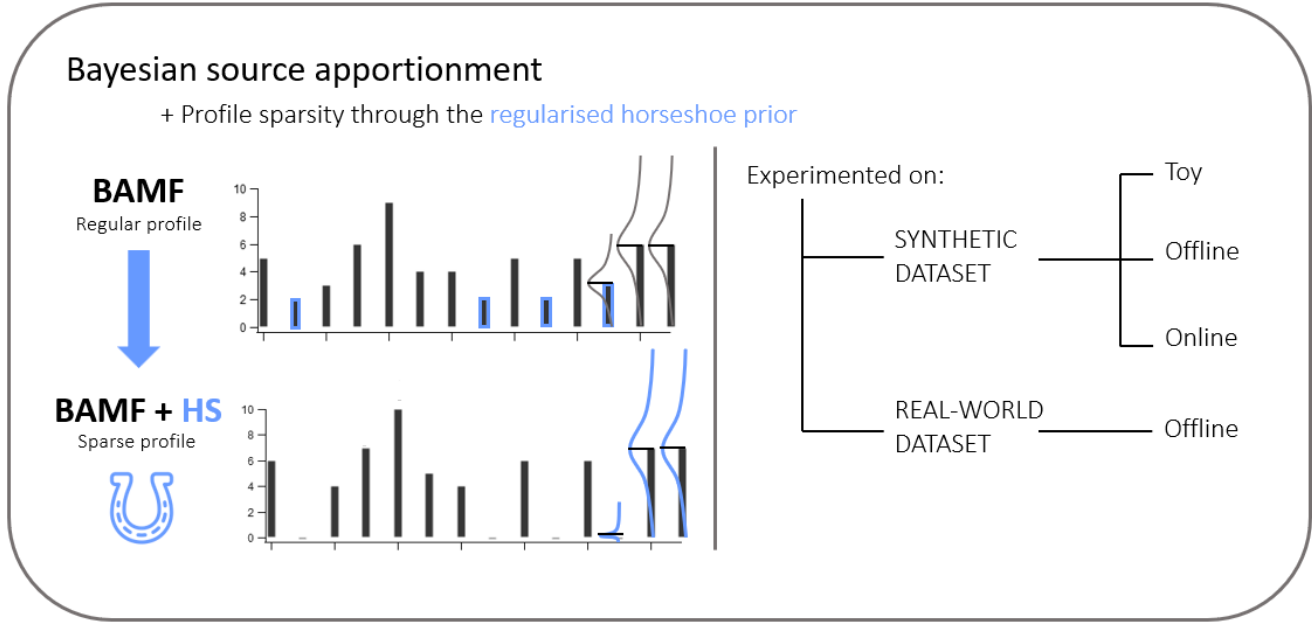
Chemical sparsity in Bayesian receptor models for aerosol source apportionment

Marta Via[1], Jure Demšar[2], Yufang Hao[3], Manousos Manousakas[3,4], Anton Rusanen[5], Jianhui Jiang[6], Stuart K. Grange[7], Jean-Luc Jaffrezo[8], Vy Ngoc Thuy Dinh[8], Gaëlle Uzu[8], Griša Močnik[1], and Kaspar R. Daellenbach[3]

[1]Center for Atmospheric Research, University of Nova Gorica, Ajdovščina 5270, Slovenia
[2]Faculty of Computer and Information Science, Tržaška Cesta 25, 1000 Ljubljana, Slovenia
[3]Laboratory of Atmospheric Chemistry, Paul Scherrer Institute, 5232 Villigen PSI, Switzerland
[4]Environmental Radioactivity Aerosol Tech. for Atmospheric Climate Impacts, INRaSTES, National Centre of Scientific Research "Demokritos", Ag. Paraskevi, 15310, Greece
[5]Atmospheric Composition Research, Finnish Meteorological Institute, 00101 Helsinki, Finland
[6] School of Ecological and Environmental Sciences, East China Normal University, 200241, Shanghai, China
[7] Climate and Environmental Physics, Physics Institute, University of Bern, Sidlerstrasse 5, 3012 Bern, Switzerland
[8] University of Grenoble Alpes, CNRS, INRAE, IRD, Grenoble INP, IGE, Grenoble 38000, France

Correspondence to: Marta Via (marta.viagonzalez@ung.si) and Kaspar R. Daellenbach (kaspar.daellenbach@psi.ch)

Abstract. Aerosol source apportionment is a key tool for understanding the origins of atmospheric particulate matter and for guiding effective air quality management strategies. However, source apportionment techniques still struggle to properly separate highly correlated sources without relying on restrictive *a priori* information, possibly skewing the solution and adding subjective operator input, with varying degrees of benefit. This study introduces sparsity into the Bayesian Autocorrelated Matrix Factorisation (BAMF) model with the aim of removing non-essential species contribution in the unconstrained profiles, which is expected to improve the separation of factors. The regularised horseshoe prior (HS) has been added to BAMF (BAMF+HS) to promote composition matrix **F** sparsity, shrinking low-signal contributions to the solutions. BAMF+HS was evaluated using three synthetic datasets designed to reflect increasing levels of data complexity (Toy, Offline, and Online), and a real-world multi-site filter dataset. The results demonstrate that BAMF+HS effectively enforces sparsity in offline datasets and that this improves accuracy in reconstructing source profiles and time series compared to BAMF and Positive Matrix Factorisation (PMF). However, its application to higher-complexity ACSM datasets revealed sensitivity to sampling instability hindering sparsification. With that, even though sparsity was not achieved, the quality of the BAMF+HS solution metrics were not deprecated compared to BAMF. Overall, this work underscores the value of incorporating profile sparsity as a solution property in Bayesian source apportionment, and positions BAMF+HS as a promising model for source apportionment.

33 **Graphical abstract**

35 ## 1. Introduction

36 Particulate Matter (PM) adversely affects human health through both short and sustained exposures (Pope and Dockerty, 1999,
37 Yang et al. 2019). The observed relationship between decreasing PM concentrations and increased life expectancy (Keuken et
38 al. 2011; Zheng et al. 2022) highlights the importance of developing mitigation plans grounded in detailed knowledge of PM
39 sources composition and concentrations. Moreover, because the toxicity of PM can be highly dependent on its sources
40 (Daellenbach et al. 2020), implementing source-specific mitigation measures contributes to more quantitative and efficient
41 abatement and a more effective protection of the population.

42

43 Source apportionment is the process of identifying PMF sources by using information about its chemical composition, and is
44 commonly conducted through receptor models (RMs) which differentiate PM sources according to the distinctness of their
45 chemical composition and time series characteristics. The most widely used RM is the Positive Matrix Factorisation model
46 (PMF, Paatero and Tapper, 1994), which decomposes the input chemical composition into the product of composition and
47 time series matrices ($\mathbf{F}$ and $\mathbf{G}$, respectively), and minimises the residuals of the fit through the weighted least squares loss. The
48 factorisation equation, hence, is written as

49 $$X = G \cdot F + E \,, \tag{1}$$

50 where $\mathbf{X}$ is the $\mathbb{R}^{n \cdot m}$ input matrix of n timepoints and m species, which is decomposed into $\mathbf{G}$ ($\in \mathbb{R}^{n \cdot p}$), and $\mathbf{F}$ ($\in \mathbb{R}^{p \cdot m}$), where p
51 is the number of factors, and $\mathbf{E}$ ($\in \mathbb{R}^{n \cdot m}$) is the residuals matrix.

52

53 Unconstrained PMF, even if it can lead to robust results, is usually insufficient when the sources are highly correlated or have

54 very similar source profiles. In such cases, guiding the model by introducing a priori knowledge has been proven beneficial

55 for the source disentanglement (Lingwall and Christensen et al. 2007, Belis et al. 2014, Dinh et al. 2025), however it can still

56 introduce substantial bias in the solution (Via et al. 2022). The extreme case of RM guidance is the Chemical Mass Balance

57 model (CMB), which factorises the initial matrices with a totally fixed $\mathbf{G}$ or $\mathbf{F}$. Globally, the RMs cover the whole range of

58 prior knowledge required (Viana et al. 2009a, Belis et al. 2013).

59

60 Bayesian models represent a probabilistic alternative to the PMF framework. The first application of Bayesian models in

61 atmospheric source apportionment was in Park et al. (2001, 2002) for Volatile Organic Compounds (VOC) source

62 apportionment. In this approach, the mass closure condition was taken to the Bayesian framework and an autocorrelation prior

63 AR(1) (the first order autoregression) was applied, improving the solution given assuming independent $\mathbf{G}$ components. The

64 autocorrelation prior importance was later reinforced in Rusanen et al. (2024) with a differently formulated autocorrelation

65 prior. The latter shows the added value of the Bayesian Autocorrelated Matrix Factorisation model (BAMF) in comparison to

66 PMF in different kinds of spectrometry-based PM synthetic datasets. The Bayesian Multivariate receptor modelling software

67 BNFA and bayesMRM (Park and Oh, 2021) were developed to provide user-friendly tools for Bayesian source apportionment.

68

69 However, studies using the Bayesian Matrix Factorisation (BMF) framework are still scarce. Some examples are Oh and Park

70 (2022), which employed a Bayesian RM to approach multi-site source apportionment, and Zhang et al. (2023), which

71 performed NH_4^+ source apportionment through the Bayesian SIMMR package (Govan et al. 2023). Bayesian models have also

72 been used as a complement to standard RMs, as in Balachandran et al. (2013) where a Bayesian model processing ensemble

73 solutions of a chemical transport model and solutions of three RMs are produced to then use it in CMB for production of final

74 results. The Bayesian model focused on attributing the proper weight to each of the ensemble components and improved the

75 correlation of sources with their markers compared to the traditional approach. Bayesian inference has also been used in Park

76 et al. (2002) and Dai et al. (2024) to generate spatially resolved source apportionment solutions adjusting the weights of each

77 location solution in a multi-site data scheme.

78

79 Thus, Bayesian Matrix Factorisation has become an effective and powerful tool for aerosol source apportionment. However,

80 to the authors knowledge, little attention has been given to improving the accuracy of chemical composition profiles, i.e. $\mathbf{F}$

81 components. This highlights the fundamental challenge in receptor modelling of obtaining chemically distinct and interpretable

82 source profiles from complex and overlapping emissions. Moreover, it has been shown in Rusanen et al. (2024) that in BAMF,

83 slight $\mathbf{F}$ differences can severely compromise the quality of $\mathbf{G}$ (Figure S2 in the mentioned article), hence, steps towards $\mathbf{F}$

84 refining should result in overall source apportionment method improvement. The enforcement of sparsity, defined as the

85 property of a dataset, model, or solution, where most elements are insignificant or inactive often represented as zeros or values

86 near zero, could be favourable for this problem. The accomplishment of sparse source fingerprints could represent "cleaner"
87 emission sources, less entangled among themselves. This work aims to implement sparsity on chemical fingerprints in BAMF
88 aiming for a more accurate source apportionment. We introduce sparsity with the regularised horseshoe prior (Piironen and
89 Vehtari, 2017), which unlike other sparsity priors, enables regularisation of the sparsity strength, and compare it with other
90 sparsity priors, such as Lasso (Tibshirani et al. 2015) and Spike-and-slab (Andersen et al. 2014). This model is then tested on
91 three synthetic datasets with different complexity degrees and one real-world dataset to depict the impact of sparsity and
92 potential benefits of its implementation.

93 **2. Methodology**

94 **2.1 Bayesian Matrix Factorisation**

95

96 Bayesian Matrix factorisation models, like other RMs, are based on the chemical mass balance equation (a). Bayesian modeling
97 approaches this problem probabilistically and bases the determination of the matrices, **F** and **G**, the main parameters to
98 determine, upon the assumptions imposed on the model, i.e. priors. Bayesian factorisation forces the decomposition through
99 modelling the **X** matrix components as a Gaussian with center on the "noise-free data matrix" **Z** ($\in \mathbb{R}^{n \cdot m}$) and a standard
100 deviation given by the positively-defined uncertainty matrix $\boldsymbol{\sigma}$ ($\in \mathbb{R}_{\geq 0}^{n \cdot m}$) (2). The matrix **Z** is, in turn, the product of the time
101 series and profiles submatrices, **G** and **F**, respectively, and (a) is rewritten as:

102 $X \sim \mathcal{N}(Z, \boldsymbol{\sigma}) = \mathcal{N}(G \cdot F, \boldsymbol{\sigma})$ (2)

103 Whilst **G** is not given any prior meaning and is sampled then by default from a uniform distribution, **F** is modelled as a Dirichlet
104 distribution to ensure positivity, with the sum of its components being equal to 1 (2):

105 $F_k \sim Dirichlet(1_m)$ (3)

106 With these **F** requirements, profiles represent the normalised contribution to the spectra of one source. Usual notation for
107 indices used hereinafter here are i, j, k for elements in the range $(1, \ldots, n)$, $(1, \ldots, m)$ and $(1, \ldots, p)$, for the timestamps, species,
108 and factors, respectively. The model configuration given by (2) and (3) will be referred to as Bayesian Factorisation model
109 (BMF) and represents the analog of PMF in the Bayesian framework. All models used in this manuscript are outlined in Table
110 1.

