



**HAL**  
open science

## Insights on impact estimation of unmeasured micropollutants from WWTP release

Rémi Servien, Kevin Bonnot, Eric Latrille, Arnaud Hélias, Dominique Patureau

### ► To cite this version:

Rémi Servien, Kevin Bonnot, Eric Latrille, Arnaud Hélias, Dominique Patureau. Insights on impact estimation of unmeasured micropollutants from WWTP release. 2023. hal-03952680v1

**HAL Id: hal-03952680**

**<https://hal.science/hal-03952680v1>**

Preprint submitted on 23 Jan 2023 (v1), last revised 21 Aug 2023 (v3)

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

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



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

# Insights on impact estimation of unmeasured micropollutants from WWTP release

R. Servien<sup>a,b,\*</sup>, K. Bonnot<sup>a,c</sup>, E. Latrille<sup>a,b</sup>, A. Hélias<sup>d,e</sup>, D. Patureau<sup>a</sup>

<sup>a</sup>INRAE, Univ. Montpellier, LBE, 102 Avenue des étangs, F-11000 Narbonne, France

<sup>b</sup>ChemHouse Research Group, Montpellier, France

<sup>c</sup> Université Paris-Saclay, INRAE, AgroParisTech, UMR ECOSYS, 78850, Thiverval-Grignon, France

<sup>d</sup>ITAP, Univ Montpellier, INRAE, Institut Agro, Montpellier, France

<sup>e</sup>ELSA, Research group for environmental life cycle sustainability assessment and ELSA-Pact industrial chair, Montpellier, France

\*corresponding author: [remi.servien@inrae.fr](mailto:remi.servien@inrae.fr)

**Abstract:** During wastewater treatment, incomplete elimination of micropollutants occurs. These micropollutants represent a potential threat to our health and aquatic environment. The potential impacts of these substances are underestimated due to the lack of available concentrations below the limit of quantifications (LOQ) for an important set of micropollutants. Here, we studied the potential impacts on Human health and aquatic environment of the release of 261 organic micropollutants at the scale of France and we showed that, even with a concentration below the LOQ, some micropollutants could have a significant impact. For the unmeasured concentrations, a global concentration distribution build on several datasets with different LOQ is used.

24 The usual convention of LOQ/2 for the unmeasured concentrations is shown to be an  
25 overestimation. Ignoring the unmeasured micropollutants led to an underestimation of  
26 the impacts of more than 300% for both human health and aquatic environment. Some  
27 substances, such as hydrazine, endrin, or 2,3,7,8-TetraCDD, could lead to a very high  
28 impact even with an unmeasured concentration. This work can be adapted to any other  
29 compartment and any other geographical context.

30

31 **Keywords:** characterization factors; limit of quantification; (eco)-toxicity; continental  
32 freshwater; human health; Life Cycle Assessment.

33

34 **Highlights:**

- 35 • Impacts for micropollutants with concentrations under LOQ are derived.
- 36 • Some micropollutants have a significant impact with an unmeasured  
37 concentration.
- 38 • Ignoring these unmeasured substances lead to a strong underestimation of  
39 the impacts.
- 40 • This could help to a prioritization of unmeasured micropollutants.

41

42

43

44

45

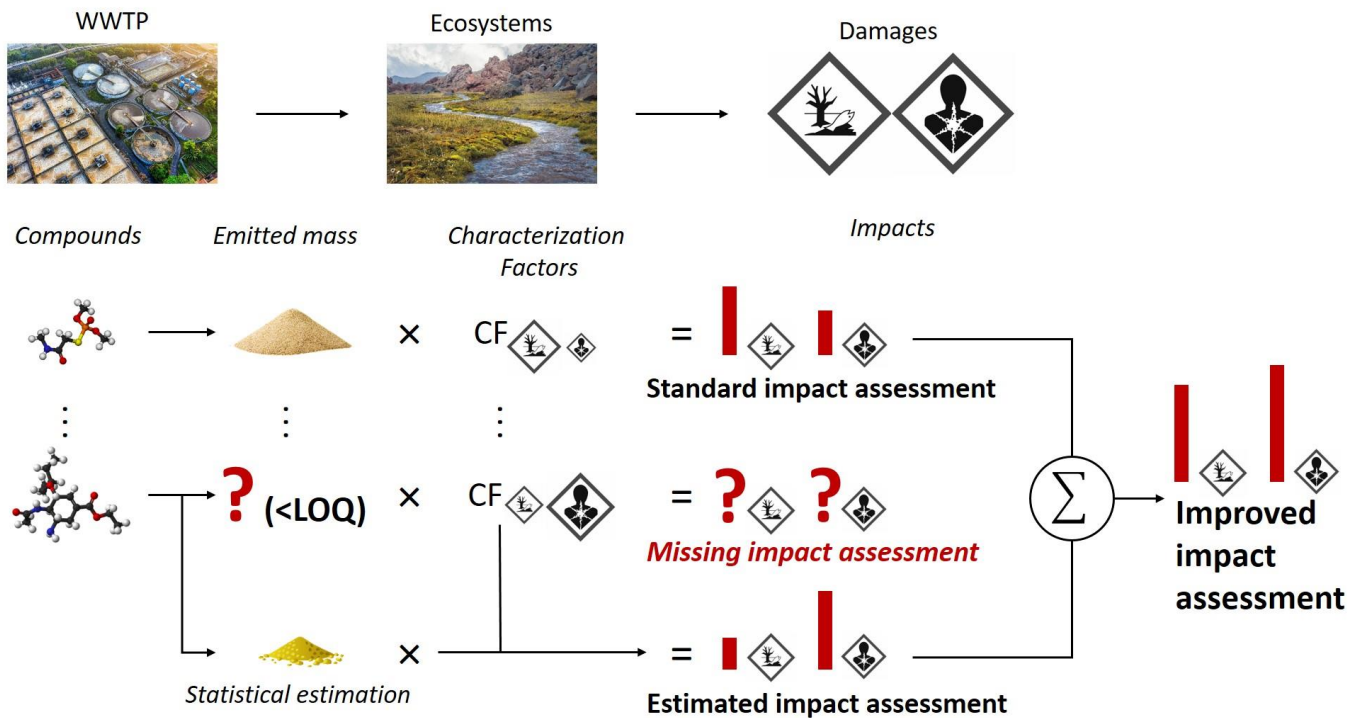
46

47

48

49 **Graphical Abstract :**

50



51

52

53 **1. Introduction**

54 Thanks to the implementation of the Water Framework Directive (Directive  
55 2008/105/CE, 2008), many European countries have launched monitoring campaigns  
56 for organic micropollutants either in the effluents of urban and industrial wastewater  
57 treatment plants (WWTP) or in the receiving rivers. Indeed, many papers refer to the  
58 quantification of thousands of compounds including pharmaceuticals, hormones,  
59 pesticides, hydrocarbons, and solvents in treated wastewaters, underlying incomplete  
60 removals of the conventional treatments (Addamo et al., 2005; Bisognin et al, 2021;  
61 Yao et al., 2021).

62 Moreover, with increasingly accurate and efficient analytical technologies, more and  
63 more compounds are detected at low concentrations ( $\text{ng.L}^{-1}$  to  $\mu\text{g.L}^{-1}$ ) in rivers,

64 groundwater, surface water, and drinking water (Bayer et al., 2014). The incomplete  
65 removal and release of micropollutants into the aquatic environment represent a  
66 potential danger to human health and the aquatic environment of continental  
67 freshwater. The identification and quantification of the potential impact of toxic  
68 substances induce the development of new sustainable process technologies for  
69 targeted substances and represent a great challenge to the safety of aquatic  
70 ecosystems (Rosenbaum et al., 2007; Zhu et al., 2015).

71 Recently, Aemig et al. (2021) identified 261 organic micropollutants in the effluents of  
72 French WWTP. The toxicological and ecotoxicological impacts of 94 and 88  
73 micropollutants, respectively, were quantified by multiplying the emitted mass of the  
74 compound in the total volume of effluents from French WWTPs by their  
75 characterization factors (CFs) (Lindim et al., 2019) as usually done in the Life Cycle  
76 Assessment (LCA) framework. Some compounds cannot be assessed because they  
77 have no known concentration or their CFs were not available. Then, Servien et al.  
78 (2022a) developed a modeling method based on machine learning approaches and 40  
79 easy-to-obtain molecular descriptors to predict missing CFs. This model allowed the  
80 prediction of toxicological and ecotoxicological CFs in continental freshwater with an  
81 acceptable margin of error. This allowed a completed assessment of the overall  
82 potential impacts (on human health and aquatic environment) of 153 organic  
83 micropollutants in continental freshwater at the scale of France. Results showed that  
84 a molecule can be highly hazardous for the aquatic environment without necessarily  
85 being an issue for human health, and vice versa. Moreover, a high impact could be  
86 driven by a high CFs, a high mass, or a combination of both. As the machine learning  
87 models allow a CFs prediction for any new compound, the missing concentrations are  
88 the only burden that prevents a complete impact estimation. Nevertheless, building

89 models based on easy-to-obtain inputs seem not easily attainable for missing  
90 concentrations and the lack of data for some micropollutants is nowadays a known  
91 problem (Leenhardt et al., 2022). Cantoni et al. (2020) tackle this problem by  
92 comparing several statistical methods to estimate concentrations below the LOQ. Their  
93 work is promising but is limited to a few number of micropollutants and cannot be  
94 applied to compounds with no value above the LOQ. Not considering the  
95 micropollutants with a high number of concentrations below the LOQ in the impact  
96 estimations is equivalent to estimating that they have null impacts, which obviously  
97 leads to a strong underestimation of the potential impacts. So, the objective of this  
98 paper was to study the potential impacts of the micropollutants for which we do not  
99 have measured concentrations or for which more than 90% of the measurements are  
100 below the limit of quantification (LOQ) threshold, including those with no measurement  
101 above the LOQ. For this, we develop an estimation procedure detailed in the following  
102 Section. This allowed a complete assessment of the individual potential impacts (on  
103 human health and aquatic environment) of each of the organic micropollutants of  
104 interest in continental freshwater at the scale of France.

