SUPPLEMENTARY MATERIAL

Criteria for choosing factor number in NMR factor analysis

Several mathematical metrics could be used to aid determination of factor number, even if metadata analysis is also used to this aim (e.g., Lanz et al. 2008):

- *Q-value Analysis*

A first standardized criterion is the calculation of **Q-value**, the total sum of the squares of scaled residuals (Paatero et al., 2002). Q is expected to decrease with the number of factor, as each additional factor introduces more degrees of freedom with a general improvement of the fit. However, spurious solutions provide only minor decreases in Q, whereas genuine factors explain a significant fraction of the total variance and their inclusion is generally reflected by a marked decrease in Q. Therefore, the visual inspection of the curve Q versus number of factors often provides a straightforward manner to highlight to number of "genuine factors" (Paatero and Tapper, 1993). In this study, the Q/Qexp-values for the NMR factor analysis (averaged between all methods, Figure S1) suggest that a number of factors higher than three does not significantly improve the goodness of fit.

- *Principal Components Analysis (PCA)*

Preliminary PCA can be run to identify the best number of factors as a function of explained variance. For the Cabauw IOP NMR-dataset, the PCA model with three factors already explains 92% of the total variance. A fourth factor explains only a further 1% indicating a probable spurious solution.

- *Uniqueness of NMR spectral profiles and contribution*

As discussed in the text, the number of factors (p) was chosen to be 3 for the NMR dataset (Figure 4) but solutions resulting from p from 2 up to 8 were explored with all 5 factor analysis algorithms listed above. Comparisons between results from different algorithms were made, both to evaluate any differences between the models and to try to determine the best number of factors that can decompose the NMR-data, supposing their best agreement around the right number of factors.

In the solution with p = 4 (Figure S2), the spectra of two of the factors were very similar and difficult to distinguish from each other (F3 and F4 in Figure S2). Table S1 showed the correlation between profiles for the p=4 solution: very high values of correlation coefficient between F2 and F3, F2 and F4, F3 and F4 (respectively 0.81, 0.78 and 0.93) suggested that, probably, division into four factors was forced based on the idea that if two profiles have a high correlation are not well defined from the factorization process. Even for p=3 the correlation between F2 and F3 was high (0.82), but a more detailed analysis of the spectrum provided a different interpretation of these two profiles.

Moreover results from the different algorithms show a significant deterioration of their agreement from the p=4 solution with respect to that from p=3, confirming the chosen number of factor.

NMR-factors correlations with aerosol components (from filters and AMS measurements)

Figure S3 shows a selection of scatter plots reporting observed correlations between specific NMR factors and other aerosol chemical components. **NMR-F2 (HULIS)** is best correlated with TC (R=0.91) and sulfate (R=0.60) (panel a) and b) of Figure S3) and with the most oxidized AMS factor for OOA (Factor 4) (Figure S6). **NMR-F3 (LINEAR ALIPHATICS)** shows positive correlations with primary components (EC, HOA) and with low-molecular weight amines, especially with TMA (trimethyl-amine) (panel c), d), e), f) of Figure S3). EC and amines often exhibited diurnal maxima in the same manner of NMR-F1 (not found for the other NMR factors).

AMS-factors correlations with reference mass spectra

Figure S4 reports correlation plots between specific AMS spectral profiles in this study and reference mass spectra derived from previous investigations (Alfarra 2004; Lanz et al., 2007; Zhang et al., 2005).

SUPPLEMENTARY FIGURES

Criteria for choosing factor number in NMR factor analysis

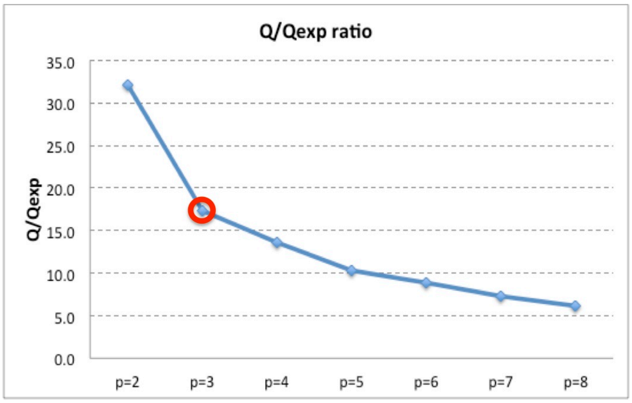

Figure S1: Q/Qexpected versus the number of factors p. Red circle denotes the chosen solution (p=3)**.**

