Supplementary Information for: "Constraining nucleation, condensation, and chemistry in oxidation flow reactors using size-distribution measurements and aerosol microphysical modelling"

Anna L. Hodshire[1], Brett B. Palm[2,a], M. Lizabeth Alexander[3], Qijing Bian[1], Pedro Campuzano-Jost[2], Eben S. Cross[4,b], Douglas A. Day[2], Suzane S. de Sá[5], Alex B. Guenther[6,7], Armin Hansel[8], James F. Hunter[4], Werner Jud[8,c], Thomas Karl[9], Saewung Kim[6], Jesse H. Kroll[3,10], Jeong-Hoo Park[11,d], Zhe Peng[2], Roger Seco[6], James N. Smith[12], Jose L. Jimenez[2], Jeffrey R. Pierce[1]

[1]Department of Atmospheric Science, Colorado State University, Fort Collins, CO, 80523 USA

[2]Dept. of Chemistry and Cooperative Institute for Research in Environmental Sciences (CIRES); University of Colorado, Boulder, CO, 80309, USA

[3]Environmental and Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland,WA, 99352, USA

[4]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA

[5]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, 02138, USA

[6]Department of Earth System Science, University of California, Irvine, Irvine, CA, 92697, USA

[7]Division of Atmospheric Sciences & Global Change, Pacific Northwest National Laboratory, Richland,WA, 99352, USA

[8]Institute of Ion and Applied Physics, University of Innsbruck, Innsbruck, 6020, Austria

[9]Institute for Atmospheric and Cryospheric Sciences, University of Innsbruck, Innsbruck, 6020, Austria

[10]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA

[11]National Center for Atmospheric Research, Boulder, CO, 80305, USA

[12]Department of Chemistry, University of California, Irvine, CA, 92697, USA

[a]Now at Department of Atmospheric Sciences, University of Washington, Seattle, WA, 98195, USA

[b]Now at Center for Aerosol and Cloud Chemistry, Aerodyne Research, Inc., Billerica, MA, 01821, USA

[c]Now at Institute of Biochemical Plant Pathology, Research Unit Environmental Simulation, Helmholtz Zentrum München, Munich, 85764, Germany

[d]Now at Climate and Air Quality Research Department, National Institute of Environmental Research (NIER), Incheon, 22689, Republic of Korea

Correspondence to: hodshire@rams.colostate.edu

S1. GoAmazon2014/5 SMPS-derived volume exposures and mean diurnal cycle of monoterpenes and isoprene

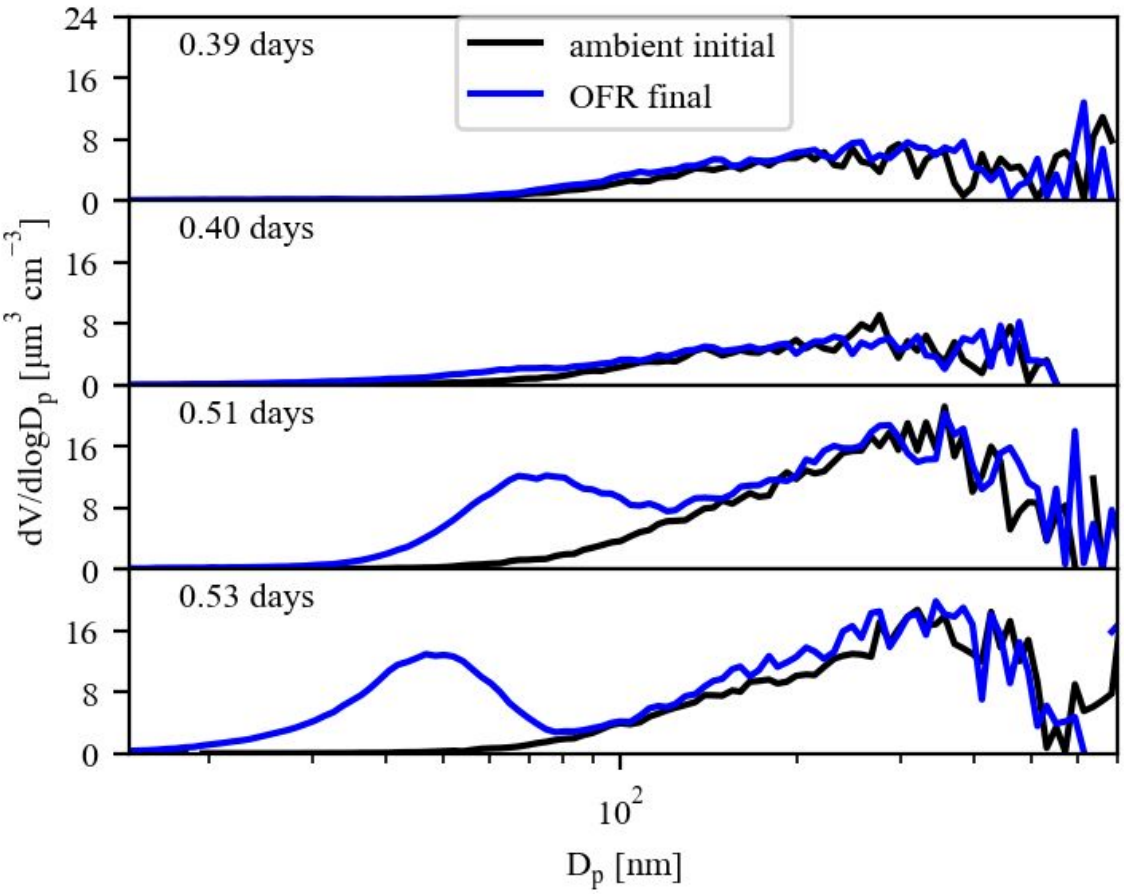

Figure S1. GoAmazon initial (i.e. ambient air, black line) and final (i.e. after OFR processing, blue line) SMPS-derived volume distributions for each individual exposure modelled in this study. The differences in SOA production between exposures of similar ages are due to the fact that the exposures were taken from different times during the campaign and thus different precursor concentrations were present (Table 2).

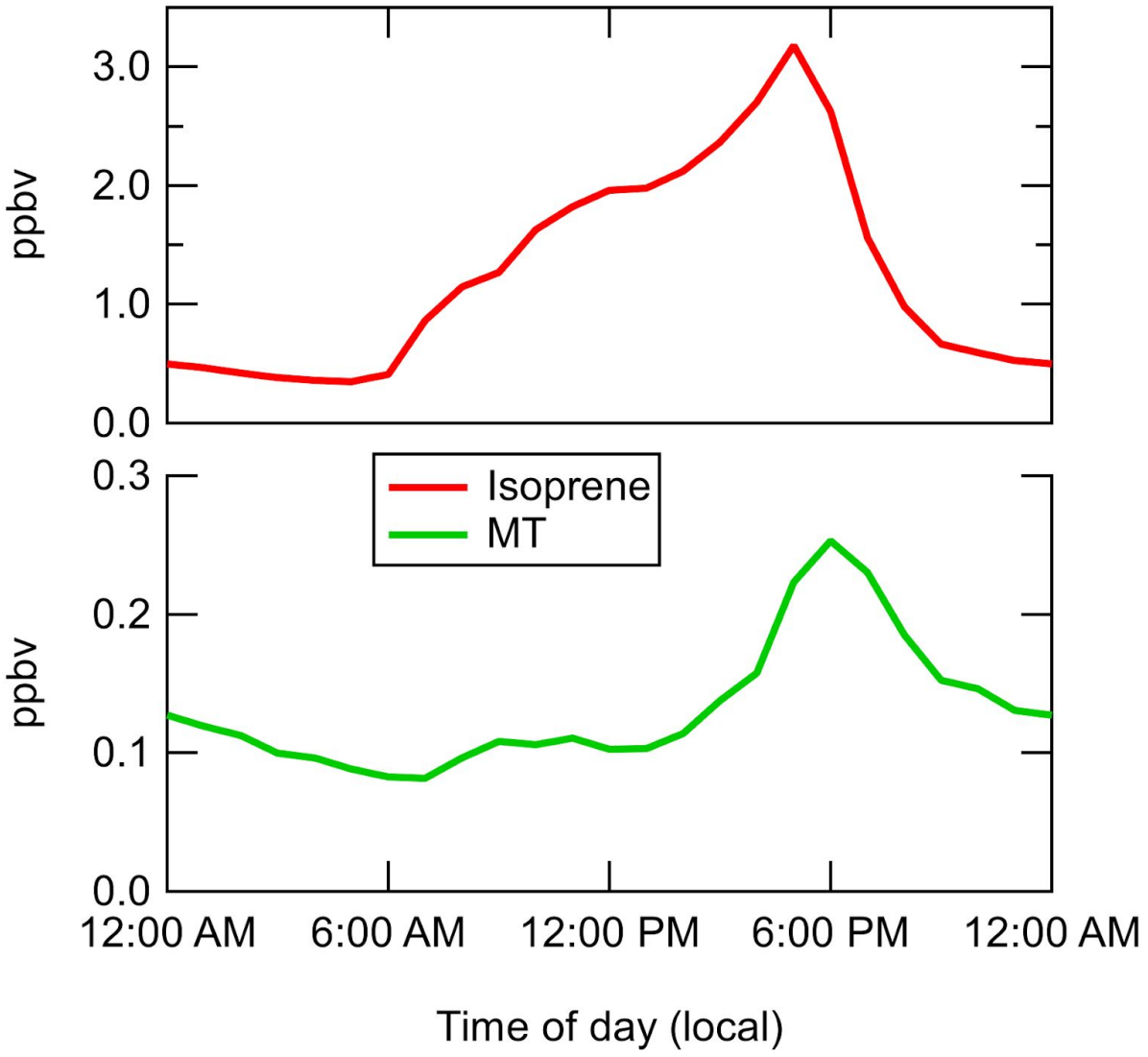

Figure S2. The mean observed diurnal cycle of isoprene (red line) and monoterpenes (MT; green line) for the dry season during the GoAmazon2014/5 campaign.

S2. BEACHON-RoMBAS individual exposures and best/worst case distributions

The following are the individual representations of the model simulations for each BEACHON-RoMBAS exposure modelled in this study. Figures S3, S5, S7, S9, S11, S13, S15, and S17 give the normalized mean error (NME) values for the parameter space that lies within (Table 3) the NUC1 nucleation scheme and base value of the reactive uptake coefficient. Figures S4, S6, S8, S10, S12, S14, S16, and S18 plot each observed final moment (solid black lines) used in computing the NME statistic (number, diameter, surface area, and volume) compared to the six TOMAS cases with the lowest (best) NME statistic (solid colored lines) and six TOMAS cases with the highest (worst) NME statistic (solid dotted lines). For comparison, the observed initial moments are also plotted for each moment (dotted black lines).

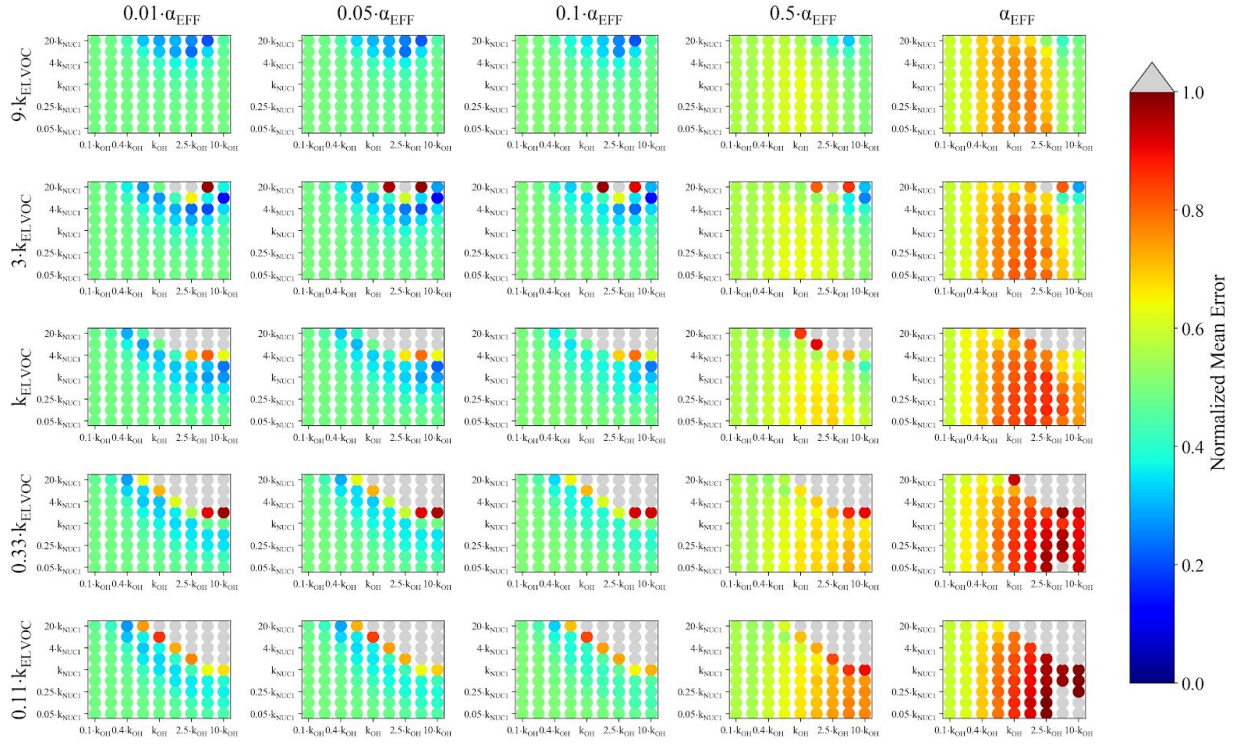

Figure S3. Representation of the parameter space for a 0.090 eq. day aging exposure from BEACHON-RoMBAS for the NUC1 nucleation scheme and base value of the reactive uptake coefficient of 0.6. The effective accommodation coefficient increases across each row of panels; the rate constant of gas-phase fragmentation increases up each column of panels. Within each panel, the rate constant of gas-phase reactions with OH increases along the x-axis and the rate constant for nucleation increases along the y-axis. The color bar indicates the normalized mean error (NME) value for each simulation, with the lowest values indicating the least error between model and measurement. Grey regions indicate regions within the parameter space whose NME value is greater than 1.

