

1 **D-OPTIMAL EXPERIMENTAL DESIGN COUPLED WITH PARALLEL FACTOR**
2 **ANALYSIS 2 DECOMPOSITION A USEFUL TOOL IN THE DETERMINATION OF**
3 **TRIAZINES IN ORANGES BY PROGRAMMED TEMPERATURE**
4 **VAPORISATION-GC/MS WHEN USING DISPERSIVE-SOLID PHASE**
5 **EXTRACTION**

6
7 A. Herrero¹, M.C. Ortiz^{1,*}, L.A. Sarabia²

8 ¹ *Department of Chemistry, University of Burgos, Faculty of Sciences,*
9 *Pza. Misael Bañuelos s/n, 09001 Burgos, Spain*

10 ² *Department of Mathematics and Computation, University of Burgos, Faculty of Sciences,*
11 *Pza. Misael Bañuelos s/n, 09001 Burgos, Spain*

12
13
14
15 **Abstract**

16
17 The determination of triazines in oranges using a GC/MS system coupled to a programmed
18 temperature vaporizer (PTV) inlet in the context of legislation is performed. Both
19 pretreatment (using a Quick Easy Cheap Effective Rugged and Safe (QuEChERS) procedure)
20 and injection steps are optimized using D-optimal experimental designs for reducing the
21 experimental effort. The relative dirty extracts obtained and the elution time shifts make it
22 necessary to use a PARAFAC2 decomposition to solve these two usual problems in the
23 chromatographic determinations. The “second-order advantage” of the PARAFAC2
24 decomposition allows unequivocal identification according to document SANCO/12495/2011
25 (taking into account the tolerances for relative retention time and the relative abundance for
26 the diagnostic ions), avoiding false negatives even in the presence of unknown co-eluent.
27 The detection limits (CC α) found, from 0.51 to 1.05 $\mu\text{g kg}^{-1}$, are far below the maximum
28 residue levels (MRLs) established by the European Union for simazine, atrazine,
29 terbuthylazine, ametryn, simetryn, prometryn and terbutryn in oranges. No MRL violations
30 were found in the commercial oranges analysed.

31
32 *Keywords: PARAFAC2, experimental design, QuEChERS, PTV-GC/MS, triazines in oranges,*
33 *SANCO/12495/2011.*

34
35
36 **1. Introduction**

37
38 The setting of low harmonized maximum residue limits (MRLs) for pesticides in food [1] and
39 the need of controlling their residues in a large number of food samples have highlighted the
40 problem of working with complex matrices which require pretreatment stages to eliminate
41 interferent compounds. Procedures including partitioning with organic solvents, adsorption
42 chromatography and gel permeation chromatography or solid-phase extraction have been
43 developed. However, the introduction of these steps causes an increase in the time of analysis,
44 high consumption of organic non-environmentally friendly solvents and a source of losses in
45 the analytical recoveries [2].

46
47 Different techniques have been developed to address this problem, among them the approach
48 known as the Quick Easy Cheap Effective Rugged and Safe (QuEChERS) multiresidue
49 methodology first reported in 2003 [3]. Compared to traditional approaches, this approach is a

* Corresponding author. e-mail: mcortiz@ubu.es. Tel.: 34 947259571. Fax: 34 947258831.

50 rapid, straightforward, and cost-effective sample preparation procedure with which a large
51 number of samples can be processed simultaneously.

52
53 The QuEChERS approach typically involves an extraction with acetonitrile followed by a
54 clean-up step (which is not always necessary [3]) which consists of a dispersive solid-phase
55 extraction (dSPE); the final determination is carried out by gas or liquid chromatography (GC
56 or LC) coupled to mass spectrometry (MS), taking the advantage of the high selectivity and
57 sensitivity provided by these techniques. Although the QuEChERS method was initially
58 developed for determining multiclass pesticides in fruits and vegetables [4], currently
59 modifications of the original method [5,6,7] and applications for a wide variety of analytes in
60 a wide variety of matrices can be found in the literature. For example, it has been applied for
61 the extraction of veterinary drugs in fish [8], beef muscle [9], milk and liver [10], for the
62 extraction of chlorinated compounds from soil samples [11] and of drugs in blood [12].

63
64 But the use of this multiresidue method increases the presence of co-extracts. Even after dSPE
65 clean-up, QuEChERS extracts are relatively dirty because of the risk of removing pesticides
66 along with other matrix compounds if refined clean-up steps are used, so the extracts can still
67 contain co-extracted compounds which could interfere with the detection and identification of
68 target analytes. The document SANCO/12495/2011 [13] recommends using “the ion that
69 shows the best signal-to-noise ratio and no evidence of significant chromatographic
70 interference” to quantify residues, specifying that an ion chromatogram that “shows evidence
71 of significant chromatographic interference must not be relied upon to quantify or identify
72 residues”. But the presence of non-target compounds can cause false negatives during
73 pesticides identification [14], since the maximum permitted tolerances for relative ion
74 abundances established in the regulations [13] for diagnostic ions will not be fulfilled if some
75 fragments of the non-target compounds contribute inadvertently to the abundance of some of
76 the m/z ratios of the pesticides.

77
78 The problem of overlapping peaks in GC/MS or LC/MS can be approached using parallel
79 factor analysis (PARAFAC), which makes discrimination possible from co-eluting matrix
80 components if the data are trilinear [14]. The PARAFAC decomposition provides the same
81 number of factors as there are compounds whose signal is higher than the expected signal-to-
82 noise ratio, as well as the mass spectrum and chromatographic profile of each compound. This
83 includes compounds that co-elute and several artifacts like baseline. This is the “second-order
84 advantage” of PARAFAC in chemical analysis. However, the PARAFAC model is greatly
85 affected by deviations from the trilinear structure of the data, in such a way that slight
86 changes in the retention time of an analyte between runs, which are usual in chromatography,
87 lead to the invalidation of the PARAFAC model. For that reason, if some deviations in the
88 chromatographic profiles have to be modelled, then the parallel factor analysis 2
89 (PARAFAC2) model, which was proposed in order to overcome this difficulty, must be used
90 [15,16]. PARAFAC2 has the “second-order advantage” if the correlation between the
91 retention times is the same in all samples. Applications of PARAFAC and PARAFAC2 to
92 chromatographic analysis can be found in refs. [17,18,19].

93
94 In this work, we describe the determination of triazines in oranges. Orange samples were
95 pretreated with a commercial kit for use with the QuEChERS method and next the extracts
96 were analysed using a GC/MS system equipped with a programmed temperature vaporizer
97 (PTV) inlet. Coupling this with large volume injections (LVI) the procedure can be improved
98 since LVI techniques are a reliable alternative to carry out the preconcentration of samples
99 inside the chromatographic system. For introducing this large volume of sample, repetitive or
100 speed controlled injection can be used, being the latter the one that leads to better results [20].

101
102 For the optimization of the analytical procedure, traditional one-factor-at-a-time experiments
103 do not address interactions among experimental factors; therefore experimental design
104 strategies can be very helpful since a large number of factors are usually involved. Among
105 them, D-optimal designs [21] make it easy the study of many experimental factors with a
106 small number of experiments, and in addition they are a general methodology for making ‘ad
107 hoc’ designs by adapting the experimental design to each analytical problem (it is possible to
108 consider independently for each factor as many levels as it is required and their interactions)
109 [22,23,24].

110
111 Most of the works found in the literature dealing with the QuEChERS method optimization
112 focused on the study of experimental conditions and/or the composition of salts, sorbents, etc.
113 employed in the extraction/partitioning and/or clean-up steps [9,25]. However, if a
114 commercial kit which contains tubes with pre-weighed sorbents and buffers for use with the
115 QuEChERS method for a certain application is employed, as it is in this work, most of these
116 parameters are already fixed. In this case, it is possible to optimize the sample preparation
117 protocol provided by the supplier with the kit for that application, since it serves just as a
118 guideline for a large number of compounds with different chemical properties. In this work,
119 several parameters of the QuEChERS procedure have been optimized using a D-optimal
120 design coupled to PARAFAC2 decomposition for selecting the best experimental conditions
121 of the pretreatment. In addition, several experimental parameters of the injection step
122 performed with the PTV inlet have also been optimized using another D-optimal design
123 coupled, in this case, to PARAFAC decomposition.

124
125 The analytical procedure used for the determination of these triazines was validated and
126 orange samples purchased from different food stores were analysed. The EU establishes
127 MRLs for simazine (SZ), atrazine (AZ), terbuthylazine (TZ), simetryn (ST), ametryn (AT),
128 prometryn (PT) and terbutryn (TT) in oranges [1,26,27], so compliance with those MRLs was
129 checked. The MRLs (0.10 mg kg^{-1} for TZ, 0.05 mg kg^{-1} for AZ, and 0.01 mg kg^{-1} for SZ, ST,
130 AT, PT and TT) are not exceeded in any case.

131
132

133 **2. Experimental**

134

135 *2.1 Chemicals and materials*

136

137 The triazines (SZ, AZ, PZ, TZ, ST, AT, PT, and TT; PESTANAL grade) were purchased
138 from Sigma-Aldrich (Madrid, Spain). Methanol, ethyl acetate and acetonitrile (HPLC grade)
139 were obtained from Merck (Darmstadt, Germany). All orange samples were purchased from
140 local food stores.

141

142 QuEChERS method was performed using a DisQuE dispersive sample preparation kit from
143 Waters (Milford, MA, USA), which consisted of 50 mL tubes containing 6 g anh. MgSO_4
144 plus 1.5 g anh. sodium acetate (DisQuE extraction tube 1) and 2 mL tubes containing 150 mg
145 anh. MgSO_4 plus 50 mg PSA sorbent and 50 mg C_{18} for d-SPE clean-up (DisQuE clean-up
146 tube 2).

147

148 *2.2 Instrumental*

149

150 The analyses were carried out on an Agilent (Agilent Technologies, Wilmington, DE, USA)
151 7890A gas chromatograph coupled to an Agilent 5975 Mass Selective Detector (MSD). The

152 injection system consisted of a septumless head and a PTV inlet (CIS 6 from Gerstel,
153 Mülheim an der Ruhr, Germany) which was equipped with an empty multi-baffled
154 deactivated quartz liner. LVI was carried out using a MultiPurpose Sampler (MPS 2XL from
155 Gerstel) with a 10 μL syringe. Analytical separations were performed on an Agilent DB-5ms
156 (30 m \times 0.25 mm i.d., 0.25 μm film thickness) column. To centrifuge the extracts, a
157 refrigerated tabletop centrifuge Sigma 2-16K (Osterode, Germany) was used. A miVac DUO
158 centrifugal concentrator (Genevac Ltd., Ipswich, UK) operating at low pressure was used for
159 faster evaporation.

160

161 2.3 GC/MS conditions

162

163 The PTV inlet was operated in the solvent vent mode. A volume of 10 μL was injected at a
164 controlled speed of $s \mu\text{L s}^{-1}$ (*injection speed*). After each injection the syringe was washed
165 several times first with acetone and next with ethyl acetate. During injection, the inlet
166 temperature was held at $T1 \text{ }^\circ\text{C}$ (*initial temperature*) for $t1 \text{ min}$ (*PTV initial time*), while the
167 column head pressure was fixed to $p \text{ psi}$ (*inlet P*) and the flow rate through the split vent was
168 set at $f \text{ mL min}^{-1}$ (*vent flow*). At a solvent vent time of $t2 \text{ min}$ (*vent time*) the split valve was
169 closed. Next, the inlet temperature was ramped at $r \text{ }^\circ\text{C s}^{-1}$ (*PTV rate*) up to $T2 \text{ }^\circ\text{C}$ (*end*
170 *temperature*), which was held for 3 min. Afterwards, the temperature reached $280 \text{ }^\circ\text{C}$ at a rate
171 of $1 \text{ }^\circ\text{C s}^{-1}$, and held at $280 \text{ }^\circ\text{C}$ for 5 min. The split valve was re-opened 2 min after injection
172 to purge the inlet at a vent flow of 60 mL min^{-1} . See Table 1 for the codification and levels of
173 experimental variables or factors.

174

175 The oven temperature program was as follows: $40 \text{ }^\circ\text{C}$ (for 2 min), temperature increase at 60
176 $^\circ\text{C min}^{-1}$ to $130 \text{ }^\circ\text{C}$ (0 min), then $7 \text{ }^\circ\text{C min}^{-1}$ to $160 \text{ }^\circ\text{C}$ (1 min), $3.5 \text{ }^\circ\text{C min}^{-1}$ to $178 \text{ }^\circ\text{C}$ (0
177 min), and finally $50 \text{ }^\circ\text{C min}^{-1}$ to $220 \text{ }^\circ\text{C}$ (2 min). A post-run step was performed for 4 min at
178 $280 \text{ }^\circ\text{C}$.

179

180 After 11 min (solvent delay), the mass spectrometer (MS) was operated in electron ionization
181 (EI) mode at 70 eV in selected ion monitoring (SIM) mode, with an acquisition window for
182 each analyte. 5 ions (ion dwell time of 80 ms) were detected for each peak: 158, 173, 186, 201
183 and 203 for SZ; 173, 200, 202, 215 and 217 for AZ; 172, 187, 214, 229 and 231 for PZ; 173,
184 214, 216, 229 and 231 for TZ; 155, 170, 198, 213 and 215 for ST; 170, 185, 212, 227 and 229
185 for AT; 184, 199, 226, 241 and 243 for PT; and 170, 185, 226, 241 and 243 for TT. The
186 transfer line temperature was set at $250 \text{ }^\circ\text{C}$, the ion source temperature at $230 \text{ }^\circ\text{C}$, and the
187 quadrupole at $150 \text{ }^\circ\text{C}$. The electron multiplier was set at 1576 V and the source vacuum at
188 10^{-5} Torr .

189

190 2.4 Samples, standards and sample preparation procedure

191

192 2.4.1 Standards

193 Stock solutions of the height triazines (100 mg L^{-1} of each triazine) were prepared in
194 methanol and stored in a refrigerator at $4 \text{ }^\circ\text{C}$. Next, solutions of each triazine were prepared in
195 ethyl acetate at a concentration of 1 mg L^{-1} from the stock solutions. Two sets of seven
196 calibration standards were needed. Low level concentration calibration standards were
197 prepared in ethyl acetate to each contain $10 \mu\text{g L}^{-1}$ of the internal standard (PZ) plus 0, 1, 2, 5,
198 10, 15 and $20 \mu\text{g L}^{-1}$ of SZ, AZ, TZ, ST, AT, PT, and TT. High level concentration
199 calibration standards were prepared in ethyl acetate to each contain $75 \mu\text{g L}^{-1}$ of the internal
200 standard (PZ) plus $20 \mu\text{g L}^{-1}$ of AZ and $70 \mu\text{g L}^{-1}$ of TZ, or $30 \mu\text{g L}^{-1}$ of AZ and $80 \mu\text{g L}^{-1}$ of
201 TZ, or $40 \mu\text{g L}^{-1}$ of AZ and $90 \mu\text{g L}^{-1}$ of TZ, or $50 \mu\text{g L}^{-1}$ of AZ and $100 \mu\text{g L}^{-1}$ of TZ, or 60

202 $\mu\text{g L}^{-1}$ of AZ and $110 \mu\text{g L}^{-1}$ of TZ, or $70 \mu\text{g L}^{-1}$ of AZ and $120 \mu\text{g L}^{-1}$ of TZ, or $80 \mu\text{g L}^{-1}$ of
203 AZ and $130 \mu\text{g L}^{-1}$ of TZ respectively.