Table S1: Pearson correlation coefficients (R) between spectral profiles of NMR-factors for p=3 and p=4 solutions.

p=3

R	F1	F2	F3
F1	1		
F2	0.64	1	
F3	0.62	0.82	1

p=4

R	F1	F2	F3	F4
F1	1			
F2	0.53	1		
F3	0.53	0.81	1	
F4	0.71	0.78	0.93	1

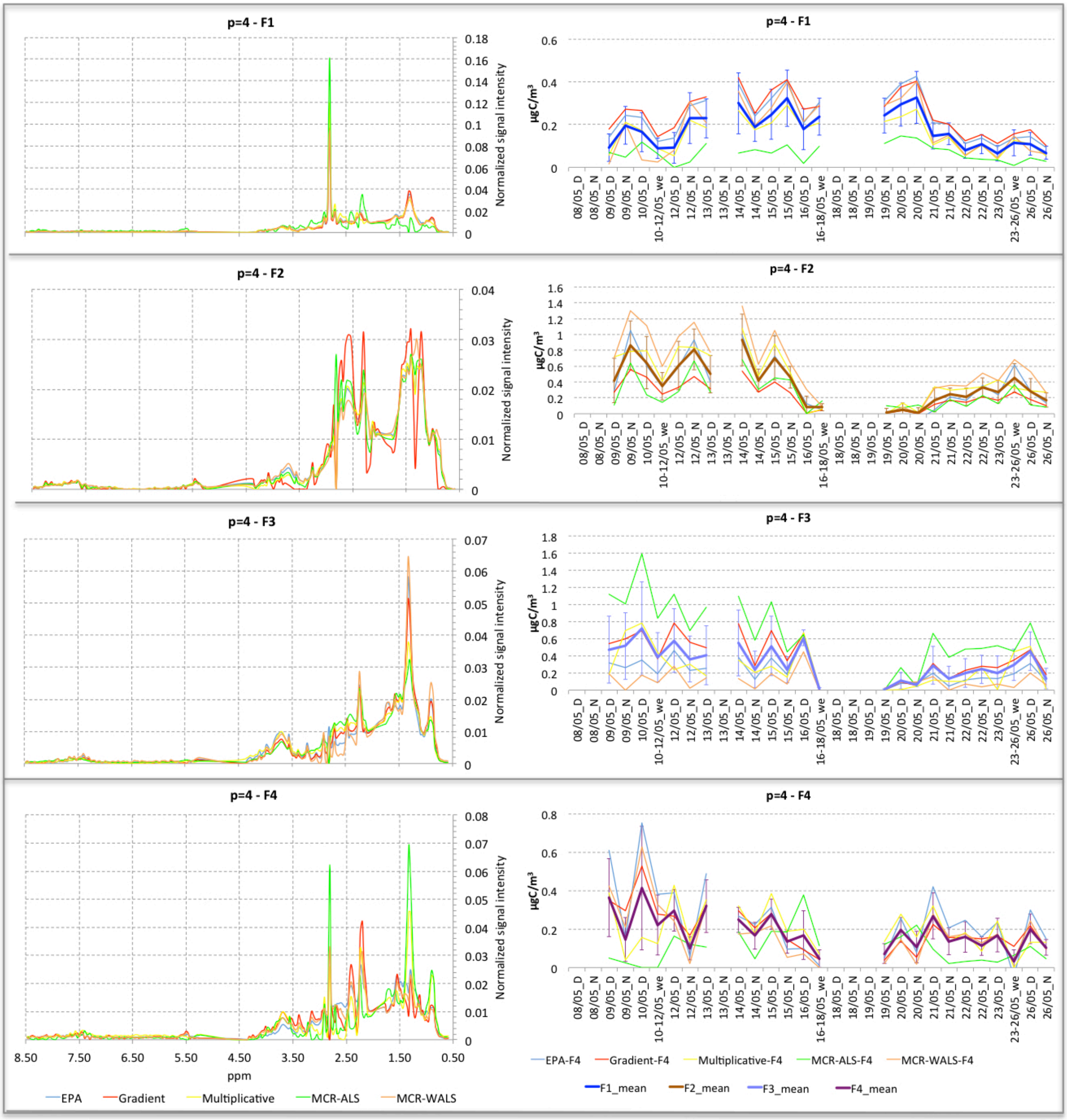

Figure S2: 4-factors solution for the NMR-dataset, spectral profiles (left side) and time series (right side). Results from all 5 different algorithms and the average between them were reported: PMF from EPA free-software (light blue line), Projected Gradient (red line), Multiplicative (yellow line), MCR-ALS (green line) and MCR-WALS (orange line) methods and average value (with standard deviation bars) for contribution (thick line in each graph).

NMR-factors correlations with aerosol components (from filters and AMS measurements)

