

## Advantages of data fusion. First multivariate curve resolution analysis of fused liquid chromatographic second-order data with dual diode array-fluorescent detection

Rocio B. Pellegrino Vidal, Gabriela Alejandra Ibanez, and Graciela Mónica Escandar

*Anal. Chem.*, **Just Accepted Manuscript** • DOI: 10.1021/acs.analchem.6b04720 • Publication Date (Web): 07 Feb 2017

Downloaded from <http://pubs.acs.org> on February 13, 2017

### Just Accepted

“Just Accepted” manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides “Just Accepted” as a free service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. “Just Accepted” manuscripts appear in full in PDF format accompanied by an HTML abstract. “Just Accepted” manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are accessible to all readers and citable by the Digital Object Identifier (DOI®). “Just Accepted” is an optional service offered to authors. Therefore, the “Just Accepted” Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the “Just Accepted” Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these “Just Accepted” manuscripts.

## REVISED VERSION

Advantages of data fusion. First multivariate curve resolution analysis of fused liquid chromatographic second-order data with dual diode array-fluorescent detection

Rocío Pellegrino Vidal, Gabriela A. Ibañez,\* Graciela M. Escandar\*

*Instituto de Química Rosario (CONICET-UNR), Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Suipacha 531 (2000) Rosario, Argentina, escandar@iquir-conicet.gov.ar*

**ABSTRACT**

For the first time, liquid chromatography-diode array detection (LC-DAD) and liquid-chromatography fluorescence detection (LC-FLD) second-order data, collected in a single chromatographic run, were fused and chemometrically processed for the quantitation of coeluting analytes. Two different experimental mixtures composed of fluorescent and non-fluorescent endocrine disruptors were analyzed. Adequate pretreatment of the matrices before their fusion was crucial to attain reliable results. Multivariate curve resolution-alternating least-squares (MCR-ALS) was applied to LC-DAD, LC-FLD and fused LC-DAD-FLD data. Although different degrees of improvement are observed when comparing the fused matrix results in relation to those obtained using a single detector, clear benefits of data fusion are demonstrated through: (1) the obtained limits of detection in the ranges 2.1–24 ng mL<sup>-1</sup> and 0.9–6.3 ng mL<sup>-1</sup> for the two evaluated systems and (2) the low relative prediction errors, below 7% in all cases, indicating good recoveries and precision. The feasibility of fusing data and its advantages in the analysis of real samples was successfully assessed through the study of spiked tap, underground and river water samples.

## INTRODUCTION

The coupling of liquid chromatography/multivariate detection with second-order chemometric calibration has proved to be a powerful analytical tool, allowing the successful quantification of strongly coeluting compounds, even in the presence of potential interferences.<sup>1-4</sup> Recently, chemometric processing of liquid chromatography (LC) data with dual diode array and fluorescence detection (DAD and FLD) allowed the validation of efficient and green methods for the analysis of sex hormones, agrochemicals, and plastics-derived endocrine disruptors in challenging matrices.<sup>5-7</sup> In the latter works, LC-DAD and LC-FLD second-order data were separately analyzed with multivariate curve resolution-alternating least-squares (MCR-ALS). Beyond the peculiar characteristics of each studied system, the general advantages of the dual detection were: 1) selection of the most appropriate type of signal for each analyte, and 2) mathematical resolution of coeluting analytes in the presence of foreign compounds, through second-order calibration.<sup>5-7</sup>

A further development in this field would be to combine the data from the detectors, instead of treating them independently. It is likely that the information content will be enhanced by a synergistic effect. Such technique involves the combination of data from different sources or detectors to produce a single model or decision, and is called data fusion.<sup>8</sup> For example, in polymer characterization with size exclusion chromatography it is common to sequentially couple several detectors in order to obtain information of complex polymer samples in real time.<sup>9,10</sup> Peré-Trepat and Tauler combined data from a DAD and a mass spectrometry detector, concluding that data fusion from both techniques improved the information available from the individual detectors.<sup>11</sup> In the latter case, fusion was performed without the need of synchronizing the matrices in the time direction, because there was a constant time difference between both detectors. In general, however, the lack of constancy of the time lag between the detectors precludes a simple analysis. In the present work this critical

1  
2  
3 problem was overcome, for the first time, through an adequate chromatographic alignment  
4  
5 before fusing the data.  
6

7 LC-DAD and LC-FLD matrices were pre-treated in order to align them in the time  
8  
9 axis, using correlation optimized warping (COW),<sup>12</sup> followed by MCR-ALS<sup>13</sup> to process the  
10  
11 obtained second-order data. Although the latter algorithm is able to take into account  
12  
13 chromatographic peak distortions from sample to sample, making it unnecessary the  
14  
15 synchronization of time profiles, in the present case the alignment proceeds between the data  
16  
17 provided by two detector for the same sample. This is necessary to preserve the bilinearity of  
18  
19 the fused data matrix for a given sample.  
20  
21

22  
23 The aim of this work was to evaluate the advantages of fusing LC-DAD and LC-FLD  
24  
25 second-order data obtained from two experimental systems. The experimental systems (Table  
26  
27 S1, Supporting Information) were constituted by fluorescent and non-fluorescent endocrine  
28  
29 disruptors (EDs) which can be found in the environment, food, and consumer products. They  
30  
31 may alter the actions of the endocrine system, resulting in adverse health effects in the live  
32  
33 organisms and their progeny. The list of EDs includes a large number of diverse  
34  
35 compounds.<sup>14</sup> Among them, we selected agrochemicals,<sup>15</sup> plastics-derived xenoestrogens,<sup>16</sup>  
36  
37 and priority polycyclic aromatic hydrocarbons (PAHs), which in addition to their  
38  
39 carcinogenic and genotoxic effects have demonstrated to be potent EDs.<sup>17</sup> Advantages and  
40  
41 drawbacks of the proposed strategy are discussed, based on a comparison with the results  
42  
43 obtained from individual LC-DAD or LC-FLD matrices. The comparison will be carried out  
44  
45 through the estimation of analytical figures of merit such as analytical sensitivity, selectivity  
46  
47 and limit of detection, and a recovery statistical indicator such as the relative error of  
48  
49 prediction for a set of validation samples. Finally, the feasibility of the fusion strategy towards  
50  
51 real water matrices containing foreign compounds is demonstrated.  
52  
53  
54  
55  
56  
57  
58  
59  
60

