



HAL
open science

A model for simultaneous evaluation of NO₂, O₃, and PM₁₀, pollution in urban and rural areas: handling incomplete data sets with multivariate curve resolution analysis

Eva Gorrochategui, Isabel Hernández, Romà Tauler

► To cite this version:

Eva Gorrochategui, Isabel Hernández, Romà Tauler. A model for simultaneous evaluation of NO₂, O₃, and PM₁₀, pollution in urban and rural areas: handling incomplete data sets with multivariate curve resolution analysis. *Atmospheric Chemistry and Physics*, 2022, 22 (13), pp.9111-9127. 10.5194/acp-22-9111-2022 . hal-03763561

HAL Id: hal-03763561

<https://hal.science/hal-03763561>

Submitted on 29 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



1 **A model for simultaneous evaluation of NO₂, O₃ and PM₁₀**
2 **pollution in urban and rural areas: handling incomplete**
3 **data sets with multivariate curve resolution analysis**

4 Eva Gorrochategui^{1,3*}, Isabel Hernandez², Romà Tauler^{3,*}

5 ¹ University of Rennes, Inserm, EHESP, Irset (Institut de Recherche en Santé, Environnement et Travail)—
6 UMR_S 1085, F-35000 Rennes, France.

7 ²Direcció General de Qualitat Ambiental i Canvi Climàtic, Generalitat de Catalunya, Spain

8 ³ Institute of Environmental Assessment and Water Research (IDAEA), Consejo Superior de
9 Investigaciones Científicas (CSIC), Barcelona, 08034, Spain.

10

11 *Correspondence to:* Eva Gorrochategui (eva.gorrochateguimatas@ehesp.fr) and Romà Tauler
12 (rtaqam@cid.csic.es)

13 **Abstract.** A powerful methodology, based on multivariate curve resolution alternating least squares (MCR-
14 ALS) with quadrilinearity constraints, is proposed to handle complex and incomplete four-way atmospheric
15 data sets, providing concise and easy interpretable results. Changes in air quality by nitrogen dioxide (NO₂),
16 ozone (O₃) and particular matter (PM₁₀) in eight sampling stations located in Barcelona metropolitan area
17 and other parts of Catalonia during the COVID-19 lockdown (2020) with respect to previous years (2018
18 and 2019) are investigated using such methodology. MCR-ALS simultaneous analysis of the 3
19 contaminants among the 8 stations and for the 3 years allows the evaluation of potential correlations among
20 the pollutants even when having missing data blocks. NO₂ and PM₁₀ show correlated profiles due to similar
21 pollution sources (traffic and industry), evidencing a decrease in 2019 and 2020 due to traffic restriction
22 policies and COVID-19 lockdown, especially noticeable in the most transited urban areas (*i.e.*, Vall
23 d'Hebron, Granollers and Gràcia). Ozone evidences an opposed inter-annual trend, showing higher
24 amounts in 2019 and 2020 respect to 2018 due to the decreased titration effect, more significant in rural
25 areas (Begur) and in the control site (Observatori Fabra).

26

27 **Keywords:** Multivariate curve resolution; MCR-ALS; Multilinear modeling; incomplete data; NO₂; O₃;
28 PM₁₀; COVID-19; lockdown.



29 1. Introduction

30 Monitoring studies of air quality have always been indispensable to assess the impact of air pollutants on
31 human health and the environment. Most evaluated air pollutants include the ones linked to industrial and
32 traffic emissions, such as tropospheric ozone (O_3), nitrogen dioxide (NO_2) and particulate matter (PM), due
33 to its potential effects on human health (Zú et al., n.d.; Khaniabadi et al., n.d.), and are the chemicals
34 evaluated in the present study.

35 The chemistry of nitrogen oxides (NO_x) and O_3 is highly complex because NO_x is the responsible for
36 tropospheric O_3 production but also for its elimination (Lerdau et al., 2000; Crutzen, 1979). On the one
37 hand, the formation of tropospheric ozone is a consequence of the photochemical reaction of the sunlight
38 with NO_x and volatile organic compounds (VOC) released by car exhausts and industries, according to the
39 following equation: $NO_x + VOC + h\nu \rightarrow O_3$. Thus, nitrogen oxides behave as catalysts in the photochemical
40 production of ozone, especially at higher solar radiation and during hours of high traffic. However, at hours
41 of low solar radiation and during nighttime, NO_x are responsible for the ozone destruction in a process
42 called titration: $NO_x + O_3 \rightarrow NO_x + O_2$. In inner rural areas with low anthropogenic activities, the latter
43 titration effect produced by NO_x emissions is generally not observed, resulting in higher average O_3
44 concentrations than in urban areas. Overall, the complex equilibrium among O_3 and NO_x species results in
45 continuous concentration changes of ozone difficult to attribute to a unique source.

46 Conversely, the chemistry of particulate matter is not directly correlated to NO_x and O_3 but it is also complex
47 due to its multiple and diverse emission sources. Different PM sources exist including city background
48 (background levels of emissions such as construction, demolition and domestic heating), traffic (motor
49 emissions and tire, pavement and brakes abrasion products), industry (high levels of sulfate, nitrate and
50 other burning products), and natural (*i.e.*, marine aerosols and air masses, especially African dust) (Querol
51 et al., 2004; Saud et al., 2011).

52 Different approaches exist to assess air quality by evaluating concentration changes of these chemical
53 pollutants. In classical air quality monitoring studies, the data treatment strategy generally involves data
54 arrangement and analysis using traditional statistics. However, these methods require extensive computer
55 calculations and their results are often limited and restricted. Instead, chemometric methods are powerful
56 data analysis tools to investigate the sources of data variance in experimentally measured environmental
57 monitoring big data sets such as air quality data sets that often contain some missing blocks. These methods
58 can be used to extract and summarize the information often hidden on these environmental big data sets.
59 Among these methods, Multivariate Curve Resolution Alternating Least Squares (MCR-ALS) (Tauler,
60 1995), originally used in the spectrochemical analysis of chemical mixtures, has been also proved to be a
61 competitive method in air pollution studies (Malik and Tauler, 2013; Alier et al., 2011). MCR-ALS is a
62 flexible soft-modelling factor analysis method that allows for the introduction of natural constraints, like
63 non-negativity of the factor solutions. Although it only requires the fulfillment of a bilinear model for the
64 factor decomposition, it can be easily adjusted to the analysis of more complex multiway data structures
65 and multilinear models, such as three-way and four-way environmental data sets (Tauler, 2021), which can
66 be analyzed using trilinear and quadrilinear MCR-ALS models, as shown in this study. The results of the
67 application of the MCR-ALS method can be used for the discovery of the main driving factors (latent



68 variables) responsible of the observed data variance, in this case, of the observed changes in the measured
69 chemical pollutants.

70 The present study is focused on promoting and extending the use of multivariate curve resolution alternating
71 least squares method, including trilinear and quadrilinear constraints, for the investigation of NO₂, O₃ and
72 PM₁₀ air pollution. In addition, this study aims at providing different strategies to deal and estimate missing
73 data also using the MCR-ALS methodology (Multivariate Curve Resolution of incomplete data multisets |
74 Elsevier Enhanced Reader, 2022). The selected chemometric strategy is ultimately used to evaluate the
75 temporal patterns of the three pollutants during 2018, 2019 and 2020 in eight monitoring stations located
76 in Catalonia (Spain), including three urban, one control site, one semi-urban and three rural. The different
77 stations were specifically selected to evaluate the influence of the geographical location on air pollution.
78 The period of time evaluated (*i.e.*, January 1st to December 31st of 2018, 2019 and 2020) was chosen to
79 cover the COVID-19 lockdown period in Catalonia and to enable a comparison respect to the same time
80 period in the previous two years. Considering that the strictest COVID-19 lockdown in Catalonia occurred
81 in the month of April 2020, a specific evaluation of air quality changes produced during this period of time
82 respect to the previous two years is provided in this study.

83 **2. Materials and methods**

84 **2.1 Air quality data**

85 The experimental data used in this work consisted on O₃, NO₂ and PM₁₀ concentrations recorded from eight
86 air quality monitoring stations operated by the Department of Environment of the Catalan Autonomous
87 Government. The selected air quality monitoring stations consisted in three urban (Gràcia, Vall d'Hebron
88 and Granollers), one semi-urban (Manlleu) and one control site (Observatori Fabra), all of them located in
89 the province of Barcelona, and three rural: Juneda and Bellver de Cerdanya in the province of Lleida, and
90 Begur (Costa Brava, NE Catalonia), in the province of Girona. More detailed information about the
91 characteristics of the stations is provided in a previous air quality monitoring study by the authors
92 (Gorrochategui et al., 2021). NO₂ concentrations were measured by means of chemiluminescence according
93 to the UNE method 77212:1993, using automatically operated MCV 30QL analyzers. Ozone concentrations
94 were measured by means of UV photometry according to ISO FDIS 139464:1998, automatically operated
95 with MCV 48 AUV analyzers. PM₁₀ concentrations were measured by means of gravimetric determination,
96 using manually operated high volume samplers MCV CAV-A/MS. The generated databases with all the
97 concentrations measured were compiled by the Department of Air Monitoring and Control Service of the
98 Generalitat de Catalunya (Xarxa de Vigilància i Previsió de la Contaminació Atmosfèrica (XVPCA).
99 Departament de Territori i Sostenibilitat, 2020).

100 **2.2. Experimental data multisets**

101 In this study, two experimental data multisets were analyzed (see Fig. 1). Both of them contained hourly
102 concentrations of NO₂, O₃ and PM₁₀ measured in the eight air quality monitoring stations but for distinct
103 periods of time. The first data multiset contained air quality data recorded in the month of April 2018, April
104 2019 and April 2020 (*i.e.*, the latter being the time when the strictest lockdown occurred in Catalonia(Real



105 Decreto 463/2020de 14 de marzo, por el que se declara el estado de alarma para la gestión de la situación
106 de crisis sanitaria ocasionada por el COVID-19., 2020)) in the different stations. The second data set
107 contained air quality data recorded in the same 8 stations but during a longer period of time: from January
108 1st to December 31st of 2018, 2019 and 2020. The latter multiset was built in order to evaluate annual trends
109 of air pollution; especially interesting in 2020, an extraordinary year due to the coronavirus pandemic.

110 As observed in Fig. 1, both data sets contained some missing data blocks, which were not included in the
111 MCR-ALS analyses of individual contaminants, a part from some spot values, which were further estimated
112 to undergo chemometric analysis.

113 In the data set of the month of April, no missing data existed for NO₂ and O₃. However, for PM₁₀, data of
114 three months of April were missing (*i.e.*, Begur 2018, Begur 2020 and Observatori Fabra 2018), as observed
115 in Fig. 1a. In the data set of the entire three years (Fig. 1b), for NO₂ and O₃, the months of January and
116 February 2018 in Observatori Fabra station were missing, respectively. For PM₁₀, data from three air quality
117 monitoring stations were missing: Gràcia (September and October, 2018), Begur (months from January to
118 October, 2018, and months from January to July, 2020) and Observatori Fabra (months from January to
119 September, 2018).

