Supplement of Atmos. Chem. Phys., 23, 10015–10034, 2023
https://doi.org/10.5194/acp-23-10015-2023-supplement
© Author(s) 2023. CC BY 4.0 License.

Atmospheric Chemistry and Physics
Open Access
EGU

Supplement of

What chemical species are responsible for new particle formation and growth in the Netherlands? A hybrid positive matrix factorization (PMF) analysis using aerosol composition (ACSM) and size (SMPS)

Farhan R. Nursanto et al.

Correspondence to: Juliane L. Fry (juliane.fry@wur.nl)

The copyright of individual parts of the supplement might differ from the article licence.

S1. Figures and tables

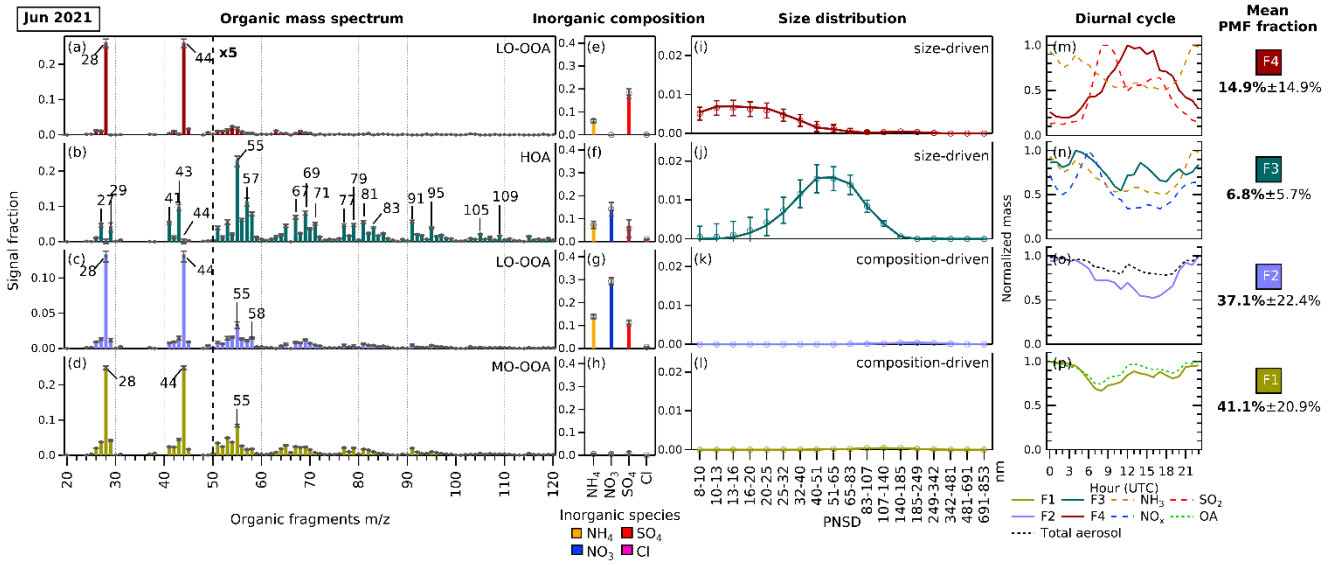


Figure S1. The profiles of 4-factor PMF solution from the combined ACSM-SMPS dataset in June 2021. Each factor is split into three matrices with their own rescaled signal fraction axes. The error bars in each variable represents standard deviation generated by performing bootstrapped run of the solution. The panel **a-d** shows the organic fragment mass spectrum from m/z 20 to 120 from ACSM (m/z < 20 not included). The panel **e-h** shows the ACSM standard inorganic aerosol species concentrations (ammonium (NH_4), nitrate (NO_3), sulfate (SO_4), and chloride (Cl)). The panel **i-l** shows the particle size distribution profiles from the SMPS. On the panel **m-p**, the diurnal cycles of the factors and related species are depicted. The mean PMF fractions and their standard deviations are shown indicating the mean contribution of each hybrid PMF factor to the "total variable reconstruction" by PMF throughout the period. The factors in May 2021 are assigned as: (F1) MO-OOA, (F2) NH_4+NO_3+SO_4+LO-OOA, (F3) size-driven NH_4+NO_3+SO_4+HOA, and (F4) size-driven NH_4+SO_4+LO-OOA.

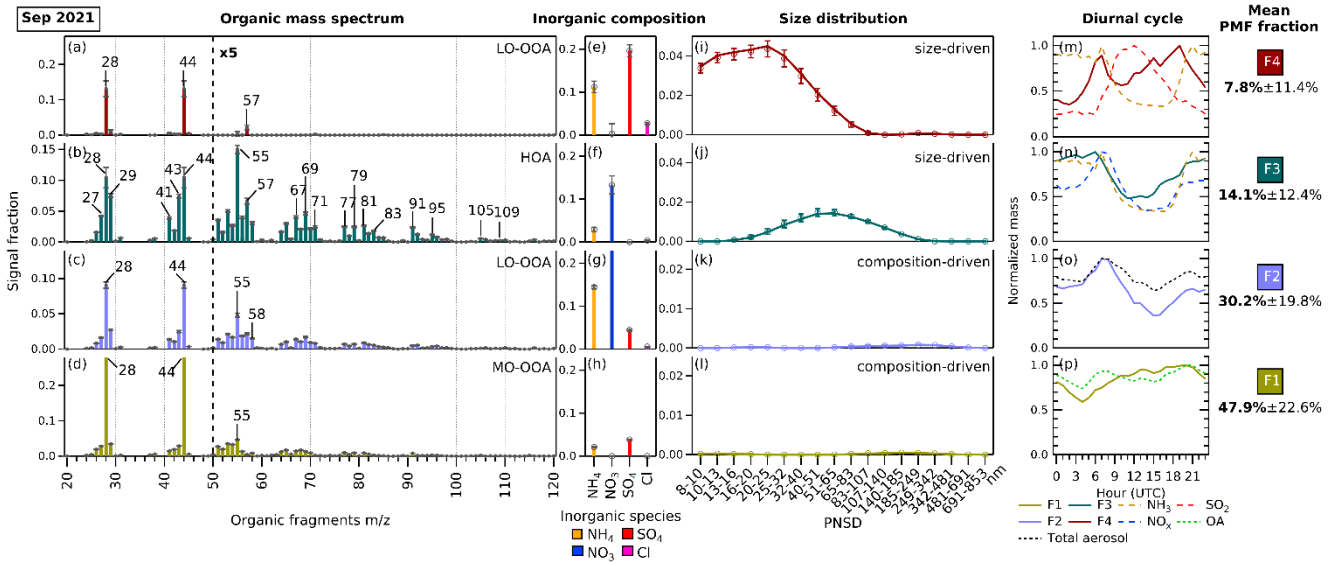


Figure S2. The profiles of 4-factor PMF solution from the combined ACSM-SMPS dataset in September 2021. Each factor is split into three matrices with their own rescaled signal fraction axes. The error bars in each variable represents standard deviation generated by performing bootstrapped run of the solution. The panel **a-d** shows the organic fragment mass spectrum from m/z 20 to 120 from ACSM (m/z < 20 not included). The panel **e-h** shows the ACSM standard inorganic aerosol species concentrations (ammonium (NH_4), nitrate (NO_3), sulfate (SO_4), and chloride (Cl)). The panel **i-l** shows the particle size distribution profiles from the SMPS. On the panel **m-p**, the diurnal cycle of the factors and related species are depicted. The mean PMF fractions and its standard deviation are shown indicating the mean contribution of each hybrid PMF factor to the "total variable reconstruction" by PMF throughout the period. The factors in May 2021 are assigned as: (F1) MO-OOA, (F2) NH_4+NO_3+ LO-OOA, (F3) size-driven NH_4+NO_3+HOA, and (F4) size-driven NH_4+SO_4+LO-OOA.