## EXPERIMENTAL SECTION

**Reagents and Solutions.** All reagents were of high-purity grade and used as received. Fluoranthene (FLT), pyrene (PYR), benzo[*a*]pyrene (BaP), benz[*a*]anthracene (BaA), benzo[*b*]fluoranthene (BbF), dibenz[*a,h*]anthracene (DBA), benzo[*g,h,i*]perylene (BghiP), carbaryl (CBL), 1-naphthol (NAP), bisphenol A (BPA), dimethyl phthalate (DMP) and dibutyl phthalate (DBP) were purchased from Sigma-Aldrich (St. Louis, MO, USA). 4-Nonylphenol (NP) and norflurazon (NOR) were provided by Fluka (Buchs, Switzerland). Methanol (MeOH) and acetonitrile (ACN) were obtained from Merck (Darmstadt, Germany). Ultrapure water was delivered by a Milli-Q apparatus (Millipore, Molsheim, France). MeOH stock solutions of all the analytes of about 1000 mg L<sup>-1</sup> were prepared and stored in dark flasks at 4 °C. From these solutions, more diluted MeOH solutions (2.00–10.0 mg L<sup>-1</sup>) were obtained. Due to their high toxicity, all reagents were handled with extreme care, using gloves and protective clothing.

**Apparatus.** Chromatographic analysis was carried out on an Agilent 1200 liquid chromatograph (Agilent Technologies, Waldbronn, Germany) equipped with a quaternary pump, a thermostated column compartment set at 35 °C, and two detectors (DAD and FLD) connected in series, with the flow first passing through the DAD and then through the FLD. A Rheodyne valve with a 20.0 µL loop was employed to inject the sample on to a Poroshell 120 EC C18 column (4.6 mm×50 mm, 2.7 µm particle size). The data were collected using the software HP ChemStation for LC Rev. HP 1990–1997.

**Chromatographic analysis.** Isocratic chromatographic conditions were employed in the evaluated systems. System I comprised DBP, FLT, PYR, BaA, NP, BbF, BaP, DBA, and BghiP. Concentrations were in the range 0-100 ng mL<sup>-1</sup> for DBP and 0-50 ng mL<sup>-1</sup> for the remaining analytes. The mobile phase was a mixture of ACN:water (85:15 v/v) and the flow rate was maintained at 1.25 mL min<sup>-1</sup> for a run time of 3.2 min. The DAD collected spectra

1  
2  
3 every 0.8 s, from 200 nm to 350 nm, each 2 nm, whereas fluorescence emission spectra were  
4  
5 collected every 2.2 s, in the range 295-450 nm, in steps of 10 nm (saved by the instrument  
6  
7 software each 1 nm), with an excitation wavelength of 222 nm.  
8

9  
10 System II was constituted by DMP, CBL, BPA, NAP and NOR, with concentrations in  
11  
12 the range 0-50 ng mL<sup>-1</sup> for BPA, CBL and NAP, and 0-100 ng mL<sup>-1</sup> for NOR and DMP. A  
13  
14 mixture of ACN:water (41:59 v/v) was used as mobile phase, and the flow rate was kept at  
15  
16 1.25 mL min<sup>-1</sup> involving a run time of 3 min per sample. Absorbance spectra were collected  
17  
18 every 0.4 s, from 200 nm to 320 nm, each 2 nm. Fluorescence emission spectra were  
19  
20 registered every 1.2 s, in the range 295-430 nm, in steps of 10 nm (saved by the instrument  
21  
22 software each 1 nm), with an excitation wavelength of 225 nm.  
23

24  
25 Since each emission data point is measured in 28 ms, the total time required for  
26  
27 collecting an emission spectrum was about 0.4 s in both systems. This latter value is  
28  
29 significantly smaller than the width at half height of a typical chromatographic band (ca. 5 s).  
30  
31 This means that the concentration change between successive data points is negligible, and  
32  
33 that the LC-FLD matrices are fairly bilinear.  
34  
35

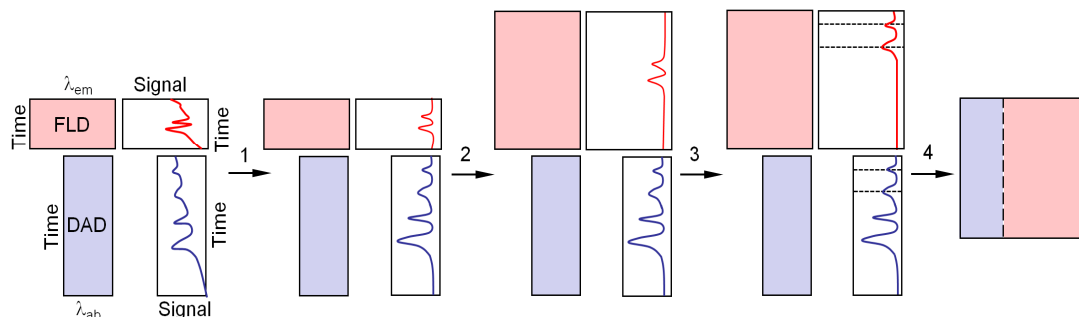
36 **Calibration and validation samples.** Concentrations of the calibration solutions for  
37  
38 systems I and II are shown in Table S2 (Supporting Information). Sixteen samples of the  
39  
40 calibration set of system I had concentrations provided by a fractional factorial design, and the  
41  
42 remaining two samples were a blank, and a solution containing the nine studied analytes at the  
43  
44 average concentration. For system II, a calibration set of ten samples was prepared: eight  
45  
46 samples of the set had concentrations provided by a fractional factorial design, the ninth one  
47  
48 contained only two of the studied analytes at average concentrations, and the last one  
49  
50 contained only the remaining three analytes, also at average concentrations. It should be  
51  
52 noticed that fractional factorial designs provide analyte concentrations with minimal  
53  
54 correlation, a condition which is necessary for succesful decomposition using MCR-ALS. For  
55  
56  
57  
58  
59  
60