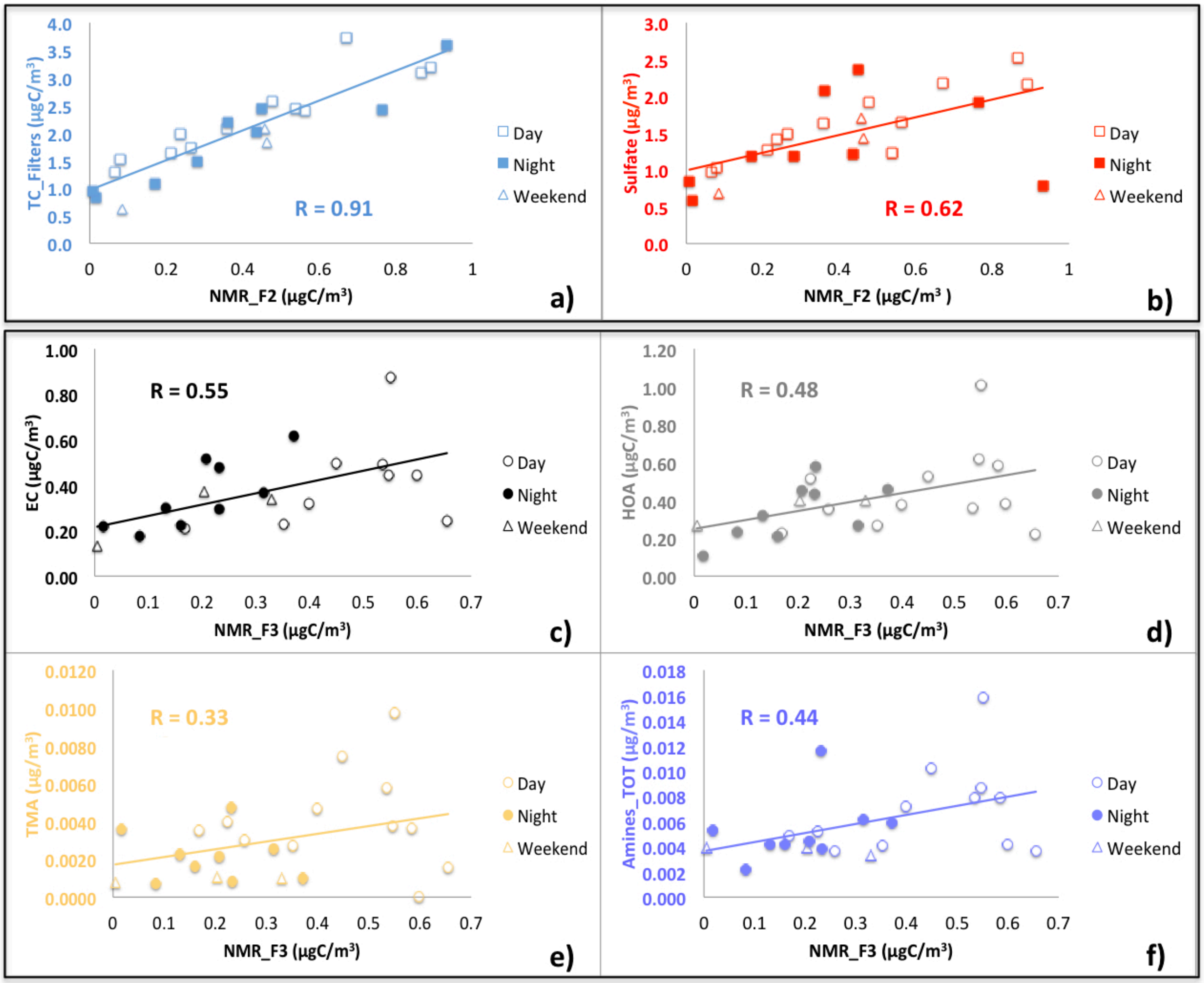

Figure S3: Main correlations between NMR-factors and other chemical data. **NMR-F2 (HULIS)** has correlation with TC and Sulphate (panel a) and b)). **NMR-F3 (LINEAR ALIPHATICS)** has positive correlations with components originated from primary combustion or forestry/agriculture sources (panel c), d), e), f)). Full circles and squares in panels represent data