105

106

## 107 **2. Materials and Methods**

### 108 **2.1 Molecules**

109 Aemig et al. (2021) identified 261 organic micropollutants representing a potential  
110 danger to human health and aquatic environment in continental freshwater. These  
111 molecules came from (i) the Water Framework Directive (WFD, Directive 2008/105/CE,  
112 2008), (ii) the RSDE national action for the survey and reduction of hazardous  
113 substances in water (INERIS, 2016), and (iii) the AMPERES French project in which

114 micropollutants (registered in the WFD and pharmaceuticals) were analyzed in  
115 influents and effluents of 15 WWTP (Martin Ruel et al., 2016). Among these 261  
116 organic micropollutants, the emitted mass of 153 was estimated with at least 10 sites  
117 of measures and 90% of the measured data above the limit of quantification. The  
118 impacts of these 153 compounds were assessed by Aemig et al. (2021) for those  
119 present in the USEtox<sup>®</sup> database (Rosenbaum et al., 2007) and by Servien et al.  
120 (2022a) for those who are not in this database.

121

## 122 **2.2 Characterization factors**

123 Among the remaining 108 compounds, 58 have a CF<sub>ET</sub> and 43 a CF<sub>HT</sub> in the USEtox<sup>®</sup>  
124 database version 2.12 with a default landscape (to remain consistent with Servien et  
125 al. (2022a,b)). USEtox<sup>®</sup> is an international consensual for characterizing human and  
126 ecotoxicological impacts of chemicals (UNEP-SETAC, 2019). It was developed by the  
127 life cycle initiative under the United Nations Environmental Programme (UNEP) and  
128 the Society for Environmental Toxicology and Chemistry (SETAC) (Henderson et al.,  
129 2011) to produce a transparent and consensus characterization model. This model  
130 gathers in one single characterization factor the chemical fate, the exposure, and the  
131 effect for each of the several thousands of organic and inorganic compounds. If the  
132 structure of the USEtox<sup>®</sup> multimedia model is always the same, to determine the CF  
133 of a molecule, numerous physicochemical parameters (such as solubility,  
134 hydrophobicity, degradability) and detailed toxicological and ecotoxicological data  
135 must be provided.

136 The remaining CFs were computed using the machine learning models developed by  
137 Servien et al. (2022a). These models predict ecotoxicological or toxicological  
138 characterization factors using 40 selected easy-to-obtain molecular descriptors (see

139 the list in Table S1). For more details on the choice and the calculation of these  
140 descriptors, the interested reader is referred to Servien et al. (2014). These models  
141 are based on a comparison between different linear or non-linear (machine learning)  
142 models and could be global or local using a cluster-then-predict approach. The  
143 modeled CFs were compared to USEtox<sup>®</sup> database, the differences between the  
144 modeled CFs and the USEtox<sup>®</sup> ones were small, and the models were therefore  
145 considered to be applicable. The best approaches were then selected for each cluster  
146 and each characterization factor. To apply the cluster-then-predict approaches, the  
147 cluster of each new compound needs to be estimated using a supervised clustering  
148 approach based on the clusters already obtained in Servien et al. (2022a) which was  
149 performed using the well-known k-nearest neighbor approach through the *knn* function  
150 of the R package *class* (Ripley, 1996). Then, the models selected in Servien et al.  
151 (2022a) were applied without any modification, assuming they had been previously  
152 tested on a large diversity of molecules covering the diversity of the compounds  
153 assessed in the present study. These previous tests have shown the reliability of the  
154 models to predict CFs that are missing in USEtox<sup>®</sup>.

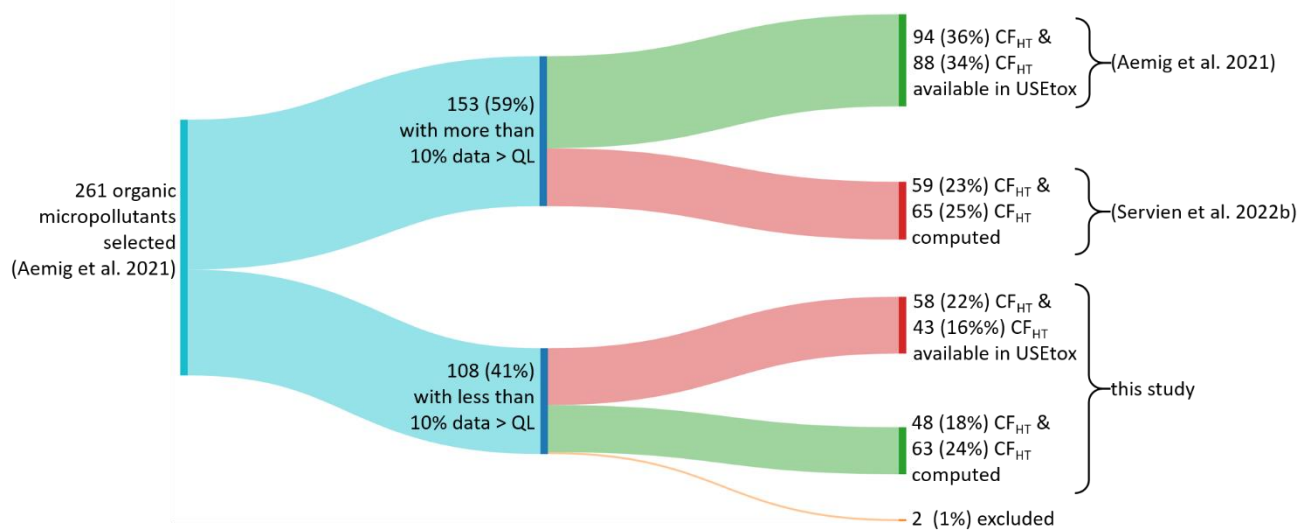
155 Note that the tributylstannanylium (CAS 36643-28-4) is made of tin and, as  
156 consequence, its molecular descriptors cannot be computed by the above-mentioned  
157 methods and, thus, its CFs. Also, chloroalkanes C10-13 (CAS 85535-84-8) is a mix of  
158 different molecular structures that can lead to different molecular descriptors and CFs  
159 and was also excluded.

160 The 106 corresponding characterization factors are gathered in Table S2. The scheme  
161 of the whole study is summarized in the following Figure 1.

162

163





164

165 **Figure 1.** Position of this study in the general context.

166

167 **2.3 Quantification of significative potential impacts**

168 The annual volume of water effluent released into the environment by WWTP was  
 169 estimated to be  $5.10^9 \text{ m}^3$ , in France (Aemig et al., 2021). This volume was determined  
 170 by multiplying the daily flow arriving at the WWTP for 365 days, assuming that it was  
 171 equal to the effluent. The emitted mass of micropollutants was estimated by multiplying  
 172 each concentration of the micropollutants, when available, by the volume of the  
 173 effluent. The concentrations and masses are those of the Supplemental Material of  
 174 Aemig et al. (2021) and are gathered in Table S3. Total impacts on human health and  
 175 aquatic environment in continental freshwater were quantified by summing the impacts  
 176 of all the compounds, as is usually done in LCA (Heijungs and Suh, 2002).

177 As explained above, the remaining 106 micropollutants are those for which not enough  
 178 data were available to correctly estimate the global concentration. So, for these  
 179 compounds, we chose to calculate the minimal concentration to reach 1% of the total  
 180 sum of impacts calculated previously for the 153 micropollutants presented in Servien  
 181 et al. (2022b). When the LOQ of a micropollutant was available, this minimal

182 concentration was also expressed in percentage of the LOQ to be comparable among  
183 different micropollutants. For 42 compounds, no data (measurements and LOQ) were  
184 available so this minimal concentration was only expressed in  $\mu\text{g.L}^{-1}$ . Obviously, as the  
185 potential impacts are linearly related to the estimated concentration, if a measured  
186 concentration is the minimal concentration multiplied by a constant  $k$ , the  
187 corresponding impacts will be  $k\%$  of the total sum of impacts calculated previously.

188

#### 189 **2.4 Estimation of values under the LOQ**

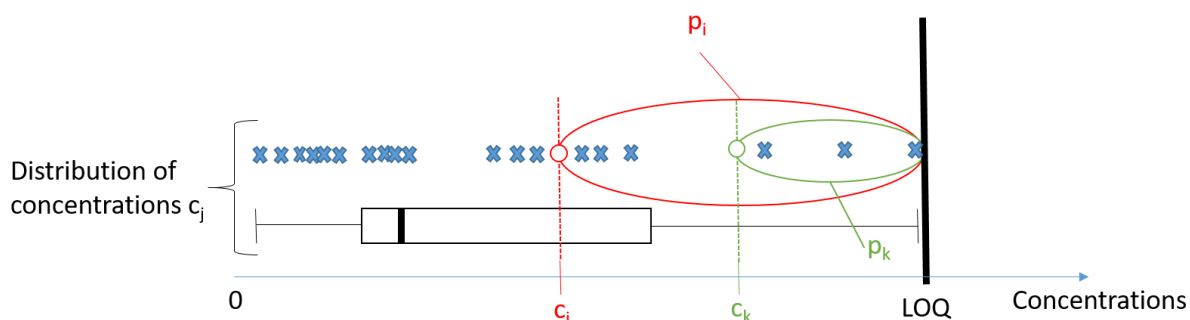
190

191 For these 106 micropollutants, according to Table S3, a very high number (more than  
192 90%) of the values are below the LOQ of the corresponding study. Data inferior to LOQ  
193 are usually estimated at half of the quantification limit as usually applied in  
194 environmental studies (INERIS, 2016). This assumption is a very strong one that could  
195 have a major impact on the estimated concentration (and thus the potential impact)  
196 when there are a lot of concentrations below LOQ. Testing this strong assumption  
197 could be of great help.

