

Supplementary Information for: ‘Monitoring pollution pathways in river water by predictive path modelling using untargeted GC-MS measurements’

Maria Cairolì^{1, +, *}, André van den Doel^{1, +}, Berber Postma¹, Tim Offermans¹, Henk Zemelink², Gerard Stroomberg^{1, 3}, Lutgarde Buydens¹, Geert van Kollenburg^{1, 4, ++}, Jeroen Jansen^{1, ++}

¹ Radboud University, Department of Analytical Chemistry & Chemometrics, Institute for Molecules and Materials (IMM), Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands.

² Rijkswaterstaat, Zuiderwagenplein 2, 8224 AD Lelystad, The Netherlands.

³ RIWA Rijn, Groenendaal 6, 3439 LV, Nieuwegein, The Netherlands.

⁴ Eindhoven University of Technology, Interconnected Resource-aware Intelligent Systems, Den Dolech 2, 5612AZ, Eindhoven, The Netherlands.

⁺ These authors contributed to this work equally: Maria Cairolì, André van den Doel.

⁺⁺ These authors contributed to this work equally: Geert van Kollenburg, Jeroen Jansen.

* Correspondence e-mail address: chemometrics@science.ru.nl

This Supplementary Information file reports Supplementary Figures and Tables for the Manuscript: 'Monitoring pollution pathways in river water by predictive path modelling using untargeted GC-MS measurements'.

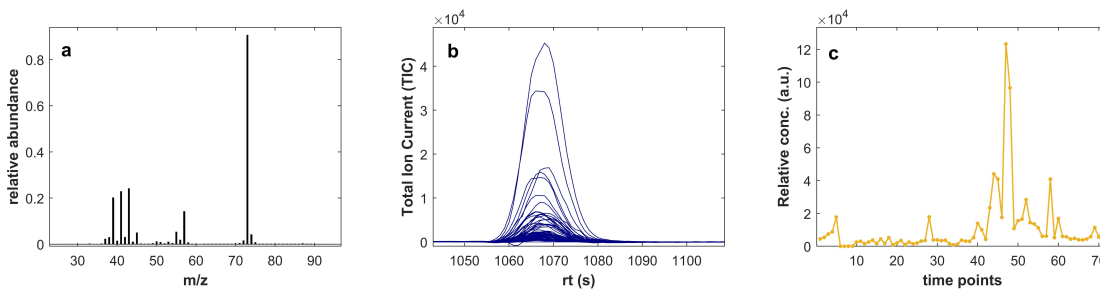
1. Supplementary Figures

Supplementary Figure 1 reports a map of the river Rhine and the tributary Lippe with the location of the sampling sites considered in the study.



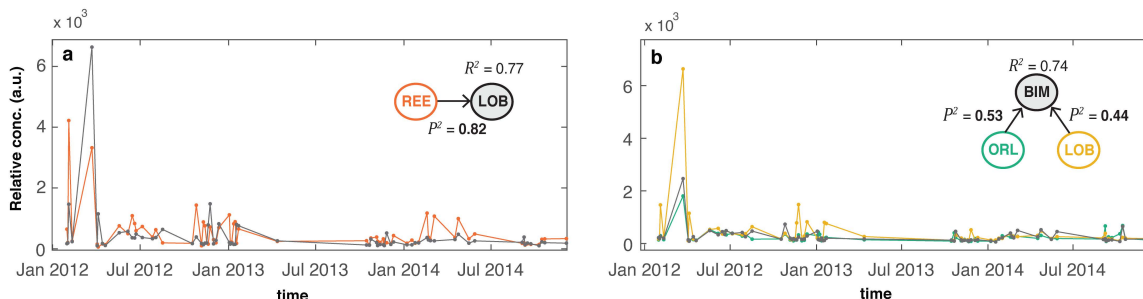
Supplementary Figure 1. Map showing the river Rhine and Lippe with the sampling sites considered in the study. Map credits: © Land NRW, dl-de/by-2-0 (www.govdata.de/dl-de/by-20) <https://www.elwasweb.nrw.de> 09/03/2023. © Bundesamt für Kartographie und Geodäsie 2023, Datenquellen: http://sg.geodatenzentrum.de/web_public/Datenquellen_TopPlus_Open_01.10.2017.pdf. The map was annotated by the authors on the ELWAS-WEB website(<https://www.elwasweb.nrw.de>) 09/03/2023

Supplementary Figure 2 shows, as example, the PARAFAC2 decomposition (PARADISE output) for the chemical tentatively annotated as MTBE at Bimmen.



