/acp-11-1865-2011, 2011.
808 Huang, X. F., Xue, L., Tian, X. D., Shao, W. W., Sun, T. L., Gong, Z. H., Ju, W. W., Jiang, B., Hu, M., and He, L.
809 Y.: Highly time-resolved carbonaceous aerosol characterization in Yangtze River Delta of China:
810 Composition, mixing state and secondary formation, Atmospheric Environment, 64, 200-207,
811 10.1016/j.atmosenv.2012.09.059, 2013.

812 Jathar, S. H., Farina, S. C., Robinson, A. L., and Adams, P. J.: The influence of semi-volatile and reactive
813 primary emissions on the abundance and properties of global organic aerosol, Atmos. Chem. Phys., 11,
814 7727-7746, 2011.
815 Jathar, S. H., Mahmud, A., Barsanti, K. C., Asher, W. E., Pankow, J. F., and Kleeman, M. J.: Water uptake by
816 organic aerosol and its influence on gas/particle partitioning of secondary organic aerosol in the United
817 States, Atmospheric Environment, 129, 142-154, 10.1016/j.atmosenv.2016.01.001, 2016.
818 Jo, D. S., Park, R. J., Kim, M. J., and Spracklen, D. V.: Effects of chemical aging on global secondary organic
819 aerosol using the volatility basis set approach, Atmos. Environ., 81, 230-244, 2013.
820 Jöckel, P., Tost, H., Pozzer, A., Bruehl, C., Buchholz, J., Ganzeveld, L., Hoor, P., Kerkweg, A., Lawrence, M.
821 G., Sander, R., Steil, B., Stiller, G., Tanarhte, M., Taraborrelli, D., Van Aardenne, J., and Lelieveld, J.: The
822 atmospheric chemistry general circulation model ECHAM5/MESSy1: consistent simulation of ozone from
823 the surface to the mesosphere, Atmos. Chem. Phys., 6, 5067-5104, 2006.
824 Karydis, V. A., Tsimpidi, A. P., Pozzer, A., Astitha, M., and Lelieveld, J.: Effects of mineral dust on global
825 atmospheric nitrate concentrations, Atmos. Chem. Phys., 16, 1491-1509, 10.5194/acp-16-1491-2016,
826 2016.
827 Karydis, V. A., Tsimpidi, A. P., Bacer, S., Pozzer, A., Nenes, A., and Lelieveld, J.: Global impact of mineral
828 dust on cloud droplet number concentration, Atmospheric Chemistry and Physics, 17, 5601-5621,
829 10.5194/acp-17-5601-2017, 2017.
830 Kerkweg, A., Buchholz, J., Ganzeveld, L., Pozzer, A., Tost, H., and Jöckel, P.: Technical Note: An
831 implementation of the dry removal processes DRY DEPosition and SEDImentation in the Modular Earth
832 Submodel System (MESSy), Atmos. Chem. Phys., 6, 4617-4632, 2006a.
833 Kerkweg, A., Sander, R., Tost, H., and Jöckel, P.: Technical note: Implementation of prescribed (OFFLEM),
834 calculated (ONLEM), and pseudo-emissions (TNUDGE) of chemical species in the Modular Earth Submodel
835 System (MESSy), Atmos. Chem. Phys., 6, 3603-3609, 10.5194/acp-6-3603-2006, 2006b.
836 Koo, B., Knipping, E., and Yarwood, G.: 1.5-Dimensional volatility basis set approach for modeling organic
837 aerosol in CAMx and CMAQ, Atmospheric Environment, 95, 158-164, 10.1016/j.atmosenv.2014.06.031,
838 2014.
839 Kuwata, M., and Lee, W. C.: 1-octanol-water partitioning as a classifier of water soluble organic matters:
840 Implication for solubility distribution, Aerosol Science and Technology, 51, 602-613,
841 10.1080/02786826.2017.1283004, 2017.
842 Lambe, A. T., Onasch, T. B., Massoli, P., Croasdale, D. R., Wright, J. P., Ahern, A. T., Williams, L. R., Worsnop,
843 D. R., Brune, W. H., and Davidovits, P.: Laboratory studies of the chemical composition and cloud
844 condensation nuclei (CCN) activity of secondary organic aerosol (SOA) and oxidized primary organic
845 aerosol (OPOA), Atmospheric Chemistry and Physics, 11, 8913-8928, 10.5194/acp-11-8913-2011, 2011.
846 Lane, T. E., Donahue, N. M., and Pandis, S. N.: Simulating secondary organic aerosol formation using the
847 volatility basis-set approach in a chemical transport model, Atmos. Environ., 42, 7439-7451, 2008.
848 Lauer, A., Eyring, V., Hendricks, J., Joeckel, P., and Lohmann, U.: Global model simulations of the impact
849 of ocean-going ships on aerosols, clouds, and the radiation budget, Atmos. Chem. Phys., 7, 5061-5079,
850 2007.
851 Liggio, J., Li, S. M., Vlasenko, A., Sjostedt, S., Chang, R., Shantz, N., Abbatt, J., Slowik, J. G., Bottenheim, J.
852 W., Brickell, P. C., Stroud, C., and Leaitch, W. R.: Primary and secondary organic aerosols in urban air
853 masses intercepted at a rural site, Journal of Geophysical Research-Atmospheres, 115,
854 10.1029/2010jd014426, 2010.
855 Martin, S. T., Andreae, M. O., Althausen, D., Artaxo, P., Baars, H., Borrmann, S., Chen, Q., Farmer, D. K.,
856 Guenther, A., Gunthe, S. S., Jimenez, J. L., Karl, T., Longo, K., Manzi, A., Muller, T., Pauliquevis, T., Petters,
857 M. D., Prenni, A. J., Poschl, U., Rizzo, L. V., Schneider, J., Smith, J. N., Swietlicki, E., Tota, J., Wang, J.,
858 Wiedensohler, A., and Zorn, S. R.: An overview of the Amazonian Aerosol Characterization Experiment