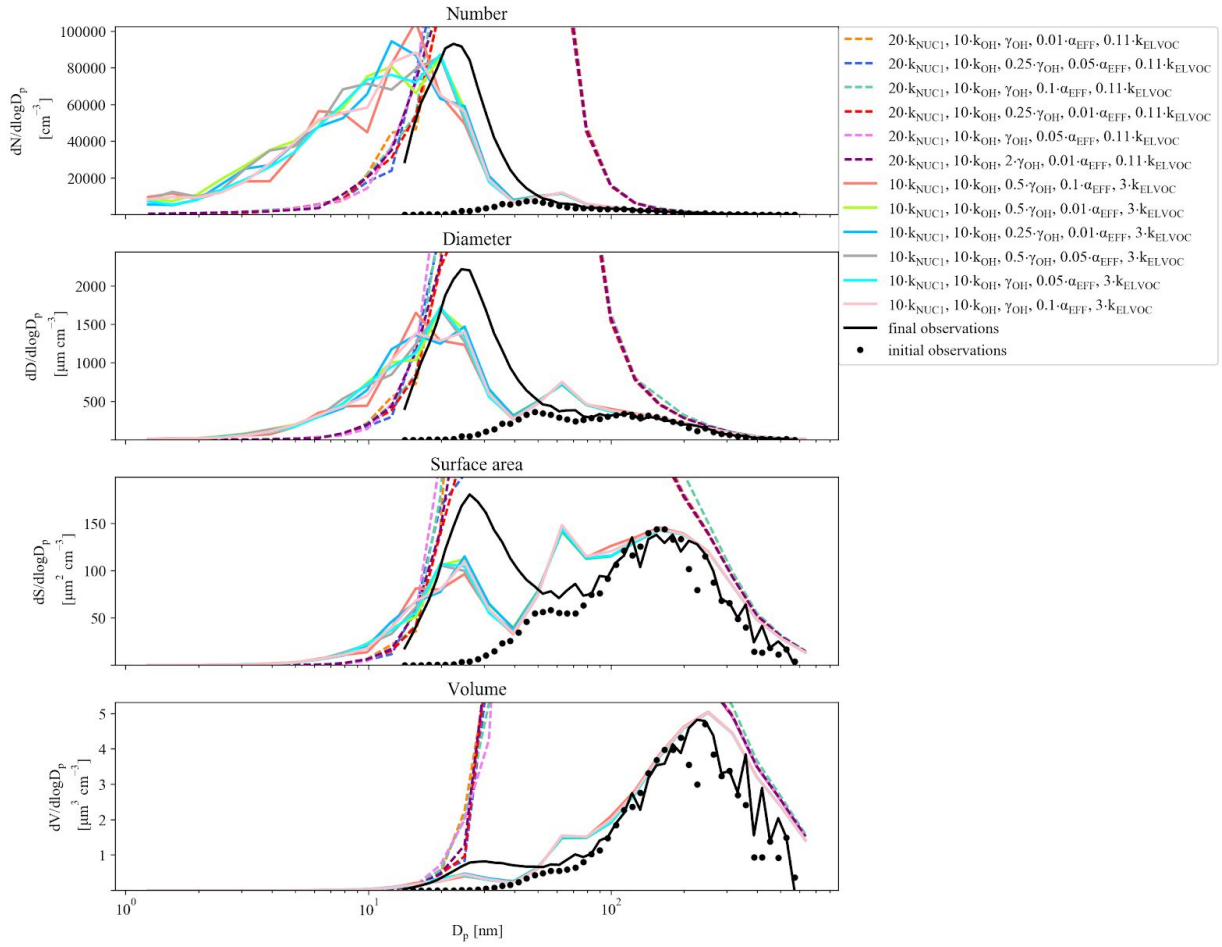

Figure S4. Example of best (solid lines) and worst (dashed lines) fit size distributions compared to the observed (black line) final OFR size distribution for a 0.090 eq days aging case from BEACHON-RoMBAS for the NUC1 nucleation scheme. The fits are determined using the mean error of moments method (see methods); each panel represents a separate moment. The top panel represents the number distribution; the second panel represents the diameter distribution; the third panel represents the surface area distribution; and the final (bottom) distribution represents the volume distribution.

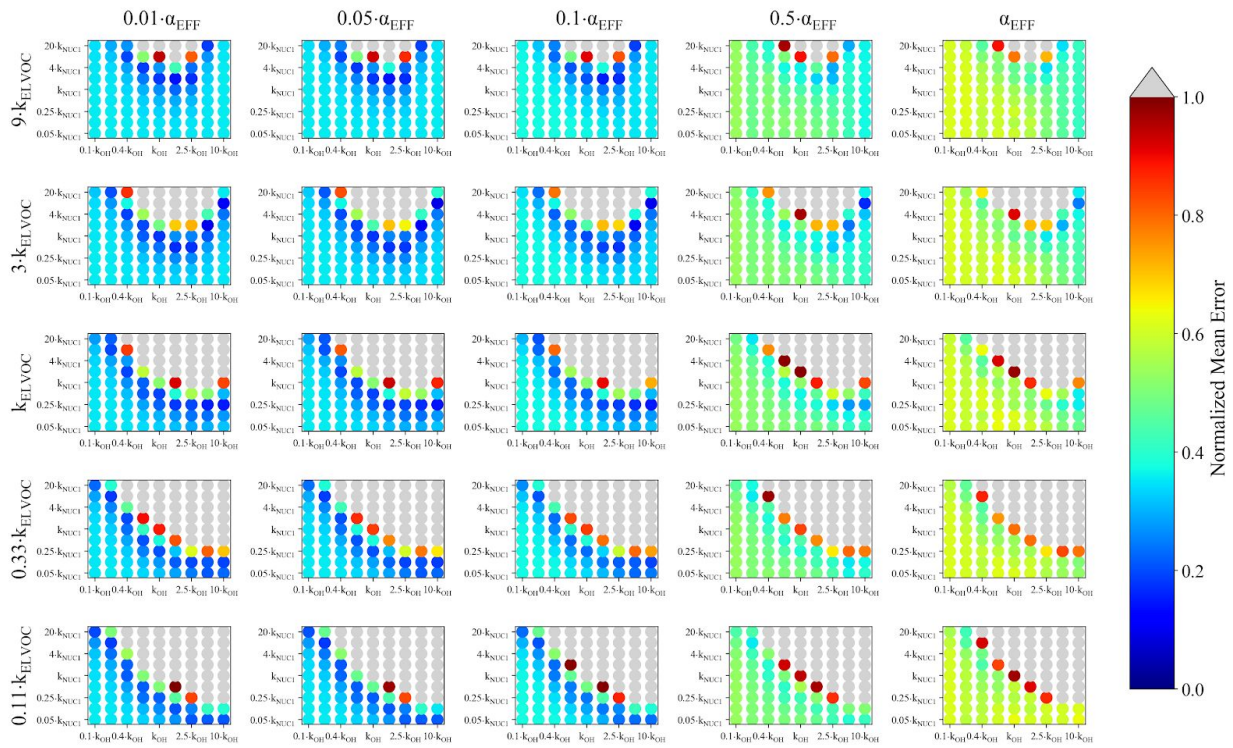

Figure S5. Representation of the parameter space for a 0.098 eq. day aging exposure from BEACHON-RoMBAS for the NUC1 nucleation scheme and base value of the reactive uptake coefficient of 0.6. The effective accommodation coefficient increases across each row of panels; the rate constant of gas-phase fragmentation increases up each column of panels. Within each panel, the rate constant of gas-phase reactions with OH increases along the x-axis and the rate constant for nucleation increases along the y-axis. The color bar indicates the normalized mean error (NME) value for each simulation, with the lowest values indicating the least error between model and measurement. Grey regions indicate regions within the parameter space whose NME value is greater than 1.

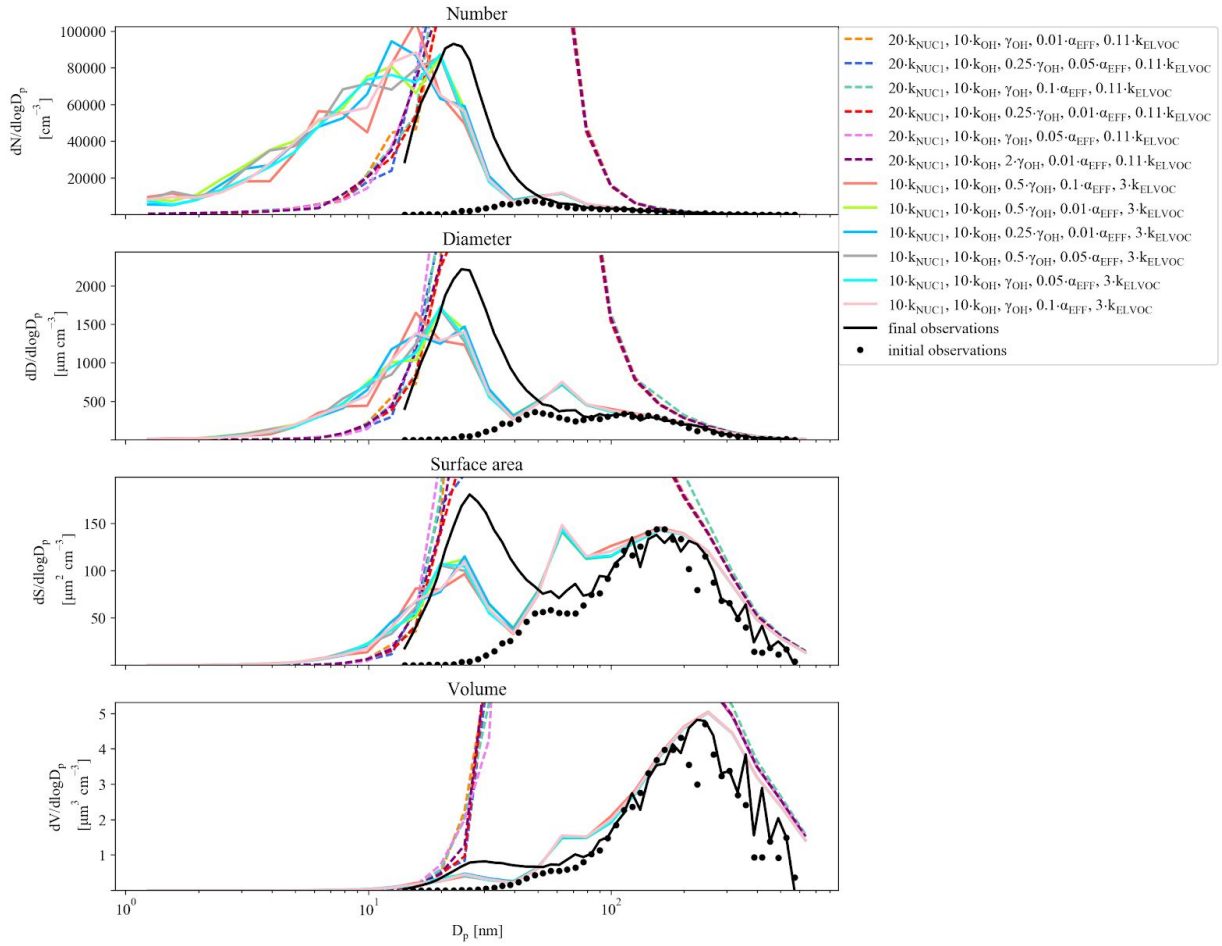

Figure S6. Example of best (solid lines) and worst (dashed lines) fit size distributions compared to the observed (black line) final OFR size distribution for a 0.098 eq days aging case from BEACHON-RoMBAS for the NUC1 nucleation scheme. The fits are determined using the mean error of moments method (see methods); each panel represents a separate moment. The top panel represents the number distribution; the second panel represents the diameter distribution; the third panel represents the surface area distribution; and the final (bottom) distribution represents the volume distribution.

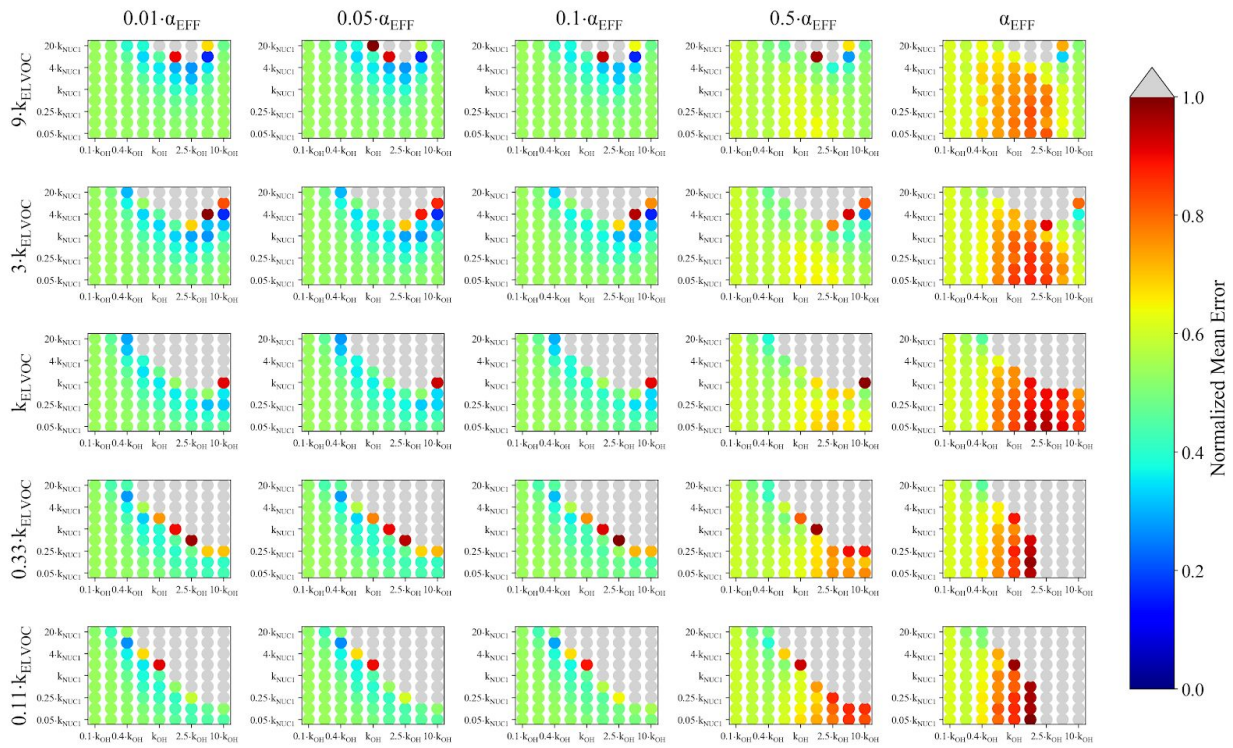

Figure S7. Representation of the parameter space for a 0.16 eq. day aging exposure from BEACHON-RoMBAS for the NUC1 nucleation scheme and base value of the reactive uptake coefficient of 0.6. The effective accommodation coefficient increases across each row of panels; the rate constant of gas-phase fragmentation increases up each column of panels. Within each panel, the rate constant of gas-phase reactions with OH increases along the x-axis and the rate constant for nucleation increases along the y-axis. The color bar indicates the normalized mean error (NME) value for each simulation, with the lowest values indicating the least error between model and measurement. Grey regions indicate regions within the parameter space whose NME value is greater than 1.

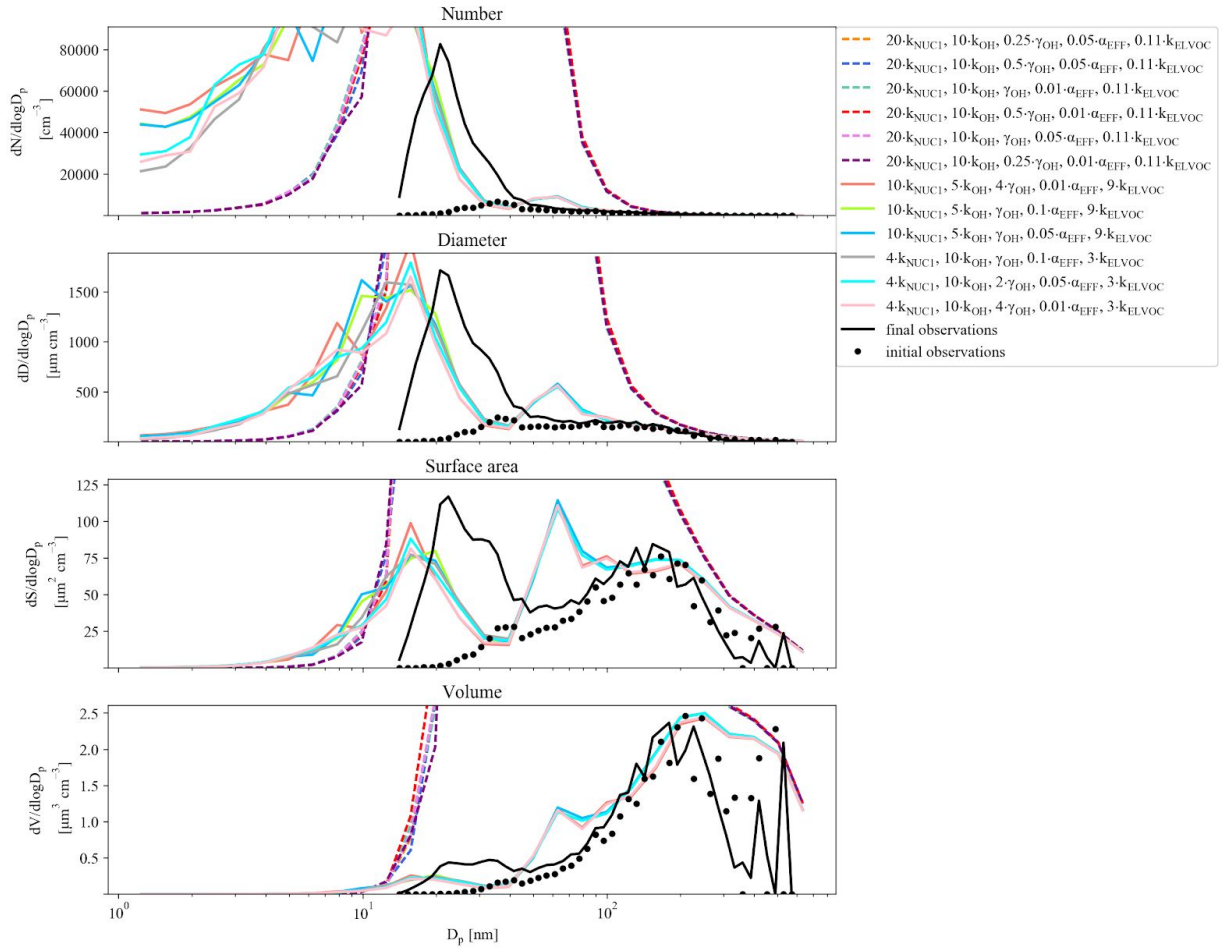

Figure S8. Example of best (solid lines) and worst (dashed lines) fit size distributions compared to the observed (black line) final OFR size distribution for a 0.16 eq days aging case from BEACHON-RoMBAS for the NUC1 nucleation scheme. The fits are determined using the mean error of moments method (see methods); each panel represents a separate moment. The top panel represents the number distribution; the second panel represents the diameter distribution; the third panel represents the surface area distribution; and the final (bottom) distribution represents the volume distribution.