198 In the RSDE datasets, the compounds are measured using routine methods so most  
199 of the LOQ are pretty high. This is not always the case for the other sources of  
200 quantification data. So, we have some quantifications that came from other sources  
201 and that are below the LOQ of the main dataset of the study, the RSDE one. These  
202 values are then scaled relative to the RSDE LOQ and are gathered for all compounds  
203 to have enough data. This gives a dataset of 141 samples with concentrations  
204 denoted  $(c_j)_{1 \leq j \leq 141}$  coming from 25 different compounds. The distribution of these  
205 samples is then used to compute the probability that the minimal concentration  $c_i$  of  
206 compound  $i$  has been reached in case of quantification below the LOQ. This probability

207  $p_i$  is computed as the number of measurements above the corresponding minimal  
 208 concentration  $c_i$  divided by the total number of measurements (i.e. 141):  $p_i =$   
 209  $\frac{\sum_{j=1}^{141} \mathbb{1}(c_i < c_j)}{141}$  and is detailed in the following Figure 2. In other terms, when the minimal  
 210 concentration  $c_i$  is near the LOQ, the probability to get a value above  $c_i$  is low.

211



212

213 **Figure 2:** The distribution of concentrations  $(c_j)_{1 \leq j \leq 141}$  is represented here by some  
 214 blue crosses and its corresponding boxplot. A new minimal concentration  $c_i$  is  
 215 estimated for a compound  $i$ , this results in the corresponding  $p_i$ , calculated by counting  
 216 all the concentrations  $c_j > c_i$ . Another minimal concentration  $c_k$  is estimated for another  
 217 compound  $k$  with  $c_k > c_i$ . Obviously, the probability  $p_k$  that the minimal concentration  
 218 for the compound  $k$  will be attained is below  $p_i$ .

219 Note that if a value is below the two LOQ (the RSDE one and the one of the other  
 220 dataset) it won't be taken into account, so the final distribution will probably be a little  
 221 overestimated. Indeed, taking into account these unknown measurements relying on  
 222 different relative LOQ values would bring a lot of uncertainties in the below LOQ  
 223 distribution calculations.

224

## 225 **2.5 Calculation of the total impacts**

226

227 As usually done in LCA, the total impacts on human health and aquatic environment  
228 in continental freshwater were quantified by summing the impacts of all the  
229 compounds. Human impact is expressed in DALY representing the number of  
230 negatively impacted human years, and the ecotoxicological impact is expressed in  
231 PDF representing the potential fraction of disappeared species. To compute these  
232 impacts, we impute values under the LOQ by the median values under the LOQ for all  
233 the compounds (see Figure 3), namely LOQ/3. For the compounds without a LOQ, the  
234 median value of the LOQ, namely  $0.05 \mu\text{g.L}^{-1}$ , was used.

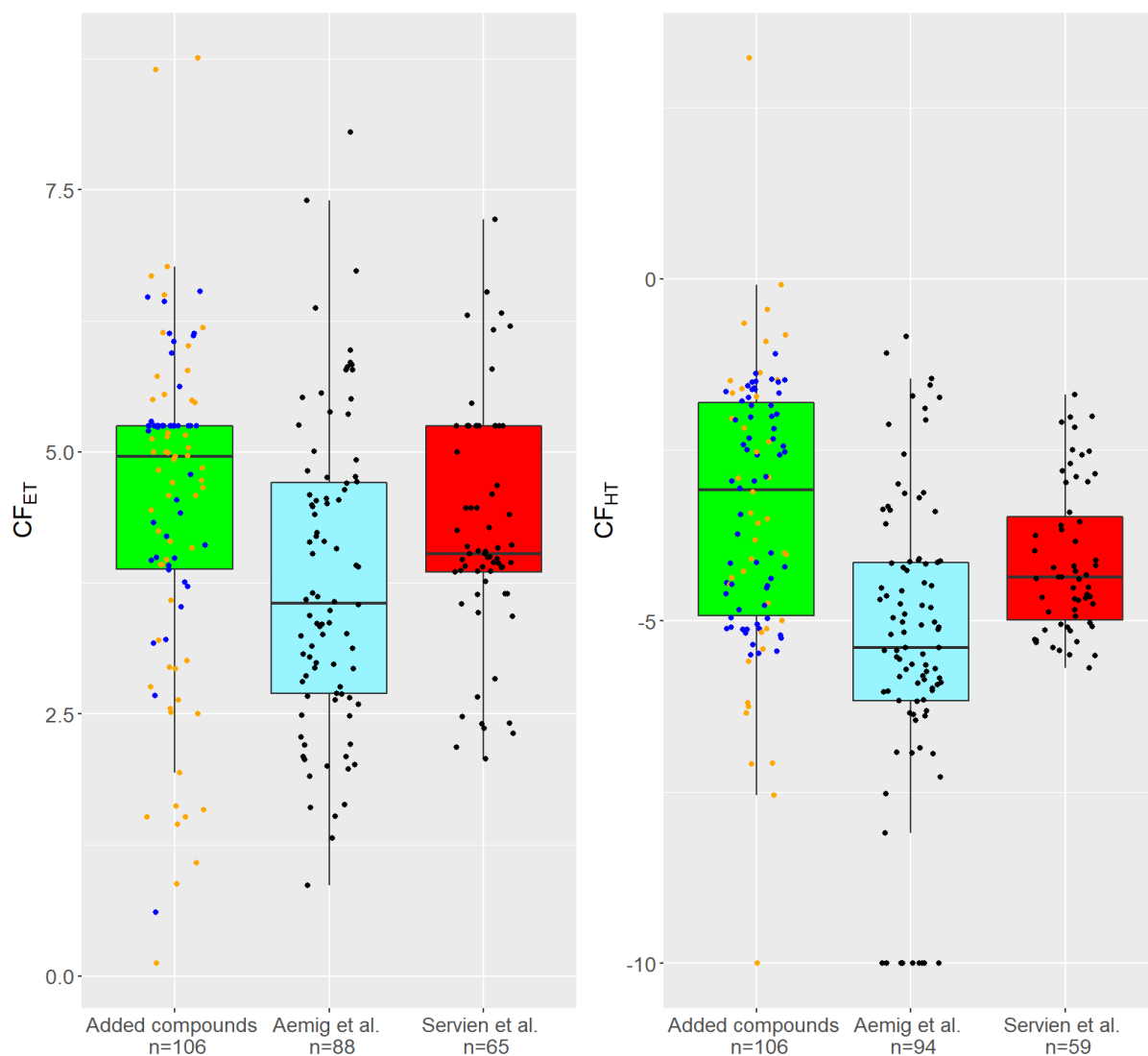
235

## 236 **3. Results and discussion**

### 237 **3.1 Global analysis of characterization factors**

238 Regarding the CFs, the different values are reported in the boxplots of Figure 3 below.  
239 As usual, the CFs are log-transformed with the addition of  $1\text{e-}10$  for the  $\text{CF}_{\text{HT}}$  to avoid  
240 a computational problem with  $\log_{10}(0)$ .  $1\text{e-}10$  has been chosen to be below the  
241 minimum of the USEtox<sup>®</sup> database ( $5\text{e-}9$ ).

242



243

244

245

246 **Figure 3.** Comparison of the  $CF_{HT}$  and the  $CF_{ET}$  (both in log10 scale) for the  
 247 compounds taken into account in previous studies and the added compounds. For the  
 248 added compounds, the CF in orange came from USEtox<sup>®</sup> whereas those in blue are  
 249 computed using Servien et al. (2022a).

