



HAL
open science

Application of a multidisciplinary and integrative weight-of-evidence approach to a 1-year monitoring survey of the Seine River

Iris Barjhoux, Lise C. Fechner, Jérémie D. Lebrun, Adriana Anzil, Sophie Ayrault, Hélène Budzinski, Jérôme Cachot, Laëtitia Charron, Arnaud Chaumot, Christelle Clérandeau, et al.

► To cite this version:

Iris Barjhoux, Lise C. Fechner, Jérémie D. Lebrun, Adriana Anzil, Sophie Ayrault, et al.. Application of a multidisciplinary and integrative weight-of-evidence approach to a 1-year monitoring survey of the Seine River. *Environmental Science and Pollution Research*, 2018, 25 (24), pp.23404-23429. 10.1007/s11356-016-6993-6 . hal-02153600

HAL Id: hal-02153600

<https://hal.science/hal-02153600v1>

Submitted on 4 May 2020

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.

Application of a multidisciplinary and integrative Weight-of-Evidence approach to a one-year monitoring survey of the Seine River

*Iris Barjhoux^{1,✉}, Lise C. Fechner^{2,3}, Jérémie D. Lebrun², Adriana Anzil⁴, Sophie Ayrault⁵,
Hélène Budzinski⁶, Jérôme Cachot⁶, Laetitia Charron¹, Arnaud Chaumot⁷, Christelle
Clérandeau⁶, Odile Dedourge-Geffard¹, Juliette Faburé^{2,3}, Adeline François⁷, Olivier
Geffard⁷, Isabelle George⁴, Pierre Labadie⁶, Yves Lévi⁸, Gabriel Munoz⁶, Patrice Noury⁷,
Lucie Oziol⁸, Hervé Quéau⁷, Pierre Servais⁴, Emmanuelle Uher², Nastassia Urien², Alain
Geffard¹*

¹ Université de Reims Champagne-Ardenne, UMR-I 02 Stress Environnementaux et Biosurveillance des milieux aquatiques (SEBIO), UFR SEN, Moulin de la Housse, BP 1039, Reims, France

² Irstea, UR Hydrosystèmes et bioprocédés Antony (HBAN), 1 rue Pierre-Gilles de Gennes, CS 10030, 97261 Antony Cedex, France

³ AgroParisTech, F-75005 Paris, France

⁴ Laboratoire Ecologie des Systèmes Aquatiques (ESA), Université Libre de Bruxelles, Bruxelles, Belgique

⁵ UMR 8212 CNRS CEA UVSQ Laboratoire des Sciences du Climat et de l'Environnement (LSCE) / Institut Pierre Simon Laplace (IPSL), Gif-sur-Yvette, France

⁶ UMR 5805 Environnements et Paléoenvironnements Océaniques et Continentaux (EPOC), Université de Bordeaux, Pessac, France

⁷ UR Milieux Aquatiques, Ecologie et Pollutions (MAEP), Laboratoire d'écotoxicologie, Irstea, Villeurbanne, France

⁸ Ecologie Systématique Evolution, Univ. Paris-Sud, CNRS, AgroParisTech, Université Paris-Saclay, Orsay, France

✉ **Corresponding author:** Iris Barjhoux (PhD)

University of Reims Champagne-Ardenne,
UMR-I 02 INERIS-URCA-ULH SEBIO,
UFR SEN, Moulin de la Housse, BP 1039,

51687 Reims Cedex 2, France.

E-mail contact: iris.barjhoux@univ-reims.fr / irisbarjhoux@hotmail.com (permanent)

Phone/Fax: +33 (0)326913719/+33 (0)326913342

Acknowledgements

This work was supported by the French research program PIREN-Seine. The authors would also like to thank the Aquitaine Region, the European Union (CPER A2E project) and the European Regional Development Fund for their financial support. This study was also carried out within the framework of the COTE Cluster of Excellence (ANR-10-LABX-45). IdEx Bordeaux (ANR-10-IDEX-03-02) provided the PhD grant allocated to G. Munoz. Finally, the authors would like to thank Annie Buchwalter for English revision of the manuscript.

1 **Abstract**

2 Quality assessment of environments under high anthropogenic pressures such as the Seine Basin,
3 subjected to complex and chronic inputs, can only be based on combined chemical and biological
4 analyses. The present study integrates and summarizes a multidisciplinary dataset acquired throughout
5 a one-year monitoring survey conducted at three workshop-sites along the Seine River (PIREN-Seine
6 program), upstream and downstream of the Paris conurbation, during four seasonal campaigns, using a
7 Weight-of-Evidence (WOE) approach. Sediment and water column chemical analyses,
8 bioaccumulation levels and biomarker responses in caged gammarids, and laboratory (eco)toxicity
9 bioassays were integrated into four lines of evidence (LOEs). Results from each LOE clearly reflected
10 an anthropogenic gradient, with contamination levels and biological effects increasing from upstream
11 to downstream of Paris, in good agreement with the variations in the structure and composition of
12 bacterial communities from the water column. Based on annual average data, the global hazard was
13 summarized as ‘moderate’ at the upstream station and as ‘major’ at the two downstream ones.
14 Seasonal variability was also highlighted; the winter campaign was least impacted. The model was
15 notably improved using previously established reference and threshold values from national-scale
16 studies. It undoubtedly represents a powerful practical tool to facilitate the decision-making processes
17 of environment managers within the framework of an environmental risk assessment strategy.

18

19 *Keywords:* Seine River, Environmental risk, Weight-of-Evidence, Pollutants, Bioavailability,
20 Biomarkers, Bioassays, Bacterial community

21

22 **1. Introduction**

23 The Seine river basin represents a catchment area of around 78,600 km² from its source at Seine-
24 Source near Dijon in north-eastern France to its mouth in the English Channel in the northwestern city
25 of Le Havre. It is supplied by a fairly regular network of tributaries. The central zone of the watershed
26 is the convergence area of the main tributaries of the Seine River, and is occupied by the large Paris
27 conurbation. In total, just over a quarter of the French population (~17.5 million) lives in this
28 watershed, mostly (85%) in urban areas. The Seine watershed also harbors very intensive agricultural
29 activities resulting in substantial diffuse sources of nutrients and pollutants such as pesticides (Billen
30 et al. 2007). The heavy urbanization and industrialization of the Paris area also result in significant
31 inputs of contaminants into the Seine River basin, including metals and persistent toxic substances
32 such as polycyclic aromatic hydrocarbons (PAHs) and polychlorinated biphenyls (PCBs) (Blanchard
33 et al. 2007; Thévenot et al. 2007). The wide variety of anthropogenic pressures that affect the Seine
34 watershed makes this area an ideal case study of contemporary environmental problems in developed
35 countries: the river's water quality and ecological status integrate and reflect the complex functioning
36 of the watershed, especially the ways humans have shaped and exploited land- and water-scapes
37 (Billen et al. 2007). However, the complexity and diversity of exogenous inputs result in an equally
38 complex, diffuse, and chronic pressure on the ecosystems of the Seine continuum. The biological
39 effects and long-term impacts of this pressure on the biota still remain difficult to evaluate directly.

40 The contamination level at a particular site can be quite readily determined through chemical analysis
41 as defined by the presence of '*substances that would not normally occur or at concentrations above*
42 *the natural background*'. However, 'pollution status' assessment additionally integrates chemical
43 bioavailability and the biological impacts of contaminants on the environment (Chapman 2007).
44 Consequently, it is now widely admitted that an efficient environmental risk assessment (ERA) should
45 be conducted through an integrated and multidisciplinary strategy to provide answers to all these
46 concerns. Moreover, such approaches are clearly recommended and even required by the European
47 Water Framework Directive 2000/60/CE (European Commission (EC) 2000).

48 Between 2011 and 2012, eight research teams collaborated on a synchronous and integrative multi-
49 marker approach aiming at a global assessment of the chemical and ecological/ecotoxicological status
50 of three workshop-sites along the Seine axis situated upstream and downstream of Paris (PIREN-Seine
51 program) (Fig. 1). This one-year monitoring program consisted of four field measurement campaigns
52 corresponding to distinct seasons. During each sampling period, a wide panel of biological and
53 chemical analyses was performed to characterize in detail the quality of the aquatic environment at
54 each sampling location based on six fundamental aspects: (i) the physico-chemical quality of the water
55 column and sediment, (ii) a comprehensive analysis of metal and organic contaminants in the same
56 two compartments, (iii) the bioaccumulation of contaminants of concern in field-transplanted

57 gammarids and river biofilms, (iv) the biological responses in these gammarids exposed *in situ*, (v) the
58 spatio-temporal variations in autochthonous bacterial community composition and metal tolerance
59 acquisition, and (vi) the (eco)toxicity of water and sediment samples in laboratory bioassays. In total,
60 about 550 parameters were monitored *per* site and *per* sampling period. Thus, one of the main
61 challenges was to find a way to summarize and interpret the large dataset issued from this multi-
62 marker study.

63 To achieve this objective, the concept of Weight of Evidence (WOE) appeared as an adequate strategy
64 for a global and integrative multidisciplinary assessment of environmental quality in the area, because
65 it is based on the packaging of a wide variety of data within several lines of evidence (LOEs). In each
66 LOE the contamination level assessed through chemical analyses is combined with bioavailability
67 analyses and biological responses from key species and/or model organisms at different levels of
68 biological organization (Chapman et al. 2002; Dagnino et al. 2008). The resulting environmental
69 diagnosis is based on the calculation of a hazard index for each LOE, which is then plotted on an
70 evaluation grid allowing for clear and rapid hazard classification (Chapman et al. 2002; Dagnino et al.
71 2008; Piva et al. 2011). We also propose a global hazard evaluation that compiles all calculated LOE
72 indices within a single one that is also finally assigned to a hazard class. This approach was
73 successfully applied to the assessment of the health status in multi-contaminated environments such as
74 harbors, urbanized and/or industrial areas (*e.g.*, Piva et al. 2011; Benedetti et al. 2012; Bebianno et al.
75 2015). These studies generally focused on sediment hazard assessment. However, the WOE approach
76 is also applicable to other matrices such as effluents, water, and soils (Chapman et al. 2002; Chapman
77 2007; Dagnino et al. 2008), and to more global environmental diagnosis such as aquatic and terrestrial
78 hazard assessment (Chapman et al. 2002; Chapman 2007; Piva et al. 2011).