1  
2  
3 each system, a validation set of 10 randomized samples was prepared in the corresponding  
4 calibration ranges. Data were saved in ASCII format, and transferred to a PC Sempron AMD  
5 microcomputer for subsequent computations.  
6  
7

8  
9  
10 **Real water samples.** Because the investigated water samples (underground, tap and  
11 river waters) did not contain the analytes of system II at levels higher than the attained  
12 detection limits, a recovery study was performed by spiking them with standard solutions of  
13 DMP, CBL, BPA, NAP, and NOR, obtaining concentration levels within their linear ranges.  
14 Underground water was obtained from Funes City (Santa Fe, Argentina), and the Paraná  
15 River water sample was collected near Rosario City (Santa Fe, Argentina). Aliquots of  
16 standard methanol solutions of the analytes were added to 5.00 mL flasks. The solvent was  
17 evaporated by using a nitrogen stream, and the sample was reconstituted, first with 3.00 mL  
18 of the corresponding real water matrix, and then acetotriline to the mark, to achieve the  
19 mobile phase composition. The samples were then filtered twice through 45  $\mu\text{m}$  pore size  
20 nylon membranes, and subjected to the same chromatographic analysis as the validation  
21 samples.  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35

36 **Data pretreatment.** Both LC-DAD and LC-FLD matrices were collected for each  
37 sample in a single run. The number of rows in these matrices was equal to the number of  
38 elution times, while the number of columns was equal to the number of recorded absorbance  
39 or fluorescence emission wavelengths. Before the fusion of the data matrices, a pretreatment  
40 was applied. The latter consisted in a series of steps to overcome the following problems: (1)  
41 baseline distortions, (2) different number of DAD and FLD time channels, and (3) shifts and  
42 distortions in chromatographic profiles between both detectors. Regarding issue (1), it is  
43 important to remark that baseline distortions could be more complex when gradient elution is  
44 applied. However, as was previously demonstrated, this fact does not represent serious  
45 problems for MCR-ALS.<sup>7</sup> The pretreatment is summarized in Figure 1.  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60





**Figure 1.** Schematic representation of the data pretreatment carried out before the LC-DAD – LC-FLD fusion: 1, baseline correction; 2, time interpolation; 3, alignment; 4, data fusion.

We now briefly describe these steps.

1) After loading the LC-DAD and LC-FLD matrices for a sample, baseline correction was applied using asymmetric least-squares.<sup>18</sup>

2) LC-DAD and LC-FLD matrices should have the same number of time channels before fusion. Since the sampling frequency is larger for the DAD than for FLD, time points were added to each column of the LC-FLD matrix by linear interpolation, using the MATLAB function 'interp1'.

3) For time alignment, two vectors ( $\mathbf{y}_{\text{DAD}}$  and  $\mathbf{y}_{\text{FLD}}$ ) were first constructed by digitally adding the columns corresponding to the absorbance or fluorescence maxima for all analytes. This produced two virtual chromatograms, containing a number of peaks equal to the total number of analytes (except for the non-fluorescent analytes in the LC-FLD matrix), which is necessary for alignment purposes. A first approximation to the time shift correction between LC-DAD and LC-FLD was performed by digitally moving the latter with respect to the former a number of sensors, based on a comparison of the two vectors mentioned above. This can be done either visually or automatically using a software for peak detection. A new  $\mathbf{y}_{\text{FLD}}$  vector is computed as described above for the shifted LC-FLD matrix.

4) Further alignment is needed owing to presence of non-uniform time shifts along the time mode. This was achieved using one-dimensional correlation optimized warping (1D-

1  
2  
3 COW).<sup>12</sup> Input parameters are segment (width of the narrowest peak) and slack (in the  
4 interval 1 to 4), and warping parameters are estimated using  $y_{DAD}$  as reference and the  
5 recalculated  $y_{FLD}$  as target, using the MATLAB function 'cow'. If the alignment is not  
6 completely satisfactory, the segment and slack are slightly modified.  
7  
8  
9

10  
11 5) Every column of the interpolated LC-FLD matrix is warped using the optimized  
12 COW parameters in the MATLAB function 'cow\_apply'.  
13

14  
15 6) If necessary, the final LC-FLD matrix is multiplied by a scaling factor, in such a  
16 way that the maximum intensity of the latter matrix is equal to the maximum intensity of the  
17 LC-DAD matrix.  
18  
19  
20

21  
22 7) LC-DAD and the corrected LC-FLD are fused to give a single LC-DAD-FLD  
23 (temporal-wise augmentation) matrix, which is saved in ASCII format for data processing  
24 with MCR-ALS.  
25  
26  
27

28  
29 Data analysis. MCR-ALS data processing involved the building of augmented fused  
30 matrices along the elution time direction, containing each validation sample in turn and the  
31 calibration data matrices. Principal component analysis (PCA) was employed to estimate the  
32 number of components present in the fused systems, obtaining results in agreement with the  
33 number of theoretically expected components. The estimates corresponded to the value at  
34 which the PCA residual fit stabilizes close to the noise level in the plot of residual fit as a  
35 function of the number of components. In all cases, the selected number of components  
36 explained at least 99% of the total variance. Initial spectral profiles employed to start MCR-  
37 ALS fitting were obtained from the so-called purest variables in the spectral domain. The  
38 following constraints were imposed during ALS fit: non-negativity in both spectral and  
39 temporal modes, unimodality in the temporal mode, correspondence between components and  
40 samples,<sup>13</sup> and a recently developed restriction called "area correlation constraint".<sup>19</sup>  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 The selected ALS convergence criterion was 0.01% (relative change in fit for  
4 successive iterations), which was achieved in less than 50 iterations. After convergence of the  
5 ALS phase, analytes were identified by their spectral profiles and MCR-ALS scores were  
6 employed for their quantification. For comparison, LC-DAD and LC-FLD data belonging to  
7 the two studied systems were separately processed by MCR-ALS following the same  
8 procedure and applying the restrictions indicated above.  
9  
10  
11  
12  
13  
14  
15