250

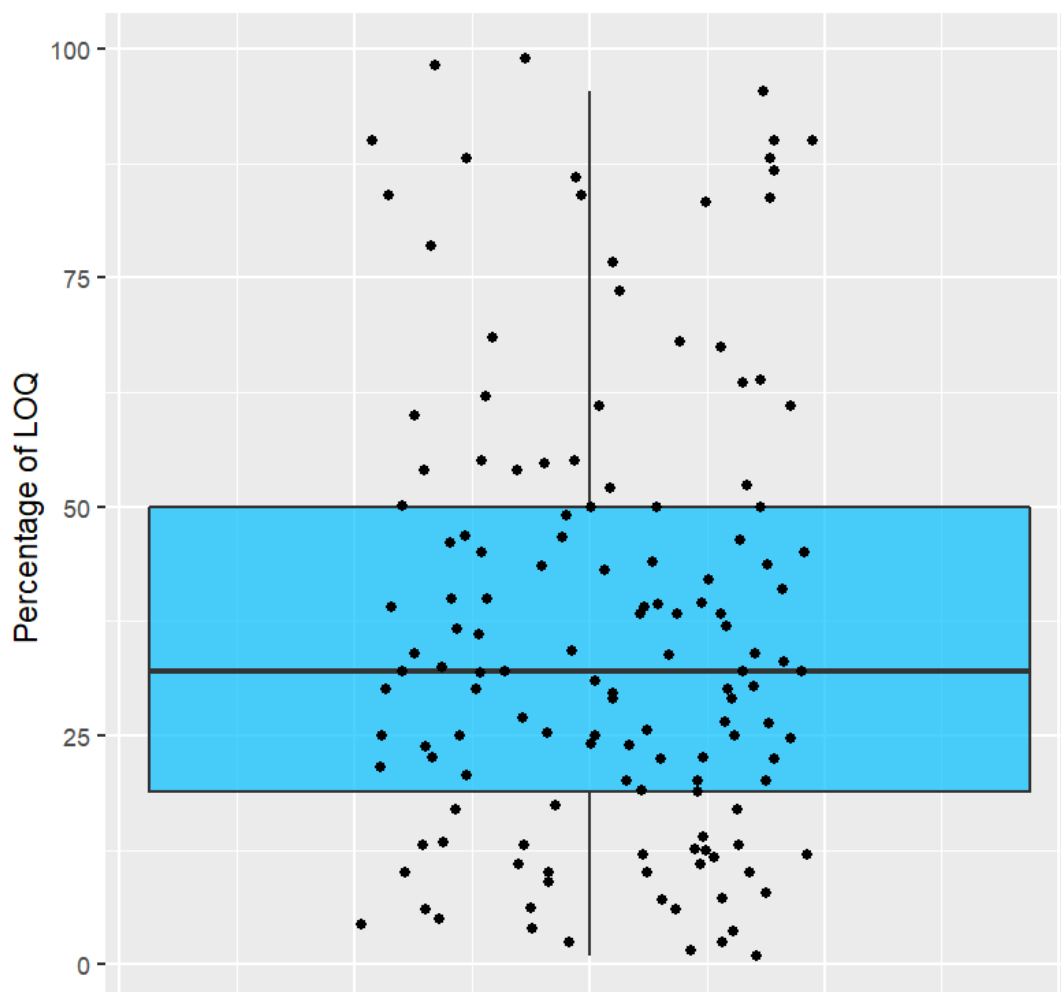
251 Figure 3 suggests that the added compounds have globally higher CF than already  
 252 included in the previous studies. In more detail, for the  $CF_{HT}$  and the  $CF_{ET}$ , the

253 maximum, the 1<sup>st</sup> and 3<sup>rd</sup> quartile, the median and the mean of the added compounds  
254 are above those of the previous studies. But there are also very low CF values largely  
255 below those of the previous study. All the descriptive values are given in Table S4. We  
256 could also see that we have extremely high values for some compounds already  
257 included in USEtox<sup>®</sup>. For the CF<sub>ET</sub>, these two extreme values are for the 1,2,3,7,8-  
258 PentaCDD (CAS 40321-76-4) and the 2,3,7,8-TetraCDF (CAS 51207-31-9) from the  
259 polychlorodibenzodioxin/furane family. Their very extreme values would lead to  
260 significant impacts if their concentrations are not too low. For the CF<sub>HT</sub>, the very  
261 extreme value is for the 2,3,7,8-TetraCDD (CAS 1746-01-6). This is the unique value  
262 of the whole USEtox<sup>®</sup> database that is log-positive for this variable. Nevertheless, the  
263 estimation of the CF<sub>HT</sub> available in the environmental footprint database (Fazio et al.,  
264 2018) seems to corroborate this very high value. This dioxin is a well-studied and  
265 recognized endocrine disrupter presenting the highest affinity to the aryl hydrocarbon  
266 receptor (AhR), a cytosolic protein present in most vertebrate tissues, that mediated  
267 the toxic and biological effects of such compounds (Van den Berg et al., 2006). Its very  
268 high CF<sub>ET</sub> and CF<sub>HT</sub> are linked to its high endocrine-disrupting potency. If we don't  
269 consider this compound, five others have CF<sub>HT</sub> above the maximal value already  
270 included in previous studies. These compounds are all pesticides: chlordane (CAS 57-  
271 74-9), 1,1-bis(4-Chlorophenyl)-2,2-Dichloroethane (CAS 72-54-8), dieldrin (CAS 60-  
272 57-1), heptachlor epoxide (CAS 1024-57-3) and toxaphene (CAS 8001-35-2).

273

### 274 **3.2 Estimation of concentrations under the LOQ**

275 In the following Figure 4, the distribution of the concentrations measured below the  
276 RSDE LOQ is represented.



277

278 **Figure 4:** Boxplot of the 141 measured values below the RSDE LOQ on another  
 279 dataset with lower LOQ (expressed in % of the RSDE LOQ).

280 We could see that the median (and so is the mean) value of this distribution seems to  
 281 be close to  $LOQ/3$ , and not  $LOQ/2$  as usually used for calculation when data are below  
 282 the LOQ. Furthermore, as the concentration values under the smaller LOQ are not  
 283 taken into account, even  $LOQ/3$  is probably an overestimation, when faced with an  
 284 under LOQ value in the RSDE dataset. This is also confirmed by the conclusions of  
 285 Cantoni et al. (2020) where the LOQ and  $LOQ/2$  are shown to be overestimation.

286

287 **3.3 Impact on the aquatic environment**

288 For the aquatic environment, all the 64 minimal concentrations that could be expressed  
 289 in percentage of LOQ are gathered in Table S6 and the smallest (i.e. that are below  
 290 the LOQ) are reported in Table 1.

291 **Table 1** – Estimation of the concentrations needed to reach 1% of the global impacts  
 292 on ecotoxicity, when the minimal concentration is below the RSDE LOQ. The  
 293 probability is calculated when the percentage of LOQ is under 100%. It is computed  
 294 using the distribution of subsection 3.2 and is the estimation that a measurement under  
 295 the LOQ is above the minimal concentration. The results for all compounds are  
 296 gathered in Table S6.

CAS number	Name	Minimal concentration ( $\mu\text{g}\cdot\text{L}^{-1}$ )	Minimal concentration (% of RSDE LOQ)	Probability
302-01-2	Hydrazine	4,0E+00	4%	0.96
72-20-8	Endrin	1,0E-02	19%	0.75
143-50-0	Chlordecone	9,0E-02	46%	0.29
2921-88-2	Chlorpyrifos	2,0E-02	88%	0.04

297  
 298 There are only 4 compounds for which a concentration under the LOQ could still cause  
 299 a significant impact on ecotoxicity and for 2 of them the minimal concentration is also  
 300 below LOQ/3. For the 60 other compounds, the minimal concentration being above the  
 301 LOQ and the measured one frequently below, their contribution to the environmental  
 302 impact could be considered as very low. For the hydrazine, a very small concentration  
 303 of only 4% of the RSDE LOQ, that have a high probability of occurrence, will lead to a  
 304 significant impact. Hydrazine and its derivatives are widely used compounds in  
 305 chemical industries and pharmaceuticals (propellant, blowing agent, pesticide,  
 306 antituberculosis...). It is a very small molecule (32 g/mol), very soluble and volatile, no  
 307 BCF is found in literature, nor DT50. It was classified by US EPA as a human  
 308 carcinogen (group B2): an intermediate inhalation minimal risk level of 4  $\mu\text{g}/\text{L}$  was  
 309 calculated and a minimum limit for cancer was fixed at 10  $\mu\text{g}/\text{L}$  (Qi et al., 2020). The



310 minimal concentration was found to be 4 µg/L and is largely below the RSDE LOQ,  
 311 which is very high (100 µg/L) by comparison to the one of endrin (0,052 µg/L) and also  
 312 others compounds. However, very recent papers refer to the development of  
 313 spectroscopic methods and sensors to analyze hydrazine in environmental samples  
 314 including water up to pico-molar level (0.03 to 3.2 µg/L, Arulraj et al., 2015; Qi et al.,  
 315 2020). However, no data are provided in real water samples.

316 For the three pesticides, due to their hydrophobicity, they are usually quantified in solid  
 317 matrices like soil, sediment, and biota. However, concentrations of endrin were  
 318 reported in river water between 0.17 and 484 ng/L (Matsumo et al., 2009) and  
 319 concentrations of chlorpyrifos between 1 and 32 ng/L (Wan et al., 2021). Their  
 320 presence in such environment is mainly due to their use in agriculture. They are also  
 321 targeted by the WFD. The Environmental Quality Standard (yearly average value) for  
 322 the sum of the 4 cyclodiene pesticides in the WFD (aldrin, isodrin, endrin, dieldrin) is  
 323 10 ng/L (inner water surface) and 5 ng/L (other water). For chlorpyrifos, the EQS is 30  
 324 ng/L. The three calculated minimal concentrations to reach 1% of the global impact are  
 325 in the same range of these environmental and regulatory values.

326 The compounds without any information on LOQ are gathered in the following Table  
 327 3.

328

329 **Table 3.** Minimal concentrations for each compound to reach 1% of the global impacts  
 330 on ecotoxicity. The compounds in this table are those without any measure or LOQ.