79 Relying on the above-mentioned promising applications of the WOE procedure to sediment hazard
80 assessment, the present study implements this multicriteria-based environmental diagnosis to a one-
81 year monitoring survey of the Seine River axis. The three model sites investigated in the present work
82 are situated along the Seine River continuum and characterized by a strong contamination gradient
83 from upstream to downstream of the Paris urban area (Priadi et al. 2011; Fechner et al. 2012; Teil et al.
84 2014). As the WOE model described by Piva et al. (2011) relies on the calculation of Ratio-to-
85 Reference (RTR) values, only part of the data obtained during the 2011-2012 campaigns was selected.
86 The upstream station is often considered as a 'reference' site in similar case studies. This station was
87 expected to be relatively unaffected by direct inputs from the Paris conurbation, but it was probably
88 impacted by the intense agricultural activities surrounding the densely urbanized area, as well as by
89 domestic or industrial inputs from the relatively small cities located upstream. The use of 'external
90 reference' values established from several field-monitoring campaigns and physiological studies on
91 the selected sentinel organisms would make it possible to classify all sites, including the upstream one.
92 To that purpose, biomarkers and bioaccumulation levels were analyzed through an active approach in

93 transplanted *Gammarus fossarum* crustaceans, a common species in the field of ecotoxicology and
94 biomonitoring. Gammarids have been reported as efficient accumulators of organic compounds and
95 metals, whether essential or not (Besse et al. 2013; Lebrun et al. 2014). Besides, they are commonly
96 used for the development of exposure biomarkers because they are easily sampled in the field and
97 handled (Besse et al. 2013; Dedourge-Geffard et al. 2013; Lebrun et al. 2014). Moreover, translocated
98 gammarid populations have been fully characterized, and reference levels are (or could be) established
99 for bioaccumulation levels and biomarker responses (Xuereb et al. 2009; Geffard et al. 2010; Coulaud
100 et al. 2011; Besse et al. 2013; Charron et al. 2013). According to the existence of reference levels or
101 the possibility to derive them for each endpoint, the dataset selected for WOE integration was the
102 following:

103 (1) Chemical hazard (LOE#1) was characterized through pesticide (PEST), alkylphenol (AKP), metal
104 element (ME), and perfluoroalkyl substance (PFAS) analysis in the water column. MEs and PFASs
105 were also measured in composite sediment samples together with more hydrophobic and (very)
106 persistent compounds such as PAHs, PCBs, polybromodiphenyl ethers (PBDEs), and organochlorine
107 pesticides (OCPs);

108 (2) Bioavailability (LOE#2) of the chemicals of concern, including PAHs, PCBs, PBDEs, OCPs, and
109 MEs, was assessed by measuring bioaccumulation levels in caged gammarids;

110 (3) Biological responses (LOE#3) in the same population of transplanted gammarids were assessed
111 using validated biomarkers such as digestive enzyme activity, feeding rate, reproductive toxicity, and
112 acetylcholinesterase (AChE) activity;

113 (4) (Eco)toxicological responses at the organism/cellular level were also investigated using laboratory
114 bioassays (LOE#4) performed on water column and sediment samples. They included genotoxicity,
115 cytotoxicity, and endocrine disruption (ED) *in vitro* bioassays, as well as a fish embryo toxicity test,
116 the Medaka embryo-larval assay (MELA).

117 The overall aim of this study was to confirm the importance and relevance of a multidisciplinary
118 survey of aquatic environment quality. Such an approach is not realistically applicable in an ERA
119 strategy without a practical tool to integrate and fruitfully interpret the resulting large dataset within a
120 global environmental context. The present work applies the WOE model, adapted from Piva et al.
121 (2011), to a practical case study on the Seine River continuum. The aim of this integrative approach is
122 to assess the overall quality of the aquatic environment and prioritize hazards at each of the three sites.
123 Such an approach could represent a promising decision-making tool for environmental managers.

124

125

126 **2. Materials and Methods**

127 *2.1. Studied area and sampling procedure*

128 The three sampling sites were previously described by Lebrun et al. (2015) and Faburé et al. (2015).
129 Briefly, these sites are situated along the Seine River in the north of France (Fig. 1). Marnay
130 (48°31'35.8" N, 3°33'29.6" E) is located approximately 200 km upstream of Paris, in a non-
131 urbanized area, and therefore expected to be at least partially free from direct inputs from the Paris
132 conurbation. Conversely, Bougival (48°52'11.2" N, 2°07'47.1" E) and Triel (48°58'55.5" N,
133 1°59'53.1" E) are both situated downstream of Paris and its conurbation, at respective distances of
134 approximately 40 and 80 km (Fig. 1). These stations are affected by various contamination inputs in
135 relation to intense anthropogenic activities (Priadi et al. 2011, Teil et al. 2014). Sampling was
136 performed at these sites during four campaigns undertaken in fall (C1 campaign, from August 31st to
137 September 27th, 2011), spring (C2 campaign, from March 2nd to April 3rd, 2012), summer (C3
138 campaign, from June 1st to July 3rd, 2012), and winter (C4 campaign, from November 13th to
139 December 18th, 2012), corresponding to contrasted temperature and flow rate conditions.

140 Dissolved metal concentrations were determined at the three sites during each sampling period (Faburé
141 et al. 2015; Lebrun et al. 2015). At the end of each seasonal campaign, water was collected at each
142 station as follows: 1 L of raw water in amber glass bottles for endocrine disruption bioassays, 10 L of
143 raw water in high-density polyethylene (HDPE) containers for MELA, 20 L of raw water in two 10-L-
144 HDPE containers for microbial community analyses. All containers were rinsed three times with river
145 water before being filled in the field; they were all brought back to the laboratory in a cool box and
146 then kept at 4 °C until further use. In addition, two 250-mL-HDPE bottles were filled in a similar way
147 and stored at -20 °C for organic contaminant analyses. Prior to analyses, samples were thawed and
148 filtered through GF/F (0.7 µm) Whatman glass microfiber filters previously ignited at 450°C for 6 h.

149 For each campaign (except C3), one composite surface (0–2 cm) bed sediment sample was collected in
150 an aluminum container, brought back to the laboratory in a cool box, and stored either at 4 °C for
151 bioassays or at -20 °C until freeze-drying, grinding and 2 mm-sieving for chemical analyses.

152

153 *2.2. Field-caged gammarid exposure*

154 The procedures are detailed in previous studies (Coulaud et al. 2011; Besse et al. 2013; Lebrun et al.
155 2015). Briefly, gammarids (*Gammarus fossarum*) were collected by kick sampling at La Tour du Pin,
156 upstream of the Bourbre River (France). This site displays good water quality according to the data
157 records of the RNB (French Watershed Biomonitoring Network). After a 15-day acclimatization

158 period in the laboratory (conditions detailed in Besse et al. (2013)) and 24 hours before *in situ* caging,
159 8 groups of 20 adult gammarids (10-11 mm) were caged in polypropylene cylinders (10 cm length, 5.5
160 cm diameter) capped at the ends with pieces of net (1 mm mesh) to ensure free water circulation. To
161 assess the effects on reproduction, three supplemental experimental systems, each containing seven
162 precopulatory pairs with D2-molt stage females (*i.e.*, hatched juveniles in brood pouches and visible
163 oocytes) were set up. A temperature probe was placed in the water to record temperature every hour
164 throughout the experiment. During the tests, gammarids were fed with the same alder (*Alnus*
165 *glutinosa*) leaves as during the acclimatization period in the laboratory, pre-conditioned for at least
166 6 ± 1 days in groundwater.

167 After 7 days of exposure, two replicates were collected and brought to the laboratory for
168 bioaccumulation measurements in whole organisms (3 pools of 5 gammarids) for each site. After 15
169 days of exposure, three replicates *per* site were collected and brought to the laboratory. Gammarids
170 from the same site were collected, counted (for survival rate assessment), then male gammarids were
171 pooled, dried, weighed, flash frozen in liquid nitrogen and stored at -80 °C until digestive enzyme
172 activity and AChE activity were analyzed. Leaf consumption was used to estimate the feeding rate for
173 each site and campaign. After 30 days of exposure, the last three replicates were collected and brought
174 to the laboratory. Gammarids from the same site were pooled together and counted (for survival rate
175 assessment); females were then selected to analyze reproduction markers (molt delay, number of
176 embryos/oocytes *per* female).

177

178 2.3. Chemical analyses

179 2.3.1. Metal element measurements

180 For metal determination in caged gammarids, 3 pools of 5 individuals were digested by HNO₃ and
181 H₂O₂, as detailed by Lebrun et al. (2015). A reference material (Mussel Tissue ERM-CE278, LGC
182 Promochem, Molsheim, France) was included in each digestion series to control the quality of
183 digestion.

184 About 0.1 g of sediment was mineralized in closed Teflon vessels under a hood using a heating block
185 (Digiprep, SCP Science). A three-step digestion was performed as described by Priadi et al. (2011). A
186 geostandard was included in each digestion series (IAEA lake sediment SL1) to control chemical
187 mineralization efficiency. All reagents used for the digestion processes were ultrapure reagents to
188 avoid contamination.

189 Major and minor element concentrations were determined in filtered acidified water and in digested
190 sediment and gammarid solutions by inductively coupled plasma quadrupolar mass spectrometry

191 (ICP-QMS; X-Series, CCT II+ Thermoelectron, France), as previously described (Faburé et al. 2015;
192 Lebrun et al. 2015, Le Pape et al. 2012, Priadi et al. 2011). Accuracy checking (SRM 1640a, NIST,
193 Gaithersburg, USA) and plasma fluctuation corrections were also performed as described in the same
194 references.

195

196 2.3.2. Organic compound analysis

197 Organic micropollutants were determined using previously established methods. Briefly, dissolved
198 (<0.7 µm fraction) pesticides and PFASs were extracted using solid phase extraction with polymeric
199 sorbents (100-500 mL samples) followed by analysis by liquid chromatography coupled to tandem
200 mass spectrometry (Dufour et al. 2015; Munoz et al. 2015), while alkylphenols were determined using
201 solid-phase microextraction (SPME) and gas chromatography coupled to mass spectrometry (Belles et
202 al. 2014). Freeze-dried sediment (1 g) or gammarid (0.2 g) samples were extracted using microwave-
203 assisted extraction followed by solid phase extraction adsorption chromatography clean-up (Budzinski
204 et al. 2000; Noura et al. 2013; Munoz et al. 2015).

205

206 2.4. Biomarker analysis

207 2.4.1. Digestive enzyme activity

208 The enzymatic activity of two carbohydrases (amylase and cellulase) and a protease (trypsin) was
209 determined as previously described by Charron et al. (2013), using starch (1%), carboxymethyl-
210 cellulose (2%), and N-benzoyl-DL-arginine 4-nitroanilide hydrochloride (3 mM) as substrates,
211 respectively.