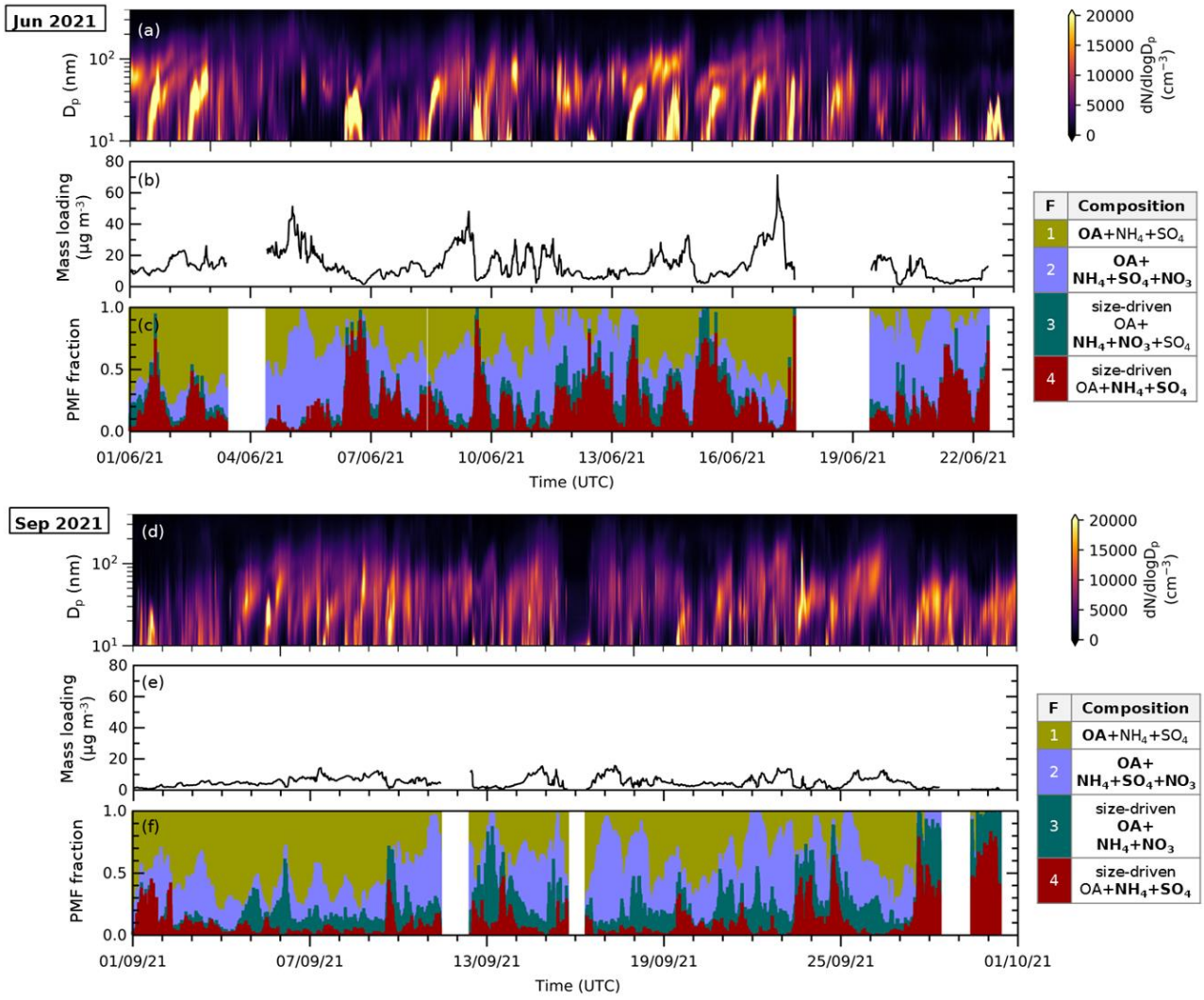


Figure S3. Time series of **(a,d)** particle size distribution ($dN/dlogD_p$) in cm^{-3} with logarithmic scale in particle size obtained from SMPS measurements, **(b,e)** total mass loading calculated from ACSM species concentration (using Tofware) in $\mu g.m^{-3}$, and **(c,f)** reconstructed PMF fraction (stacked) from analyses in June and September 2021. Episodes of F4 and F3 coincide with relatively low total aerosol mass conditions and high fine particle concentrations.

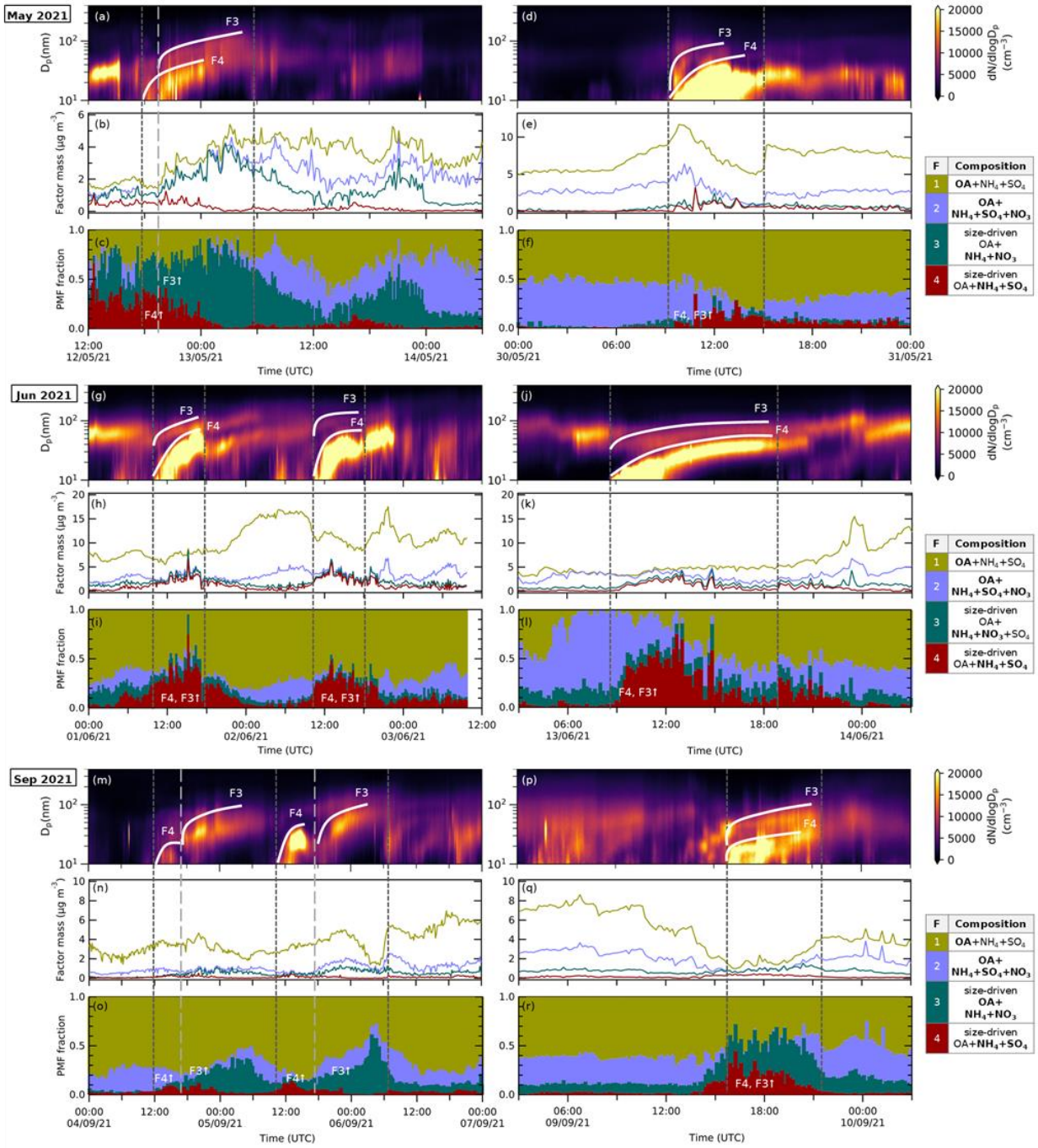


Figure S4. Selected time series windows during which new particle formation (NPF) events were detected by the scanning mobility particle sizer (SMPS) resembling 'banana' shapes in May 2021 **(a-f)**, June 2021 **(g-k)** and September 2021 **(m-r).** Rapid increases of fine new particles were observed. During these events, the increase in size-driven factors F4 and F3 can be seen equally in the reconstructed PMF fraction time series. The results suggest the events occur with either simultaneous episodes of F4 and F3 (growth-mode) **(d-f, g-i, j-l, p-r)**, or sequential episodes of both F4 and F3 **(a-c, m-o)**.