CAS number	Name	Minimal concentration (µg.L <sup>-1</sup> )
40321-76-4	1,2,3,7,8-P5CDD	9,62E-05
51207-31-9	2,3,7,8-T4CDF	1,23E-04
1746-01-6	2,3,7,8-T4CDD	1,16E-02
57653-85-7	1,2,3,6,7,8-H6CDD	3,61E-02
72-54-8	DDD 44'	4,04E-02
28159-98-0	Cybutryne	5,31E-02
72-55-9	DDE 44'	1,56E-01

19408-74-3	1,2,3,7,8,9-H6CDD	3,08E-01
39227-28-6	1,2,3,4,7,8-H6CDD	3,08E-01
57117-44-9	1,2,3,6,7,8-H6CDF	3,08E-01
70648-26-9	1,2,3,4,7,8-H6CDF	3,08E-01
72918-21-9	1,2,3,7,8,9-H6CDF	3,08E-01
57117-41-6	1,2,3,7,8-P5CDF	3,09E-01
3268-87-9	1,2,3,4,6,7,8,9-O8CDD	3,10E-01
35822-46-9	1,2,3,4,6,7,8-H7CDD	3,10E-01
39001-02-0	1,2,3,4,6,7,8,9-O8CDF	3,10E-01
446255-20-5	Heptabromodiphenylethers	3,10E-01
55673-89-7	1,2,3,4,7,8,9-H7CDF	3,10E-01
57117-31-4	2,3,4,7,8-P5CDF	3,10E-01
60851-34-5	2,3,4,6,7,8-H6CDF	3,10E-01
67562-39-4	1,2,3,4,6,7,8-H7CDF	3,10E-01
789-02-6	DDT 24'	5,58E-01
2032-65-7	Methiocarb	6,08E-01
25167-83-3	Tetrachlorophenols	6,50E-01
131860-33-8	Azoxystrobin	1,43E+00
139968-49-3	Metaflumizone	1,59E+00
118-79-6	2,4,6-tribromophenol	1,98E+00
564-25-0	Doxycycline	2,11E+00
153719-23-4	Thiamethoxam	2,61E+00
111988-49-9	Thiacloprid	4,24E+00
2303-17-5	Triallate	5,88E+00
21312-10-7	Acetyl sulfamethoxazole	6,67E+00
5466-77-3	Octinoxate	7,36E+00
615-58-7	2,4-dibromophenol	9,62E+00
135410-20-7	Acetamiprid	1,05E+01
210880-92-5	Clothianidin	1,65E+01
14297-93-9	4-epi-chlortetracycline	3,37E+01
25167-80-0	Monochlorophenol	3,67E+01
25167-82-2	Trichlorophenols	5,41E+01
128-37-0	Butylhydroxytoluene	6,19E+01
57-62-5	Chlortetracycline	1,27E+02
95-56-7	2-bromophenol	1,55E+02

331

332 For these compounds, only the minimal concentration in the usual unity ( $\mu\text{g.L}^{-1}$ ) is  
333 available. It leads to results that are more difficult to interpret but we could see that  
334 there are 2 compounds from the dioxin/furane family (1,2,3,7,8-PentaCDD and 2,3,7,8-  
335 TetraCDF) for which a concentration above 0,01 and 0,1  $\text{ng.L}^{-1}$  will be enough to lead  
336 to a significant impact. The 15 other congeners of this family present minimal

337 concentration between 11 and 310 ng/L. These compounds, produced during waste  
338 incineration, and chemical manufacturing, are classified as very hydrophobic and  
339 persistent compounds, and accumulate in environmental compartments such as soil,  
340 and sediment as well as in biota (Srogi, 2008). As an example, Villanneau et al (2011)  
341 found soil concentrations of 23,8 ng I-TEQ/kgDM for the  $\Sigma$ 17 PCDD/F. The regulatory  
342 limit in meats for the sum of dioxin and furan is 1 to 5 ng I-TEQ / kg of fat. In such  
343 organic and fat matrices, low concentrations can be found, but the probability to find  
344 such low concentrations in water matrices may be very low. Dioxins and furans are  
345 also on the list of priority substances of the WFD but no EQS for water is provided only  
346 EQS for the biota is fixed at 6,5 ng/kg DW (I-TEQ for 7 PCDD+10 PCDF+12 PCB).  
347 Such low concentrations could probably be reached locally (see Lu et al., 2016, for  
348 PCDD/F concentrations in raw and treated drinking water and Ellis et al., 2008, for data  
349 on lake Chelan).

350 DDD 44' and DDE 44' are transformation products of DDT that is no more used in  
351 western countries since the 90th. All congeners are hydrophobic and persistent  
352 compounds with accumulation in biota (Garrison et al., 2014). Total DDT, including  
353 several DDT isomers, DDE, and DDD, is listed on the WFD with an EQS of 25 ng/L  
354 implying its monitoring in the French RSDE program.

355 Cybutryne is a biocide from the triazine family and is used mainly in anti-fouling paints  
356 for the protection of boat hulls. It was included on the WFD priority list for monitoring  
357 and control in 2012 with an EQS of 2.5 ng/L (yearly average) and a maximal admissible  
358 concentration of 16 ng/L. In Switzerland, Ollers et al. (2001) have quantified cybutryne  
359 in effluent at very low frequencies around 15 ng/l, which is below the calculated minimal  
360 concentration (53 ng/L).

361 At the bottom of Table 3, the listed compounds present high minimal concentrations  
362 ranging from 10 to 155 µg/L. As such, chlorophenols, present on EU and US EPA  
363 priority lists, are widely used for dyestuff, pharmaceutical and pesticide synthesis and  
364 are also by-products of their degradation in the environment (Ganigar et al., 2010).  
365 These compounds were mainly monitored in industrial wastewaters where they are  
366 treated by biological and physicochemical processes before discharge in environment  
367 or domestic sewage; this could explain why no concentration of such compound was  
368 found in the domestic effluent of the French studied WWTP. Gao et al. (2008) have  
369 quantified 2,4,6-trichlorophenol in 54.4 % of their river water samples (600 sites) at a  
370 median value of 2 ng/L. These low measured values suggest also the existence of low  
371 limit of quantifications as the one given by Suliman et al. (2006) around µg/L for TCP.  
372 These data suggest that it may not be realistic to get such minimal concentration in  
373 effluent. To reinforce this idea, we can have a look on compounds like chlortetracycline  
374 and its metabolite, 4-epi-chlortetracycline, whose minimal concentrations to reach 1%  
375 of the total ecotoxicological impact are respectively 127 and 33 µg/L: these compounds  
376 were found in the effluent of many WWTP but at very lower concentrations than the  
377 minimal one. For example, Li and Zhang (2011) measured chlortetracycline levels  
378 around 24 and 33 ng/L in secondary effluents of 2 WWTP in Hong Kong.  
379 Chlortetracycline concentrations in effluents of American WWTPs (<8.1–60 ng L<sup>-1</sup>)  
380 (Yang et al., 2005; Spongberg and Witter, 2008) and Australian WWTPs (5–250 ng L<sup>-</sup>  
381 <sup>1</sup>) (Watkinson et al., 2009) were also found.

382

### 383 **3.4 Impact on Human health**

384 For human toxicity, all the 64 minimal concentrations that could be expressed in  
 385 percentage of LOQ are gathered in Table S5 and the smallest (i.e. that are below their  
 386 LOQ) are reported in the following Table 2.

387 **Table 2** – Estimation of the concentrations needed to reach 1% of the global impacts  
 388 on human toxicity, when the minimal concentration is below the RSDE LOQ. The  
 389 probability is calculated when the percentage of LOQ is under 100%. It is computed  
 390 using the distribution of subsection 3.2 and is the estimation that a measurement under  
 391 the LOQ is above the minimal concentration. When the probability is equal to one, it  
 392 means that no measured value is below this percentage of LOQ. The results for all  
 393 compounds are gathered in Table S5.

CAS number	Name	Minimal concentration ( $\mu\text{g.L}^{-1}$ )	Minimal concentration (% of RSDE LOQ)	Probability
302-01-2	Hydrazine	2,1E-01	0.2%	1
1024-57-3	Heptachlor epoxide	8,2E-05	0.4%	1
8001-35-2	Toxaphene	2,9E-04	0.6%	1
143-50-0	Chlordecone	1,6E-03	0.8%	1
60-57-1	Dieldrin	4,4E-04	0.9%	1
57-74-9	Chlordane	5,5E-04	2.8%	0.97
309-00-2	Aldrin	2,0E-03	4.0%	0.96
2385-85-5	Mirex	2,7E-03	5.4%	0.94
72-20-8	Endrin	3,5E-03	7.0%	0.91
62-53-3	Aniline	6,6E+00	13%	0.79
76-44-8	Heptachlor	3,1E-03	15%	0.77
50-29-3	DDT 44'	1,0E-02	20%	0.71
50-32-8	Benzo(a)pyrene	2,1E-03	21%	0.7
470-90-6	Chlorfenvinphos	2,3E-02	23%	0.67
67-56-1	Methanol	2,4E+03	24%	0.67
36483-60-0	Hexabromodiphenylethers	6,2E-03	31%	0.52
36355-01-8	Hexabromobiphenyl	6,9E-03	35%	0.44
75-01-4	Chloroethene	1,9E+00	38%	0.42
32536-52-0	Octabromodiphenylethers	2,5E-02	50%	0.26
120-12-7	Anthracene	2,3E-02	56%	0.18
39635-31-9	PCB 189	2,9E-03	59%	0.18
124495-18-7	Quinoxifen	6,0E-02	60%	0.18
118-74-1	Hexachlorobenzene	7,2E-03	72%	0.12
3424-82-6	DDE 24'	7,6E-02	76%	0.11
87-68-3	Hexachlorobutadiene	4,3E-01	86%	0.06

32774-16-6 PCB 169 4,7E-03 94% 0.02

394

395 More compounds (26) could have a significant impact here if we compare with the  
 396 previous subsection on environmental impact and 16 have a minimal concentration  
 397 below LOQ/3 and will let to a significant increase in the total sum of impacts if they  
 398 were taken into account. Among the four previous compounds (hydrazine, endrin,  
 399 chlordecone, chlorpyrifos) with a significant ET impact, only chlorpyrifos is not present  
 400 here. Many of these compounds are pesticides, hydrocarbons, and fire retardants that  
 401 are persistent and hydrophobic compounds usually quantified in soil, sediment, and  
 402 biota for which the minimal concentration is between 10  $\mu\text{g.L}^{-1}$  (heptachlor epoxide)  
 403 and 100  $\text{ng.L}^{-1}$  (hexachlorobutadiene). There are also smaller molecules like aniline,  
 404 chloroethene, and methanol, used as solvents and intermediate chemicals for dye and  
 405 pharmaceutical industries whose minimal concentrations are higher between  $\mu\text{g.L}^{-1}$   
 406 and  $\text{mg.L}^{-1}$  (methanol). These compounds are present in industrial wastewater at very  
 407 high concentrations and are quite well degraded in WWTP (Chaturvedi and Katoch,  
 408 2020), suggesting low residual concentrations in urban effluents. The compounds  
 409 without any information on LOQ are gathered in Table 4.