111 On top of this structure, Rusanen et al (2024) proposed an autocorrelation prior for **G** which should account for the inherent
112 autocorrelation of air pollutant sources in time. The formulation of the autocorrelation prior for **G** is given by (4) and includes
113 two more modelling parameters, $\boldsymbol{\alpha}$ ($\in \mathbb{R}_{\geq 0}^{p}$) and $\boldsymbol{\beta}$ ($\in \mathbb{R}_{\geq 0}^{p}$), which regulate the similarity of one **G** component with the
114 previous one as follows:

115 $G_{i+1,k} \sim C^+(G_{ik}, \alpha_k \cdot \Delta t_i + \beta_k)$ (4)

116 where $i \in (2, \cdots, n-1)$, C represents the Cauchy distribution and $^+$ represents $\mathbb{R}_{\geq 0}$. This prior centers the $i+1$'th

117 component distribution in the ith component with a distribution width that linearly depends on the temporal gap between these

118 two timestamps. Hence, the more temporally-separated two consecutive points are, the less correlated they are expected to be.

119 The Cauchy distribution was chosen due to its heavier tails which enable more probable jumps between consecutive i's than a

120 Gaussian distribution. This flexibility could be convenient for real-world datasets which are affected by measurement gaps.

121 The coefficients α and β are source dependent to allow for source-dependent correlation degrees. The model which introduces

122 this prior to BMF is called Bayesian Autocorrelation Matrix Factorisation model (BAMF, Rusanen et al. 2024).

123

2.1.1   The horseshoe prior

125 The introduction of sparsity in BAMF entails the addition of several hyperpriors in the **F** prior to implement the shrinkage

126 mechanism. In this study, we used the regularised horseshoe prior (Piironen and Vehtari, 2017), which is a global-local

127 complex of hyperpriors, i.e. the shrinkage power is both regulated globally source-wise and **F**-component-wise. The

128 regularised horseshoe (HS) prior implemented in **F** in the BAMF scheme as

129 $$F_{kj} = \mu_{kj} \cdot \tilde{\lambda}_{kj} \cdot \tau_j \,, \tag{5}$$

130 where μ ($\in \mathbb{R}^{n \cdot m}$) represents the **F** matrix without the horseshoe prior. μ, in turn is defined as a standard Cauchy distribution

131 prior as

132 $$\mu_{kj} \sim C^+(0,1), \tag{6}$$

133 where τ ($\in \mathbb{R}_{\geq 0}^m$) represents the global shrinkage parameters and

134 $$\tau \sim C^+(0, \tau_0 \cdot \sigma_{HS}), \tag{7}$$

135 where the parameter τ_0 can be regulated by the user to regulate the overall shrinkage power and σ_{HS} is sampled from an uniform

136 distribution. The hyperparameter $\tilde{\lambda}_{kj}$ applies the local shrinkage to λ ($\in \mathbb{R}^{n \cdot m}$) **as**

137 $$\tilde{\lambda}_{kj} = \sqrt{\frac{c^2 \cdot \lambda^2}{c^2 + \lambda^2 \cdot \tau^2}} \,, \tag{8}$$

138 where

139 $$c^2 \sim \Gamma^{-1}(0.5 \cdot slab\_df, 0.5 \cdot slab\_df) \tag{9}$$

140 $$\lambda \sim slab_{scale} \cdot C^+(0,1) \tag{10}$$

141 both combined providing the characteristic shrinking horseshoe shape. Further description of the horseshoe implementation

142 on BAMF can be found in Section S2.1, and the prior derivation and details, in Piironen and Vehtari (2018). The distribution

143 parameters τ_0, σ_{HS}, and *slab_df*, *slab_scale* were tested and results did not show significant sensitivity to their variations, so

144 we keep the defaults as provided in Piironen and Vehtari (2018) as can be found in the available shared codes. The models

145 with the horseshoe (HS) priors are hereinafter marked with "+HS". Figure S1 shows a schematic diagram of the matrix

146 decomposition through BAMF.

147

148 In order to assess the amount of sparsity of a dataset or a solution, we used the Gini coefficient (Gini et al., 1936), which

149 assesses the inequality over a distribution as follows:

150 $$Gini = \frac{\sum_{i=1}^{n}(2 \cdot i - n - 1) \cdot x_i}{2 \cdot n \cdot \sum_{i=1}^{n} x_i} \tag{11}$$

151 where x values are sorted in ascending order and n is the number of elements in x. Since it quantifies the unevenness, it can be

152 a proxy for sparsity; if some values are high and the others are zero, Gini $\approx$ 1, if all values are equal, Gini = 0. Also, the

153 solution-to-truth Gini values ratio will be discussed throughout the analysis, referred to as "Gini ratio". To evaluate if the

154 sparsity is enforced precisely where it should, an additional metric has been applied called "zero truth sum". This metric sums

155 up the modelled contributions of the null species in the truth profiles.

2.1.2   Alternative factorisation methodologies

156

157 **BAMF-AR1.** There is an alternative formulation for the autocorrelation prior as introduced in Bayesian models by Park et al.

158 (2001). The AR(1) autocorrelation prior is the first degree polynomial expansion of the autoregressive models and it proposes

159 a linear progression of $G_{i+1,k}$ from $G_{i,k}$. We introduce AR(1) in the Bayesian framework as

160 $$G_{i+1,k} \sim \mathcal{N}(\alpha_k \cdot G_{ik} + \beta_k, \gamma_k) \tag{12}$$

161 In this formulation, the $i+1$-th point stems from a Gaussian distribution centered linear combination on the i-th point with

162 source-dependent slopes (α) and intercept (β), and width γ. Although, unlike (4), it disregards the decrease of correlation

163 between gapped consecutive points, this prior allows for source-specific time series trends, which would be beneficial for

164 certain source description. The model which introduces this prior to BMF will be called BAMF-AR1.

165

166 **BAMF-GS.** Another formulation is introduced, switching the Dirichlet distribution to the matrix **G** instead of **F** (6). This

167 swap should allow **F** to retain the **X** matrix mass and could potentially help disentangling profiles due to the upweighting of

168 the chemical profiles.

169 $$G_k \sim Dirichlet(1_n) \tag{13}$$

170 Thus, the G presents two priors, the dirichlet distribution and the autocorrelation prior, whilst F is sampled from the default

171 uniform distribution. This model will be called hereinafter BAMF-GS as short from BAMF-G simplex, since a simplex is the

172 set of positive vectors that sum to one hence it is the natural geometric structure for the Dirichlet distribution to sit on. This

173 model structure, nevertheless, does not allow for a horseshoe prior application, since due to the factors mass now incorporated

174 in **F**, the coefficients will be very distinct from zero and the horseshoe prior will not perceive them as potential signals to

175 sparsify.

176

177 **CMB.** Lastly, a Bayesian formulation of the CMB model was employed in order to test the horseshoe prior capacities with the

178 most proper factorisation possible. This model is mainly analogous to CMB in the Bayesian framework, but the **G** matrix was

179 fixed with the truth time series. Hence the model only had to determine the **F** components distributions to match the
180 factorisation condition (2) given the truth **G**.

2.1.3 Solver and Hamiltonian-Monte Carlo Markov Chain

182 All Bayesian models were compiled and run in STAN (Carpenter et al. 2017), a probabilistic programming language developed
183 for Bayesian modelling. STAN solves Bayesian inference through the Hamiltonian Monte Carlo (HMC) algorithm based on
184 Markov Chain Monte Carlo methods (MCMC). HMC uses an approximate Hamiltonian dynamics simulation with the
185 Metropolis acceptance/rejection criterion and a no-U-turn sampler (NUTS, Hoffman and Gelman, 2014). For the sake of
186 brevity, we present only the essential concepts here, directing readers to Carpenter et al. (2017), Gelman et al. (2014), STAN
187 Manual (2025) and references therein for comprehensive information.

188 The parameters of the model, primarily **F** and **G** but also all the other defined hyperparameters, are sampled from their posterior
189 distributions, constructed from the priors and the data introduced. In the Hamiltonian analogy, the evolution of these parameters
190 across samples is computed as the trajectory of a fictitious particle. This particle moves through the parameter space driven by
191 random momentum in all directions. The trajectory is hence simulated using a discretized approximation, and candidate
192 positions are accepted or rejected according to the Metropolis criterion (Metropolis et al. 1953). This process provides a
193 distribution over samples of possible solutions from which confidence intervals of each of the model (hyper)parameters can
194 be extracted. A set of samples is called a chain, each of them initialised with a different seed in order to explore the solution
195 space more broadly. In order to initialise the model parameters more effectively, the maximum a posteriori (MAP) point
196 parameters solution estimated by STAN is used through the LBFGS algorithm (Liu and Nocedal, 1989). Even if this approach
197 makes the parameter sampling process much more efficient, solutions might have multiple local maxima and MAP will
198 initialise the models based only on one of those. This highlights the importance of using different seeds to explore the solution
199 space more widely. Since the early iterations of each Markov chain are typically influenced by the starting values and may not
200 represent samples from the true posterior distribution, we discarded the first half of the samples from each chain. Different
201 settings were used according to the type of experiment, shown in Table S1.

202 In order to evaluate the convergence of a solution to the target posterior distribution, the potential scale reduction factor ($\hat{R}$,
203 Gelman and Rubin, 1992) is used. This coefficient compares the variance within chains and between chains of the **Z** matrix,
204 hence if chains converge, $\hat{R} \approx 1$, values of $\hat{R} \gg 1$ imply chain divergence and values of $\hat{R} \ll 1$ imply sampling divergence in chains.

2.2 PMF

206 In PMF (Paatero and Tapper, 1994), equation (1) is solved through the ME-2 solver (ME2, Paatero, 1999) based on the
207 weighted means squares minimization of the quantity:

208 $$Q = \sum_{j=1}^{m} \sum_{i=1}^{n} \frac{e_{ij}^2}{\sigma_{ij}^2} \qquad\qquad (12)$$

209 PMF was implemented on all datasets unconstrainedly through the Source Finder software (SoFi version 9.5, Canonaco et al.

210 2013) with 100 sorted runs which are posteriorly sorted as in BAMF. The number of runs may seem compromising the PMF

211 quality in comparison to the 6000 samples per chain used in Bayesian models. However, this comparison is misleading, since

212 the factorisation space is indeed better explored by PMF, with 100 different sampling seeds, while only 4 seeds (chains) were

213 used in BAMF-like models as usual procedure in Bayesian modelling for the sake of computational resources.

214

215 **2.3 Pre- and post- processing for all models**

216 Previously to model running, **X** and **σ** are normalised to use consistent scales of all priors and posteriors for Bayesian models.

217 The normalisation is based on ensuring a mean of X = 1.

218 $$X^* = X / f_{norm} \ , \ \ \sigma^* = \sigma / f_{norm} \qquad\qquad \text{where} \ \ \ f_{norm} = \sum_{i,j} X_{ij} / (n \cdot m) \qquad\qquad (13)$$

219 After the factorisation this normalisation is reverted converting the normalised matrices, hereinafter referred as $\overline{G}$, $\overline{F}$, to the

220 properly-scaled **G, F** matrices.

221

222 The factor ordering in the matrices is random in the model results, hence the sorting process is, hence, the solutions factors

223 must be sorted. Here, as in Rusanen et al. (2024), we used the Hungarian algorithm (Kuhn, 1955) to sort the $\underline{Z_k}$ components

224 $(\overline{Z_k} = \overline{G_{ik}} \cdot \overline{F_{kj}}$, i.e. each factor's normalised **Z** submatrix). The metric to sort the components is the Manhattan distance (i.e.

225 the sum of the absolute differences of two Cartesian coordinates). All factors in each chain of samples are then reordered upon

226 the factor order of a small group of samples of that chain (the last 5, arbitrarily chosen) and, subsequently, one all-samples-

227 averaged F_k and G_k are retrieved for each of the chains. Then, the order of factors of each of the chains is sorted again in the

228 same way in relation to the truth F, to have all sources equally sorted in all chains. Median and quantiles are computed over

229 samples and chains to produce the final solutions and uncertainties. This sorting process is also used for the PMF solution

230 despite not being its usual sorting approach for the sake of homogeneity in comparison to the Bayesian models.

231

232 The last step of the experimental process is to assess the model performance on the given dataset. The evaluation of the

233 performance should be based on: i. reconstruction performance, or the difference between **X** and **Z**; ii. similarity to truth, or

234 environmental sensibility based on the apportionment of source tracers in case the truth is not available; iii. computational

235 performance. The reconstruction performance will be assessed by checking the cell-wise correlation between **X** and **Z** and

236 checking the median and maximum of the absolute value of relative deviations of Z and X with respect to the measurement

237 uncertainty matrix **σ** (|**X-Z**|/**σ**). The similarity to truth, when available, is tackled by comparing the median ratio between

238 modelled **G** and truth, the Pearson correlation for the **G** matrix and the Spearman correlation for **F** amongst models. The

239 Spearman correlation coefficient for the factor profiles was chosen due to the expected non-linearity of the comparison and

240 likely presence of outliers. These comparisons, and especially when the ground truth is not available, need to be accompanied
241 by visual inspection of the solution quality, looking for resemblance with known environmental sources. The models
242 accounting for sparsity will be also compared upon the aforementioned Gini metric and, when truth is available, the Gini ratio
243 with truth and the "zero-truth sum".