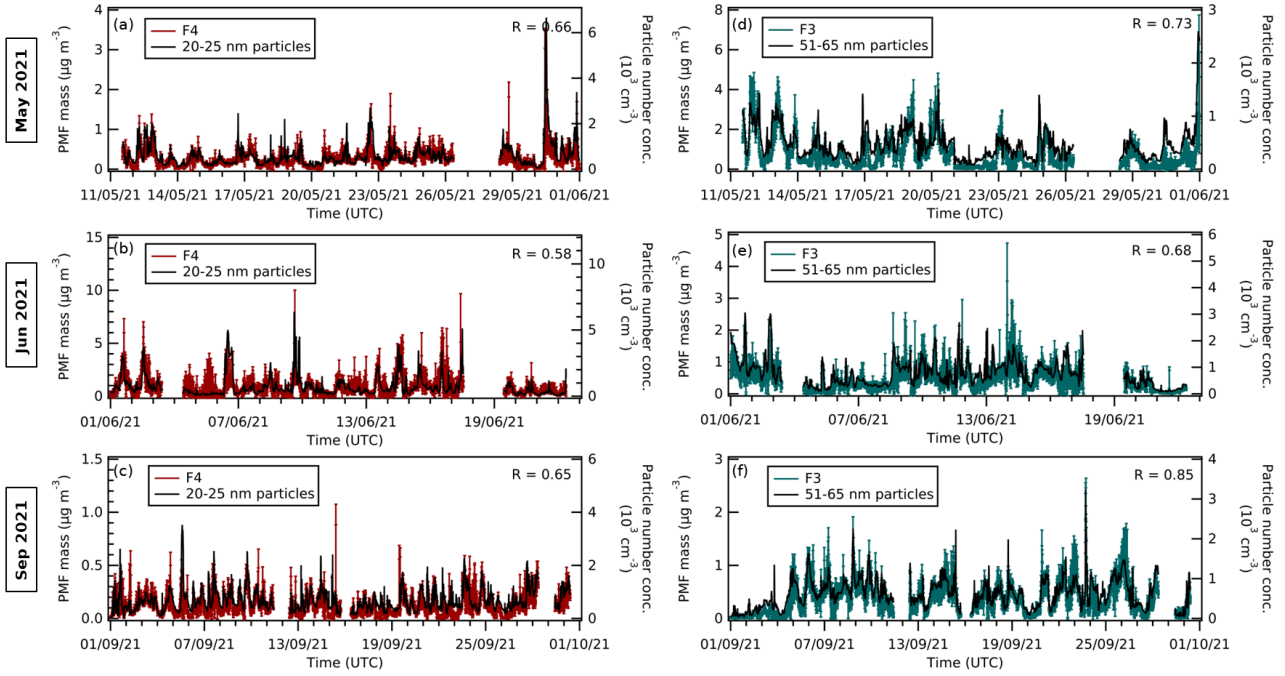


Figure S5. (a-c) The bootstrapped time series of F4 and number concentration of particles in the size bin of 20-25 nm and **(d-f)** the bootstrapped time series of F3 and number concentration of particles in the size bin of 51-65 nm in May 2021, June 2021, and September 2021. Both size bins are the median size of the size-driven factors' particle size distribution. The uncentered R Pearson coefficient shown on the right top of each graph shows the correlation between the time series of the factor and the particle number concentration.

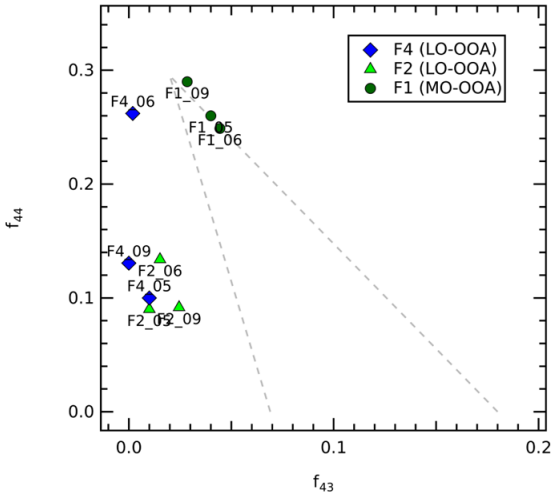


Figure S6. The triangle plot showing f_{44} vs f_{43} signal for resolved LO-OOA in F4 (dark blue, diamond), LO-OOA in F2 (light green, triangle), and MO-OOA in F1 (dark green, dots). The number indicated after the factor's number corresponds to the period (05 for May, 06 for June, and 09 for September). We note that the June F4 signature is closer to the MO-OOA ratios (see Sect. 3.3.3 for discussion). Moving to the top of the triangle means a more oxidized OA. As the measurements were carried on using capture vaporizer (CV) and the triangle plot was first developed for mass spectrometry with standard vaporizer (SV), some of the f_{44}/f_{43} values are outside the triangle.

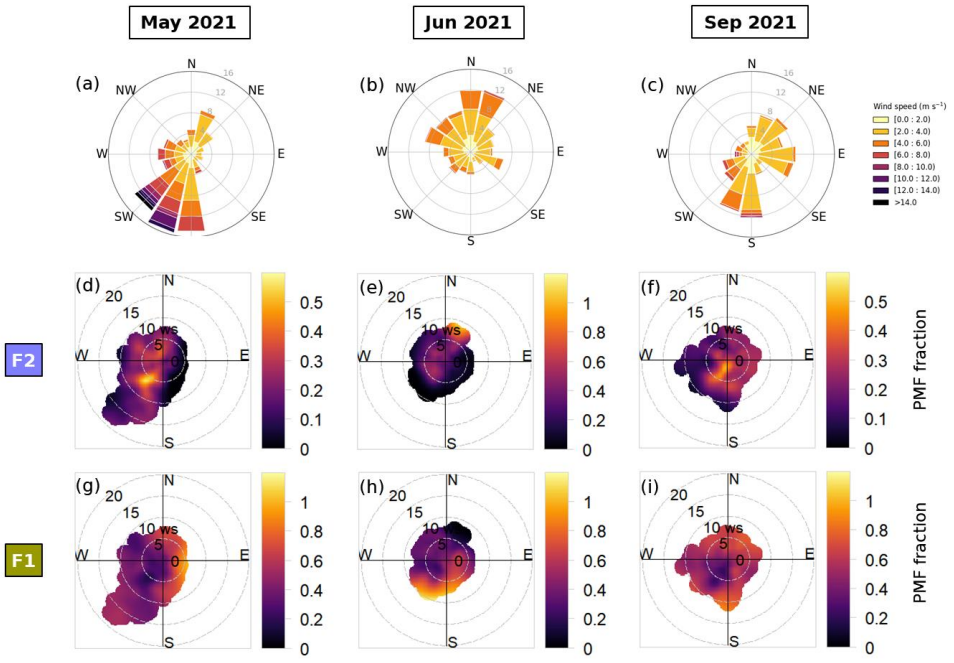


Figure S7. (a-c) Wind roses and **(d-i)** bivariate polar plots of composition-driven factor mass fraction (F2 and F1) by wind speed and wind direction measured in Cabauw in May, June, and September 2021. The bulk OA+IA factor (F2, **d-f**) is shown to be originated from different directions, while the bulk OA factor (F1, **g-i**) is mainly coming from the easterly and southerly sector.