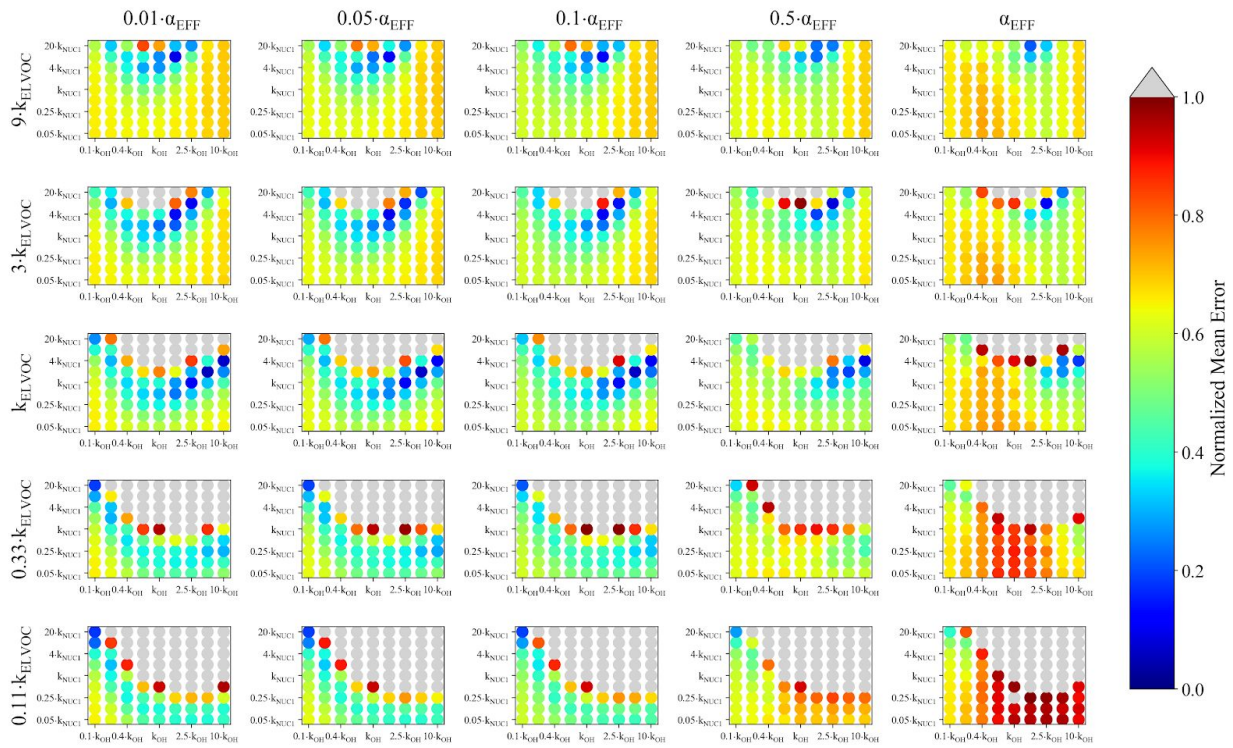

Figure S9. Representation of the parameter space for a 0.23 eq. day aging exposure from BEACHON-RoMBAS for the NUC1 nucleation scheme and base value of the reactive uptake coefficient of 0.6. The effective accommodation coefficient increases across each row of panels; the rate constant of gas-phase fragmentation increases up each column of panels. Within each panel, the rate constant of gas-phase reactions with OH increases along the x-axis and the rate constant for nucleation increases along the y-axis. The color bar indicates the normalized mean error (NME) value for each simulation, with the lowest values indicating the least error between model and measurement. Grey regions indicate regions within the parameter space whose NME value is greater than 1.

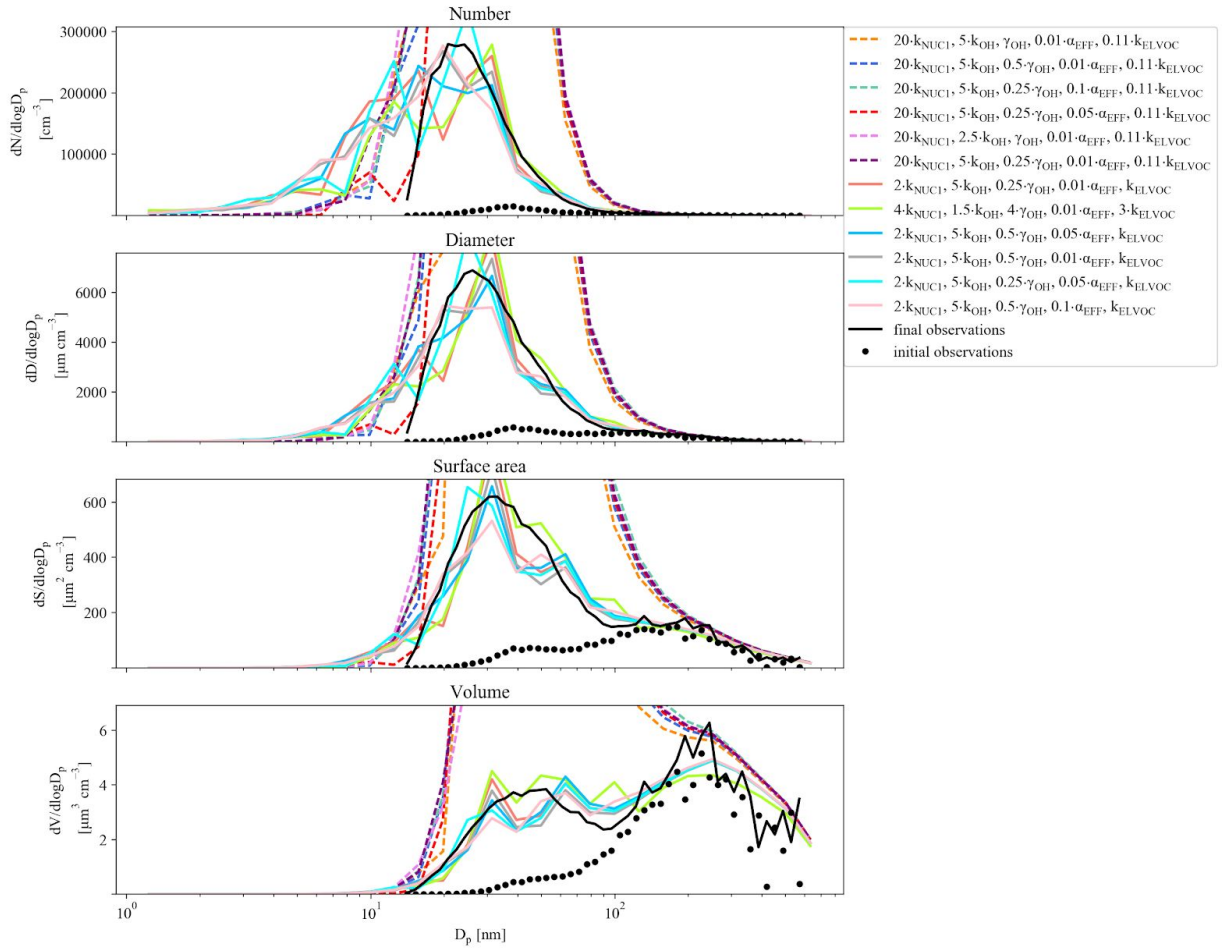

Figure S10. Example of best (solid lines) and worst (dashed lines) fit size distributions compared to the observed (black line) final OFR size distribution for a 0.23 eq days aging case from BEACHON-RoMBAS for the NUC1 nucleation scheme. The fits are determined using the mean error of moments method (see methods); each panel represents a separate moment. The top panel represents the number distribution; the second panel represents the diameter distribution; the third panel represents the surface area distribution; and the final (bottom) distribution represents the volume distribution.

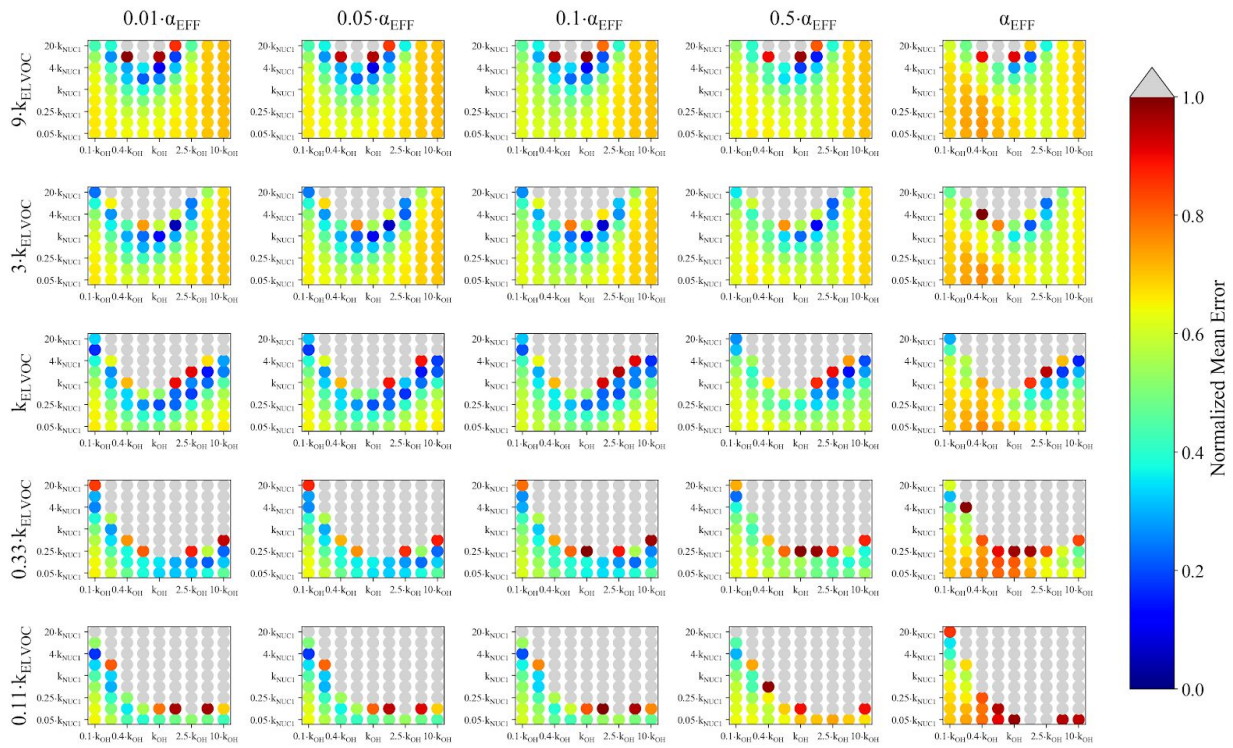

Figure S11. Representation of the parameter space for a 0.27 eq. day aging exposure from BEACHON-RoMBAS for the NUC1 nucleation scheme and base value of the reactive uptake coefficient of 0.6. The effective accommodation coefficient increases across each row of panels; the rate constant of gas-phase fragmentation increases up each column of panels. Within each panel, the rate constant of gas-phase reactions with OH increases along the x-axis and the rate constant for nucleation increases along the y-axis. The color bar indicates the normalized mean error (NME) value for each simulation, with the lowest values indicating the least error between model and measurement. Grey regions indicate regions within the parameter space whose NME value is greater than 1.

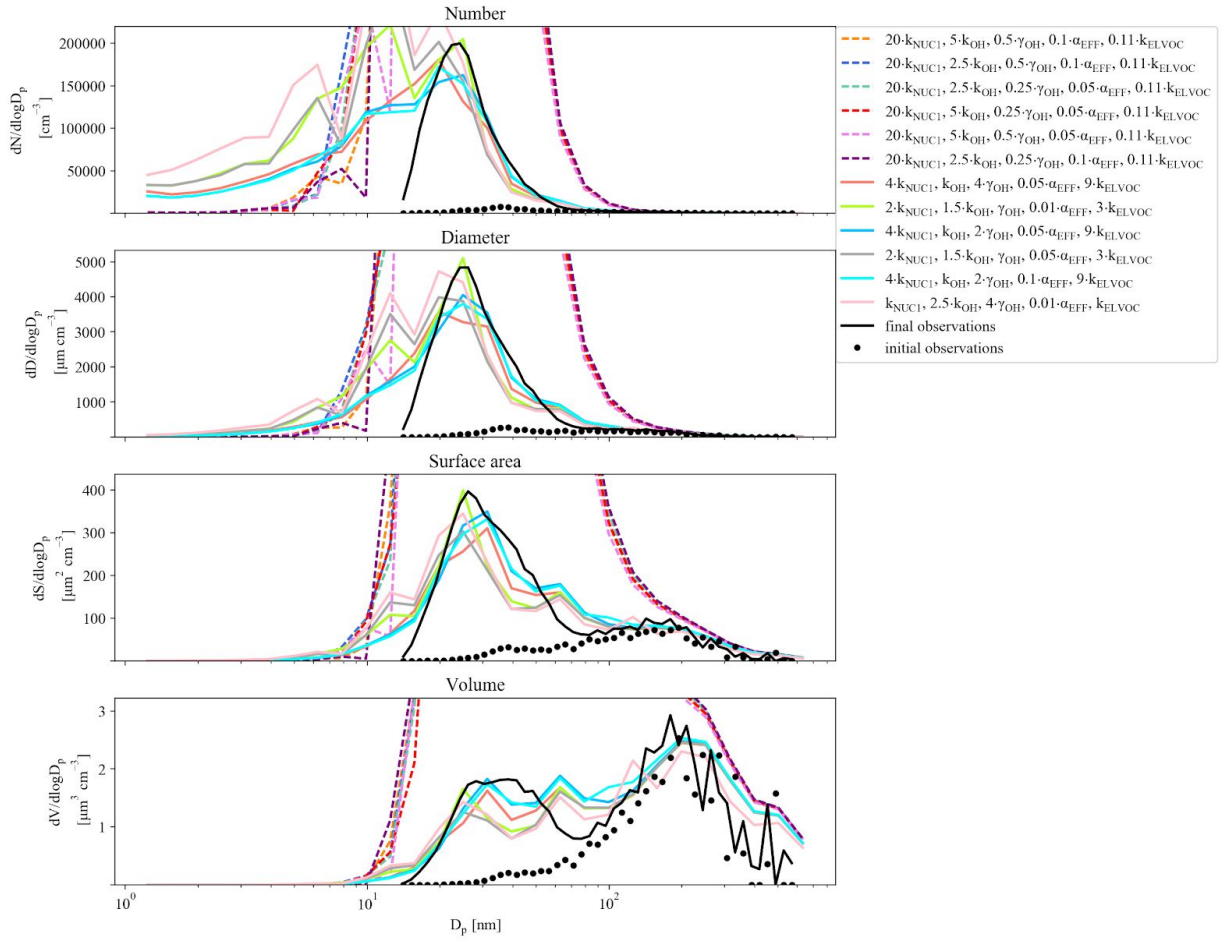

Figure S12. Example of best (solid lines) and worst (dashed lines) fit size distributions compared to the observed (black line) final OFR size distribution for a 0.27 eq days aging case from BEACHON-RoMBAS for the NUC1 nucleation scheme. The fits are determined using the mean error of moments method (see methods); each panel represents a separate moment. The top panel represents the number distribution; the second panel represents the diameter distribution; the third panel represents the surface area distribution; and the final (bottom) distribution represents the volume distribution.