244

2.4 Datasets

246 The datasets created for model experimentation can be divided into synthetic and real-world datasets. Synthetic datasets are
247 artificially created with the purpose of knowing the **F**, **G**, to test model accuracy retrieving these matrices with respect to the
248 *truth* and these have been widely used for source apportionment validation in the last decades (Park et al. 2002, Brinkman et
249 al. 2006; Belis et al. 2015; Via et al. 2022; Rusanen et al. 2024). In order to challenge the models gradually, we created three
250 synthetic datasets with increasing degrees of complexities (toy, offline, online ACSM synthetic datasets). Additionally, a real-
251 world chemically sparse dataset was also used to test the results. Despite the truth's factorisation is unknown and results cannot
252 be directly verified, the model's factorisation can be assessed environmentally or based on indicators on the goodness of the
253 fit. The different datasets have different levels of sparsity, as can be seen in Table 2, that the models with the horseshoe prior
254 should aim to replicate.

2.3.1 Chemically-sparse toy dataset

256 A simplistic synthetic toy dataset was designed to perform basic control and performance tests. It was devised by creating
257 three very simple and sparse profiles and using three time series (HOA, SOAbio, BBOA) from modelled source time series of
258 the city of Zurich (Rusanen et al. 2024) in order to test how sparsity priors act on very uneven species contribution. Then, **F**
259 and **G** were multiplied to generate **Z**, and some gaussian error with standard deviation **σ** was added to each component to
260 generate a realistic **X** matrix. The uncertainties matrix **σ** was designed as a sixth of the X values plus Gaussian noise. With this
261 arrangement, the models can be applied conventionally to the **X**, **σ** matrices and the modelled **F** and **G**, can be compared to
262 the original truth **F**, **G**, which will be referred hereinafter as $\mathbf{F_0}$ and $\mathbf{G_0}$, displayed in Figure S2.

2.3.2 Chemically-sparse synthetic offline dataset

264 We created a synthetic offline filter dataset, mimicking the filter-based measurements input matrices, in order to test the models
265 accuracy in these kinds of datasets. This dataset mimics the concentrations on the coarse fraction (PM_{10} - $PM_{2.5}$) as collected
266 by a high-volume sampler on the Zurich-Kaserne site (Grange et al. 2021) including the following chemical species: OC, Al,
267 Na, Mg, Cl, K, Ca, S, Fe, Cu, Zn, Mn, Sb, Ba, mannitol, arabitol. In the original real-world dataset, data obtained with two
268 series of samples (PM_{10} and $PM_{2.5}$) were subtracted in order to focus on the coarse source apportionment, since the main
269 emission sources of these elements and organic species stem from mechanical processes leading to major coarse models. It
270 was created by crafting first the **F** and **G,** then multiply them and creating **X** and **σ**. The **F** matrix was slightly modified from

271 that proposed in Manousakas et al. (2025), making the chemical profiles slightly sparse by zeroing the non-relevant species in

272 each of the factors (dust, traffic, salt, coarse biological). The **G** matrix was composed of the time series of:

- Dust: modelled PM_{10} dust (Vasilakos et al. in prep.) converted to coarse with the Al_{PM10} vs. PM_{10} ratio from Grange et al. (2021).

- Traffic: modelled PM10 copper (Upadhyay et al. 2025) converted to coarse with the Cu_{PM10} vs. PM_{10} ratio from Grange et al. (2021).

- Salt: coarse Na+Cl (Grange et al. 2021) converted to PM concentrations and multiplied by an arbitrary number (3 in this case match the concentrations of the sea salt factor in the original dataset).

- Coarse biological: coarse Arabitol+Mannitol (Grange et al. 2021) converted to PM concentrations and multiplied by 3, similarly as for the salt factor.

281 This dataset will be called "offline synthetic dataset". Another more simplistic dataset was prepared similarly but using Al and

282 Cu for dust and traffic factors, respectively, in the same way as in the salt or coarse biological factors, i.e. omitting the use of

283 modelled data. This dataset will be hereinafter named "Purely-measurement-based offline synthetic dataset" and its modelling

284 results will be described in section 3.2.

285 Once the **F** and **G** matrices were created, **X** was calculated by their multiplication and the addition of Gaussian noise with

286 amplitude **σ**. The uncertainties matrix **σ** was generated as in Grange et al. (2021) multiplied by 2 to balance the signal-to-noise

287 ratio to the datasets in Manousakas et al. (2025). The matrices **F**, **G** of this dataset are displayed in Figure S3.

288

2.3.3   Chemically sparse real-world offline dataset

290 A real-world dataset was employed to test the current models applicability in campaign measurements. This dataset was

291 originally used for source apportionment in Manousakas et al. (2025) and Grange et al. (2021) and consists of $PM_{10-2.5}$ samples

292 at five Swiss National Air Pollution Monitoring Network (NABEL): Basel, Bern, Magadino, Payerne, and Zurich. The

293 measurements were taken in the June 2018 - July 2019 period every fourth day and using Digitel high-volume samplers. During

294 the sampling campaign PM_{10} and $PM_{2.5}$ were collected and the respective concentrations were subtracted to generate the coarse

295 ($PM_{10-2.5}$) concentrations. These samples include: i. OC concentrations, measured through the thermal optical transmission

296 (TOT) EN16909 method with the EUSAAR2 temperature protocol; ii. elemental concentrations (Al, Fe, Cu, Zn, Mn, Sb, Ba,

297 Sr, Bi, Pb) measured by inductively coupled plasma atomic emission spectrometry (ICP-AES) and inductively coupled plasma

298 mass spectroscopy (ICP-MS); iii. water soluble inorganic ion concentrations (Ca^+, Cl^+, Mg^+, K^+, Na^+), determined by ion

299 chromatography (IC); iv. Organic species (mannitol, arabitol) determined by a high-performance liquid chromatographic

300 method followed by pulsed amperometric detection (HPLC-PAD). The uncertainties of these species were calculated as in

301 Grange et al. (2021).

302   **2.3.4   Chemically less sparse synthetic online ACSM datasets**

303   With the aim of recreating more complex real-world datasets to test the models we generated 6 datasets for four European

304   cities: Krakow, Milan, Paris, and Zurich. The objective is to recreate OA matrices as given by a mass spectrometer instrument

305   like Q-ACSM, for which there are plenty of real-world source apportionment studies in the literature. The **G** matrix was created

306   from OA sources time series generated through the regional air quality model CAMx (Comprehensive Air Quality Model with

307   Extensions) as previously published by Jiang et al. (2019). The five sources of these datasets were hydrocarbon-like OA

308   (HOA), related to traffic emissions, biomass burning OA (BBOA), biogenic SOA (SOA$_{bio}$), biomass burning SOA (SOA$_{bb}$),

309   and traffic SOA (SOA$_{tr}$). To ensure seasonal representativity while keeping computational costs low, datasets included the

310   first two weeks of every second month of 2011 (January, March, …). The relative concentrations of these datasets are shown

311   in Figure S5. This figure shows the highest seasonal OA variation for the city of Milan and the lowest for Zurich. In terms of

312   sources, the most seasonally stable sources, overall, are HOA and SOA$_{tr}$ in contrast to the remarkable variability of BBOA

313   and SOA$_{bio}$. The profiles used to create the species matrix **F** were those in Table S2 for primary sources (HOA, BBOA). For

314   secondary sources, the profiles from the European megacity dataset presented in Rusanen et al. (2024) were used for the Zurich

315   city, which were slightly perturbed for the other cities due to the limited availability of these sources' profiles in the literature.

316

317   The **X** matrix was obtained by multiplying the **F** and **G** submatrices and adding Gaussian noise. The procedure to calculate the

318   error matrix for such datasets is described in Via et al. (2022) and the dataset used to calculate the error matrix is that from the

319   Zurich site, which ranges from February 2011 until December 2011.

320

321   Lastly, a sensitivity analysis was carried out by modifying the original **F**, **G** matrices slightly upon which the **X**, **σ** matrices

322   were subsequently created. The first Zurich dataset (period 01/09/2011 - 14/09/2011) was used for this purpose and we chose

323   to perturbate one factor only (HOA). The **F**, **G** submatrices were perturbed independently upon the expression:

324   $$G_{HOA}' = G_{HOA} \cdot \mathcal{N}(1, \sigma') \qquad\qquad F_{HOA}' = F_{HOA} \cdot \mathcal{N}(1, \sigma') \qquad\qquad (14)$$

325   where we used σ' = [0, 0.1, 0.2, 0.3, 0.4, 0.5] to create different degrees of perturbation. The profiles in **F** were normalised

326   after that process. It must be noted that the perturbation is more relevant on **F** than in **G** since a given σ' in the aforementioned

327   range is more comparable and impactful on the profile contributions, bounded to 1, than on the unbounded time series

328   timepoints. Consequently, with this framework, we obtained 6 **G**-perturbed and 6 **F**-perturbed input matrices. Both BAMF

329   and BAMF+HS models were run with all these input matrices and their subsequent HOA results were compared to the original

330   truth in order to grasp the sensitivity of the models upon time series and profile perturbations.

3.   Results

3.1 Chemically sparse toy dataset

Here, we introduce the evaluated models relying on unrealistically simplified toy data with the purpose of showcasing the performance of the horseshoe prior introduction to BAMF (Figure S2) and the alternative factorisation methodologies, which are discussed in SI Section C.1.

In the first evaluation step, we assess the performance of the horseshoe prior under the assumption that the source matrix $\mathbf{G}$ is known, in order to isolate its effect on the estimation of $\mathbf{F}$. Figure 1 shows the distribution of each $\mathbf{F}$ component for CMB with and without the horseshoe prior (CMB, CMB+HS, respectively, Table 1). Figure 1 showcases the horseshoe prior power to generate sparsity in $\mathbf{F}$ components, shrinking the lowest signals to zero and, as a consequence, enlarging the most prominent signals. Table 3 shows how the Gini metric is consistently higher for CMB+HS with respect to CMB, supported by a higher Gini ratio and lower zero truth metric reflecting the sparsification of profiles and higher similarity to truth. The RMSE compared to the truth for the profiles improved with the horseshoe prior applied for all three factors (for CMB and CMB+HS, respectively: 1.2e-04, 3.8e-05 for F1; 1.86e-04, 5e-05 for F2; 3.3e-05, 1e-05 for F3). Hence, the sparsity introduced in $\mathbf{F}$ through the regularised horseshoe prior successfully improved the profile description of the solution.

In the next evaluation step, we test the various models assuming no prior knowledge. Figure 2 shows the results of PMF, BAMF, and BAMF+HS models on the toy dataset and Table 3 shows their factorisation and comparison to truth performances. In terms of factorisation, median relative errors are better for BAMF+HS and BAMF than for PMF, but their maximum errors are higher and the Pearson coefficients slightly lower, all this entailing comparable factorisation performances. All models generally adapt well to the truth features, but they present non-negligible differences. PMF results better resemble the truth in terms of $\mathbf{G}$ factors, both in concentrations and feature adaptation. However, in terms of profiles, the BAMF+HS model is the closest to the truth, especially for the second and third factors for which the sparsity introduction results are advantageous with respect to BAMF results. This is confirmed by the Gini ratio with truth which is the largest for BAMF+HS and with the lower truth zero metric, indicating a higher F similarity to truth. The sparsity effects can also be seen in Figure S5, in which the horseshoe shrinkage is evident for the low m/zs allowing in turn the larger m/zs to retain more mass, hence resembling better the truth profiles. However, the BAMF+HS could not shrink down the lowest signals in Factor 1, likely because their contribution estimated by the mass balance and the autocorrelation restrictions of this model made it unclear for the horseshoe to shrink them down completely. With this result, this toy dataset depicts the capacities and limitations of the horseshoe implementation on BAMF: it is capable to sparsify effectively only the signals which are close enough to zero as given by the restrictions of the BAMF model.

363   While other sparsity priors exist (e.g. Lasso and Spike-and-slab priors (Figure S6, Table 3, Table S3)), our tests show that the
364   BAMF+HS model is most effective in shrinking unnecessary contributors to F. Hence this prior will be used onwards. This is
365   evidently portrayed by the Gini ratio, for which neither Lasso nor Spike-Slab achieve the signal shrinkage that the BAMF+HS
366   does. Also, neither BAMF+Lasso nor BAMF+Spike-and-slab managed to sparsify the first factor. Additionally, different
367   autocorrelation formulations were tried with and without the horseshoe prior, showing worse performance than BAMF or
368   BAMF+HS, respectively, as discussed in section SI C.1. This supports using the BAMF autocorrelation prior instead of the
369   alternative AR(1) prior, G simplex formulation or lack of autocorrelation prior models, although these models are also tried
370   on the other datasets to prove this further.