204
205

2.4.2 Samples

206 The oranges were purchased from four different food stores (P1, P2, P3 and P4). Each orange
207 was cut with a knife and put into a freezer overnight. The sample was blended while frozen
208 until it reaches homogeneous texture. Then 15 mL 5% acetic acid (v/v) in acetonitrile and 15
209 g of the homogenized sample were added into the 50 mL DisQuE extraction tube 1 and after
210 vortex mixing for 2 min, the homogenate was centrifuged at 1500 rcf, 10°C for 5 min. 1.2 mL
211 of the acetonitrile extract was transferred into the DisQuE clean-up tube 2; the tube was
212 shaken for 60 s and next centrifuged at 1500 rcf, 10°C for 1 min. 0.5 mL of the supernatant
213 were transferred into a tube and evaporated to dryness under vacuum in a centrifugal
214 concentrator during 20 min at 50°C . The residue was reconstituted with 500 μL of ethyl
215 acetate containing $10 \mu\text{g L}^{-1}$ of PZ as internal standard, filtered through Simplepure nylon
216 membranes (13-mm diameter, 0.22- μm , Membrane Solutions, Spring View Lane Plano, TX)
217 and transferred into a vial with insert for analysis.

218
219

2.4.3 Fortified samples

220 Fortified orange samples were prepared following the procedure described in Section 2.4.2 for
221 orange samples but in this case homogenized orange samples were fortified and vortex mixed
222 for 30 s before extraction. A set of orange samples was fortified to each contain 0 or $10 \mu\text{g L}^{-1}$
223 of SZ, AZ, TZ, ST, AT, PT, and TT for low level concentration analysis, and another set was
224 fortified with $0 \mu\text{g L}^{-1}$ of AZ and TZ, or $50 \mu\text{g L}^{-1}$ of AZ plus $100 \mu\text{g L}^{-1}$ of TZ for high level
225 concentration analysis. The residue obtained through evaporation was reconstituted with 500
226 μL of ethyl acetate containing $10 \mu\text{g L}^{-1}$ of PZ as internal standard for low level concentration
227 analysis and $75 \mu\text{g L}^{-1}$ of PZ for high level concentration analysis

228

229 Another set of 13 fortified orange samples was prepared to optimize the QuEChERS
230 procedure. These samples were fortified with $10 \mu\text{g L}^{-1}$ of SZ, AZ, TZ, ST, AT, PT, and TT
231 and prepared following the procedure described previously but with some modifications
232 according to the experimental design shown in Section 4.2 (see Table 2 for the codification
233 and levels of experimental variables or factors). In this case, 15 mL of m % of acetic acid
234 (v/v) in acetonitrile (*modifier*) and 15 g of the homogenized sample were added into the tube
235 1 and after vortex mixing for m1 min (*mix_t1*), the homogenate was centrifuged at 1500 rcf
236 and 10°C for m2 min (*centr_t1*). v mL of the acetonitrile extract (*volume*) was transferred
237 into the tube 2 that was shaken for m3 s (*mix_t2*) and centrifuged at 1500 rcf and 10°C for 1
238 min. 0.5 mL of the supernatant were evaporated to dryness during 20 min at e $^\circ\text{C}$ (*evap_T*).
239 The residue was reconstituted with 500 μL of ethyl acetate containing $10 \mu\text{g L}^{-1}$ of PZ,
240 filtered and transferred into a vial with insert for analysis.

241
242

2.4.4 Matrix-matched standards

243 10 matrix-matched standards were prepared following the procedure described in Section
244 2.4.2 but in this case the residues obtained through evaporation were fortified with the desired
245 amount of each pesticide and a known amount of internal standard solution to each contain 10
246 $\mu\text{g L}^{-1}$ of the internal standard (PZ) plus 0, 0.5, 1, 2, 3, 4, 5, 10, 15 and $20 \mu\text{g L}^{-1}$ of SZ, AZ,
247 TZ, ST, AT, PT, and TT.

248
249

2.5 Software

250
251

252 MSD ChemStation E.02.01.1177 (Agilent Technologies, Inc.) and Gerstel Maestro 1 (version
 253 1.3.20.41/3.5) were used for data acquisition and processing. PARAFAC and PARAFAC2
 254 models were performed with the PLS_Toolbox [28] for use with MATLAB version 7.10 (The
 255 MathWorks). The least squares regression models were built and validated with
 256 STATGRAPHICS Centurion XVI [29] and the least median of squares (LMS) regression
 257 models were fit with PROGRESS [30]. Decision limit, $CC\alpha$, and capability of detection,
 258 $CC\beta$, were determined using the DETARCHI program [31], and $CC\alpha$ and $CC\beta$ at the
 259 maximum residue limit (MRL) were estimated using NWAYDET (a program written in-
 260 house that evaluates the probabilities of false non-compliance and false compliance for n-way
 261 data). Both D-optimal experimental designs were built and analysed with NEMRODW [32]
 262
 263

264 3. Theory

265 3.1 PARAFAC and PARAFAC2 models

266
 267 The PARAFAC decomposition is a method that decomposes a data tensor, $\underline{\mathbf{X}}$, into trilinear
 268 factors [33,34], each consisting of three loading vectors. GC/MS data can be arranged for
 269 each chromatographic peak in a three-way array or data tensor. In this case, the PARAFAC
 270 structural model for the abundance of each sample (slab k-th of $\underline{\mathbf{X}}$), is described in Eq. (1)
 271
 272

$$273 \quad x_{ijk} = \sum_{f=1}^F a_{if} b_{jf} c_{kf} + e_{ijk}, \quad i=1, \dots, I; \quad j=1, \dots, J; \quad k=1, \dots, K \quad (1)$$

274
 275 where F is the number of factors (i.e. the total number of co-eluent analytes), a_{if} , b_{jf} and c_{kf}
 276 are the elements of the three loadings matrices \mathbf{A} , \mathbf{B} and \mathbf{C} , and e_{ijk} is the residual non
 277 explained by the trilinear model. In Eq. (1), \mathbf{a}_f is the chromatographic profile, \mathbf{b}_f is the
 278 spectral profile (the mass spectrum) and \mathbf{c}_f is the sample profile of the f-th analyte. Written in
 279 matrix notation the PARAFAC model, for the k-th slab of $\underline{\mathbf{X}}$, reads
 280

$$281 \quad \mathbf{X}_k = \mathbf{A} \mathbf{D}_k \mathbf{B}^T + \mathbf{E}_k, \quad k=1, \dots, K \quad (2)$$

282
 283 where \mathbf{D}_k is a diagonal matrix that holds the k-th row of matrix \mathbf{C} in its diagonal, and \mathbf{E}_k is the
 284 matrix of the residuals.
 285

286 The PARAFAC model assumes that the chromatographic and spectral profiles do not change
 287 shape in different experiments (only their magnitude) and the model fitted is highly affected if
 288 the structure of the data deviates considerably from this assumption. This trilinearity
 289 assumption can be relaxed in the chromatographic mode if the PARAFAC2 decomposition is
 290 used [35,36]. This is a slightly different decomposition method where the chromatographic
 291 profile depends also on the k-th sample, this solution being more accurate in this case. It is
 292 very rare to have alignment problems with MS data, but changes in retention times are very
 293 usual in chromatography [37]. The PARAFAC2 model can be expressed as follows:
 294

$$295 \quad \mathbf{X}_k = \mathbf{A}_k \mathbf{D}_k \mathbf{B}^T + \mathbf{E}_k = \mathbf{P}_k \mathbf{H} \mathbf{D}_k \mathbf{B}^T + \mathbf{E}_k, \quad k=1, \dots, K \quad (3)$$

296
 297 where \mathbf{D}_k , \mathbf{B}^T , and \mathbf{E}_k are defined as in Eq. (2), \mathbf{A}_k ($I \times F$) are the chromatographic mode
 298 loadings estimated for the k-th sample, \mathbf{P}_k is an orthogonal matrix of the same size as \mathbf{A}_k and
 299 \mathbf{H} is a small quadratic matrix with dimension equal to the number of components. In contrast

300 to PARAFAC, PARAFAC2 does not assume that \mathbf{A} is the same for all k but the cross-product
 301 matrix $\mathbf{A}_k^T \mathbf{A}_k$, which allows some deviation in the chromatographic profiles. PARAFAC2 has
 302 the “second-order advantage” if the correlation between the retention times is the same in all
 303 samples, which is a weaker condition than the equality of chromatographic profiles imposed
 304 by the PARAFAC model.

305

306

307 4. Results and discussion

308

309 4.1 Optimization of injection parameters

310

311 It is well known that the introduction of the sample into the chromatographic system has a
 312 great influence on sensitivity, trueness and precision, especially when LVI techniques are
 313 used. For that reason in the first part of this work, the optimization of some injection
 314 parameters was carried out in order to obtain the highest chromatographic responses. The
 315 eight experimental variables detailed in Section 2.3 (*injection speed, initial temperature, PTV*
 316 *initial time, inlet P, vent flow, vent time, PTV rate and end temperature*), which may influence
 317 the PTV injection, were optimized. Table 1 shows the levels considered for the studied
 318 factors; there were 7 factors at two levels and one factor at 3 levels, such that 384 experiments
 319 would be necessary in a full design. Such a number of experiments was unaffordable; hence
 320 the experimental effort was reduced using a D-optimal design [22].

321

322 The aim of this experimental design procedure is to reduce the experimental effort just
 323 enough to estimate with suitable precision effects and interactions previously established. For
 324 that, once factors and their levels are established and a mathematical model is proposed (i.e.
 325 the search space is defined) an exchange algorithm based on the D-optimality criterion is used
 326 to look for experimental matrices with good quality [22,38]. In our case the search space was
 327 defined by the 384 experiments of the full factorial design. For each “ n ” (number of
 328 experiments to be done) the exchange algorithm searched among the 384 experiments the “ n ”
 329 experiments that led to the joint confidence region for the coefficients of the model with the
 330 smallest volume. The final number of experiments of the D-optimal design, “ n ”, was chosen
 331 in such a way that the maximum of the variance inflation factors (VIFs) was close to 1 to
 332 guarantee the smallest possible variance for the estimated coefficients.

333

334 The mathematical reference-state model proposed for the response y as a function of the
 335 studied factors was:

336

$$337 \quad y = \beta_0 + \beta_{1A}x_{1A} + \beta_{2A}x_{2A} + \beta_{3A}x_{3A} + \beta_{4A}x_{4A} + \beta_{5A}x_{5A} + \beta_{6A}x_{6A} + \beta_{7A}x_{7A} + \beta_{8A}x_{8A} + \beta_{8B}x_{8B} +$$

$$338 \quad \beta_{1A3A}x_{1A}x_{3A} + \varepsilon$$

$$339 \quad (4)$$

340

341 where x_{ij} ($i = 1, 2, \dots, 8$ and $j = A, B$) are binary variables equal to 1 when the i -th factor is at
 342 the j -th level, and 0 in any other case; β_0 is the intercept, β_{ij} are the coefficients of the model,
 343 and β_{1A3A} is an additional coefficient to estimate the possible interaction between factors *vent*
 344 *flow* and *vent time*. The highest level was considered as the reference level for all the factors
 345 of the model (level C for the eighth factor and level B for the rest of factors); so the
 346 coefficients of this model measures the effect on the response when each factor changes from
 347 the reference level to another one. All the coefficients were estimated by least squares.

348 The model of Eq. (4) had 11 coefficients, therefore at least 11 experiments were necessary to
349 fit the model. In this case, the 384 possible experiments of the full factorial design were
350 reduced to the 16 experiments (plus 3 replicates) of the D-optimal design shown in Table 3
351 (experiments 12, 13 and 14 are replicates of experiment 11), for which a first minimum of the
352 maximum VIF was reached. The VIFs of the coefficients of the reduced model ranged
353 between 1.00 and 1.24, which meant that this design provides sufficiently precise estimates
354 for the coefficients of the model.

355
356 Injections of the calibration standard of $10 \mu\text{g L}^{-1}$ of SZ, AZ, TZ, ST, AT, PT and TT,
357 containing $10 \mu\text{g L}^{-1}$ of PZ as internal standard, were carried out according to the
358 experimental plan (in random order) in Table 3. Fig. 1a shows the total ion chromatograms
359 (TIC) acquired for experiments 2 and 11; absence of co-eluent and no shifting of the
360 chromatographic peaks can be seen. Table 3 also contains the experimental responses, which
361 are the standardized loadings of the sample mode calculated for each triazine through the
362 PARAFAC decomposition of a data tensor obtained from the 19 experiments of the D-
363 optimal design.

364
365 For obtaining these responses, GC/MS data from injections of the calibration standard
366 according to the experimental design were arranged in a data tensor, and next PARAFAC
367 decompositions were performed with them. For each experiment of the experimental plan, 5
368 m/z ratios were acquired at a range of J times around the retention time of each triazine and
369 the internal standard, in such a way that a data tensor with dimensions $I \times 5 \times 19$ was obtained
370 for the chromatographic peak of each triazine after baseline correction. The first dimension of
371 the datasets refers to the number of scans (I was 24, 21, 20, 18, 13, 10, 13 and 11 for SZ, AZ,
372 PZ, TZ, ST, AT, PT and TT respectively), the second dimension to the number of m/z ratios
373 at which abundance was acquired, and the third dimension to the number of experiments. A
374 PARAFAC model was built by applying the ALS algorithm to the tensor of each triazine and
375 to the tensor of the internal standard; the non-negativity constraint was enforced for the three
376 modes. Only one factor was necessary in all the PARAFAC decompositions, as expected
377 since there were no interferences; the variance explained with the multi-way models ranged
378 from 98.0 to 99.3 %.

379
380 For identification, the ratios of the loadings of the spectral profile of four diagnostic ions were
381 calculated for each triazine (expressed as a percentage of the loading of each ion with respect
382 to the highest loading, which corresponds to the base peak), and then the ratios were checked
383 to see if they were within the tolerance intervals established for relative ion abundances
384 according to the document SANCO/12495/2011 [13]. To calculate the permitted tolerance
385 intervals (see Table 4) a standard which contained $10 \mu\text{g L}^{-1}$ of all the triazines was used as
386 reference sample. For each triazine, there were at least three relative ion abundances (in fact
387 four) within the tolerance intervals, as the regulation requires when working with a standard
388 mass resolution detector in the SIM mode. Besides, the relative retention time (the ratio of the
389 retention time of the chromatographic profile of each triazine to that of the one of the internal
390 standard) corresponded to that of the reference sample with a tolerance of $\pm 0.5 \%$ as
391 document SANCO states. Since mass spectral and chromatographic profiles were
392 unequivocally identified for the samples of the experimental plan. The use of the PARAFAC
393 decomposition guaranteed a direct relation between the loadings of the sample profile and the
394 amount of each triazine present in each sample. As usual in chromatography, the raw
395 responses were standardized; the three way procedure for this task was developed and
396 discussed in ref. [39]. The loadings of the sample profile were standardized by dividing each
397 loading by the corresponding loading estimated for the internal standard (PZ) in the

398 PARAFAC decomposition performed with the data tensor of PZ. These standardized loadings
 399 were the responses of the D-optimal design showed in Table 3.

400
 401 The standardized loadings were used to fit the model of Eq. (4) for the seven triazines.
 402 Except for TZ, all the models were significant at 10% significance level (p-values < 0.10; null
 403 hypothesis: the linear regression model is not significant) and did not have significant lack of
 404 fit at 5 % significance level (p-values > 0.2; null hypothesis: the regression model adequately
 405 fits data). Data 12 was outlier in the models fit for PT and TT; therefore the final models for
 406 these triazines were fit without it (the VIFs of the coefficients of these models were between
 407 1.05 and 1.33 after outliers elimination). In all cases, the residuals followed a normal
 408 distribution. The coefficients of determination ranged from 0.80 to 0.93.