410

411 **Table 4.** Minimal concentrations for each compound to reach 1% of the global HT  
 412 impacts. The compounds in this table are those without any measure or LOQ.

CAS number	Name	Minimal concentration ( $\mu\text{g.L}^{-1}$ )
1746-01-6	2,3,7,8-T4CDD	3,97E-08
72-54-8	DDD 44'	1,87E-04
39001-02-0	1,2,3,4,6,7,8,9-O8CDF	8,37E-04
3268-87-9	1,2,3,4,6,7,8,9-O8CDD	1,60E-03
55673-89-7	1,2,3,4,7,8,9-H7CDF	1,93E-03
35822-46-9	1,2,3,4,6,7,8-H7CDD	2,02E-03
67562-39-4	1,2,3,4,6,7,8-H7CDF	2,10E-03
57653-85-7	1,2,3,6,7,8-H6CDD	2,14E-03
39227-28-6	1,2,3,4,7,8-H6CDD	2,18E-03

72918-21-9	1,2,3,7,8,9-H6CDF	2,41E-03
19408-74-3	1,2,3,7,8,9-H6CDD	2,65E-03
70648-26-9	1,2,3,4,7,8-H6CDF	2,73E-03
57117-44-9	1,2,3,6,7,8-H6CDF	2,78E-03
57117-41-6	1,2,3,7,8-P5CDF	3,11E-03
60851-34-5	2,3,4,6,7,8-H6CDF	3,56E-03
57117-31-4	2,3,4,7,8-P5CDF	4,12E-03
40321-76-4	1,2,3,7,8-P5CDD	4,72E-03
51207-31-9	2,3,7,8-T4CDF	1,03E-02
446255-20-5	Heptabromodiphenylethers	1,45E-02
72-55-9	DDE 44'	1,59E-02
789-02-6	DDT 24'	2,09E-02
118-79-6	2,4,6-tribromophenol	3,53E-01
25167-83-3	Tetrachlorophenols	9,29E-01
131860-33-8	Azoxystrobin	9,47E-01
139968-49-3	Metaflumizone	1,08E+00
2303-17-5	Triallate	1,56E+00
564-25-0	Doxycycline	1,59E+00
615-58-7	2,4-dibromophenol	1,82E+00
14297-93-9	4-epi-chlortetracycline	1,92E+00
57-62-5	Chlortetracycline	2,05E+00
25167-82-2	Trichlorophenols	2,66E+00
28159-98-0	Cybutryne	8,56E+00
2032-65-7	Methiocarb	8,65E+00
95-56-7	2-bromophenol	8,79E+00
111988-49-9	Thiacloprid	9,89E+00
153719-23-4	Thiamethoxam	1,07E+01
210880-92-5	Clothianidin	1,17E+01
25167-80-0	Monochlorophenol	1,46E+01
128-37-0	Butylhydroxytoluene	1,70E+01
5466-77-3	Octinoxate	1,81E+01
135410-20-7	Acetamiprid	1,95E+01
21312-10-7	Acetyl sulfamethoxazole	2,05E+01

413

414 As seen previously, the very high value for the  $CF_{HT}$  for the 2,3,7,8-TetraCDD lead to  
415 a very low value for its minimal concentration. Many other compounds of the same  
416 family only need around  $1 \text{ ng.L}^{-1}$  to have significant impacts. However, as already  
417 discussed for environmental impact, such low concentration will be measured with low  
418 probability in aqueous media. Other compounds as doxycycline, chlortetracycline, and  
419 acetyl sulfamethoxazole were already quantified in effluents of WWTP but at

420 concentrations largely below the calculated minimal concentrations (Verlicchi et al.,  
421 2012).

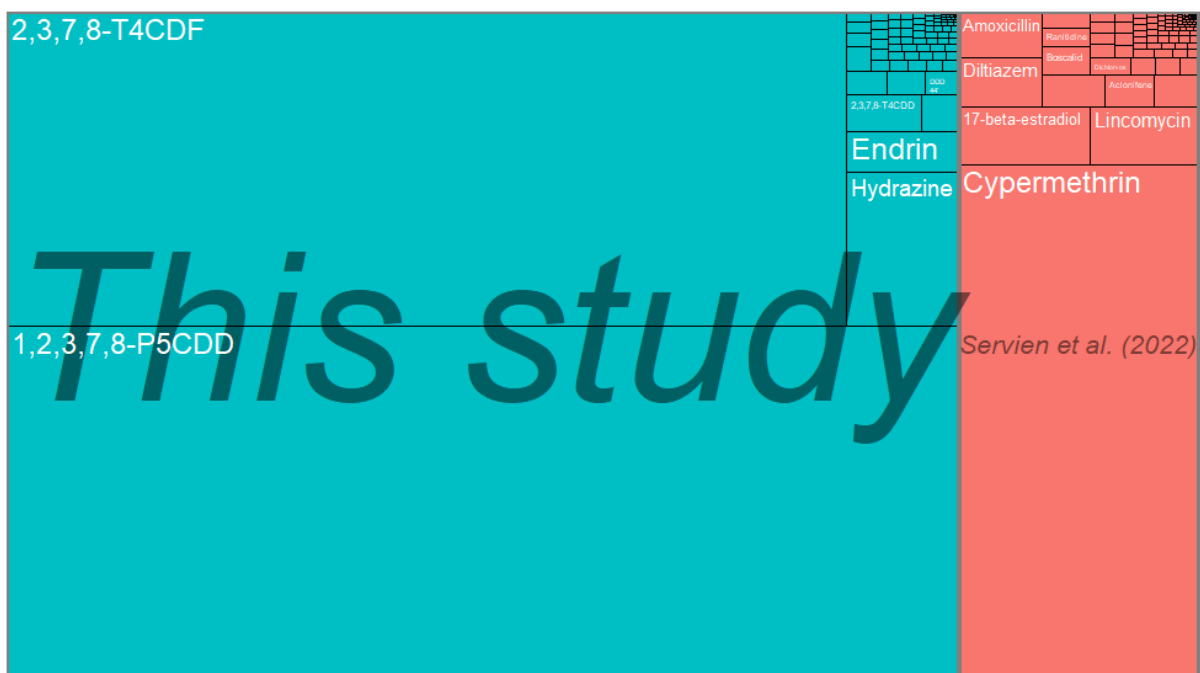
422

### 423 3.5 Influence on the sum of the impacts

424

425 According to all the assumptions made previously we were able to compare the sum  
426 of the impacts of the 106 micropollutants of this study to the 153 of Servien et al.  
427 (2022b). For the environmental impacts, the sum of the impacts of this study was nearly  
428 4 times higher than the previous one. This could be seen in the following Figure 5  
429 where the environmental impacts of the 259 compounds are represented.

430



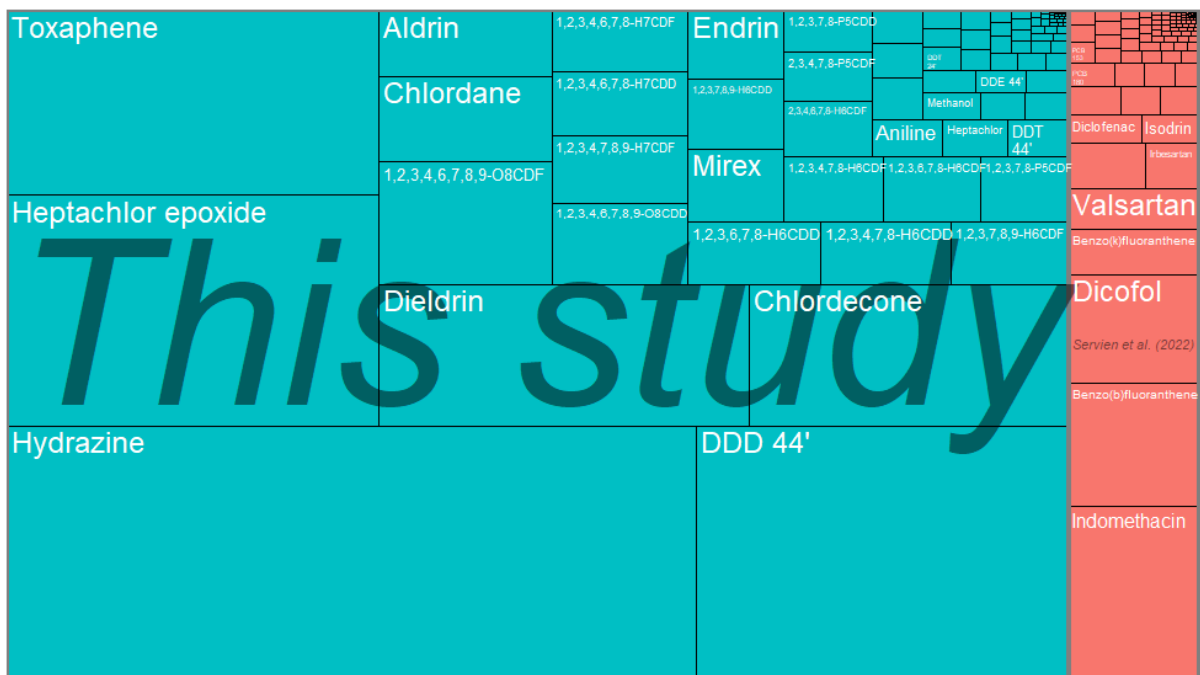
431

432 **Figure 5.** Graphical representation of the environmental impacts of the 259  
433 compounds studied in Servien et al. (2022b) (in red) and this study (in blue). Each  
434 compound has an estimated environmental impact proportional to its rectangular area.