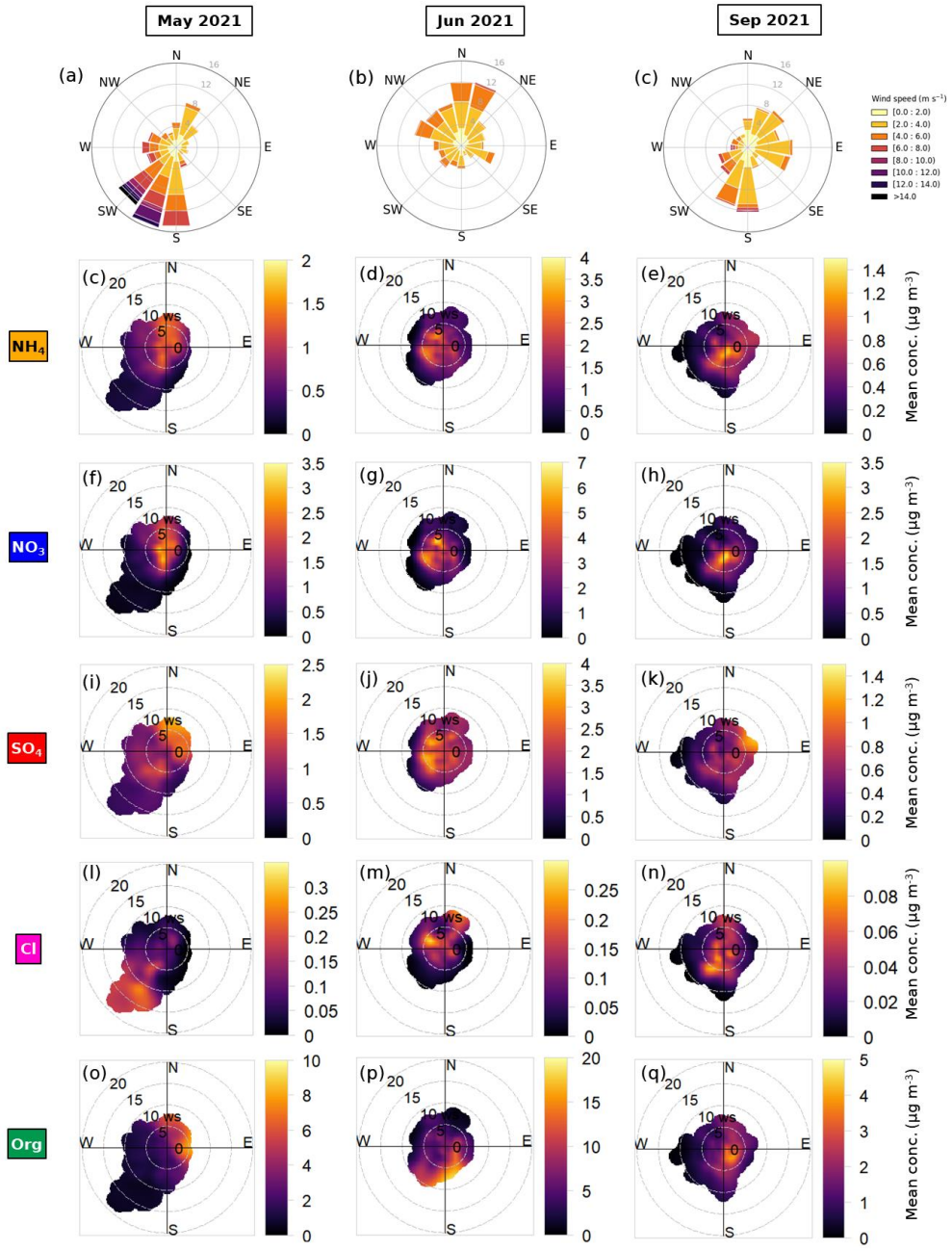


Figure S8. (a-c) Wind roses and **(d-q)** bivariate polar plots of ACSM-generated chemical species by wind speed and wind direction measured in Cabauw in May, June, and September 2021. The species include ammonium (NH_4, **c-e**), nitrate (NO_3, **f-h**), sulfate (SO_4, **i-k**), chloride (Cl, **l-n**), and organics (Org, **o-q**) in 10-minute resolution. NO_3 and SO_4 are found to be correlated to NH_3 source direction (see Fig. S9) across seasons due to formation of ammonium nitrate and sulfate. NO_3 are correlated with the southern and western sector of the site. SO_4 is correlated to westerlies and north-easterlies, influenced by both NH_3 and SO_2 source direction since they react forming ammonium sulfate (Fig. S9).

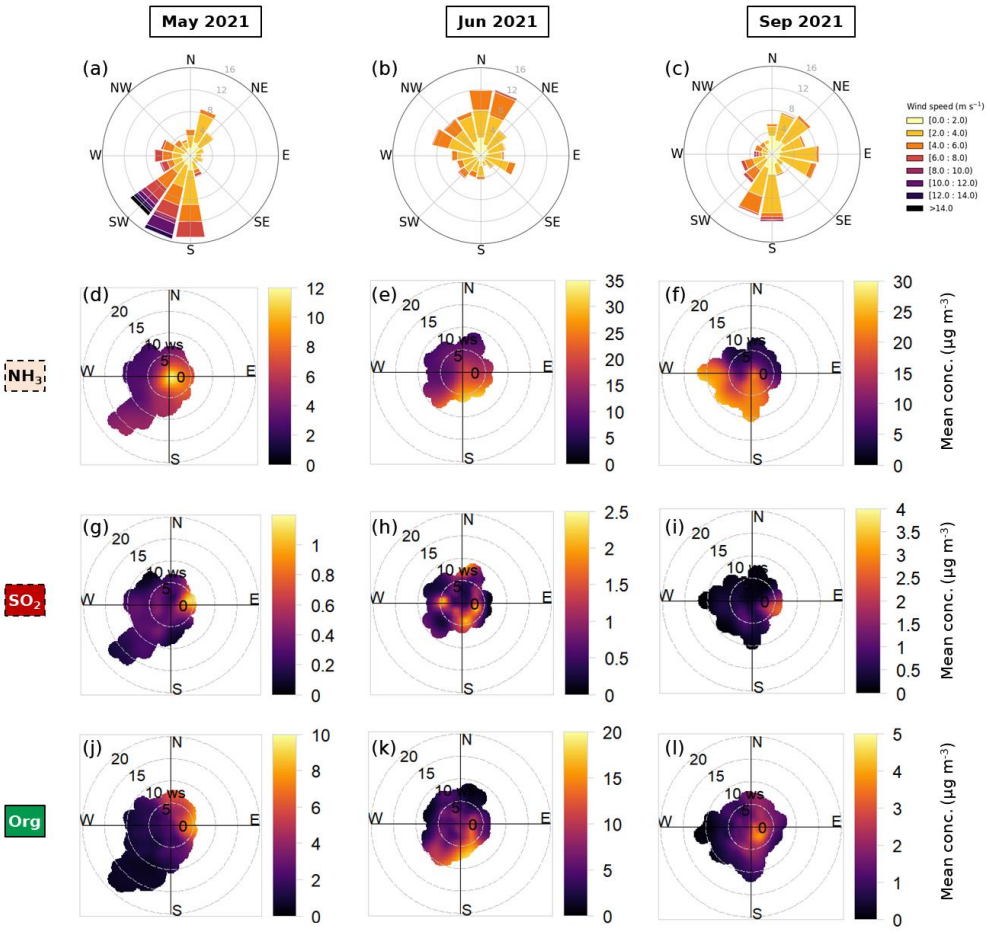


Figure S9. (a-c) Wind roses and **(d-l)** bivariate polar plots of species concentrations related to F4 by wind speed and wind direction measured in Cabauw in May, June, and September 2021. The species include ammonia (NH_3, **d-f**) and sulfur dioxide (SO_2, **g-i**) in the gaseous phase in hourly resolution, and organics (Org, **j-l**) in the aerosol phase in 10-minute resolution. NH_3 is predominant around the site, particularly from the southern sector. SO_2 is prevalent from the east and west of the site. Org species are predominantly coming from the eastern sector explaining the correlation of nucleation-mode factor (F4) episodes with easterlies in May and June 2021 when OA participates in early NPF.