409
 410 Coefficients of the reference-state model of Eq. (4) were estimated by least squares and
 411 depend on the reference state chosen. But for the analysis of effects, since the selection of the
 412 reference state is arbitrary, makes the interpretation of coefficients difficult because if the
 413 selected reference state changes, the estimated coefficients change too. This problem is
 414 avoided if a presence-absence model is used for that, so the reference-state model of Eq. (4)
 415 was converted into the equivalent presence-absence model of Eq (5).

$$\begin{aligned}
 & y = \beta'_0 + \beta'_{1A} x_{1A} + \beta'_{1B} x_{1B} + \beta'_{2A} x_{2A} + \beta'_{2B} x_{2B} + \beta'_{3A} x_{3A} + \beta'_{3B} x_{3B} + \beta'_{4A} x_{4A} + \beta'_{4B} x_{4B} + \beta'_{5A} x_{5A} + \\
 417 & \beta'_{5B} x_{5B} + \beta'_{6A} x_{6A} + \beta'_{6B} x_{6B} + \beta'_{7A} x_{7A} + \beta'_{7B} x_{7B} + \beta'_{8A} x_{8A} + \beta'_{8B} x_{8B} + \beta'_{8C} x_{8C} + \beta'_{1A3A} x_{1A} x_{3A} + \\
 & \beta'_{1A3B} x_{1A} x_{3B} + \beta'_{1B3A} x_{1B} x_{3A} + \beta'_{1B3B} x_{1B} x_{3B} + \varepsilon
 \end{aligned}
 \tag{5}$$

418
 419
 420 The variables and coefficients in Eq. (5) have the same meaning as in Eq. (4) but now j = A,
 421 B, C; therefore it includes all the levels of the factors. In this case, each coefficient of Eq. (5)
 422 estimates the effect of the factor at the corresponding level on the response.

423
 424 But the coefficients of the model of Eq. (5) cannot be estimated by least squares because for
 425 each factor 'i' the sum $x_{iA} + x_{iB}$ (or $x_{iA} + x_{iB} + x_{iC}$, when there are three levels) equals 1.
 426 However, once coefficients of Eq. (4) have been calculated, as a relationship exists between
 427 the coefficients of both presence-absence and reference state models, the coefficients of Eq.
 428 (5) can be computed. Ref. [22] contains more details about the equations that relate the
 429 coefficients of both models. Fig. 2 shows the graphic study of the effects of the different
 430 injection conditions on standardized loadings. The bars show through the differences of the
 431 coefficients, the expected change of the responses as effect of changes of the levels of each
 432 factor. The significant effects (at 5% significance level) are those that are not within the
 433 interval depicted by the dash-dotted lines.

434
 435 More significant effects were found for the most volatile triazines (SZ and AZ). The effect of
 436 the factor *end temperature* (coefficient b_{7A}) was significant for the six triazines and had a
 437 negative sign for level A and a positive one for level B; i.e. the highest responses were
 438 achieved when injection was carried out at level B of this factor, namely this experimental
 439 parameter was set at 320 °C. The factor *injection speed* was also significant for three of the
 440 six models; in this case the maximum response was obtained for level C of the factor, i.e.
 441 when the sample was injected at 3.4 $\mu\text{L s}^{-1}$. And also *initial temperature* and *PTV rate* were
 442 significant for some models; the lowest level of both factors led to the best responses, i.e. the
 443 optimum conditions for these two parameters were 40 °C and 10 °C s^{-1} respectively. The rest
 444 of the parameters under optimization as well as the interaction between *vent flow* and *vent*
 445 *time* had no significant influence on the considered response. However the working

446 conditions were chosen bearing in mind the sign and size of the corresponding coefficients in
 447 Fig. 2 (taking into account the value of the majority) that maximized the response, so factors
 448 *vent flow*, *inlet P*, *solvent vent time* and *PTV initial time* were set to 100 mL min⁻¹, 8 psi, 0.3
 449 min, and 0.6 min, respectively. The model fit for TZ was not significant because the
 450 variability of the standardized loadings of the replicates was of the same order as the one in
 451 the rest of experiments in the design, even if it was low. Therefore it was concluded that these
 452 experimental parameters do not interfere significantly the standardized loadings obtained for
 453 TZ in the experimental domain; the optimal conditions found for the rest of triazines can be
 454 applied also to the analysis of TZ.

455

456 4.2 Optimization of the extraction parameters

457

458 Once the injection parameters were optimized, the analysis of orange samples was tackled.
 459 Obviously, a pretreatment step was needed to put the samples into the chromatographic
 460 system; in this case the QuEChERS procedure was chosen. Taking as reference the procedure
 461 proposed by the supplier of the QuEChERS kit for the analysis of pesticides in oranges, six
 462 variables of the QuEChERS extraction procedure explained in Section 2.4.3 and Table 2
 463 (*mix_t1*, *centr_t1*, *volume*, *mix_t2*, *evap_T* and *modifier*) were optimized.

464

465 The reference-state model fitted in this case was:

466

$$467 \quad y = \beta_0 + \beta_{1A}x_{1A} + \beta_{2A}x_{2A} + \beta_{3A}x_{3A} + \beta_{4A}x_{4A} + \beta_{5A}x_{5A} + \beta_{6A}x_{6A} + \beta_{6B}x_{6B} + \varepsilon \quad (6)$$

468

469 where x_{ij} ($i = 1, 2, \dots, 6$ and $j = A, B$) were binary variables that equals 1 when the i th factor is
 470 at the j th level, and 0 otherwise. β_0 is the intercept and β_{ij} are the coefficients of the model.
 471 The model also shows that the percentage of acetic acid in acetonitrile (*modifier*) was studied
 472 at three levels, and the remaining five factors were studied at two levels. The 96 possible
 473 experiments of the full factorial design were reduced to only 10 experiments (plus 3
 474 replicates) using the D-optimal design methodology previously described. The corresponding
 475 design, shown in Table 5, had the VIFs of the coefficients with values between 1.09 and 1.11,
 476 so precise enough estimates for the coefficients of the model could be expected. Like in
 477 Section 4.1, the reference-state model of Eq. (6) was converted into the equivalent presence-
 478 absence model of Eq (7).

479

$$480 \quad y = \beta'_0 + \beta'_{1A}x_{1A} + \beta'_{1B}x_{1B} + \beta'_{2A}x_{2A} + \beta'_{2B}x_{2B} + \beta'_{3A}x_{3A} + \beta'_{3B}x_{3B} + \beta'_{4A}x_{4A} + \beta'_{4B}x_{4B} + \beta'_{5A}x_{5A} + \\ 481 \quad \beta'_{5B}x_{5B} + \beta'_{6A}x_{6A} + \beta'_{6B}x_{6B} + \beta'_{6C}x_{6C} + \varepsilon \quad (7)$$

482

483 Different pretreatments were carried out on the fortified orange samples (described in Section
 484 2.4.3) setting the experimental conditions according to the experimental plan in Table 5. The
 485 obtained extracts were injected into the chromatographic system and the corresponding
 486 GC/MS signals were acquired (Fig. 1b shows the TIC acquired for experiment 4). There were
 487 shifts in the retention times and many co-eluent due to the relatively dirty extracts obtained
 488 with the QuEChERS procedure (a known disadvantage of using this pretreatment procedure).
 489 The high degree of both spectral and chromatographic interferences can cause false negatives
 490 during analyte identification, since the maximum permitted tolerances for relative ion
 491 abundances established in document SANCO/12495/2011 will not be fulfilled for at least 3
 492 diagnostic ions. As an example, Fig. 3 shows, in dotted lines, the abundances acquired for the
 493 5 ions of AZ obtained for the 13 experiments in Table 5. More than one chromatographic
 494 peak appeared in all cases, except for m/z 202 apparently, with apparent shifting in the

495 retention times from experiment to experiment when orange extracts were injected (dotted
496 lines). There were shifts with respect to the standards (continuous lines in Fig. 3) too. For
497 solving these problems, the multi-way approach is useful.

498
499 Like in Section 4.1, data tensors were built up for PARAFAC decomposition in order to
500 calculate the experimental design responses. However, when complex matrix are analysed, as
501 in this case, it has been proved [40] that the estimations of three-way models are more precise
502 if both standards and fortified samples are included in the decomposition step. In this case, the
503 7 low level concentration calibration standards described in Section 2.4.1 were analysed (the
504 sample containing $10 \mu\text{L}^{-1}$ of each triazine was 5 times replicated) and the corresponding
505 matrices were arranged together with those of the fortified samples of the experimental design
506 into data tensors in such a way that a data tensors with dimensions $I \times 5 \times 25$ were obtained for
507 each analyte. The first dimension of the datasets refers to the number of scans (I was 31, 22,
508 28, 26, 15, 10, 20 and 15 for SZ, AZ, PZ, TZ, ST, AT, PT and TT respectively), the second to
509 the number of ions monitored, and the third to the number of samples (7 standards + 5
510 replicated standard + 13 experiments of the D-optimal design). A PARAFAC model was built
511 for each analyte (the non-negativity constraint was enforced for the three ways), but the
512 changes in the retention time between samples caused the PARAFAC models to fail. From
513 the loadings of chromatographic and spectral modes, the relative retention time of no factor
514 matched that of the reference standard and the relative ion abundances were not within the
515 tolerance intervals established in ref. [13], i.e. the PARAFAC decomposition was not capable
516 of successfully extracting the contribution of the signals corresponding to each triazine. This
517 was due to the fact that PARAFAC model assumes that the chromatographic profiles do not
518 change shape in different samples and the model fitted is highly affected if the structure of the
519 data deviates considerably from the trilinear structure.

520
521 To solve this problem, PARAFAC2 models were built for each triazine and the internal
522 standard by applying the ALS algorithm with unimodality and non-negativity constraints in
523 the chromatographic mode and non-negativity constraint in spectral and sample modes,
524 respectively. One important difference as against PARAFAC is that constraints in the first
525 mode do not apply to the estimated profiles, \mathbf{A}_k , themselves but only to \mathbf{H} (Eq. (3)). Although
526 it is generally advised not to use constrains in the chromatographic mode, successful results
527 were obtained by using them.

528
529 Table 6 summarizes the characteristics of the PARAFAC2 models built for this second D-
530 optimal design for the optimization of QuEChERS procedure. The final models were chosen
531 by comparing models with different number of factors, taking into account the explained
532 variance, the CORCONDIA index, the unequivocal identification of each triazine
533 (verification of compliance with the maximum permitted tolerances for the relative
534 abundances and with the relative retention time tolerances), and the degree of agreement of
535 the loadings of the sample mode. The CORCONDIA index is a measure of the degree of
536 trilinearity of the data tensor and was developed by Bro and Kiers [41]; values over 70 are
537 adequate. In the case of PT, the initial PARAFAC2 decomposition of the data tensor led to
538 models with a CORCONDIA index far below 70 due to high interferent effect of co-eluent,
539 which suggested that the trilinearity assumption was not fulfilled. In this case, the method
540 required alignment to facilitate a suitable decomposition in the final model, with an adequate
541 CORCONDIA value. The alignment was readily achieved by using the TIC, and next a
542 smaller window of scans centred on the retention time of PT to avoid as far as possible co-
543 eluents was considered.

544

545 Three-factor models were necessary in the PARAFAC2 decomposition for SZ, AZ, TZ and
546 TT; for PZ, ST and PT, a two-factor model was necessary and a one-factor model for AT. In
547 all the cases, the explained variances were larger than 96.5%. In addition, the CORCONDIA
548 index was always greater than 83%, which indicated that the trilinearity hypothesis was
549 assumable. The detection of outliers was carried out by calculating Q residual and Hotelling's
550 T^2 indices and removing those objects whose values exceed the corresponding threshold at a
551 given confidence level. Only experiment 13 in Table 5 was considered as an outlier in the
552 model built for PZ since those two indices exceeded their threshold values at the 99%
553 confidence level; for that reason this object was removed from the data tensor in the final
554 PARAFAC2 decomposition for the internal standard.

555
556 The factor related to each triazine in the corresponding model was confirmed provided the
557 pesticides were identified according to the requirements established in document
558 SANCO/12495/2011 [13], as previously. Table 4 shows the relative abundances calculated
559 with the loadings of the spectral profiles. The number of diagnostic ions that verified the
560 compliance is shown in Table 6. At least 3 diagnostic ions verified the compliance in all
561 cases, as SANCO/12495/2011 establishes for unequivocal identification of a compound. In
562 cases where the fourth diagnostic ion did not verify the compliance, it was very close to the
563 corresponding interval.

564
565 The loadings of spectral, chromatographic and sample modes obtained for AZ are showed in
566 Fig. 4 as an example (the loadings of the chromatographic mode in PARAFAC2 models are
567 referred throughout the paper to loadings scaled by the last mode loadings [28]). The loadings
568 of the first factor were coherent with AZ, which enabled the unequivocal identification of this
569 triazine (Fig. 4a shows, through the chromatographic profiles, how the shift in the retention
570 time was pronounced). The loadings of the spectral mode of this factor (Fig. 4b) matched the
571 spectra obtained from the spectrum of a $10 \mu\text{g L}^{-1}$ standard used as reference. And it was also
572 confirmed that the relative retention time of the chromatographic profile obtained for each
573 sample was within the tolerance intervals estimated from the reference sample. This means
574 that the PARAFAC2 decomposition was capable of suitably extracting and differentiating the
575 information related to AZ and co-eluents, which were related to spectral and chromatographic
576 profiles of factors 2 and 3. As it can be clearly seen in the estimated sample profile, the
577 loadings of the first factor increased with the concentration of AZ as expected for the first 12
578 samples (standards) in Fig. 4c, whereas those for the other two factors remained nearly
579 constant for these samples and increased for orange extracts.

580
581 The loadings of the sample mode estimated for the internal standard (PZ) are shown in Fig.
582 5a. The second factor was identified as that the one related to PZ and the first corresponded to
583 a co-eluent, which only had significant loadings for the orange extracts. The loadings
584 obtained for the internal standard had practically constant values for the standards (samples
585 from 1 to 12) and higher values for the oranges extracts. This was due to the matrix-induced
586 response enhancement effect (co-extracts fill active sites, causing higher analyte transfer
587 efficiency and thus greater signals in the presence of the matrix [42]), which is reduced by the
588 use of the PTV injector but not completely eliminated [43]. This matrix-induced enhancement
589 was observed for the chromatographic peaks of all triazines when orange extracts were
590 injected.