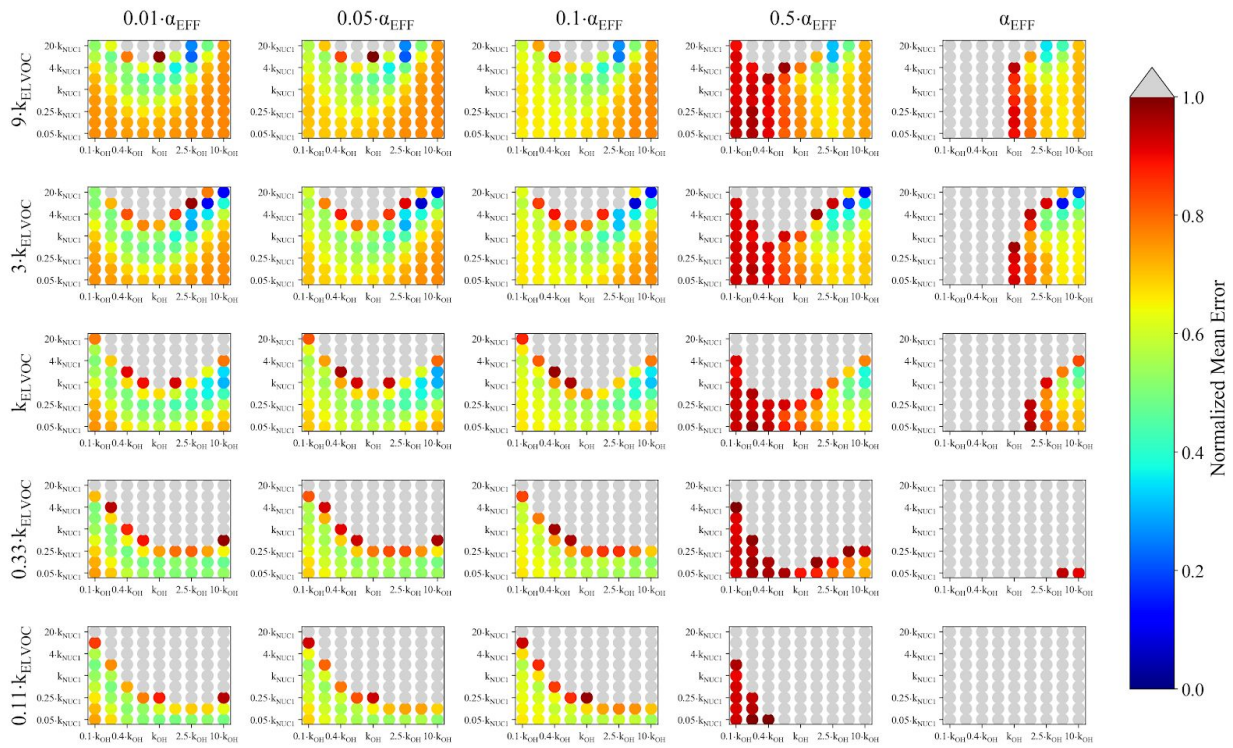

Figure S13. Representation of the parameter space for a 0.77 eq. day aging exposure from BEACHON-RoMBAS for the NUC1 nucleation scheme and base value of the reactive uptake coefficient of 0.6. The effective accommodation coefficient increases across each row of panels; the rate constant of gas-phase fragmentation increases up each column of panels. Within each panel, the rate constant of gas-phase reactions with OH increases along the x-axis and the rate constant for nucleation increases along the y-axis. The color bar indicates the normalized mean error (NME) value for each simulation, with the lowest values indicating the least error between model and measurement. Grey regions indicate regions within the parameter space whose NME value is greater than 1.

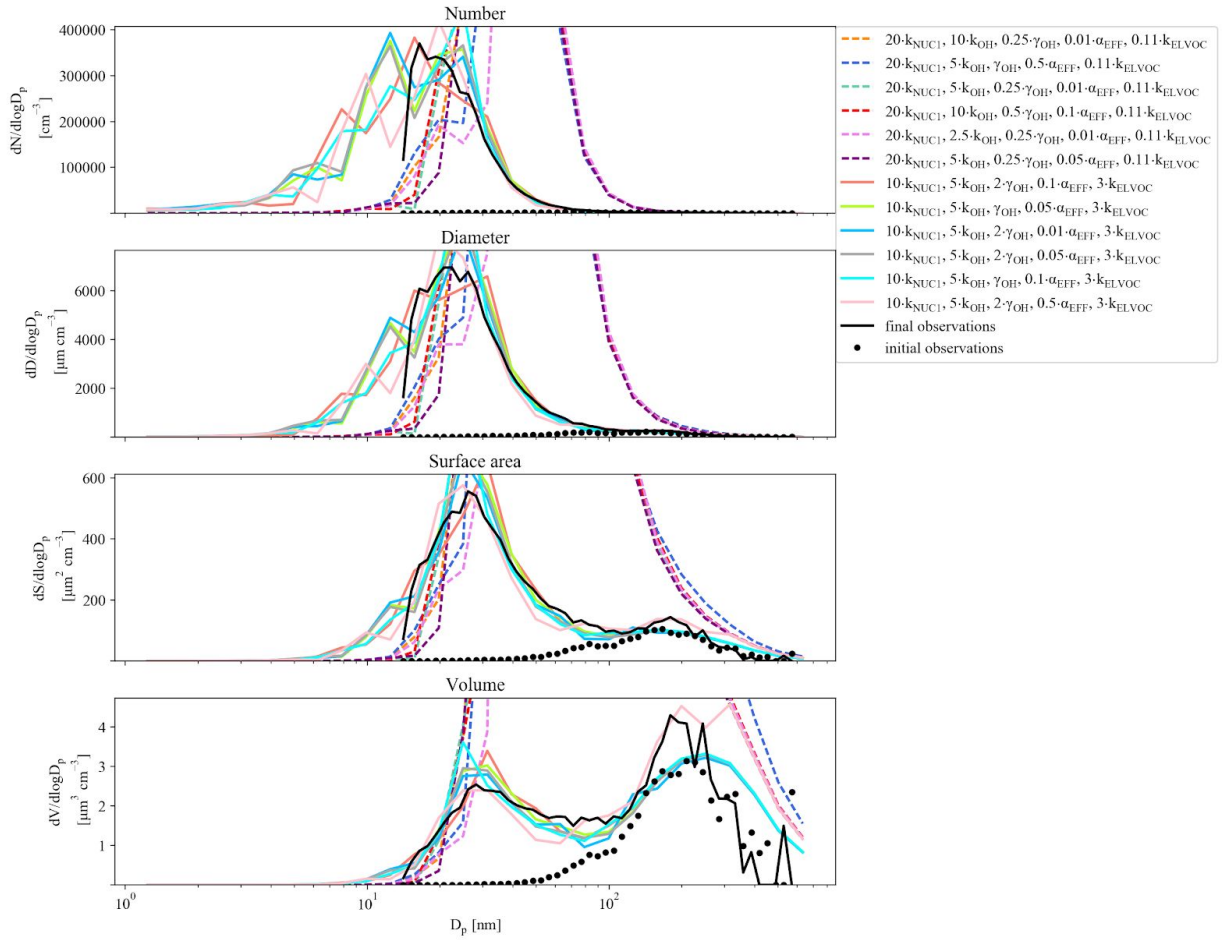

Figure S14. Example of best (solid lines) and worst (dashed lines) fit size distributions compared to the observed (black line) final OFR size distribution for a 0.77 eq days aging case from BEACHON-RoMBAS for the NUC1 nucleation scheme. The fits are determined using the mean error of moments method (see methods); each panel represents a separate moment. The top panel represents the number distribution; the second panel represents the diameter distribution; the third panel represents the surface area distribution; and the final (bottom) distribution represents the volume distribution.

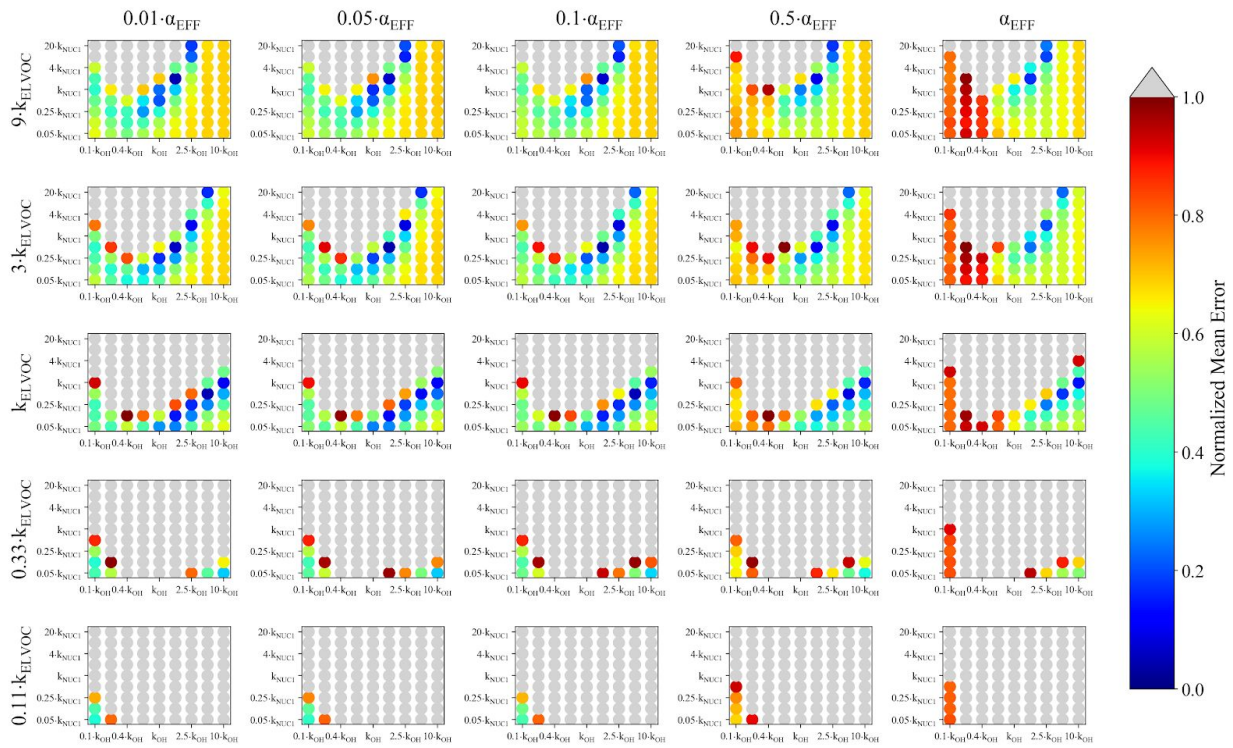

Figure S15. Representation of the parameter space for a 0.82 eq. day aging exposure from BEACHON-RoMBAS for the NUC1 nucleation scheme and base value of the reactive uptake coefficient of 0.6. The effective accommodation coefficient increases across each row of panels; the rate constant of gas-phase fragmentation increases up each column of panels. Within each panel, the rate constant of gas-phase reactions with OH increases along the x-axis and the rate constant for nucleation increases along the y-axis. The color bar indicates the normalized mean error (NME) value for each simulation, with the lowest values indicating the least error between model and measurement. Grey regions indicate regions within the parameter space whose NME value is greater than 1.

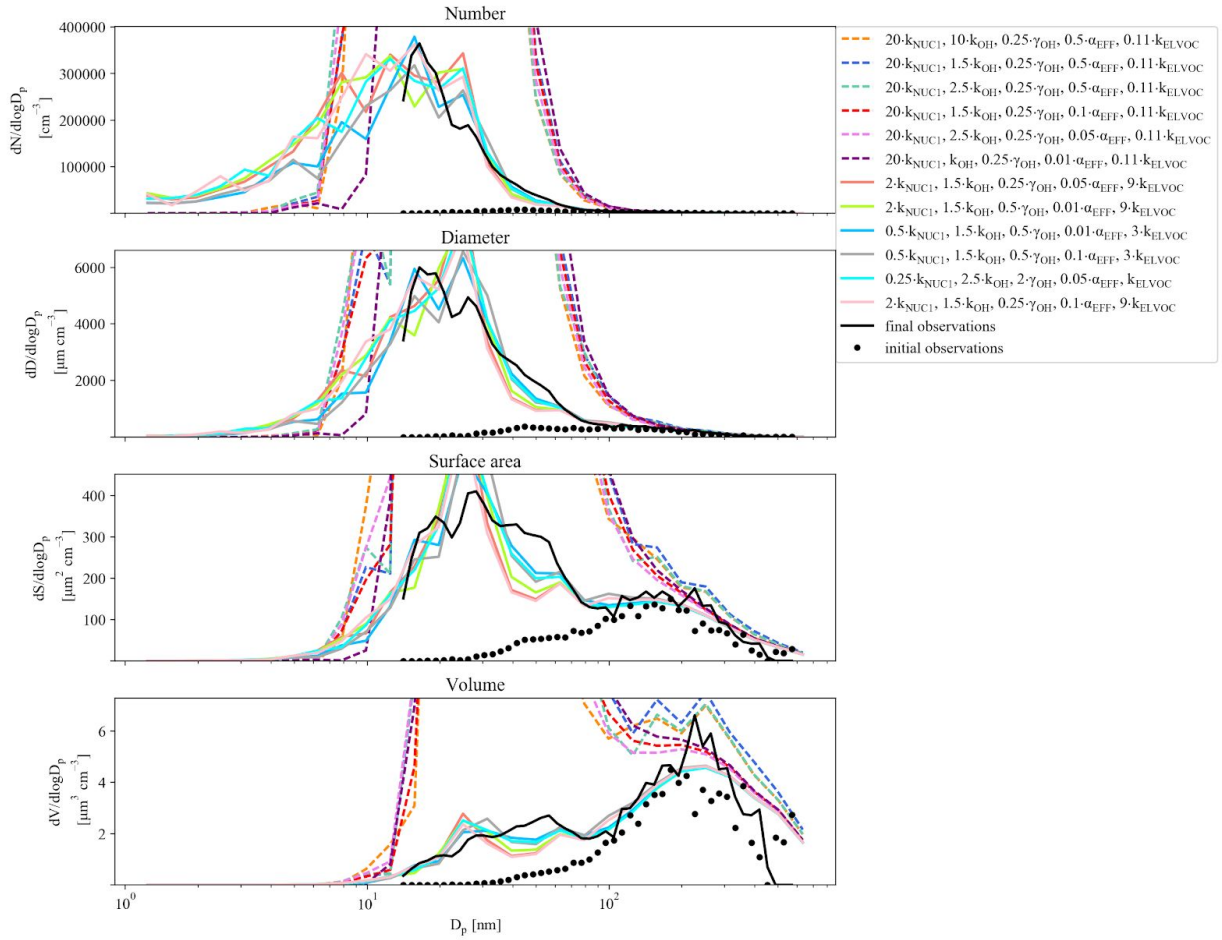

Figure S16. Example of best (solid lines) and worst (dashed lines) fit size distributions compared to the observed (black line) final OFR size distribution for a 0.82 eq days aging case from BEACHON-RoMBAS for the NUC1 nucleation scheme. The fits are determined using the mean error of moments method (see methods); each panel represents a separate moment. The top panel represents the number distribution; the second panel represents the diameter distribution; the third panel represents the surface area distribution; and the final (bottom) distribution represents the volume distribution.