16 **Chemometric algorithm and software.** A brief theoretical description of MCR-ALS  
17 is given in the Supporting Information. The routines employed are written in MATLAB 7.0.  
18 (Mathworks, MA, USA). MCR-ALS was implemented using the graphical interface of the  
19 MVC2 toolbox,<sup>20</sup> freely available on the Internet.<sup>21</sup> COW was applied using the MATLAB  
20 codes freely provided in the Internet.<sup>22</sup>  
21  
22  
23  
24  
25  
26  
27  
28  
29

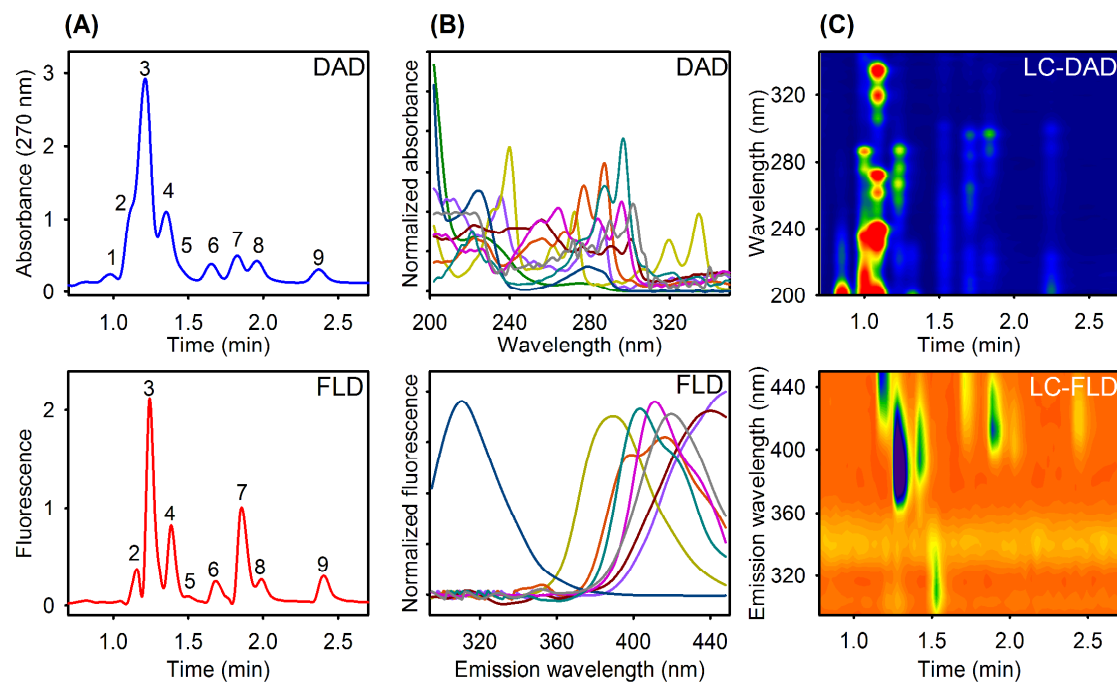
## 30 **RESULTS AND DISCUSSION**

31 **General considerations.** Chromatographic conditions for both evaluated systems  
32 were set to achieve suitable results under isocratic regimes and involving the shortest possible  
33 overall run time. LC-DAD and LC-FLD matrices for calibration and validation samples of the  
34 studied systems were first recorded, and predictions were then obtained from each detector  
35 separately (DAD only if the analyte was non-fluorescent). Subsequently, individual matrices  
36 were pretreated in order to properly fuse them.  
37  
38  
39  
40  
41  
42  
43  
44

45 MCR-ALS was the algorithm of choice for processing all matrix data, because of the  
46 lack of repeatability in the elution time profiles between successive runs, i.e., from sample to  
47 sample.<sup>13,23</sup> Selection of the number of MCR-ALS components and the applied restrictions in  
48 the studied systems are indicated in the experimental section.  
49  
50  
51  
52

53 **System I.** This system is constituted by two xenoestrogens (DBP and NP) and seven  
54 PAHs. Figure 2 shows typical DAD and FLD chromatograms for a calibration sample, the  
55  
56  
57  
58  
59  
60

corresponding DAD and FLD spectra, and the contour plots of LC-DAD and LC-FLD data. The low resolution of the DAD and FLD chromatographic bands and the large degree of overlap in both spectral modes can be appreciated.



**Figure 2.** DAD (blue) and FLD (red) chromatograms of a selected calibration sample (sample number 2 in Table S2) of system I (A), normalized absorption and fluorescence emission spectra for the assayed analytes (B), and the corresponding two-dimensional contour plots (C). For FLD, excitation and emission wavelengths were 222 and 410 nm, respectively. DBP (1, green), FLT (2, violet), PYR (3, dark yellow), BaA (4, orange), NP (5, dark blue), BbF (6, burgundy), BaP (7, pink), DBA (8, dark cyan) and BghiP (9, gray).

DBP (non-fluorescent) and NP (weakly fluorescent) were only quantified through their UV signals, but PAHs were determined using both types of detectors because they display intense absorbance and fluorescence signals. Table S3 of Supporting Information shows the number of estimated MCR components for LC-DAD and LC-FLD matrices and the fitting parameters after convergence of the ALS optimization, and Table S4 displays the obtained concentration recoveries.

The MCR-ALS statistical results (relative error of prediction, REP) obtained from each detector, and figures of merit ( $\gamma$ , analytical sensitivity; Sel, selectivity; and LOD, limits of detection) calculated from rigorous approaches,<sup>24</sup> are shown in Table 1. The values of

analytical sensitivity and selectivity using LC-DAD data, in the ranges of 0.09–1.6 ng<sup>-1</sup> mL and 0.1–0.4, respectively, are significantly larger than those calculated from LC-FLD data, which were in the ranges 0.01–0.5 ng<sup>-1</sup> mL and 0.001–0.2, respectively. The REP values, in the range 2–7% for LC-DAD data, are also notably better than those obtained with LC-FLD data, with values up to 22%. This result can be ascribed to the fact that all investigated PAHs fluoresce in the same spectral range (Figure 2B), with bands severely overlapped. Moreover, LODs estimated for FTL, BbF, DBA and BghiP with the standard expression<sup>24</sup> were unreasonably large, due to the low sensitivity values. Further work is in progress to analyze this inconsistency.