859   2008 (AMAZE-08), Atmospheric Chemistry and Physics, 10, 11415-11438, 10.5194/acp-10-11415-2010,
860   2010.
861   May, A. A., Levin, E. J. T., Hennigan, C. J., Riipinen, I., Lee, T., Collett, J. L., Jimenez, J. L., Kreidenweis, S. M.,
862   and Robinson, A. L.: Gas-particle partitioning of primary organic aerosol emissions: 3. Biomass burning,
863   Journal of Geophysical Research-Atmospheres, 118, 11327-11338, 10.1002/jgrd.50828, 2013.
864   Mohr, C., DeCarlo, P. F., Heringa, M. F., Chirico, R., Slowik, J. G., Richter, R., Reche, C., Alastuey, A., Querol,
865   X., Seco, R., Penuelas, J., Jimenez, J. L., Crippa, M., Zimmermann, R., Baltensperger, U., and Prevot, A. S.
866   H.: Identification and quantification of organic aerosol from cooking and other sources in Barcelona using
867   aerosol mass spectrometer data, Atmospheric Chemistry and Physics, 12, 1649-1665, 10.5194/acp-12-
868   1649-2012, 2012.
869   Murphy, B. N., and Pandis, S. N.: Simulating the formation of semivolatile primary and secondary organic
870   aerosol in a regional chemical transport model, Environ. Sci. Technol., 43, 4722-4728, 2009.
871   Murphy, B. N., Donahue, N. M., Fountoukis, C., and Pandis, S. N.: Simulating the oxygen content of
872   ambient organic aerosol with the 2D volatility basis set, Atmos. Chem. Phys., 11, 7859-7873, 2011.
873   Murphy, B. N., Donahue, N. M., Fountoukis, C., Dall'Osto, M., O'Dowd, C., Kiendler-Scharr, A., and Pandis,
874   S. N.: Functionalization and fragmentation during ambient organic aerosol aging: application of the 2-D
875   volatility basis set to field studies, Atmospheric Chemistry and Physics, 12, 10797-10816, 10.5194/acp-12-
876   10797-2012, 2012.
877   Napier, W. J., Ensberg, J. J., and Seinfeld, J. H.: Insight into the numerical challenges of implementing 2-
878   dimensional SOA models in atmospheric chemical transport models, Atmospheric Environment, 96, 331-
879   344, 10.1016/j.atmosenv.2014.07.048, 2014.
880   Ng, N. L., Canagaratna, M. R., Zhang, Q., Jimenez, J. L., Tian, J., Ulbrich, I. M., Kroll, J. H., Docherty, K. S.,
881   Chhabra, P. S., Bahreini, R., Murphy, S. M., Seinfeld, J. H., Hildebrandt, L., Donahue, N. M., DeCarlo, P. F.,
882   Lanz, V. A., Prévôt, A. S. H., Dinar, E., Rudich, Y., and Worsnop, D. R.: Organic aerosol components observed
883   in Northern Hemispheric datasets from Aerosol Mass Spectrometry, Atmos. Chem. Phys., 10, 4625-4641,
884   10.5194/acp-10-4625-2010, 2010.
885   Paciga, A., Karnezi, E., Kostenidou, E., Hildebrandt, L., Psichoudaki, M., Engelhart, G. J., Lee, B. H., Crippa,
886   M., Prevot, A. S. H., Baltensperger, U., and Pandis, S. N.: Volatility of organic aerosol and its components
887   in the megacity of Paris, Atmospheric Chemistry and Physics, 16, 2013-2023, 10.5194/acp-16-2013-2016,
888   2016.
889   Pankow, J. F., and Asher, W. E.: SIMPOL.1: a simple group contribution method for predicting vapor
890   pressures and enthalpies of vaporization of multifunctional organic compounds, Atmospheric Chemistry
891   and Physics, 8, 2773-2796, 10.5194/acp-8-2773-2008, 2008.
892   Pathak, R. K., Presto, A. A., Lane, T. E., Stanier, C. O., Donahue, N. M., and Pandis, S. N.: Ozonolysis of
893   alpha-pinene: parameterization of secondary organic aerosol mass fraction, Atmospheric Chemistry and
894   Physics, 7, 3811-3821, 2007.
895   Pozzer, A., Zimmermann, P., Doering, U. M., van Aardenne, J., Tost, H., Dentener, F., Janssens-Maenhout,
896   G., and Lelieveld, J.: Effects of business-as-usual anthropogenic emissions on air quality, Atmos. Chem.
897   Phys., 12, 6915-6937, 10.5194/acp-12-6915-2012, 2012.
898   Pringle, K. J., Tost, H., Message, S., Steil, B., Giannadaki, D., Nenes, A., Fountoukis, C., Stier, P., Vignati, E.,
899   and Leieved, J.: Description and evaluation of GMXe: a new aerosol submodel for global simulations (v1),
900   Geoscientific Model Development, 3, 391-412, 2010.
901   Pye, H. O. T., and Seinfeld, J. H.: A global perspective on aerosol from low-volatility organic compounds,
902   Atmos. Chem. Phys., 10, 4377-4401, 2010.
903   Robinson, A. L., Donahue, N. M., Shrivastava, M. K., Weitkamp, E. A., Sage, A. M., Grieshop, A. P., Lane, T.
904   E., Pierce, J. R., and Pandis, S. N.: Rethinking organic aerosols: Semivolatile emissions and photochemical
905   aging, Science, 315, 1259-1262, 2007.

Roeckner, E., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kornblueh, L., Manzini, E., Schlese, U., and Schulzweida, U.: Sensitivity of simulated climate to horizontal and vertical resolution in the ECHAM5 atmosphere model, Journal of Climate, 19, 3771-3791, 10.1175/jcli3824.1, 2006.

Saarikoski, S., Carbone, S., Decesari, S., Giulianelli, L., Angelini, F., Canagaratna, M., Ng, N. L., Trimborn, A., Facchini, M. C., Fuzzi, S., Hillamo, R., and Worsnop, D.: Chemical characterization of springtime submicrometer aerosol in Po Valley, Italy, Atmospheric Chemistry and Physics, 12, 8401-8421, 10.5194/acp-12-8401-2012, 2012.

Sander, R., Baumgaertner, A., Gromov, S., Harder, H., Joeckel, P., Kerkweg, A., Kubistin, D., Regelin, E., Riede, H., Sandu, A., Taraborrelli, D., Tost, H., and Xie, Z. Q.: The atmospheric chemistry box model CAABA/MECCA-3.0, Geoscientific Model Development, 4, 373-380, 2011.

Shiraiwa, M., Berkemeier, T., Schilling-Fahnestock, K. A., Seinfeld, J. H., and Poschl, U.: Molecular corridors and kinetic regimes in the multiphase chemical evolution of secondary organic aerosol, Atmospheric Chemistry and Physics, 14, 8323-8341, 10.5194/acp-14-8323-2014, 2014.

Shiraiwa, M., Li, Y., Tsimpidi, A. P., Karydis, V. A., Berkemeier, T., Pandis, S. N., Lelieveld, J., Koop, T., and Poschl, U.: Global distribution of particle phase state in atmospheric secondary organic aerosols, Nature Communications, 8, 10.1038/ncomms15002, 2017.

Sun, Y., Zhang, Q., Macdonald, A. M., Hayden, K., Li, S. M., Liggio, J., Liu, P. S. K., Anlauf, K. G., Leaitch, W. R., Steffen, A., Cubison, M., Worsnop, D. R., van Donkelaar, A., and Martin, R. V.: Size-resolved aerosol chemistry on Whistler Mountain, Canada with a high-resolution aerosol mass spectrometer during INTEX-B, Atmospheric Chemistry and Physics, 9, 3095-3111, 10.5194/acp-9-3095-2009, 2009.

Sun, Y. L., Zhang, Q., Schwab, J. J., Demerjian, K. L., Chen, W. N., Bae, M. S., Hung, H. M., Hogrefe, O., Frank, B., Rattigan, O. V., and Lin, Y. C.: Characterization of the sources and processes of organic and inorganic aerosols in New York city with a high-resolution time-of-flight aerosol mass apectrometer, Atmospheric Chemistry and Physics, 11, 1581-1602, 10.5194/acp-11-1581-2011, 2011.

Thomson, A. M., Calvin, K. V., Smith, S. J., Kyle, G. P., Volke, A., Patel, P., Delgado-Arias, S., Bond-Lamberty, B., Wise, M. A., Clarke, L. E., and Edmonds, J. A.: RCP4.5: a pathway for stabilization of radiative forcing by 2100, Climatic Change, 109, 77-94, 2011.

Timonen, H., Carbone, S., Aurela, M., Saarnio, K., Saarikoski, S., Ng, N. L., Canagaratna, M. R., Kulmala, M., Kerminen, V. M., Worsnop, D. R., and Hillamo, R.: Characteristics, sources and water-solubility of ambient submicron organic aerosol in springtime in Helsinki, Finland, Journal of Aerosol Science, 56, 61-77, 10.1016/j.jaerosci.2012.06.005, 2013.

Tost, H., Jockel, P. J., Kerkweg, A., Sander, R., and Lelieveld, J.: Technical note: A new comprehensive SCAVenging submodel for global atmospheric chemistry modelling, Atmos. Chem. Phys., 6, 565-574, 2006.