371   **3.2 Chemically sparse synthetic offline dataset**

372   This synthetic offline dataset is intended to assess the performance of different models on a proxy representation of atmospheric
373   aerosol data, while maintaining the verifiability property inherent to synthetic datasets as described in Section 2.4.2. We
374   performed source apportionment of the $\mathbf{X}$ matrix through the aforementioned Bayesian models and PMF, obtaining 4 factors
375   fingerprints and time series. The dataset used in this source apportionment is expected to be much more sparse than ACSM-
376   like datasets, hence it could better expose the capabilities and added value of the sparsity prior.

377

378   BAMF was run by initialising $\mathbf{F}$ as a normal distribution to make the sampling more sturdy to avoid the initialization failure
379   otherwise. A t-test was run comparing the $\mathbf{F, G}$ factors from this slightly modified model and BAMF to ensure their similarity.
380   Its results passed the t-test for all factors except for one factor, although it presented a $R^2=0.9990$ correlation and only a 20%
381   of quantitative difference with that BAMF factor. Hence, one can assume that the model provides an acceptable level of
382   agreement with BAMF, capturing the essential structure of the factors with only very minor deviations.

383

384   Figure 3 presents the (a) time series (b) auto-correlation (c) profiles of the source apportionment solution for PMF, BAMF,
385   BAMF+HS, (d) additional comparison to truth metrics, and Figure 4 shows the histograms of the models $\mathbf{F}$ components
386   estimation. The time series and autocorrelation show only slight differences between the models, the PMF being the most
387   different to the truth in all factors except the salt one, as supported in Figure 3 (d). Amongst factors, the coarse biological
388   source is the most poorly reconstructed. If accounting for the sum of all factors $\mathbf{G}$ R^2s and $\mathbf{G/G_0}$, in the last row of Figure 3(d),
389   the most accurate model is the BAMF+HS, followed by PMF and then BAMF. In terms of profiles, the best overall model
390   performance depends on the metric, $\mathbf{F}$ Spearman correlation coefficient being highest for BAMF and R^2 and cosine similarity
391   correlation coefficients for BAMF+HS. This fact, accompanied by Gini being the highest for BAMF+HS and the closest to 1
392   Gini ratio, indicates that the extreme values of the profile (i.e. maximum and zeros species contributions) are closer to truth
393   for BAMF+HS, whose extreme contributions would be less relevant in the Spearman correlation coefficient. Considering the
394   Truth $\mathbf{F}$ zeros sum metric, the horseshoe shrinkage is visibly sparsifying most of the low signals whilst BAMF and PMF

395 present non-zero contributions for species whose contribution in this factor is null. Hence, the BAMF+HS model would
396 effectively promote the profiles sparsity which it was intended for.

397

398 However, the favourable results of BAMF+HS in comparison to the other models could be a dataset-dependent finding, related
399 to the properties of the created synthetic dataset. The purely-measurement-based offline synthetic dataset, whose performance
400 statistics are shown in Table S4, shows that PMF overperforms BAMF+HS, presenting slightly higher $\mathbf{F}$ and $\mathbf{G}$ R^2 and better
401 G/G_0. This could indicate that the optimal model selection might be dataset dependent. However, the source time series of this
402 very simplistic dataset are fully correlated with some species time series, since they are used to generate factor time series,
403 which makes it a very redundant dataset. In this scenario, the source apportionment comparison might still be valid, but it is
404 not the perfect showcase for RMs testing due to the excessive source correlation with species. However, we found it valuable
405 to present different model performances on different datasets, which in atmospheric measurements, can suffer from artefacts
406 complicating the behaviour of some models.

407

408 In the same way, the alternative autocorrelation priors models were also tried and will be thoroughly discussed in Section SI
409 C.II. However, overall, the BAMF+HS model is the one providing the best source apportionment results for this offline dataset,
410 taking advantage of the sparsity to upgrade both profiles and time series accuracy.

411

3.3 Real-world offline dataset

413 To test the models on real-world data and identify their limitations for more complex datasets, we tested the models in the real-
414 world offline PM_{10}-$PM_{2.5}$ offline dataset described in Section 2.4.3. Since the truth is not accessible, the model performance
415 can only be assessed upon environmental, factorisation-related, and computational criteria. For this dataset, BAMF and
416 BAMF-AR1 models presented initialisation issues incapacitating them to properly launch the models. To avoid this issue, we
417 made them more robust by initialising $\mathbf{F}$ components as Gaussian distributions centered at zero and with a standard deviation
418 of 1 so that we restrict values to be bounded to 1. This modification was not needed for the other models, which did not present
419 initialisation issues.

420

421 Source apportionment results for PMF, BAMF, and BAMF+HS are shown in Figure 5 and Table 4. Figure 6 shows the $\mathbf{F}$
422 distributions for these models, as a detail of Figure S9 (a). Figure S6, S9 (a) display very similar results for PMF, BAMF,
423 BAMF+HS both in terms of $\mathbf{F}$, $\mathbf{G}$, and reconstruction metrics, and only some differences can be perceived for PMF, while
424 BAMF and BAMF+HS histograms are almost overlapping in Figure 6. However, the BAMF+HS profiles present a remarkable
425 difference in terms of sparsity as seen in the $\mathbf{F}$ Gini metric, which is mostly the highest for BAMF+HS or equal, except for the
426 biological factor for which PMF is slightly higher. For some species, the relative $\mathbf{F}$ components apportionment is more strongly
427 suppressed by BAMF+HS than by BAMF or PMF, hence, their contribution on other profiles can be larger. This is clearly

428 visible, for instance, for OC, Mg⁺, K⁺, S⁺, or mannitol, which are zeroed in the Salt factor and consequently are larger on the

429 factors where these species are relevant. This is more evidently depicted in Figure S9 (a) and Figure 6, where the distribution

430 of **F** components is shown. For the aforementioned species, the horseshoe effect can be seen in the distribution, whilst BAMF

431 and PMF are further from zero. This result thus highlights the potential benefits of sparsity introduction in matrix factorisation.

432

433 The application of other autocorrelation priors was not advantageous with respect to the regular BAMF autocorrelation and

434 even worsened the shrinkage power of the horseshoe prior as discussed in SI C.III.

3.4 Chemically less sparse synthetic online ACSM datasets

436 The next step to test these models is their trial on more realistic synthetic datasets. For that purpose, 6 datasets for 4 European

437 cities (a total of 24 datasets) were designed with 5 factors in each of them (section 2.4). We applied the 8 models under

438 discussion (PMF, BMF, BMF+HS, BAMF, BAMF+HS, BAMF-AR1, BAMF-GS) to the 24 synthetic datasets and computed

439 the summary statistics (the median of the ratios of **G** over the truth **G**, **G**/**G₀**, the Pearson correlation of **G** with truth, **G** r, and

440 the Spearman correlation of **F** with truth, **F** ρ). All metrics over cities, datasets and sources are presented in Table S5.

442 Figure 7 shows the model summary statistics over the 6 generated datasets for the four cities and Figure S11 shows the factor-

443 dependent statistics. In this case, the Pearson correlation coefficient (not squared) was used to compare the results of these

444 ACSM-like datasets more easily to those presented in Rusanen et al. (2024), in which this metric was used. Broadly speaking,

445 Figure 7 shows a good agreement between models and the truth, with mostly all solutions with correlations with truth for **F**

446 and **G** above 0.7, similarly to Rusanen et al. (2024). However, there are clear differences amongst models and cities. PMF is

447 performing worse in comparison to the Bayesian models, including BMF, the Bayesian analog to PMF in all datasets except

448 for Milan. As shown in Table S5, PMF presents the highest |**Z-X**|/**σ**, the highest overestimations of **G**, and correlations of **G**

449 and **F** are the lowest in comparison to other models except for the Milan dataset. In terms of **G**/**G₀** , the model providing the

450 best results are BAMF, BAMF-AR1, BAMF-GS, followed by their horseshoe versions. The BAMF+HS, presents slightly

451 lower F ρ, F R², and the sparsity Gini metric ratio is not close to one, entailing the horseshoe prior did not successfully

452 implement sparsity and the **F** accuracy did not improve. In terms of correlations with **F** and **G**, the models including the

453 horseshoe prior present higher dispersion within a city with respect to the models without sparsity terms. Considering all the

454 parameters, the models with the best overall performance are BAMF, BAMF-GS, and BAMF+HS.

456 Figure S12 shows the autocorrelation for lags 0-168 h (half of the monthly measurement period) for all the sources and sites,

457 displaying the cyclicity of the selected sources. In all cases, the short-term lags present very high autocorrelation, entailing

458 that the similarity on adjacent timestamps is very high and decays over longer periods. Typically, and as presented on the

459 figure, the autocorrelation of primary sources, with more marked daily cycles, decays faster than secondary sources, which

460 evolve more steadily due to their slower reaction to emissions. Whilst HOA and BBOA present a very steady intradaily

461 structure, with one or two maxima per day, the biogenic SOA presents one peak per day and the other two secondary sources

462 may or may not present marked daily cycles. This different intra- and inter-daily structure amongst sources certainly challenges

463 the models to resolve the source-dependent characteristic.

464

465 Figure 8 shows the modelled vs. truth Pearson correlation coefficients for the autocorrelation of each source in the 4 cities.

466 Each dot represents one of the 6 datasets for each site, and colors represent the different sources. The results show that all

467 models present very high Pearson coefficient ranges for **G** autocorrelations in comparison to truth except for PMF, which

468 struggles with this dataset aspect due to the lack of accounting of self-correlation. In general terms, the best captured correlation

469 by all models is that of SOA_{Bio}, with the most regular cyclical patterns. The SOA_{BB} and SOA_{Tr} autocorrelations seem to

470 challenge the models further due to more irregular patterns, and for some datasets, their autocorrelation is poorly modeled.

471 POA sources are generally accurately modelled, with HOA patterns slightly better captured than those from BBOA. Regarding

472 models, the ones with better performance are BAMF-GS, BAMF, and BAMF+HS, with only slight differences between the

473 last two. This observation suggests that the horseshoe prior addition does not significantly reduce the autocorrelation power of

474 the BAMF.

475

476 Regarding sparsity, Figure S13 depicts the lack of sparsity both for input and modelled data. This figure shows the truth's 5

477 lowest m/z components as well as BAMF, BAMF+HS outcomes. The reference (truth) profiles do not present zeros but very

478 small signals, as do many ACSM-like profiles in the AMS spectral database (Ulbrich et al. 2009). Both BAMF, BAMF+HS

479 reflect this lack of sparsity, however, it could be expected that BAMF+HS would decrease the contributions of the lowest

480 components. However, the sparsity introduction was not achieved as seen before in the lack of improvement of the Gini ratios.

481 This lack of sparsity despite the enforcement through the horseshoe prior can be explained by the complexity of the data, which

482 due to chain divergence, hinders the models performance. Figure S14 shows the model $\hat{R}$, a typical Bayesian metric to evaluate

483 the precision of Hamiltonian chains, computing the ratio between inter- and intra-chain variabilities. In any case results are

484 very close to the ideal value, 1, so the validity of all models' solutions is assured. However, this plot reflects the solution

485 deprecation of the solution with models when the horseshoe prior is applied. The horseshoe prior adds more complexity to the

486 **F** with three more parameters compared to non-sparsity models which could be the cause of the increased model instability

487 across chains.

488

489 Finally, a sensitivity analysis was run for the first Zurich dataset perturbating independently the original **F**, **G** matrices to

490 different degrees, monitoring the correlation of the modelled **F**, **G** matrices to the original truth (Figure 9). Subfigures (a) and

491 (b) show how both in the case of the original **F** and **G** perturbations, the **F** accuracy drops immediately and analogously for

492 both models, with a more sudden decay for **G** perturbations. Contrarily, the affectations in G (subplots (c) and (d)) are different

493 for both models, with a steady decay for BAMF with **G** perturbations and a non-clear trend for **F** perturbations, whilst

494 BAMF+HS correlation rests insensitive to **F**, **G** perturbations with an increasing/decreasing erratic behaviour. This result

495 shows the reduced precision in **G** of BAMF+HS in comparison to BAMF due to the chain divergence issue, which, in any
496 case, does not severely compromise its accuracy. This finding also explains justifies the bigger variations for BAMF+HS with
497 respect to BAMF in all the metrics shown in Figure 8. Additionally, it showcases the general strong sensitivity of **F**
498 determination opposite to the general robustness of **G** upon general **X** matrix perturbation.

3. Discussion

500 This study aims to explore further BAMF capabilities and the benefits introduced through additional priors and/or
501 modifications of the current model structure as given by Rusanen et al. (2024). The introduction of sparsity in source
502 apportionment models was of particular interest to provide more distinct and concise source profiles which can, in turn,
503 improve the time series accuracy. However, in real-world applications, it may also remove small but relevant signals along
504 with noise. Therefore, comparison with BAMF results is recommended, leaving it to the user to decide whether the method's
505 use is appropriate for their case.