212

213 2.4.2. AChE activity

214 AChE activity was analyzed as described in Xuereb et al. (2009) according to the colorimetric method
215 initially developed by Ellman et al. (1961), with DTNB (5,5'-dithiobis(2-nitrobenzoic acid) as a
216 substrate.

217

218 2.4.3. Feeding rate assessment

219 Feeding rates were calculated according to the method described by Coulaud et al. (2011).
220 Calculations were based on leaf disc scanning and expressed as consumed surface *per day per* living
221 gammarid ($\text{mm}^2 \text{d}^{-1} \text{organism}^{-1}$).

222

223 2.4.4. Reproduction markers

224 At the end of the exposure period (30 days), the size, molting stage, number of oocytes and embryos
225 *per* female were determined according to Geffard et al. (2010). To accurately assess the females' molt
226 stages, the third and fourth pereopod pairs (dactilopodite and protopodite) of females were cut off,
227 mounted on a microscope slide with a coverslip, and their integumental morphogenesis was observed
228 ($\times 200$) to discriminate among the five molt stages (AB, C1, C2, D1, and D2). The number of oocytes
229 *per* female in C2/D1-molt stage was determined by *in vivo* observation of the two ovaries under a
230 binocular microscope. In the same way, embryos of females bearing a brood in 2nd, 3rd, or 4th
231 embryonic stage were manually recovered from the marsupium, placed on a slide with water, and
232 counted under a binocular microscope. Desynchronization between female molt stage and embryonic
233 development stage was also recorded to assess delays in female molt cycle (Geffard et al. 2010).

234

235 2.5. Bioassays

236 2.5.1. Endocrine disruption *in vitro* bioassays

237 Endocrine disruption (ED) bioassays were conducted on organic extracts prepared in
238 dimethylsulfoxide (DMSO) from water column samples (1 L) or from freeze-dried sediment samples
239 (1 g) according to Jugan et al. (2009) and Kinani et al. (2010), respectively. Three luciferase reporter
240 bioassays were used to evaluate the ED potential of organic extracts from sediment or water column:
241 using MELN cells (Balaguer et al. 1999), PC-DR-LUC cells (Jugan et al. 2007), or MDA-kb2 cells
242 (Wilson et al. 2002), we measured disruptions of the transcriptional activity of the estrogen receptor
243 $\text{ER}\alpha$ (ER), of the thyroid receptor $\text{TR}\alpha 1$ (TR), and of the androgen (AR) and glucocorticoid (GR)
244 receptors, respectively, by bioluminescence.

245 The results were expressed as fold induction in relative luminescence units (RLUs) as compared to
246 luciferase activity of the solvent control (DMSO 0.1%). Only RLU values significantly different from
247 that of the solvent control (Student's *t*-test, $p < 0.05$) were considered as above the LD. Any detectable
248 RLU levels above the bottom value of the sigmoidal dose-response curves of reference ligands were
249 considered as above the LQ. This threshold value of the sigmoid was obtained by nonlinear regression
250 of the Hill equation (GraphPad Prism 5 Software, San Diego, CA, USA). Furthermore, only RLU

251 levels significantly different from that of the corresponding blank value (Student's *t*-test, $p < 0.05$)
252 were taken into account.

253

254 2.5.2. Microtox® and SOS Chromotest procedures

255 The two bioassays were performed on sediment elutriates to measure the toxicity of water-extractable
256 pollutants. After thawing overnight at 4 °C, 6 g wet weight of sediment were mixed with 24 mL of
257 deionized water for 10 min at 300 rpm. The solid phase was pelleted at 1,800 × g for 10 min, and the
258 supernatant was immediately collected and stored at 4 °C in the dark prior to toxicity testing within
259 24 h.

260 For the Microtox® assay, the standard procedure of the Acute Toxicity Basic Test was used (AZUR
261 Environmental 1998; ISO 1999). Bioluminescence was measured after 30 min on duplicated series of
262 elutriate serial dilutions using a Microtox Model 500 analyzer (Azur Environmental).

263 The SOS-Chromotest developed by Quillardet and Hofnung (1985) was miniaturized in microplates.
264 Briefly, *E. coli* PQ37 strain was exposed to the elutriate (3% final, v/v) for 3 h at 37 °C, in triplicate,
265 with and without the liver S9 fraction (10% final, v/v) from β-naphthoflavone- and phenobarbital-
266 treated rats (Trinova-Biochem). Following exposure, beta galactosidase (BG) and alkaline phosphatase
267 (AP) activity levels were measured colorimetrically at 420 nm (Fluo Star Optima, BMG Labtech). The
268 SOS control-relative induction factor (IF) was calculated by dividing the BG/AP activity ratio of the
269 sample by the solvent control BG/AP ratio, as described by Quillardet and Hofnung (1985). Results
270 were expressed as mean induction factor ± standard deviation (three replicates).

271

272 2.5.3. Medaka embryo-larval assay (MELA)

273 Japanese Medaka (*Oryzias latipes*) embryos of the CAB strain were provided by the UMS Amagen (Gif-
274 sur-Yvette, France) 1 day post fertilization (dpf).

275 Whole sediment toxicity was evaluated by the Medaka Embryo-Larval Assay in sediment contact
276 (MELAc), using the protocol described by Barhoumi et al. (2016). Reference non-contaminated
277 sediment (Yville-sur-Seine) was used as a negative control (Vicquelin et al. 2011). Briefly, 25 embryos
278 *per* replicate were laid onto a Nitex® mesh at the sediment surface and immersed into egg-rearing
279 solution (ERS). To avoid hypoxia at the sediment-water interface, ERS was thoroughly renewed and
280 dissolved oxygen was measured daily.

281 The toxicity of water samples was evaluated by the Medaka Embryo-Larval Assay in 96-well
282 microplates, adapted from Helmstetter and Alden (1995). Before testing, the water samples were filtered
283 through 0.8 µm-filters (Millipore) to remove particles. Twenty-five embryos *per* condition were
284 individually incubated in 300 µL of water sample. Water was renewed daily, and spring water
285 (Cristaline) was used as a negative control.

286 The procedure was similar for the two assays, and followed previously published protocols (Vicquelin
287 et al. 2011; Barjhoux et al. 2012; Barhoumi et al. 2016). In summary, exposure was performed at
288 26 ± 0.3 °C, and stopped at the first hatching peak in one of the test conditions (10-11 dpf). Hatchlings
289 and unhatched embryos were transferred to clean water or ERS, respectively, for three additional days.
290 Viability, time to hatch, hatching success, body and head length, and developmental abnormalities
291 were recorded in embryos and larvae according to Barjhoux et al. (2012).

292

293 *2.6. Bacterial community composition*

294 An aliquot (from 0.8 L to 5.0 L) of each water sample was filtered through a 0.2 µm pore-size, 47-
295 mm-diameter polycarbonate filter (Millipore, MA). All filters were stored at -20 °C until use. DNA
296 was extracted using phenol-chloroform-isoamyl alcohol, following an enzymatic cell lysis stage in
297 the presence of lysozyme, mutanolysine and sodium dodecyl sulfate. Bacterial community structure
298 (number and relative abundance of the different taxa) was assessed by pyrosequencing of the V1-V3
299 region of the bacterial 16S rRNA gene, and downstream sequence analysis was performed using the
300 software program MOTHUR (full procedure described in García-Armisen et al. 2014).

301 Using the PRIMER v6 software program, we compared bacterial community structures among
302 samples based on Bray-Curtis coefficient matrix, after square-root transformation of the data. This
303 coefficient evaluates the dissimilarity between each pair of samples in terms of species abundance.
304 The resulting matrix was used as a basis for a graphic representation of dissimilarities in a non-metric
305 multi-dimensional scaling (NMDS) graph, where each sample was represented by a dot; the more
306 different the structures of two bacterial communities, the further apart the two corresponding dots on
307 the graph.

308

309 *2.7. Data integration within the Weight-of-evidence (WOE) approach*

310 The data selected to characterize contamination levels (*i.e.*, chemical analyses) in the area,
311 contaminant bioavailability (*i.e.*, bioaccumulation levels in caged gammarids), and *in situ* biological
312 responses (*i.e.*, biomarkers in gammarids) and following laboratory exposure (*i.e.*, bioassays) were

313 integrated into a WOE approach according to Piva et al. (2011). Slight modifications and/or
314 adaptations were made and are described below.

315

316 2.7.1. Line of evidence 1: sediment and water column chemistry (LOE#1)

317 Among the 210 metal and organic compounds analyzed in the abiotic compartment, we selected
318 chemicals to be included in the WOE approach according to their mention in reference studies
319 (MacDonald et al. 2000; Piva et al. 2011), and French and European regulatory documents (EC 2000,
320 2013; French Ministry of Ecology Energy Sustainable Development and Planning (MEDAD) 2007,
321 2015). The reference values used in LOE#1 calculations were environmental quality standards (EQSs)
322 or environmental guideline values (EGVs) of the Ineris (French National Institute for Environmental
323 Technology and Hazards) when available. Otherwise, Predicted No Effect Concentrations (PNECs)
324 were gathered from environmental institutes recognized at the European level (Environment Agency,
325 Ineris, Anses, and European Chemical Agency) and key reports (MacDonald et al. 2000; European
326 Union (EU) 2005; MEDAD 2006; Dulio and Andres 2014). Then, the geometric mean of the PNECs
327 was used as the reference value. Details on the selected reference values and the list of chemicals are
328 given in Table S1 (for the water column) and Table S2 (for the sediment) in the Electronic
329 Supplementary Material (ESM). Note that for both metal and organic compound analysis, data below
330 the limit of detection (LD) were set at LD/2 before being integrated into calculations. Moreover, when
331 the reference value was lower than the corresponding LD/2, measured concentrations below the LD
332 were removed from the dataset.

333 The detailed calculation procedure implemented in LOE#1 is presented in ESM Fig. S1. As described
334 by Piva et al. (2011), the elaboration of the chemical data into the corresponding hazard quotient (HQ)
335 was based on the calculation of a ratio-to-reference (RTR) for each chemical and its weighting (RTR_w)
336 according to chemical status within the Water Framework Directive (WFD) 2013/39/EU (EC 2013)
337 (see ESM Tables S1 and S2).

338 The global Chemical Hazard Quotient (ChemHQ) was calculated by averaging the RTR_w values for
339 chemicals whose measured concentrations were below or equal to the reference level (*i.e.*, RTR ≤ 1)
340 and summing the RTR_w values for chemicals whose concentrations exceeded the reference (*i.e.*,
341 RTR > 1). With this calculation procedure, the resulting ChemHQ value increases with the number
342 and the magnitude of exceeding endpoints, but is not strongly influenced by the number of chemicals
343 whose concentrations are below the respective reference levels (Piva et al. 2011). This quotient was
344 calculated for each site and each campaign, for the water column (ChemHQ_{water}) and the sediment
345 (ChemHQ_{sed}). A hazard class was finally assigned to each calculated ChemHQ value according to the
346 hazard classification grid established by Piva et al. (2011).