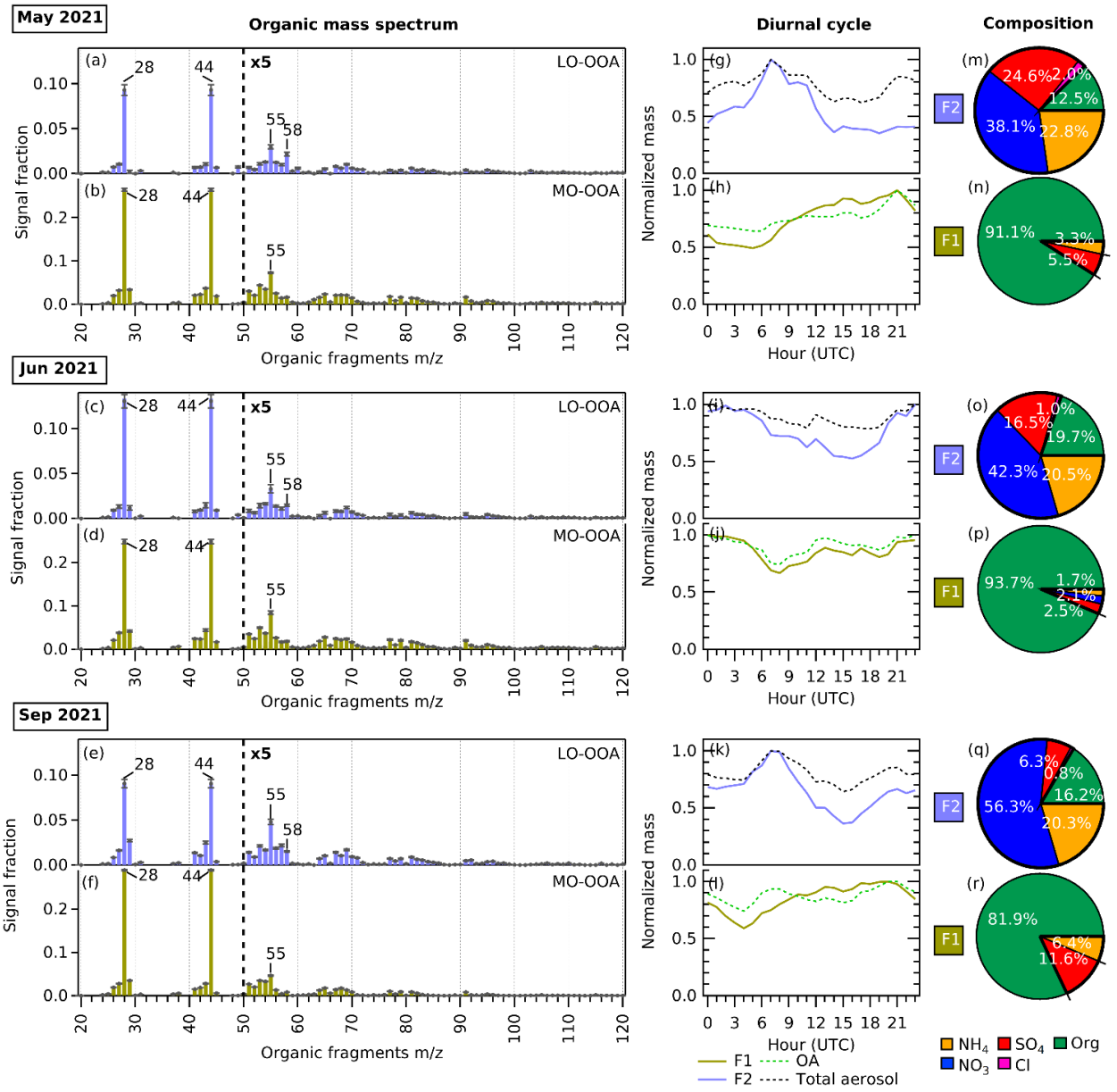


Figure S10. (a-f) Organic mass spectrum from m/z 20 to 120 (m/z < 20 not included) of F2 (purple) showing LO-OOA factor profiles and F1 (lime green) showing MO-OOA factor profiles in May 2021, June 2021, and September 2021. The error bars in each m/z were generated from bootstrap run. **(g-l)** The diurnal cycles of F2, consisting of OA and IA mixture, follows the diurnal pattern of total bulk aerosol pattern, while F1, mainly consisting of OA, follows the diurnal pattern of bulk OA measured by ToF-ACSM (Org). The MO-OOA F1 concentration increase slightly during the day due to photochemical activity and then concentrated in the boundary layer after sunset. It decreases around the midnight until morning due to absence of photochemistry. MO-OOA in F1 shows a more stable diurnal cycle pattern with some increase from the morning to afternoon due to photochemistry. The concentration is accumulated in the evening due to shallow boundary layer and then diminishes in the early morning due to absence of photochemistry. **(m-n)** Pie charts showing mass percentage of each aerosol species contributing to each composition-driven factor in May 2021, June 2021, and September 2021. Green represents organics (Org), orange represents ammonium (NH_4), dark blue represents nitrate (NO_3), dark red represents sulfate (SO_4), and pink represents chloride (Cl). F4 are dominated by ammonium sulfate while F3 are dominated by ammonium nitrate.

S2. Atmospheric composition and aerosol formation

Ion balance ratio/ammonium balance

The ion balance ratio, or also called ammonium balance ($NH_{4\_bal} = n_{NH_4}/(n_{NO_3}+2\times n_{SO_4}+n_{Cl})$), is used to assess the ion-pairing between cations and anions forming inorganic aerosol (IA). When the $NH_{4\_bal} < 1$, nitrate or sulfate may be present in compounds other than ammonium salts. The anions can form refractory compounds with other cations (e.g., $NaNO_3$, Na_2SO_4, $Ca(NO_3)_2$) and thus be less efficiently detected by the spectrometer, or exist in form of organic nitrates and sulfates, yielding an additional amount of nominal nitrate and sulfate measured by ACSM (Farmer et al., 2010; Docherty et al., 2011). It also may be caused by excess aerosol acidity (Farmer et al., 2010; Docherty et al., 2011) although the clear relationship between NH_{4_bal} and high aerosol acidity (pH < 0) is observed for mass spectrometry measurements only when $NH_{4\_bal} < 0.65$ (Schueneman et al., 2021). On the other hand, when the $NH_{4\_bal} > 1$, uncharged amines may be detected as ammonium by the ACSM (Farmer et al., 2010; Docherty et al., 2011), or unmeasured anions such as carboxylates may be present. The ion balance ratio in spring (May), summer (June), and autumn (September) are 0.997, 0.986, and 1.066 respectively, which can be regarded as close to unity. Therefore, we can infer that the bulk aerosol charge is fully neutralized for all periods.

Table S1. Mean bulk atmospheric chemical composition in the three period as detected by the ToF-ACSM and co-located gas measurements in Cabauw.

Species/ratio	Mean concentration in May 2021 (spring)		Mean concentration in Jun 2021 (summer)		Mean concentration in Sep 2021 (autumn)	
	Mass ($\mu g\ m^{-3}$)	Molar (10^{-4} mol m^{-3})	Mass ($\mu g\ m^{-3}$)	Molar (10^{-4} mol m^{-3})	Mass ($\mu g\ m^{-3}$)	Molar (10^{-4} mol m^{-3})
NH_3[(a)]	7.03	41.37	15.25	89.56	12.67	74.37
$NO_x = NO+NO_2$	9.55	21.79	10.90	24.90	14.78	34.08
SO_2	0.34	0.54	0.85	1.33	0.59	0.91
Total aerosol[(b)]	6.60	-	14.12	-	5.15	-
Org	2.49	-	6.36	-	2.18	-
NO_3^-	1.84	2.97	3.42	5.52	1.60	2.58
NH_4^+	1.01	5.61	1.90	10.54	0.75	4.17
SO_4^{2-}	1.15	1.20 (×2 = 2.40)	2.29	2.39 (×2 = 4.78)	0.56	0.58 (×2 = 1.16)
Cl^-	0.11	0.30	0.15	0.42	0.05	0.14
NH_{4_bal}[(c)]	-	0.997 ± 0.001	-	0.986 ± 0.001	-	1.066 ± 0.001
m_{Org}/m_{IA}[(d)]	0.61	-	0.82	-	0.74	-

[(a)] Data taken from Zegveld-Oude Meije measurement station

[(b)] Total mass of aerosol detected by ToF-ACSM ($m_{Org}+m_{NO_3}+m_{NH_4}+m_{SO_4}+m_{Cl}$)

[(c)] Ion balance ratio/ammonium ratio is the ratio between the measured ammonium (n_{NH_4}) and the total ammonium needed to neutralize the major anions ($n_{NO_3}+2\times n_{SO_4}+n_{Cl}$). The value is obtained from linear regression.