Trail, M., Tsimpidi, A. P., Liu, P., Tsigaridis, K., Hu, Y., Nenes, A., and Russell, A. G.: Downscaling a global climate model to simulate climate change over the US and the implication on regional and urban air quality, Geoscientific Model Development, 6, 1429-1445, 10.5194/gmd-6-1429-2013, 2013.

Trail, M., Tsimpidi, A. P., Liu, P., Tsigaridis, K., Rudokas, J., Miller, P., Nenes, A., Hu, Y., and Russell, A. G.: Sensitivity of air quality to potential future climate change and emissions in the United States and major cities, Atmospheric Environment, 94, 552-563, 10.1016/j.atmosenv.2014.05.079, 2014.

Tsigaridis, K., Daskalakis, N., Kanakidou, M., Adams, P. J., Artaxo, P., Bahadur, R., Balkanski, Y., Bauer, S. E., Bellouin, N., Benedetti, A., Bergman, T., Berntsen, T. K., Beukes, J. P., Bian, H., Carslaw, K. S., Chin, M., Curci, G., Diehl, T., Easter, R. C., Ghan, S. J., Gong, S. L., Hodzic, A., Hoyle, C. R., Iversen, T., Jathar, S., Jimenez, J. L., Kaiser, J. W., Kirkevag, A., Koch, D., Kokkola, H., Lee, Y. H., Lin, G., Liu, X., Luo, G., Ma, X., Mann, G. W., Mihalopoulos, N., Morcrette, J. J., Mueller, J. F., Myhre, G., Myriokefalitakis, S., Ng, N. L., O'Donnell, D., Penner, J. E., Pozzoli, L., Pringle, K. J., Russell, L. M., Schulz, M., Sciare, J., Seland, O., Shindell, D. T., Sillman, S., Skeie, R. B., Spracklen, D., Stavrakou, T., Steenrod, S. D., Takemura, T., Tiitta, P., Tilmes, S., Tost, H., van Noije, T., van Zyl, P. G., von Salzen, K., Yu, F., Wang, Z., Wang, Z., Zaveri, R. A., Zhang, H.,

Zhang, K., Zhang, Q., and Zhang, X.: The AeroCom evaluation and intercomparison of organic aerosol in global models, Atmospheric Chemistry and Physics, 14, 10845-10895, 10.5194/acp-14-10845-2014, 2014.

Tsimpidi, A. P., Karydis, V. A., Zavala, M., Lei, W., Molina, L., Ulbrich, I. M., Jimenez, J. L., and Pandis, S. N.: Evaluation of the volatility basis-set approach for the simulation of organic aerosol formation in the Mexico City metropolitan area, Atmos. Chem. Phys., 10, 525-546, 2010.

Tsimpidi, A. P., Karydis, V. A., Zavala, M., Lei, W., Bei, N., Molina, L., and Pandis, S. N.: Sources and production of organic aerosol in Mexico City: insights from the combination of a chemical transport model (PMCAMx-2008) and measurements during MILAGRO, Atmos. Chem. Phys., 11, 5153-5168, 2011.

Tsimpidi, A. P., Karydis, V. A., Pozzer, A., Pandis, S. N., and Lelieveld, J.: ORACLE (v1.0): module to simulate the organic aerosol composition and evolution in the atmosphere, Geoscientific Model Development, 7, 3153-3172, 10.5194/gmd-7-3153-2014, 2014.

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.

van der Gon, H., Bergstrom, R., Fountoukis, C., Johansson, C., Pandis, S. N., Simpson, D., and Visschedijk, A. J. H.: Particulate emissions from residential wood combustion in Europe revised estimates and an evaluation, Atmospheric Chemistry and Physics, 15, 6503-6519, 10.5194/acp-15-6503-2015, 2015.

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., Morton, D. C., DeFries, R. S., Jin, Y., and van Leeuwen, T. T.: Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997-2009), Atmos. Chem. Phys., 10, 11707-11735, 2010.

Xu, L., Guo, H. Y., Boyd, C. M., Klein, M., Bougiatioti, A., Cerully, K. M., Hite, J. R., Isaacman-VanWertz, G., Kreisberg, N. M., Knote, C., Olson, K., Koss, A., Goldstein, A. H., Hering, S. V., de Gouw, J., Baumann, K., Lee, S. H., Nenes, A., Weber, R. J., and Ng, N. L.: Effects of anthropogenic emissions on aerosol formation from isoprene and monoterpenes in the southeastern United States, Proceedings of the National Academy of Sciences of the United States of America, 112, 37-42, 10.1073/pnas.1417609112, 2015.

Zhang, Q., Jimenez, J. L., Canagaratna, M. R., Allan, J. D., Coe, H., Ulbrich, I., Alfarra, M. R., Takami, A., Middlebrook, A. M., Sun, Y. L., Dzepina, K., Dunlea, E., Docherty, K., DeCarlo, P. F., Salcedo, D., Onasch, T., Jayne, J. T., Miyoshi, T., Shimono, A., Hatakeyama, S., Takegawa, N., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Williams, P., Bower, K., Bahreini, R., Cottrell, L., Griffin, R. J., Rautiainen, J., Sun, J. Y., Zhang, Y. M., and Worsnop, D. R.: Ubiquity and dominance of oxygenated species in organic aerosols in anthropogenically-influenced Northern Hemisphere midlatitudes, Geophys. Res. Lett., 34, doi: L13801 10.1029/2007gl029979, 2007.

Zhao, B., Wang, S. X., Donahue, N. M., Chuang, W. N., Hildebrandt Ruiz, L., Ng, N. L., Wang, Y. J., and Hao, J. M.: Evaluation of One-Dimensional and Two-Dimensional Volatility Basis Sets in Simulating the Aging of Secondary Organic Aerosol with Smog-Chamber Experiments, Environmental Science & Technology, 49, 2245-2254, 10.1021/es5048914, 2015.

Table 1. Statistical evaluation of EMAC POA ~~(sum of fPOA and bbPOA)~~ against AMS POA (sum

995   of HOA and BBOA) from 61 data sets worldwide during 2001-2010.

996

Site Type	Number of datasets	Mean Observed (μg m^{-3})	Mean Predicted (μg m^{-3})	MAGE (μg m^{-3})	MB (μg m^{-3})	NME (%)	NMB (%)	RMSE (μg m^{-3})
Urban Downwind	15	0.82	0.65	0.39	-0.17	47	-21	0.51
Rural/Remote	46	0.43	0.45	0.38	0.02	88	6	0.50
Season								
Winter	6	1.18	0.72	0.66	-0.45	56	-38	0.78
Spring	30	0.42	0.52	0.40	0.10	95	24	0.50
Summer	14	0.50	0.45	0.30	-0.05	60	-12	0.40
Autumn	11	0.49	0.41	0.28	-0.08	57	-18	0.38
Total	61	0.53	0.50	0.38	-0.03	72	-5	0.50

997

998

999 **Table 2.** Statistical evaluation of EMAC SOA against AMS OOA from 61 data sets worldwide

1000 during 2001-2010.

1001

Site Type	Number of datasets	Mean Observed ($\mu g\ m^{-3}$)	Mean Predicted ($\mu g\ m^{-3}$)	MAGE ($\mu g\ m^{-3}$)	MB ($\mu g\ m^{-3}$)	NME (%)	NMB (%)	RMSE ($\mu g\ m^{-3}$)
Urban Downwind	15	2.98	2.76	1.32	-0.22	44	-7	1.82
Rural/Remote	46	2.72	2.53	1.52	-0.18	56	-7	2.11
Season								
Winter	6	2.81	0.67	2.14	-2.14	76	-76	2.52
Spring	30	2.22	2.48	1.05	0.25	47	11	1.44
Summer	14	4.30	3.85	2.23	-0.45	52	-10	2.99
Autumn	11	2.35	2.34	1.29	-0.01	55	0	1.60
Total	61	2.78	2.59	1.47	-0.19	53	-7	2.04

1002
1003
1004

1005   **Table 3.** Statistical evaluation of EMAC aged SOA against AMS LVOOA from 41 data sets

1006   worldwide during 2001-2010.