347 The contribution of each chemical and class of substance to the ChemHQ value was also calculated.

348

349 2.7.2. Line of evidence 2: bioavailability (LOE#2)

350 Bioavailability was assessed through bioaccumulation measurements in whole gammarids.

351 Concentrations measured in gammarids following the caging procedure with control food supply have
352 to be regarded as mainly proceeding from the water column rather than from the trophic route. Besse
353 et al. (2013) used the same experimental conditions to study the bioaccumulation levels of 11 MEs and
354 38 hydrophobic organic compounds (including PAHs, PCBs, PBDEs, and OCPs) in gammarids *G.*
355 *fossarum* of the same geographical origin as those used in the present study. In particular, the authors
356 established bioaccumulation thresholds for 35 substances permitting to reveal bioavailable
357 contamination of the environment when values go beyond these reference levels. These threshold
358 tissue concentrations were used to derive the reference levels for bioaccumulation data in the present
359 study. The list of the selected chemicals for LOE#2 and their respective reference values are presented
360 in ESM Table S3.

361 Based on the procedure of Piva et al. (2011), the calculation method applied to LOE#2 was quite
362 similar to LOE#1, with the calculation of an RTR value for each chemical and the weighting of these
363 values according to the status of that chemical within the WFD (see Section 2.7.1). One of the
364 differences between the two LOEs is related to the addition of a correction function ($Z(i)$) to take into
365 account the significance of the deviations from the reference values (Piva et al. 2011). As three
366 replicates were not available for every compound and reference value, we set the correction function
367 $Z(i)$ as a fixed factor, as proposed by the authors. A hazard class was attributed to each resulting RTR_w
368 value, as described by the authors.

369 The Bioavailability Hazard Quotient (BioavHQ) for each site and each campaign was calculated by
370 averaging the RTR_w values whose relative hazard was classified as 'slight' and summing the RTR_w
371 values with a 'moderate' to 'severe' hazard class, following the same reasoning as applied in LOE#1
372 (see Section 2.7.1). A global hazard class for bioavailability was attributed to each BioavHQ, as
373 described by Piva et al. (2011). As in the case of LOE#1, the contribution of each chemical and class
374 of substance to the BioavHQ value was calculated. Details on the complete calculation procedure
375 implemented in LOE#2 are presented in ESM Fig. S2.

376

377 2.7.3. Line of evidence 3: biomarkers (LOE#3)

378 The calculation method described by Piva et al. (2011) for biomarker LOE required a reference value
379 (or control value) and an effect (inhibition and/or induction) threshold for each biomarker.

380 Only unilateral differences in comparison to reference values were taken into account, in agreement
381 with their biological significance. As a result, only inhibition responses were taken into account for
382 AChE activity, feeding rate, digestive enzyme activity levels, and the number of oocytes and embryos
383 *per female*. In contrast, only induction effects were taken into account for mortality and molt delay
384 endpoints. However, bilateral differences could be taken into account in the case of biomarkers for
385 which both induction and inhibition responses have an (eco)toxicological/biological significance, as in
386 LOE#4 for the time to hatch of embryos during the MELA (see Section 2.7.4).

387 The calculations of reference and threshold values for gammarid **AChE activities and feeding rates**
388 were adapted from Xuereb et al. (2009) and Coulaud et al. (2011). They were based on gammarid
389 weight for AChE activity, and on the size of encaged gammarids and the mean temperature during *in*
390 *situ* exposure for feeding rates. The thresholds (Th) were calculated according to the unilateral lower
391 limit of the 95% confidence interval (CI95%) of the corresponding reference value (A. Chaumot,
392 *personal communication*):

$$393 \quad Th (\%) = \frac{Ref. value - CI95\%lower limit}{Ref. value} \times 100$$

394 For **digestive enzyme activity levels**, mean reference values were taken from Charron et al. (2013).
395 The corresponding thresholds were calculated as described above for AChE activity levels and feeding
396 rates.

397 The reference value for the **number of oocytes/embryos *per female*** after normalization of female
398 size was adapted from Geffard et al. (2010). The corresponding thresholds were calculated as
399 described above.

400 For the percentage of females with a **molt delay**, we expected the reference value to be 0%, and
401 calculated the threshold value as the percentage representing the presence of 2 asynchronous females
402 within a batch of 15 females *per site* and *per campaign* (*i.e.*, 13.3%). As a value equal to '0' is not
403 acceptable in the calculations of Biomarkers Hazard Quotient (BiomHQ; see details for calculations
404 below), '100' was added to the reference value (thus equal to 100%) and each measured molt delay
405 (the threshold remained unchanged).

406 The reference value, threshold, and effect retained for each biomarker are reported in ESM Table S4.
407 The complete calculation procedure was adapted from Piva et al. (2011), with a few modifications,
408 and is described in ESM Fig. S3.

409 Briefly, for each biomarker response, we calculated the percentage of variation relatively to the
410 reference (%VAR). The %VAR is then supposed to be corrected according to the statistical

411 significance of the difference between the reference value and the mean biomarker value ($Z(i)$)
412 function; Piva et al. 2011), resulting in an effect value ($E(i)$) for each endpoint. However, as the
413 threshold values (Th) were pre-established using a statistical approach (*vs.* evaluated by ‘expert
414 judgement’ in Piva et al. (2011)), we manually set the $Z(i)$ function, as we did for bioavailability data
415 (ESM Fig. S3). As mentioned above, only unilateral differences were taken into account. In other
416 words, when only inhibition was considered as ‘ecotoxicologically relevant’ for a given biomarker,
417 any induction effect was considered as ‘within the reference range’, resulting in an effect $E(i)$ set at
418 ‘0’, and *vice versa*. Moreover, to take into account the fact that the reference value of some biomarkers
419 could vary depending on exposure conditions over the year, it appeared more accurate to evaluate the
420 annual average response of a biomarker by averaging the $E(i)$ values calculated for each campaign
421 (*per station*) rather than averaging the biomarker responses directly.

422 A hazard class was attributed to each $E(i)$ value (ESM Table S5) according to the gradation scale of
423 Piva et al. (2011) (ESM Fig. S3). The effect value was then weighted ($E_w(i)$) against the biological
424 significance of the biomarker response (ESM Table S6). The Biomarker Hazard Quotient (BiomHQ)
425 for each site and each campaign was calculated by averaging the $E_w(i)$ values for which $E(i)$ relative
426 hazard was classified as ‘moderate’, and summing the $E_w(i)$ values for which $E(i)$ belonged to a
427 ‘major’ to ‘severe’ hazard class (ESM Fig. S3). The procedure was based on the reasoning applied in
428 LOE#1 (see Section 2.7.1) and LOE#2.

429 Finally, a global hazard class for biomarkers was attributed to the BiomHQ value for each site and
430 campaign, as proposed by Piva et al. (2011).

431

432 2.7.4. Line of evidence 4: bioassays (LOE#4)

433 The calculation method applied to derive the Bioassay Hazard Quotient (ToxHQ) was quite similar to
434 the one used for BiomHQ (Piva et al. 2011), and is described in ESM Fig. S4.

435 The reference values used in the present study were the measurements from the negative control
436 treatment of each bioassay. Responses from *in vitro* bioassays are usually expressed as induction or
437 inhibition factors in comparison to the control. Thus, these data were just slightly modified to
438 correspond to the percentage of variation relatively to the control value (%VAR(i)) defined in ESM
439 Fig. S4. Afterwards the effect $E(i)$ was calculated for each endpoint as described for biomarkers, using
440 threshold (Th) values and a correction factor $Z(i)$ (ESM Fig. S4).

441 For **MELA** results, Th values were calculated by examining the variability of the data from the
442 negative control treatment and the associated CI95%, as described above for biomarkers such as AChE

443 activity (see Section 2.7.3). The Z(i) function was consequently set similarly to BiomHQ calculations
444 (see Section 2.7.3).

445 For *in vitro* bioassays, the Th values were established based on expert judgement. For **Microtox**[®]
446 results, we set the threshold and the Z(i) function according to acute toxicity levels based on the
447 inhibition percentage (= % VAR(i)) of the bioluminescence recorded at the highest concentration.
448 Thus, the Th value set at 10% represented the ‘not toxic’/‘moderately toxic’ limit, according to our
449 laboratory expertise and adapted from Bennett and Cabbage (1992) and Brouwer et al. (1990). As a
450 result, data below this value were weighted by a Z(i) factor equal to 0.2 in the effect E(i) calculation.
451 Similarly, bioluminescence inhibition factors above 10% and below 50% resulted in a weighting Z(i)
452 value equal to 0.5. When the % VAR(i) was above 50% (*i.e.*, considered as ‘clearly cytotoxic’), the
453 Z(i) function was equal to 1.

454 A similar methodology was implemented for **SOS Chromotest** data using the threshold value
455 established by Mersch-Sundermann et al. (1992). As a result, the Th value was set at 50%, in
456 agreement with the induction factor threshold of 1.5 established as the ‘not genotoxic’/‘marginally
457 genotoxic’ limit by the authors. Moreover, the Z(i) function was set at (i) 0.2 for a non-significant
458 value in comparison to the blank and/or for an induction factor below 1.5, (ii) 0.5 for an induction
459 factor above 1.5 but strictly below 2, and (iii) 1 for an induction factor equal or superior to 2 (value
460 above which effects could be considered as clearly ‘genotoxic’ according to Mersch-Sundermann et al.
461 (1992)).

462 For **ED bioassays**, Th values of 50% and 10% were set for agonist (ER, TR, AR/GR) and antagonist
463 (anti-AR) activities, respectively, according to our laboratory expertise (Lucie Oziol, *personal*
464 *communication*). The Z(i) function was also manually set at (i) 0.2 for data below the LD or LQ
465 values, (ii) 0.5 for data not significantly different from the blank value (according to Student *t*-test
466 results with a 5% risk), (iii) 1 in all other cases.

467 Similarly to the reasoning applied in LOE#3 calculations, only unilateral differences in comparison to
468 the reference value were taken into account, in agreement with the biological significance of each
469 endpoint. Thus, any response with an effect other than the one defined as ‘ecotoxicologically relevant’
470 led to an effect E(i) set at 0, except for the time to hatch of medaka embryos for which bilateral
471 differences were taken into account.