591
592 Fig. 5b shows the standardized loadings of the sample mode calculated for AZ by dividing
593 each loading of Fig. 4c by the corresponding loading of Fig. 5a. In this case, as well as
594 compensating for small fluctuations in injection volume and changes in detector response, the
595 use of internal standard corrected to a certain extent the effect on the loadings of the matrix-

596 induced enhancement (this will be even clearer below, in Section 4.3). As the whole mass
597 spectra and chromatographic peak were taken into account to calculate the loadings of both
598 the analyte and the internal standard, the correction might be more effective [39].
599

600 The standardized loadings calculated for the 7 triazines were the responses for the D-optimal
601 experimental design being performed for optimizing the QuEChERS procedure (see Table 5).
602 Since experiment 13 was an outlier removed from the data set for PZ, therefore the final D-
603 optimal design had only 12 experiments. This implied that the VIFs of the coefficients of the
604 model of Eq. (6) ranged from 1.20 to 1.42, values which estimates precise enough guaranteed;
605 the removing of one experiment of the experimental plan did not worsen significantly the
606 quality of the estimates.
607

608 Due to high variability of the replicates and since the models were not significant at 95%
609 confidence level (only two models were significant at 90%), in this case the statistical
610 analysis based on the statistical significance of the models was not useful. As in addition the
611 design was almost saturated since there are only 9 different experiments to estimate 8
612 coefficients, the interpretation of the size of the coefficients was more suitable. Fig. 6 shows
613 the graphic study of the effects of the experimental factors studied on the responses.
614

615 In general, the effects of the experimental factors followed similar patterns for all the
616 triazines. Taking into account these patterns all the factors were set at the corresponding high
617 level for obtaining the largest responses. The vortex mixing time (*mix_t1*), the centrifugation
618 time (*centr_t1*) and the percentage of acetic acid in acetonitrile (*modifier*) of the extraction
619 step were 2 min, 5 min and 5% respectively; the volume of the extract (*volume*) and the
620 mixing time (*mix_t2*) of the clean-up step were 1.2 mL and 1 min; and the evaporation
621 temperature (*evap_T*) was 40°C. These levels are in solid bars in Fig. 6.
622

623 *4.3 Analysis of samples: identification and quantification of triazines in oranges.*

624

625 Once the analytical procedure was optimized, the analysis of pesticides in orange samples was
626 dealt with. The “second-order advantage” makes PARAFAC2 decomposition especially
627 useful for quantifying and identifying analytes in complex samples where unknown
628 interferences are present, as it has been seen above as it is the case here. This is of great interest
629 in identifying and quantifying substances for which a permitted limit has been established
630 (EU established MRLs for these pesticides: 0.10 mg kg⁻¹ for TZ and 0.05 mg kg⁻¹ for AZ [1];
631 and 0.01 mg kg⁻¹ for SZ [26], ST, AT, PT and TT [27]).
632

633 European guidelines recommend the use of matrix-matched standards whenever matrix-
634 enhancement is demonstrated to minimize errors associated with it [44]. Therefore in addition
635 to standards in ethyl acetate, matrix-matched standards which had the same concentration of
636 co-extracted matrix components were also used in this analysis for building the PARAFAC2
637 models. In this case, a data tensor of dimension I×5×44 was obtained for every triazine after
638 baseline correction. The first dimension of the data tensors refers to the number of scans (see
639 4th column in Table 6), the second dimension to the number of diagnostic ions acquired for
640 each compound, and the third dimension to the number of samples. The first 12 samples of
641 the data tensor correspond to the 7 standards in Section 2.4.1 with concentrations from 0 to 20
642 µg L⁻¹, plus 5 replicates of the 10 µg L⁻¹ standard. The next 6 samples correspond to the
643 fortified orange samples in Section 2.4.3 for low level concentration analysis; the first sample
644 was a blank sample and the next 5 were fortified orange samples for recovery studies. The
645 next 16 samples correspond to the 4 samples of Section 2.4.2, whose analysis was performed

646 in quadruplicate. And finally, the last 10 samples correspond to the 10 matrix-matched
647 standards in Section 2.4.4.

648
649 PARAFAC2 models were built from the decomposition of these data tensors; the
650 characteristics of these models are summarized in Table 6. In some cases the final models
651 have 43 samples instead of 44 because some outlier data were found. Chromatographic peaks
652 of 3 and 10 $\mu\text{g L}^{-1}$ matrix-matched standards were outliers for ST and TT respectively and
653 they were removed from the corresponding data tensor. On the other hand, in the models built
654 for TZ, AT and PT, Q residual and Hotelling's T^2 indices which exceeded their threshold
655 values at the 99% confidence level were obtained for the 20 $\mu\text{g L}^{-1}$ matrix-matched standard,
656 so this sample was removed from their data tensors and the three PARAFAC2 models were
657 built again. In the case of PT, the method still required alignment of the chromatographic
658 peaks.

659
660 The PARAFAC2 models had the same number of factors as (or greater than) the models built
661 from standards and fortified orange samples in Section 4.2 (first rows in Table 6). This can be
662 due to the fact that orange samples from different suppliers were arranged into the data tensor,
663 so different co-eluent related to new factors could appear in some chromatographic peaks. In
664 addition, these models had similar explained variance percentages and higher CORCONDIA
665 indexes than the previous ones. Probably, the adding up of matrix-matched standards to the
666 data tensors increased the trilinear structure of the data, what made the CORCONDIA indexes
667 increased.

668
669 In all the PARAFAC models built for the triazines, except for AT, the presence of other co-
670 eluting substances was observed, i.e. GC/MS signals were non-specific, and for that reason
671 more than one factor was necessary. In most cases this meant that the identification of the
672 triazines according to the legislation could not be done either through the chromatogram or
673 through the spectrum. The maximum permitted tolerances for the qualifier ions were
674 exceeded in many cases when the analysis was carried out in the "usual way", taking into
675 account the abundance acquired at a certain retention time. But this problem was avoided
676 when the PARAFAC2 decomposition was used, leading to the unequivocal identification of
677 all the triazines according to the requirements of legislation in force regarding both the
678 relative retention time tolerances and the maximum permitted tolerances (see Table 4) for the
679 relative ion abundance.

680
681 Again by way of example, a detailed analysis of the model built for AZ is shown. Fig. 7
682 shows the loadings obtained with the four-factor model. As in the previous case, the first
683 factor was unequivocally identified as AZ, through the loadings of the chromatographic and
684 spectral modes in Figs. 7a and 7b (see the verification of compliance in Table 6). The
685 impossibility of doing the identification by the "usual way" was clear in these figures. The
686 interference of co-eluent were highly significant in Fig. 7b, so it was very unlikely that the
687 corresponding relative abundances were within the maximum permitted tolerance intervals for
688 the 3 diagnostic ions stated by the legislation. In addition, it was also probable that any co-
689 eluent would modify severely the relative retention time. In any case, this would have led to
690 the wrong conclusion that there was not AZ in the sample, i.e. to false negatives.

691
692 The loadings of the sample mode for the first factor, Fig. 7c, followed the expected pattern;
693 the higher the concentration of AZ the higher the values were obtained. Only the first factor
694 had non null values of the loadings for standards (samples 1 to 7), which means that the rest
695 of factors were related to other co-eluent of the extract and they only had non null values in
696 the rest of samples (from extracts). Next, in the samples for recovery studies (8 to 13), the

697 loadings of the first factor reflected the first blank sample (sample 8) and the next five
698 fortified samples (samples 9 to 13) perfectly; they took zero value for the first object and
699 higher values for the rest. Next, for the orange samples whose concentration was being
700 determined (samples 14 to 34), the loadings had so low values that either there was no AZ in
701 the samples or it was at very low concentrations. And finally, for the matrix-matched
702 standards, the loadings increased their values with the concentration of AZ, as expected, but
703 higher values were obtained when comparing to samples 1 to 7 due to the matrix-induced
704 enhancement. However, this effect was compensated if standardized loadings were
705 considered; see Fig. 7d.

706
707 Once standardized loadings for the sample mode were calculated for all the triazines,
708 calibration lines “standardized loading vs. concentration” were built with the matrix-matched
709 standards. Table 7 shows the parameters of the regression models and the concentrations
710 estimated for the orange samples of Section 2.4.2 together with the confidence intervals at a
711 95% confidence level. The IUPAC recommendation has been followed; negative values of
712 calculated concentration were neither substituted by zero nor removed. Most of the calculated
713 concentrations for the samples of oranges were negative or statistically equal to zero
714 (confidence intervals contains zero), and in those cases where a significant positive result was
715 obtained, some of the values found were below the corresponding decision limits and all of
716 them were far below of the corresponding MRLs. That is, no MRL violations were found.

717 718 *4.4 Figures of merit of the analytical procedure*

719
720 Some figures of merit such as accuracy (trueness and precision), recovery and repeatability,
721 limit of decision and capability of detection both at null concentration and at the MRL were
722 calculated.

723 724 *4.4.1 Accuracy: trueness and precision*

725 In order to determine trueness, a least squares regression line between the concentration
726 calculated with the calibration model, c_{calc} , and the known concentration of the matrix-
727 matched standards, c_{true} , was fitted. Outlier data were detected using the least median of
728 squares (LMS) regression and then removed if their absolute values of standardized residual
729 were higher than 2.5, in such a way that a reweighted least squares (RLS) regression model
730 was built with the rest of data. Table 8 shows the parameters of the RLS models together with
731 the number of standards, the calibration range (those which includes concentration equal to
732 zero correspond to the quantification of the triazines in oranges) and some figures of merit. In
733 order to compare the results obtained, the data corresponding to the standards (in ethyl
734 acetate) are also included in the table. Similar results were obtained in both cases; despite
735 small differences between the values of the standard deviation of regression (s_{yx}) can be
736 found. The precision of an analytical procedure can be estimated for the studied concentration
737 range from the residual deviation standard of the regression “ c_{calc} vs. c_{true} ”; this value can be
738 regarded as an estimate of the intermediate repeatability in the analysed concentration range
739 [45]. Table 8 also shows the s_{yx} values estimated for the different triazines.

740
741 But to guarantee trueness, the joint hypotheses “the slope is 1 and the intercept is zero” has to
742 be jointly checked. Fig. 8 shows the joint confidence regions for slope and intercept estimated
743 for both matrix-matched standards and standards. In all the cases, the analytical procedures
744 fulfilled the property of trueness because the confidence ellipse contained the point (0,1). The
745 ellipse with the smallest size corresponded to PT in both cases (figs. 8(a) and 8(b)), and the
746 size of the remaining ellipses varied depending on the standard deviation of each regression
747 line. No significant differences were found between matrix-matched standards and standards.

748 The differences observed in the orientation of the ellipses were only due to the number of data
749 used to build the regression models.

750

751 In addition, Table 8 shows the mean of the absolute value of the relative errors in calibration.
752 As expected, in all cases larger errors were obtained for the matrix-matched standards, but
753 always within an acceptable range.

754

755 4.4.2 Recovery

756 Recovery was calculated from the 6 fortified orange samples (samples from 13 to 18 in the
757 PARAFAC2 decomposition in Section 4.3) whose pretreatment was described in Section
758 2.4.3. The first sample was a blank sample which only contained the internal standard and the
759 next 5 were orange samples fortified to contain $10 \mu\text{g L}^{-1}$ of the 7 analysed triazines. In Table
760 8, the average recovery rate is expressed as the percentage of the amount of each triazine
761 initially added in each sample that was found with the analytical procedure. The found
762 recovery rates ranged from 57% to 83%, except for SZ, for which the recovery only reached
763 36.7%. It is possible that the commercial kit used in the pretreatment step was not the most
764 suitable for SZ.

765

766 4.4.3 Repeatability

767 This figure of merit was calculated as the standard deviation of the concentration calculated
768 for the 5 samples fortified at $10 \mu\text{g L}^{-1}$ (samples from 14 to 18 of the PARAFAC2
769 decomposition in Section 4.3) for the orange samples, and from the 5 replicates of the
770 standard of $10 \mu\text{g L}^{-1}$ (standards from 8 to 12). Values between 0.11 and $1.9 \mu\text{g L}^{-1}$ were
771 obtained; higher values were also obtained for the fortified samples, as can be seen in Table 8.
772 This is reasonable taking the pretreatment step after the fortification of these samples into
773 account; this increases the uncertainty of the analytical results.

774

775 4.4.4 Decision limit and capability of detection

776 According to ISO 11843 [46] the decision limit, $CC\alpha$, is “*the value of the net concentration*
777 *the exceeding of which leads, for a given error probability α , to the decision that the*
778 *concentration of the analyte in the analysed material is larger than that in the blank*
779 *material”*. Whereas the capability of detection, x_d or $CC\beta$, for a given probability of false
780 positive α , is “*the true net concentration of the analyte in the material to be analysed which*
781 *will lead, with probability $1-\beta$, to the correct conclusion that the concentration in the*
782 *analysed material is larger than that in the blank material”*. The need of assessing both the
783 probability of false positive, α , and of false negative, β , has also been recognized by IUPAC
784 [47].

785

786 In multivariate or multi-way analysis, the decision limit and the capability of detection can be
787 calculated from slope, intercept and s_{yx} of the regression “ c_{calc} vs. c_{true} ” [48,49], using the
788 following equations:

789

$$790 \quad CC\alpha = \frac{t_{\alpha, N-2} w_{x_0} \hat{\sigma}}{\hat{b}} \quad (8)$$

791

$$792 \quad CC\beta = \frac{\Delta(\alpha, \beta) w_{x_0} \hat{\sigma}}{\hat{b}} \quad (9)$$

793

794 where $\Delta(\alpha, \beta)$ is the non-central parameter of a non-central Student's t -distribution related to
795 the probabilities α and β , w_{x_0} is a parameter related to the distribution of the matrix-matched
796 standards on the x -axis, and $\hat{\sigma}$ and \hat{b} are the standard deviation of regression and the slope of
797 the regression " c_{calc} vs. c_{true} " respectively.
798

799 The decision limit and the detection capability were calculated for each triazine with
800 probabilities of false positive, α , and false negative, β , equal to 0.05, and considering only
801 one replicate. Table 8, column 11, shows the values of $CC\alpha$ obtained, which ranged from 0.51
802 to $1.15 \mu\text{g kg}^{-1}$. Decision limits did not differ very much for the different triazines, neither
803 between standards nor for matrix-matched standards. The same can be concluded for
804 detection capability, column 12, which ranged from 0.99 to $2.21 \mu\text{g kg}^{-1}$. In addition, the
805 decision limit values are close to the detection capability values both in standards and in
806 matrix-matched standards. Besides, all the values are far below the MRLs established for
807 these triazines.
808

809 However since these triazines have MRLs, it is mandatory to calculate these figures of merit
810 at the MRLs too. The MRL for SZ, ST, AT, PT and TT is $10 \mu\text{g kg}^{-1}$, so the calculation of the
811 decision limit and detection capability of these triazines can be made from the regression
812 models " c_{calc} vs. c_{true} " previously built provided that the calibration range includes the MRL
813 (in the case of ST and AT in the matrix-matched standards, new regression models were built
814 in the ranges from 1 to $20 \mu\text{g L}^{-1}$ and from 1 to $15 \mu\text{g L}^{-1}$ respectively; whose parameters are
815 shown in Table 8). The values of $CC\alpha$ and $CC\beta$ at the MRL for these five triazines are
816 detailed in Table 8. For SZ, for example, a value of $CC\beta$ of $11.51 \mu\text{g kg}^{-1}$ was obtained,
817 which means that the analytical procedure is capable to distinguish $11.51 \mu\text{g kg}^{-1}$ from $10 \mu\text{g}$
818 kg^{-1} with probabilities of false non-compliance and false compliance equal to 0.05. $CC\alpha$
819 values are very close to the MRL, and $CC\beta$ values are very close to $CC\alpha$ values for each
820 triazine, both in standards and in matrix-matched standards (see last two columns in table 8).
821

822 The MRLs of AZ and TZ are 50 and $100 \mu\text{g kg}^{-1}$, they are far above the range of
823 concentrations of the previous regression models so these cannot be used to estimate these
824 figures of merit. For that reason, in order to estimate $CC\alpha$ and $CC\beta$ at the MRL the 7 high
825 level concentration standards of Section 2.4.1, which contained $75 \mu\text{g L}^{-1}$ of the internal
826 standard plus 20, 30, 40, 50, 60, 70 and $80 \mu\text{g L}^{-1}$ of AZ and 70, 80, 90, 100, 110, 120 and
827 $130 \mu\text{g L}^{-1}$ of TZ, respectively, were introduced into the chromatographic system. The
828 GC/MS signals were acquired together with the signals of five replicates of the standard
829 which contained $50 \mu\text{g L}^{-1}$ of AZ and $100 \mu\text{g L}^{-1}$ of TZ, and the 6 samples corresponding to
830 the fortified orange samples in Section 2.4.3 for high level concentration analysis.
831

832 After baseline correction, these signals were arranged in a data tensor with dimensions
833 $I \times 5 \times 18$ for each chromatographic peak. I refers to the number of scans and was 26, 26 and 27
834 for AZ, TZ and PZ respectively. The second dimension refers to the number of ions acquired
835 for each compound; those specified in Section 2.4. And the third dimension is the number of
836 samples of the data tensor: the first 7 samples corresponded to the 7 standards, samples 8 to
837 12 corresponded to the replicates of one of these standards, and samples 13 to 18 to the 6
838 fortified orange samples (the first sample was a blank sample and the next 5 were fortified
839 orange samples). These last six samples provided information about the matrix-matched
840 enhancement to the system for a more reliable PARAFAC2 decomposition.
841

842 Fig. 1c shows the TIC acquired for one of these last samples, where the magnitude of the
843 coeluent's peaks with respect to those of triazines is far lower than in Fig. 1(b). In fact, only
844 one factor was necessary in the three PARAFAC2 decompositions carried out (the models
845 built for AZ, TZ and PZ explained a 99.05, 99.63 and 99.17% of variance). This meant that, at
846 these concentration levels, the interferences of coeluent on the signals of triazines were
847 unimportant, neither for AZ and TZ nor for the internal standard.