507 Firstly, the use of the simplistic toy dataset highlighted the added value of the sparsity introduction through the horseshoe prior
508 in the totally constrained experiment. However, for an unconstrained experiment, sparsity was proven remarkably
509 advantageous, but subject to the underlying matrix factorisation results. That is, the horseshoe prior in BAMF+HS effectively
510 suppresses weak signals of F contributions as determined by BAMF, yet it fails to guide the model toward a more accurate or
511 sparser solution when the initial BAMF estimate is suboptimal. Other sparsity priors, like Lasso and Spike-and-slab, were tried
512 out but did not improve the regularised horseshoe performance.

514 The regularised horseshoe prior introduction in BAMF improved apportionment of offline synthetic and real-world datasets,
515 promoting sparser profiles. The synthetic dataset comparison to truth was maximal for BAMF+HS, with sparser profiles and
516 consequently better **G** accuracy. Its application also proved advantageous for the real-world dataset, despite not being able to
517 be compared to the truth. The results show a sparsity effect which provides more distinct profiles in comparison to PMF and
518 BAMF. This result encourages the usage of the horseshoe prior for sparsity introduction in datasets of expected strong sparsity,
519 such as elemental datasets.

521 Subsequently, in the more realistic European datasets, the sparsity introduction could not be effectively enforced. Although
522 solution quality was not substantially compromised, the profiles remained non-sparse after applying the prior. This is likely
523 due to model instability arising from the higher complexity of these datasets, which is further aggravated by the addition of
524 the horseshoe prior, as it requires sampling a larger number of parameters. Moreover, the inherent nature of ACSM datasets—
525 characterized by highly correlated species—might also contribute to this limitation, since the model struggles to disentangle
526 overlapping sources when variables are strongly interdependent. The higher chain divergence found for the *horseshoed* models,

527 of lower precision due to different landings on the solution space depending on the chain, which could be reduced by selecting

528 chains a-posteriori upon user-defined criterion as is practiced in PMF. This is further confirmed by the insensitivity to **G** or **F**

529 perturbations that the BAMF+HS suffers in comparison to BAMF. Nonetheless, given that ACSM-like factor profiles exhibit

530 low sparsity in the literature, hence the use of sparsity priors in these datasets is less justified.

531

532 The sparsity conceptual framework could also be brought into PMF through the pulling equations, which can shrink down

533 manually the expectedly low signals in a factor. However, this methodology requires that the user indicates the species that

534 are intended to be zeroed, which introduces user-subjectivity to the problem. The BAMF+HS method, contrarily, acts globally,

535 shrinking those species with lowest signals in favour of the matrix factorisation, hence no user intervention is needed. However,

536 if the purpose were to enforce a shrinkage of a certain species as in the PMF case, this feature could also be implemented

537 through the horseshoe method with minimal code modification.

538

539 The other tried out models (BAMF-AR1, BAMF-GS) did not show a significant improvement with respect to BAMF. The

540 BAMF-AR1 contains another autocorrelation to parametrisation (STAN Team, 2025) which should allow for trend

541 consideration, although this matter was not tackled in the current work and remains to be validated in future studies. The

542 BAMF-GS seemed to capture slightly better the G variability in comparison to BAMF in the online datasets, but led to worse

543 correlation to truth in the offline synthetic dataset. Nonetheless, it does not support enforcing sparsity in F, thereby reducing

544 its effectiveness for profile adjustments.

545 **4. Conclusions**

546 This study presents a sparsity introduction technique for the Bayesian Autocorrelated Matrix Factorisation model (BAMF)

547 which intends to essentialise source apportionment profiles removing noisy signals. The regularised horseshoe prior is

548 introduced in BAMF (BAMF+HS) in order to narrow down the lowest signals in factor profiles while keeping the most

549 significant ones regularised. The BAMF+HS model is built in STAN, an open-source framework for statistical modelling with

550 MCMC sampling. In order to test the capabilities of the developed model, we generated three kinds of synthetic datasets to

551 compare the model factorisation outputs to the truth factors, namely Toy, offline, and online synthetic datasets, each

552 representing a progressively increasing level of complexity. Likewise, to confirm its usability to real-world data, BAMF+HS

553 was also applied to a multi-site filter dataset. Given the opportunity to explore source apportionment with different types of

554 datasets, we also tested other receptor models such as Positive Matrix Factorisation (PMF) and other BAMF-like Bayesian

555 models. In the Bayesian framework, we essayed a different formulation of the autocorrelation term (BAMF-AR1) and a

556 permutation on the factorisation matrix logic (BAMF-GS).

557 The main result highlights can be summarised as:

- BAMF+HS has been shown to be advantageous to introduce sparsity in factor profiles for offline datasets and to not deprecate the solution for the more complex ACSM-like datasets. Other sparsifying priors tried out were not as effective in low-signal shrinkage.

- The BAMF+HS performance towards truth profile reconstruction was higher than for BAMF and PMF in the toy and offline synthetic datasets. Improving **F** typically led to a more accurate determination of **G**, highlighting the strong interdependence between the two factorisation matrices.

- The real-world dataset also shows a better description of sources through BAMF+HS in terms of matrix factorisation metrics and profile sparsification achievement.

- As shown in the toy dataset, the introduction of sparsity did not solve factorisation issues inherent to the underlying factorisation model.

- The BAMF+HS model does not create sparsity in ACSM-like datasets, which are originally, indeed, non-sparse. BAMF+HS is more unstable than BAMF for these more complex datasets as a result the higher chain divergence during HMC sampling as suggested by the $\hat{R}$ metric. However, the effects of the horseshoe prior do not affect the overall performance of BAMF or its autocorrelation accuracy.

- The alternative formulations for BAMF, BAMF-AR1 and BAMF-GS, did not show a significant improvement with respect to BAMF.

With all that, profile sparsity has been shown to substantially enhance the accuracy of source apportionment analyses, improving the separation of the chemical composition of sources. The BAMF+HS model succeeds in incorporating this property in profile fingerprints, especially in filter-based datasets. With the aim of improving further source apportionment techniques, future research should be directed to enhance the robustness and generalisability of the BAMF+HS model across diverse data types. Moreover, continued exploration of the underlying properties of solution spaces — such as profiles sparsity, time series autocorrelation — may provide valuable insights into disentangling complex source contributions through receptor modelling. In this regard, the Bayesian source apportionment framework offers a particularly suitable foundation, allowing for the integration of prior knowledge and uncertainty quantification in the inference process.

Code and data availability

The models and datasets can be found at https://github.com/martavia0/BAMF-horseshoe.git

Author contribution

MV: Conceptualisation, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, validation, visualisation, writing (original draft preparation). YH: Formal analysis, investigation, software; JD: investigation, resources, software, validation. MM: Data curation. AR: Data curation, formal

588 analysis, methodology, investigation, resources, software. JJ: Data curation. SKG: Data curation. J-LJ: Data curation; VNTD:
589 Data curation. GU: Data curation. GM: conceptualisation, funding acquisition, investigation, supervision, validation. KRD:
590 Conceptualisation, data curation, formal analysis, funding acquisition, investigation, methodology, supervision, validation. All
591 co-authors participated in the revision and edition of the manuscript.
592

593 **Competing interests**

594 The authors declare that they have no conflict of interest.
595

596 **Disclaimer**
597

598 Co - funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not
599 necessarily reflect those of the European Union or European Research Exacutive Agency. Neither the European Union nor the
600 granting authority can be held responsible for them.

601 **Acknowledgements**

602 This publication is co- funded by/ has received funding from/ the European Union's Horizon Europe research and innovation
603 program under the Marie Sklodowska-Curie COFUND Postdoctoral Programme grant agreement No.101081355- SMASH
604 and by/from /the Republic of Slovenia and the European Union from the European Regional Development Fund. Work was
605 partly supported by ARIS program P1-0385. All the chemical measurements on the PM10 and PM2.5 Swiss samples used in
606 this study were performed on the Air O Sol analytical plateau at IGE (France). K.R.D. acknowledges support by the SNSF
607 Ambizione grant PZPGP2_201992. The authors thank Jimeng Wu and Imad El Haddad for their input related to
608 autocorrelation.

609

610 **References**

611 Andersen, M. R., Winther, O., & Hansen, L. K. (2014). Bayesian inference for structured spike and slab priors. Advances in
612 Neural Information Processing Systems, 27.
613 Belis, C. A., Karagulian, F., Larsen, B. R., & Hopke, P. K. (2013). Critical review and meta-analysis of ambient particulate
614 matter source apportionment using receptor models in Europe. Atmospheric Environment, 69, 94-108.
615 Belis, C. A., Larsen, B. R., Amato, F., El Haddad, I., Favez, O., Harrison, R. M., ... & Viana, M. (2014). European guide on
616 air pollution source apportionment with receptor models. JRC reference reports EUR26080 EN.

617 Belis, C., Pernigotti, D., Karagulian, F., Pirovano, G., Larsen, B., Gerboles, M., and Hopke, P.: A new methodology to assess

618 the performance and uncertainty of source apportionment models in intercomparison exercises, Atmospheric Environment,

619 119, 35–44, 2015.

620 Brinkman, G., Vance, G., Hannigan, M. P., and Milford, J. B.: Use of synthetic data to evaluate positive matrix factorization

621 as a source apportionment tool for PM2. 5 exposure data, Environmental science & technology, 40, 1892–1901, 2006.

622 Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... & Riddell, A. (2017). Stan: A

623 probabilistic programming language. Journal of statistical software, 76, 1-32.

624 Crippa, M., DeCarlo, P. F., Slowik, J. G., Mohr, C., Heringa, M. F., Chirico, R., ... & Baltensperger, U. (2013). Wintertime

625 aerosol chemical composition and source apportionment of the organic fraction in the metropolitan area of Paris. Atmospheric

626 Chemistry and Physics, 13(2), 961-981.

627 Dai, T., Dai, Q., Yin, J., Chen, J., Liu, B., Bi, X., ... & Feng, Y. (2024). Spatial source apportionment of airborne coarse

628 particulate matter using PMF-Bayesian receptor model. Science of The Total Environment, 917, 170235.

629 Daellenbach, K. R., Uzu, G., Jiang, J., Cassagnes, L. E., Leni, Z., Vlachou, A., ... & Prévôt, A. S. (2020). Sources of particulate-

630 matter air pollution and its oxidative potential in Europe. Nature, 587(7834), 414-419.

631 Daellenbach, K. R., Manousakas, M., Jiang, J., Cui, T., Chen, Y., El Haddad, I., ... & Prévôt, A. S. H. (2023). Organic aerosol

632 sources in the Milan metropolitan area–Receptor modelling based on field observations and air quality modelling. Atmospheric

633 Environment, 307, 119799.

634 Dinh, V. N. T., Uzu, G., Dominutti, P., Sauvage, S., Elazzouzi, R., Darfeuil, S., Voiron, C., Samaké, A., Zhang, S., Socquet,

635 S., Favez, O., and Jaffrezo, J.-L.: Toolbox for accurate estimation and validation of PMF solutions in PM source apportionment,

636 EGUsphere [preprint], https://doi.org/10.5194/egusphere-2025-1968, 2025.

637 Elser, M., Huang, R. J., Wolf, R., Slowik, J. G., Wang, Q., Canonaco, F., ... & Prévôt, A. S. (2016). New insights into PM 2.5

638 chemical composition and sources in two major cities in China during extreme haze events using aerosol mass spectrometry.

639 Atmospheric Chemistry and Physics, 16(5), 3207-3225.

640 Gelman, A., & Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. Statistical Science, 7(4),

641 457–472. https://doi.org/10.1214/ss/1177011136

642 Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B.: Bayesian data analysis, 3rd edn., CRC

643 Press, ISBN 9781439898208, 2014.

644 Gini, Corrado (1936). "On the Measure of Concentration with Special Reference to Income and Statistics", Colorado College

645 Publication, General Series No. 208, 73–79

646 Govan, E., Jackson, A. L., Inger, R., Bearhop, S., & Parnell, A. C. (2023). simmr: A package for fitting stable isotope mixing

647 models in R. arXiv preprint arXiv:2306.07817.

648 Grange, S. K., Fischer, A., Zellweger, C., Alastuey, A., Querol, X., Jaffrezo, J. L., ... & Hueglin, C. (2021). Switzerland's

649 PM10 and PM2. 5 environmental increments show the importance of non-exhaust emissions. Atmospheric environment: X,

650 12, 100145.

651 Hoffman, M. D. and Gelman, A.: The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo, J.

652 Mach. Learn. Res., 15, 1593–1623, 2014.

653 Jiang, J., Aksoyoglu, S., El-Haddad, I., Ciarelli, G., Denier van der Gon, H. A., Canonaco, F., ... & Prévôt, A. S. (2019).

654 Sources of organic aerosols in Europe: A modelling study using CAMx with modified volatility basis set scheme. Atmospheric

655 Chemistry and Physics Discussions, 2019, 1-35.

656 Keuken, M. P., Moerman, M., Voogt, M., Blom, M., Weijers, E. P., Röckmann, T., & Dusek, U. (2013). Source contributions

657 to PM2. 5 and PM10 at an urban background and a street location. Atmospheric Environment, 71, 26-35.