472 The reference value, threshold, and effect of each endpoint are reported in ESM Table S7. As in the
473 case of biomarkers, the annual average response of a bioassay endpoint was obtained by averaging the
474 E(i) values calculated for each campaign at each station. Each effect value E(i) was then weighted
475 ($E_w(i)$) according to the corresponding bioassay endpoint. The weight of each response was defined as
476 proposed by Piva et al. (2011), with slight modifications. The WOE approach was first proposed by
477 these authors to assess sediment hazard in particular, with a low coefficient (0.3) for bioassays using

478 the water column as a test matrix. In the present study, we apply the procedure to both sediment and
479 water column hazard assessment. Consequently, we chose to set the coefficient for water column
480 testing similarly to what was used for sediment testing (*i.e.*, equal to 1 for total water and 0.8 for the
481 water-dissolved fraction). Moreover, as an ‘ecotoxicologically relevant’ effect was identified for each
482 marker and as ‘contrary’ effects were discarded from analysis, it seemed superfluous to weight
483 endpoints according to the possibility of hormetic responses. Details on weighting calculations for
484 each endpoint are given in ESM Table S8. Finally, the cumulative ToxHQ was calculated as the sum
485 of the $E_w(i)$, and a global hazard class for bioassays was attributed as described by Piva et al. (2011)
486 (ESM Fig. S4). This hazard quotient was elaborated for each site and each campaign, for the water
487 column ($ToxHQ_{water}$) and for the sediment ($ToxHQ_{sed}$).

488

489 2.7.5. Weight of Evidence integration

490 The complete calculation procedure implemented for WOE integration is detailed in ESM Fig. S5. As
491 described by Piva et al. (2011), the first step of the integration of the HQs derived from the four LOEs
492 within a global index (WOE index) consisted in normalizing HQ values to a common scale. The
493 authors also proposed to ascribe different weightings to LOE results according to their environmental
494 relevance. Thus, they chose to multiply BioavHQ indices by 1.2× to give greater importance to
495 bioavailability data as compared to the presence of chemicals in the abiotic compartment (*i.e.*,
496 ChemHQs, weighted by 1.0×). Similarly, they suggested to apply a 1.2×-coefficient to the data
497 acquired using bioassays (ToxHQ indices) because they reflected acute effects at the organism level,
498 whereas biomarker responses (BiomHQ indices) describing sublethal effects at the molecular scale
499 remained weighted by 1×. The situation was somewhat different in the present study, since biomarkers
500 included both responses at the molecular level (*e.g.*, enzyme activity levels) and life history traits (*e.g.*,
501 feeding behavior and reproduction ability). As a result, it seemed more relevant to apply greater
502 weightings to the results of the LOE related to disturbances of organisms exposed *in situ* than to
503 organisms exposed under laboratory conditions. We thus chose to weight the BioavHQ and BiomHQ
504 indices by 1.2×, whereas ChemHQs and ToxHQ indices were still weighted by 1×.

505 The resulting HQ indices from the four LOEs were summed up and normalized to 100% to yield an
506 overall WOE hazard index. Finally, each WOE value was assigned to a hazard class, as described by
507 Piva et al. (2011) (ESM Fig. S5).

508

509

510

511 3. Results and discussion

512 3.1. Water column and sediment chemistry: LOE#1

513 3.1.1. Chemical hazard quotient for the water column (ChemHQ_{water})

514 The chemical hazard relative to water column contamination was evaluated according to the
515 concentrations of 15 pesticides, 2 AKPs, 1 PFAS (PFOS), and 9 MEs measured in the dissolved
516 fraction. They are listed in ESM Table S1.

517 The concentration of each chemical (ESM Table S9) was used to calculate an integrative
518 contamination index, ChemHQ_{water} (Table 1). The contribution of each class of compounds to the
519 global chemical hazard is presented in Fig. 2a. These results show that perfluorooctanesulfonic acid
520 (PFOS) was omnipresent in the area and contributed to 51% (Marnay C4) to 99% (Bougival C2) of the
521 ChemHQ_{water} values. Average PFOS concentrations were above the EQS value of 6.5×10^{-4} $\mu\text{g/L}$ (EC
522 2013) at all sites and for all sampling campaigns as well as for calculated annual means; the RTR
523 values increased along the anthropogenic gradient from 1.30-4.37 in Marnay, 4.85-22.8 in Bougival,
524 and up to 5.88-35.2 in Triel (ESM Tables S1 and S9). At each site, the lowest RTR value was recorded
525 during the C4 campaign, and the highest during the C1 campaign. In contrast, the contribution of
526 pesticides increased at each site in the C4 campaign, with values around 48% at Marnay, 24% at
527 Bougival and 22% at Triel. Among the pesticides, one compound –metazachlor, a chloroacetanilide
528 herbicide- accounted for almost the total contribution of this class of chemicals (21% to 46%; data not
529 shown). The contamination gradient previously mentioned for PFOS was also recorded in the case of
530 metazachlor winter concentrations, with RTR values increasing from 1.55 in Marnay to 2.29 and 2.38
531 in Bougival and Triel, respectively (ESM Tables S1 and S9). Metazachlor is an herbicide commonly
532 used in rapeseed crops and usually applied in late August/early September. This substance is
533 considered as ‘moderately sorbing’, and several months might go by between its application date and
534 its release in the surrounding waters, depending on the intensity of the rain events and the
535 hydrodynamic characteristics of the watershed (Passeport et al. 2013).

536 As shown in Fig. 2a, metal elements only contributed noticeably to ChemHQ_{water} values at the
537 downstream sites, with the highest contributions recorded in the C4 campaign (around 11%), as shown
538 for pesticides. Despite a clear contamination gradient along the Seine river, the dissolved
539 concentrations of metals did not exceed, or only slightly exceeded their respective EQS at the two
540 downstream sites, as previously described (Faburé et al. 2015; Lebrun et al. 2015), and in agreement
541 with previous studies at the same sites (Fechner et al. 2012). Values exceeding the corresponding
542 reference value were limited and almost strictly related to copper concentrations, with a maximal RTR
543 value below 1.35 noted in Triel during the autumn campaign (ESM Tables S1 and S9).

544 Following the integration of overall contamination data measured in the water column, the resulting
545 ChemHQ_{water} values obviously reflected the anthropogenic gradient pressure along the Seine River
546 axis, with values increasing from the Marnay upstream site to the Bougival and Triel downstream sites
547 (Table 1). As a result, the chemical hazard for the annual average value was classified as ‘moderate’ at
548 Marnay and as ‘severe’ at the two stations downstream of the Paris agglomeration. Seasonal variations
549 in water column contamination were also evidenced, with lower ChemHQ_{water} values at all sites in
550 winter (C4), thus downgrading the hazard class for the most impacted sites from ‘severe’ to ‘major’. In
551 contrast, the highest ChemHQ_{water} values were recorded in the fall season (C1) at each station (Table
552 1), possibly as a consequence of the lower dilution of point source discharges under low flow
553 conditions in the River Seine.

554

555 3.1.2. Chemical hazard quotient for sediment (ChemHQ_{sed})

556 Sediment contamination was assessed based on the concentrations of 1 PFAS (PFOS), 18 PAHs, 8
557 PCBs, 7 PBDEs, 12 OCPs, and 15 MEs. They are listed in ESM Table S2, and detailed in ESM Tables
558 S10 and S11).

559 Considering that for some chemicals such as PAHs, PBDEs and DDTs, the reference values were
560 available both for individual substances and for the total concentrations of the classes of compounds, it
561 was possible to calculate ChemHQ_{sed} values in two different ways (Table 1). The calculation method
562 strongly influenced the relative contribution of each class of compounds to the calculated global index
563 (Fig. 2b and 2c). When ‘individual concentrations’ were used, the most contributive chemical family
564 was PAHs at each site, and sampling time (except for Bougival C1 and Triel C1/C4 samples) with
565 total contributions varying from 75% to 99% in Marnay, 52% to 74% in Bougival, and 75% to 99% in
566 Triel (Fig. 2b). Among this class of compounds, the corresponding reference values were widely
567 exceeded for phenanthrene and pyrene, with RTR values respectively between 3.0-18.9 and 2.9-15.5
568 in Marnay, 15.6-57.3 and 18.6-75.4 in Bougival, and between 5.7-836 and 12.1-672 in Triel (ESM
569 Tables S2 and S10). RTR values around 200 were also noted for anthracene and benzo[a]anthracene in
570 Triel C3 samples (ESM Tables S2 and S10). As in the case of PAHs, OCPs were omnipresent in the
571 area, with overall contributions reaching 25% at Marnay, 57% at Bougival, and 74% at Triel (Fig. 2b).
572 These high contribution levels were mainly attributable to heptachlor concentrations that exceeded the
573 reference value of 0.02 µg/kg dw 7.7- to 16.9-fold in Marnay, 46.7- to 136-fold in Bougival, and 109-
574 to 487-fold in Triel (ESM Tables S2 and S10).

575 When PAH concentrations were summed (ΣPAHs) and compared to the reference value for total
576 PAHs, the resulting RTR values were much lower than those described above for individual
577 substances. They only varied from 0.35 to 1.70 in Marnay, 2.34 to 9.75 in Bougival, and 1.33 to 65.5

578 in Triel samples (ESM Tables S2 and S10). As a result, the contribution of PAHs to ChemHQ_{sed}
579 values was also low, with values around 10% or lower for all samples, except Triel C3 for which
580 ΣPAHs accounted for about 73% of the calculated hazard quotient (Fig. 2c). Consequently, the global
581 relative contributions of OCPs and MEs logically increased with this calculation method (Fig. 2c).

582 The reference values for MEs in sediment were only exceeded in downstream samples. These overruns
583 were systematic and particularly substantial for Cd, Cu, Pb, and Zn, with RTR values reaching 23.8,
584 34.1, 20.7, and 39.1, respectively (Triel and Bougival samples combined; ESM Tables S2 and S10).
585 Higher enrichment factors (in relation to the geochemical background) have been reported for these
586 elements in sediment cores sampled downstream of the Paris conurbation as compared to upstream
587 and Oise River sites (Le Cloarec et al. 2011). The sites downstream of Paris receive and integrate all
588 kinds of pollutants that affect the rest of the Seine Basin. They result from industrial (*e.g.*, foundries,
589 wire factories) and agricultural activities (*e.g.*, the use of CuSO₄ as a fungicide and bactericide in
590 vineyards), but also intense urbanization (*e.g.*, the use of leaded gasoline, leaching of old Zn roofs
591 following rainfalls, effluents from waste water treatment plants (WWTPs)) (Le Cloarec et al. 2011;
592 Ayrault et al. 2012). The situation is particularly worsened at sites such as Triel, situated downstream
593 of the Oise river confluence: Triel is not only affected by the inputs from the Paris suburbs and
594 upstream activities, but also by the high industrialization of the Oise basin and one of the most
595 important sewage plants of the Paris agglomeration, the Seine-Aval WWTP of Achères located on the
596 banks of the Seine River, between the Bougival and Triel sampling sites (Le Cloarec et al. 2011;
597 Ayrault et al. 2012).