848

849 The relative ion abundances of the diagnostic ions were in all the cases within the permitted
850 tolerance intervals and also the relative retention time corresponded to that of a reference
851 sample ($50 \mu\text{g L}^{-1}$ of AZ, $100 \mu\text{g L}^{-1}$ of TZ and $75 \mu\text{g L}^{-1}$ of PZ) with a tolerance of $\pm 0.5 \%$,
852 therefore mass spectral and chromatographic profiles of the unique factor were unequivocally
853 identified. Next, the loadings of the sample profile were standardized by dividing each
854 loading by the corresponding loading of PZ, and a calibration line between them and the
855 concentration of each triazine was built. And finally, a regression model " c_{calc} vs. c_{true} " was
856 fit; the corresponding parameters are summarized in Table 8.

857

858 From these last models, like in the previous case, the decision limit (for a probability of false
859 non-compliance, α , equal to 0.05) and the capability of detection (for probabilities of false
860 noncompliance, α , and false compliance, β , equal to 0.05) at the MRL were calculated. The
861 obtained values for AZ and TZ are shown in Table 8. Also in this case, $CC\alpha$ values are very
862 close to the MRLs, and $CC\beta$ values to $CC\alpha$ values for both AZ and TZ. Although these
863 figures of merit were not available in matrix-matched standards, the results obtained for the
864 rest of triazines indicated that values close to those calculated for standards would be obtained
865 also in matrix-matched standards.

866

867

868 5. Conclusions

869

870 The use of D-optimal designs in the optimizations steps have meant an important saving in
871 the optimization cost, since the number of experiments were reduced from 384 to 16 in the
872 first case and from 96 to 10 in the second one.

873

874 The QuEChERS procedure used caused relatively dirty extracts, which interfere in the
875 quantification and in the unequivocal identification according to legislation in force for the
876 analysis of triazines in oranges. Nevertheless, the problems with co-eluting interferences and
877 with shifts in the retention time have been solved taking into account the three-way structure
878 of the GC/MS data and using the PARAFAC2 decomposition (since it is less restrictive than
879 PARAFAC decomposition).

880

881 Suitable validation results have been obtained for the analytical procedure proposed. No MLR
882 violations have been found in the commercial oranges analysed.

883

884

885 6. Acknowledgements

886

887 The authors thank the financial support provided by projects of Ministerio de Economía y
888 Competitividad (CTQ2011-26022) and Junta de Castilla y León (BU108A11-2).

889

890

891 7. References

- [1] Commission Regulation (EC) No. 149/2008 of 29 January 2008 amending Regulation (EC) No. 396/2005 of the European Parliament and of the Council by establishing Annexes II, III and IV setting maximum residue levels for products covered by Annex I thereto. Official Journal of the European Union. L58, 1.3.2008, p. 1.
- [2] D. Barceló, A.R. Fernández-Alba, C.L. Wilson (Eds.), Chromatographic-mass spectrometric food analysis for trace determination of pesticide residues, in Wilson & Wilsoo's comprehensive analytical chemistry, Vol. 43, Elsevier, Amsterdam, 2005.
- [3] M. Anastassiades, S.J. Lehotay, D. Stajnbaher, F.J. Schenck, J. AOAC Int. 86 (2003) 412A.
- [4] P. Payá, M. Anastassiades, D. Mack, I. Sigalova, B. Tasdelen, J. Oliva, A. Barba, Anal. Bioanal. Chem. 389 (2007) 1697.
- [5] S.J. Lehotay, K.A. Son, H. Kwon, U. Koesukwiwat, W. Fu, K. Mastovska, E. Hoh, N. Leepipatpiboon, J. Chromatog. A 1217 (2010) 2548.
- [6] A. Wilkowska, M. Biziuk, Food Chem. 125 (2011) 803.
- [7] G. Liu, L. Rong, B. Guo, M. Zhang, S. Li, Q. Wu, J. Chen, B. Chen, S. Yao, J. Chromatog. A 1218 (2011) 1429.
- [8] R. Pereira Lopesa, R. Cazorla Reyes, R. Romero-González, J.L. Martínez Vidal, A. Garrido Frenich, J. Chromatog. B 895-896 (2012) 39.
- [9] R. Pérez-Burgos, E.M. Grzelak, G. Gokce, J. Saurina, J. Barbosa, D. Barrón, J. Chromatog. B 899 (2012) 57.
- [10] B. Kinsella, S.J. Lehotay, K. Mastovska, A.R. Lightfield, A. Furey, M. Danaher, Anal. Chim. Acta 637 (2009) 196.
- [11] C. García Pinto, M.E. Fernández Laespada, S. Herrero Martín, A.M. Casas Ferreira, J.L. Pérez Pavón, B. Moreno Cordero, Talanta 81 (2010) 385.
- [12] F. Plössl, M. Giera, F. Bracher, J. Chromatogr. A 1135 (2006) 19.
- [13] Document N° SANCO/12495/2011, Method Validation and Quality Control Procedures for Pesticide Residues Analysis in Food and Feed. (2011) EU, Brussels.
- [14] L. Rubio, L.A. Sarabia, A. Herrero, M.C. Ortiz, Anal. Bioanal. Chem. 403-4 (2012) 1618.
- [15] M.C. Ortiz, L.A. Sarabia, J. Chromatogr. A 1158 (2007) 94.
- [16] J.M. Amigo, T. Skov, J. Coello, S. MasPOCH, R. Bro, Trends Anal. Chem. 27 (2008) 714.
- [17] B.D. Real, M.C. Ortiz, L.A. Sarabia, J. Chromatogr. B 910 (2012) 122.
- [18] J.M. Amigo, M.J. Popielarz, R.M. Callejón, M.L. Morales, A.M. Troncoso, M.A. Petersen, T.B. Toldam-Andersen, J. Chromatogr. A 1217 (2010) 4422.
- [19] R. Morales, M.C. Ortiz, L.A. Sarabia, Anal. Chim. Acta 754 (2012) 20.
- [20] H.G.J. Mol, M. Althuizen, H.G. Janssen, C.A. Cramers, J. High Resol. Chromatogr. 19 (1996) 69.
- [21] P.F. De Aguiar, B. Bourguignon, M.S. Khost, D.L. Massart, R. Phan-Tan-Luu, Chemom. Intell. Lab. Systems 30 (1995) 199.

- [22] G.A. Lewis, D. Mathieu, R. Phan-Tan-Luu, *Pharmaceutical and Experimental Designs*, Marcel Dekker, New York, 1999.
- [23] D. Arroyo, M.C. Ortiz, L.A. Sarabia, *J. Chromatog. A* 1218 (2011) 4487.
- [24] N. Rodríguez, M.C. Ortiz, L.A. Sarabia, A. Herrero, *Anal. Chim. Acta* 657 (2010) 136.
- [25] N. Li, L. Nian, R. Zhang, S. Wu, R. Ren, Y. Wang, H. Zhang, A. Yu, *Talanta* 105 (2013) 219.
- [26] Commission Regulation (EU) No. 310/2011 of 28 March 2011 amending Annexes II and III to Regulation (EC) No 396/2005 of the European Parliament and of the Council as regards maximum residue levels for aldicarb, bromopropylate, chlorfenvinphos, endosulfan, EPTC, ethion, fenthion, fomesafen, methabenzthiazuron, methidathion, simazine, tetradifon and triforine in or on certain products. *Official Journal of the European Union* L86, 1.4.2011, p. 1.
- [27] Regulation (EC) No. 396/2005 of the European Parliament and of the Council of 23 February 2005 on maximum residue levels of pesticides in or on food and feed of plant and animal origin and amending Council Directive 91/414/EEC. *Official Journal of the European Union* L70, 16.3.2005, p. 1.
- [28] B.M. Wise, N.B. Gallagher, R. Bro, J.M. Shaver, W. Windig, R.S. Koch, *PLS Toolbox 5.8.2. Eigenvector Research Inc., Manson, WA*, 2010.
- [29] *STATGRAPHICS Centurion XVI* (version 16.1.11), StatPoint Technologies, Inc. Herndon, VA, 2010.
- [30] P.J. Rousseeuw, A.M. Leroy, *Robust Regression and Outliers Detection*, John Wiley and Sons, New Jersey, 2001.
- [31] L.A. Sarabia, M.C. Ortiz, *TrAC: Trends Anal. Chem.* 13 (1994) 1.
- [32] D. Mathieu, J. Nony, R. Phan-Thao-Lu, *NemrodW* (Version 2007_03), L.P.R.A.I. Marseille, France, 2007.
- [33] C.A. Anderson, R. Bro, *Chemom. Intell. Lab. Syst.* 52 (2000) 1.
- [34] R. Bro, *Chemom. Intell. Lab. Syst.* 46 (1999) 133.
- [35] H.A.L. Kiers, J.M.F. Ten Berge, R. Bro, *J. Chemom.* 13 (1999) 275.
- [36] R. Bro, C.A. Andersson, H.A.L. Kiers, *J. Chemom.* 13 (1999) 295.
- [37] D. Arroyo, M.C. Ortiz, L.A. Sarabia, *J. Chromatogr. A* 1157 (2007) 358.
- [38] B.D. Real, M.C. Ortiz, L.A. Sarabia, *Talanta* 71 (2007) 1599.
- [39] I. García, L. Sarabia, M.C. Ortiz, J.M. Aldama, *Anal. Chim. Acta* 526 (2004) 139.
- [40] M.C. Ortiz, L.A. Sarabia, I. García, D. Giménez, E. Meléndez, *Anal. Chim. Acta* 559 (2006) 124.
- [41] R. Bro, H.A.L. Kiers, *J. Chemometrics* 17 (2003) 274.
- [42] D.R. Erney, A.M. Gillespie, D.M. Gilvydis, C.F. Poole, *J. Chromatogr.* 638 (1993) 57.
- [43] C.F. Poole, *J. Chromatogr. A* 1158 (2007) 241.
- [44] CEN standard method EN 15662:2008. Foods of plant origin. Determination of pesticide residues using GC-MS and/or LC-MS/MS following acetonitrile extraction/partitioning and clean-up by dispersive SPE – QuEChERS method.

- [45] M.B. Sanz, L.A. Sarabia, A. Herrero, M.C. Ortiz, *Talanta* 56 (2002) 1039.
- [46] International Organization for Standardization, ISO 11843, Capability of detection. Part 1: Terms and definitions, 1997; and Part 2: Methodology in the linear calibration case, 2000. Geneva, Switzerland.
- [47] J. Inczédy, T. Lengyel, A.M. Ure, A. Gelencsér, A. Hulanicki, *Compendium of Analytical Nomenclature IUPAC*, Pot City Press Inc, Baltimore, 3rd ed., 2000.
- [48] M.C. Ortiz, L.A. Sarabia, A. Herrero, M.S. Sánchez, M.B. Sanz, M.E. Rueda, D. Giménez, M.E. Meléndez, *Chemom. Intell. Lab. Syst.* 69 (2003) 21.
- [49] M.C. Ortiz, M.S. Sánchez, L.A. Sarabia, in: S.D. Brown, R. Tauler, B. Walczak (Eds.), *Comprehensive Chemometrics*, Elsevier, Amsterdam, 2009, pp. 127.

Table 1 Factors and experimental domain for optimization of the PTV injection.

Factors	Codified variable	Level A	Level B	Level C
Vent flow (mL min ⁻¹)	x ₁	100	150	–
Inlet P (psi)	x ₂	8	9	–
Vent time (min)	x ₃	0.3	0.45	–
Initial temperature (°C)	x ₄	40	50	–
PTV initial time (min)	x ₅	0.6	0.5	–
PTV rate (°C s ⁻¹)	x ₆	10	5	–
End temperature (°C)	x ₇	280	320	–
Injection speed (μL s ⁻¹)	x ₈	0.85	1.7	3.4

Table 2 Factors and experimental domain for optimization of the QuEChERS procedure.

Factors	Abbreviated name	Codified variable	Level A	Level B	Level C
Time of vortex mixing 1 (min)	<i>mix_t1</i>	x ₁	1	2	–
Time of centrifugation (min)	<i>centr_t1</i>	x ₂	1	5	–
Acetonitrile extract volume (mL)	<i>volume</i>	x ₃	1	1.2	–
Time of vortex mixing 2 (s)	<i>mix_t2</i>	x ₄	30	60	–
Evaporation temperature (°C)	<i>evap_T</i>	x ₅	40	50	–
Percent (v/v) of acetic acid in acetonitrile (%)	<i>modifier</i>	x ₆	0	1	5

Table 3 Experimental plan and responses of the D-optimal design for the optimization of the PTV injection.