1007

Site Type	Number of datasets	Mean Observed ($\mu g\ m^{-3}$)	Mean Predicted ($\mu g\ m^{-3}$)	MAGE ($\mu g\ m^{-3}$)	MB ($\mu g\ m^{-3}$)	NME (%)	NMB (%)	RMSE ($\mu g\ m^{-3}$)
Urban Downwind	8	1.77	0.81	1.26	-0.95	71	-54	1.57
Rural/Remote	33	1.65	0.95	1.18	-0.70	72	-42	1.71
Season								
Winter	3	2.36	0.17	2.19	-2.19	93	-93	2.40
Spring	18	1.06	0.73	0.84	-0.33	79	-32	1.05
Summer	11	2.64	1.46	1.74	-1.18	66	-45	2.47
Autumn	9	1.49	0.92	0.92	-0.57	62	-38	1.14
Total	41	1.68	0.93	1.20	-0.75	72	-45	1.68

1008

1009

1010 **Table 4.** Statistical evaluation of EMAC fresh SOA against AMS SVOOA from 41 data sets

1011 worldwide during 2001-2010.

1012

Site Type	Number of datasets	Mean Observed ($\mu g\ m^{-3}$)	Mean Predicted ($\mu g\ m^{-3}$)	MAGE ($\mu g\ m^{-3}$)	MB ($\mu g\ m^{-3}$)	NME (%)	NMB (%)	RMSE ($\mu g\ m^{-3}$)
Urban Downwind	8	0.81	1.22	0.75	0.41	92	50	0.98
Rural/Remote	33	1.03	1.35	0.82	0.32	79	31	1.03
Season								
Winter	3	0.87	0.36	0.51	-0.51	59	-59	0.63
Spring	18	0.54	0.88	0.66	0.34	122	64	0.91
Summer	11	1.89	2.22	1.08	0.33	57	17	1.19
Autumn	9	0.83	1.44	0.86	0.61	104	73	1.11
Total	41	0.99	1.32	0.80	0.33	81	34	1.02

1013

1014

Table 5. Worldwide O:C ratio predictions and observations of total OA.

Location	Lat	Lon	Time period	Measured O:C	Simulated O:C	Reference
Riverside, US	33.95	-117.4	14/07/05-13/08/05	0.35	0.57	(Docherty et al. (2011))
Mexico City , MX	19.48	-99.15	10/03/06-30/03/06	0.41	0.55	(Aiken et al. (2008))
Whistler, CA	50.01	-122.95	19/04/06-16/05/06	0.83	0.57	(Sun et al. (2009))
Egbert, CA	44.23	-79.78	11/05/07-15/06/07	0.45	0.53	(Liggio et al. (2010))
London, UK	51.53	-0.15	17/10/07-15/11/07	0.27	0.29	(Allan et al. (2010))
Manaus, BR	-2.58	-60.2	07/02/08-14/03/08	0.44	0.50	(Martin et al. (2010))
Po Valley, IT	44.65	11.62	30/03/08-20/04/08	0.47	0.56	(Saarikoski et al. (2012))
Mace Head, IE	53.3	-9.8	16/05/08-13/06/08	0.53	0.65	(Murphy et al. (2012))
Cabauw, NL	51.97	4.93	28/04/08-30/05/08	0.54	0.53	Murphy et al. (2012)
Beijing, CN	40.0	116.0	24/07/08-20/09/08	0.33	0.49	Huang et al. (2010)
West coast, CL	-20	-80	21/10/08-30/11/08	0.62	0.89	(Hawkins et al. (2010))
Kaiping, CN	22.32	112.53	12/10/08-18/11/08	0.47	0.41	(Huang et al. (2011))
Helsinki, FI	60.2	24.95	09/01/09-13/03/09	0.45	0.25	(Carbone et al. (2014))
Barcelona, ES	41.39	2.12	25/02/09-26/03/09	0.36	0.50	(Mohr et al. (2012))
Mace Head, IE	53.3	-9.8	26/02/09-26/03/09	0.58	0.65	Murphy et al. (2012)
Cabauw, NL	51.97	4.93	25/02/09-25/02/09	0.34	0.32	Murphy et al. (2012)
Helsinki, FI	60.2	24.95	09/04/09-08/05/09	0.51	0.46	(Timonen et al. (2013))
Pasadena, US	34.14	-118.12	22/05/09-12/06/09	0.44	0.51	(Guo et al. (2017))
Pasadena, US	34.14	-118.12	18/06/09-07/07/09	0.55	0.56	(Guo et al. (2017))
New York, US	40.74	-73.92	13/07/09-03/08/09	0.36	0.55	(Sun et al. (2011))
Pasadena, US	34.14	-118.12	10/07/09-04/08/09	0.48	0.55	(Guo et al. (2017))
Paris, FR	48.83	2.36	01/07/09-31/07/09	0.38	0.54	(Crippa et al. (2013))
Alps, AT	47.05	12.95	10/07/09-31/07/09	0.38	0.58	(Holzinger et al. (2010))
Shenzhen, CN	22.6	113.9	25/10/09-02/12/09	0.30	0.41	(He et al. (2011))
Fresno, US	36.81	-119.78	09/01/10-23/01/10	0.27	0.46	(Ge et al. (2012))
Cool, US	38.88	-121.0	02/06/10-28/06/10	0.44	0.51	(Xu et al. (2015))
Jiaxing, CN	30.8	120.8	29/06/10-15/07/10	0.28	0.42	(Huang et al. (2013))
Guangzhou, CN	22.71	112.93	13/11/10-01/12/10	0.40	0.42	(Gong et al. (2012))
Beijing, CN	40.0	116.0	22/11/10-22/12/10	0.32	0.20	(Hu et al. (2016))
Jiaxing, CN	30.8	120.8	11/12/10-23/12/10	0.33	0.24	(Huang et al. (2013))

1015

1016

1017 **Table 6.** Worldwide O:C ratio predictions and observations of OOA.