598 Similarly to what was observed for the water column, the integrative ChemHQ_{sed} values clearly
599 illustrated the expected contamination gradient; however, values calculated using individual substance
600 concentrations were significantly higher than those calculated using total concentrations (Table 1).
601 While the hazard class remained 'severe' for the two downstream stations using either calculation
602 method (for all campaigns and the annual average value), the hazard status at Marnay varied from
603 'severe' to 'absent' for the C2 sample and from 'moderate' to 'absent' for the C4 sample when the
604 ChemHQ_{sed} value was calculated using individual substances or total concentrations, respectively
605 (Table 1). The annual average hazard quotient for sediment chemistry was calculated using the mean
606 concentration of each chemical from the three composite sediment samples collected in the field
607 during the C1, C3, and C4 campaigns. Depending on the calculation method, the annual average
608 hazard class at Marnay varied from 'severe' to 'major' (Table 1).

609 According to MacDonald et al. (2000), ΣPAHs can be efficiently used to predict sediment toxicity,
610 with no substantial difference with toxicity predictions based on individual PAH concentrations.
611 However, in the approach developed here, the use of the ΣPAHs did not allow us to get access to
612 information on the individual compounds involved in exceeding the reference value. Identifying them

613 could nonetheless be valuable to identify the specific chemicals involved in the biological effects
614 highlighted in LOE#3 and LOE#4. The information could also be exploited during further
615 investigations aimed at identifying the source(s) of this pollution. Moreover, we preferred to adopt the
616 most conservative and protective approach for aquatic biota as regards the environmental hazard. As a
617 result, the ChemHQ_{sed} values calculated using individual concentrations were kept for the subsequent
618 step, (*i.e.*, final integration into the WOE index).

619

620 3.2. Bioavailability of chemicals: LOE#2

621 Bioavailability Hazard Quotient (BioavHQ) values were calculated according to the bioaccumulated
622 concentrations of 17 PAHs, 8 PCBs, 1 PBDE, 7 OCPs, and 4 MEs (listed in ESM Table S3, and
623 detailed in ESM Table S12) in caged gammarids.

624 The relative contributions of each class of chemicals to the global BioavHQ are illustrated in Fig. 3a.
625 Accumulation levels of organic compounds were not analyzed during the C4 campaign, and PAH and
626 PBDE data are missing for the C1 campaign, so contributions to the BioavHQ values are only
627 discussed based on annual average data.

628 Among trace metals, Ni was noticeably accumulated by gammarids exposed *in situ*, namely 1.7- to
629 5.4-fold higher than the reference level, with no specific variation among sites attributable to the
630 anthropogenic gradient (ESM Tables S3 and S12). This could result from a non-identified diffuse
631 contamination source, and/or Ni geochemical background differences between the native region of
632 gammarids (the Rhône-Alpes region) and the Seine Basin. In the water column, the Ni background
633 seemed to be slightly lower in the Rhône watershed (mainly between 0.58 and 2.51 µg/L, and more
634 locally up to 3.93 µg/L) than in the Seine Basin (2.51-3.93 µg/L), according to the FOREGS
635 Geochemical Baseline Mapping Program (<http://weppi.gtk.fi/publ/foregsatlas/>). Thus, while Ni locally
636 reached similar background levels in the Rhône watershed as in the Seine Basin, the globally lower Ni
637 geochemical background in the Rhône Basin may partially explain the higher Ni bioaccumulation
638 levels in transplanted gammarids. Yet excess Ni as compared to the reference value reflected an
639 increase in the Ni bioavailable fraction between the two areas whatever the exact origin (higher
640 geochemical background and/or anthropogenic activities).

641 In contrast, whereas RTR values for Pb accumulation at Marnay remained around or below 1 (annual
642 average: 0.94), they ranged between 1.7 and 7.8 in Bougival, and between 1.8 and 3.8 in Triel (ESM
643 Tables S3 and S12), reflecting a significant increase in Pb bioaccumulation in gammarids downstream
644 of Paris. Nevertheless, metals contributed only little (< 10%; Fig. 3a) to the annual average BioavHQ
645 values of the downstream sites. As a result, ME accumulation in gammarids exposed along the Seine
646 axis represented a limited hazard in comparison to organic compounds, as showed by the RTR_w-based

647 hazard status mainly evaluated as ‘absent’ or ‘slight’ for MEs (only one ‘major’ status and one
648 ‘moderate’ status were recorded for Ni accumulation; ESM Table S3).

649 The annual average RTR_w-based hazard classification was ‘absent’ for all organic compounds at
650 Marnay (except for PCB 118 classified as ‘slight’), whereas bioavailability-related hazard reached the
651 ‘major’ to ‘severe’ grade at the downstream stations (ESM Table S3). More specifically, PAHs
652 accounted for up to 87% and 72% in the calculation of annual average BioavHQ values in Bougival
653 and Triel, respectively. Still considering annual average data, all PAHs significantly accumulated (at
654 least 2.5-fold; ESM Tables S3 and S12) in gammarids exposed at Bougival as compared to the
655 reference levels. Among these compounds, respective reference bioaccumulation levels were exceeded
656 around 10-fold or just above for acenaphthene, anthracene, benzo[e]pyrene and phenanthrene, nearly
657 20-fold for benzo[a]anthracene and chrysene/triphenylene, and an extreme >30-fold for fluoranthene
658 and pyrene (ESM Tables S3 and S12). Annual average RTR values were lower in Triel-exposed
659 gammarids, mainly falling between 1.2 and 9.2 (ESM Tables S3 and S12). Bioaccumulation levels
660 more than 10-fold the reference values were only recorded for fluoranthene (11.1×) and pyrene
661 (15.4×) at that site (ESM Tables S3 and S12), confirming that these compounds were the most
662 bioaccumulated ones when compared to the reference levels.

663 These results are in overall agreement with the PAH concentrations measured in sediment, since
664 particularly high RTR values were recorded for anthracene, phenanthrene, benzo[a]anthracene, and
665 pyrene (see Section 3.1.2). However, PAH bioaccumulation levels were higher at Bougival than at
666 Triel, whereas PAHs were globally more abundant in Triel sediment than in Bougival sediment
667 (organic carbon-normalized concentrations increased 1.7-fold to 5.4-fold between the two sites
668 depending on the sampling campaign; ESM Table S11). This was likely related to variations in PAH
669 bioavailability between the two sites.

670 PCBs ranked second among the chemicals contributing to the annual average BioavHQ values at the
671 downstream sites, *i.e.* 11% in Bougival and 15% in Triel (Fig. 3a). The annual average RTR values
672 ranged between 2.7 and 6.7 in Bougival, and between 2.5 and 5.1 in Triel, depending on the congener
673 (ESM Tables S3 and S12). As a result, the RTR_w-based hazard was ‘moderate’ for PCB 50+28, PCB
674 52 and PCB 101, and ‘major’ for PCB 118 in Bougival (ESM Table S3). As regards Triel, only PCB
675 101 and PCB 118 accumulation levels represented a substantial hazard (*i.e.*, above the ‘slight’ status)
676 evaluated as ‘moderate’ and ‘major’, respectively.

677 The overall LOE#2 results tend to suggest that PAHs were the main problematic class of compounds
678 regarding their potential bioavailability/bioaccumulation in aquatic organisms in the two stations
679 located downstream of Paris.

680 The integration of overall bioavailability data resulted in lower BioavHQ values for the winter
681 campaign (C4) at all sites (Table 2). The hazard associated with bioaccumulation levels was
682 consequently evaluated as ‘slight’ for this campaign. However, the conclusion on the C4 campaign
683 should be interpreted carefully because it was only based on ME accumulation, and ME accumulation
684 proved to be limited as compared to organic compounds at the other sampling periods.

685 With BioavHQ values varying between 1.8 and 13.4, the related hazard was also classified as ‘slight’
686 in Marnay for each sampling period as well as for the annual average value (Table 2). In contrast, the
687 highest BioavHQ values were recorded in Bougival (except for C4) and the associated hazard was thus
688 evaluated as ‘major’ (Table 2). The hazard was also identified as ‘major’ during the fall campaign
689 (C1) at Triel; however, it decreased to ‘moderate’ for the other campaigns (C2 and C3) as well as for
690 the annual average BioavHQ (Table 2).

691

692 3.3. Biomarker responses in gammarids exposed *in situ*: LOE#3

693 Biological responses in gammarids exposed *in situ* along the Seine axis were investigated using
694 several biomarkers of various physiological impairments, including neurotoxicity, energy acquisition
695 disturbance, feeding behavior impairment, reproduction dysfunctioning/failure, and survival (ESM
696 Table S13). These markers have been studied in our laboratories for several years and are commonly
697 used in laboratory and field experiments (Dedourge-Geffard et al. 2013; Xuereb et al. 2009; Geffard et
698 al. 2010; Coulaud et al. 2011; Charron et al. 2013; Chaumot et al. 2015). The substantial insights
699 gained from these numerous studies more particularly allowed us to (i) fully determine the basal level
700 and variation range of each marker, (ii) identify and characterize the confounding factors that may
701 modulate biomarker responses, especially under field conditions, and thus (iii) determine specific
702 reference levels and effect thresholds for these biomarkers adapted to *in situ* deployment conducted in
703 the context of an environmental survey (ESM Table S4). These reference values and thresholds were
704 used to calculate biomarker Hazard Quotients (BiomHQs) applied to the Seine axis case study (Table
705 3).

706 The contribution of each category of markers to the calculated BiomHQ values is illustrated in Fig. 3b.
707 This analysis is completed by the hazard class attributed to each biomarker response, presented in
708 ESM Table S5.

709 No sign of neurotoxicity was highlighted: no significant inhibition of AChE activity was recorded in
710 comparison to the established reference value (ESM Tables S4 and S13); the hazard relative to the
711 AChE marker was ‘absent’ for all sites and all sampling campaigns (ESM Table S5).