from day samples, empty ones for night samples and empty tringles for weekend samples.

AMS-factors correlations with reference mass spectra

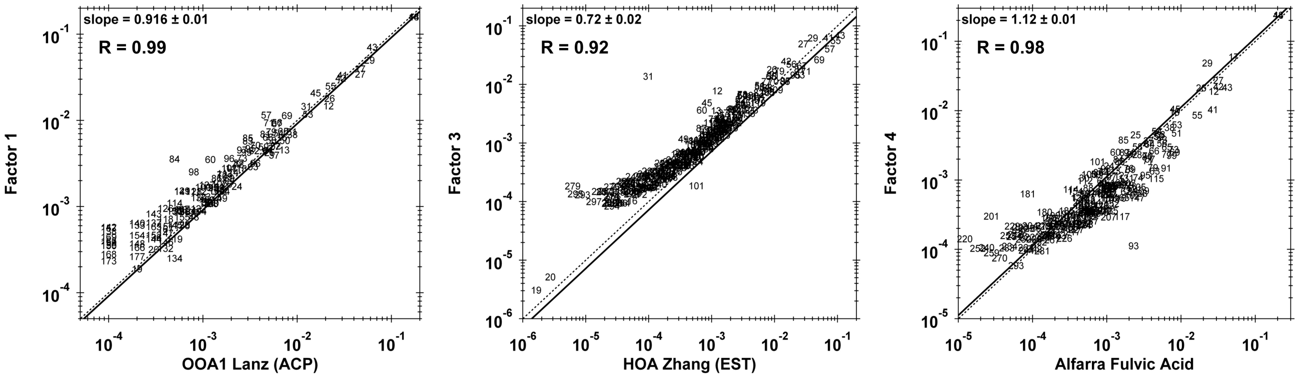

Figure S4: Correlation of PMF factor 1 mass spectrum (MS) to LV-OOA/OOA1 MS from Lanz et al., 2007 (left), factor 3 MS to HOA MS from Zhang et al., 2005 (middle), and factor 4 MS to fulvic acid MS from Alfarra 2004 (right).