Location	Lat	Lon	Time period	Measured O:C	Simulate O:C	Reference
Cheju Island, KR	33.51	126.50	11/04/01-30/04/01	0.84	0.64	(Ng et al. (2010))
Manchester, UK	53.5	-2.22	14/06/01-25/06/01	0.81	0.72	Ng et al. (2010)
New York, US	40.74	-73.92	30/06/01-05/08/01	0.59	0.64	Ng et al. (2010)
Vancouver, CA	49.25	-123.13	11/08/01-24/08/01	0.56	0.59	Ng et al. (2010)
Manchester, UK	53.5	-2.22	17/01/02-28/01/02	0.47	0.61	(Ng et al. (2010))
Hohenpeissenberg, DE	47.8	11.0	19/05/02-31/05/02	0.50	0.61	(Ng et al. (2010))
East coast, US	37.95	-74.3	18/07/02-26/07/02	0.58	0.67	(Ng et al. (2010))
Jungfraujoch , CH	46.3	7.6	28/06/02-17/07/02	0.70	0.63	(Ng et al. (2010))
East coast, US	37.95	-74.3	29/07/02-10/08/02	0.60	0.70	(Ng et al. (2010))
Pittsburgh, US	40.44	-79.94	06/09/02-22/09/02	0.51	0.60	(Ng et al. (2010))
Fukue Island , JP	32.69	128.84	18/03/03-16/04/03	0.65	0.67	(Ng et al. (2010))
Hyytiala , FI	61.8	24.3	31/03/05-15/04/05	0.47	0.64	(Ng et al. (2010))
Boulder, US	40.02	-105.27	07/06/03-20/06/03	0.44	0.60	(Ng et al. (2010))
Tokyo , JP	35.67	139.75	23/07/03-14/08/03	0.49	0.70	(Ng et al. (2010))
NE London, UK	51.7	0.4	29/07/03-31/08/03	0.48	0.72	(Ng et al. (2010))
Okinawa , JP	26.87	128.25	03/10/03-24/12/03	0.82	0.77	(Ng et al. (2010))
Tokyo , JP	35.67	139.75	20/01/04-10/02/04	0.56	0.59	(Ng et al. (2010))
New York, US	40.74	-73.92	07/01/04-06/02/04	0.45	0.58	(Ng et al. (2010))
Norfolk coast, UK	53.0	1.1	25/04/04-26/05/04	0.70	0.69	Ng et al. (2010)
Wiesbaden, DE	50.22	8.45	14/07/04-04/08/04	0.57	0.63	Ng et al. (2010)
Pinnacle Park, US	43.0	-76	14/07/04-05/08/04	0.64	0.61	Ng et al. (2010)
Nova Scotia, CA	43.76	-66.1	07/07/04-14/08/04	0.65	0.68	Ng et al. (2010)
Mainz, DE	49.98	8.23	16/09/04-01/10/04	0.66	0.62	Ng et al. (2010)
Duke Forest, US	35.97	-79.1	13/09/04-21/09/04	0.46	0.60	Ng et al. (2010)
Roveredo, CH	46.23	9.12	01/03/05-15/03/15	0.43	0.58	Ng et al. (2010)
Harkingen, CH	47.32	7.82	12/05/05-30/05/15	0.55	0.63	Ng et al. (2010)
Riverside, US	33.95	-117.4	14/07/05-13/08/05	0.48	0.62	Ng et al. (2010)
Zurich, CH	47.4	8.5	14/07/05-04/08/05	0.51	0.65	Ng et al. (2010)
Thompson, US	43.11	-70.95	09/07/05-15/08/05	0.58	0.62	Ng et al. (2010)
Roveredo, CH	46.23	9.12	25/11/05-15/12/05	0.56	0.56	Ng et al. (2010)
Zurich, CH	47.4	8.5	06/01/06-25/01/06	0.65	0.54	Ng et al. (2010)
Reiden, CH	47.25	7.97	27/01/06-13/02/06	0.64	0.55	Ng et al. (2010)
Mexico City , MX	19.48	-99.15	10/03/06-30/03/06	0.53	0.69	Ng et al. (2010)
Payerne, CH	46.8	6.95	31/05/06-03/07/06	0.59	0.63	Ng et al. (2010)
Beijing, CN	40.0	116.0	09/07/06-21/07/06	0.50	0.65	Ng et al. (2010)
Massongex, CH	46.24	6.14	23/11/06-17/12/06	0.88	0.55	Ng et al. (2010)
Payerne, CH	46.8	6.95	12/01/07-17/02/07	0.50	0.61	Ng et al. (2010)
Rhine valley, CH	46.29	9	16/02/07-22/02/07	0.34	0.57	Ng et al. (2010)
Egbert, CA	44.23	-79.78	14/05/07-15/06/07	0.61	0.64	Ng et al. (2010)
Po Valley, IT	44.65	11.62	30/03/08-20/04/08	0.65	0.65	(Saarikoski et al. (2012))
Finokalia , GR	35.33	25.66	08/05/08-04/06/08	0.80	0.73	Murphy et al. (2012)
Jungfraujoch , CH	46.3	7.6	01/05/08-29/05/08	0.70	0.65	Ng et al. (2010)
Beijing, CN	40.0	116.0	24/07/08-20/10/08	0.48	0.63	(Huang et al. (2010))
Kaiping, CN	22.32	112.53	12/10/08-18/11/08	0.52	0.57	(Huang et al. (2011))
Grenoble, FR	45.18	5.73	14/01/09-30/01/09	0.65	0.53	Ng et al. (2010)
Helsinki, FI	60.2	24.95	09/01/09-13/03/09	0.60	0.54	(Carbone et al. (2014))
Barcelona, ES	41.39	2.12	25/02/09-26/03/09	0.58	0.61	(Mohr et al. (2012))
Finokalia , GR	35.33	25.66	25/02/09-25/03/09	0.50	0.67	Murphy et al. (2012)
Helsinki, FI	60.2	24.95	09/04/09-08/05/09	0.58	0.63	(Timonen et al. (2013))
New York, US	40.74	-73.92	13/07/09-03/08/09	0.50	0.64	(Sun et al. (2011))
Shenzhen, CN	22.6	113.9	25/10/09-02/12/09	0.51	0.58	(He et al. (2011))
Fresno, US	36.81	-119.78	09/01/10-23/01/10	0.42	0.62	(Ge et al. (2012))
Pasadena, US	34.14	-118.12	15/05/10-16/06/10	0.58	0.61	(Hayes et al. (2013))
Jiaxing, CN	30.8	120.8	29/06/10-15/07/10	0.41	0.60	(Huang et al. (2013))
Guangzhou, CN	22.71	112.93	13/11/10-01/12/10	0.47	0.58	(Gong et al. (2012))
Beijing, CN	40.0	116.0	22/11/10-22/12/10	0.53	0.57	(Hu et al. (2016))
Jiaxing, CN	30.8	120.8	11/12/10-23/12/10	0.59	0.50	(Huang et al. (2013))

1018 ________
1019 ____________
1020

1021

1022   **Table 7.** Statistical evaluation of EMAC O:C ratio of total OA against observations from 18 data

1023   sets worldwide during 2001-2010.

Season	Number of datasets	Mean Observed ($\mu g\ m^{-3}$)	Mean Predicted ($\mu g\ m^{-3}$)	MAGE ($\mu g\ m^{-3}$)	MB ($\mu g\ m^{-3}$)	NME (%)	NMB (%)	RMSE ($\mu g\ m^{-3}$)
Winter	4	0.42	0.35	0.10	-0.07	24	-17	0.12
Spring	8	0.53	0.53	0.09	0.00	16	1	0.11
Summer	4	0.37	0.51	0.14	0.14	38	38	0.15
Autumn	2	0.51	0.65	0.14	0.14	28	28	0.19
Total	18	0.47	0.50	0.11	0.03	23	7	0.13

1024

1025   **Table 8.** Statistical evaluation of EMAC O:C ratio of OOA against observations from ~~18~~32 data
1026   sets worldwide during 2001-2010.