658 Kuhn, H. W. (1955). The Hungarian method for the assignment problem. Naval research logistics quarterly, 2(1-2), 83-97.

659 Lingwall, J. W., & Christensen, W. F. (2007). Pollution source apportionment using a priori information and positive matrix

660 factorization. Chemometrics and Intelligent Laboratory Systems, 87(2), 281-294.

661 Liu, D. C. and Nocedal, J.: On the Limited Memory BFGS Method for Large Scale Optimization, Math. Program., 45, 503–

662 528, https://doi.org/10.1007/BF01589116, 1989.

663 Manousakas, M. I., Rausch, J., Jaramillo Vogel, D., Schneider-Beltran, K., Alastuey, A., Jaffrezo, J. L., ... & Dällenbach, K.

664 R. Comparison of PM Source Profiles Identified by Different Techniques and the Potential of Utilizing Single-Particle

665 Analysis Data in Source Apportionment. Available at SSRN 5323830.

666 Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by

667 fast computing machines. The Journal of Chemical Physics, 21(6), 1087–1092. https://doi.org/10.1063/1.1699114

668 Oh, M. S., & Park, C. K. (2022). Regional source apportionment of PM2. 5 in Seoul using Bayesian multivariate receptor

669 model. Journal of Applied Statistics, 49(3), 738-751.

670 Paatero, P., & Tapper, U. (1994). Positive matrix factorization: A non-negative factor model with optimal utilization of error

671 estimates of data values. Environmetrics, 5(2), 111-126.

672 Paatero, P. (1999). The multilinear engine—a table-driven, least squares program for solving multilinear problems, including

673 the n-way parallel factor analysis model. Journal of Computational and Graphical Statistics, 8(4), 854-888.

674 Park, E. S., Guttorp, P., & Henry, R. C. (2001). Multivariate receptor modeling for temporally correlated data by using

675 MCMC. Journal of the American Statistical Association, 96(456), 1171-1183.

676 Park, E. S., Spiegelman, C. H., & Henry, R. C. (2002). Bilinear estimation of pollution source profiles and amounts by using

677 multivariate receptor models. Environmetrics, 13(7), 775-798.

678 Park, E. S., Lee, E. K., & Oh, M. S. (2021). Bayesian multivariate receptor modeling software: BNFA and

679 bayesMRM. Chemometrics and Intelligent Laboratory Systems, 211, 104280.

680 Piironen, J. and Vehtari, A.: Sparsity information and regularization in the horseshoe and other shrinkage priors, Electronic

681 Journal of Statistics, 11, 5018–5051, https://doi.org/10.1214/17-EJS1337SI, 2017.

682 Rasmussen, M. A., & Bro, R. (2012). A tutorial on the Lasso approach to sparse modeling. Chemometrics and Intelligent

683 Laboratory Systems, 119, 21-31.

Pope III, C. A., & Dockery, D. W. (1999). Epidemiology of particle effects. In Air pollution and health (pp. 673-705). Academic Press.

Rusanen, A., Bjorklund, A., Manousakas, M. I., Jiang, J., Kulmala, M. T., Puolamaki, K., and Daellenbach, K. R.: A novel probabilistic source apportionment approach: Bayesian auto-correlated matrix factorization, Atmospheric Measurement Techniques, 17, 1251–1277, https://doi.org/10.5194/amt-17-1251-2024, 2024.

Sage, A. M. (2007). Evolving mass spectra of the oxidized component of organic aerosol mass spectrometer analysis of aged diesel emissions. Atmos. Chem. Phys. Discuss., 7, 10065-10096.

STAN Development Team. (2025). *Stan user's guide* (Version 2.36). https://mc-stan.org/docs/stan-users-guide/. Accessed July 2025.

Tobler, A. K., Skiba, A., Canonaco, F., Močnik, G., Rai, P., Chen, G., ... & Prevot, A. S. (2021). Characterization of non-refractory (NR) PM 1 and source apportionment of organic aerosol in Kraków, Poland. Atmospheric chemistry and physics, 21(19), 14893-14906.

Tibshirani, R., & Wasserman, L. (2015). Sparsity and the lasso. Statistical machine learning, 1-15.

Ulbrich, I. M., Handschy, A., Lechner, M., and Jimenez, J. L.: AMS Spectral Database, http://cires.colorado.edu/jimenez-group/AMSsd/, accessed: 2025-05-19, n.d.

Upadhyay, A., Jiang, J., Cheng, Y., Vasilakos, P., Chen, Y., Banos, D. T., ... & El-Haddad, I. (2025). High-resolution modelling of particulate matter chemical composition over Europe: brake wear pollution. Environment International, 109615.

Via, M., Minguillón, M. C., Reche, C., Querol, X., & Alastuey, A. (2021). Increase in secondary organic aerosol in an urban environment. Atmospheric chemistry and physics, 21(10), 8323-8339.

Via, M., Chen, G., Canonaco, F., Daellenbach, K. R., Chazeau, B., Chebaicheb, H., Jiang, J., Keernik, H., Lin, C., Marchand, N., et al.: Rolling vs. seasonal PMF: real-world multi-site and synthetic dataset comparison, Atmospheric measurement techniques, 15, 5479–5495, 2022.

Viana, M., Kuhlbusch, T. A., Querol, X., Alastuey, A., Harrison, R. M., Hopke, P. K., ... & Hitzenberger, R. (2008). Source apportionment of particulate matter in Europe: a review of methods and results. Journal of aerosol science, 39(10), 827-849.

Yang, Y., Ruan, Z., Wang, X., Yang, Y., Mason, T. G., Lin, H., & Tian, L. (2019). Short-term and long-term exposures to fine particulate matter constituents and health: A systematic review and meta-analysis. Environmental pollution, 247, 874-882.

Zhang, Y., Albinet, A., Petit, J. E., Jacob, V., Chevrier, F., Gille, G., ... & Favez, O. (2020). Substantial brown carbon emissions from wintertime residential wood burning over France. Science of the Total Environment, 743, 140752.

Zhang, Y., Ma, X., Tang, A., Fang, Y., Misselbrook, T., & Liu, X. (2023). Source Apportionment of Atmospheric Ammonia at 16 Sites in China Using a Bayesian Isotope Mixing Model Based on δ15N–NH x Signatures. Environmental Science & Technology, 57(16), 6599-6608.

Zheng, Y., Xue, T., Zhao, H., & Lei, Y. (2022). Increasing life expectancy in China by achieving its 2025 air quality target. Environmental science and ecotechnology, 12, 100203.

718

719

Figures

Table 1. Priors used in the models tested in this manuscript.

Model	G priors	F priors
BMF	None	None
BMF+HS	None	Regularised horseshoe
BAMF	Rusanen et al. (2024)	None
BAMF+HS	Rusanen et al. (2024)	Regularised horseshoe
BAMF-AR1	AR(1)	None
BAMF-AR1+HS	AR(1)	Regularised horseshoe
BAMF-GS	Rusanen et al. (2024)	None
PMF	None	None
CMB	Fixed a-priori.	None
CMB+HS	Fixed a-priori.	Regularised horseshoe

723

Table 2. Sparsity of the F factors of the synthetic datasets.

725

Dataset	Factor	F Gini	% zeros
Chemically-sparse synthetic toy dataset	Factor 1	0.67	75.0
	Factor 2	0.67	75.0
	Factor 3	0.5	25.0
Chemically-sparse synthetic offline dataset	Dust	0.74	37.5
	Traffic	0.86	12.5

	Salt		0.78	37.5
	Coarse biological		0.88	25.0
Less chemically-sparse synthetic online ACSM datasets	HOA	Krakow	0.74	5.0
		Milan	0.67	0.0
		Paris	0.68	0.0
		Zurich	0.68	0.0
	BBOA	Krakow	0.47	1.2
		Milan	0.72	0.0
		Paris	0.74	0.0
		Zurich	0.58	0.0
	SOA_{bio}	Krakow	0.52	0.0
		Milan	0.50	0.0
		Paris	0.50	0.0
		Zurich	0.67	0.0
	SOA_{BB}	Krakow	0.55	0.0
		Milan	0.53	0.0
		Paris	0.53	0.0
		Zurich	0.73	0.0

	SOA_{TR}	Krakow	0.45	0.0
		Milan	0.42	0.0
		Paris	0.46	0.0
		Zurich	0.60	0.0

726

Table 3. Toy experiment statistics of (a) Factorisation performance. (b) Comparison to truth. Green sequential colorscales represent variables whose larger value leans to a better performance and the blue-to-red divergent colorscales (centered at 1, in white) represent G/G_0 divergence with respect to 1. Red bars in (a) depict deviations from the ideal 0 value.

731

(a)

Model	R^2	Median(|Z-X|/sigma)	Max(|Z-X|/sigma)
CMB	0.9985	0.2542	0.5755
CMB+HS	0.9996	0.2648	0.4774
PMF	0.9979	0.2658	0.5847
BMF	0.9650	0.2125	1.2917
BMF+HS	0.9689	0.2107	1.2863
BAMF	0.9818	0.1222	0.7247
BAMF+HS	0.9820	0.1208	0.7180
BAMF-AR1	0.9806	0.0947	1.0257
BAMF-AR1+HS	0.9790	0.1171	1.0297
BMF-GS	0.9657	0.2024	1.3031
BAMF-GS	0.9818	0.1576	0.9082

(b)

Model	Sources	G G/G_0	G R^2	F ρ	F R^2	F Gini	F Gini ratio	F Zeros sum
CMB	Source1	1.00	1.00	0.82	1.00	0.61	0.91	0.06
CMB	Source2	1.00	1.00	0.82	1.00	0.60	0.90	0.08
CMB	Source3	1.00	1.00	0.96	1.00	0.44	0.92	0.03
CMB+HS	Source1	1.00	1.00	0.82	1.00	0.63	0.95	0.03
CMB+HS	Source2	1.00	1.00	0.82	1.00	0.63	0.95	0.04
CMB+HS	Source3	1.00	1.00	0.96	1.00	0.47	0.96	0.02
PMF	Source1	1.00	0.96	0.82	0.95	0.35	0.52	0.36
PMF	Source2	3.29	0.97	0.82	0.88	0.30	0.45	0.42
PMF	Source3	0.65	0.74	0.79	0.81	0.31	0.64	0.2
BMF	Source1	1.08	0.87	0.82	0.76	0.25	0.37	0.48
BMF	Source2	4.10	0.84	0.51	0.43	0.16	0.23	0.58
BMF	Source3	0.60	0.46	0.79	0.80	0.21	0.44	0.25
BMF+HS	Source1	1.11	0.88	0.82	0.77	0.25	0.38	0.48
BMF+HS	Source2	3.92	0.85	0.82	0.50	0.15	0.23	0.57
BMF+HS	Source3	0.60	0.44	0.79	0.84	0.22	0.46	0.23
BAMF	Source1	1.66	0.98	0.82	0.78	0.27	0.41	0.46
BAMF	Source2	1.99	0.95	0.82	0.97	0.33	0.50	0.37
BAMF	Source3	0.54	0.50	0.96	1.00	0.38	0.79	0.08
BAMF+HS	Source1	1.96	0.98	0.82	0.76	0.27	0.40	0.47
BAMF+HS	Source2	1.28	0.95	0.82	0.99	0.58	0.86	0.11
BAMF+HS	Source3	0.51	0.48	0.96	1.00	0.44	0.93	0.02
BAMF-AR1	Source1	1.70	0.98	0.82	0.79	0.27	0.41	0.46
BAMF-AR1	Source2	3.05	0.92	0.82	0.89	0.37	0.55	0.36
BAMF-AR1	Source3	0.42	0.50	0.96	0.94	0.43	0.90	0.08
BAMF-AR1+HS	Source1	1.92	0.99	0.82	0.77	0.26	0.39	0.46
BAMF-AR1+HS	Source2	2.58	0.89	0.82	0.90	0.48	0.72	0.36
BAMF-AR1+HS	Source3	0.38	0.49	0.96	0.97	0.49	1.03	0.08
BMF-GS	Source1	0.88	0.92	0.82	0.63	0.21	0.31	0.52
BMF-GS	Source2	1.12	0.86	0.82	0.71	0.20	0.30	0.53
BMF-GS	Source3	1.02	0.29	0.79	0.91	0.22	0.46	0.22
BAMF-GS	Source1	0.89	0.94	0.82	0.66	0.21	0.32	0.53
BAMF-GS	Source2	1.12	0.95	0.82	0.70	0.19	0.28	0.53
BAMF-GS	Source3	1.02	0.36	0.79	0.90	0.22	0.46	0.22
BAMF-Lasso	Source1	1.76	0.98	0.82	0.76	0.27	0.40	0.47
BAMF-Lasso	Source2	2.02	0.95	0.82	0.95	0.34	0.50	0.37
BAMF-Lasso	Source3	0.49	0.47	0.96	1.00	0.41	0.86	0.05
BAMF-Spike-Slab	Source1	1.57	0.97	0.82	0.77	0.26	0.40	0.47
BAMF-Spike-Slab	Source2	1.78	0.95	0.82	0.99	0.34	0.52	0.34
BAMF-Spike-Slab	Source3	0.62	0.50	0.96	0.98	0.33	0.69	0.12