120

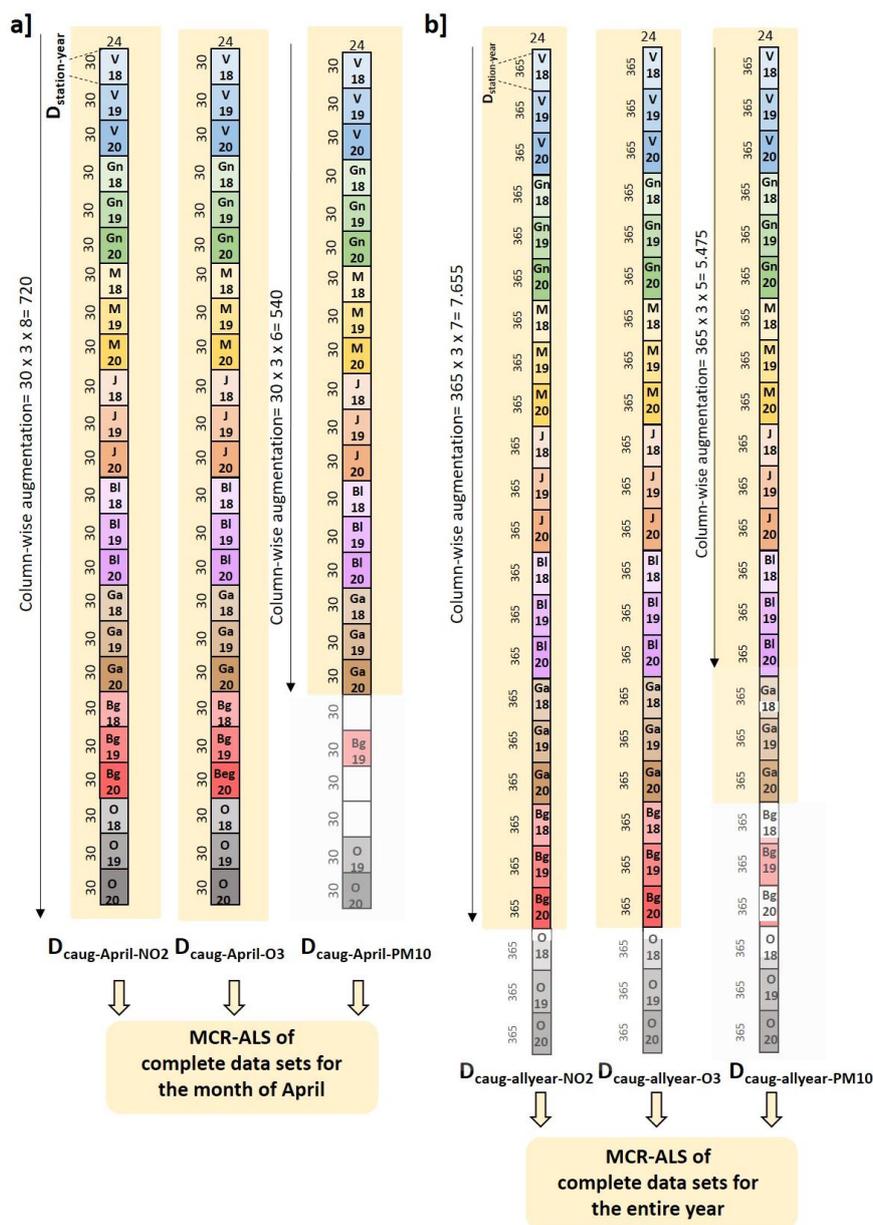
121 2.3. Data sets arrangement

122 In this study, the two data multisets were separately arranged to further undergo MCR-ALS individual
123 analyses of the complete experimental data sets (Fig. 1).

124 To conduct the analysis of the month of April, data matrices for NO₂, O₃ and PM₁₀ were separately arranged
125 in a first step. For each contaminant, a total of 24 data matrices, one per year (three years) and per
126 monitoring station (eight stations), of size 30 x 24 (month days' x hourly measurements), were obtained.

127 As observed in Fig. 1a these 24 data matrices were labeled as **D**_{station-year}; with the name of the corresponding
128 air quality station (V: Vall d'Hebron, Gn: Ganollers, M: Manlleu, J: Juneda, Bl: Bellver, Ga: Gràcia, Bg:
129 Begur and O: Observatori Fabra) and the two last digits of the year (2018, 2019 and 2020). These 24 data
130 matrices were then arranged using a column-wise augmentation, obtaining three augmented data matrices:
131 **D**_{caug-April-NO₂}, **D**_{caug-April-O₃} and **D**_{caug-April-PM₁₀}. The two first augmented matrices (NO₂ and O₃) contained
132 concentration measures of the month of April for each station and each year folded one on top of the other,
133 first performing the augmentation for the 3 years (30 x 3) and then for the eight stations (30 x 3 x 8), as
134 shown in Fig. 1a. The resulting dimensions of these two column-wise augmented data matrices for further
135 MCR-ALS analysis were (720 x 24). However, as previously stated, for PM₁₀, data of three months were
136 missing and thus, the final column-wise augmented matrix was built only with the six stations containing
137 no missing data (30 x 3 x 6), resulting in a (540 x 24) matrix (yellow-shaded area in Fig. 1a).

138 To conduct the analysis of the entire three years, data matrices for NO₂, O₃ and PM₁₀ were also separately
139 arranged in a second step. For each contaminant, a total of 24 data matrices, one per year and per monitoring
140 station, of size 365 x 24 (year days x hourly measurements), were obtained.



141

142 **Figure 1.** Data arrangement for individual analysis of completed data sets. NO₂, O₃ and PM₁₀ concentrations in
 143 a) the month of April and b) the entire year, for one year and one station can be arranged in a data matrix with
 144 the days in the rows and the 24 hour measurements in the columns $D_{\text{station-April}}$ or $D_{\text{station-year}}$. These individual
 145 data matrices are arranged in three (one per each pollutant) column-wise augmented data matrices. For the
 146 data set of April in 2018-2019-2020 (a), the augmented matrices are $D_{\text{caug-April-NO}_2}$, $D_{\text{caug-April-O}_3}$ (720,24) and $D_{\text{caug-}}$
 147 $D_{\text{caug-April-PM}_{10}}$ (540,24). For the data set of the entire year in 2018-2019-2020 (b), the augmented matrices are $D_{\text{caug-}}$
 148 $D_{\text{caug-allyear-NO}_2}$, $D_{\text{caug-allyear-O}_3}$ (7,655,24) and $D_{\text{caug-allyear-PM}_{10}}$ (5,475,24). Stations containing missing data (white-gaps in
 149 the figure) are excluded for further MCR-ALS analysis (yellow-shaded area). VH: Vall d’Hebron, Grn:
 150 Granollers, Mn: Manlleu, Jun: Juneda, Bell: Bellver, Gra: Gràcia, Beg: Begur, OF: Observatori Fabra.



151 These 24 data matrices were then arranged using a column-wise arrangement, obtaining three augmented
152 data matrices: $\mathbf{D}_{\text{caug-allyear-NO}_2}$, $\mathbf{D}_{\text{caug-allyear-O}_3}$ and $\mathbf{D}_{\text{caug-allyear-PM}_{10}}$ (Fig. 1b). In this case, for the three
153 contaminants, some data were missing (white gaps in the figure). For NO_2 and O_3 , some data from
154 Observatori Fabra station was missing. Thus, the resulting augmented matrices, $\mathbf{D}_{\text{caug-allyear-NO}_2}$, $\mathbf{D}_{\text{caug-allyear-}}$
155 o_3 , contained information of the whole year, seven stations and the three years ($365 \times 3 \times 7$), resulting in
156 (7655×24) matrices, as shown in the figure. For PM_{10} , data of three stations were missing (white gaps in
157 the figure) corresponding to Gràcia, Begur and Observatori Fabra. Thus, in order to perform the MCR-ALS
158 analysis, the resulting PM_{10} column-wise augmented matrix only contained information of five stations
159 with no missing data ($365 \times 3 \times 5$), resulting in a (5475×24) matrix (yellow-shaded area in Fig. 1b).
160 Data arrangement for the simultaneous study of the three pollutants considering the whole incomplete
161 multiblock experimental data sets is further described in Sect. 2.7.

162

163 2.4. Estimation of missing data

164 Estimation of missing data was used for the case when failures of stations and/or malfunction of them
165 caused the absence of measurements for few hours or few days. In order to estimate such missing data, the
166 nearest-neighbor method (Peterson, 2009) (*i.e.*, knn imputation) was used. In this study, the function
167 *mdcheck* (*i.e.*, missing data checker and infiller) of PLS Toolbox version 8.9.1 (Eigenvector Inc., WA) was
168 utilized to perform the imputation. This function checks for missing data and infills them using a PCA
169 model imputation from distinct algorithms. In our case, three algorithms were tested consisting on 'svd'
170 (Singular Value Decomposition), 'NIPALS' (Nonlinear Iterative Partial Least Squares) and 'knn', the latter
171 providing the better estimation results in our case, and thus, the one that was finally used in this study.
172 It is important to mention that estimation of missing data was not performed in cases where the entire month
173 was missing. For those cases, the station was not included in the MCR-ALS analysis of the complete data
174 set. For the analysis of incomplete multiblock data sets, an especial arrangement was performed using a
175 particular data fusion strategy, as further explained in Sect. 2.7.

176

177 2.5. MCR-ALS analysis of the experimental data

178 Different chemometric methods have been proposed in the literature for the analysis of environmental
179 monitoring data. MCR-ALS is a frequently used method in spectrochemical mixture data analysis, which
180 can also be easily extended to the analysis of environmental source apportionment data sets (Alier et al.,
181 2011). MCR-ALS is a flexible soft-modelling factor analysis tool which allows for the application of natural
182 constraints (see below) and it can be easily adapted to the analysis of complex multiway (multimode) data
183 structures, such as three- and four-way environmental data sets using trilinear and quadrilinear model
184 constraints (De Juan et al., 1998; Smilde et al., 2004; Malik and Tauler, 2013).

185 The simplest application of the MCR-ALS method is based on a bilinear model that performs the factor
186 decomposition of a two-way data set (*i.e.* a data table or a data matrix). Eq. (1) summarizes this bilinear
187 model in its element-wise way, while Eq. (2) presents the same model in a matrix linear algebra format:

$$188 \quad d_{ij} = \sum_{n=1}^N x_{i,n} y_{j,n} + e_{i,j} \quad i=1,..I \text{ (days)}, j=1,..J \text{ (hours)} \quad (1)$$



189
$$\mathbf{D} = \mathbf{X} \mathbf{Y}^T + \mathbf{E} \tag{2}$$

190 In Eq. (1), the individual data values, $\mathbf{d}_{i,j}$ elements (in this case the O_3 , NO_2 or PM_{10} concentration values
 191 measured one day (i) at a particular hour (j)) are decomposed as the sum of a reduced number of
 192 contributions (components), $n=1,..N$, each one of them defined by the product of two factors, $\mathbf{x}_{i,n}$ (scores)
 193 and $\mathbf{y}_{j,n}$ (loadings). In addition, the term $\mathbf{e}_{i,j}$ is the residual part of $\mathbf{d}_{i,j}$, which cannot be explained by these N
 194 components and accounts as experimental noise and uncertainties. In Eq. (2), the data matrix, \mathbf{D} , of
 195 dimensions $I \times J$ is decomposed into the scores factor matrix \mathbf{X} ($I \times N$) and the loadings factor matrix, \mathbf{Y}^T
 196 ($N \times J$). The number of components, N, is selected to explain as much as possible the data variance, while
 197 the unexplained small contributions of data variance and experimental noise are in \mathbf{E} . Multivariate Curve
 198 Resolution (Tauler, 1995) performs the bilinear model factor decomposition shown in Eq. (1) and (2) using
 199 an alternating least squares (MCR-ALS) algorithm under a set of constraints which reduce the extent of the
 200 bilinear model rotation ambiguities (Abdollahi and Tauler, 2011) and allow the physical identification and
 201 interpretation of the factor matrices \mathbf{X} and \mathbf{Y}^T , as for example, the application of non-negativity constraints
 202 to the elements of the factor matrices \mathbf{X} and \mathbf{Y}^T (Bro and De Jong, 1997; De Juan and Tauler, 2003). Models
 203 with different number of components can be tested and a final decision is taken considering the data fit and
 204 the shapes and reliability of the resolved profiles. The ALS algorithm also needs initial estimates of either
 205 \mathbf{X} or \mathbf{Y}^T factor matrices. These initial estimates can be obtained from the more ‘dissimilar’ rows or columns
 206 of the original data matrix (Jackson et al., 2002). Eq. (2) for \mathbf{D} is solved iteratively, which updates the
 207 solutions (vector profiles in \mathbf{X} and \mathbf{Y}^T matrices) until they fit the data optimally and fulfill the proposed
 208 constraints