Season	Number of datasets	Mean Observed (μg m^{-3})	Mean Predicted (μg m^{-3})	MAGE (μg m^{-3})	MB (μg m^{-3})	NME (%)	NMB (%)	RMSE (μg m^{-3})
Winter	7	0.58	0.56	0.14	-0.02	23	-4	0.17
Spring	12	0.63	0.65	0.08	0.02	13	4	0.10
Summer	10	0.58	0.65	0.09	0.07	15	12	0.11
Autumn	3	0.58	0.65	0.10	0.07	17	11	0.10
Total	32	0.60	0.63	0.10	0.03	16	5	0.12

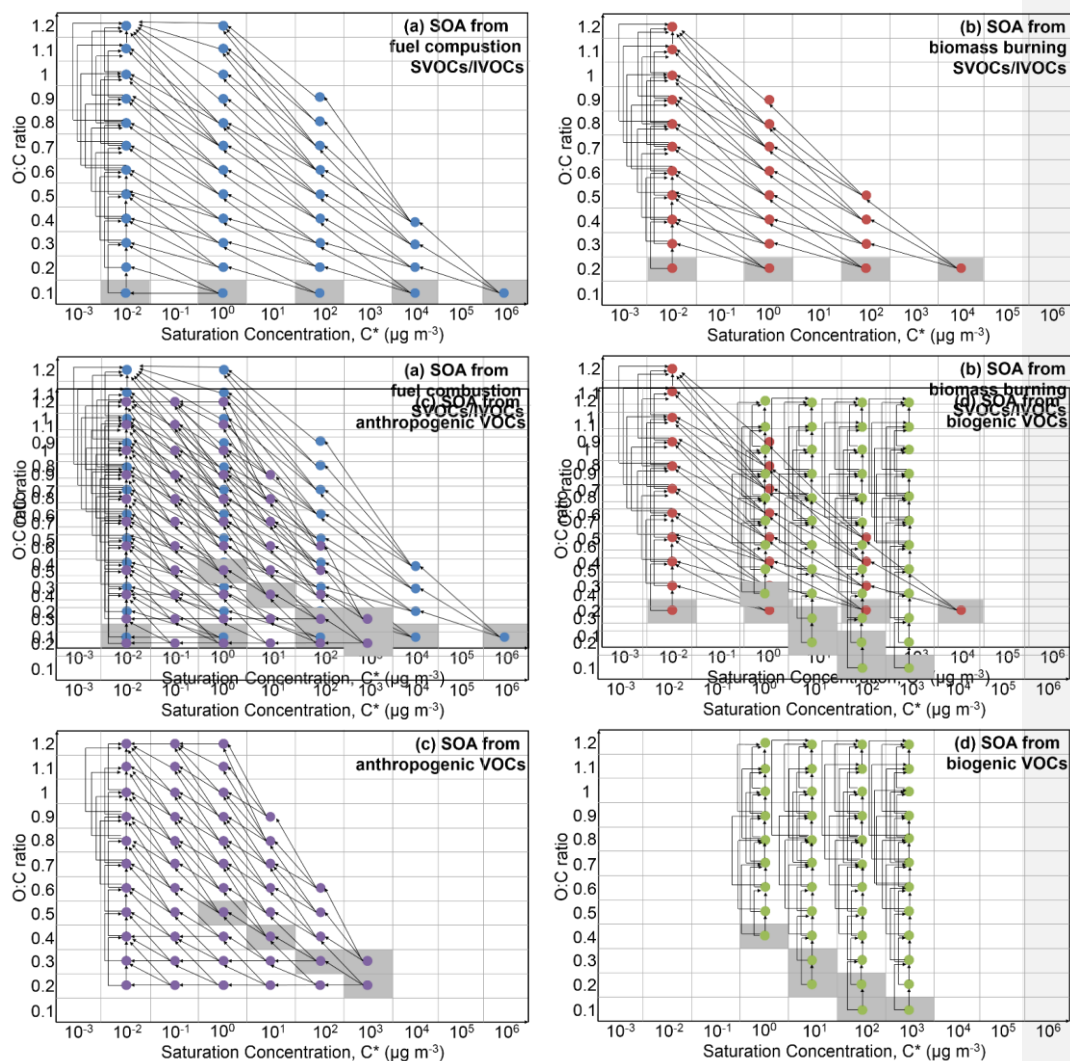
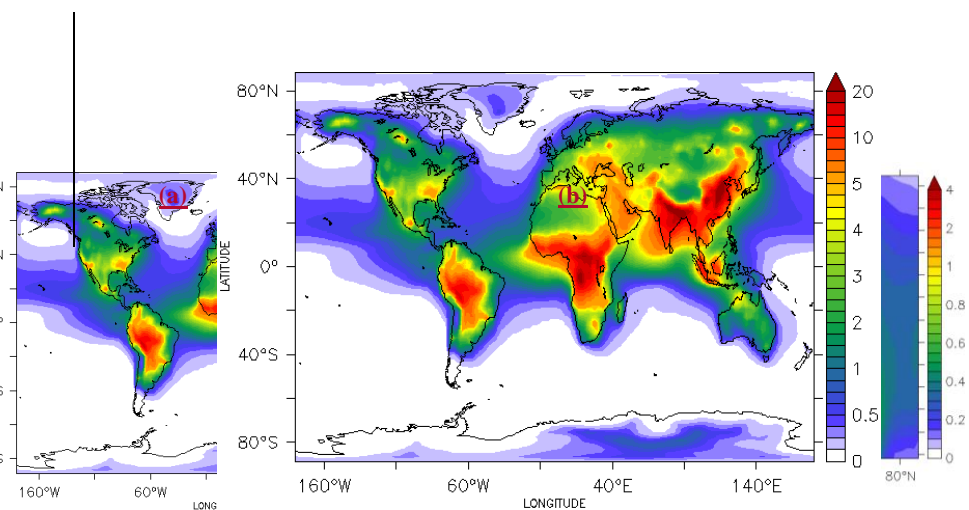
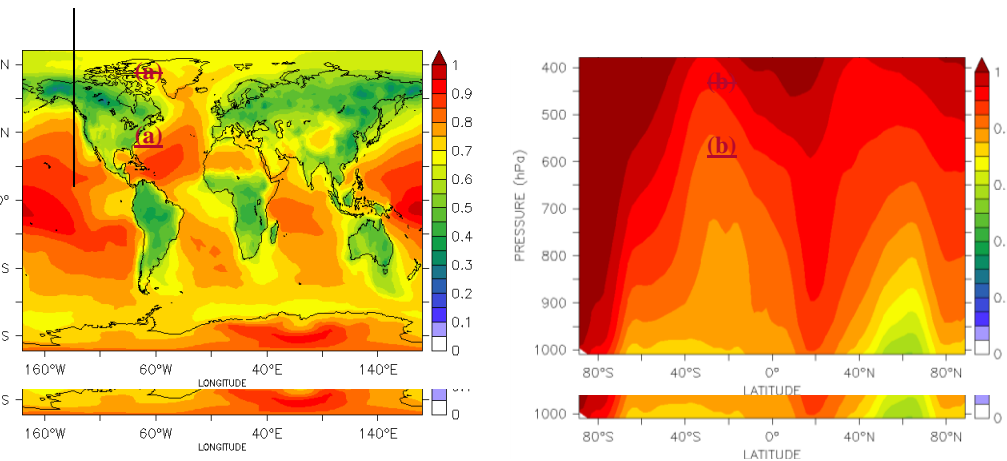
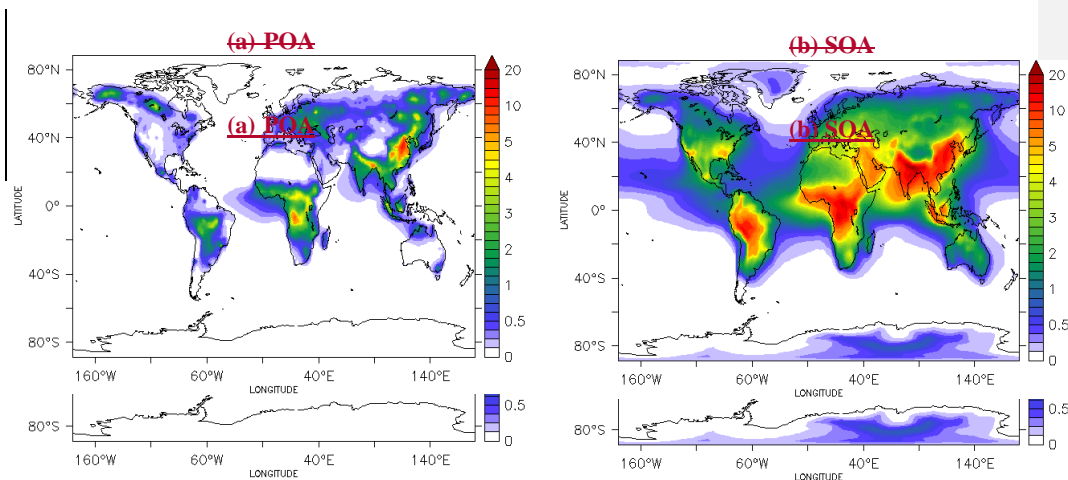