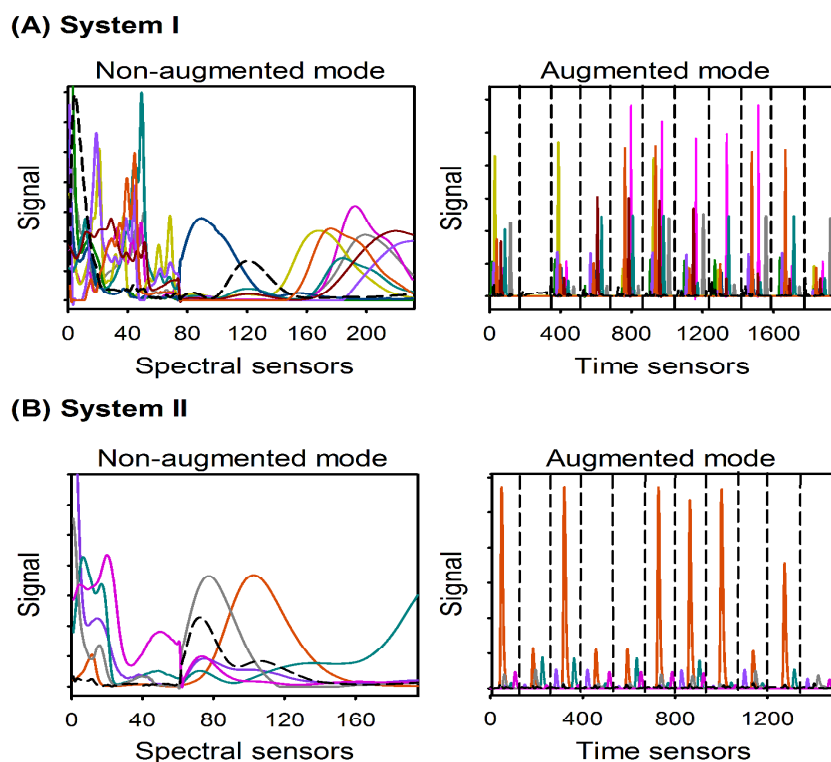
**Table 1 Statistical results for non-fluorescent (nf) and fluorescent (f) analytes of system I determined by LC-DAD, LC-FLD and LC-DAD-FLD data<sup>a</sup>**

	nf	f							
	DBP	FLT	PYR	BaA	NP <sup>b</sup>	BbF	BaP	DBA	BghiP
<b>LC-DAD</b>									
γ	0.09	0.5	1.6	1.5	0.1	0.6	0.7	0.7	0.4
Sel	0.1	0.2	0.4	0.3	0.3	0.2	0.2	0.2	0.1
LOD	38	7.6	2.3	2.5	26	6.2	5.3	5.0	8.6
REP	7	4	2	3	6	6	2	3	7
<b>LC-FLD</b>									
γ		0.01	0.5	0.2		0.01	0.3	0.03	0.03
Sel		0.03	0.2	0.01		0.001	0.1	0.002	0.002
LOD		<sup>c</sup>	7.8	16		<sup>c</sup>	13	<sup>c</sup>	<sup>c</sup>
REP		22	4	6		12	8	17	11
<b>LC-DAD-FLD</b>									
γ	0.2	0.4	1.8	1.3	0.3	0.6	0.9	1.0	0.4
Sel	0.4	0.4	0.6	0.4	0.8	0.3	0.3	0.5	0.3
LOD	24	8.3	2.1	2.8	11	6.6	4.1	3.6	9.1
REP	7	6	2	5	5	5	4	4	5

<sup>a</sup> γ (ng<sup>-1</sup> mL), analytical sensitivity; Sel, selectivity, LOD (ng mL<sup>-1</sup>), limit of detection calculated according to ref. 21; REP (%), relative error of prediction. <sup>b</sup> NP was not determined by FLD due to lack of selectivity (see text). <sup>c</sup> The LOD values estimated for these analytes were unreasonably large (see text).

In conclusion, in system I and under the employed working conditions, dual detection may not bring additional advantages to the analysis. However, the potential improvement of

1  
2  
3 data fusion remains to be analyzed. As an example, Figure S1A (Supporting Information)  
4 illustrates the procedure followed for the alignment and fusion of experimental DAD and FLD  
5 chromatograms of system I. Tables S3 and S4 shows the MCR-ALS fitting parameters and  
6 the predicted concentration values, respectively. MCR-ALS was able to retrieve satisfactory  
7 DAD and FLD spectral profiles from the fused matrices (Figure 3).  
8  
9  
10  
11  
12



40  
41  
42  
43  
44  
45  
46  
47

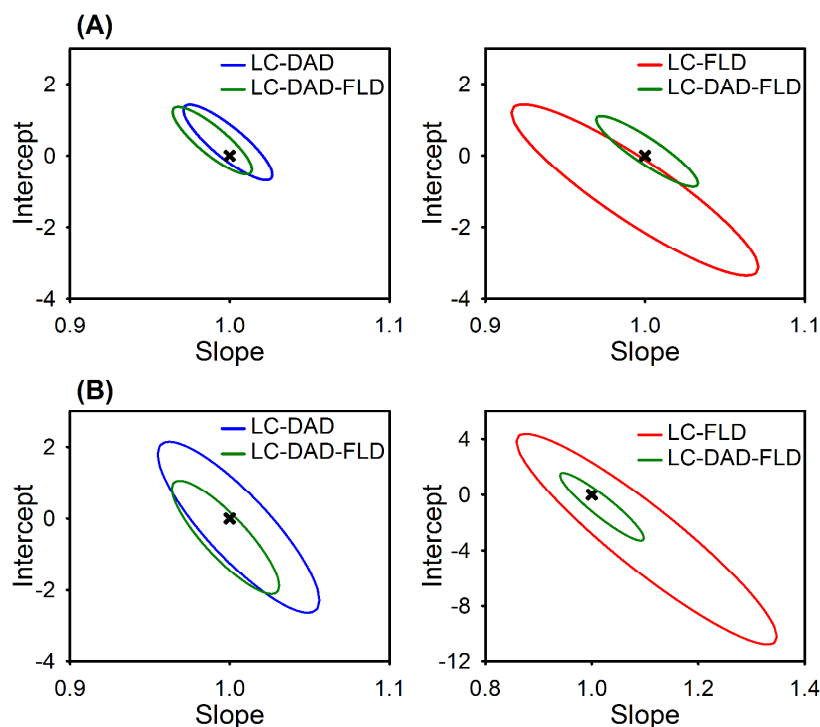
**Figure 3.** Spectral and time profiles retrieved by MCR-ALS when processing fused data for a validation sample of systems I (sample S1 in Table S4) (A) and II (sample S3 in Table S5) (B). In (A): DBP (green), FLT (violet), PYR (dark yellow), BaA (orange), NP (dark blue), BbF (burgundy), BaP (pink), DBA (dark cyan), BghiP (gray), and blank (black dotted-line). In (B) DMP (violet), CBL (orange), BPA (gray), NAP (dark cyan), NOR (pink), and blank (black dotted-line). The dotted vertical lines in the time profiles separate, from left to right, the studied validation sample and the successive calibration samples (In (A) only 10 of the 18 calibration samples were included in the plot).