[(d)] Organic-to-inorganic mass ratio is the ratio between the measured organic aerosol mass (m_{Org}) and the measured inorganic aerosol mass ($m_{NO_3}+m_{NH_4}+m_{SO_4}+m_{Cl}$). The concentration is obtained from ToF-ACSM measurements.

Ammonium sulfate and nitrate aerosol formation

The ammonium sulfate and nitrate aerosol formation can be explained from the chemistry of a HNO_3-NH_3-H_2SO_4-H_2O system (Ansari and Pandis, 1998; Seinfeld and Pandis, 2016; Weber et al., 2016; Pye et al., 2020). The formation of ammonium sulfate and nitrate aerosols involves the buffering capacity of semi-volatile NH_3 partitioning between the gas and particle phase, reacting with H_2SO_4 and HNO_3 and forming aerosols. Coming from SO_x emission, H_2SO_4 reacts with NH_3 forming a mixture of ammonium bisulfate (NH_4HSO_4), ammonium sulfate ($(NH_4)_2SO_4$), and leaving some free gas-phase NH_3 in the equilibrium. As sulfate has an extremely low vapor pressure and ammonium are semi-volatile, NH_4^+/NH_3 acts as a mobile species between gas and aerosol phase, establishing equilibrium between the two phases. With the decrease of SO_x emissions and steady NH_3 emissions, NH_4^+ will be converted back to gas-phase NH_3 favoring the formation of NH_4HSO_4 and the release H^+, keeping the aerosol pH stable (0–3) (Weber et al., 2016). The buffering effect of NH_4^+/NH_3 is also reflected in the stable ion balance ratio around unity as observed across seasons in this study (see Table S1). The equilibrium mechanism to keep the aerosol pH from rising is expected to occur despite SO_x emission reduction, unless the atmospheric sulfate concentration goes below 0.3 μg m^{-3}, where sulfate starts to be associated with non-volatile cations (Weber et al., 2016). With the addition of HNO_3 in the system coming from NO_x emission, ammonium nitrate (NH_4NO_3) is also formed by either the association of free gas-phase NH_3 or NH_4^+ in the aerosol phase moving to the gas-phase. Since HNO_3 is more volatile than H_2SO_4, ammonium nitrate is moves more freely between the gas and aerosol phase to buffer the aerosol acidity and reach equilibrium. As the buffering system relies on chemical species volatility, meteorological conditions like temperature and relative humidity also play role in the formation of ammonium sulfate and nitrate aerosol.

In summary, from the chemistry perspective, ammonium sulfate is more likely to become condensation nuclei due to low volatility nature of sulfuric acid favouring the aerosol phase, whereas ammonia or amines (as bases) stabilize it. In consequence, they are more likely to occur in smaller particles. Nitric acid plays role in particle growth as it condenses to the particle phase at smaller rate, and serve more as bridge to facilitate the association ammonium sulfate particle clusters in early NPF (Liu et al., 2018). The association of semi-volatile organic vapors to the condensation nuclei occurs through acid-base chemistry and accretion of organic molecules (Zhang et al., 2004; Hodshire et al., 2016), building up the particle mass throughout aerosol formation and growth.

S3. PMF analysis

PMF variables downweighting

Downweighting is an important step in hybrid PMF input matrix preparation to ensure the variables have similar magnitude over others, or to remove duplicated variables in the matrix. Downweighting is done first to the species mass concentrations and particle number concentrations. The values and errors of species mass concentrations and particle number concentrations

in the matrix ($X_{ij,input}$) were downweighted by dividing the dataset with a downweighting constant (DWC) to get the final values and errors ($X_{ij,DW}$),

$$X_{ij,DW} = \frac{X_{ij,input}}{DWC}$$

The goal of downweighting is to normalize the magnitude of newly introduced variables in comparison to the organic mass spectrum, by decreasing their magnitude in reference to the spectrum. The dataset from September 2021 is used for the calculation as it represents the lowest concentration of all chemical species among analyzed periods. The 95% percentile concentration is selected for the calculation to avoid using outlier peak in the dataset. The concentration of m/z 44 fragment (C_{f44}) is chosen to represent the organic mass spectrum dataset for the calculation of DWC as it generally has the highest average peak among organic fragments. For the inorganic mass concentration, nitrate concentration (C_{NO_3}) is selected as the reference for the inorganic chemical species since it generally has the highest concentration among measured inorganics. The PNSD in the 51-65 nm size bin ($C_{p51\text{-}65}$) is chosen for the particle size distribution, also for the same reason.

The PNSD has different unit compared to other variables (cm^{-3} instead of $\mu g\ m^{-3}$). However, we disregard the unit as we are only interested in seeing the particle size concentration variation over time, and how the sizes are being distributed in the profiles in the PMF solution. In this study, the DWC for PNSD variables was multiplied by 100 to tune down the number distribution values relative to the inorganic species. This empirically tuned multiplier factor is chosen because it gives convergent PMF solutions.

Table S2. The determination of downweighting constant (DWC) for inorganic species mass species and particle number size distribution (PNSD) in the PMF input matrix prior to analysis. The values are obtained from September 2021 dataset and applied for all analyzed periods.

Variable	Value
$C_{f44;95\%}$ ($\mu g\ m^{-3}$)	1.43
$C_{NO_3;95\%}$ ($\mu g\ m^{-3}$)	4.95
$C_{p51\text{-}65;95\%}$ (cm^{-3})	1198.34
$\boldsymbol{DWC\ inorganic = \frac{C_{NO_3;95\%}}{C_{f44;95\%}}}$	**3.45**
$\boldsymbol{DWC\ PNSD = 100 \times \frac{C_{p51\text{-}65;95\%}}{C_{f44;95\%}}}$	**83527.07**

The information in m/z 44 is duplicated among different organic fragments in the default fragmentation table (Allan et al., 2004) and therefore also need to be downweighted. The downweighting procedure is applied for m/z 44, 28, 18, 17, and 16 signals in the organic mass spectrum as provided by PETv3.08 during PMF input matrix preparation (Ulbrich et al., 2009). A