209 In this work the MCR-ALS method has been applied either to the individual data matrices $\mathbf{D}_{\text{station,year}}$
 210 described in previous section and in Fig. 1 for every pollutant (O_3 , NO_2 or PM_{10}), at one period of time
 211 (April or full year) and at one monitoring station or to the augmented matrices of the same three pollutants
 212 for the three years ($k=1,..3$) simultaneously and for the different stations ($l=1,..,8$) concatenated vertically
 213 in $\mathbf{D}_{\text{caug-April}}$ or $\mathbf{D}_{\text{caug-allyear}}$ (see Fig. 1). In the case of the individual data matrices described above for the
 214 period of time (April or full year), $\mathbf{D}_{\text{station,april}}$ or $\mathbf{D}_{\text{station,year}}$, the factor (scores) matrix \mathbf{X} will have
 215 respectively the April or full year day profiles of the components and the factor (loadings) matrix \mathbf{Y}^T will
 216 have the corresponding hour profiles of these components. In the case of the column-wise augmented data
 217 matrices $\mathbf{D}_{\text{caug-April}}$ or $\mathbf{D}_{\text{caug-allyear}}$, bilinear model Eq. (2) was extended as:

218
$$\mathbf{D}_{\text{caug}} = \mathbf{X}_{\text{caug}} \mathbf{Y}^T + \mathbf{E}_{\text{caug}} \tag{3}$$

219 Where \mathbf{X}_{caug} is now the augmented factor (scores) matrix with the augmented day profiles concatenated
 220 vertically for the different years and stations, and \mathbf{Y}^T is the matrix of the hour profiles again, which are
 221 common for all the concatenated matrices in \mathbf{X}_{caug} . During the ALS optimization of the bilinear model in
 222 Eq. (3), constraints can be also applied and the same aspects in relation to number of components and
 223 convergence as for solving Eq. (2) are considered.

224 **2.6 MCR-ALS analysis of the complete experimental data sets using trilinear and quadrilinear**
 225 **constraints**

226 Solving Eq. (3) using bilinear MCR-ALS does not take into account the temporal and spatial structure of



227 the data in the vertical concatenated mode which includes the information of the day, year and station. This
 228 data structure can be considered in the trilinear and specially in the quadrilinear extensions of the bilinear
 229 models described in Eq. (1-3).

230 The factor decomposition model given before can be extended to a three-way dataset, $\underline{\mathbf{D}}$, or to a four-way
 231 data set, $\underline{\mathbf{D}}$, expressed individually for every data value as given by Eq. (4) and (5).

$$232 \quad d_{i,j,kl} = \sum_{n=1}^N x_{i,n} y_{j,n} z_{kl,n} + e_{i,j,kl} \quad (\text{trilinear model}) \quad (4)$$

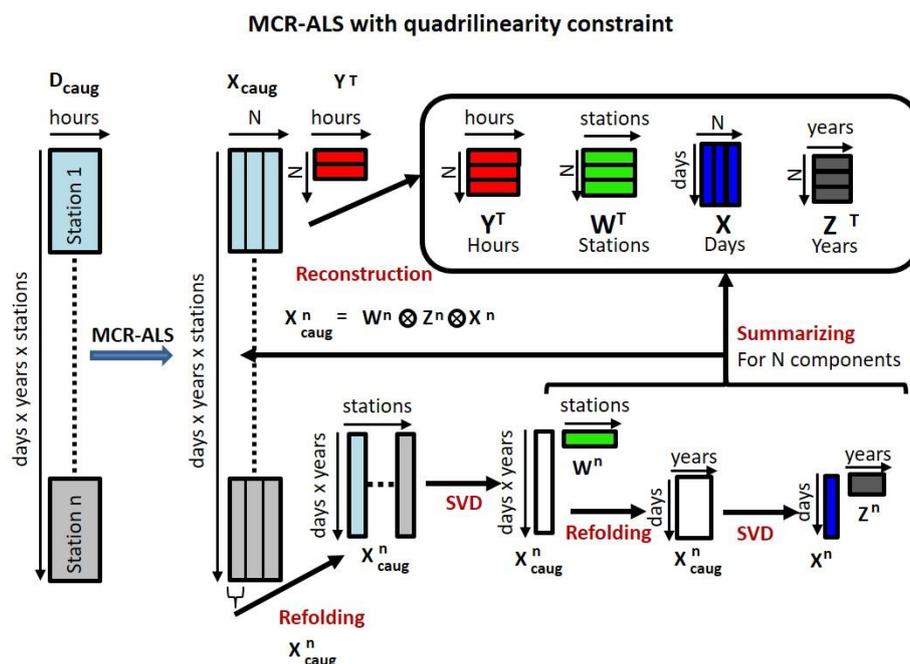
$$233 \quad d_{i,j,k,l} = \sum_{n=1}^N x_{i,n} y_{j,n} z_{k,n} w_l + e_{i,j,k,l} \quad (\text{quadrilinear model}) \quad (5)$$

234 where $d_{i,j,kl}$ are the individual data values (concentrations of O_3 , NO_2 or PM_{10}) in the four experimental data
 235 modes: the day of April or of the full year $i=1, \dots, 30$ or $i=1, \dots, 365$, the hour of the day $j=1, \dots, 24$, and the
 236 year-station $kl=1, \dots, 24$ in the case of the trilinear model, and $d_{i,j,k,l}$ has the year-station third mode separated
 237 in year $k=1, 2, 3$, and station, $l=1, \dots, 8$ in the quadrilinear model. These data values are modeled as the sum
 238 of a number of components (contributions), $n=1, \dots, N$, defined by the product of three factors $x_{i,n}$, $y_{j,n}$, and
 239 $z_{kl,n}$, in the case of the trilinear model and in four factors in the case of the quadrilinear model, $x_{i,n}$, $y_{j,n}$, $z_{k,n}$
 240 and w_l . These factors are related with the three and four data modes respectively (day, hour and year-
 241 station or day, hour, year and station). $e_{i,j,kl}$ and $e_{i,j,k,l}$ are the part of $d_{i,j,kl}$ and $d_{i,j,k,l}$ not explained by the
 242 contribution of these N components. These trilinear and quadrilinear models can be written in a matrix form
 243 using the decomposition of every individual \mathbf{D}_{kl} data slice (every individual matrix \mathbf{D}_{kl}), as shown in Eq.
 244 (6) and (7).

$$245 \quad \mathbf{D}_{kl} = \mathbf{X} \mathbf{Z}_{kl} \mathbf{Y}^T + \mathbf{E}_{kl} \quad (\text{trilinear model}) \quad (6)$$

$$246 \quad \mathbf{D}_{kl} = \mathbf{X} \mathbf{Z}_k \mathbf{W}_l \mathbf{Y}^T + \mathbf{E}_{kl} \quad (\text{quadrilinear model}) \quad (7)$$

247 Under the trilinear model, all individual data matrices, $\mathbf{D}_{kl,(I,J)}$ are simultaneously decomposed with the
 248 same number of components N and the same daily, \mathbf{X} (I,N) and hourly \mathbf{Y}^T (N,J) profiles. Thus, they differ
 249 only in a diagonal matrix \mathbf{Z}_{kl} , (N,N) different for every one of the $kl=1, \dots, 24$ year-stations (year-station
 250 profiles), which gives the relative amounts of the N components in every data matrix (year-station), \mathbf{D}_{kl} .
 251 These N diagonal elements of the \mathbf{Z}_{kl} can also be grouped in the third factor matrix \mathbf{Z} ($K \times L, N$). Under the
 252 quadrilinear model, all individual data matrices, $\mathbf{D}_{kl,(I,J)}$ are simultaneously decomposed with the same
 253 number of components N and the same daily, \mathbf{X} (I,N) and hourly \mathbf{Y}^T (N,J) profiles. Thus, they differ in the
 254 diagonal matrices \mathbf{Z}_k , (N,N) and \mathbf{W}_l (N,N), which are different for every year (k) and station (l), which give
 255 the relative amounts of the N components in every data matrix \mathbf{D}_{kl} respectively. These N diagonal elements
 256 of the \mathbf{Z}_k and \mathbf{W}_l matrices can also be grouped in the third and four factor matrices \mathbf{Z} (K,N) and \mathbf{W} (L,N).
 257 Therefore, the proposed trilinear and quadrilinear models take advantage of the natural structure of the
 258 analyzed data sets, especially in relation to their different temporal modes (*i.e.* hourly, daily, yearly) and to
 259 the different type of monitoring stations analyzed simultaneously. The implementation of trilinear and
 260 quadrilinear models as a constraint in the MCR-ALS method has been described in previous works (Tauler,
 261 2021; Malik and Tauler, 2013; Alier et al., 2011) Here only a brief explanation of the case of the
 262 implementation of the quadrilinear model constraint for the case of study is shown.



263

264 **Figure 2. MCR-ALS with the quadrilinearity constraint. Graphical description of the implementation of the**
 265 **quadrilinearity constraint during the Alternating Least Squares optimization. See Eq. (4-7) and their**
 266 **explanation in the manuscript.**

267 Figure 2 shows the practical implementation of the quadrilinear constraint in the MCR-ALS analysis of the
 268 four-way data set obtained in the two type of data, when the April data of the three parameters (O_3 , NO_2
 269 and PM_{10}) were studied over the three years (2018-2019 and 2020) and over the different monitoring
 270 stations described above, and also for the analogous four-way data set when instead of April data, the full
 271 year data were considered for the same years and stations.