Figure 1: The 2-D grid space in ORACLE 2-D module with saturation concentration (in µg m^{-3}) on the x-axis and the O:C ratio on the y-axis. For each cell an organic surrogate compound is defined with a specific carbon number calculated as a function of effective saturation concentration and O:C ratio. The formation and evolution of SOA from: **(a)** fuel combustion SVOCs and IVOCs, **(b)** biomass burning SVOCs and IVOCs, **(c)** anthropogenic VOCs and, **(d)** biogenic VOCs are shown. The arrows correspond to the aging reactions and the grey grids to the initial chemical state of the species prior to aging.

1034 **Figure 2:** Predicted average **(a)** surface ~~concentration~~and **(b)** zonal concentrations of total OA (in

1035 µg m^{-3}) during the years 2001-2010.

1036
1037
1038
1039
1040

1041

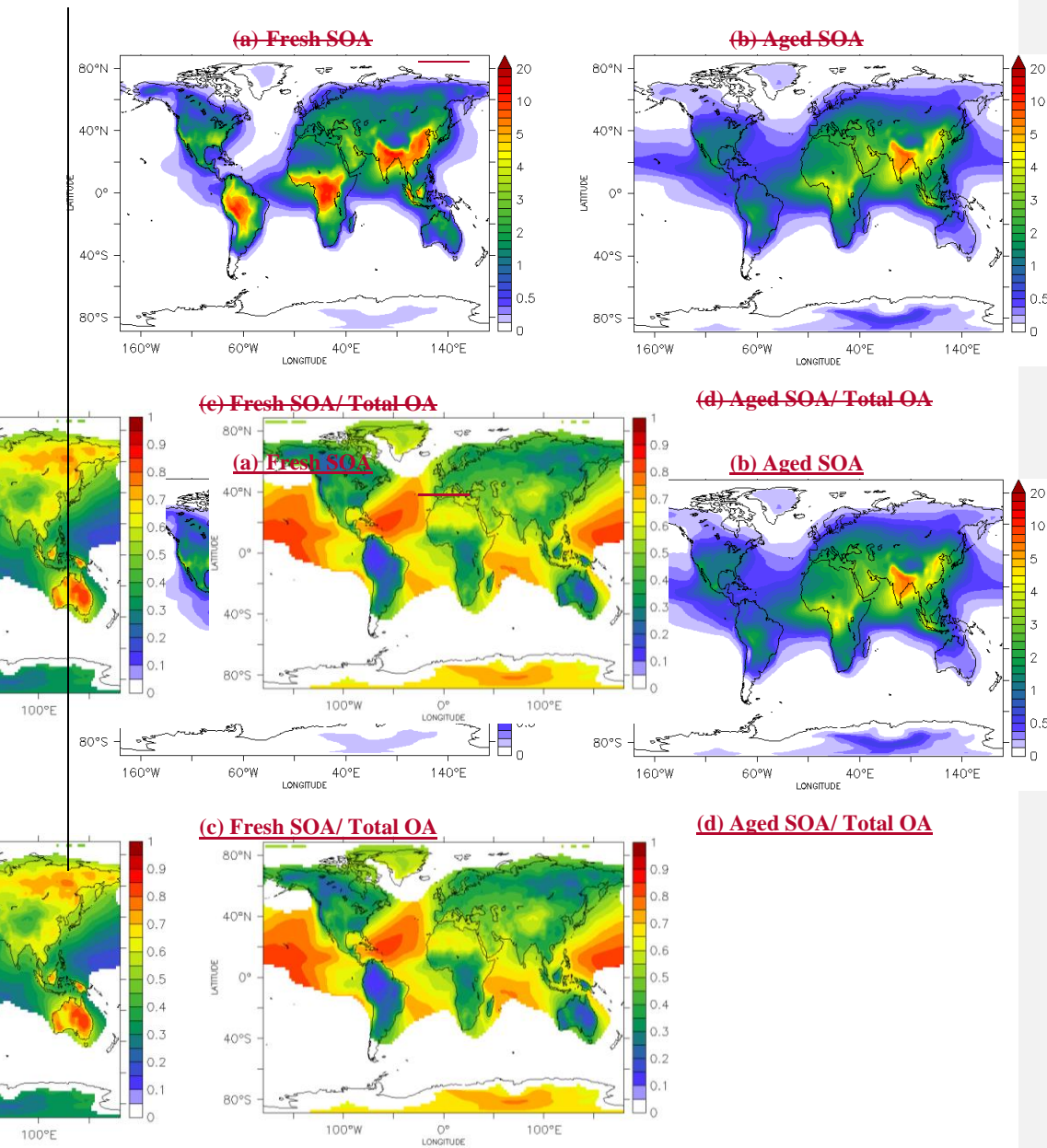
1042 **Figure 3:** Predicted (a) average **(a)** surface and **(b)** average zonal concentrations O:C of total

1043 O:COA during the years 2001-2010.

1044   **Figure 4:** Predicted average surface concentration (in μg m⁻³) of **(a)** POA (μg m⁻³) and **(b)** SOA

1045   (μg m⁻³) during the years 2001-2010.

(a) Fresh SOA

(b) Aged SOA

(c) Fresh SOA/ Total OA

(d) Aged SOA/ Total OA

1046

1047 **Figure 5:** Predicted average surface concentration (in μg m^{-3}) of **(a)** fresh SOA and **(b)** aged SOA

1048 and surface fraction of **(c)** fresh SOA and **(d)** aged SOA to total OA during the years 2001-2010.

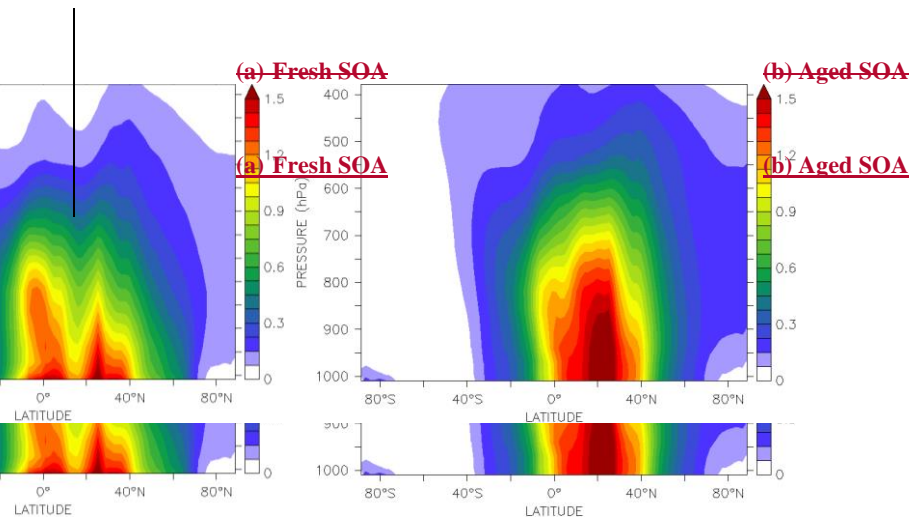



Figure 6: Predicted average zonal concentration (in μg m^{-3}) of **(a)** fresh SOA and **(b)** aged SOA during the years 2001-2010.