Run	Vent flow (mL min ⁻¹)	Inlet P (psi)	Vent time (min)	Initial temperature (°C)	PTV initial time (min)	PTV rate (°C s ⁻¹)	End temperature (°C)	Injection speed (μL s ⁻¹)	Responses (standardized loadings)						
									SZ	AZ	TZ	ST	AT	PT	TT
1	150	8	0.45	40	0.6	10	280	0.85	0.45	0.76	1.61	0.79	0.79	1.24	1.08
2	150	8	0.30	50	0.6	5	280	0.85	0.43	0.74	1.60	0.75	0.76	1.19	1.07
3	100	9	0.45	40	0.5	5	280	0.85	0.44	0.74	1.60	0.78	0.79	1.26	1.11
4	100	9	0.45	50	0.6	10	320	0.85	0.48	0.77	1.63	0.95	0.91	1.42	1.24
5	100	8	0.30	50	0.5	10	320	0.85	0.46	0.75	1.63	0.87	0.84	1.32	1.13
6	100	9	0.30	40	0.5	5	320	0.85	0.47	0.77	1.60	0.90	0.87	1.34	1.17
7	100	8	0.45	50	0.5	10	280	1.70	0.46	0.77	1.62	0.80	0.79	1.25	1.11
8	100	8	0.30	40	0.6	5	280	1.70	0.46	0.76	1.60	0.85	0.82	1.28	1.13
9	150	9	0.45	50	0.5	5	280	1.70	0.44	0.75	1.61	0.79	0.79	1.27	1.10
10	150	9	0.45	40	0.6	10	320	1.70	0.50	0.79	1.62	0.93	0.87	1.35	1.19
11	150	8	0.30	40	0.5	10	320	1.70	0.50	0.78	1.65	1.00	0.93	1.44	1.27
12	150	8	0.30	40	0.5	10	320	1.70	0.47	0.79	1.59	0.89	0.85	1.29	1.12
13	150	8	0.30	40	0.5	10	320	1.70	0.49	0.78	1.61	0.94	0.90	1.35	1.19
14	150	8	0.30	40	0.5	10	320	1.70	0.49	0.78	1.64	0.97	0.93	1.40	1.25
15	150	9	0.30	50	0.6	5	320	1.70	0.48	0.77	1.70	0.90	0.87	1.32	1.18
16	100	9	0.30	50	0.6	10	280	3.40	0.46	0.77	1.63	0.86	0.83	1.31	1.16
17	150	9	0.30	40	0.5	10	280	3.40	0.47	0.77	1.63	0.84	0.82	1.30	1.13
18	100	8	0.45	40	0.6	5	320	3.40	0.50	0.79	1.63	0.95	0.89	1.36	1.19
19	150	8	0.45	50	0.5	5	320	3.40	0.48	0.77	1.62	0.93	0.89	1.35	1.19

Table 4 Detected ions (the most intense ones are in bold), relative abundance and tolerance intervals for the reference sample of 10 $\mu\text{g L}^{-1}$ of the eight triazines, and relative abundances calculated with the loadings of the spectral profiles of the PARAFAC2 models built from both the samples of the D-optimal design of the QuEChERS procedure optimization and from the quantitative determination samples.

Analyte	m/z	Reference sample		D-optimal design samples (QuEChERS)	Quantitative determination samples
		Relative abundance (%)	Tolerance interval (%)	Relative abundance (%)	Relative abundance (%)
SZ	158	23.74	(20.18-27.30)	27.25	29.33
	173	44.55	(37.87- 51.24)	42.88	38.53
	186	71.89	(64.70-79.08)	81.85	66.79
	201	100.00	–	100.00	100.00
	203	31.95	(27.16-36.74)	32.82	32.45
AZ	173	24.55	(20.86-28.23)	25.39	21.09
	200	100.00	–	100.00	100.00
	202	33.86	(28.78-38.94)	35.33	35.28
	215	57.57	(51.82-63.33)	51.66	53.42
	217	19.07	(16.21-21.93)	19.85	14.51
PZ	172	57.57	(51.81-63.33)	52.75	52.07
	187	26.51	(22.53-30.49)	50.87	29.12
	214	100	–	100.00	100.00
	229	60.63	(54.56-66.69)	56.45	50.44
	231	19.36	(16.45-22.26)	17.54	19.25
TZ	173	35.50	(30.17-40.82)	36.22	33.84
	214	100	–	100.00	100.00
	216	31.03	(26.38-35.69)	35.29	33.80
	229	24.24	(20.60-27.87)	23.12	24.26
	231	7.40	(3.70-11.09)	6.96	6.96
ST	155	33.88	(28.80-38.96)	30.19	24.14
	170	31.92	(27.13-36.71)	34.27	32.86
	198	18.44	(14.75-22.13)	17.64	15.84
	213	100.00	–	100.00	100.00
	215	4.74	(2.37-7.11)	5.04	5.08
AT	170	24.21	(20.58-27.84)	26.95	32.60
	185	18.04	(15.34-20.75)	19.61	20.51
	212	59.45	(53.50-65.39)	56.47	60.84
	227	100.00	–	100.00	100.00
	229	5.53	(2.77-8.30)	5.87	6.08
	184	84.13	(75.71-92.54)	91.02	87.39
	199	26.22	(22.29-30.16)	26.56	24.29

PT	226	61.88	(55.69-68.06)	65.90	63.44
	241	100.00	-	100.00	100.00
	243	5.12	(2.56-7.68)	6.91	5.60
	170	57.20	(51.48-62.92)	62.16	60.00
	185	70.64	(63.58-77.71)	69.97	68.62
TT	226	100.00	-	100.00	100.00
	241	53.46	(48.11-58.80)	52.38	55.11
	243	4.30	(2.15-6.45)	1.57	4.91

Table 5 Experimental plan and responses of the D-optimal design for the optimization of the QuEChERS procedure.

Run	<i>mix_t1</i> (min)	<i>centr_t1</i> (min)	<i>volume</i> (mL)	<i>mix_t2</i> (s)	<i>evap_T</i> (°C)	<i>modifier</i> (%)	Responses (standardized loadings)						
							SZ	AZ	TZ	ST	AT	PT	TT
1	2	1	1.2	30	40	0	0.28	0.70	0.72	0.98	0.65	0.73	0.88
2	1	1	1.0	60	40	0	0.34	0.70	0.96	0.89	0.97	0.89	0.72
3	1	5	1.0	30	50	0	0.35	0.45	0.69	0.67	0.75	0.76	0.61
4	1	5	1.0	30	50	0	0.26	0.34	0.53	0.53	0.59	0.59	0.38
5	1	5	1.0	30	50	0	0.32	0.52	0.63	0.67	0.69	0.68	0.31
6	1	5	1.0	30	50	0	0.38	0.84	0.93	1.11	0.91	1.06	0.78
7	2	5	1.2	60	50	0	0.47	0.88	1.20	1.19	1.21	1.21	0.45
8	1	5	1.2	30	40	1	0.44	0.97	1.08	1.20	0.99	1.10	0.56
9	2	5	1.0	60	40	1	0.49	1.01	1.27	1.22	1.14	1.24	0.86
10	2	1	1.0	30	50	1	0.44	0.81	1.04	1.02	1.00	1.04	0.50
11	1	1	1.2	60	50	1	0.50	0.75	0.96	1.02	1.17	1.10	0.83
12	2	5	1.0	30	40	5	0.50	0.81	1.18	1.23	1.07	1.25	1.08
13	1	1	1.2	60	50	5	–	–	–	–	–	–	–

Table 6 Characteristics of PARAFAC2 models: number of factors, data tensor size, explained variance, CORCONDIA index and verification of compliance with the maximum permitted tolerances for the identification of each triazine (number of diagnostic ions that are within the tolerance intervals calculated for each sample using a reference sample containing $10 \mu\text{g L}^{-1}$ of SZ, AZ, PZ, TZ, ST, AT, PT, and TT).

Study	Analyte	Factors	Data tensor dimension ^a I × J × K	Explained variance (%)	CORCONDIA index (%)	Verified compliance
D-optimal design to optimize QuEChERS procedure	SZ	3	31 × 5 × 25	96.53	85.91	3
	AZ	3	22 × 5 × 25	99.28	83.61	3
	PZ	2	28 × 5 × 24	97.33	100.00	3
	TZ	3	26 × 5 × 25	99.43	86.56	4
	ST	2	15 × 5 × 25	99.32	100.00	3
	AT	1	10 × 5 × 25	98.96	– ^b	4
	PT	2	6 × 5 × 25	98.3	100	4
	TT	3	15 × 5 × 25	97.96	83.30	3
Quantitative and qualitative determination of triazines in oranges	SZ	3	28 × 5 × 44	96.5	99.58	3
	AZ	4	21 × 5 × 44	98.76	97.86	3
	PZ	2	28 × 5 × 44	98.71	100.00	3
	TZ	3	26 × 5 × 43	98.67	90.78	4
	ST	3	15 × 5 × 43	99.17	99.67	3
	AT	1	10 × 5 × 43	97.4	– ^b	3
	PT	3	6 × 5 × 43	99.97	98.8	4
	TT	3	15 × 5 × 25	98.5	99.07	4

^(a) I refers to the number of scans, J refers to the number of ions, and K refers to the number of samples

^(b) It was not possible to calculate the CORCONDIA index as there was only one factor

Table 7 Parameters of the calibration lines “standardized loadings vs. true concentration” (intercept, slope, standard deviation of regression (s_{yx}) and correlation coefficient) and calculated concentrations, c_{calc} , for the four commercial orange samples (95% confidence intervals are in brackets). The calculated concentrations are the mean of the four replicates analysed for each sample.

		SZ	AZ	TZ	ST	AT	PT	TT
Regression parameters	Intercept	0.0459	-0.0037	0.0793	-0.0256	0.0198	0.1225	0.0685
	Slope	0.0460	0.0823	0.1116	0.1122	0.0801	0.1387	0.1181
	s_{yx}	0.0172	0.0289	0.0259	0.0280	0.0249	0.0340	0.0596
	Correlation coefficient	0.9989	0.9987	0.9939	0.9928	0.9911	0.9989	0.9978
C_{calc} ($\mu\text{g L}^{-1}$)	P1	0.62 (-0.01,1.23)	1.14 (0.61,1.65)	0.27 (-0.34,0.76)	0.24 (-0.35,0.73)	1.16 (0.58,1.67)	-0.23 (-0.64,0.16)	-0.14 (-0.98,0.66)
	P2	0.00 (-0.65,0.62)	0.63 (0.09,1.15)	0.04 (-0.59,0.55)	0.26 (-0.32,0.75)	1.34 (0.78,1.84)	-0.34 (-0.75,0.06)	0.04 (-0.79,0.83)
	P3	0.15 (-0.50,0.76)	0.99 (0.46,1.51)	0.13 (-0.49,0.63)	0.24 (-0.35,0.73)	0.87 (0.26,1.39)	-0.6 (-1.01,-0.20)	0.19 (-0.64,0.98)
	P4	0.01 (-0.64,0.63)	1.24 (0.72,1.76)	0.01 (-0.62,0.53)	0.24 (-0.34,0.73)	1.20 (0.63,1.71)	-0.47 (-0.88,-0.07)	-0.01 (-0.85,0.78)

Table 8 Parameters of reweighted regression models “ c_{calc} vs. c_{true} ”: intercept, slope and standard deviation of regression (s_{yx}) and number of standards (n). Some figures of merit: error (mean of the absolute value of relative errors in calibration), average recovery rate, repeatability and detection limit (CC α) and capability of detection (CC β) at $x_0 = 0$ and at the maximum residue level ($x_0 = \text{MRL}$) for both standards and matrix-matched standards.

	Analyte	n	Concentration range ($\mu\text{g L}^{-1}$)	Intercept	Slope	s_{yx}	Error ^a (%)	Recovery (%)	Repeatability ($\mu\text{g L}^{-1}$)	x = 0		x = MRL	
										CC α ($\mu\text{g kg}^{-1}$)	CC β ($\mu\text{g kg}^{-1}$)	CC α ($\mu\text{g kg}^{-1}$)	CC β ($\mu\text{g kg}^{-1}$)
Matrix-matched standards	SZ	8	0 - 20	6.26×10^{-5}	1	0.3736	5.94	36.73	0.57	0.81	1.55	10.78	11.51
	AZ	10	0 - 20	7.74×10^{-5}	1	0.3532	8.35	61.13	1.15	0.72	1.39	–	–
	TZ	6	0 - 5	2.29×10^{-4}	1	0.2320	9.29	57.24	1.77	0.61	1.16	–	–
	ST	6	0 - 5	-1.79×10^{-5}	1	0.2497	10.10	73.83	1.42	0.63	1.19	–	–
		6	1 - 20	3.19×10^{-7}	1	0.7277	–	–	–	–	–	10.37	10.71
	AT	7	0 - 5	-2.85×10^{-4}	1	0.3114	10.60	83.17	1.90	0.74	1.41	–	–
		7	1 - 15	1.16×10^{-7}	1	0.4851	–	–	–	–	–	11.25	12.38
	PT	9	0 - 15	3.47×10^{-5}	1	0.2454	12.60	64.06	1.47	0.51	0.99	10.52	11.01
TT	9	0 - 15	1.69×10^{-4}	1	0.7211	5.40	81.43	1.74	1.05	2.04	11.46	12.83	
Standards	SZ	7	0 - 20	-9.09×10^{-6}	1	0.4998	3.51	–	0.37	1.15	2.21	11.08	12.08
	AZ	7	0 - 20	2.14×10^{-6}	1	0.4648	3.77	–	0.16	1.07	2.21	–	–
		7	20 - 80	1.79×10^{-5}	1	0.2384	–	–	–	–	–	50.51	50.99
	TZ	7	0 - 20	2.41×10^{-5}	1	0.3958	3.35	–	0.11	0.91	1.75	–	–
		7	70 - 130	3.57×10^{-5}	1	1.7329	–	–	–	–	–	103.7	107.2
	ST	7	0 - 20	4.81×10^{-6}	1	0.4125	4.20	–	0.57	0.95	1.82	10.9	11.72
	AT	7	0 - 20	6.95×10^{-6}	1	0.3901	7.29	–	0.57	0.9	1.73	10.85	11.63
	PT	7	0 - 20	8.56×10^{-6}	1	0.2322	2.42	–	0.31	0.53	1.03	10.5	10.97
TT	7	0 - 20	3.26×10^{-5}	1	0.2596	1.28	–	0.26	0.6	1.15	10.56	11.08	

FIGURE CAPTIONS

- Fig. 1 Total ion chromatograms (TICs) from the injection of: (a) a standard of $10 \mu\text{g L}^{-1}$ of SZ, AZ, PZ (internal standard), TZ, ST, AT, PT and TT, injected in both the injection conditions of experiments 2 and 11 (chromatogram with the largest abundances) in Table 2; (b) the extract of a orange sample fortified with $10 \mu\text{g L}^{-1}$ of SZ, AZ, TZ, ST, AT, PT and TT and extracted according to the experiment 4 in Table 3 (the extract also contained $10 \mu\text{g L}^{-1}$ of PZ); and (c) the extract of a orange sample fortified with $50 \mu\text{g L}^{-1}$ of AZ and $100 \mu\text{g L}^{-1}$ of TZ and extracted in the optimal conditions (it also contained $75 \mu\text{g L}^{-1}$ of PZ). Peak labels: 1: SZ; 2: AZ; 3: PZ; 4: TZ; 5: ST; 6: AT; 7: PT; and 8: TT.
- Fig. 2 Graphic analysis of the effects of PTV injection factors on the response (standardized loadings of PARAFAC models). Factors: 1, *Vent flow*; 2, *Inlet P*; 3, *Vent time*; 4, *Initial temperature*; 5, *PTV initial time*; 6, *PTV rate*; 7, *End temperature* and 8, *Injection speed*. The dash-dotted lines represent the confidence interval of the calculated effects at 95% confidence level; significant effects (light orange) and non-significant effects (dark blue) .
- Fig. 3 Abundances (counts) acquired for AZ for standards (blue continuous line) and orange extracts of experiments in Table 5 (orange dotted line). Diagnostic ions: (a) 173, (b) 200, (c) 202, (d) 215 and (e) 217.
- Fig. 4 Loadings of the (a) chromatographic, (b) spectral, and (c) sample modes of the PARAFAC2 model built for AZ in the optimization of the QuEChERS procedure (chromatographic loadings are scaled loadings). First factor is in blue continuous line (blue solid bars in the spectral mode and blue points in the sample mode), second factor is in green dashed line (green dashed bars in the spectral mode and green triangles in the sample mode), and third factor is in red dotted line (red pointed bars in the spectral mode and red squares in the sample mode).
- Fig. 5 (a) Loadings of the sample mode of the PARAFAC2 model built for PZ (first factor, blue points; and second factor, green triangles) and (b) standardized loadings of the sample mode calculated for AZ in the optimization of the QuEChERS procedure.
- Fig. 6 Graphic analysis of effects of factors of QuEChERS procedure on the response (standardized loadings of PARAFAC2 models). Factors: 1, *Vent flow*; 2, *Inlet P*; 3, *Vent time*; 4, *Initial temperature*; 5, *PTV initial time*; and 6, *PTV rate*. For each factor the levels chosen are in solid bars.
- Fig. 7 Loadings of the (a) chromatographic, (b) spectral, and (c) sample modes and (d) standardized loadings of the sample mode of the PARAFAC2 model built for AZ in the quantitative and qualitative determination of triazines in oranges (chromatographic loadings are scaled loadings). First factor is in blue continuous line (blue solid bars in the spectral mode and blue points in the sample mode), second factor is in green dashed line (green dashed bars in the spectral mode and

green triangles in the sample mode), third factor is in red dotted line (red pointed bars in the spectral mode and red squares in the sample mode), and fourth factor is in cyan dash-dot-dot line (cyan transparent bars in the spectral mode and cyan stars in the sample mode).