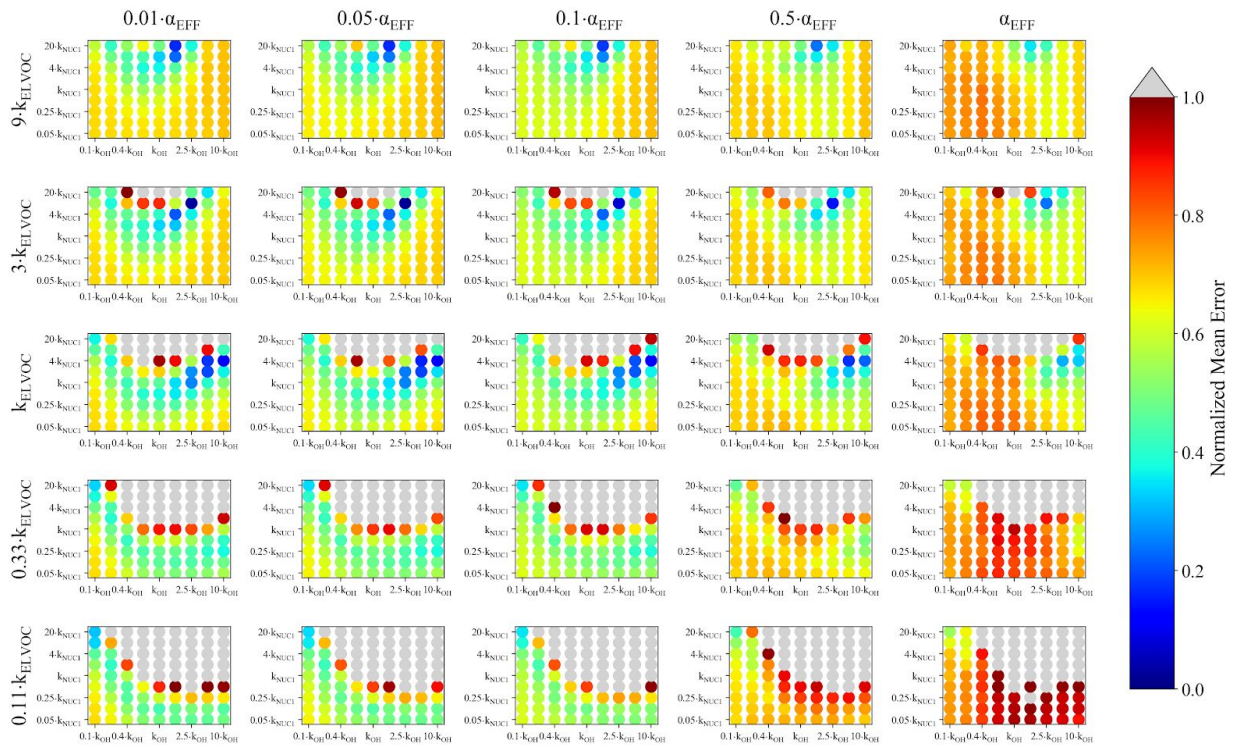

Figure S17. Representation of the parameter space for a 0.91 eq. day aging exposure from BEACHON-RoMBAS for the NUC1 nucleation scheme and base value of the reactive uptake coefficient of 0.6. The effective accommodation coefficient increases across each row of panels; the rate constant of gas-phase fragmentation increases up each column of panels. Within each panel, the rate constant of gas-phase reactions with OH increases along the x-axis and the rate constant for nucleation increases along the y-axis. The color bar indicates the normalized mean error (NME) value for each simulation, with the lowest values indicating the least error between model and measurement. Grey regions indicate regions within the parameter space whose NME value is greater than 1.

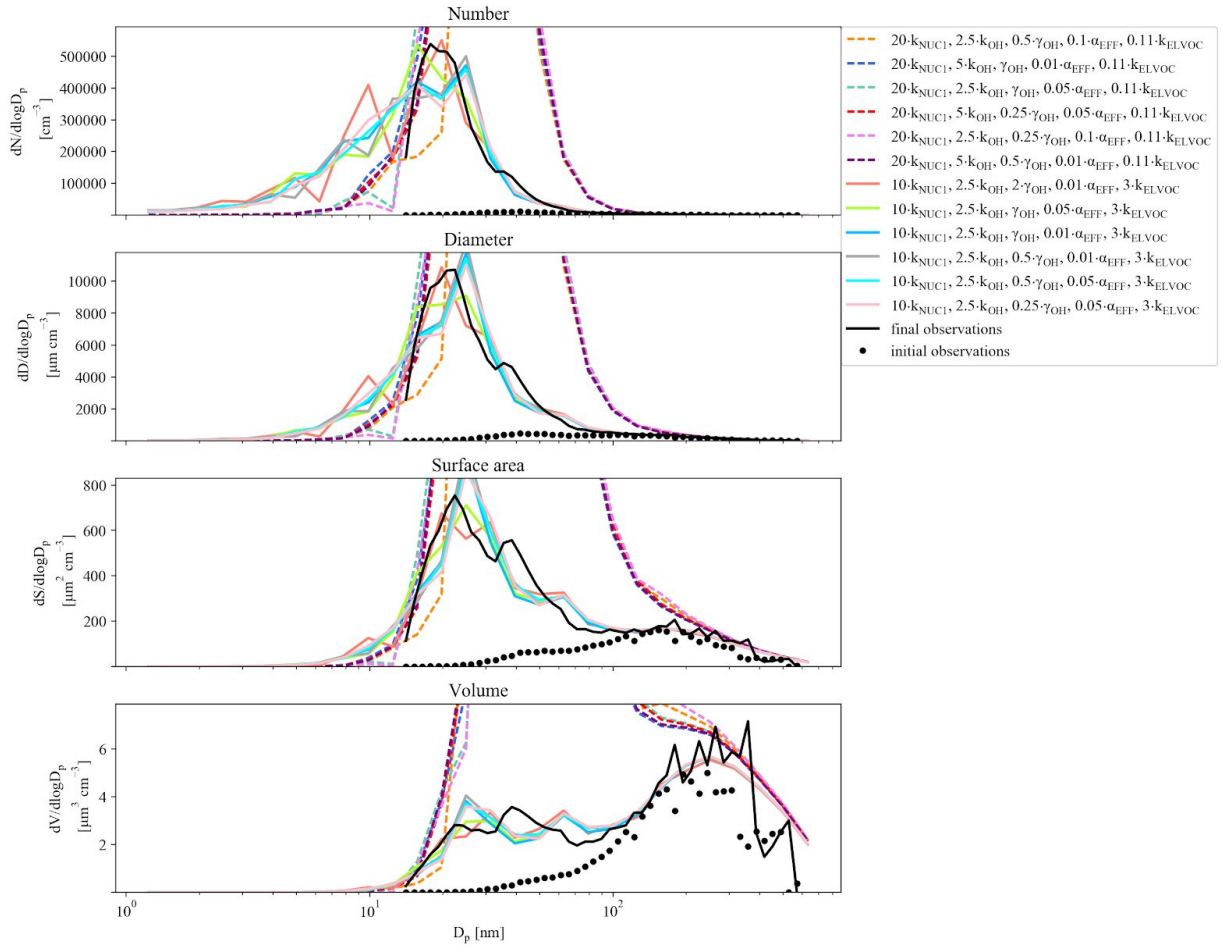

Figure S18. Example of best (solid lines) and worst (dashed lines) fit size distributions compared to the observed (black line) final OFR size distribution for a 0.91 eq days aging case from BEACHON-RoMBAS for the NUC1 nucleation scheme. The fits are determined using the mean error of moments method (see methods); each panel represents a separate moment. The top panel represents the number distribution; the second panel represents the diameter distribution; the third panel represents the surface area distribution; and the final (bottom) distribution represents the volume distribution.

S3. Sensitivity of the model to the NUC2 nucleation scheme for BEACHON-RoMBAS

Figures S19 gives the average NME values across the 0.09-0.9 day eq. aging exposures from BEACHON-RoMBAS examined in this study for the parameter space that lies within (Table 3) the NUC2 nucleation scheme and base value of the reactive uptake coefficient. This figure is qualitatively similar to Fig. 5, which shows the same parameter space except for the NUC1 nucleation scheme.

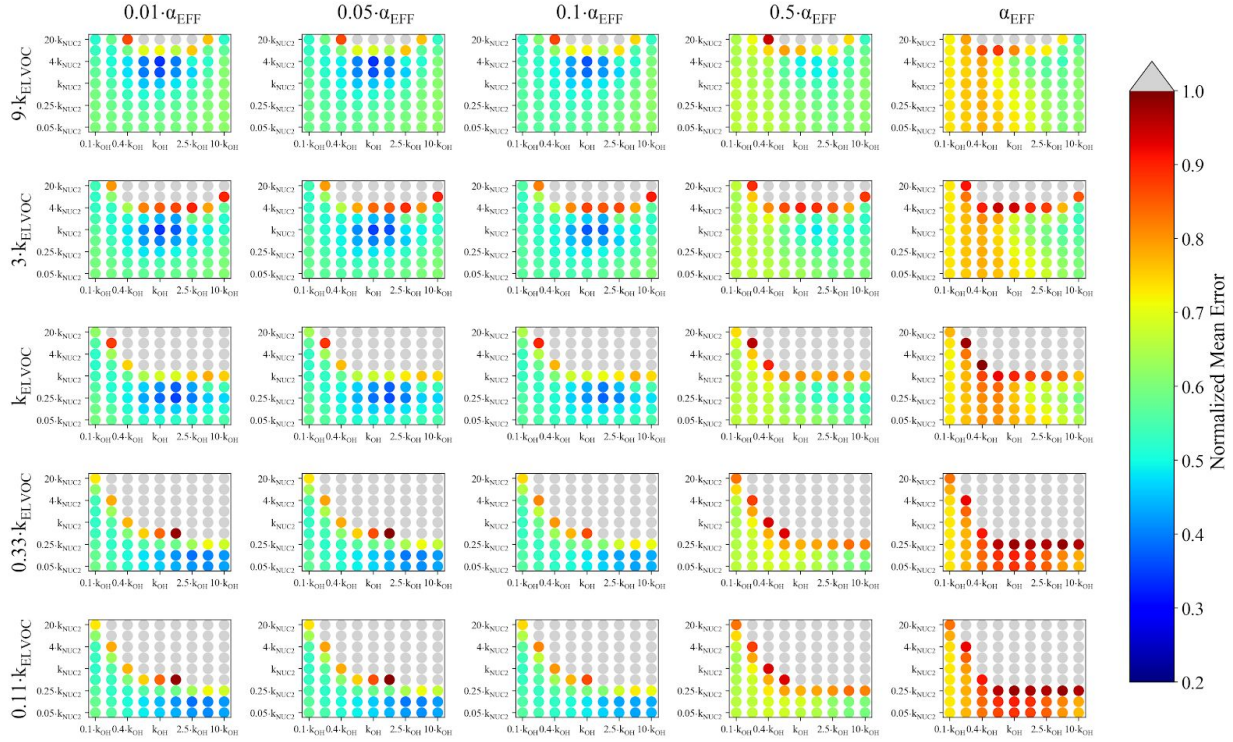

Figure S19. Representation of the parameter space for the average across the 0.09-0.9 day eq. aging exposures from BEACHON-RoMBAS examined in this study for the NUC2 nucleation scheme and base value of the reactive uptake coefficient of 0.6. The effective accommodation coefficient increases across each row of panels; the rate constant of gas-phase fragmentation increases up each column of panels. Within each panel, the rate constant of gas-phase reactions with OH increases along the x-axis and the rate constant for nucleation increases along the y-axis. The color bar indicates the normalized mean error (NME) value for each simulation, with the lowest values indicating the least error between model and measurement. Grey regions indicate regions within the parameter space whose NME value is greater than 1. No averaged case had a NME value less than 0.2 for the cases shown here.

S4. Sensitivity of the model to the heterogeneous chemistry parameter for BEACHON-RoMBAS

In order to determine the model sensitivity to the reactive uptake coefficient, γ_{OH}, we replot the results of Fig. 4 in Fig. S20 to be a function of γ_{OH} (increasing across each column) for the base value of the gas-phase fragmentation rate constant.

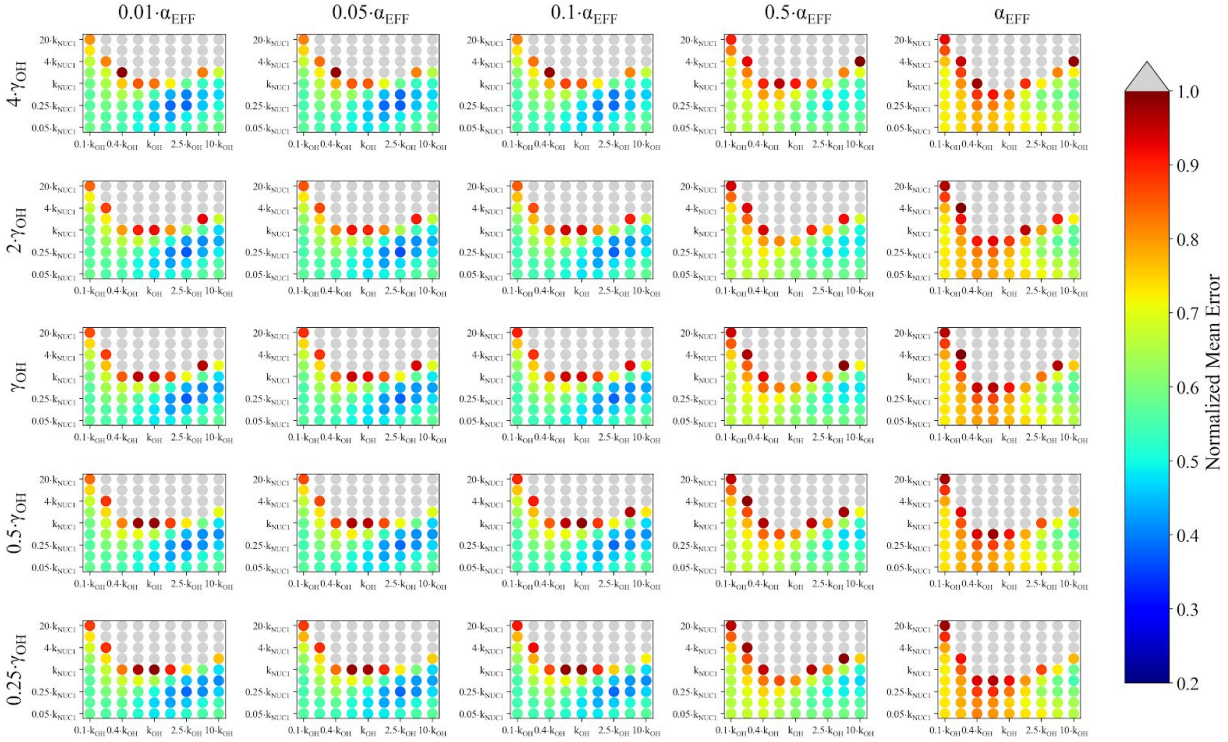

Figure S20. Representation of the parameter space for the average across the 0.09-0.9 day eq. aging exposures from BEACHON-RoMBAS examined in this study for the NUC1 nucleation scheme and base value of the gas-phase fragmentation rate constant (k_{ELVOC}). The effective accommodation increases across each row; the reactive uptake coefficient increases up each column of panels. Within each panel, the rate constant of gas-phase reactions with OH increases along the x-axis and the rate constant for nucleation increases along the y-axis. The color bar indicates the normalized mean error (NME) value for each simulation, with the lowest values indicating the least error between model and measurement. Grey regions indicate regions within the parameter space whose NME value is greater than 1. No averaged case had a NME value less than 0.2 for the cases shown here. This figure shows that for a fixed value of reactive uptake coefficient (each column), the NME values for each set of nucleation rate constants/gas-phase functionalization constants/effective accommodation coefficients do not show significant change.

S5. GoAmazon sensitivity to S/IVOCs

To model the GoAmazon OFR distributions, we initially run TOMAS with the S/IVOC input concentration derived by using the campaign average ratio of MT:S/IVOCs from BEACHON-RoMBAS (Table 2), as S/IVOCs were not measured during the GoAmazon campaign. Figures S21 and S22 show that for this initial assumed S/IVOC concentration, we cannot replicate the observed growth (Fig. S22, black solid lines) and instead get poor NME values for all combinations of parameters (Fig. S21).