435



436 For the impacts on human health it was more than 8 times higher, even without taking  
 437 into account the molecule 2,3,7,8-TetraCDD. This unique molecule has an impact of  
 438 more than 500 000 higher than all the other ones, if we consider it, highlighting its very  
 439 high potential impact, even at a very small concentration. The other impacts are  
 440 represented in the following Figure 6.



441  
 442 **Figure 6.** Graphical representation of the human impacts of the 258 compounds  
 443 studied in Servien et al. (2022b) (in red) and this study (in blue). Note that the 2,3,7,8-  
 444 T4CDD has been removed. Each compound has an estimated environmental impact  
 445 proportional to its rectangular area.

446  
 447 The choice of the imputation value for the measured concentrations below the LOQ is  
 448 also crucial. Indeed, a choice of LOQ/2 instead of LOQ/3 would have led to an  
 449 overestimation of 50.0% of the overall impacts on human health and 49.7% on  
 450 environmental impacts.

451

#### 452 **4 Conclusion**

453 In a previous study, 261 organic micropollutants were selected to study their potential  
454 impacts on human health and aquatic environment in continental freshwater at the  
455 scale of France. However, the lack of data allows quantifying the impacts for only 1/3  
456 of these substances (88 for aquatic environment and 94 for human health). Using a  
457 new modeling approach, we were then able to estimate the impact of 153 organic  
458 micropollutants, *i.e.* for all with an estimated mass released in the environment  
459 (Servien et al., 2022). The hundred remaining compounds without an accurate  
460 concentration measure were not taken into account. It means that their impacts were  
461 considered null, even if their characterization factors were high and could be worrying.  
462 In this work, the potential impacts of all these compounds are considered. We derive  
463 individual compound results, that could help to select substances in which a special  
464 effort should be made on the tertiary treatments to implement in WWTP. It was  
465 particularly possible to obtain a rough estimation of the probability that a compound  
466 produces a non-negligible impact, even if a measured concentration is below a LOQ.  
467 Obviously, due to the lack of concentration data, no precise results on the total sum of  
468 the impacts could be obtained. Nevertheless, with some strong assumptions, it was  
469 possible to conduct such a study and to show that ignoring compounds with  
470 unmeasured concentrations could lead to strong underestimations of total impacts. A  
471 future work could be to balance the individual impacts of these compounds with a  
472 probability of occurrence. Also, it has to be underlined that this whole methodology  
473 (from the computation of missing CFs using machine learning models to the individual  
474 compound impact estimations with missing concentrations) can be adapted to any  
475 other compartment and any other geographical context, with predictive models still to  
476 develop.

477

478 **Bibliography**

479

480 Addamo, M. Augugliaro, V. Di Paola, A. García-López, E. Loddo, V. Marci, G.,  
481 Palmisano, L., 2005. Removal of drugs in aqueous systems by photoassisted  
482 degradation. *Journal of Applied Electrochemistry*, 35(7-8):765–774,  
483 <https://doi.org/10.1007/S10800-005-1630-Y>.

484 Aemig, Q., Hélias, A., Patureau, D., 2021. Impact assessment of a large panel of  
485 organic and inorganic micropollutants released by wastewater treatment plants at the  
486 scale of France. *Water Research*, 188, 116524,  
487 <https://doi.org/10.1016/j.watres.2020.116524>.

488 Arulraj, A.D., Vijayan, M., Vasantha, V.S., 2015. Spectrophotometric determination of  
489 pico-molar level of hydrazine by using Alizarin red in water and urine samples,  
490 *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 148, 355-361,  
491 <https://doi.org/10.1016/j.saa.2015.03.092>.

492 Bayer, A., Asner, R., Schüssler, W., Kopf, W., Weiß, K., Sengl, M., Letzel, M., 2014.  
493 Behavior of sartans (antihypertensive drugs) in wastewater treatment plants, their  
494 occurrence and risk for the aquatic environment. *Environmental Science and Pollution*  
495 *Research*, 21(18):10830–10839, <https://doi.org/10.1007/s11356-014-3060-z>.

496 Bisognin, R.P., Wolff, D.B., Carissimi, E., Prestes, O.D., Zanella, R., 2021. Occurrence  
497 and fate of pharmaceuticals in effluent and sludge from a wastewater treatment plant  
498 in Brazil, *Environmental Technology*, 42(15), 2292-2303,  
499 <https://doi.org/10.1080/09593330.2019.1701561>.

500 Cantoni, B., Delli Compagni, R., Turolla, A., Epifani, I., Antonelli, M., 2020. A statistical  
501 assessment of micropollutants occurrence, time trend, fate and human health risk  
502 using left-censored water quality data, *Chemosphere*, 257, 127095,  
503 <https://doi.org/10.1016/j.chemosphere.2020.127095>.

504 Chaturvedi, N.K. and Katoch, S.S., 2020. Remedial Technologies for Aniline and  
505 Aniline Derivatives Elimination from Wastewater, *Journal of Health & Pollution*, 10(25),  
506 200302, <https://doi.org/10.5696/2156-9614-10.25.200302>.

507 Ellis, S.G., Booij, K., Kaputa, M., 2008. Comparison of semipermeable membrane  
508 device (SPMD) and large-volume solid-phase extraction techniques to measure water  
509 concentrations of 4,4'-DDT, 4,4'-DDE, and 4,4'-DDD in Lake Chelan, Washington,  
510 *Chemosphere*, 72(8), 1112-1117, <https://doi.org/10.1016/j.chemosphere.2008.04.040>.

511 Fazio, S., Castellani, V., Sala, S., Schau, E., Secchi, M., Zampori, L., Diaconu, E.,  
512 2018. Supporting information to the characterisation factors of recommended EF Life  
513 Cycle Impact Assessment methods: New methods and differences with ILCD, EUR  
514 28888 EN, *Publications Office of the European Union*, Luxembourg,  
515 <https://doi.org/10.2760/671368>.

516 Directive 2008/105/CE, 2008. Directive 2008/105/CE du 16/12/08 établissant des  
517 normes de qualité environnementales dans le domaine de l'eau, modifiant et  
518 abrogeant les directives du Conseil 82/176/CEE, 83/513/CEE, 84/156/CEE,  
519 84/491/CEE, 86/280CEE et modifiant la directive 20 0 0/60/CE,  
520 <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000020127782/>.

521 Ganigar, R., Rytwo, G., Gonen, Y., Radian, A., Mishael, Y.G., 2010. Polymer–clay  
522 nanocomposites for the removal of trichlorophenol and trinitrophenol from water,  
523 *Applied Clay Science*, 49(3), 311-316, <https://doi.org/10.1016/j.clay.2010.06.015>.

524 Gao, J., Liu, L., Liu, X., Zhou, H., Huang, S., Wang, Z., 2008. Levels and spatial  
525 distribution of chlorophenols – 2,4-Dichlorophenol, 2,4,6-trichlorophenol, and  
526 pentachlorophenol in surface water of China, *Chemosphere*, 71(6), 1181-1187,  
527 <https://doi.org/10.1016/j.chemosphere.2007.10.018>.

528 Garrison, A.W., Cyterski, M., Roberts, K.D., Burdette, D., Williamson, J., Avants, J.K.,  
529 2014. Occurrences and fate of DDT principal isomers/metabolites, DDA, and o,p'-DDD  
530 enantiomers in fish, sediment and water at a DDT-impacted superfund site,  
531 *Environmental Pollution*, 194, 224-234, <https://doi.org/10.1016/j.envpol.2014.07.025>.

532 Henderson, A.D., Hauschild, M.Z., Van De Meent, D., Huijbregts, M.A.J., Larsen, H.F.,  
533 Margni, M., McKone, T.E., Payet, J., Rosenbaum, R.K., Jolliet O., 2011. USEtox® fate  
534 and ecotoxicity factors for comparative assessment of toxic emissions in life cycle  
535 analysis: sensitivity to key chemical properties, *The International Journal of Life Cycle*  
536 *Assessment*, 16, pp. 701-709 <https://doi.org/10.1007/s11367-011-0294-6>.

537 INERIS, 2016. Les substances dangereuses pour le milieu aquatique dans les re- jets  
538 des stations de traitement des eaux usées urbaines - action nationale de recherche et  
539 de réduction des rejets de substances dangereuses dans l'eau par les stations de  
540 traitement des eaux,  
541 [https://www.ineris.fr/sites/ineris.fr/files/contribution/Documents/rapport-drc-15-](https://www.ineris.fr/sites/ineris.fr/files/contribution/Documents/rapport-drc-15-136871-11867e-rsde-steu-v-publique-1466157070.pdf)  
542 [136871-11867e-rsde-steu-v-publique-1466157070.pdf](https://www.ineris.fr/sites/ineris.fr/files/contribution/Documents/rapport-drc-15-136871-11867e-rsde-steu-v-publique-1466157070.pdf).

543 Leenhardt, S., et al., 2022. Impacts des produits phytopharmaceutiques sur la  
544 biodiversité et les services écosystémiques, Synthèse du rapport d'ESCo, INRAE-  
545 Ifremer (France), 124 pages. [https://www.inrae.fr/actualites/impacts-produits-](https://www.inrae.fr/actualites/impacts-produits-phytopharmaceutiques-biodiversite-services-ecosystemiques-resultats-lexpertise-scientifique-collective-inrae-ifremer)  
546 [phytopharmaceutiques-biodiversite-services-ecosystemiques-resultats-lexpertise-](https://www.inrae.fr/actualites/impacts-produits-phytopharmaceutiques-biodiversite-services-ecosystemiques-resultats-lexpertise-scientifique-collective-inrae-ifremer)  
547 [scientifique-collective-inrae-ifremer](https://www.inrae.fr/actualites/impacts-produits-phytopharmaceutiques-biodiversite-services-ecosystemiques-resultats-lexpertise-scientifique-collective-inrae-ifremer)

548 Li, B. and Zhang, T., 2011, Mass flows and removal of antibiotics in two municipal  
549 wastewater treatment plants, *Chemosphere*, 83(9), 1284-1289,  
550 <https://doi.org/10.1016/j.chemosphere.2011.03.002>.