712 Unlike AChE activity, the decrease in gammarid survival represented a ‘severe’ hazard at all sites and
713 all sampling periods (ESM Table S5), except in the C4 campaign during which no significant
714 mortality was noted (<4% in all stations; ESM Table S13). This acute toxicity endpoint was the only
715 that showed an effect E(i) above 1 at Marnay (representing a 1.8- to 3.2-fold increase in comparison to
716 the reference value, ESM Tables S4 and S13), therefore it contributed 100% to the calculated
717 BiomHQ value for that station (Fig. 3b). The impact on gammarid survival at Marnay raises the
718 question whether some early (sublethal) responses of other physiological functions, not addressed in
719 the present study, could potentially exist. For instance, studying the impacts of exposure on the
720 immune system and the inflammatory mechanisms could be of great interest by bringing supplemental
721 data to the currently investigated biomarkers. Due to their direct implications in individual fitness,
722 population and ecosystem health (Bols et al. 2001), immunomarkers are considered as attractive non-
723 specific markers that could be consistently integrated into ERA and biomonitoring surveys (Bado-
724 Nilles et al. 2015).

725 Mortality was modulated in the same order of magnitude at the downstream sites as at Marnay, with a
726 1.5- to 2.5-fold rise at Bougival, and a 1.6- to 3.0-fold rise at Triel (ESM Tables S4 and S13).
727 However, it resulted in contributions to BiomHQ values that were more limited than at Marnay and
728 ranged between 43% to 61% at Bougival, and 32% to 45% at Triel (Fig. 3b). Several other markers
729 were significantly modulated in the gammarids exposed at the two downstream sites.

730 Digestive enzyme activities were particularly down-regulated in the C1 and C3 sampling periods, but
731 were not significantly modulated during the winter (C4) campaign (‘absent’ hazard for all markers;
732 ESM Table S5). Overall, inhibition of enzymatic activities ranged between 17% and 37% at Bougival
733 and Triel during the fall (C1) and summer (C3) campaigns (ESM Tables S4 and S13). Special cases
734 were recorded at Bougival: no significant inhibition of cellulase activity was noted during the C3
735 campaign, whereas trypsin inhibition increased by 60% of the reference value at the same time (ESM
736 Tables S4 and S13). Still based on the digestive enzyme activities, the annual hazard (assessed by
737 averaging the effect E(i) of the three sampling periods) related to energy acquisition parameters was
738 evaluated as ‘moderate’ at Bougival, and ranked from ‘moderate’ to ‘severe’ at Triel (ESM Table S5).
739 However, this class of markers only slightly contributed to the global hazard related to biomarker
740 responses, with a maximal contribution barely exceeding 20% for the Triel C3 sampling point (Fig.
741 3b).

742 In contrast to energy acquisition markers, feeding rates were the most severely inhibited in gammarids
743 exposed to the downstream stations during the winter (C4) campaign, with values representing 92%
744 and 79% decreases in comparison to the corresponding reference value at Bougival and Triel,
745 respectively (ESM Tables S4 and S13). These effects accounted for 100% of the calculated BiomHQ
746 at the downstream sites during the C4 sampling period, and represented a ‘severe’ hazard (Fig. 3b,

747 ESM Table S5). Gammarid feeding activity was also repressed during the summer campaign, up to
748 21% at Bougival ('moderate' hazard) and 49% at Triel ('severe' hazard'). As a result, a 'major' annual
749 hazard was attributed to the effects on feeding behavior in gammarids exposed downstream of Paris
750 (ESM Table S5).

751 Among the reproductive impairment biomarkers, the number of oocytes *per* female did not decrease in
752 gammarids exposed along the Seine axis. Inversely, *in situ* exposure yielded values up to 2-fold higher
753 than the expected value (ESM Tables S4 and S13). Therefore the environmental hazard associated to
754 this marker was 'absent' for all sites and all campaigns (ESM Table S5). However, the number of
755 embryos *per* female decreased by 33% at Bougival and 71% at Triel during the C1 campaign (ESM
756 Tables S4 and S13), representing a 'severe' potential hazard to aquatic organisms (ESM Table S5).
757 Even more pronounced impacts were highlighted on molt delay, which noticeably increased during all
758 campaigns, by 27% to 57% at Bougival and by 40% to 50% at Triel (ESM Tables S4 and S13). The
759 overall environmental hazard was classified as 'severe' for both stations for this reproductive marker
760 (ESM Table S5). Reproductive impairments consequently represented the second most contributive
761 class of biomarkers to BiomHQ values for the C1 and C3 campaigns and the annual average value,
762 with contributions varying between 20% and 40-45% at Bougival and Triel (Fig. 3b).

763 All biomarkers were finally integrated into a global hazard quotient. The resulting BiomHQ values
764 clearly depicted the expected anthropogenic gradient, with values increasing from Marnay to Bougival
765 and then Triel, for each sampling campaign as well as for the annual average value (Table 3). The
766 lowest BiomHQ was recorded in the C4 campaign at all sites; the hazard was classified as 'slight'
767 (Marnay) or 'major' (Bougival, Triel). However, as in the case of bioavailability, these results should
768 be interpreted carefully because reproduction markers and cellulase activity were not investigated
769 during that campaign. The C4 data did not fulfil the minimum requisite advised by Piva et al. (2011) to
770 calculate the cumulative BiomHQ: only two markers had a weighting above 1.2, while the
771 recommended number is 3. Conversely, the highest BiomHQ values were noted in the fall (C1)
772 campaign (Table 3). This observation is consistent with the water column and sediment ChemHQ
773 values (using 'total concentrations'), which were also higher in C1 samples. In agreement with the
774 hazard class established from the other sampling periods, the overall level of risk (based on annual
775 average estimation) was identified as 'moderate' at Marnay, 'major' at Bougival, and 'severe' at Triel
776 (Table 3), reflecting a noticeable increase of physiological disturbances in gammarids exposed along
777 the Seine axis. A similar gradient of biological effects was reported in caged zebra mussels following
778 exposure at the same sampling sites in winter, spring, and summer (Michel et al. 2013). Genotoxicity
779 markers (DNA strand breaks and micronucleus frequency) significantly increased from Marnay to
780 Bougival and then Triel. Seasonal variations of the responses were also highlighted for DNA strand
781 breaks, with the lowest levels recorded in winter as compared to summer and spring (Michel et al.
782 2013). All these observations are in good agreement with the conclusions drawn from our LOE#3

783 integration results, suggesting that the effect gradient and seasonal trends (with lower biological
784 disturbances in winter) are constant year in, year out.

785 Overall, our results demonstrate that the selected set of biomarkers efficiently reflected biological
786 disturbances in gammarids following exposure along the Seine axis. The use of markers at different
787 levels of biological organization, from the molecular level to life history traits, allowed us to
788 discriminate among sites and observations. It also highlighted that *in situ* exposure differentially
789 affected various physiological mechanisms depending on the level of anthropogenic pressure and the
790 sampling period, thus demonstrating the complementarity of the selected endpoints. The reference
791 values and biomarker thresholds used in the present study were specifically established according to
792 the characteristics of the transplanted gammarid population (*e.g.*, size and weight) and some abiotic
793 exposure parameters (*e.g.*, temperature). The use of such reference values and thresholds improved the
794 reliability of the environmental diagnosis by integrating response variations due to the physiological
795 state of gammarids and/or to field exposure conditions other than chemical pressure. Such finer
796 characterization of the reference state is clearly needed when biological responses following
797 seasonally varying contamination are studied. It can therefore be assumed that site qualification in
798 terms of environmental hazard/risk is more relevant and robust using these adaptive references rather
799 than more generic ones. For example, a 20-30% inhibition threshold for AChE activity is generally
800 admitted in the literature for freshwater and marine invertebrates (Escartín and Porte 1996; Owen et al.
801 2002), but the value established and used in the present study was substantially lower (12%) (Xuereb
802 et al. 2009). This type of methodology could be applied to responses analyzed within other LOEs such
803 as bioavailability and bioassays, to refine and adjust the conclusions of the strategy (*e.g.*, a WOE
804 approach) implemented to describe the ecological state of an aquatic environment.

805

806 3.4. Laboratory bioassays: LOE#4

807 The ecotoxicological diagnosis of the area was completed using a battery of laboratory bioassays
808 performed on water column and sediment samples whose responses were integrated into the WOE
809 approach to calculate cumulative hazard indices (ToxHQs) in the two abiotic compartments.

810

811 3.4.1. Bioassay hazard quotient in the water column (ToxHQ_{water})

812 Aqueous samples were tested for embryotoxicity and teratogenicity using MELA (dissolved fraction)
813 and for endocrine-disrupting potency using cellular *in vitro* bioassays on organic extracts. Responses
814 were selected among the various endpoints monitored during the MELA according to their reliability,
815 relevance and sensitivity to characterize survival (embryonic and larval survival rates), *in ovo*

816 development (hatching success and time to hatch) and growth (total body length and head size at
817 hatching), and teratogenicity (total percentage of abnormal larvae). These responses (ESM Table S14)
818 were integrated into the WOE approach to calculate $ToxHQ_{water}$ indices for each site and each
819 sampling period (Table 4). The contributions of the different classes of endpoints to the global hazard
820 quotient are shown in Fig. 4a.

821 Exposure of medaka ELS early life stages (ELS) to the dissolved fraction of the water samples
822 collected at Marnay did not induce strong deleterious effects. Only a slight decrease (around 7%) in
823 embryonic and larval survival rates exceeding the established thresholds was recorded in the C1
824 campaign (ESM Tables S7 and S14). Similarly, a slight reduction of the total body length of larvae at
825 hatching was noted in medaka exposed to C1 and C3 samples, representing less than a 4% fall in
826 comparison to the reference value (ESM Tables S7 and S14). The greatest modulations as compared to
827 the reference at Marnay were highlighted by ER induction factors exceeding the corresponding Th
828 value by 4.3- to 5.6-fold (ESM Tables S7 and S14). However, ER agonist activity was only evaluated
829 as significant (in comparison to the blank) in the C3 samples (ER induction factor: 3.26 \times ; ESM Table
830 S14). As a result, survival, growth, and endocrine disruption were the most contributive classes of
831 endpoints to the global hazard quotient in bioassays on water column samples from Marnay (Fig. 4a).
832 In agreement with this limited effect, the $ToxHQ_{water}$ values were low (< 5) at Marnay, and the
833 resulting hazard class for bioassays was 'absent' for all sampling periods as well as for the annual
834 average value (Table 4).

835 Similar results were obtained for the Bougival C3 and C4 samples, with only slight effects on survival,
836 growth, and estrogenic potency (ESM Table S14), so that the resulting hazard was classified as
837 'absent' for these periods (Table 4). In contrast, the C1 water sample from that station proved much
838 more problematic as it strongly increased embryonic mortality (46%), and to a lesser extent larval
839 mortality (14%) (ESM Table S14). Similarly, hatching success was reduced by 50% in comparison to
840 the control. Teratogenicity also increased 4.4-fold as compared to the reference value, so that more
841 than 80% of the hatchlings exhibited developmental abnormalities (ESM Tables S7 and S14).
842 Additionally, ER activity was between 6.1 and 9.2 times higher than the established Th value (ESM
843 Tables S7 and S14). However, these endocrine-disrupting effects did not contribute much to the global
844 hazard (2%) because survival, development and teratogenicity accounted for 63%, 16% and 12% of
845 the calculated $ToxHQ_{water}$, respectively (Fig. 4a). These effects account for the downgrading of the
846 hazard associated to bioassay responses to the 'moderate' status (very close to the 'major' class) for
847 the C1 campaign at Bougival. Nevertheless, in relation to the biological responses recorded in the fall
848 (C1) campaign, the annual average hazard class was classified as 'slight' at Bougival (Table 4).