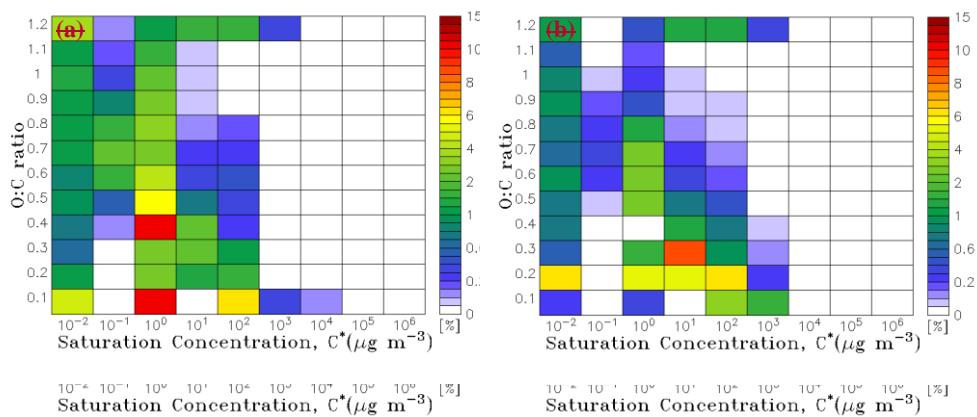



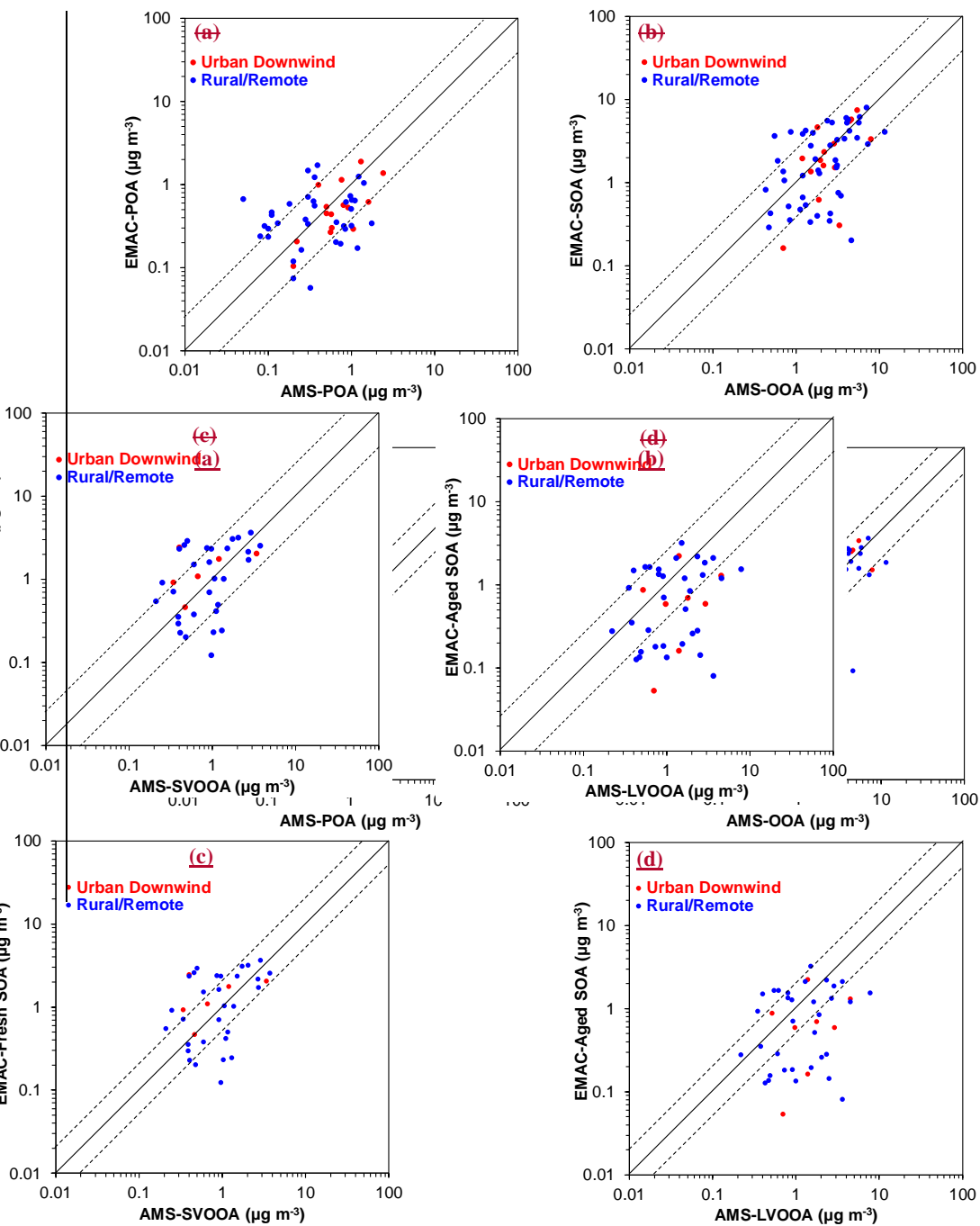
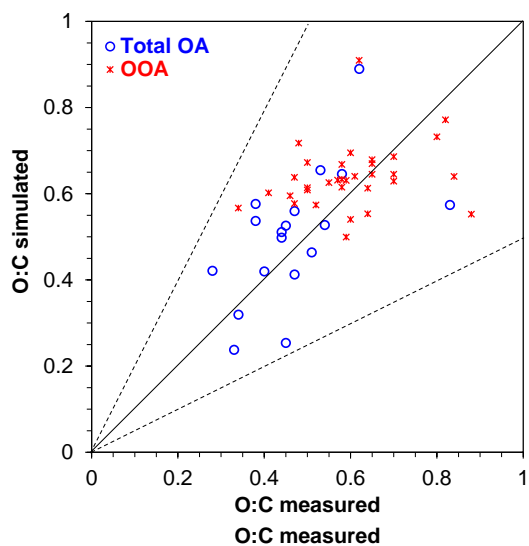
Figure 7: ~~The~~Predicted average ~~predicted~~ fraction of OA concentration in each cell of the ORACLE 2-D grid space to total OA concentration over **(a)** Europe and **(b)** the Amazon basin during the years 2001-2010.

1054

Figure 8: Scatterplots comparing model predictions to AMS measurements and their PMF
analysis for: **(a)** POA, **(b)** OOA, **(c)** ~~SV-OOA~~SVOOA, and **(d)** ~~LV-OOA~~LVOOA concentrations
(in μg m^{-3}) in the Northern Hemisphere during 2001-2010. Each point represents the data set
average value over urban downwind (in red) and rural/remote (in blue) sites. Also shown are the
1:1, 2:1, and 1:2 lines.

1060

1061 **Figure 9:** Scatterplot comparing model predictions to measurements for O:C ratio of total OA (in
1062 blue) and OOA (in red) over non-urban areas in the Northern Hemisphere during 2001-2010. Each
1063 point represents the data set average value over a specific measurement site. Also shown are the
1064 1:1, 2:1, and 1:2 lines.
1065
1066