48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

The figures of merit for the fused data (Table 1) suggest that the analytical performance for the non-fluorescent DBP improves with the fusion. Indeed, the selectivity towards this analyte increases four-fold (from 0.1 to 0.4) in going from LC-DAD to fused data. The data fusion favors the detection of DBP from the remaining sample constituents, demonstrating the positive effect of the fusion in the selectivity of the method. Likewise, the

LOD decreases from 38 to 24 ng mL<sup>-1</sup>. For a similar reason, the selectivity towards the PAHs significantly improves: 0.1–0.4 (LC-DAD), 0.001–0.2 (LC-FLD) and 0.3–0.6 (fused data), with a remarkable decrease in the LOD of NP: from 26 ng mL<sup>-1</sup> (LC-DAD) to 11 ng mL<sup>-1</sup> (fused data). The obtained REP values (all below 7%) indicate good recoveries and precision.

Figure 4A compares the elliptical joint confidence region (EJCR)<sup>25</sup> test for the slopes and intercepts of the plots of predicted concentrations vs. nominal ones obtained from individual LC-DAD and LC-FLD data, and from LC-DAD-FLD fused data. Because all ellipses include the theoretically expected values of (1,0) for slope and intercept, respectively, the accuracy of the applied methodologies can be claimed. The size of the LC-DAD ellipse is similar to that for LC-DAD-FLD, suggesting similar precision, whereas a significant size difference is verified between fluorescence and fused data.

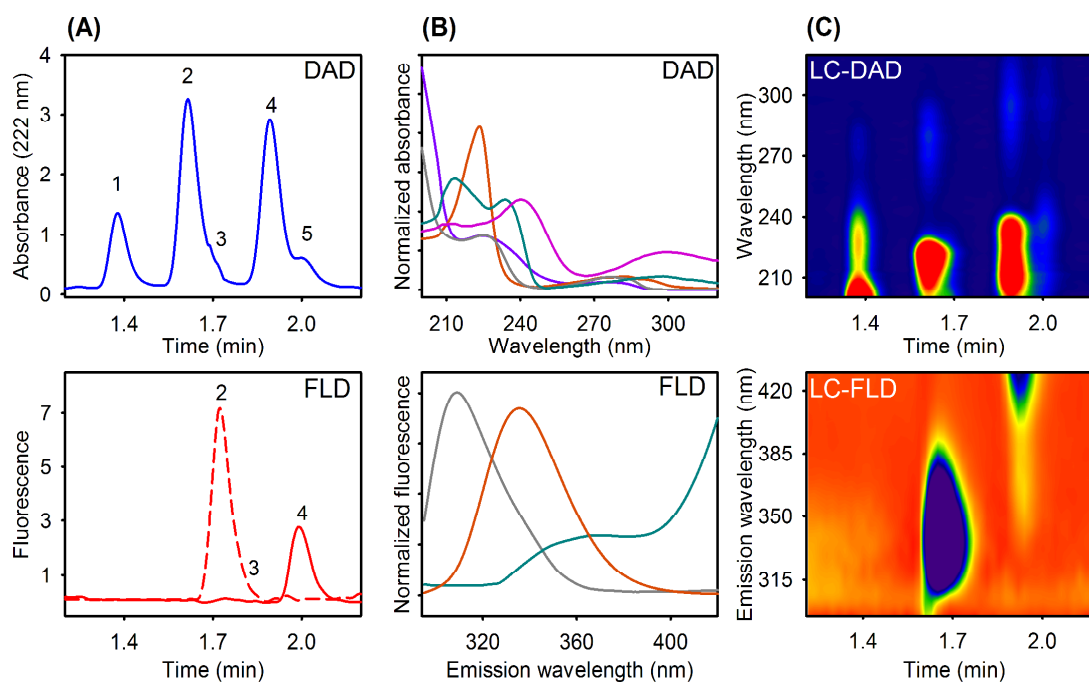


**Figure 4.** Elliptical joint regions (at 95% confidence level) for slope and intercept of the MCR-ALS regressions of individual LC-DAD (blue) and LC-FLD (red) data, and fused LC-DAD-FLD data (green) for system I (A) and system II (B). Crosses mark the theoretical (intercept = 0, slope = 1) point.

**System II.** This system comprises two non-fluorescent (DMP and NOR) and three fluorescent (CBL, BPA and NAP) compounds (Figure 5). Figure 1SB (Supporting

Information) illustrates the alignment and fusion of experimental DAD and FLD chromatograms for this system. The number of estimated responsive constituents, residual fits and explained variances for MCR-ALS in this system using LC-DAD, LC-FLD and LC-DAD-FLD data are shown in Table S3. Table S5 displays the corresponding concentration predictions.

All five analytes were quantified through individual LC-DAD data, and CBL and NAP were also determined by the analysis of their LC-FLD matrices. Although BPA is fluorescent, its recoveries from the FLD data were very poor (the results were not included in Table S5).