272 The individual data sets with the concentrations of the three parameters (one per year and station), were
 273 arranged in the column-wise augmented data matrix D_{caug} of dimensions 30 (April) or 365 days x 3 years
 274 x 8 stations, giving a total number of 720 rows for April data or of 8760 rows for the full year data, and 24
 275 hourly measures in columns. These number of row elements is for the case of no-missing data, however
 276 they will be lowered for the cases of missing data, especially in the case of the full year data (see previous
 277 section in missing data). The application of the quadrilinearity constraint implies that the augmented
 278 profiles of every component n , x^n_{aug} , having the vertically concatenated information of days times years
 279 times stations is first refolded in the data matrix X^n_{aug} of dimensions 30 (April) or 365 (all year) times 3
 280 (years) rows by 8 (stations). This augmented factor matrix is decomposed by SVD considering only the
 281 first singular component into the product of two vector profiles, one long vector profile x^n_{caug} (90 or 1095
 282 x 1) of the combined day-year profile by a vector profile w^n (8x1) describing differences among the
 283 different stations for the component n . x^n_{caug} long vector day-year profile can be further refolded in a matrix
 284 and decomposed by SVD into the product of two new vector profiles, one related with the year profile, z^n ,
 285 and another with the day profile x^n , for the considered component n . In this way, for every component



286 (contribution), the concentration of any one of the three parameters (O_3 , NO_2 and PM_{10}) is decomposed in
 287 the product of four profiles, one related with the day (of April or of the whole year), \mathbf{x}^n , another related with
 288 the hour of the day, \mathbf{y}^n , another with the considered year, \mathbf{z}^n , and another with the monitoring station, \mathbf{w}^n .
 289 This factor decomposition allows a detangled interpretation of the temporal and spatial sources of variation
 290 of the observed concentrations of the three pollutants. Therefore, the application of this quadrilinearity
 291 constraint implies that for every component, the daily changes are described by the same single \mathbf{x}^n vector
 292 profile which changes over the years and station by station by the corresponding scalars values in \mathbf{z}^n and
 293 \mathbf{w}^n . Once the three profiles in the three augmented modes, \mathbf{x}_n , \mathbf{z}_n and \mathbf{w}_n , are obtained, they can be multiplied
 294 using the kronecker product (Soloveyichik and Trushin, 2016) to reconstruct the long vector profile, $\mathbf{x}^{n, aug}$,
 295 (see Fig. 2) and rebuild the bilinear model in the next iteration of the general ALS optimization. Finally,
 296 the vector profiles for every component n in the different modes, can be grouped in the corresponding factor
 297 matrices \mathbf{X} , \mathbf{Z} and \mathbf{W} , which together \mathbf{Y}^T give the full quadrilinear decomposition of the four-way data set,
 298 \mathbf{D} . See previous works for a more detailed description of the algorithm used for the practical implementation
 299 of the quadrilinearity constraint in MCR-ALS (De Juan and Tauler, n.d.; De Juan et al., 1998).

300 2.7 MCR-ALS simultaneous analysis of incomplete multiblock experimental data

301 The simultaneous analysis of the NO_2 , O_3 and PM_{10} experimental data can be done one step forward using
 302 a data fusion multiblock strategy. This would imply building a single MCR model for the whole multiset
 303 data obtained for the 3 pollutants, NO_2 , O_3 and PM_{10} in April or in the whole year, for the three years, 2018,
 304 2019 and 2020, and for the different monitoring stations. This is expressed in the following data matrix
 305 equation (see also Fig. 3):

$$306 \quad \mathbf{D}_{eraug} = [\mathbf{D}_{caugNO_2}, \mathbf{D}_{caugO_3}, \mathbf{D}_{caugPM_{10}}] = \mathbf{X}_{caug} [\mathbf{Y}_{NO_2}^T, \mathbf{Y}_{O_3}^T, \mathbf{Y}_{PM}^T] = \mathbf{X}_{caug} \mathbf{Y}_{raug}^T \quad (8)$$

307 In this Equation the column-wise augmented data matrices, \mathbf{D}_{caugNO_2} , \mathbf{D}_{caugO_3} , and $\mathbf{D}_{caugPM_{10}}$ previously
 308 described (and analyzed separately by MCR-ALS with different factor decomposition models, bilinear,
 309 trilinear and quadrilinear), are now concatenated horizontally giving the new single row and column-wise
 310 super-augmented data matrix \mathbf{D}_{eraug} which is decomposed in the two new augmented factor matrices, \mathbf{X}_{caug}
 311 and \mathbf{Y}_{raug}^T using the MCR bilinear model and constraints, like it was described in Sect. 2.4. In the
 312 augmented rows of the new \mathbf{Y}_{raug}^T will be the resolved hour profiles for the three contaminants $\mathbf{Y}_{NO_2}^T$, $\mathbf{Y}_{O_3}^T$
 313 and \mathbf{Y}_{PM}^T . In addition, if the trilinearity/quadrilinearity constraints are applied to the columns of the
 314 resolved factor matrix \mathbf{X}_{caug} as described above in Sect. 2.6 using matrix decompositions of Eq. (6) and (7),
 315 the common day, year and station profiles will be separately recovered and analyzed.

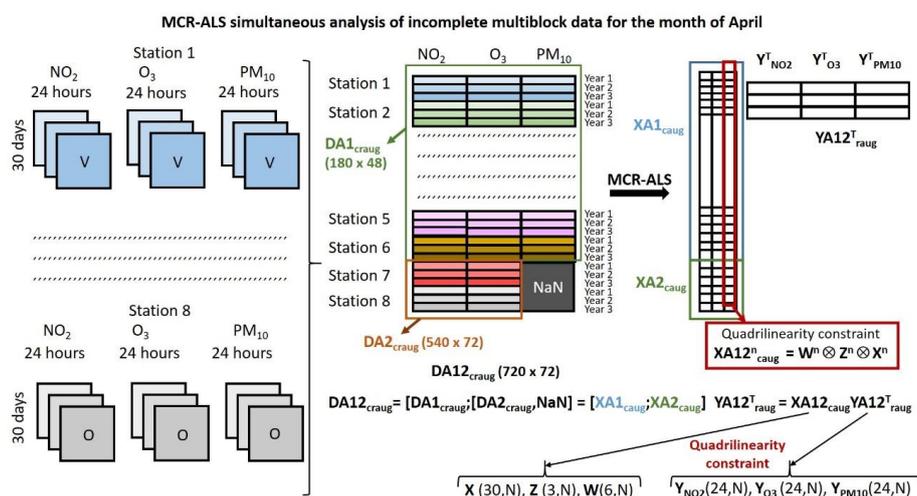
316 However as previously described, April and the whole year individual data sets were not obtained for all
 317 stations, years and pollutants, and therefore they together could not be fitted in a rectangular super-
 318 augmented data matrix containing all the data for all the years and stations as shown in Eq. (8) for \mathbf{D}_{eraug} .
 319 Some of the individual data sets were missing (see Sect. 2.3 and 2.4). In particular, in the case of April, two
 320 different data blocks could be arranged. First, the NO_2 , O_3 and PM_{10} concentrations data for 3 years and 6
 321 stations were arranged in the complete row- and column-wise augmented April data block, \mathbf{DAI}_{eraug} , with
 322 540 rows (30 days x 3 years x 6 stations) and 72 columns (24 hours for NO_2 + 24 hours for O_3 + 24 hours
 323 for PM_{10}). Secondly, the additional NO_2 and O_3 concentration data for 3 years and 2 stations were arranged



324 in complete row- and column-wise augmented April data block $\mathbf{DA2}_{\text{craug}}$ with 180 rows (30 days x 3 years
 325 x 2 stations) and 48 columns (24 hours for NO_2 + 24 hours for O_3). These two April data blocks can be
 326 analyzed independently, but a new dataset can be built concatenating the two data blocks as shown in Fig.
 327 S1, which will be reformulated and analyzed as shown in next Equation.

328 $\mathbf{DA12}_{\text{craug}} = [\mathbf{DA1}_{\text{craug}}; [\mathbf{DA2}_{\text{craug}}, \text{NaN}(180,24)]] = \mathbf{XA12}_{\text{caug}} \mathbf{YA12}_{\text{raug}} =$

329 $= \mathbf{XA12}_{\text{caug}} [\mathbf{Y}_{\text{NO}_2}, \mathbf{Y}_{\text{O}_3}, \mathbf{Y}_{\text{PM10}}] \quad (9)$



330

331 **Figure 3. MCR-ALS with the quadrilinearity constraint for the simultaneous analysis of the three contaminants in the**
 332 **incomplete multiblock data set of the month of April. See Eq. (9) and their explanation in the manuscript.**

333 This new incomplete data set $\mathbf{DA12}_{\text{craug}}$ is built using the two data blocks previously defined, $\mathbf{DA1}_{\text{craug}}$ and
 334 $\mathbf{DA2}_{\text{craug}}$, both concatenated vertically and filling the empty data block corresponding to the unknown
 335 concentrations of PM_{10} for two missing stations with the NaN notation. The application of MCR-ALS to
 336 this incomplete dataset is decomposed using a bilinear model (see in Fig. 3), giving the two factor matrices
 337 $\mathbf{XA12}_{\text{caug}}$ and $\mathbf{YA12}_{\text{raug}}$. $\mathbf{XA12}_{\text{caug}}$ factor matrix has the column-wise augmented day x year x station
 338 profiles in its columns and $\mathbf{YA12}_{\text{raug}}$ factor matrix has the row-wise augmented hour profiles for NO_2
 339 (\mathbf{Y}_{NO_2}), O_3 (\mathbf{Y}_{O_3}) and PM_{10} (\mathbf{Y}_{PM10}) in its rows. As previously, the trilinear/quadrilinear constraints can be
 340 applied during the ALS factor decomposition to the $\mathbf{XA12}_{\text{caug}}$ factor matrix and allow the separate recovery
 341 of the day, year and station profiles, a part of the hour profiles for NO_2 , O_3 and PM_{10} obtained in $\mathbf{YA12}_{\text{raug}}$.
 342 Analogous equations can be described for the NO_2 , O_3 and PM_{10} experimental data measured not only in
 343 April but during the whole year. In this case however the two data blocks, $\mathbf{DY1}_{\text{craug}}$ and $\mathbf{DY2}_{\text{craug}}$, will have
 344 different sizes than for the only April month data because they are for all the days of the whole year.
 345 Different data sets were missing in this case. $\mathbf{DY1}_{\text{craug}}$ has the data for 5 stations with 5475 rows (365 days
 346 x 3 years x 5 stations) and 72 columns (24 hours for NO_2 + 24 hours for O_3 + 24 hours for PM), and $\mathbf{DY2}_{\text{craug}}$
 347 has the additional data for 2 stations but only for NO_2 and O_3 concentrations, with 2190 rows (365 days x 3
 348 years x 2 stations) and 48 columns (24 hours for NO_2 + 24 hours for O_3) (see Fig. S2). For the whole year



349 data, the bilinear factor decomposition can be described by the new Eq. (10).

$$\begin{aligned}
 350 \quad \mathbf{DY12}_{\text{craug}} &= [\mathbf{DY1}_{\text{craug}}; [\mathbf{DY2}_{\text{craug}}, \mathbf{NaN}(2190,24)]] = \mathbf{XY12}_{\text{caug}} \mathbf{YY12}_{\text{raug}} = \\
 351 \quad &= \mathbf{XY12}_{\text{caug}} [\mathbf{Y}_{\text{NO}_2}, \mathbf{Y}_{\text{O}_3}, \mathbf{Y}_{\text{PM10}}] \quad (11)
 \end{aligned}$$

352 Where now $\mathbf{DY12}_{\text{craug}}$ is the new incomplete data set built with the two data blocks $\mathbf{DY1}_{\text{craug}}$ and $\mathbf{DY2}_{\text{craug}}$
 353 concatenated vertically and $\mathbf{NaN}(2190, 24)$ is for the missing PM_{10} concentrations during 3 years in the
 354 missing 2 stations (see Supplement Fig. S2). $\mathbf{XY12}_{\text{caug}}$ and $\mathbf{YY12}_{\text{raug}}$ are now the two factor matrices
 355 obtained in the bilinear decomposition of $\mathbf{DY12}_{\text{craug}}$. The first factor matrix $\mathbf{XY12}_{\text{caug}}$ have the column-wise
 356 augmented day x year x station profiles in its columns and the second factor matrix $\mathbf{YY12}_{\text{raug}}$ has the hour
 357 profiles for NO_2 , O_3 and PM_{10} in its rows. And also as previously, the trilinear/quadrilinear constraints
 358 applied during the ALS factor decomposition to the $\mathbf{XY12}_{\text{caug}}$ factor matrix will allow the separate recovery
 359 of the day, year and station profiles, apart from the hour profiles for NO_2 (\mathbf{Y}_{NO_2}), O_3 , (\mathbf{Y}_{O_3}), and PM_{10}
 360 (\mathbf{Y}_{PM10}) obtained in $\mathbf{YY12}_{\text{raug}}$. The difference with the results of April data is that now the column-wise
 361 augmented profiles in $\mathbf{XY12}_{\text{caug}}$ will have information about the 365 days of the whole year and not only
 362 for the 30 days of April. Supplement Fig. S3 is given to illustrate graphically the bilinear model applied to
 363 the incomplete two-blocks data set.