551 Lindim, C., de Zwart, D., Cousins, I.T., Kutsarova, S., Kühne, R., Schüürmann, G.,  
552 2019. Exposure and ecotoxicological risk assessment of mixtures of top pre- scribed  
553 pharmaceuticals in Swedish freshwaters. *Chemosphere*, 220, 344–352,  
554 <https://doi.org/10.1016/j.chemosphere.2018.12.118>.

555 Lu, F., Jiang, Y., Wu, D., Zhou, J., Li, S., Zhang, J., 2016. Levels and profiles of  
556 polychlorinated dibenzo-p-dioxin and dibenzofurans in raw and treated water from  
557 water treatment plants in Shenzhen, China, *Environmental Pollution*, 211, 233-240,  
558 <https://doi.org/10.1016/j.envpol.2015.12.062>.

559 Martin Ruel, S, Choubert, J.-M.M., Budzinski, H., Miège, C., Esperanza, M., Coquery,  
560 M., 2012. Occurrence and fate of relevant substances in wastewater treatment plants  
561 regarding water framework directive and future legislations. *Water Science and*  
562 *Technology*, 65, 1179–1189, <https://doi.org/10.2166/wst.2012.943>.

563 Matsumoto, E., Kawanaka, Y., Yun, S.J., Oyaizu, H., 2009. Bioremediation of the  
564 organochlorine pesticides, dieldrin and endrin, and their occurrence in the  
565 environment. *Applied Microbiology & Biotechnology*, 84, 205–216.  
566 <https://doi.org/10.1007/s00253-009-2094-5>.

567 Ollers, S., Singer, H.P., Fassler, P., Muller, S.R., 2001. Simultaneous Quantification of  
568 Neutral and Acidic Pharmaceuticals and Pesticides at the Low-ng/1 Level in Surface  
569 and Waste Water, *Journal of Chromatography*, 911, 225-234,  
570 [https://doi.org/10.1016/S0021-9673\(01\)00514-3](https://doi.org/10.1016/S0021-9673(01)00514-3).

571 Qi, Y.-L., Chen, J., Zhang, B., Li, H., Li, D.-D., Wang, B.-Z., Yang, Y.-S., Zhu, H.-L.,  
572 2020. A turn-on fluorescent sensor for selective detection of hydrazine and its  
573 application in *Arabidopsis thaliana*, *Spectrochimica Acta Part A: Molecular and*  
574 *Biomolecular Spectroscopy*, 227, 117707, <https://doi.org/10.1016/j.saa.2019.117707>.

575 Ripley, B.D., 1996. *Pattern Recognition and Neural Networks*. Cambridge.

576 Rosenbaum, R.K., Margni, M., Jolliet, O., 2007. A flexible matrix algebra framework for  
577 the multimedia multipathway modelling of emission to impacts, *Environment*  
578 *International*, 33(5), 624-634, <https://doi.org/10.1016/j.envint.2007.01.004>.

579 Servien, R., Mamy, L., Li, Z., Rossard, V., Latrille, E., Bessac, F., Patureau, D., Benoit,  
580 P., 2014. TyPol - a new methodology for organic compounds clustering based on their  
581 molecular characteristics and environmental behaviour, *Chemosphere*, 111, 613–622,  
582 <https://doi.org/10.1016/j.chemosphere.2014.05.020>.

583 Servien, R., Latrille, E., Patureau, D., Hélias, A., 2022. Machine learning models based  
584 on molecular descriptors to predict human and environmental toxicological factors in  
585 continental freshwater. *Peer Community Journal*, 2, e15,  
586 <https://doi.org/10.24072/pcjournal.90>.

587 Servien, R., Leenknecht, C., Bonnot, K., Rossard, V., Latrille, E., Mamy, L., Benoit, P.,  
588 Hélias, A., Patureau, D., 2022. Improved impact assessment of micropollutants release  
589 from WWTPs, *Case Studies in Chemical and Environmental Engineering*, 5, 100172.  
590 <https://doi.org/10.1016/j.cscee.2021.100172>.

591 Spongberg, A. L. and Witter, J. D., 2008. Pharmaceutical compounds in the  
592 wastewater process stream in Northwest Ohio, *Science of the Total Environment*, 397,  
593 148-157, <https://doi.org/10.1016/j.scitotenv.2008.02.042>.

594 Srogi, K., 2008. Levels and congener distributions of PCDDs, PCDFs and dioxin-like  
595 PCBs in environmental and human samples: a review, *Environmental Chemistry*  
596 *Letters*, 6, 1–28. <https://doi.org/10.1007/s10311-007-0105-2>.

597 Suliman, F., Al-Kindi, S., Al-Kindy, S., Al-Lawati, H., 2006. Analysis of phenols in water  
598 by high-performance liquid chromatography using coumarin-6-sulfonyl chloride as a  
599 fluorogenic precolumn label, *Journal of Chromatography A*, 1101(1–2), 179-184,  
600 <https://doi.org/10.1016/j.chroma.2005.09.094>.

601 UNEP-SETAC, 2019. Global Guidance for Life Cycle Impact Assessment Indicators:  
602 Volume 2, [https://www.lifecycleinitiative.org/training-resources/global-guidance-for-](https://www.lifecycleinitiative.org/training-resources/global-guidance-for-life-cycle-impact-assessment-indicators-volume-2/)  
603 [life-cycle-impact-assessment-indicators-volume-2/](https://www.lifecycleinitiative.org/training-resources/global-guidance-for-life-cycle-impact-assessment-indicators-volume-2/).

604 Van den Berg, M., Birnbaum, L.S., Denison, M., De Vito, M., Farland, W., Feeley, M.,  
605 Fiedler, H., Hakansson, H., Hanberg, A., Haws, L., Rose, M., Safe, S., Schrenk, D.,  
606 Tohyama, C., Tritscher, A. Tuomisto, J., Tysklind, M., Walker, N., Peterson, R.E., 2006.  
607 The 2005 world health organization reevaluation of human and mammalian toxic  
608 equivalency factors for dioxins and dioxin-like compounds, *Toxicological Sciences*,  
609 93(2), 223-241, <https://doi.org/10.1093/toxsci/kfl055>.

610 Verlicchi, P., Al Aukidy, M., Zambello, E., 2012. Occurrence of pharmaceutical  
611 compounds in urban wastewater: removal, mass load and environmental risk after a  
612 secondary treatment--a review, *Science of the Total Environment*, 429, 123-55. doi:  
613 <https://doi.org/10.1016/j.scitotenv.2012.04.028>.

614 Villanneau, E., Saby, N., Marchant, B., Jolivet, C., Boulonne, L., Caria, G., Barriuso,  
615 E., Bispo, A., Briand, O., Arrouays, D., 2011. Which persistent organic pollutants can  
616 we map in soil using a large spacing systematic soil monitoring design? A case study



617 in Northern France, *The Science of the total environment*, 409, 3719-31,  
618 <https://doi.org/10.1016/j.scitotenv.2011.05.048>.

619 Wan, Y., Tran, T.M., Nguyen, V.T., Wang, A., Wang, J., Kannan, K., 2021,  
620 Neonicotinoids, fipronil, chlorpyrifos, carbendazim, chlorotriazines, chlorophenoxy  
621 herbicides, bentazon, and selected pesticide transformation products in surface water  
622 and drinking water from northern Vietnam, *Science of The Total Environment*, 750,  
623 141507, <https://doi.org/10.1016/j.scitotenv.2020.141507>.

624 Watkinson, A. J., Murby, E. J., Kolpin, D. W., Costanzo, S. D., 2009. The occurrence  
625 of antibiotics in an urban watershed: From wastewater to drinking water, *Science of*  
626 *the Total Environment*, 407, 2711-2723,  
627 <https://doi.org/10.1016/j.scitotenv.2008.11.059>.

628 Yang, S. W., Cha, J. M., Carlson, K., 2005. Simultaneous extraction and analysis of  
629 11 tetracycline and sulfonamide antibiotics in influent and effluent domestic wastewater  
630 by solid-phase extraction and liquid chromatography-electro spray ionization tandem  
631 mass spectrometry, *Journal of Chromatography A*, 1097, 40-53,  
632 <https://doi.org/10.1016/j.chroma.2005.08.027>.

633 Yao, L., Chen, Z.-Y., Dou, W.-Y., Yao, Z.-K., Duan, X.-C., Chen, Z.-F., Zhang, L.-J.,  
634 Nong, Y.-J., Zhao, J.-L., Ying, G.-G., 2021. Occurrence, removal and mass loads of  
635 antiviral drugs in seven wastewater treatment plants with various treatment processes,  
636 *Water Research*, 207, 117803, <https://doi.org/10.1016/j.watres.2021.117803>.

637 Zhu, B., Zonja, B., Gonzalez, O., Sans, C., Pérez, S., Barceló, C., Esplugas, S., Xu,  
638 K., and Qiang, Z., 2015. Degradation kinetics and pathways of three calcium channel  
639 blockers under UV irradiation. *Water Research*, 86:9–16,  
640 <https://doi.org/10.1016/j.watres.2015.05.028>.