732

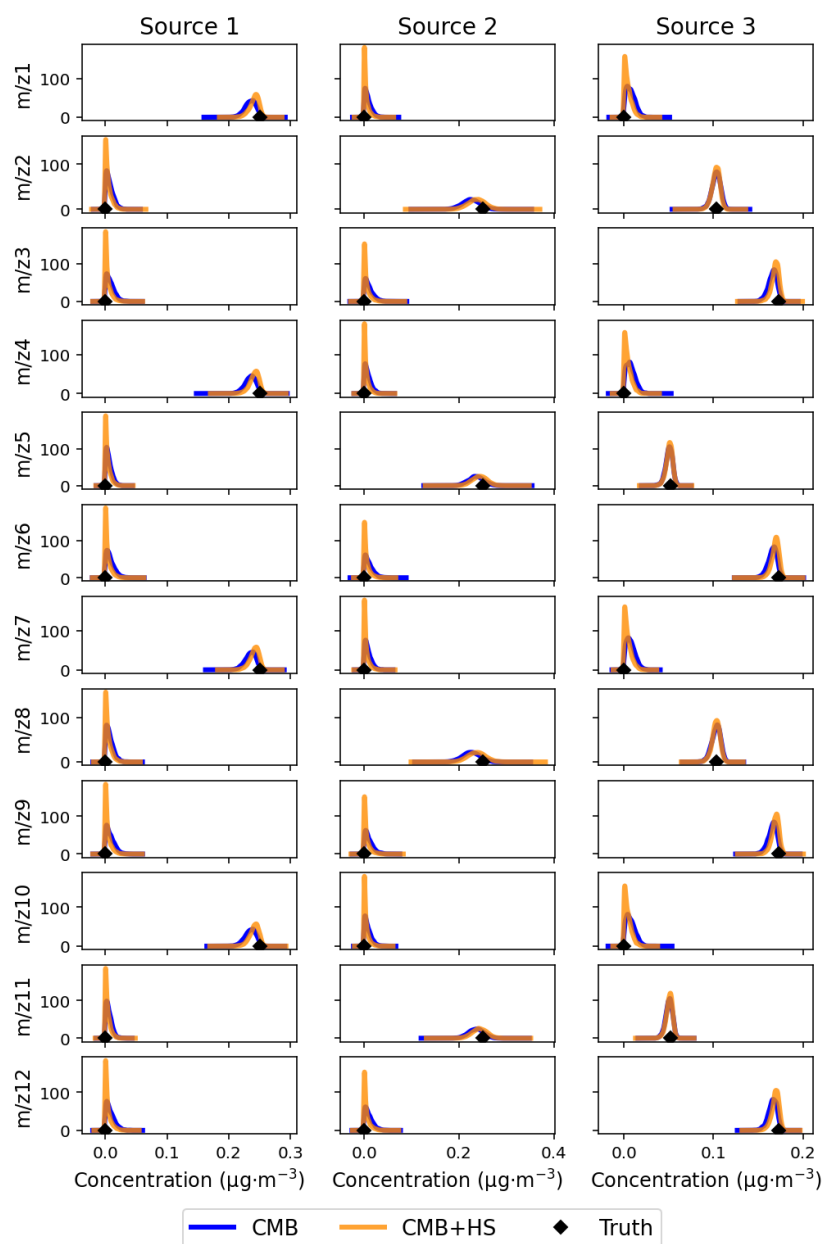
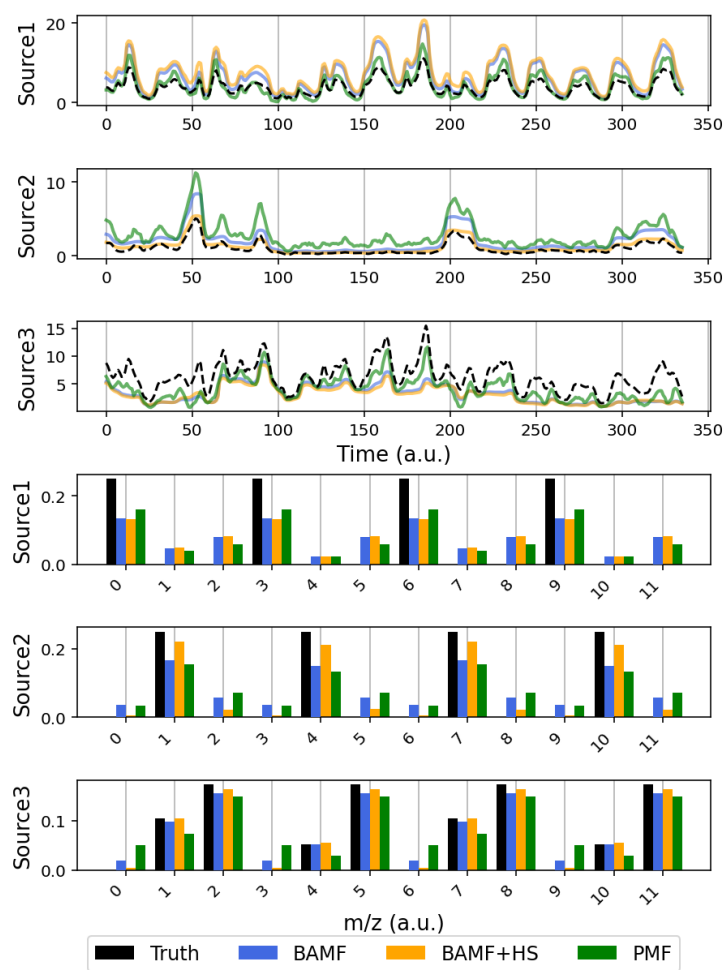



Figure 1. F components distributions for CMB and CMB+HS compared to truth.

736

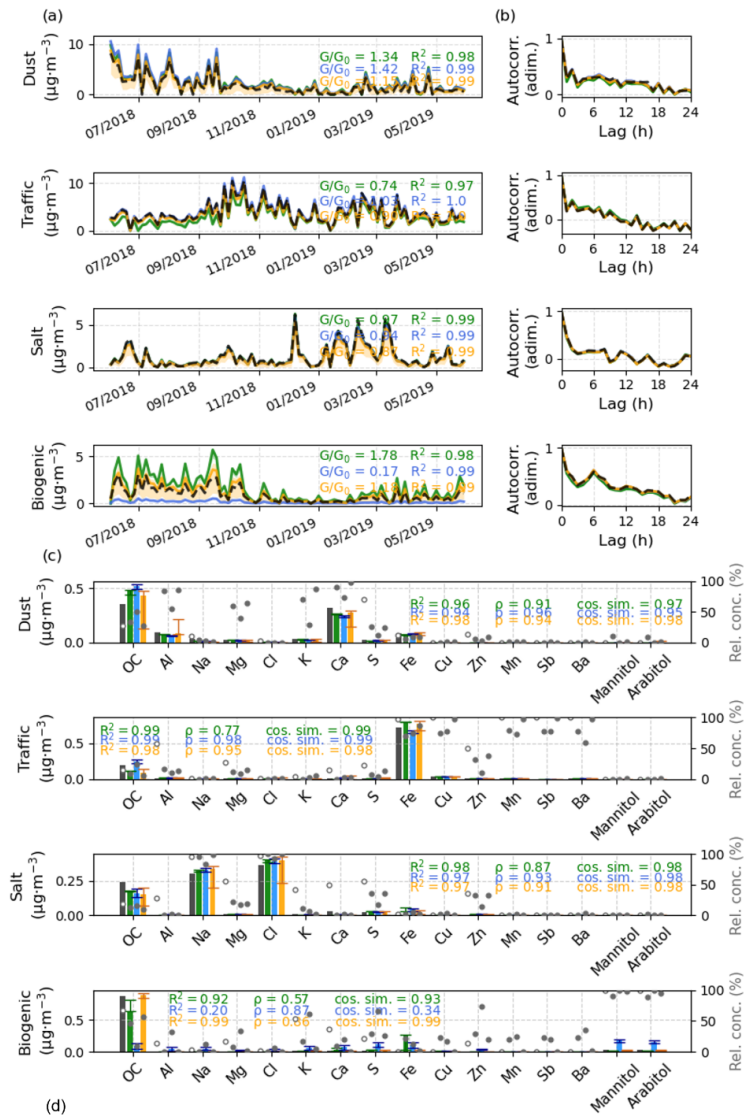
737 **Figure 2. Toy dataset source apportionment results for PMF, BAMF, and BAMF+HS models. (a) Time Series. (b) Profiles.**

Factor	Model	F contr. R^2	F contr. ρ	F Gini	F Gini ratio (*)	Truth zeros sum
Dust	PMF	0.95	0.84	0.79	0.90	0.0035
	BAMF	0.77	0.90	0.80	1.08	0.0025
	BAMF+HS	0.94	0.90	0.79	1.07	0.0011
Traffic	PMF	0.99	0.90	0.89	1.19	0.0000
	BAMF	0.93	0.94	0.85	0.99	0.0005
	BAMF+HS	0.995	0.93	0.88	1.02	0.0004
Salt	PMF	0.95	0.84	0.79	0.90	0.0012
	BAMF	0.95	0.88	0.80	1.01	0.0118
	BAMF+HS	0.98	0.87	0.82	1.04	0.0031
Biogenic	PMF	0.87	0.43	0.85	1.08	0.0035
	BAMF	0.53	0.51	0.50	0.56	0.0118
	BAMF+HS	0.96	0.84	0.90	1.01	0.0007
$\sum^k$ (*)	PMF	3.80	3.04	3.34	0.43	0.0198
	BAMF	3.18	3.23	2.95	0.53	0.0172
	BAMF+HS	**3.88**	**3.54**	**3.39**	**0.15**	**0.0031**

Figure 3. Synthetic offline dataset source apportionment results for PMF, BAMF, and BAMF+HS models. (a) Time Series. (b) Autocorrelation. (c) Profiles. (d) Table with additional metrics for comparison to truth. Bold numbers reflect the highest value amongst models. F contr. represents here the percentage of each factor into a given species. The sum row reflects the overall performance of the model for all sources for each statistic metric except for the ones marked with (*), in which the difference to 1 in absolute value is summed up.

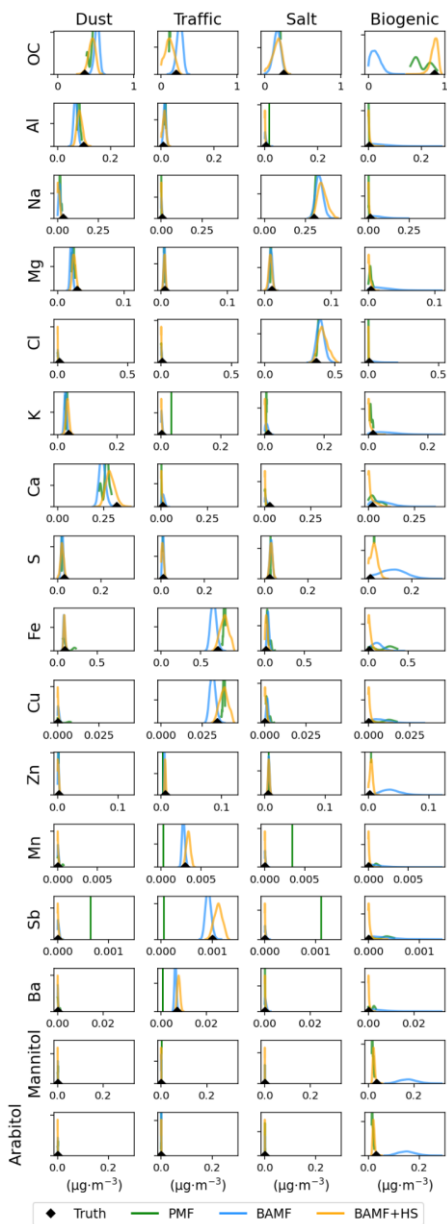



Figure 4. Profile components distribution for PMF, BAMF, BAMF+HS on the real-world filters dataset