Fig. 8 Joint confidence ellipses, at 95% confidence level, for the slope and the intercept of the regression models “calculated concentration vs. true concentration” in Table 8 for (a) matrix-matched standards and (b) standards. SZ: solid blue line; AZ: red triangles; TZ: green circles; ST: black dotted line; AT: magenta dash-dot line; PT: cyan squares; and TT: yellow dashed line.

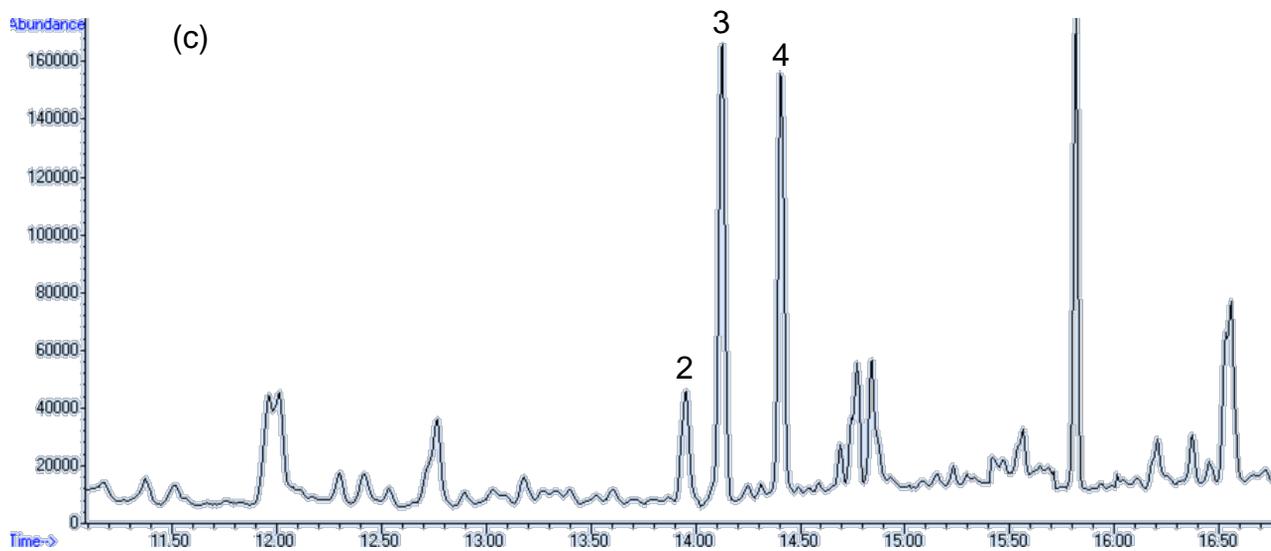
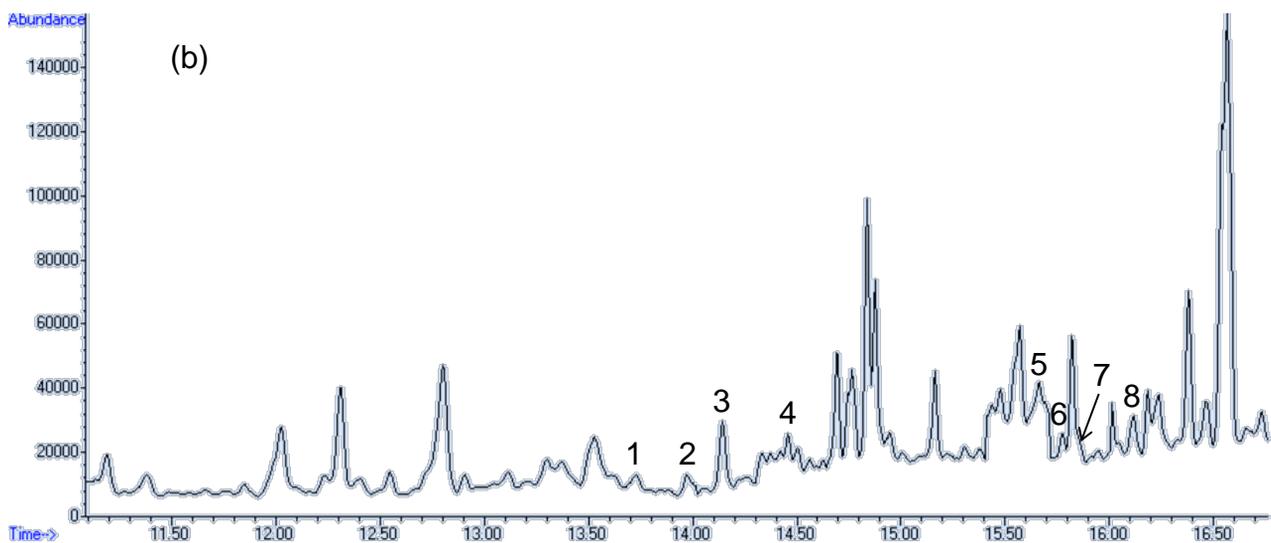
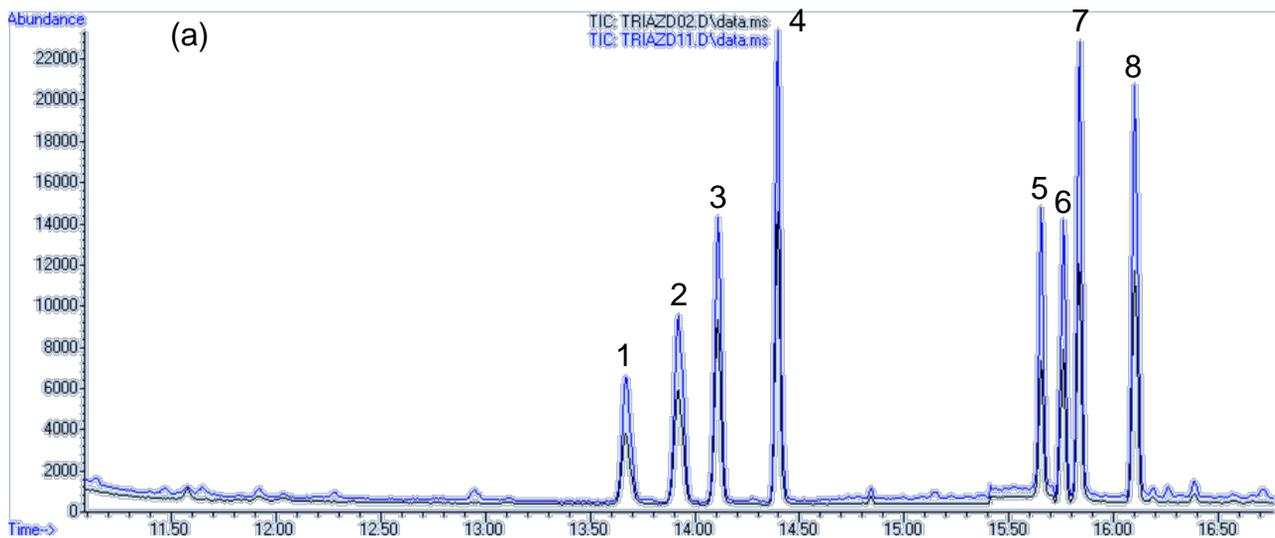


FIGURE 1

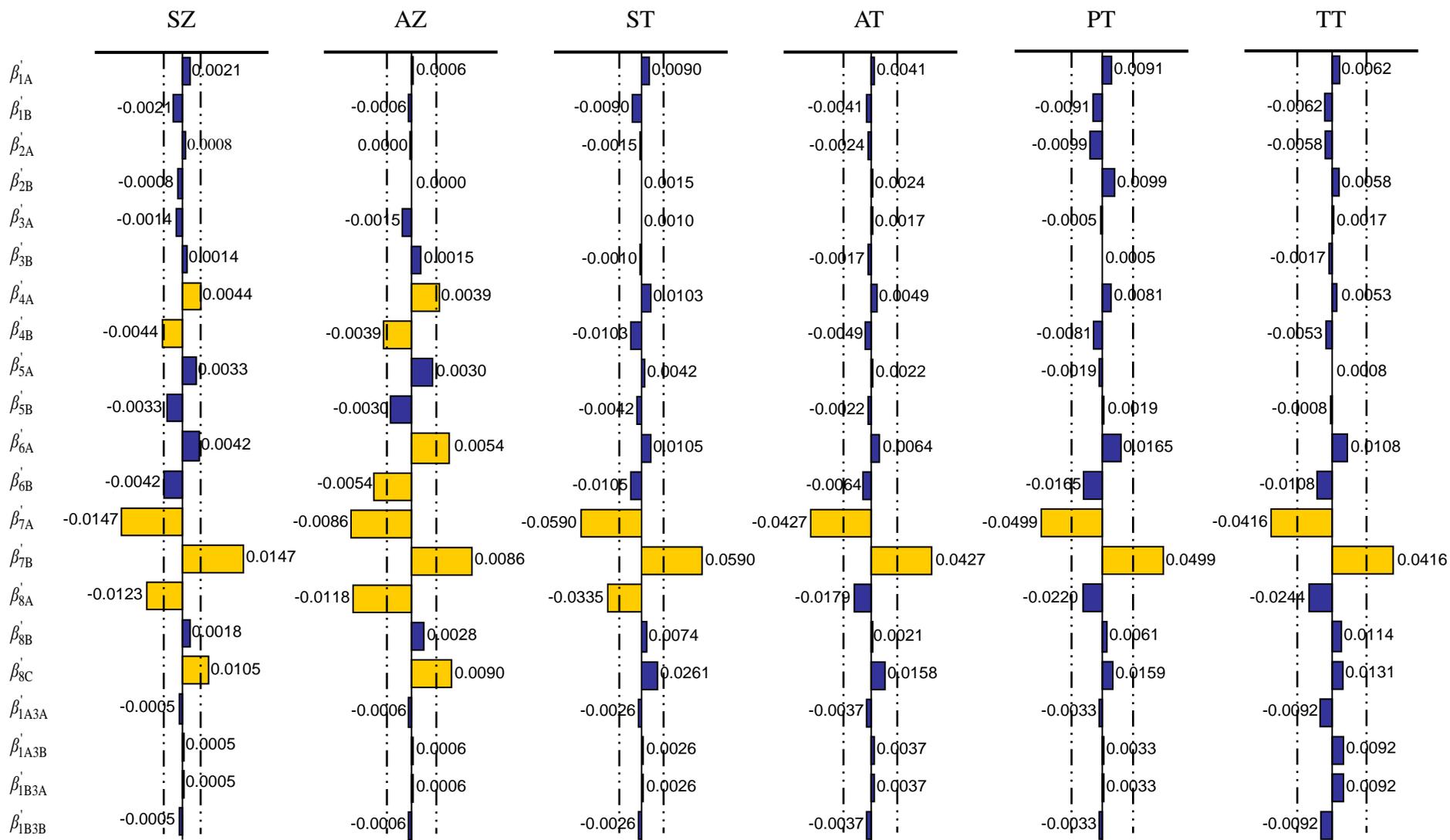


FIGURE 2

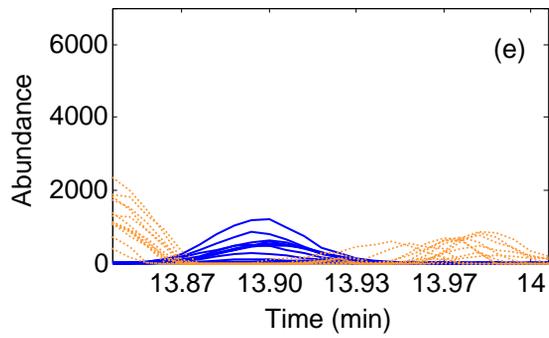
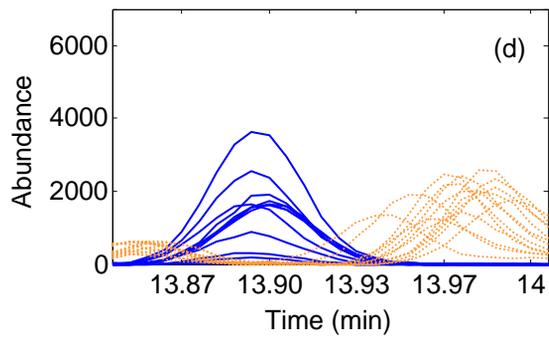
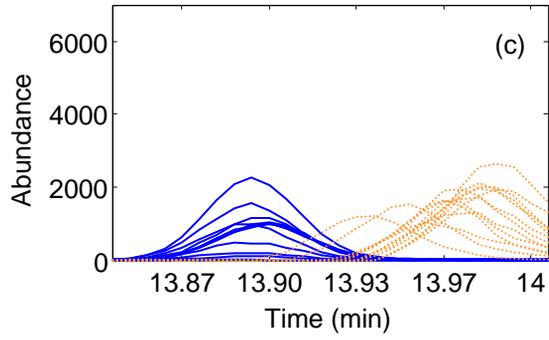
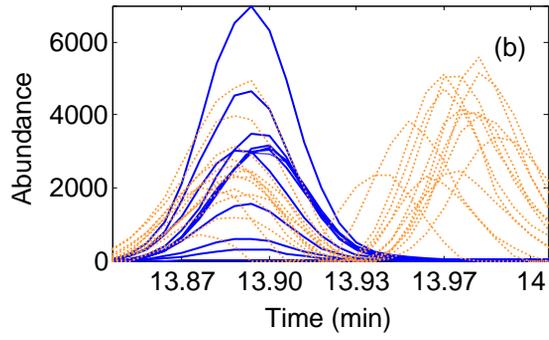
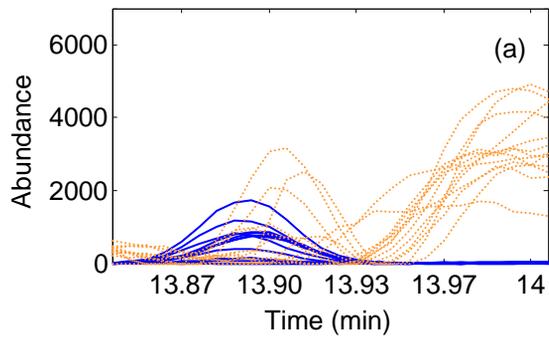