364

365 2.8 Evaluation of MCR-ALS results

366 The final evaluation of the MCR-ALS fitting results is performed calculating the explained data variances
 367 (R^2) using Eq. (11),

$$368 \quad R^2 = 100 \times \left(1 - \frac{\sum_{i=1}^m \sum_{j=1}^n (d_{ij} - \hat{d}_{ij})^2}{\sum_{i=1}^m \sum_{j=1}^n d_{ij}^2} \right) \quad (11)$$

369 where d_{ij} are the experimental predicted O_3 , NO_2 or PM_{10} concentrations, and \hat{d}_{ij} are the corresponding
 370 calculated values by MCR-ALS using either the bilinear (Eq. (1-3)), trilinear (Eq. (4) and (6)) or
 371 quadrilinear (Eq. (5) and (7)) models. Apart from the global fitting with the full model (all components),
 372 the explained variances can be also calculated individually for every MCR-ALS component, where now
 373 the calculated values, \hat{d}_{ij} , take only into account one of the n components of the model. In this way, the
 374 relative importance of the different contributions can be evaluated, as well as their overlapping degree with
 375 the other contributions or components.

376 2.8 Software

377 MATLAB 9.10.0 R2021a (The MathWorks, Inc., Natick, MA, USA) was used as the development platform
 378 for data analysis and visualization. The new graphical interface MCR-ALS GUI 2.0 (Malik and Tauler,
 379 2013), freely available as a toolbox at the web address <http://www.mcrals.info/>, was used for bilinear and
 380 trilinear data sets. Statistics Toolbox™ for MATLAB and PLS Toolbox 8.9.1 (Eigenvector Research Inc.,
 381 Wenatchee, WA, USA) were also used in this work. New specific MCR-ALS command line code for
 382 incomplete multiblock data sets is under final development and it can be requested to one of the authors
 383 (RT, email:roma.tauler@idaea.csic.es).



384 **3. Results and discussion**

385 Results of MCR-ALS will be shown separately for the analysis of the month of April and for the analysis
 386 of the entire years. In the study of the month of April, the individual analysis of the three contaminants per
 387 separate is firstly performed, using only data from stations with no missing blocks (*i.e.*, data matrices $D_{\text{caug-}}$
 388 April-NO_2 , $D_{\text{caug-April-O}_3}$ and $D_{\text{caug-April-PM}_{10}}$, yellow-shaded area of Fig. 1a). Then, a simultaneous analysis of
 389 the three contaminants containing incomplete data is performed (*i.e.*, data matrix $D_{A12\text{craug}}$, Fig. S1). In the
 390 study of the entire years, again the individual analysis of the three contaminants per separate is initially
 391 performed, using only data from stations with no missing blocks (*i.e.*, data matrices $D_{\text{caug-allyear-NO}_2}$, $D_{\text{caug-}}$
 392 allyear-O_3 and $D_{\text{caug-allyear-PM}_{10}}$, yellow-shaded area of Fig. 1b). Then, a simultaneous analysis of the three
 393 contaminants containing incomplete data is performed (*i.e.*, data matrix $D_{Y12\text{craug}}$, Fig. S3). In all cases the
 394 selection of the number of components and the initial estimates for MCR-ALS were performed as described
 395 in Sect. 2.5. A summary of the explained variances of the MCR-ALS analyses for the different data sets
 396 with non-negativity and either bilinear, tri-linear or quadrilinear modeling and with different number of
 397 components is given in **Table 1**.

398 **Table 1.** MCR-ALS decomposition and explained variances for the different models.

	Explained variances: April 2018-2019-2020.		
	MCR-ALS bilineal	MCR-ALS quadrilinear	MCR-ALS trilinear
$D_{\text{caug-April-NO}_2}^d$	All 94.4%	All 78.4%	All 79.2%
(4 comp)	Sum 125.2%	Sum 109.8%	Sum 113.5%
$D_{\text{caug-April-O}_3}^d$	All 98.4%	All 92.9%	All 93.5%
(3 comp)	Sum 143.5%	Sum 118.3%	Sum 126.4%
$D_{\text{caug-April-PM}_{10}}^d$	All 91.8%	All 78.4%	All 79.0%
(3 comp)	Sum 126.5%	Sum = 112.9	Sum= 113.0%
$D_{A12\text{craug}}^d$	All 96.2%	All 90.7%	All 91.2%
(5 comp)	Sum 135.5%	Sum 111.3%	Sum 114.4
Explained variances: All year 2018-2019-2020.			
$D_{\text{caug-allyear-NO}_2}^d$	All 95.1.0%	All 80.3%	All 80.5%
(4 comp)	Sum 131.0%	Sum 116.2%	Sum 115.7%
$D_{\text{caug-allyear-O}_3}^d$	All 97.5%	All 90.1%	All 90.6%
(3 comp)	Sum 132.2%	Sum 130.2%	Sum 129.2%
$D_{\text{caug-allyear-PM}_3}^d$	All 88.1%	All 72.4%	All 72.6%
(3 comp)	Sum 116.6%	Sum 105.0%	Sum 103.2%
$D_{Y12\text{craug}}^d$	All 94.7%	All 86.4%	All 86.8%
(5 comp)	Sum 126.8%	Sum 125.8%	Sum 126.6%

399 ^a MCR-ALS for raw data with non-negativity constraint

400 ^b MCR-ALS for raw data with non-negativity and quadrilinear constraint

401 ^c MCR-ALS for raw data with non-negativity and trilinear constraint

402 ^d Augmented data matrices and number of components (see Fig. 1, Eq. (3), (9) and (10), and explanation in section Data sets arrangement)

403

404 MCR-ALS bilinear analysis of April data in the $D_{\text{caug-April-NO}_2}$, $D_{\text{caug-April-O}_3}$ and $D_{\text{caug-April-PM}_{10}}$ data matrices
 405 with non-negativity constraints explained respectively 94.40%, 98.4% and 91.8% of the total variance when
 406 four, three and three components were considered (Table 1). These values indicate the higher complexity
 407 of the NO_2 data compared to O_3 data as will be shown also below. When the quadrilinear constraint was
 408 applied these values decreased to 78.4%, 92.9% and 78.4% respectively, confirming again the less complex
 409 and more regular changes of ozone concentrations in the three years at the different monitoring stations.
 410 Variances explained by the individual components are given in the Figures shown below. The amount of



411 variance overlap (also given in Table 1) in every case can be obtained subtracting the sum of the individual
412 variances with the variance obtained with all the components simultaneously. This difference is again larger
413 in the case of NO₂. In Table 1 also, the variances obtained when the trilinearity constraint was applied,
414 instead of the quadrilinearity constrain, are also given, with similar results to those obtained by both
415 multilinear models. In the case of MCR-ALS of all-year data of the $D_{\text{caug-allyear-NO}_2}$, $D_{\text{caug-allyear-O}_3}$ and $D_{\text{caug-}}$
416 allyear-PM_{10} data matrices, rather similar results to those from April were obtained in terms of explained
417 variances for all three type of models (see Table 1), reflecting again the higher complexity of the NO₂ data
418 over the years and stations compared to O₃ data, and the intermediate behavior of PM₁₀ data, although the
419 later more similar to the NO₂ data.

420 Possible correlations between NO₂, O₃ and PM₁₀ data sets during the month of April of 2018, 2019 and
421 2020 and in the eight stations were investigated using the incomplete data arrangement described in Sect.
422 2.7 and Fig. S1. MCR-ALS analysis of $DA12_{\text{eraug}}$ with five components and with only negativity constraints
423 gave a total explained variance of 96.2% (Table 1). When the quadrilinearity constraint was applied, the
424 total explained variance decreased down to 90.7% (91.2% for the MCR-ALS trilinear). Such decrease
425 between bilinear and quadrilinear MCR-ALS models of only 5 % indicated a good quadrilinear behavior
426 of the whole system in April. The explained variances of each component individually are given below
427 with the corresponding Figures of the resolved profiles. In the case of the simultaneous study of NO₂, O₃
428 and PM₁₀ profiles along all the three years (*i.e.*, 2018, 2019 and 2020) in the seven stations using the
429 incomplete data arrangement described in Sect. 2.7 and Fig. S3. Results using five components indicated
430 also a rather good quadrilinear behavior of the system. A more detailed description of the profiles describing
431 the concentration changes of the three pollutants and of the behavior of the whole systems formed by all of
432 them in the different stations and during the three years, separately for April and for the all year, is given
433 below.

434

435 3.1. Study of the month of April

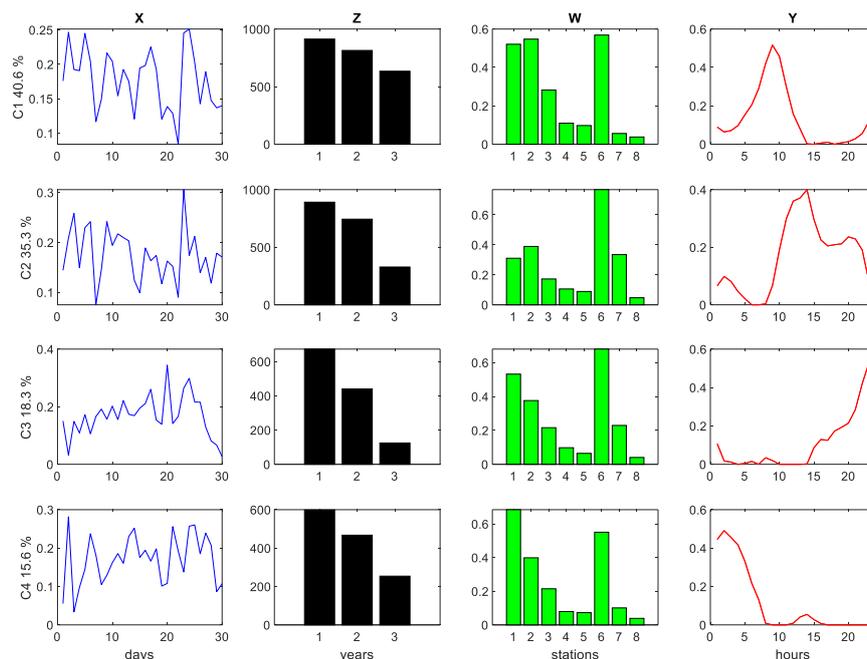
436 3.1.1. NO₂ study ($D_{\text{caug-April-NO}_2}$ data matrix)

437 In Fig. 4, from left to right, the profiles of the different modes of the four components are shown: x-day
438 (blue), z-year (black), w-station (green) and y-hours (red). Component profiles in the four modes obtained
439 by MCR-ALS when using non-negativity and quadrilinearity constraints are shown in Fig. 4.