748

749

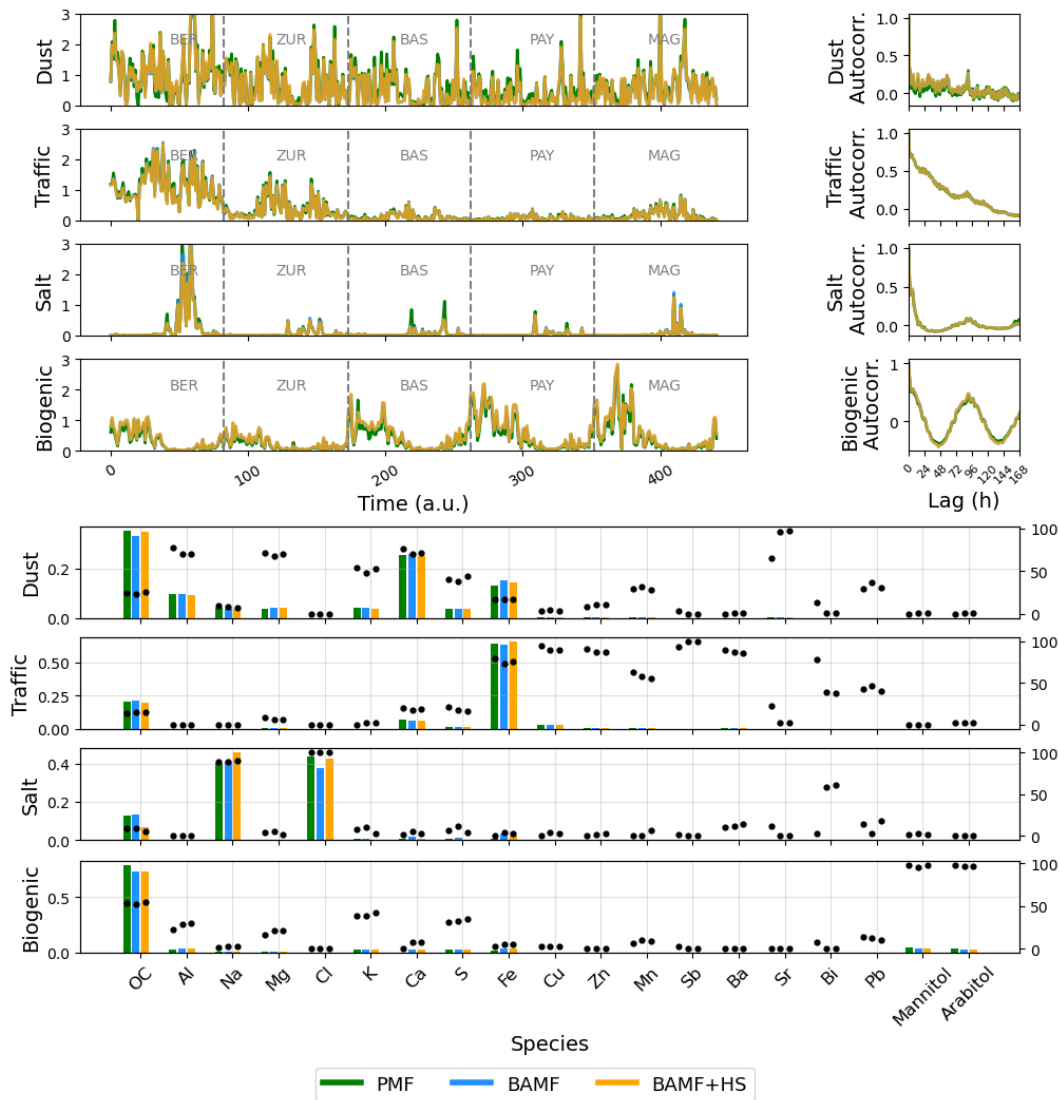



Figure 5. Comparison of PMF, BAMF, BAMF+HS for the real-world filters dataset. From left to right and top to bottom: time series, autocorrelation, and profile plots. The dots in the profiles (right axis) show the contribution of each species to the source.

Table 4. Offline real-world dataset reconstruction and sparsity statistics. Bold numbers reflect the highest value amongst models.

Model	R^2 (Z, X)	Median |X-Z|/σ	Median |X-Z|/σ	Factor	F Gini
PMF	0.68	0.77	10.52	Dust	**0.77**

				Traffic	**0.87**
				Salt	0.86
				Biogenic	**0.87**
BAMF	0.67	0.75	11.13	Dust	0.76
				Traffic	**0.87**
				Salt	0.84
				Biogenic	0.83
BAMF+HS	0.67	0.75	11.12	Dust	**0.77**
				Traffic	**0.87**
				Salt	**0.87**
				Biogenic	0.83

754

755

756

757

758

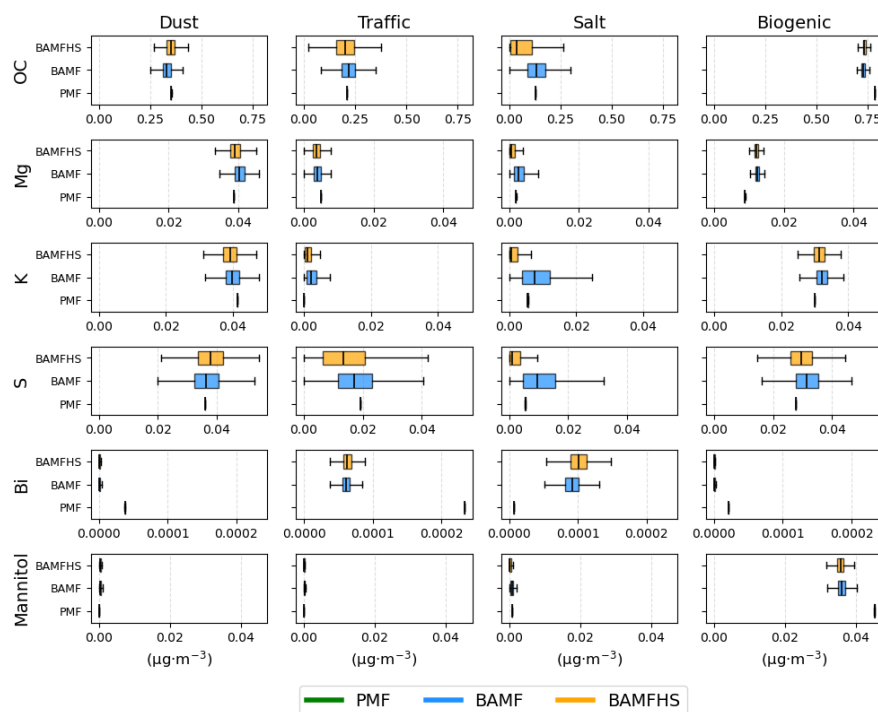



Figure 6. Distribution of individual profile components derived from PMF, BAMF, and BAMF+HS analyses for the real-world filter dataset. A complete comparison of all profiles is presented in Figure S9.

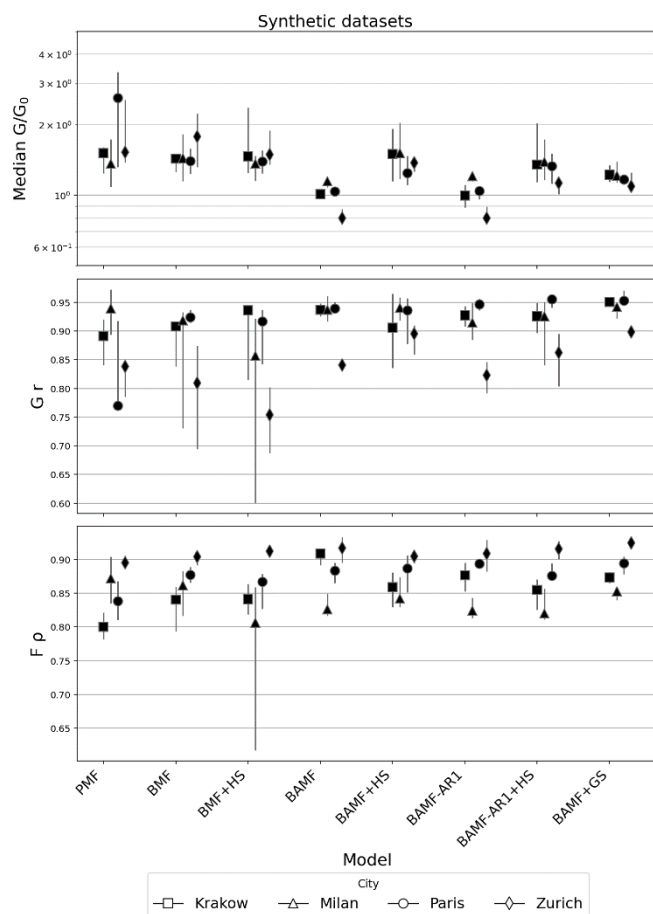



Figure 7. European cities synthetic datasets summary statistics; from top to bottom, median ratio time series with truth, Pearson correlation coefficient of G with truth, Spearman correlation coefficient of F with truth. The axis of the G/G_0 plot is in logarithmic scale.

767

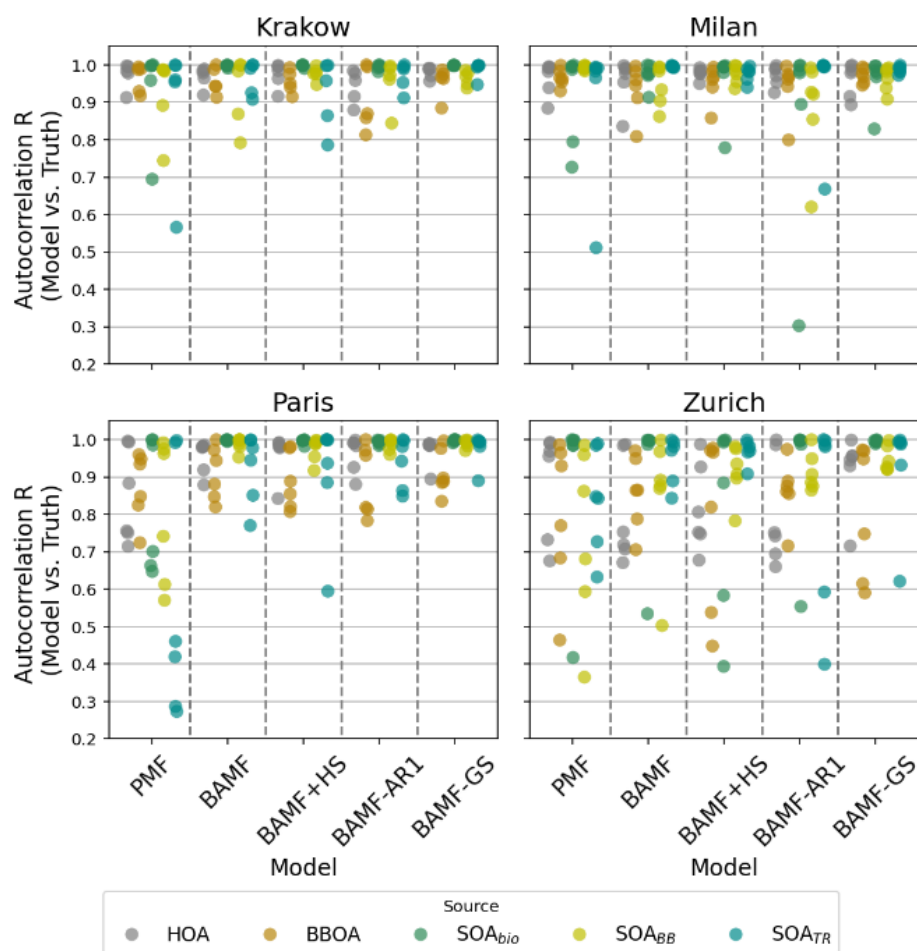



Figure 8. Pearson correlation of the autocorrelations of model solutions with the truth for all factors and all cities.

768

769

770

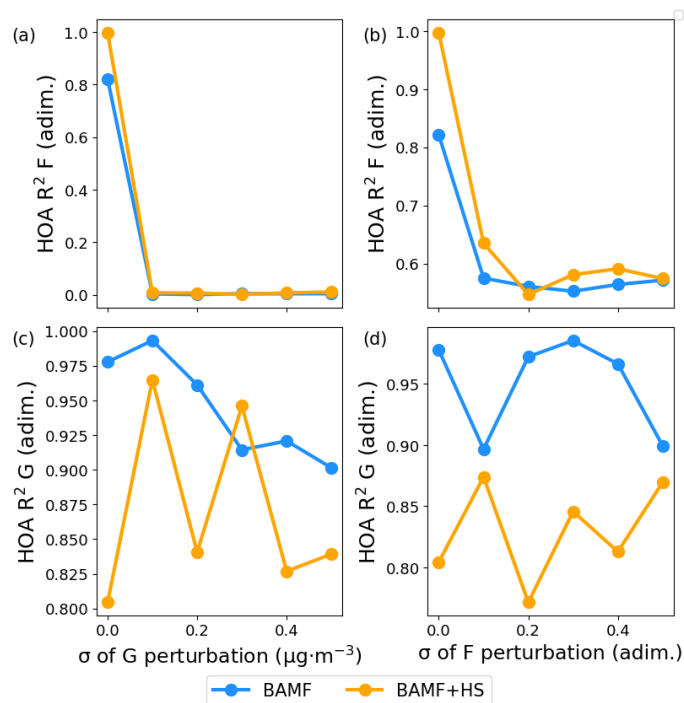



Figure 9. Squared Pearson coefficient of F, G matrix with original truth F, G matrices of the BAMF, BAMF+HS models with the degrees of perturbation in F and G.