Comparison PM1 filters/AMS

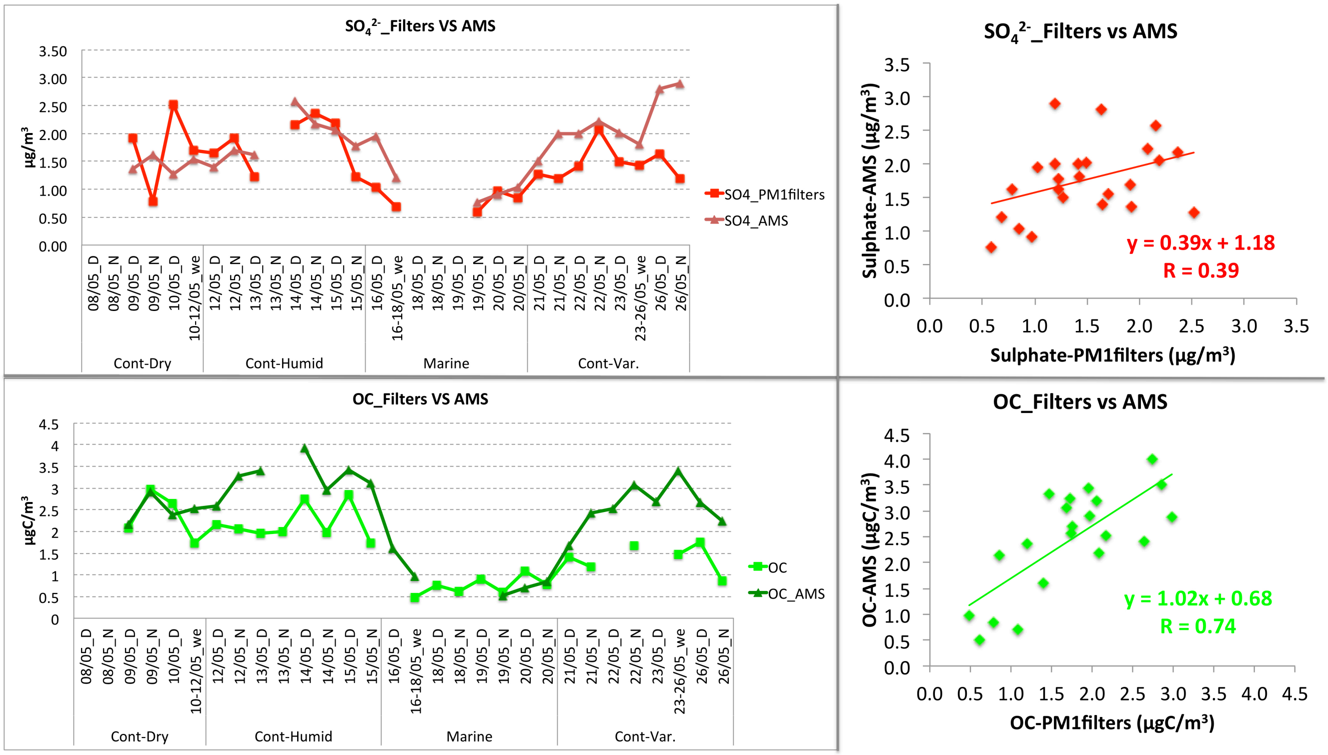

Figure S5: Comparison between PM1filters & AMS chemical data: on the left panels were reported sulfate and OC time series; on the right their correlations.

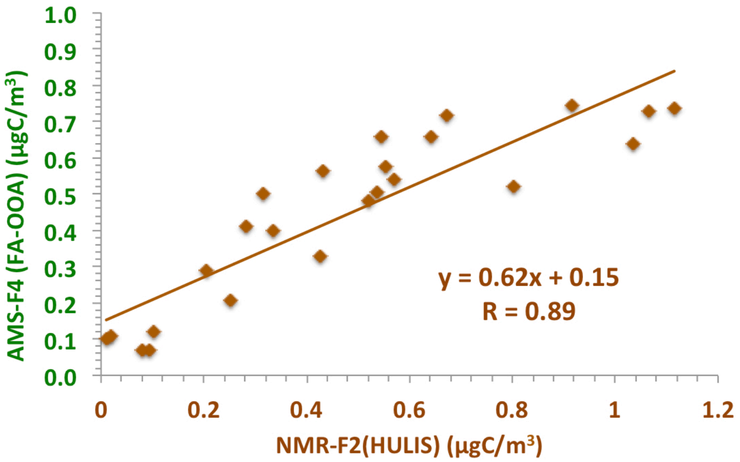

Figure S6: Comparison between AMS-F4 (FA-OOA) factor & NMR-F2 (HULIS) factor.