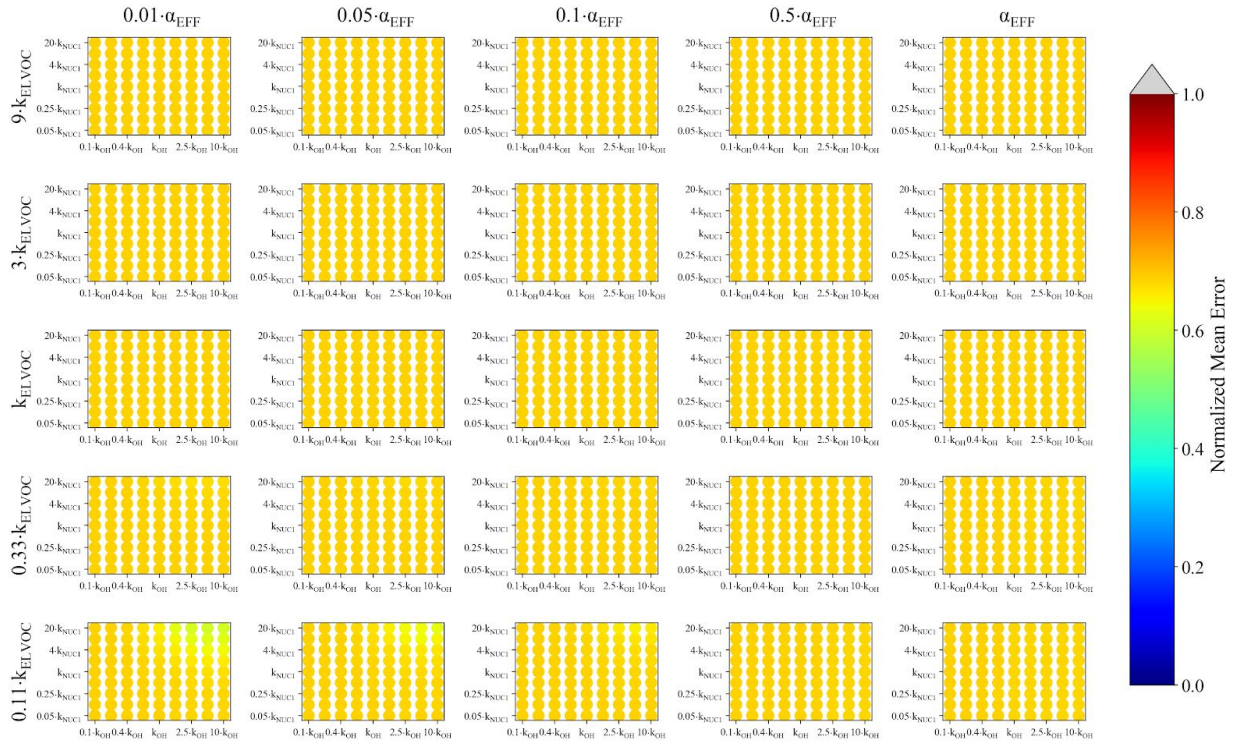

Figure S21. Representation of the parameter space for a 0.53 eq. day aging exposure from GoAmazon for the NUC1 nucleation scheme and base value of the reactive uptake coefficient of 0.6. The effective accommodation coefficient increases across each row of panels; the rate constant of gas-phase fragmentation increases up each column of panels. Within each panel, the rate constant of gas-phase reactions with OH increases along the x-axis and the rate constant for nucleation increases along the y-axis. The color bar indicates the normalized mean error (NME) value for each simulation, with the lowest values indicating the least error between model and measurement. Grey regions indicate regions within the parameter space whose NME value is

greater than 1. In order to see enough growth in the Aitken mode, an increase of between 20-40 times the original S/IVOC input is required. See text and Figs. 6 and S23-S29 for more details.

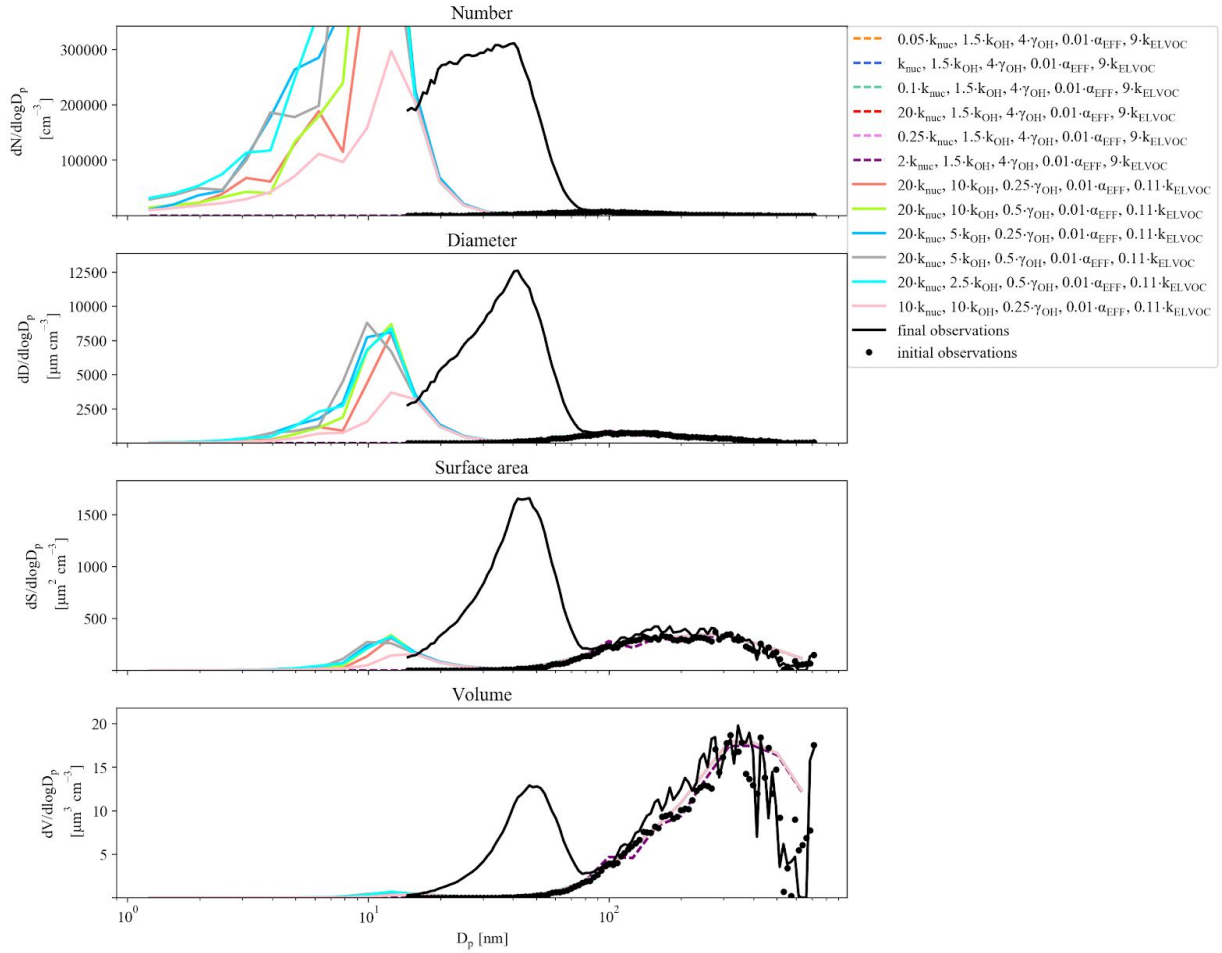

Figure S22. Example of best (solid lines) and worst (dashed lines) fit size distributions compared to the observed (black line) final OFR size distribution for a 0.53 eq days aging case from GoAmazon for the NUC1 nucleation scheme. The fits are determined using the mean error of moments method (see methods); each panel represents a separate moment. The top panel represents the number distribution; the second panel represents the diameter distribution; the third panel represents the surface area distribution; and the final (bottom) distribution represents the volume distribution. In order to see enough growth in the Aitken mode, an increase of between 20-40 times the original S/IVOC input is required. See text and Figs. 6 and S23-S29 for more details.

S6. Sensitivity of the model to the NUC2 and ACT nucleation schemes for GoAmazon

Figures S23 and S24 gives the average NME values across the 0.3-0.6 day eq. aging exposures from GoAmazon examined in this study for the parameter space that lies within the NUC2 nucleation scheme and ACT nucleation scheme, respectively, for base value of the reactive uptake coefficient. These figures use the assumption that the initial S/IVOC concentration is 30 times that of the base S/IVOC concentration (Table 2). Figure S22 is qualitatively similar to Fig. 6, which shows the same parameter space except for the NUC1 nucleation scheme. Figure S23, using the ACT nucleation scheme, shows shifted regions of best-fits as compared to the NUC1 (Fig. 6) and NUC2 (Fig. S22) nucleation schemes.

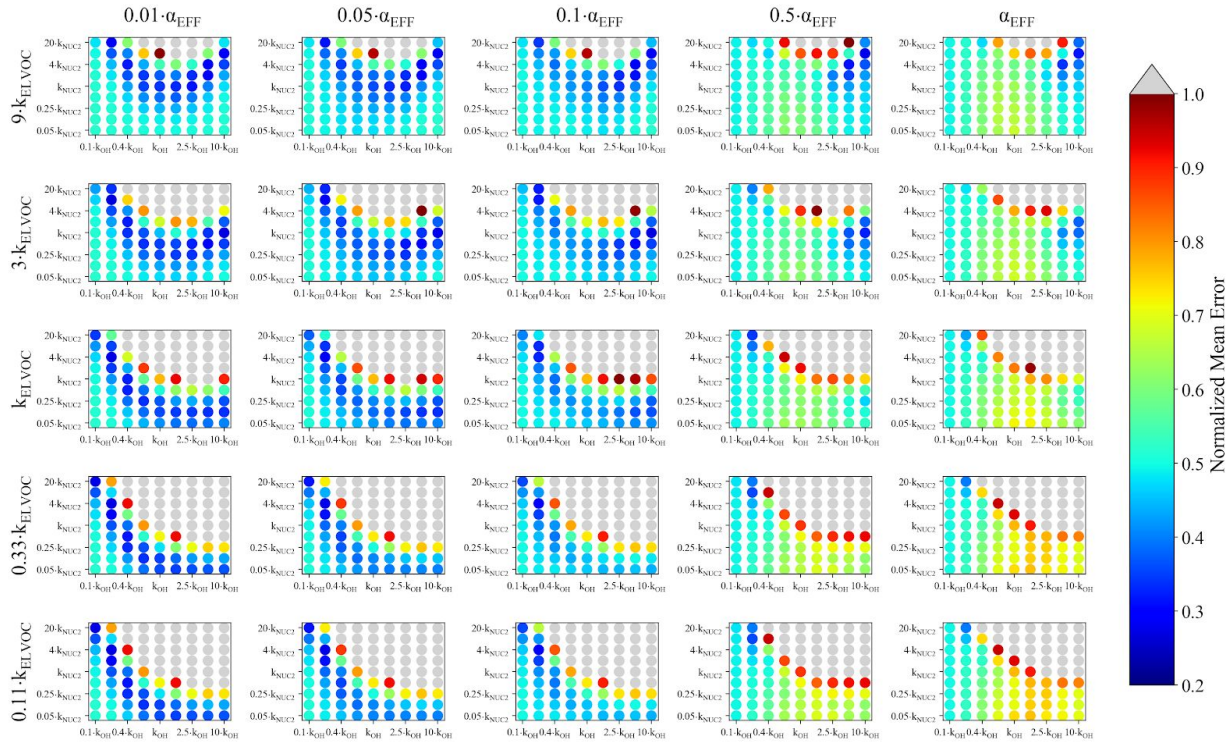

Figure S23. Representation of the parameter space for the average across the 0.3-0.6 day eq. aging exposures from GoAmazon examined in this study for the NUC2 nucleation scheme, base value of the reactive uptake coefficient of 0.6, and assumed S/IVOC concentrations of 30 times that of the baseS/IVOC concentrations. The effective accommodation coefficient increases across each row of panels; the rate constant of gas-phase fragmentation increases up each column of panels. Within each panel, the rate constant of gas-phase reactions with OH increases along the x-axis and the rate constant for nucleation increases along the y-axis. The color bar indicates the normalized mean error (NME) value for each simulation, with the lowest values indicating the least error between model and measurement. Grey regions indicate regions within the

parameter space whose NME value is greater than 1. No averaged case had a NME value less than 0.2 for the cases shown here.

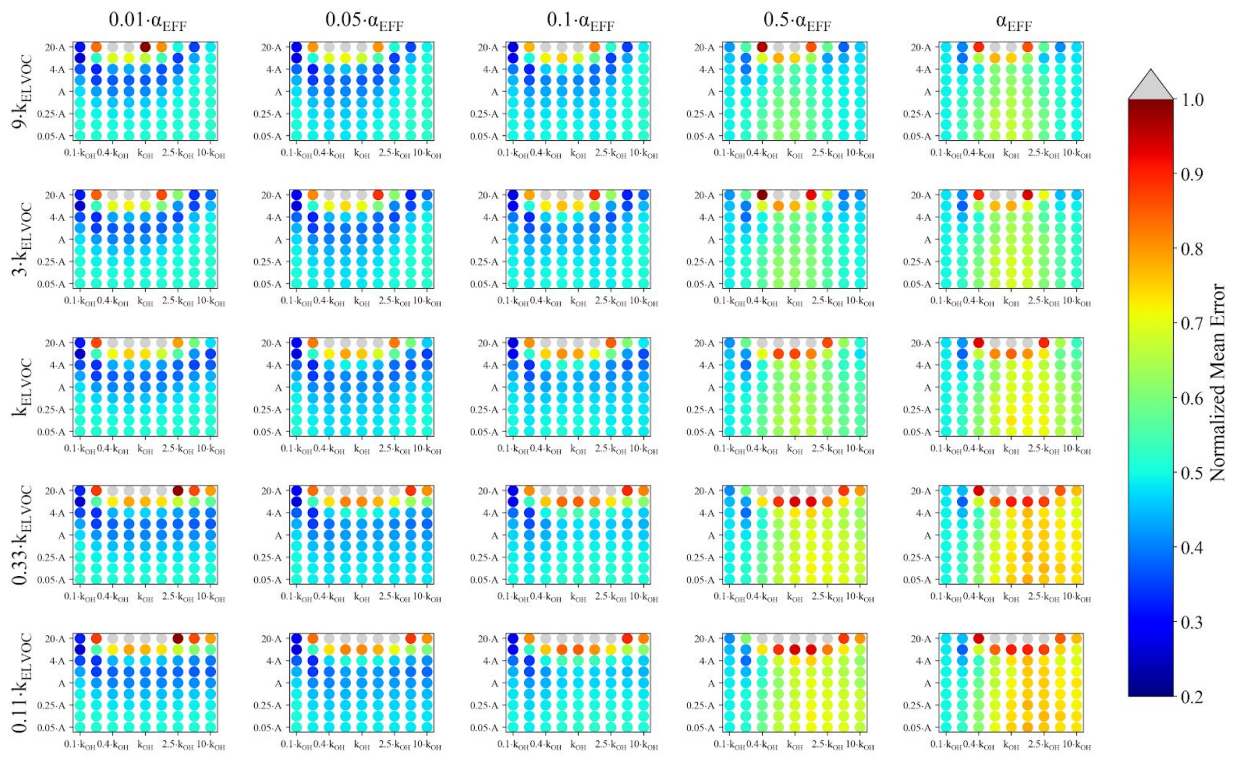

Figure S24. Representation of the parameter space for the average across the 0.3-0.6 day eq. aging exposures from GoAmazon examined in this study for the ACT nucleation scheme, base value of the reactive uptake coefficient of 0.6, and assumed S/IVOC concentration of 30 times that of the base S/IVOC concentration. The effective accommodation coefficient increases across each row of panels; the rate constant of gas-phase fragmentation increases up each column of panels. Within each panel, the rate constant of gas-phase reactions with OH increases along the x-axis and the rate constant for nucleation increases along the y-axis. The color bar indicates the normalized mean error (NME) value for each simulation, with the lowest values indicating the least error between model and measurement. Grey regions indicate regions within the parameter space whose NME value is greater than 1. No averaged case had a NME value less than 0.2 for the cases shown here.