**Figure 5.** DAD (blue) and FLD (red) chromatograms of a selected calibration sample (sample number 4 in Table S2) of system II (A), normalized absorption and fluorescence emission spectra for the assayed analytes (B), and the corresponding two-dimensional contour plots (C). Emission wavelengths for the FLD detection were 312 nm (dotted line) and 425 nm (solid line), and the excitation wavelength was 225 nm. DMP (1, violet), CBL (2, orange), BPA (3, gray), NAP (4, dark cyan) and NOR (5, pink).

From the analysis of the MCR-ALS statistical results for the individual and combined detection modes (Table 2), the general conclusion is that the fusion improves the analytical sensitivity, selectivity and LODs for the five analytes, with ranges of 0.6–35 ng<sup>-1</sup> mL, 0.4–



0.9, and 0.9–6.3 ng mL<sup>-1</sup>, respectively. Adequate REPs, in the range 4–7%, are also obtained with the fusion. Specifically, it is noteworthy that data fusion is critical for the analytes DMP and BPA. The absorbance spectra of both analytes are very similar (Figure 5B), producing degeneracy in the LC-DAD resolution and, therefore, poor statistical values (see Table 2). Besides, the fluorescence detection for BPA is also unsuccessful because this analyte partially coelutes with the highly fluorescent CBL (Figure 5A). Data fusion can overcome the serious problem of selectivity in both cases. The analytical sensitivity increases six-fold and five-fold for DMP and BPA, respectively, and selectivity increases four-fold in both cases. LOD conveniently decreases from 28 (LC-DAD) to 6.3 ng mL<sup>-1</sup> (fused data) for DMP, and from 25 (LC-DAD) to 3.9 ng mL<sup>-1</sup> (fused data) for BPA.

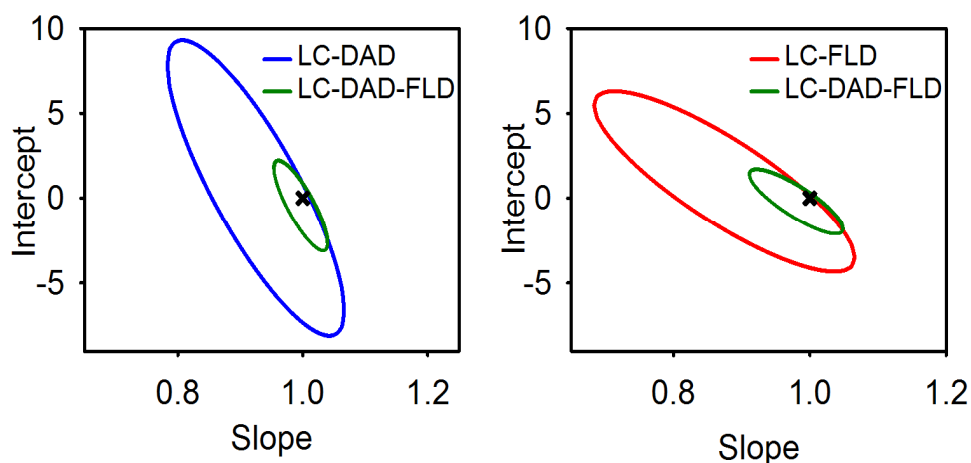
**Table 2 Statistical results for non-fluorescent (nf) and fluorescent (f) analytes of system II determined by LC-DAD, LC-FLD and LC-DAD-FLD data<sup>a</sup>**

	nf		f		
	DMP	NOR	CBL	BPA <sup>b</sup>	NAP
LC-DAD					
γ	0.1	0.6	1.6	0.2	1.0
Sel	0.1	0.4	0.4	0.1	0.3
LOD	28	6.7	2.5	25	4.0
REP	11	5	6	8	7
LC-FLD					
γ			5.4		1.7
Sel			0.8		1.0
LOD			1.2		2.4
REP			5		5
LC-DAD-FLD					
γ	0.6	1.0	35	1.0	4.1
Sel	0.4	0.7	0.9	0.4	0.7
LOD	6.3	3.7	0.9	3.9	1.3
REP	6	4	4	7	5

<sup>a</sup> γ (ng<sup>-1</sup> mL), analytical sensitivity; Sel, selectivity, LOD (ng mL<sup>-1</sup>), limit of detection calculated according to ref. 21; REP (%), relative error of prediction. <sup>b</sup> BPA was not determined by FLD due to lack of selectivity (see text).

Furthermore, the fluorescent detections of both CBL and NAP are better than the DAD ones, and fusion improves even more the already good results obtained from the LC-FLD data, with LODs of 0.9 and 1.3 ng mL<sup>-1</sup> for CBL and NAP, respectively. Although the ellipses calculated from the results obtained with the individual detectors and with data fusion include the 1,0 ideal point, the sizes of the latter are significantly smaller, suggesting a better precision when data fusion is employed (Figure 4B).

**Real water samples.** In view of the above results, system II was selected to probe the fusion procedure in real water samples through a recovery study. As expected, due the presence of foreign species in the real matrices, either one or two additional MCR-ALS components were required during the chemometric processing. The explained variance after convergence of the ALS optimization was always higher than 99%. Table S6 of Supporting Information displays the obtained concentration recoveries. As in validation samples, the advantage of fusion is remarkable in the recoveries of DMP and BPA.



**Figure 6.** Elliptical joint regions (at 95% confidence level) for slope and intercept of the MCR-ALS regressions of individual LC-DAD (blue) and LC-FLD (red) data, and fused LC-DAD-FLD data (green) for real water samples. Crosses mark the theoretical (intercept = 0, slope = 1) point.