440 NO₂ hour profile of first component (C1) showed a narrow maximum between 9:00-11:00 hours, coincident
441 with the rush traffic hour and due to fuel combustion by vehicles. In second component (C2) this hour
442 profile presented a much wider peak during daily hours (10:00-20:00h), again potentially attributed to the
443 combined effects of traffic emissions and ozone formation (see below). Third component (C3) reached an
444 hourly maximum in the late evening, approximately at 22:00 hours whereas forth component (C4) showed
445 a maximum between 00:00 and 05:00 hours, describing the NO₂ night behavior. As observed in the year
446 profiles (z-mode), for all the components, NO₂ contributions showed a significant decrease in 2019 and
447 even higher in 2020 respect to 2018; the latter possibly attributed to the COVID-19 curfew and mobility
448 restrictions. Moreover, as observed in the station profiles (w-mode), such depletion was consequently more
449 notorious in the three urban stations Vall d'Hebron (1), Granollers (2) and Gràcia (6), which were the



450 stations with higher NO₂ concentration levels. Considering that the principal emission source of NO₂ is
 451 traffic, it is reasonable that the four MCR-ALS resolved components evidenced a decline in the year-mode,
 452 corresponding to a diminution in April 2020 (under the strictest lockdown), compared to 2019 (under no
 453 pandemia) and to 2018 (under no pandemia and no other traffic restrictions in Barcelona, such as the low
 454 emission zones(LEZ - Àrea Metropolitana de Barcelona, 2020)). Moreover, as stated in a previous study of
 455 the authors(Gorrochategui et al., 2021), in April 2020 an historical record of rainfall was registered in the
 456 control site of Observatory Fabra. Therefore, the highly rainy conditions of April 2020 favored the
 457 cleansing of the atmosphere, including NO₂ gases. Finally, the day profiles (x-mode) for the different
 458 components did not show any particular pattern for the different days of the month.
 459



460

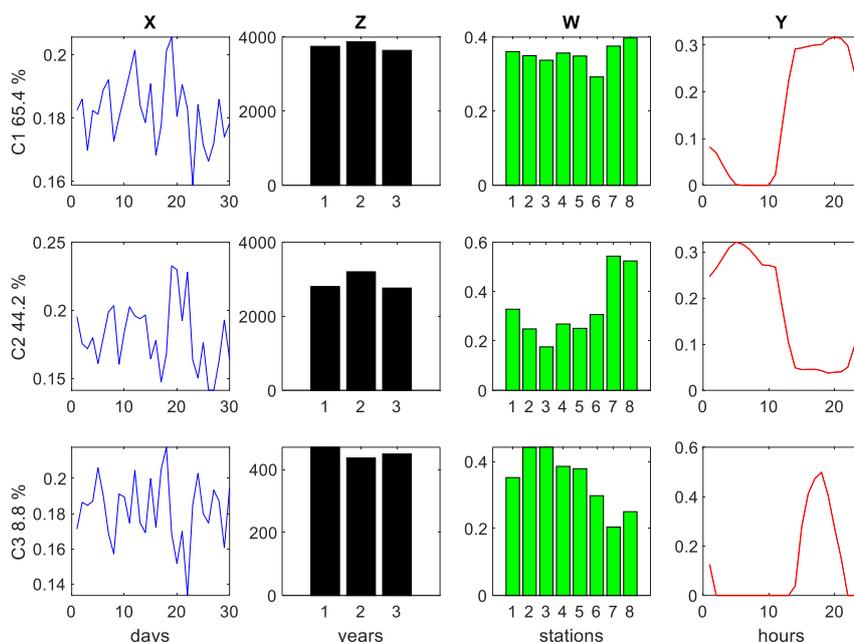
461 **Figure 4. MCR-ALS analysis of NO₂ concentrations in the column-wise super-augmented data matrix $D_{\text{caug-April-NO}_2}$ (Eq. (3)) using non-negativity and quadrilinearity constraints. Profiles of the four different data modes are given in different colors: (X) in blue days of April; (Z) in black year 1=2018, 2=2019 and 3= 2020; (W) in green stations 1: Vall d’Hebron, 2: Granollers, 3: Manlleu, 4: Juneda, 5: Bellver, 6: Gràcia, 7: Begur, 8: Observatori Fabra; and (Y) in red, hours of the day.**

466 **3.1.2. O₃ study ($D_{\text{caug-April-O}_3}$ data matrix)**

467 Profiles obtained by MCR-ALS for the three components using non-negativity and quadrilinearity
 468 constraints are shown in Fig. 5. MCR-ALS hourly (y-mode) resolved profiles of the first component (C1)
 469 showed a maximum between 14:00 and 22:00h, due to the cumulative solar radiation. There was practically
 470 no difference on this component among stations, among years nor among the days of the month. The MCR
 471 hourly resolved profile of the second component (C2) showed a different O₃ profile, corresponding to the
 472 concentration at night. As observed in the w-mode, O₃ concentration at night was higher in the rural station



473 of Begur and in the control site Observatori Fabra, the latter emplaced in Collserola mountain and receiving
 474 only some impact from Barcelona's city. The higher O₃ nightly concentration observed in these stations is
 475 due to the fact that in inner rural areas, as well as in the control site, with low anthropogenic activities, the
 476 titration effect (*i.e.*, ozone destruction under no solar radiation) produced by NO₂ emissions is generally not
 477 observed, resulting in higher average O₃ concentrations than in urban areas. Finally, third component (C3)
 478 showed again a maximum between 16:00 and 21:00 h, similar to the behavior described by C1, but narrower
 479 and with a pattern among stations different to C1.

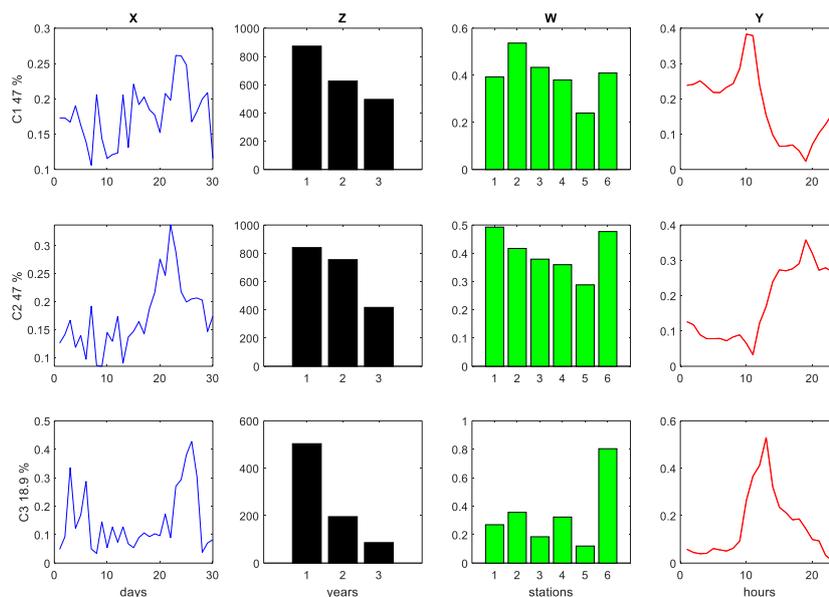


480

481 **Figure 5.** MCR-ALS analysis of O₃ concentrations in the column-wise super-augmented data matrix $D_{\text{caug-April-O}_3}$ (Eq. (3)) using non-negativity and quadrilinearity constraints. Profiles of the four different data modes are
 482 given in different colors: (X) in blue days of April; (Z) in black year 1=2018, 2=2019 and 3= 2020; (W) in green
 483 stations 1: Vall d'Hebron, 2: Granollers, 3: Manlleu, 4: Juneda, 5: Bellver, 6: Gràcia, 7: Begur, 8: Observatori
 484 Fabra; and (Y) in red, hours of the day.
 485

486 3.1.3. PM₁₀ study ($D_{\text{caug-April-PM}_{10}}$ data matrix)

487 Profiles obtained by MCR-ALS for these three components using non-negativity and quadrilinearity
 488 constraints are shown in Fig. 6. MCR hourly resolved profiles in the y-mode for the three resolved
 489 components indicated a wide maximum between 00:00 and 15:00h (C1), between 15:00 and 22:00h (C2)
 490 and between 10:00 and 20:00h (C3). As observed in the year profile (z-mode), PM₁₀ contribution decreased
 491 in 2019 but most significantly in 2020, probably due to the COVID-19 lockdown. This behavior was the
 492 same observed for NO₂ (Fig. 4), and it is due to the fact that among the PM₁₀ sources, traffic should be also
 493 included. Moreover, such depletion was more evident in the urban stations profile (w-mode) of Vall
 494 d'Hebron, Granollers and Gràcia.



495

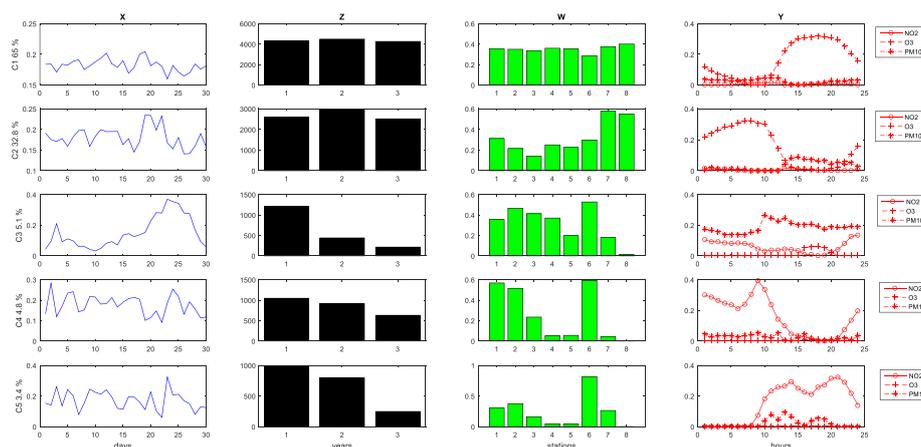
496 **Figure 6. MCR-ALS analysis of PM₁₀ concentrations in the column-wise super-augmented data matrix D_{caug}**
 497 **April-PM₁₀ (Eq. (3)) using non-negativity and quadrilinearity constraints. Profiles of the four different data modes**
 498 **are given in different colors: (X) in blue days of April; (Z) in black year 1=2018, 2=2019 and 3= 2020; (W) in**
 499 **green stations 1: Vall d’Hebron, 2: Granollers, 3: Manlleu, 4: Juneda, 5: Bellver, 6: Gràcia; and (Y) in red,**
 500 **hours of the day.**

501 **3.1.4. NO₂, O₃ and PM₁₀ simultaneous study (DA12_{cravg} data matrix)**

502 MCR-ALS resolved profiles of the DA12_{cravg} data matrix (see Method Sect. 2.7) are given in Fig. 7. Results
 503 obtained for the hour profiles (y-mode, in red) for the three pollutants, NO₂, O₃ and PM₁₀, are overlaid in
 504 the same plot with the same hour time axis. In this way, possible correlations among the different pollutants
 505 can be better explored in these plots. Profiles of components 1 (C1) and 2 (C2) mostly described the O₃
 506 pollution: C1 hour profile showed an ozone day-time profile with a wide maximum between 12:00 and
 507 22:00h and C2 described the ozone night-time profile, again with a large maximum between 00:00 and
 508 10:00h. Component 3 described both PM₁₀ and NO₂ correlated pollution sources, with particulate matter
 509 having the highest contribution. The correlation between NO₂ and PM₁₀ can be due to the common sources
 510 of these contaminants (*i.e.*, traffic and industry). Component 4 described the night-time profile of NO₂ and
 511 last component 5 showed the daily NO₂ profile with two maxima, one in the morning (10:00-15:00h) and
 512 another at the late evening (20:00- 22:00h), probably associated to the traffic. From the year profiles in z-
 513 mode, the evolution of the pollution in the month of April along 2018, 2019 and 2020 could be elucidated.
 514 Interestingly, for components 1 and 2 (mostly describing O₃ pollution), the variation remained rather
 515 constant for the month of April during these three years. Moreover, the variation among stations in the
 516 profiles (w-mode) for the two first components was very little. Only in component 2, the stations of Begur
 517 and Observatori Fabra showed a higher O₃ contribution, probably due to the lower titration effect produced
 518 in rural areas and in the Observatori Fabra control site, as previously observed in the individual MCR-ALS