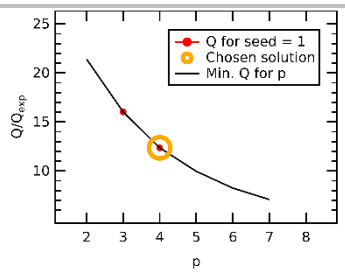
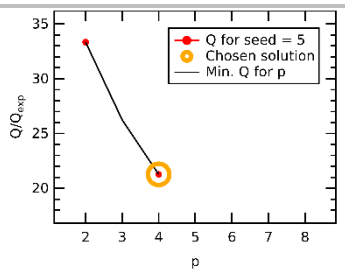
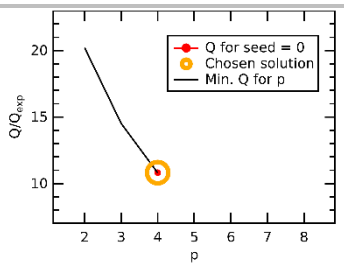
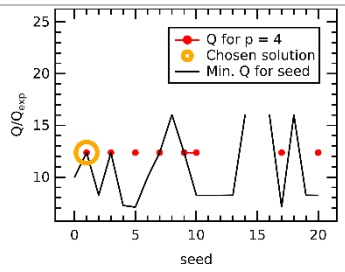
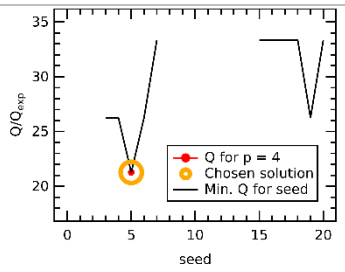
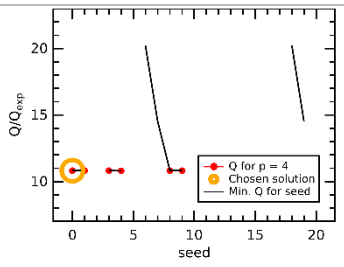
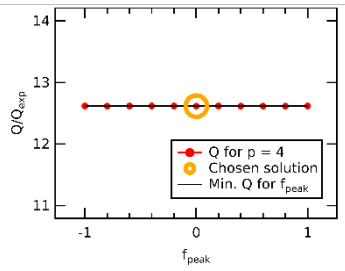
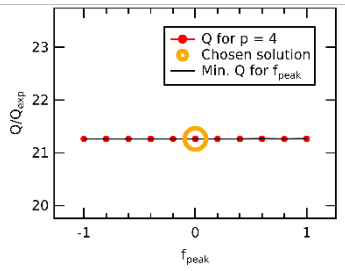
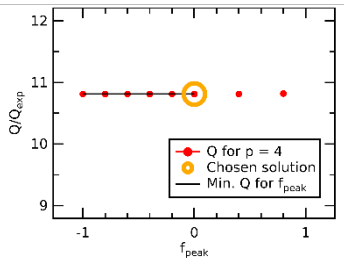
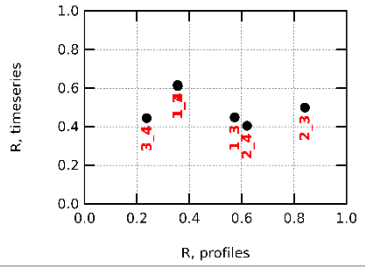
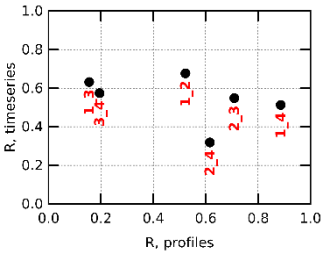
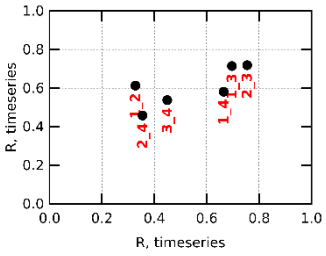
correction calculation for capture vaporizer (CV) is also opted at the end of the PMF analysis to account the additional thermal decomposition in smaller fragments.

Determination of PMF solution

PMF analysis in each period were run unconstrained by varying the number of factors (*p*) from 2 to 8 factors and using varied seed values (min = 0; max = 20; delta = 1). The optimum *p* is selected first based on the lowest residuals and local minima (Q/Q_{exp}) of the PMF solutions. With the introduction of inorganic species mass concentration and particle number size distribution (PNSD) variables, we suggest that the PMF solution also must include at least two factors that show a significant signal of particle size distribution, or size-driven factors. A minimum of two size-driven is required in order to study new particle formation and growth. Lastly, the lowest Q/Q_{exp} should not be accepted if it contains redundant factors with very similar organic and inorganic profile. The diurnal cycle of each factor and correlation with other chemical species (gas and aerosol species) are also done to confirm the PMF solution. After the optimum p and seed value are chosen, the rotationality of the solution is explored by performing PMF run with varied rotation (f_{peak}) values (min = -1, max = +1, delta = 0.2).

For each period, a 4-factor solution was picked as the best fit (see Table S3 for the statistical summary). The measured and total reconstructed PMF mass time series, total residuals time series, and scaled residuals of PMF variables are shown in Fig. S11 (May), Fig. S12 (June), and Fig. S13 (September). We can observe that the PMF solutions that were solved by PMF analysis from hybrid PMF input have reduced number of converging solutions unlike regular PMF analysis with only organic spectrum. As we introduce extra variables to the input, more constraints are introduced to the PMF analysis, and thus increasing the number of variables that needs to be considered into the PMF model fit, limiting the number of convergences. NO_3 and SO_4 represent the largest contributor of residuals as can be seen in the scaled residual plots in Fig. S11 (May), Fig. S12 (June), and Fig. S13 (September).

Table S3. Diagnostic plots of hybrid PMF analyses in May 2021, June 2021, and September 2021, combining CV-ToF-ACSM organic mass spectrum, CV-ToF-ACSM inorganic species concentration, and TROPOS-SMPS PNSD measurements: (a) Q/Q_{exp} vs. number of factors, (b) Q/Q_{exp} vs. seed value, (c) Q/Q_{exp} vs. f_{peak} value, and (d) correlation of time series and mass spectra among PMF factors (R time series vs. R profiles). The number of factors, seed value, f_{peak} value, Q/Q_{exp}, and the reasoning of the chosen PMF solutions are listed below the plots. If not stated, the default values for seed and f_{peak} are zero.

Plots*/ values chosen	**May 2021**	**Jun 2021**	**Sep 2021**
(a) Q/Q_{exp} vs. *p*			
(b) Q/Q_{exp} vs. seed			
(c) Q/Q_{exp} vs. f_{peak}			
(d) Un-centered R time series vs. R profiles			
p	4	4	4
seed	1	5	0
f_{peak}	0 (no significant change in rotational changes)	0 (no significant change in rotational changes)	0 (no significant change in rotational changes)
Q/Q_{exp}	12.62	21.27	10.82

Comments	Although solutions with p > 4 have lower Q/Q_{exp}, 4-factor solution was chosen because it does not generate redundant factors with same organic and inorganic compositions unlike factors with p > 4. It resolves two size-driven factors, and every factor has distinctive particle size distribution and inorganic composition. One POA (HOA) and three OOA factors (one MO-OOA and two LO-OOA) are also generated.	This PMF solution was chosen because it is solution with the highest number of factors that converged. It resolves two size-driven factors, and every factor has distinctive particle size distribution and inorganic composition. One POA (HOA) and three OOA factors (one MO-OOA and two LO-OOA) are also generated.	This PMF solution was chosen because it is solution with the highest number of factors that converged. It resolves two size-driven factors, and every factor has distinctive particle size distribution and inorganic composition. One POA (HOA) and three OOA factors (one MO-OOA and two LO-OOA) are also generated.