FIGURE 3

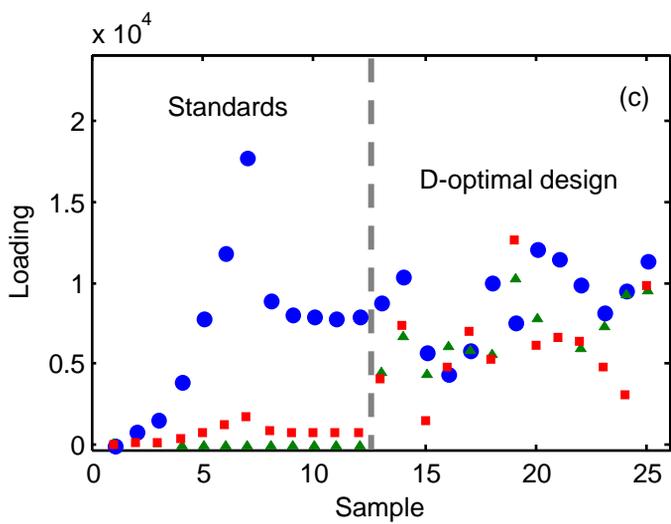
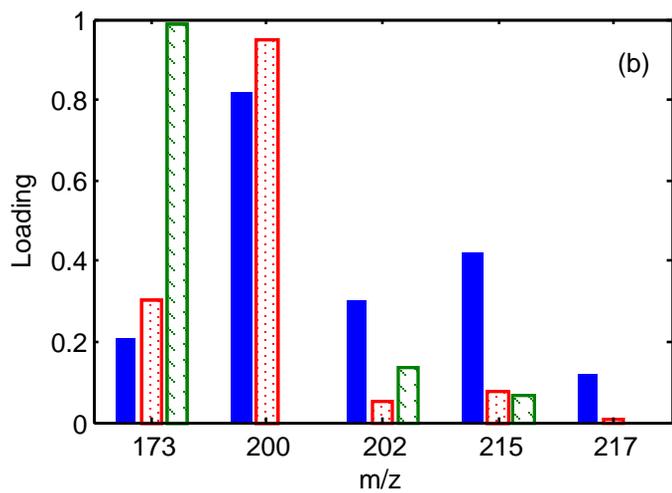
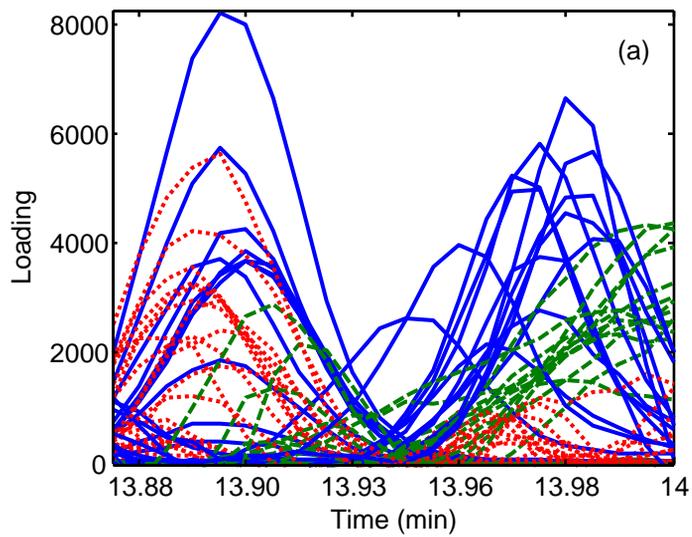


FIGURE 4

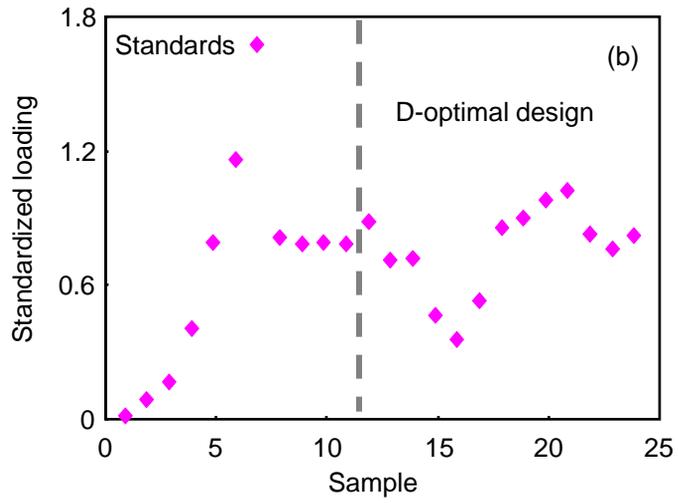
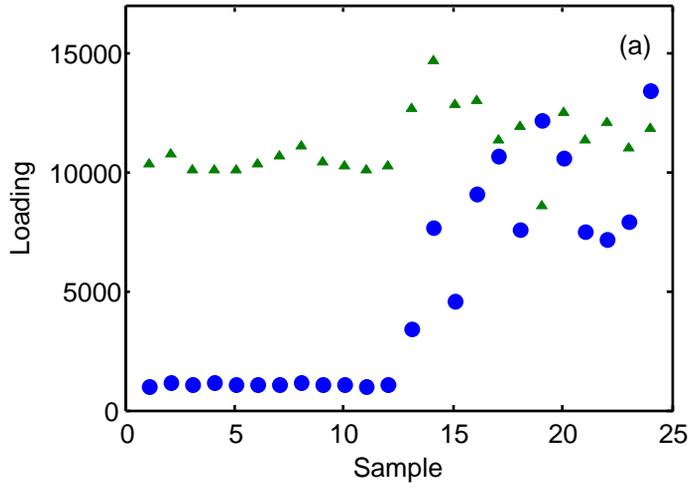


FIGURE 5

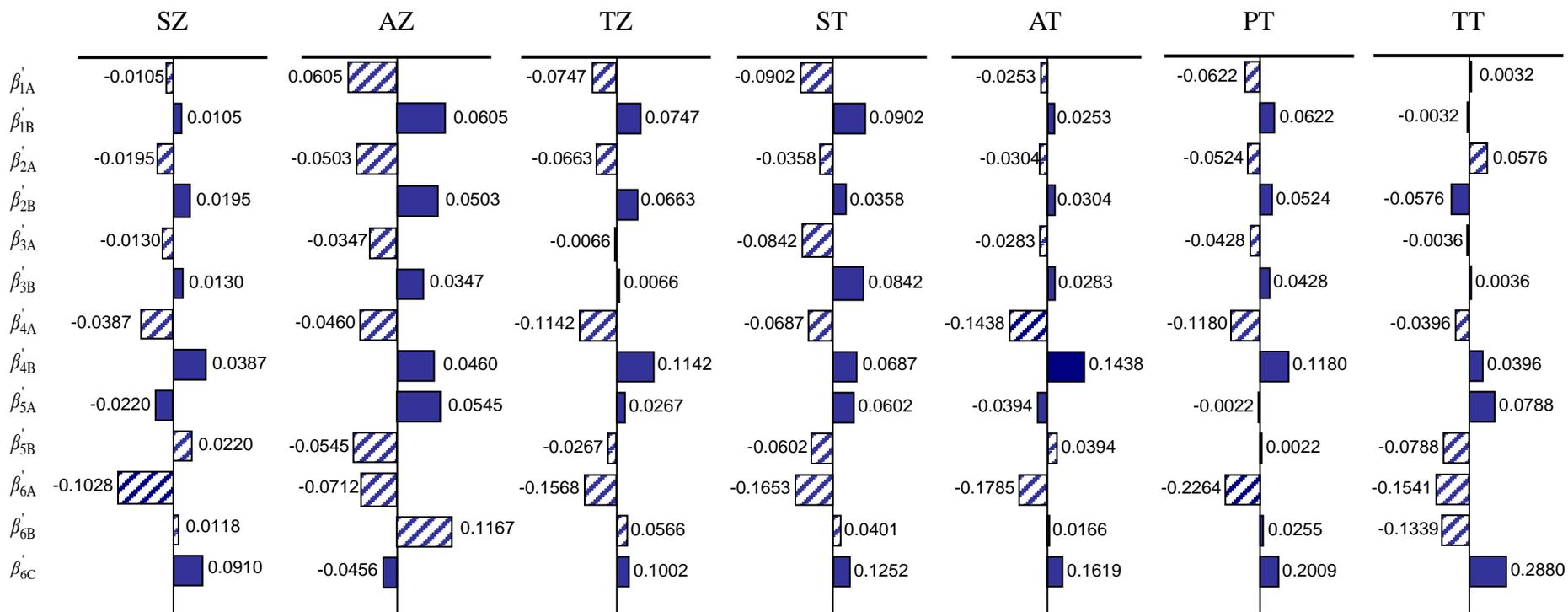


FIGURE 6

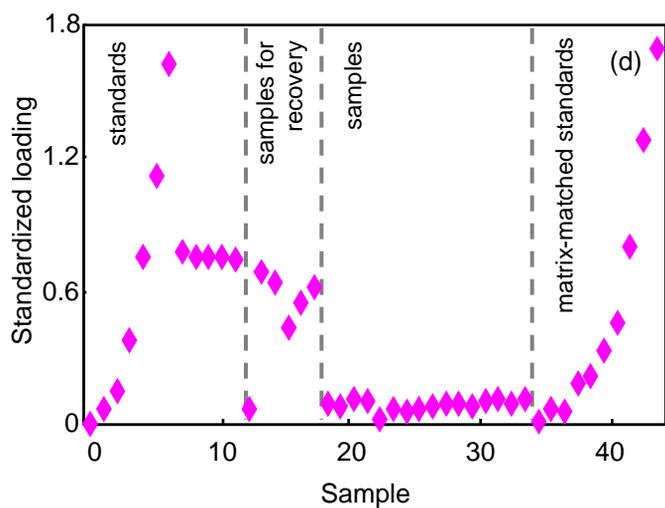
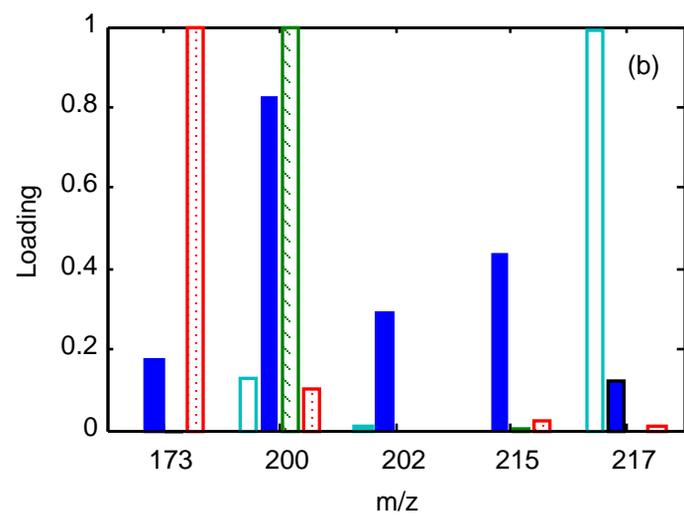
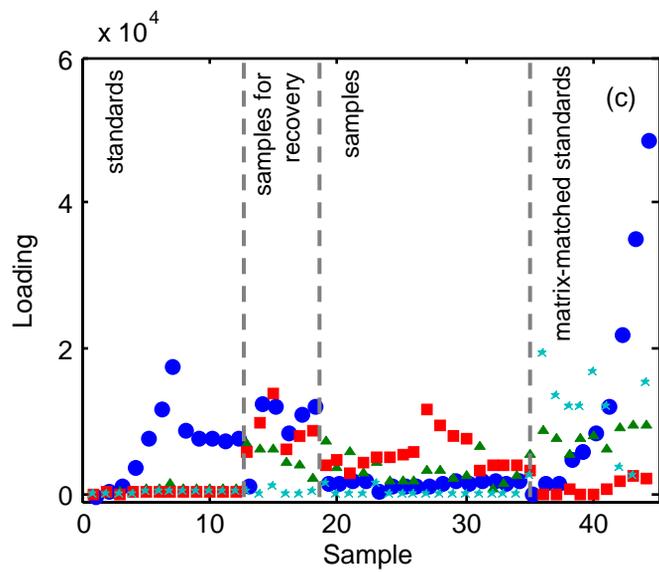
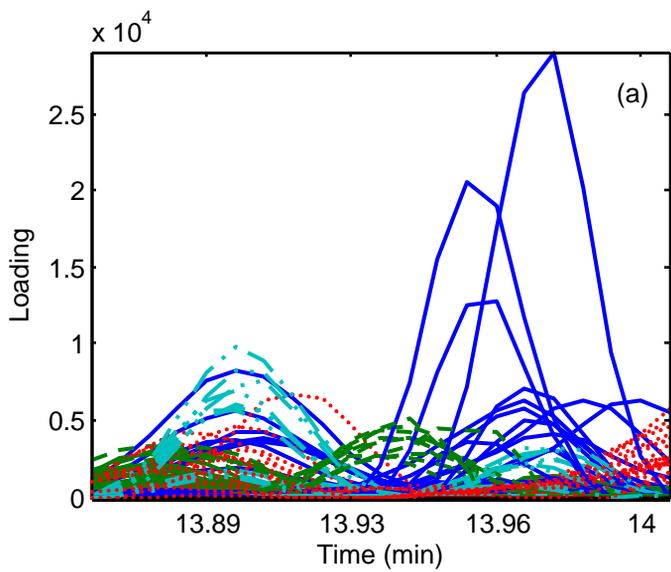


FIGURE 7

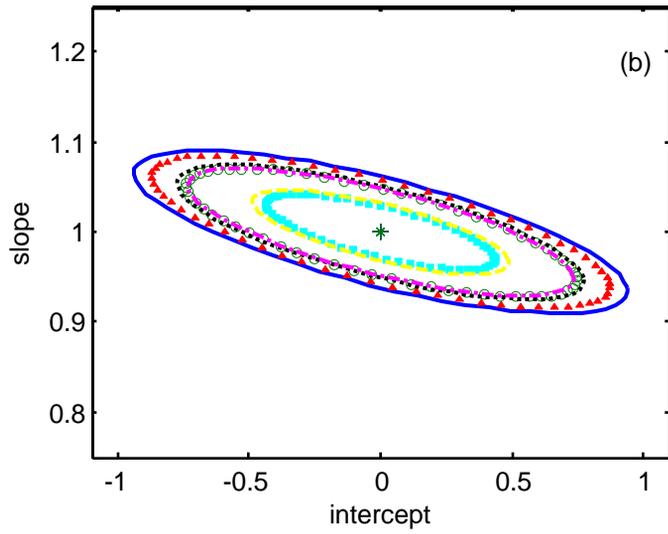
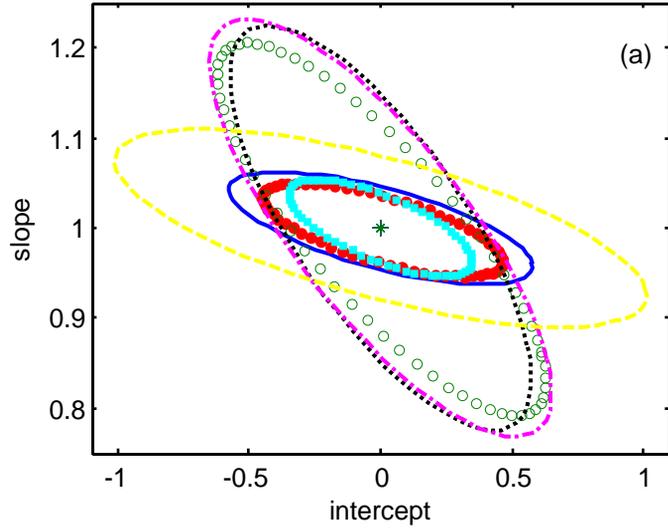


FIGURE 8