S7. GoAmazon individual exposures and best/worst case distributions

The following are the individual representations of the model simulations for each GoAmazon exposure modelled in this study for the assumption that the initial S/IVOC concentration is 30 times that of the base S/IVOC concentration (Table 2). Figures S25, S27, S29, and S31 give the NME values for the parameter space that lies within (Table 3) the NUC1 nucleation scheme and base value of the reactive uptake coefficient. Figures S26, S28, S30, and S32 plot each observed final moment (solid black lines) used in computing the NME statistic (number, diameter, surface area, and volume) compared to the six TOMAS cases with the lowest (best) NME statistic (solid colored lines) and six TOMAS cases with the highest (worst) NME statistic (solid dotted lines). For comparison, the observed initial moments are also plotted for each moment (dotted black lines).

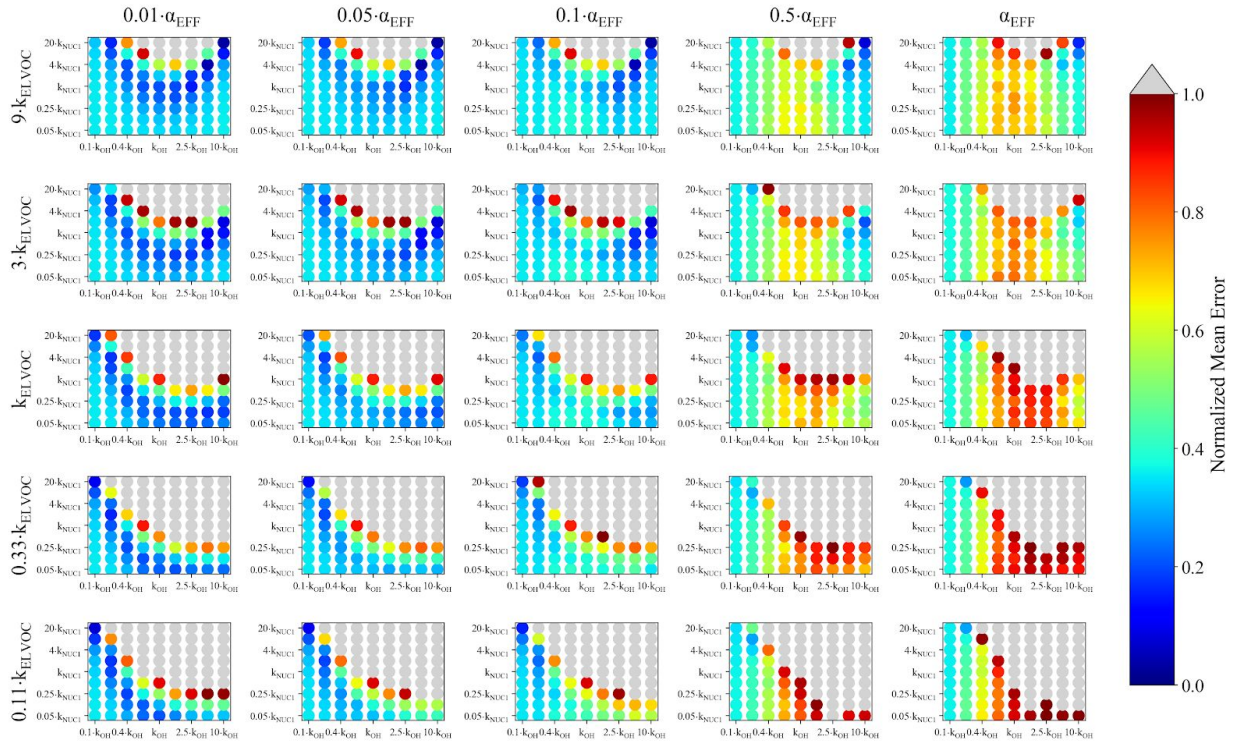

Figure S25. Representation of the parameter space for a 0.39 eq. day aging exposure from GoAmazon for the NUC1 nucleation scheme, base value of the reactive uptake coefficient of 0.6, and 30 times the base S/IVOC input concentration. The effective accommodation coefficient increases across each row of panels; the rate constant of gas-phase fragmentation increases up each column of panels. Within each panel, the rate constant of gas-phase reactions with OH increases along the x-axis and the rate constant for nucleation increases along the y-axis. The color bar indicates the normalized mean error (NME) value for each simulation, with the lowest values indicating the least error between model and measurement. Grey regions indicate regions within the parameter space whose NME value is greater than 1.

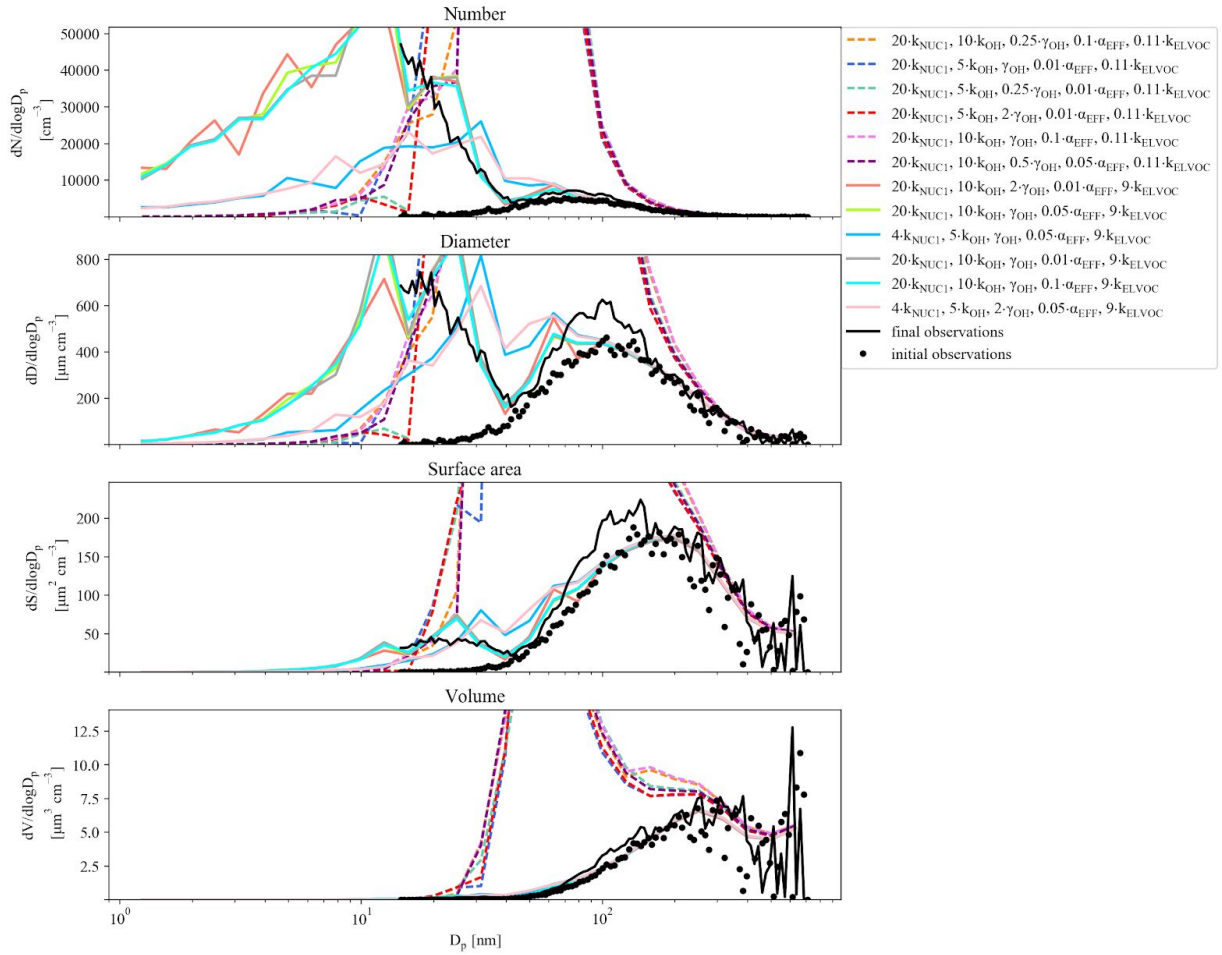

Figure S26. Example of best (solid lines) and worst (dashed lines) fit size distributions compared to the observed (black line) final OFR size distribution for a 0.39 eq. days aging case from GoAmazon for the NUC1 nucleation scheme, using 30 times the base S/IVOC input concentration. The fits are determined using the mean error of moments method (see methods); each panel represents a separate moment. The top panel represents the number distribution; the second panel represents the diameter distribution; the third panel represents the surface area distribution; and the final (bottom) distribution represents the volume distribution.

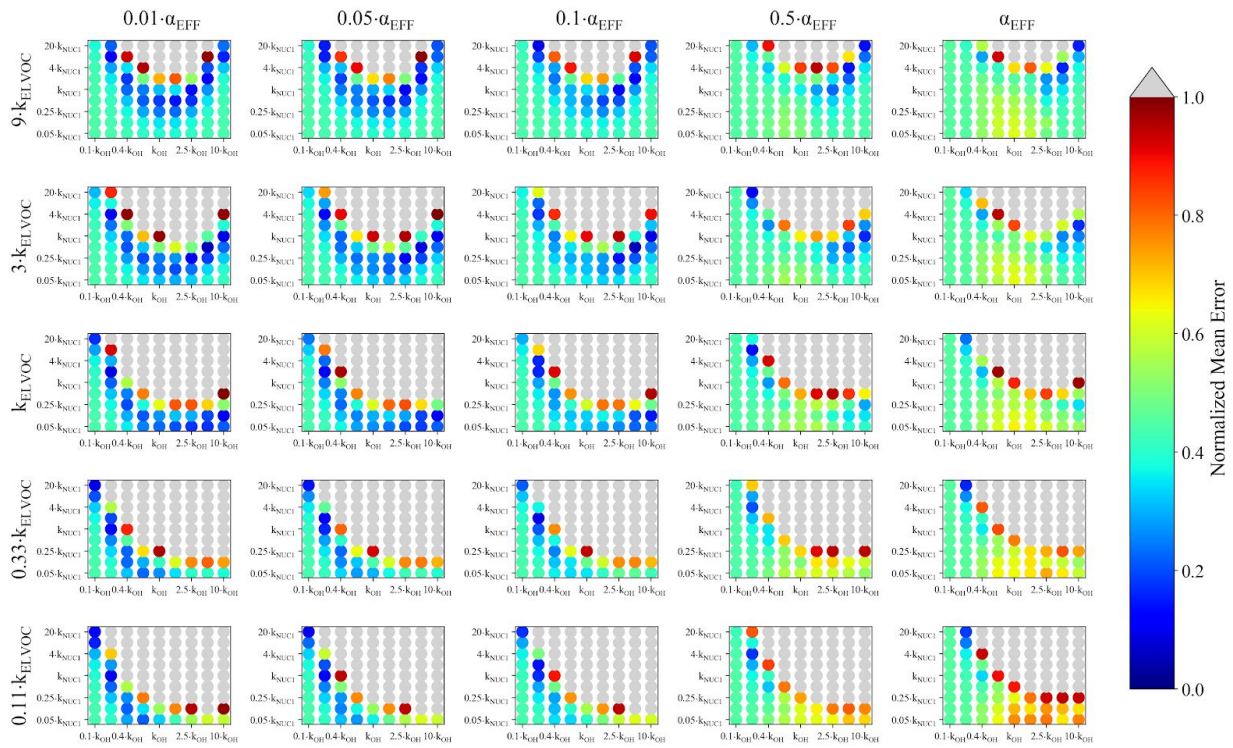

Figure S27. Representation of the parameter space for a 0.40 eq. day aging exposure from GoAmazon for the NUC1 nucleation scheme, base value of the reactive uptake coefficient of 0.6, and 30 times the base S/IVOC input concentration. The effective accommodation coefficient increases across each row of panels; the rate constant of gas-phase fragmentation increases up each column of panels. Within each panel, the rate constant of gas-phase reactions with OH increases along the x-axis and the rate constant for nucleation increases along the y-axis. The color bar indicates the normalized mean error (NME) value for each simulation, with the lowest values indicating the least error between model and measurement. Grey regions indicate regions within the parameter space whose NME value is greater than 1.

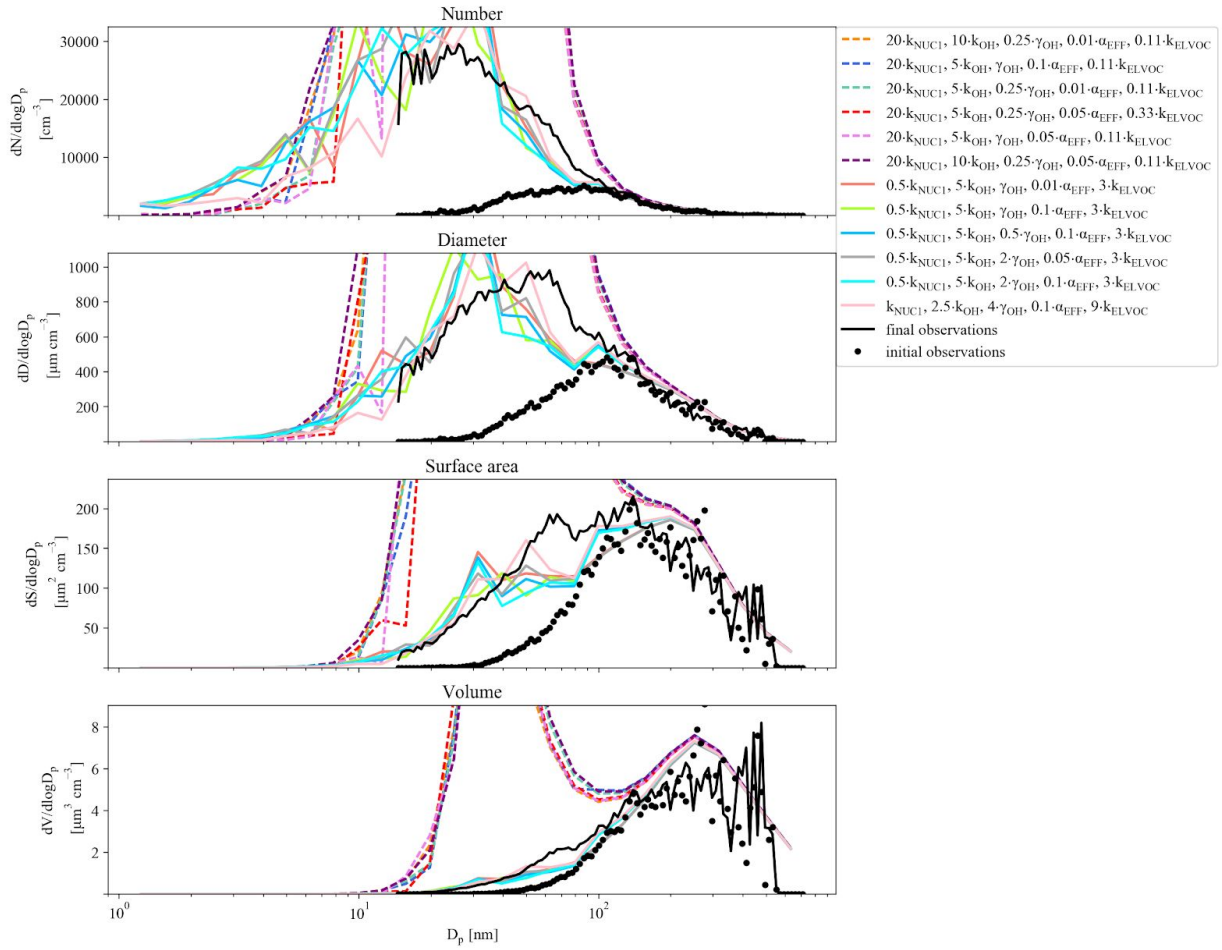

Figure S28. Example of best (solid lines) and worst (dashed lines) fit size distributions compared to the observed (black line) final OFR size distribution for a 0.40 eq days aging case from GoAmazon for the NUC1 nucleation scheme, using 30 times the base S/IVOC input concentration. The fits are determined using the mean error of moments method (see methods); each panel represents a separate moment. The top panel represents the number distribution; the second panel represents the diameter distribution; the third panel represents the surface area distribution; and the final (bottom) distribution represents the volume distribution.

