Text S1: The PMF model process

The contribution of various sources to rainwater ions was quantified by using a positive matrix factorization (PMF) model [86–88]. In a PMF model, the source contribution matrix (g_{ik}) and source profile matrix (f_{kj}) are decomposed by an ion concentration dataset matrix (x_{ij}):

$$x_{ij} = \sum_{k=1}^{p} g_{ik} f_{kj} + e_{ij} \quad (1)$$

where i, j, p, and e_{ij} represent the sample number, ion components, source number, and sample error, respectively.

An uncertainty (u_{ij}) value is needed in the PMF modeling procedure, which is dependent on the ion contents of a sample. If x_{ij} is less than the method detection limit (MDL; see Section 2.2 of main text), $u_{ij} = \frac{5}{6} \times$MDL, while $u_{ij} = [(\text{error fraction} \times x_{ij})^2 + \text{MDL}^2]^{1/2}$ if x_{ij} exceeds MDL. The error fraction is presented as measurement uncertainty (in %). The final goal of the PMF model is to obtain a minimum objective function (Q):

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{m} \left[\frac{x_{ij} - \sum_{k=1}^{p} g_{ij} f_{ij}}{u_{ij}} \right]^2 \quad (2)$$

where x_{ij}, u_{ij}, n, m, and p represent the ion concentration, uncertainty, sample number, ion species, and source number, respectively.

Briefly, the estimation of the concentration uncertainty and the determination of the method detection limit and the uncertainties of the variables were firstly performed, and dozens of model runs were subsequently conducted. After choosing a model with the lowest Q (robust) as the optimal solution, the number of factors ranging from 2 to 5 was examined by checking the Q value, the residual analysis, and the correlation between the observed and predicted values. In this case, a two-factor solution was invariably optimal for its most stable results and reasonable factors, because the Q true value was minimum.

In this study, according to the signal-to-noise value of the selected variables, TZ (total ion equivalent concentrations, TZ^++TZ^-) was set as total variable (weak), and other variables including Cl^-, F^-, NO_3^-, SO_4^{2-}, K^+, Na^+, Ca^{2+}, and Mg^{2+} were set as "strong." The F_{peak} values of 0.5, −0.5, 1.0, −1.0, and 1.5 were selected for rotation, and the F_{peak} of −0.5 was selected based on the minimum Q change. This study used two uncertainty-check tools including "bootstrapping" (BS) and "displacement method" (DISP) tests. We set the number of bootstraps to 100 and the minimum correlation R-value as 0.6. The BS result of the base run and F_{peak} indicated a very good reproducibility of the bootstrap factors. The DISP analysis of the base run indicated that the solution was stable and had very small amounts of rotational ambiguity; thus, the results can be considered valid.

References

86. Guo, W.; Zhang, Z.; Zheng, N.; Luo, L.; Xiao, H.; Xiao, H. Chemical characterization and source analysis of water-soluble inorganic ions in PM2.5 from a plateau city of Kunming at different seasons. Atmospheric Research 2020, 234, 104687, doi: 10.1016/j.atmosres.2019.104687.
87. Hien, P.D.; Bac, V.T.; Thinh, N.T.H. PMF receptor modelling of fine and coarse PM10 in air masses governing monsoon conditions in Hanoi, northern Vietnam. Atmospheric Environment 2004, 38, 189-201, doi: 10.1016/j.atmosenv.2003.09.064.
88. Lee, D.H.; Kim, J.H.; Mendoza, J.A.; Lee, C.H.; Kang, J.-H. Characterization and source identification of pollutants in runoff from a mixed land use watershed using ordination analyses. Environmental Science and Pollution Research 2016, 23, 9774-9790, doi:10.1007/s11356-016-6155-x.