519 analysis of O₃ data. In contrast, the variation among stations and among years was more significant for the
 520 rest of components (C3-C5), mainly describing NO₂ and PM₁₀ contamination. As observed in z-modes,
 521 contamination by NO₂ and PM₁₀ was lower in 2019 and even lower in 2020. Considering that the most
 522 important source of NO₂ is traffic, the decrease in 2019 can be explained by the implementation of the low
 523 emission zones (LEZs(LEZ - Àrea Metropolitana de Barcelona, 2020)) in Barcelona, as a traffic restriction
 524 policy, first implemented on 2017 and finally put into permanent effect on January 1, 2020. However, the
 525 decrease observed in 2020 might be mostly associated to the COVID-19 lockdown restrictions, being April
 526 2020, the time when the strictest confinement was declared in Catalonia (Real Decreto-ley 10/2020, de 29
 527 de marzo, por el que se regula un permiso retribuido recuperable para las personas trabajadoras por cuenta
 528 ajena que no presten servicios esenciales, con el fin de reducir la movilidad de la población en el contexto
 529 de la 1, 2020; Real Decreto 463/2020de 14 de marzo, por el que se declara el estado de alarma para la
 530 gestión de la situación de crisis sanitaria ocasionada por el COVID-19., 2020). Regarding the variation
 531 among stations, components C3 to C5 showed a higher NO₂ and PM₁₀ contribution in three urban stations
 532 (Vall d’Hebron, Granollers and Gràcia), which is in accordance to the higher traffic density registered on
 533 these sites. The results observed for components C3 to C5 regarding NO₂ and PM₁₀ pollution were in
 534 concordance to those of their respective individual models evidencing the good performance of the MCR-
 535 ALS simultaneous analysis of the incomplete multiblock data sets and the confirmation of the reliability of
 536 the proposed approach.



537 **Figure 7.** MCR-ALS analysis of NO₂, O₃ and PM₁₀ concentrations in the column-wise super-augmented
 538 incomplete April data matrix DA12_{caug} (see Eq. (9)) using non-negativity and quadrilinearity constraints.
 539 Profiles of the four different data modes are given in different colors: (X) in blue days of the year; (Z) in black
 540 year 1=2018, 2=2019 and 3= 2020; (W) in green stations 1: Vall d’Hebron, 2: Granollers, 3: Manlleu, 4: Juneda,
 541 5: Bellver, 6: Gràcia, 7: Begur, 8: Observatori Fabra; and (Y) in red, hours of the day.

543 **3.2. Study of the entire years**

544 **3.2.1. NO₂ study (D_{caug-allyear-NO2} data matrix)**



545 Profiles obtained by MCR-ALS using non-negativity and quadrilinearity constraints are shown in Fig. S4.
546 The hour profiles of the 4 resolved components in the analysis of the entire year were similar to those
547 obtained in the analysis of the month of April: C1 hour profile in April's model was equivalent to C3 hour
548 profile in all years' model and C2 and C4 hour profiles were equivalent in both models. Also, the diminution
549 observed in z-mode profile in 2019 and in a bigger extent in 2020 in the month of April was also produced
550 when analyzing all the year, but in a lesser extent. This might be due to the fact that the traffic restriction
551 policies were mostly implemented during the strictest confinement (from March 14th to May 4th in
552 Catalonia) and were gradually removed in the de-escalation phases (Gorrochategui et al., 2021). Also, the
553 extraordinary rainy conditions registered in April 2020 (Gorrochategui et al., 2021) were not registered for
554 the rest of the months, making the NO₂ depletion less noticeable in the analysis of the whole year. Regarding
555 stations, the ones showing the higher contribution were the same three urban stations (*i.e.*, Vall d'Hebron,
556 Granollers and Gràcia) observed in the study of the month of April. Observe also that in the day-of-year x-
557 profiles some seasonal tendencies can be observed in C1 and C2, with their lower intensities in the middle
558 of the profile corresponding to the warmer seasons with higher sunlight radiation and higher NO₂ depletion
559 due to the photochemical reaction to form O₃. C3 and C4 year profiles did not show major differences over
560 the year.

561 3.2.2. O₃ study (D_{caug-allyear-O3} data matrix)

562 Only few differences between the analysis of the entire years versus that of the month of April were
563 observed in the inter-year z-mode (Fig. S5). Component 2 in all years' model, corresponding to a late
564 evening peak of O₃, suffered a slightly significant increase in 2019 and 2020 respect to 2018, which was
565 not observed in the analysis of the month of April. Such increment can be explained by the reduction of the
566 titration effect, which was a little higher when considering all the year. The diminution of C3 in 2019 and
567 in 2020 was also more evident when analyzing the entire years. In this case, this component was associated
568 to the daily maximum of O₃, coincident with the sunlight hours and summer and spring seasons, when the
569 photochemical reactions with NO_x take place to form ozone. The reason why the changes in O₃ were more
570 evident when considering all year instead of when analyzing just the month of April, opposed to what
571 happened with NO₂, might be due to the fact that despite the traffic restrictions were gradually removed in
572 the de-escalation phases, the curfew policies remained, causing a potential cumulative suppression of the
573 titration effect.

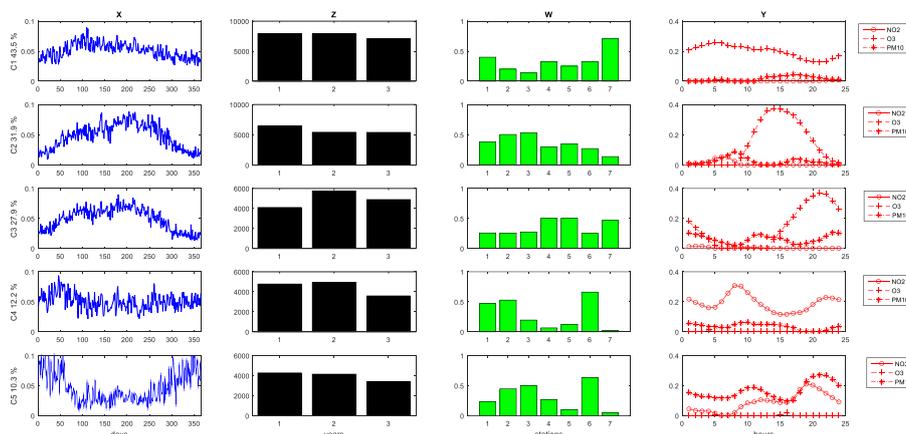
574 3.2.3. PM₁₀ study (D_{caug-allyear-PM10} data matrix)

575 As occurred with NO₂ and O₃, the profiles of the components in PM₁₀ MCR-ALS analysis of all year were
576 similar to those obtained in the analysis of the month of April (Fig. S6). C1 and C3 hours profile in April's
577 model were equivalent to C3 and C1 in all years' model, respectively, and C2 described the same PM₁₀
578 profile in both models. Also, the diminution observed in 2019 and in a bigger extent in 2020 in the month
579 of April was also produced when analyzing all the year, but in a lesser extent, as stated for NO₂. Moreover,
580 the meteorological stations with higher contribution in the model were the same than in the model of April,
581 except for Manlleu, which showed a significant contribution in C2 of this model for the first time when the
582 entire year PM₁₀ data were investigated.



583 **3.2.4. NO₂, O₃ and PM₁₀ simultaneous study (DY12_{craug} data matrix)**

584 MCR-ALS resolved profiles of **DY12_{craug}** are given in Fig. 8. As observed in the y-mode profiles,
 585 components 1 and 2 mostly described O₃ pollution: C1 showed an ozone profile with little daily variation
 586 whereas C2 described a wide O₃ maximum between 14:00 and 20:00h. Moreover, the seasonal trend of C2
 587 (x-mode) showed a wide maximum coincident with the solar radiation registered in summer and spring
 588 months. C2 was higher in the urban stations and lower in the rural station of Begur, which could indicate
 589 that such O₃ resulted from the photochemical reaction among NO_x in the presence of sunlight in highly
 590 transited areas. Component 3 clearly showed the nighttime profile of O₃, with a wide maxim between 17:00
 591 and 00:00h. Interestingly, this component was the only one showing clearly an increase in 2019 and 2020
 592 respect to 2018. As explained in the individual model, such increase is due to the diminution of the titration
 593 effect. Component 4 described the NO₂ profile, with a first maximum between 09:00 and 12:00h and a
 594 second but lower maximum at late evening (20:00-00:00h). Component 5 described the simultaneous
 595 contribution of NO₂ and PM₁₀, with higher contribution of PM₁₀, having again the same two-maxima profile
 596 observed in component 5 for NO₂. Interestingly, both components 5 and 6 presented maximums in the
 597 urban stations (Vall d’Hebron, Granollers and Gràcia) and a decrease in 2020, due to the traffic diminution
 598 registered during the COVID-19 lockdown.



599
 600 **Figure 8. MCR-ALS results of the simultaneous analysis of NO₂, O₃ and PM₁₀ for the entire years (incomplete**
 601 **super-augmented data matrix $D_{Y12craug}$) using non-negativity and quadrilinearity constraints. Profiles of the**
 602 **four different data modes are given in different colors: days of the month April (X) in blue, year (Z) in black,**
 603 **stations (W) in green and hours (Y) in red. In the y-mode, the hourly profiles of the three contaminants are**
 604 **overlapped. Capital letters in the figure indicate the different stations: 1: Vall d’Hebron, 2: Granollers, 3:**
 605 **Manlleu, 4: Juneda, 5: Bellver, 6: Gràcia, 7: Begur.**

606 **4. Conclusions**

607 MCR-ALS with quadrilinearity constraints has demonstrated to be a powerful tool to resolve the principal
 608 contamination profiles of four-way environmental datasets, even when containing missing data blocks. The
 609 main advantage provided by the use of quadrilinearity constraints is the better and easier interpretability of
 610 the profiles, which appear more condensed and concise.



611 In this study, resolved MCR profiles using quadrilinearity constraints have been shown to describe
612 adequately the different patterns and evolution of NO₂, O₃ and PM₁₀ contamination during the different
613 hours of the day, during the different days (hourly and daily variations) for the two periods of time
614 evaluated: month of April versus the entire year for 2018, 2019 and 2020. For each period of time studied,
615 the individual models of the contaminants together with their simultaneous analysis have been performed.
616 The simultaneous analysis of the incomplete multiblock data sets allowed the exploration of the potential
617 correlations among the three contaminants, which was easily interpretable with the representation of
618 overlapped NO₂, O₃ and PM₁₀ hourly profiles. Interestingly, both in the study of the month of April and the
619 study of the entire years, the simultaneous analysis of the three contaminants evidenced a correlation
620 between NO₂ and PM₁₀, due to their common pollution sources (*i.e.*, traffic and industry). Moreover, the
621 profiles of these two contaminants showed an inter-year decrease, due to the introduction of LEZs (LEZ -
622 Àrea Metropolitana de Barcelona, 2020) in 2019 and due to the COVID-19 lockdown restrictions and to
623 the high amount of rainfall registered in April 2020 (Real Decreto 463/2020 de 14 de marzo, por el que se
624 declara el estado de alarma para la gestión de la situación de crisis sanitaria ocasionada por el COVID-19,
625 2020). Such decrease was consistently higher in the three most transited urban stations studied: Vall
626 d'Hebron, Granollers and Gràcia.
627 On the other hand, MCR-ALS ozone profiles both in individual and simultaneous models presented an
628 opposite inter-year trend, especially when analyzing the entire years. Globally, O₃ profiles showed an
629 increase in 2019 and in 2020 respect to 2018, which can be attributed to the diminution of the titration effect
630 linked to the lockdown and curfew restrictions. Such effect was more evident in inner rural areas and in the
631 control site (*i.e.*, Begur and Observatori Fabra), where the amount of NO_x necessary to react with ozone
632 and to produce its suppression is lower compared to urban areas due to the smaller traffic density and
633 industrial activity.
634 Overall, this work contributes to the better knowledge of the evolution of NO₂, O₃ and PM₁₀ contamination
635 in eight rural and urban areas of Catalonia during the two years before the COVID-19 (*i.e.*, 2018 and 2019)
636 and the year itself of the pandemic (*i.e.*, 2020). The work also highlights: (a) the capacity of MCR-ALS
637 with quadrilinearity constraints to perform simultaneous analysis of different contamination sources to
638 study potential correlations among them and (b) the good performance of this approach in the analysis of
639 complex four-way environmental data sets containing missing data blocks, providing concise and easy
640 interpretable results.