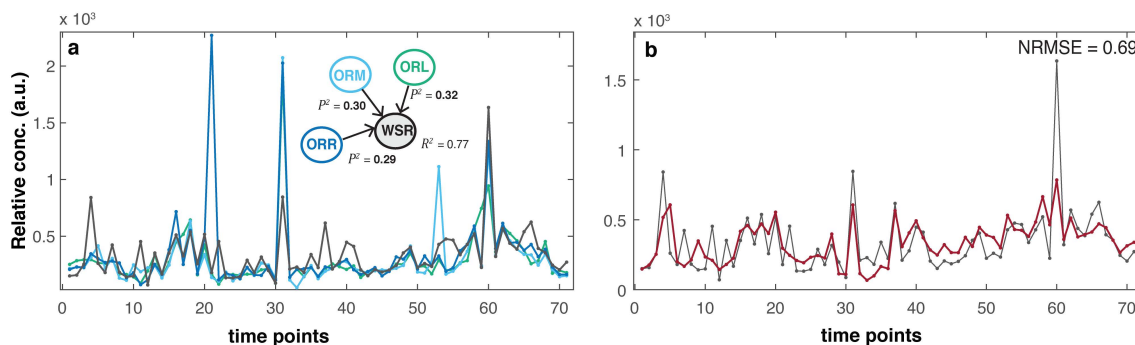
Supplementary Figure 2. PARAFAC2 decomposition for the chemical tentatively annotated as MTBE at Bimmen. **a** PARAFAC2 mass spectrum. **b** PARAFAC2 elution profiles for all the 71 samples measured at Bimmen, in retention time window 1050-1100 s. **c** PARAFAC2 relative concentrations of MTBE over the 71 time points (samples).

Supplementary Figure 3 shows, as example, the concentration profiles with the actual temporal spacing for the chemical tentatively annotated as cumene at Rees and Lobith (**Supplementary Figure 3a**) and at Orsoy Left, Lobith and Bimmen (**Supplementary Figure 3b**).



Supplementary Figure 3. Measured concentration profiles of cumene on the right and left sides of the river with actual temporal spacing. a Cumene concentration profile measured by PARAFAC2 in Rees (orange) and Lobith (black). **b** Measured cumene concentration profile in Orsoy Left (green), Lobith (yellow), and Bimmen (black).

Supplementary Figure 4 shows the concentration profiles measured by PARAFAC2 in the three sampling sites in Orsoy and in Wesel Rhine (**Supplementary Figure 4a**) of the chemical tentatively identified as benzene, 1,2,4,5-tetramethyl-. The three sites in Orsoy predict the chemical concentration profile in Wesel Rhine with NRMSE = 0.69 (**Supplementary Figure 4b**).



Supplementary Figure 4. Measured concentration profile of untargeted chemical in the three sites in Orsoy and in Wesel Rhine, Process PLS prediction in Wesel Rhine. a Chemical concentration profile measured by PARAFAC2 in Orsoy Left (green), Orsoy Middle (light blue), Orsoy Right (blue), and Wesel Rhine (black). **b** Measured (black) chemical concentration profile in Wesel Rhine and predicted profile by the three sampling sites in Orsoy (red).

2. Supplementary Tables

Supplementary Table 1 reports information on the water samples selected after the temporal synchronization step.

Supplementary Table 1. Synchronized water samples.

Sample	Date (dd/mm/yyyy HH:MM)
1	30/01/2012 17:40
2	01/02/2012 15:15
3	03/02/2012 15:00
4	10/02/2012 15:10
5	22/03/2012 19:00
6	02/04/2012 19:00
7	04/04/2012 19:00
8	13/04/2012 15:15
9	19/04/2012 15:00
10	18/05/2012 14:45
11	05/06/2012 15:10
12	14/06/2012 15:00
13	19/06/2012 15:12
14	22/06/2012 15:00
15	05/07/2012 15:05
16	26/07/2012 19:00
17	02/08/2012 15:08
18	16/08/2012 19:00
19	17/10/2012 15:00
20	25/10/2012 21:00
21	05/11/2012 19:00
22	09/11/2012 15:05
23	12/11/2012 19:00
24	16/11/2012 15:00
25	19/11/2012 15:05
26	23/11/2012 15:05
27	30/11/2012 19:00
28	05/12/2012 19:00
29	11/12/2012 19:00
30	02/01/2013 15:00
31	04/01/2013 15:00
32	07/01/2013 15:05
33	11/01/2013 19:00
34	14/01/2013 15:00
35	17/01/2013 15:00
36	21/01/2013 19:00

37	12/04/2013 11:15
38	14/10/2013 23:00
39	20/10/2013 23:00
40	23/10/2013 23:00
41	04/11/2013 23:00
42	06/11/2013 23:00
43	11/11/2013 23:00
44	18/11/2013 23:00
45	20/11/2013 23:00
46	25/11/2013 19:00
47	04/12/2013 11:00
48	10/12/2013 03:00
49	06/01/2014 23:00
50	17/01/2014 07:00
51	24/01/2014 03:00
52	31/01/2014 07:00
53	17/02/2014 19:00
54	20/02/2014 15:00
55	24/02/2014 19:00
56	04/03/2014 23:00
57	16/04/2014 19:00
58	23/04/2014 19:00
59	15/05/2014 19:00
60	27/05/2014 23:00
61	03/09/2014 15:00
62	07/09/2014 15:00
63	09/09/2014 15:00
64	14/09/2014 15:00
65	21/09/2014 15:00
66	08/10/2014 15:00
67	14/10/2014 15:00
68	20/10/2014 15:00
69	04/12/2014 15:00
70	14/12/2014 10:35
71	17/12/2014 11:15