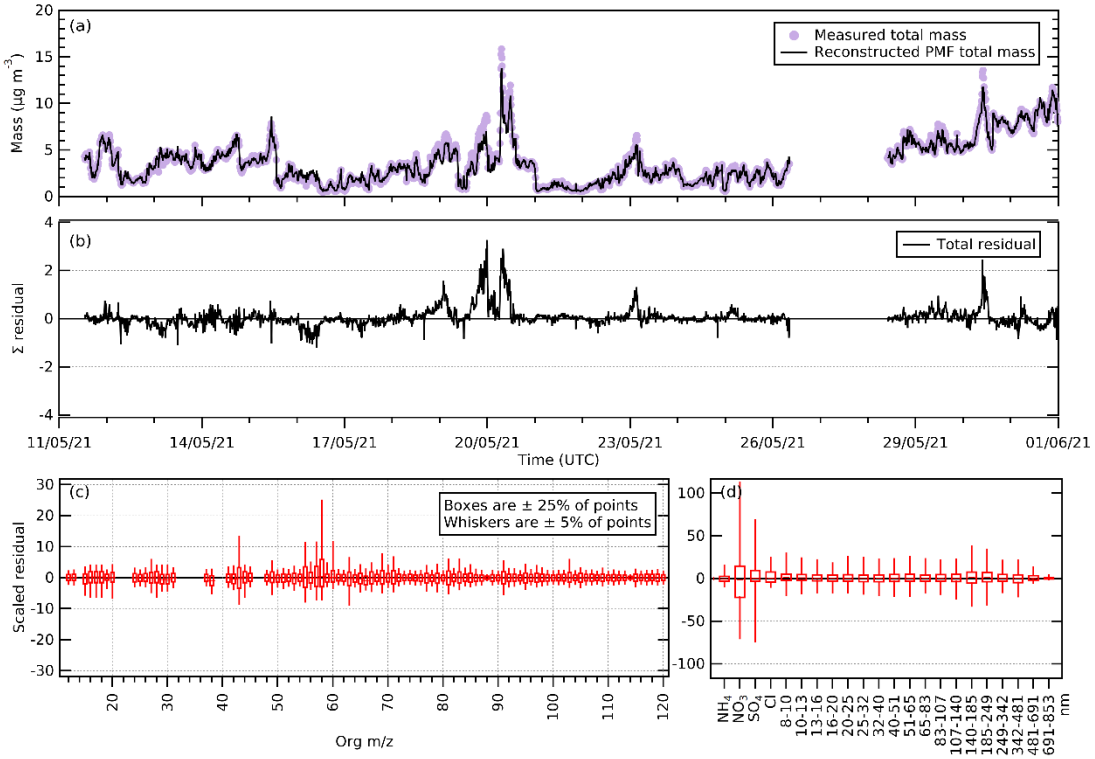


Figure S11. Diagnostic plots of hybrid PMF analysis combining CV-ToF-ACSM organic mass spectrum, CV-ToF-ACSM inorganic species mass concentration, and TROPOS-SMPS PNSD measurements in May 2021: (a) time series of the measured and reconstructed PMF mass, (b) time series of residual of the least-square-fit, (c) distribution of scaled residuals for each m/z, and (d) distribution of scaled residuals for each inorganic species and size-binned particle number concentration.

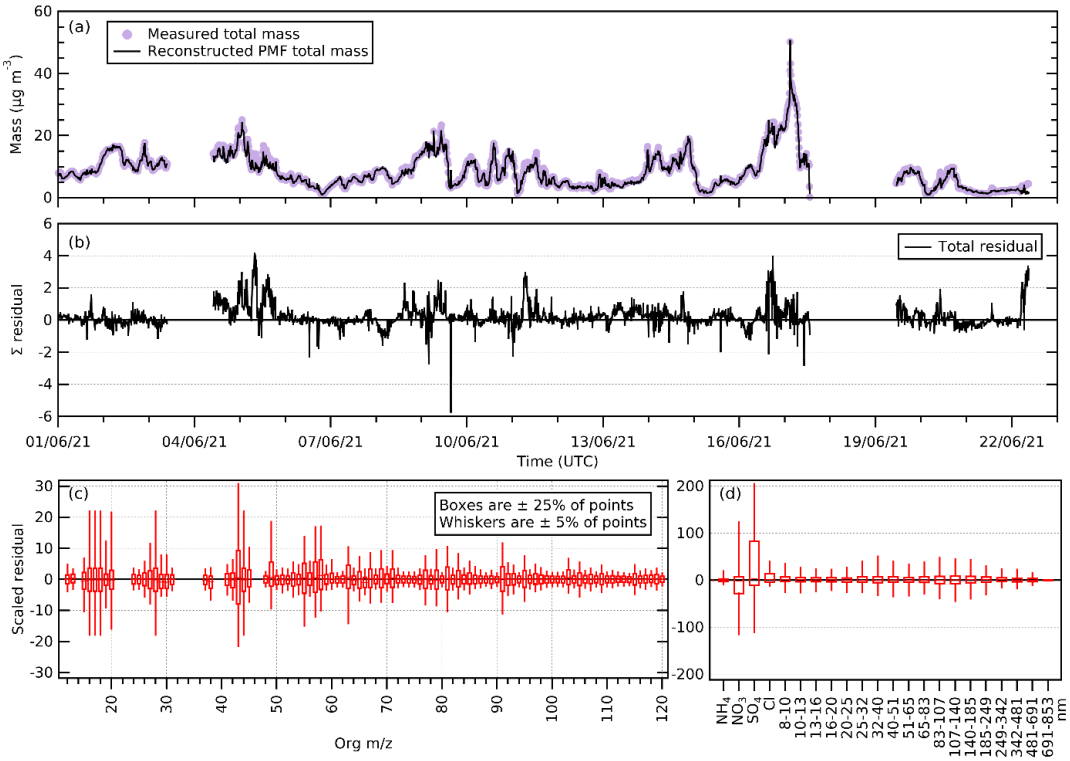
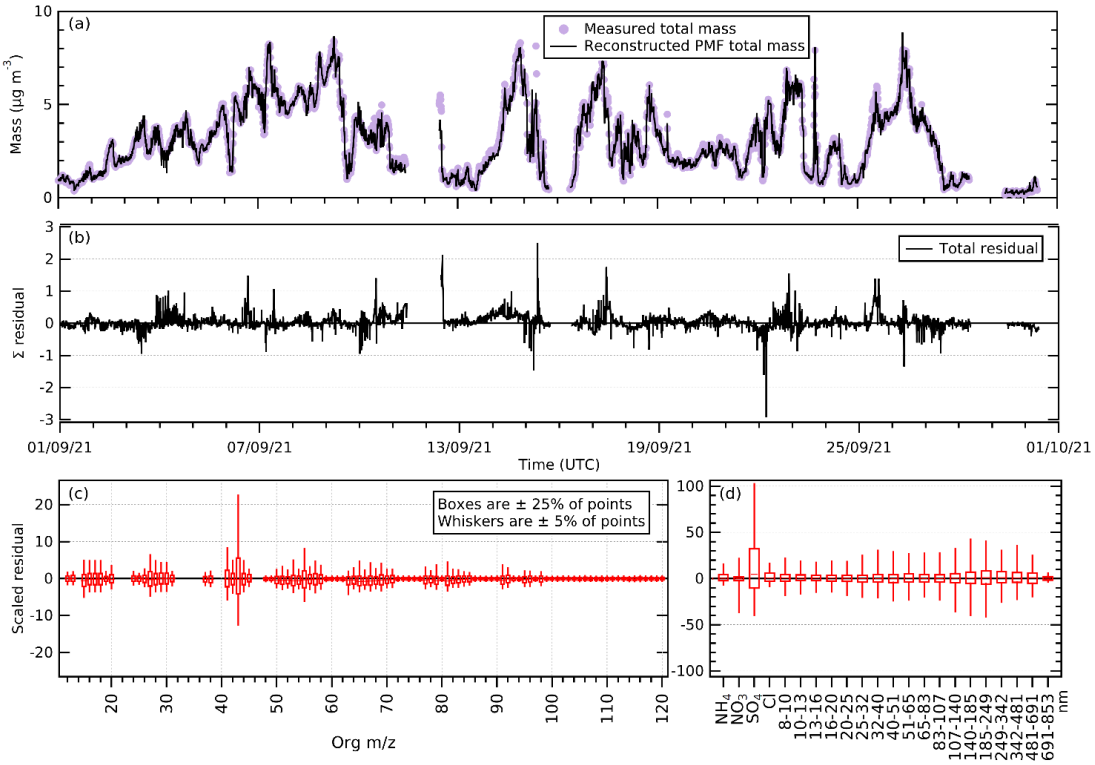


Figure S12. Diagnostic plots of hybrid PMF analysis combining CV-ToF-ACSM organic mass spectrum, CV-ToF-ACSM inorganic species mass concentration, and TROPOS-SMPS PNSD measurements in June 2021: (a) time series of the measured and reconstructed PMF mass, (b) time series of residual of the least-square-fit, (c) distribution of scaled residuals for each m/z, and (d) distribution of scaled residuals for each inorganic species and size-binned particle number concentration.



195 **Figure S13.** Diagnostic plots of hybrid PMF analysis combining CV-ToF-ACSM organic mass spectrum, CV-ToF-ACSM inorganic species mass concentration, and TROPOS-SMPS PNSD measurements in September 2021: (a) time series of the measured and reconstructed PMF mass, (b) time series of residual of the least-square-fit, (c) distribution of scaled residuals for each m/z, and (d) distribution of scaled residuals for each inorganic species and size-binned particle number concentration.