641

642 **Author contributions**

643 EGM performed data curation, formal analysis and writing. IH provided air quality data. RT contributed to
644 data curation and global supervision.

645

646 **Competing interests**

647 The contact author has declared that neither they nor their co-authors have any competing interests.

648 **Financial support**



649 This study was supported by the Ministry of Science and Innovation of Spain under the project PID2019-
650 105732GB-C21.

651

652 References

653 Abdollahi, H. and Tauler, R.: Uniqueness and rotation ambiguities in Multivariate Curve Resolution
654 methods, *Chemom. Intell. Lab. Syst.*, 108, 100–111, <https://doi.org/10.1016/J.CHEMOLAB.2011.05.009>,
655 2011.

656 Alier, M., Felipe, M., Hernández, I., and Tauler, R.: Trilinearity and component interaction constraints in
657 the multivariate curve resolution investigation of NO and O₃ pollution in Barcelona, *Anal. Bioanal.*
658 *Chem.*, 399, <https://doi.org/10.1007/s00216-010-4458-1>, 2011.

659 LEZ - Àrea Metropolitana de Barcelona: [https://www.zbe.barcelona/en/zones-baixes-emissions/la-](https://www.zbe.barcelona/en/zones-baixes-emissions/la-zbe.html)
660 [zbe.html](https://www.zbe.barcelona/en/zones-baixes-emissions/la-zbe.html), last access: 12 December 2020.

661 Multivariate Curve Resolution of incomplete data multisets | Elsevier Enhanced Reader:
662 [https://reader.elsevier.com/reader/sd/pii/S0169743913001007?token=DF6B30679DAB7F46A10398BAD](https://reader.elsevier.com/reader/sd/pii/S0169743913001007?token=DF6B30679DAB7F46A10398BAD3FD3C9CC1EEBD01DCC607B7B12D43E21198292040FE9D46BA56DA57D3E6AEFF9F11AC35&originRegion=eu-west-1&originCreation=20220215102437)
663 [3FD3C9CC1EEBD01DCC607B7B12D43E21198292040FE9D46BA56DA57D3E6AEFF9F11AC35&ori-](https://reader.elsevier.com/reader/sd/pii/S0169743913001007?token=DF6B30679DAB7F46A10398BAD3FD3C9CC1EEBD01DCC607B7B12D43E21198292040FE9D46BA56DA57D3E6AEFF9F11AC35&originRegion=eu-west-1&originCreation=20220215102437)
664 [ginRegion=eu-west-1&originCreation=20220215102437](https://reader.elsevier.com/reader/sd/pii/S0169743913001007?token=DF6B30679DAB7F46A10398BAD3FD3C9CC1EEBD01DCC607B7B12D43E21198292040FE9D46BA56DA57D3E6AEFF9F11AC35&originRegion=eu-west-1&originCreation=20220215102437), last access: 15 February 2022.

665 Bro, R. and De Jong, S.: A FAST NON-NEGATIVITY-CONSTRAINED LEAST SQUARES
666 ALGORITHM, *Journal of Chemometrics*, 393–401 pp., 1997.

667 Crutzen, P. J.: THE ROLE OF NO AND NO₂ IN THE CHEMISTRY OF THE TROPOSPHERE AND
668 STRA TOSPHERE, *Ann. Rev. Earth Planet. Sci.*, 7, 443–72, 1979.

669 Xarxa de Vigilància i Previsió de la Contaminació Atmosfèrica (XVPCA). Departament de Territori i
670 Sostenibilitat:

671 [http://mediambient.gencat.cat/ca/05_ambits_dactuacio/atmosfera/qualitat_de_laire/avaluacio/xarxa_de_vi-](http://mediambient.gencat.cat/ca/05_ambits_dactuacio/atmosfera/qualitat_de_laire/avaluacio/xarxa_de_vigilancia_i_previsio_de_la_contaminacio_atmosferica_xvzca/)
672 [gilancia_i_previsio_de_la_contaminacio_atmosferica_xvzca/](http://mediambient.gencat.cat/ca/05_ambits_dactuacio/atmosfera/qualitat_de_laire/avaluacio/xarxa_de_vigilancia_i_previsio_de_la_contaminacio_atmosferica_xvzca/), last access: 28 November 2020.

673 Real Decreto-ley 10/2020, de 29 de marzo, por el que se regula un permiso retribuido recuperable para las
674 personas trabajadoras por cuenta ajena que no presten servicios esenciales, con el fin de reducir la
675 movilidad de la población en el contexto de la I: [https://www.boe.es/buscar/doc.php?id=BOE-A-2020-](https://www.boe.es/buscar/doc.php?id=BOE-A-2020-4166)
676 [4166](https://www.boe.es/buscar/doc.php?id=BOE-A-2020-4166), last access: 20 December 2020.

677 Real Decreto 463/2020de 14 de marzo, por el que se declara el estado de alarma para la gestión de la
678 situación de crisis sanitaria ocasionada por el COVID-19.: [https://www.boe.es/buscar/doc.php?id=BOE-](https://www.boe.es/buscar/doc.php?id=BOE-A-2020-3692)
679 [A-2020-3692](https://www.boe.es/buscar/doc.php?id=BOE-A-2020-3692), last access: 11 December 2020.

680 Gorrochategui, E., Hernandez, I., Pérez-Gabucio, E., Lacorte, S., and Tauler, R.: Temporal air quality
681 (NO₂, O₃, and PM₁₀) changes in urban and rural stations in Catalonia during COVID-19 lockdown: an
682 association with human mobility and satellite data, *Environ. Sci. Pollut. Res. Int.*,
683 <https://doi.org/10.1007/S11356-021-17137-7>, 2021.

684 Jackson, J. E., Inti ; Jackson, J. E., Morris, H. J., Mudholkar, G. S., Fisher, M. T., Lee, J., Mara, M. K.,
685 and Murphy, B. J.: Interactive self-modeling mixture analysis, *Anal. Chem.*, 63, 1425–1432,
686 <https://doi.org/10.1021/AC00014A016>, 2002.

687 De Juan, A. and Tauler, R.: Chemometrics applied to unravel multicomponent processes and mixtures
688 Revisiting latest trends in multivariate resolution, *Anal. Chim. Acta*, 500, 195–210,



- 689 [https://doi.org/10.1016/S0003-2670\(03\)00724-4](https://doi.org/10.1016/S0003-2670(03)00724-4), 2003.
- 690 De Juan, A. and Tauler, R.: Comparison of three-way resolution methods for non-trilinear chemical data
691 sets, <https://doi.org/10.1002/cem.548>, n.d.
- 692 De Juan, A., Rutan, S. C., Tauler, R., and Luc Massart, D.: Comparison between the direct trilinear
693 decomposition and the multivariate curve resolution-alternating least squares methods for the resolution
694 of three-way data sets, *Chemom. Intell. Lab. Syst.*, 40, 19–32, [https://doi.org/10.1016/S0169-7439\(98\)00003-3](https://doi.org/10.1016/S0169-7439(98)00003-3), 1998.
- 696 Khaniabadi, Y. O., Goudarzi, G., Seyed, &, Daryanoosh, M., Borgini, A., Tittarelli, A., and De Marco,
697 A.: Exposure to PM 10 , NO 2 , and O 3 and impacts on human health, <https://doi.org/10.1007/s11356-016-8038-6>, n.d.
- 699 Lerda, M. T., Munger, J. W., and Jacob, D. J.: The NO₂ flux conundrum, *Science* (80-.), 289, 2291–
700 2293, <https://doi.org/10.1126/SCIENCE.289.5488.2291>/ASSET/DFAD0B0C-9C2B-4556-B374-
701 1D8E5FF627FB/ASSETS/GRAPHIC/2291-1.GIF, 2000.
- 702 Malik, A. and Tauler, R.: Extension and application of multivariate curve resolution-alternating least
703 squares to four-way quadrilinear data-obtained in the investigation of pollution patterns on Yamuna
704 River, India—A case study, *Anal. Chim. Acta*, 794, 20–28, <https://doi.org/10.1016/J.ACA.2013.07.047>,
705 2013.
- 706 Peterson, L. E.: K-nearest neighbor, 4, 1883, <https://doi.org/10.4249/SCHOLARPEdia.1883>, 2009.
- 707 Querol, X., Alastuey, A., Viana, M. M., Rodriguez, S., Artiñano, B., Salvador, P., Garcia Do Santos, S.,
708 Fernandez Patier, R., Ruiz, C. R., De La Rosa, J., Sanchez De La Campa, A., Menendez, M., and Gil, J.
709 I.: Speciation and origin of PM₁₀ and PM_{2.5} in Spain, *J. Aerosol Sci.*, 35, 1151–1172,
710 <https://doi.org/10.1016/j.jaerosci.2004.04.002>, 2004.
- 711 Saud, T., Mandal, T. K., Gadi, R., Singh, D. P., Sharma, S. K., Saxena, M., and Mukherjee, A.: Emission
712 estimates of particulate matter (PM) and trace gases (SO₂, NO and NO₂) from biomass fuels used in rural
713 sector of Indo-Gangetic Plain, India, *Atmos. Environ.*, 45, 5913–5923,
714 <https://doi.org/10.1016/j.atmosenv.2011.06.031>, 2011.
- 715 Smilde, A., Bro, R., and Geladi, P.: Multi-way Analysis with Applications in the Chemical Sciences i
716 Multi-way Analysis in Chemistry and Related Fields, 2004.
- 717 Soloveychik, I. and Trushin, D.: Gaussian and robust Kronecker product covariance estimation: Existence
718 and uniqueness, *J. Multivar. Anal.*, 149, 92–113, <https://doi.org/10.1016/J.JMVA.2016.04.001>, 2016.
- 719 Tauler, R.: Multivariate curve resolution applied to second order data, *Chemom. Intell. Lab. Syst.*, 30,
720 133–146, [https://doi.org/10.1016/0169-7439\(95\)00047-X](https://doi.org/10.1016/0169-7439(95)00047-X), 1995.
- 721 Tauler, R.: Multivariate curve resolution of multiway data using the multilinearity constraint, *J.*
722 *Chemom.*, 35, <https://doi.org/10.1002/cem.3279>, 2021.
- 723 Zú, J., Niga, ~, Tarajia, M., Herrera, V., Urriola, W., Gómez, B., and Motta, J.: Assessment of the
724 Possible Association of Air Pollutants PM 10 , O 3 , NO 2 With an Increase in Cardiovascular,
725 Respiratory, and Diabetes Mortality in Panama City Data Analysis,
726 <https://doi.org/10.1097/MD.0000000000002464>, n.d.
- 727