The significant differences in the size of ellipses when the statistical EJCR test is applied to individual LC-DAD and LC-FLD data, and to LC-DAD-FLD fused data (Figure 6)

1  
2  
3 support the gain in precision when fusion is applied. The obtained results also suggest that the  
4  
5 presence of foreign compounds which are present in the studied matrices does not produce a  
6  
7 significant interference in our analysis.  
8  
9

## 10 11 **CONCLUSIONS**

12  
13  
14 The presently proposed fusion approach provides a useful and reliable way of  
15  
16 improving the analytical quality of the results in second-order chromatographic analysis. The  
17  
18 benefit of fusion is highlighted both in validation and in real matrices when analytes cannot be  
19  
20 quantitatively estimated from individual detectors, and in cases of low selectivity. It is  
21  
22 noteworthy that this resource continues to add advantages to those already achieved with the  
23  
24 coupling of dual chromatographic detection to chemometric analysis.  
25  
26  
27

## 28 **ACKNOWLEDGMENT**

29  
30 Authors thanks Prof. Alejandro Olivieri for helpful discussions. Universidad Nacional  
31  
32 de Rosario (Project BIO 237), CONICET (Consejo Nacional de Investigaciones Científicas y  
33  
34 Técnicas, Project PIP 0163), and ANPCyT (Agencia Nacional de Promoción Científica y  
35  
36 Tecnológica, Project PICT 2013-0136) are gratefully acknowledged for financial support.  
37  
38 R.P.V thanks CONICET for a doctoral fellowship.  
39  
40  
41  
42  
43

44 **Supporting Information.** A brief explanation of the theory of MCR-ALS algorithm,  
45  
46 additional tables and a figure.  
47  
48  
49

## 50 **References**

51  
52  
53 (1) Arancibia, J. A.; Damiani, P. C.; Escandar, G. M.; Ibañez, G. A.; Olivieri, A. C. *J.*  
54  
55 *Chromatogr. B* **2012**, *910*, 22–30.  
56  
57  
58  
59  
60

- 1  
2  
3  
4 (2) Parastar, H. In *Data Handling in Science and Technology, Fundamentals and Analytical*  
5  
6 *Applications of Multiway Calibration*; Muñoz de la Peña, A.; Goicoechea, H. C.; Escandar, G.  
7  
8 M.; Olivieri, A. C., Eds.; Elsevier, Amsterdam, 2015, vol. 29, ch. 6, pp. 314-324.  
9  
10 (3) Olivieri, A. C.; Escandar, G. M., *Practical Three-Way Calibration*, Elsevier, Waltham,  
11  
12 USA, 2014.  
13  
14 (4) Pérez, R. L.; Escandar, G. M. *Sustain. Chem. Pharm.* **2016**, *4*, 1–12.  
15  
16 (5) Pérez, R. L.; Escandar, G. M. *Environ. Pollut.* **2016**, *209*, 114–122.  
17  
18 (6) Carabajal, M. D.; Arancibia, J. A.; Escandar, G. M. *Microchem. J.* **2016**, *128*, 34–41.  
19  
20 (7) Pellegrino Vidal, R. B.; Ibañez, G. A.; Escandar, G. M. *Talanta* **2016**, *159*, 336–343.  
21  
22 (8) Liu, Y.; Brown, S. D. *Anal. Bioanal. Chem.* 2004, *380*, 445–452.  
23  
24 (9) Brun, Y.; Gorenstein, M. V.; Hay, N. *J. Liq. Chromatogr. Relat. Technol.* **2000**, *23*, 2615–  
25  
26 2639.  
27  
28 (10) Guerre, M.; Wahidur Rahaman, S. M.; Améduri, B.; Poli, R.; Ladmiral, V. *Polym. Chem.*  
29  
30 **2016**, *7*, 6918–6933.  
31  
32 (11) Peré-Trepat, E.; Tauler, R. *J. Chromatogr., A* **2006**, *1131*, 85–96.  
33  
34 (12) Nielsen, N. P. V.; Carstensen, J. M.; Smedsgaard, J. *J. Chromatogr., A* **1998**, *805*, 17–35.  
35  
36 (13) Tauler, R.; Maeder, M.; de Juan, A. In *Comprehensive Chemometrics*; Brown, S.; Tauler,  
37  
38 R.; Walczak, B., Eds.; Elsevier, Oxford, 2009, pp. 473–505.  
39  
40 (14) Chang, H. S.; Choo, K. H.; Lee, B.; Choi, S. J. *J. Hazard. Mater.* **2009**, *172*, 1–12.  
41  
42 (15) Sun, H.; Shen, O. X.; Xu, X. L.; Song, L.; Wang, X. R. *Toxicology* **2008**, *249*, 238–242.  
43  
44 (16) LaFleur, A. D.; Schug, K. A. *Anal. Chim. Acta* **2011**, *696*, 6–26.  
45  
46 (17) Zhang, Y.; Dong, S.; Wang, H.; Tao, S.; Kiyama, R. *Environ. Pollut.* **2016**, *213*, 809–  
47  
48 824.  
49  
50 (18) Eilers, P. H. C.; Currie, I. D.; Durbán, M. *Comput. Stat. Data Anal.* **2006**, *50*, 61–76.  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4 (19) de Oliveira Neves, A. C.; Tauler, R.; Gomes de Lima, K. M. *Anal. Chim. Acta* **2016**, *937*,  
5  
6 21–28.  
7

8 (20) Olivieri, A. C.; Wu, H. L.; Yu, R. Q. *Chemom. Intell. Lab. Syst.* **2009**, *96* (2), 246–251.  
9

10 (21) <http://www.iquir-conicet.gov.ar/descargas/mvc2.rar> (accessed November 2016).  
11

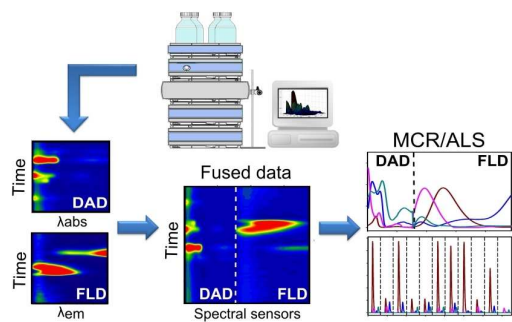
12 (22) <http://www.models.life.ku.dk/go?filename=WarpingTB.zip> (accessed November 2016).  
13

14 (23) Olivieri, A. C.; Escandar, G. M.; Muñoz de la Peña, A. *Trends Anal. Chem.* **2011**, *30*,  
15  
16 607–617.  
17

18 (24) Bauza, M. C.; Ibañez, G. A.; Tauler, R.; Olivieri, A. C. *Anal. Chem.* **2012**, *84*,  
19  
20 8697–8706.  
21  
22

23 (25) González, A. G.; Herrador, M. A.; Asuero, A. G. *Talanta* **1999**, *48*, 729–736.  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

for TOC only



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60