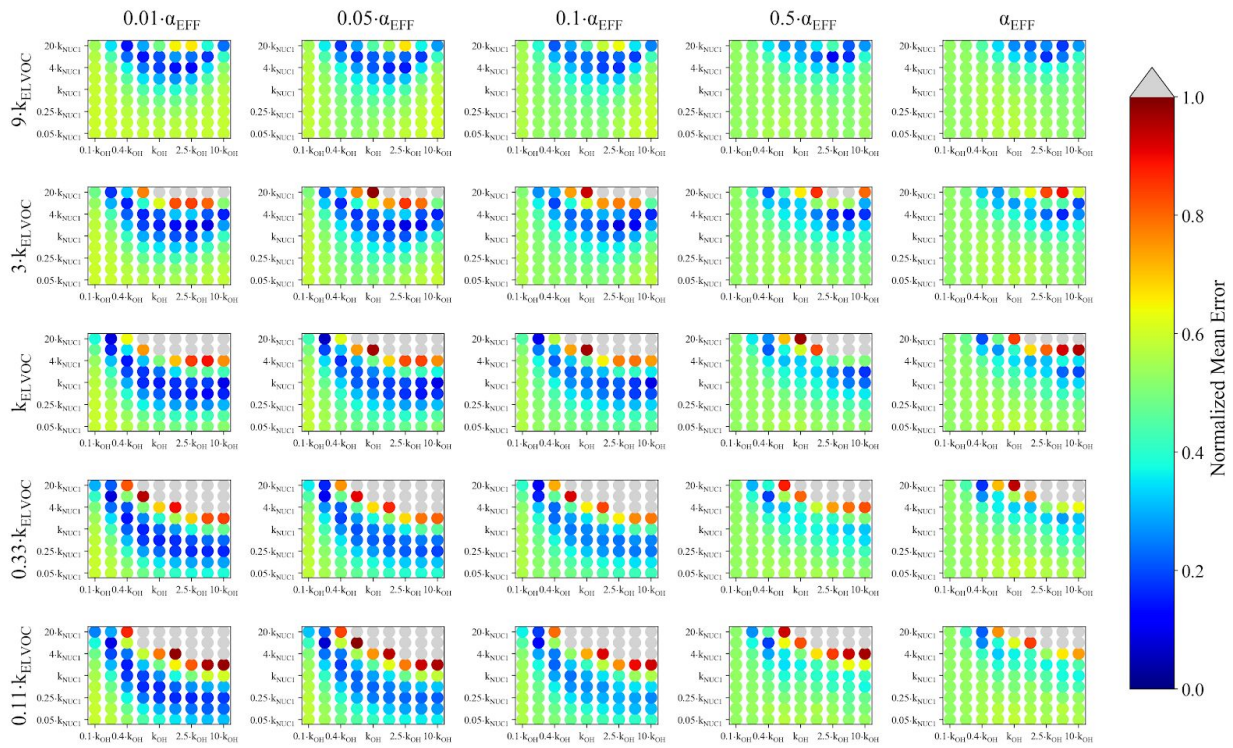

Figure S29. Representation of the parameter space for a 0.51 eq. day aging exposure from GoAmazon for the NUC1 nucleation scheme, base value of the reactive uptake coefficient of 0.6, and 30 times the base S/IVOC input concentration. The effective accommodation coefficient increases across each row of panels; the rate constant of gas-phase fragmentation increases up each column of panels. Within each panel, the rate constant of gas-phase reactions with OH increases along the x-axis and the rate constant for nucleation increases along the y-axis. The color bar indicates the normalized mean error (NME) value for each simulation, with the lowest values indicating the least error between model and measurement. Grey regions indicate regions within the parameter space whose NME value is greater than 1.

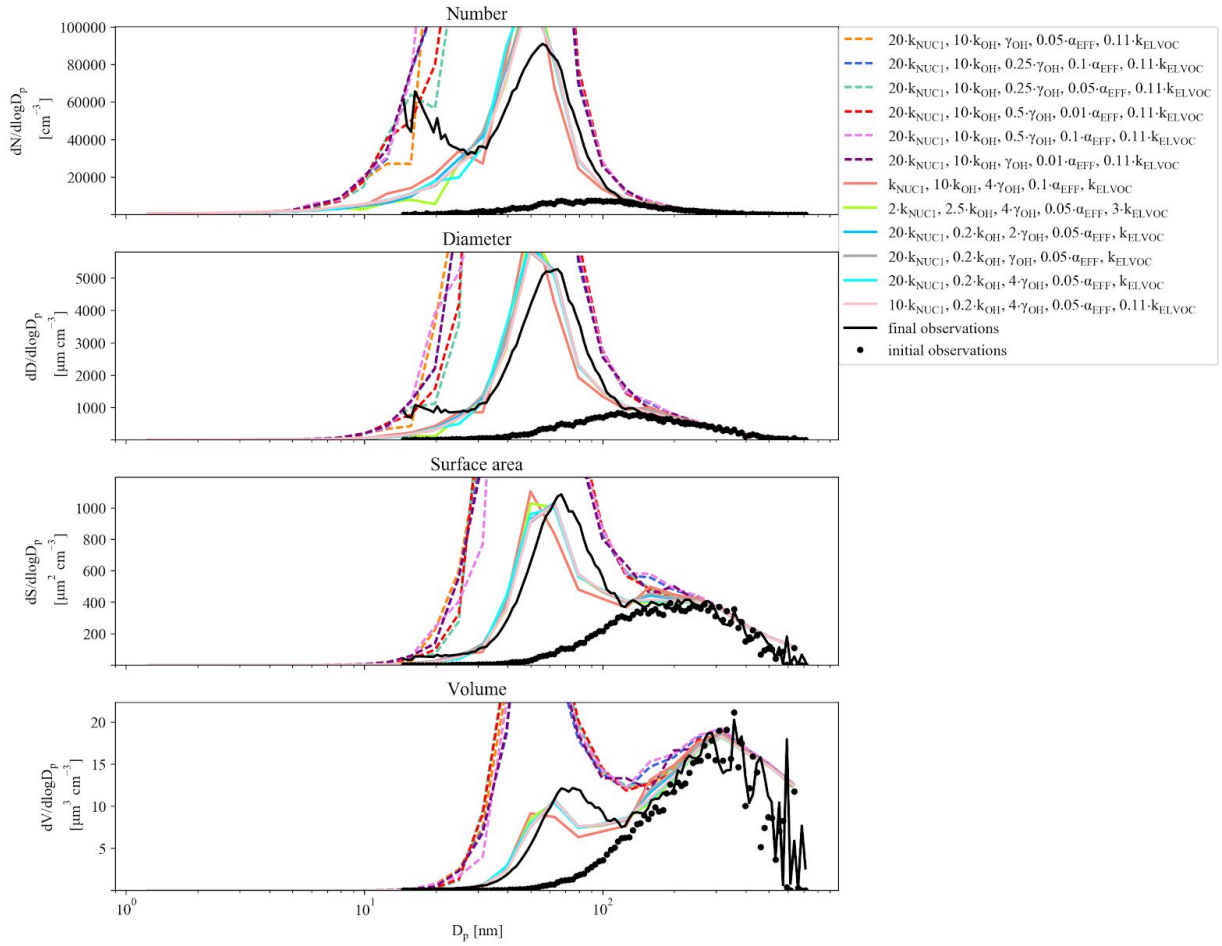

Figure S30. Example of best (solid lines) and worst (dashed lines) fit size distributions compared to the observed (black line) final OFR size distribution for a 0.51 eq days aging case from GoAmazon for the NUC1 nucleation scheme, using 30 times the base S/IVOC input concentration. The fits are determined using the mean error of moments method (see methods); each panel represents a separate moment. The top panel represents the number distribution; the second panel represents the diameter distribution; the third panel represents the surface area distribution; and the final (bottom) distribution represents the volume distribution.

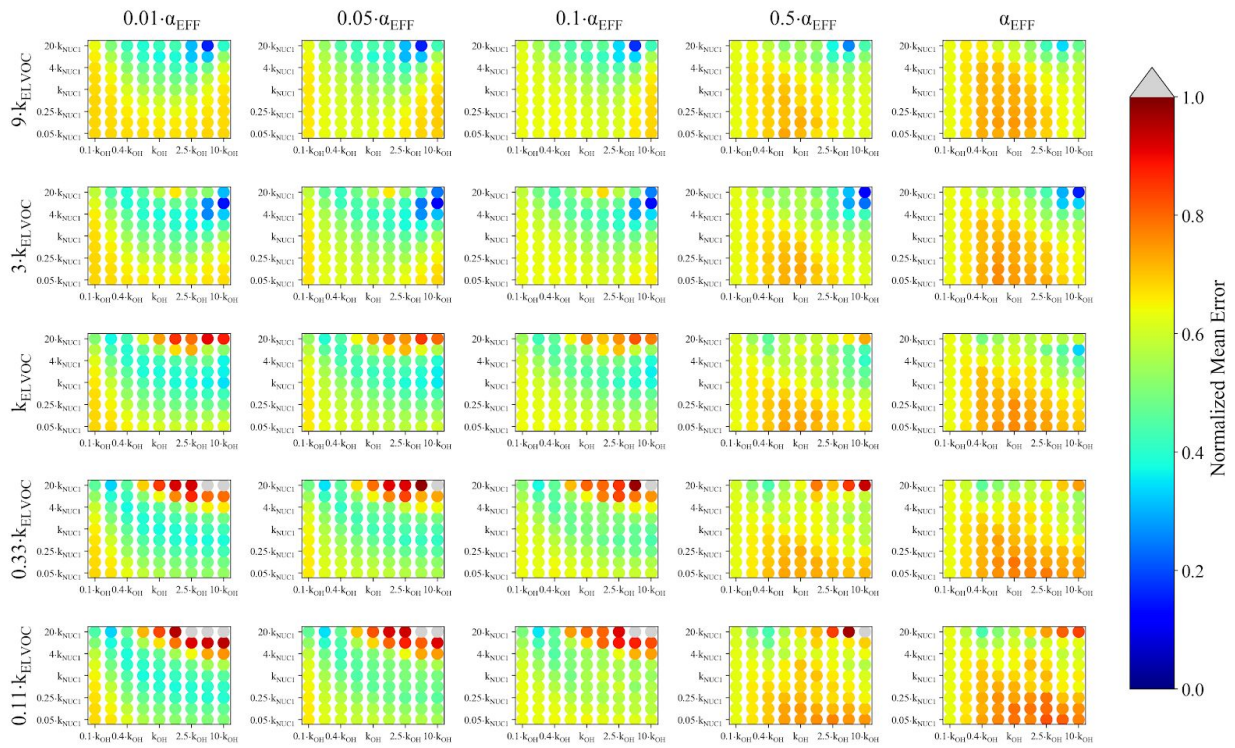

Figure S31. Representation of the parameter space for a 0.53 eq. day aging exposure from GoAmazon for the NUC1 nucleation scheme, base value of the reactive uptake coefficient of 0.6, and 30 times the base S/IVOC input concentration. The effective accommodation coefficient increases across each row of panels; the rate constant of gas-phase fragmentation increases up each column of panels. Within each panel, the rate constant of gas-phase reactions with OH increases along the x-axis and the rate constant for nucleation increases along the y-axis. The color bar indicates the normalized mean error (NME) value for each simulation, with the lowest values indicating the least error between model and measurement. Grey regions indicate regions within the parameter space whose NME value is greater than 1.

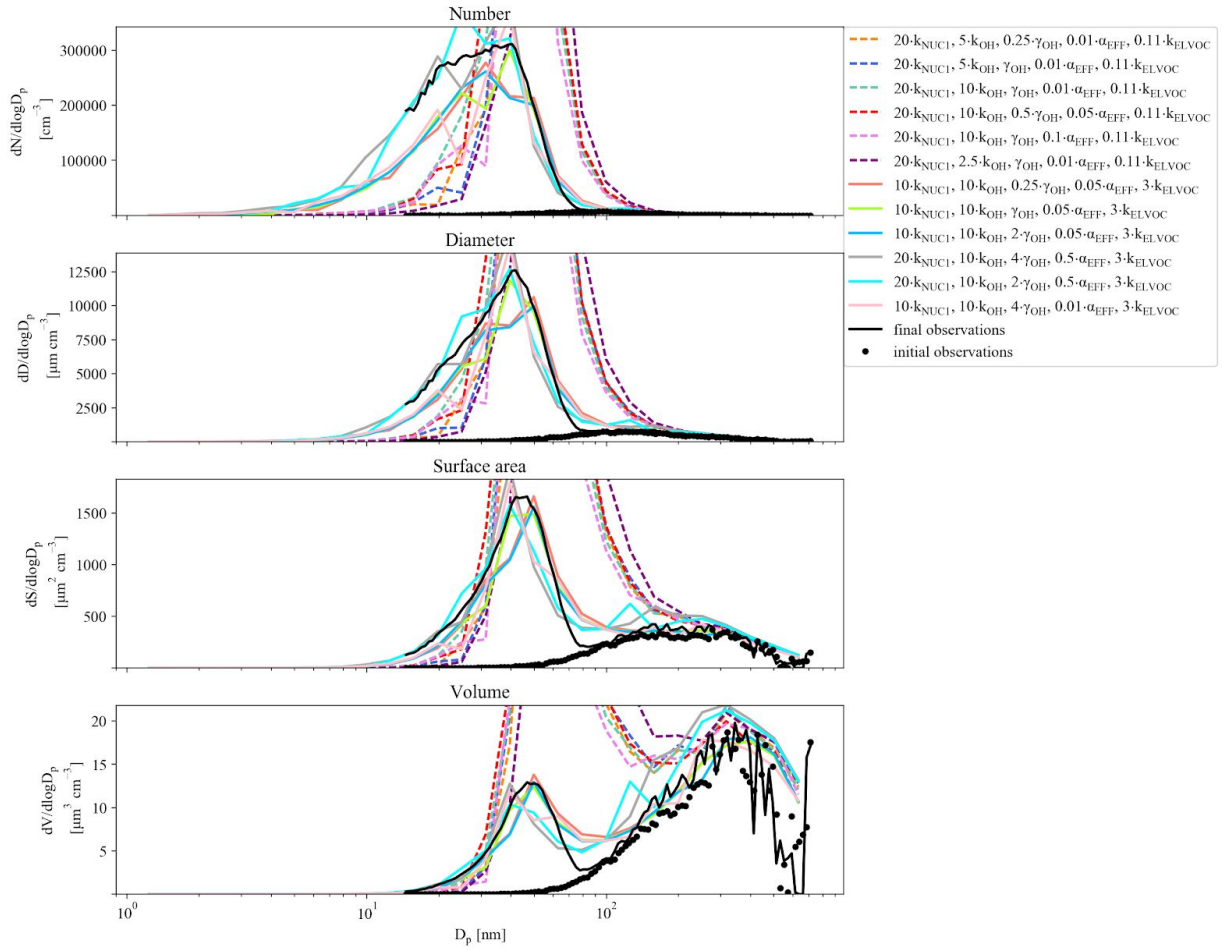

Figure S32. Example of best (solid lines) and worst (dashed lines) fit size distributions compared to the observed (black line) final OFR size distribution for a 0.53 eq days aging case from GoAmazon for the NUC1 nucleation scheme, using 30 times the base S/IVOC input concentration. The fits are determined using the mean error of moments method (see methods); each panel represents a separate moment. The top panel represents the number distribution; the second panel represents the diameter distribution; the third panel represents the surface area distribution; and the final (bottom) distribution represents the volume distribution.