849 The Triel response profile was quite similar to that of Bougival, yet more pronounced. Bioassays on
850 water samples from the winter (C4) campaign only revealed an impact on ED parameters, with

851 induction of ER activity 6.5-fold higher than the reference value (11 times the Th level; ESM Tables
852 S7 and S14). A slight induction of the AR/GR agonist activity was also noted, but it only represented a
853 1.83-fold increase as compared to the reference level (ESM Tables S7 and S14). These responses
854 accounted for 74% of the cumulative ToxHQ_{water} in Triel-C4 (Fig. 4a), but resulted in a global hazard
855 classified as ‘absent’ (Table 4). ER and AR/GR agonist activities were induced with the same order of
856 magnitude following cell exposure to Triel-C1 water organic extract (10.4× and 1.9× the
857 corresponding Th value, respectively; ESM Tables S7 and S14). However, endocrine-disrupting
858 effects only accounted for 7% of the calculated ToxHQ_{water} at that site (Fig. 4a), as stronger impacts
859 were detected in medaka ELS following exposure to the dissolved fraction of the water sample
860 collected at Triel during the fall campaign. As observed for Bougival during the same sampling
861 campaign, exposure led to a strong increase of the mortality rate of exposed embryos (only one third
862 of the embryos were still alive at the end of the experiment; 13.3 times the Th value; ESM Tables S7
863 and S14). Moreover, hatching success was reduced by more than 70% in comparison to the control (~7
864 times the Th value), and the mean time to hatch was delayed by about 25.5% (1.6 times higher than
865 the Th value) as compared to the reference (ESM Tables S7 and S14). An impact on medaka *in ovo*
866 development was also reflected by the total body length of hatchlings, which was reduced by 5.4% in
867 comparison to control organisms (about 4 times the Th value; ESM Tables S7 and S14). However, this
868 effect on *in ovo* growth was quite limited as it only accounted for 7% of the global hazard quotient
869 ToxHQ_{water}, whereas survival and development alterations respectively accounted for 61% and 25%
870 (Fig. 4a). The resulting hazard attributed to the ToxHQ_{water} index for Triel-C1 was identified as
871 ‘major’ (Table 4). The growth of medaka embryos exposed to the Triel-C3 water sample was also
872 slightly reduced: larvae were 3.4% shorter than the controls (2.4 times the Th value; data not shown).
873 Nevertheless, the summer (C3) sample from Triel was particularly marked by a strong induction of
874 both ER and TR activities, exceeding the corresponding Th values by 21.8× and 25.6× respectively
875 (equivalent to respective induction factors of 11.9 and 13.8 as compared to the control; ESM Tables
876 S7 and S14). As a result, endocrine-disrupting effects contributed to 80% of the cumulative
877 ToxHQ_{water} calculated for Triel-C3 (Fig. 4a), and the global hazard evaluated using bioassays was
878 summarized as ‘slight’ (Table 4). Annual averaging of the effects observed during bioassays resulted
879 in a global ToxHQ_{water} classifying the hazard as ‘slight’ (but very close to the ‘moderate’ class) for
880 Triel (Table 4). The ToxHQ_{water} value was mainly due to impacts on survival (44%) and *in ovo*
881 development (19%) of medaka ELS, as well as to endocrine-disrupting effects (27%) on specific cell
882 lines (Fig. 4a).

883

884 3.4.2. Bioassay hazard quotient of sediment (ToxHQ_{sed})

885 As in the case of the water column, ED assays were performed on sediment organic extracts, and the
886 MELAc (MELA adapted to sediment testing by direct contact with particles) was implemented on
887 whole sediment samples. In addition, cytotoxicity and genotoxicity of sediment elutriates were
888 investigated using the Microtox and SOS Chromotest procedures, respectively. As previously
889 mentioned, the responses of all the bioassays (ESM Table S14) were integrated into a common hazard
890 quotient, $ToxHQ_{sed}$ (Table 4), and the contribution of each endpoint class to the hazard index was
891 calculated (Fig. 4b).

892 Medaka exposure to Marnay sediments only resulted in limited sublethal effects on fish ELS. A slight
893 delay in the average time to hatch of embryos was noted following exposure to the sediment sampled
894 during the three campaigns (C1, C3 and C4), representing 2 to 5% variation in comparison to the
895 control values (ESM Tables S7 and S14). Moreover, a significant increase in the percentage of
896 abnormal larvae at hatching (around 2 times higher than the control, *i.e.*, around 35.5% of abnormal
897 individuals; ESM Tables S7 and S14) was noticed after exposure to Marnay sediment collected during
898 the fall (C1) period. *In ovo* growth of embryos exposed to C3 and C4 samples also slightly decreased:
899 biometric measurements were lower than the reference values, especially hatchling head size for
900 which the highest percentage of variation relative to the control was recorded (~8%; ESM Tables S7
901 and S14). The cytotoxicity of sediment elutriates remained moderate for Marnay C3 and C4 samples
902 (less than 30% inhibition of bioluminescence), but was greater during the fall (C1) campaign with
903 50% inhibition (ESM Tables S7 and S14). The strongest effects were noted for ER agonist activity of
904 sediment organic extracts, which noticeably increased by 8.5- and 9-fold, following exposure to C3 and
905 C1 organic extracts, respectively, and to a lesser extent following exposure to the C4 sample (4×)
906 (ESM Tables S7 and S14). As a result, development, growth, endocrine disruption, and cytotoxicity
907 were the most contributive endpoint classes to the global hazard quotient $ToxHQ_{sed}$. They accounted
908 for 26%, 28%, 18%, and 18% of the annual average calculation of $ToxHQ_{sed}$, respectively (Fig. 4b).
909 The resulting hazard was summarized as ‘slight’ at Marnay for all the sampling periods as well as for
910 the annual average value (Table 4).

911 The bioassay response profiles at the Marnay and Bougival stations were very similar, as illustrated by
912 the endpoint contributions to the $ToxHQ_{sed}$ values (Fig. 4b). However, the response was clearly greater
913 downstream of Paris. For example, the time to hatch of embryos exposed to Bougival C3 sediment
914 was 24% longer than in the control treatment, and more than 50% of the larvae showed developmental
915 abnormalities on average following exposure to Bougival C1 sediment (ESM Tables S7 and S14).
916 Whereas elutriates from the Bougival fall (C1) and winter (C4) samples inhibited bioluminescence by
917 around 40%, the Bougival C3 elutriate reduced it by up to 96% as compared to the reference value
918 (ESM Tables S7 and S14). Similarly, TR and AR/GR agonist activities as well as AR antagonist
919 activity were enhanced with Bougival samples as compared to Marnay samples: they exceeded the
920 established respective Th values by 2.1× to 4.3×. As observed for Marnay, ER agonist activity showed

921 the greatest induction as compared to the reference level: between 7.7-fold in the winter (C4)
922 campaign and 32-fold in the summer (C3) campaign (ESM Tables S7 and S14). As a result, the
923 $ToxHQ_{sed}$ values for Bougival were higher than for Marnay. While the global hazard remained
924 classified as ‘slight’ for the C1 and C4 campaigns (although close to the ‘moderate’ hazard limit;
925 Table 4), it reached the ‘moderate’ grade for Bougival C3 due to the strong effects highlighted by
926 bioassays. In agreement with these observations, the annual global hazard associated with bioassay
927 responses was summarized as ‘moderate’ for Bougival (Table 4).

928 According to bioassay endpoints, the greatest toxicity was recorded at Triel. Signs of acute toxicity
929 were highlighted in medaka ELS following exposure to C1 and C3 sediment: embryonic survival
930 decreased by 13% as compared to the reference value in Triel C1, and hatching success did not exceed
931 1.3% and 0% in Triel C1 and C3, respectively (ESM Tables S7 and S14). Moreover, the few resulting
932 hatchlings finally died before the end of the experiment (data not shown). Consequently, teratogenicity
933 and *in ovo* growth were only evaluated for the Triel winter (C4) sample. Forty-nine percent of the
934 larvae exposed to this sediment displayed developmental deformities, and their growth was reduced by
935 8% to 9% depending on the endpoint (ESM Tables S7 and S14). The cytotoxicity of Triel elutriates
936 was very high in the C1 and C3 samples: bioluminescence was inhibited by 96% and 94%.,
937 respectively. In contrast, cytotoxicity was very limited in the Triel C4 sample: only 25% inhibition of
938 bacteria bioluminescence was noted (ESM Table S14). Modulation of endocrine-disrupting responses
939 following cell line exposure to Triel sediment organic extracts was of the same order of magnitude as
940 in Bougival sediment. TR, AR/GR and Anti-AR activity induction exceeded the respective Th values
941 by 1.9- to 5.3-fold, depending on the endpoint and the campaign (ESM Tables S7 and S14). In
942 addition, ER agonist activity modulation was the strongest, with induction factors ranging between
943 4.1-fold in the winter (C4) campaign and 30-fold in the summer (C3) campaign (ESM Tables S7 and
944 S14). Unsurprisingly, the hazard for toxicity assessed through laboratory bioassays was evaluated as
945 ‘severe’ for Triel C1 and C3 sediment (Table 4). The survival and development endpoint classes
946 respectively accounted for 51% and 28% of the global $ToxHQ_{sed}$ value in Triel C1. Development
947 (47%), cytotoxicity (23%), and endocrine disruption (20%) were the most contributive endpoint
948 classes to the Triel C3 $ToxHQ_{sed}$ value, whereas growth (41%) and teratogenicity (30%) accounted for
949 the main part of the Triel C4 $ToxHQ_{sed}$ value (Fig. 4b). Due to the limited effects of this latter sample,
950 the corresponding hazard was evaluated as ‘slight’. Finally, based on bioassay responses we concluded
951 that Triel sediment represented a ‘major’ yearly hazard (Table 4).

952 Overall, the $ToxHQ$ values calculated for the water column and the sediment clearly reflected the
953 anthropogenic gradient between the upstream and downstream sites, with values increasing from
954 Marnay to Bougival and Triel (Table 4), similarly to biomarker responses (see Section 3.3). Moreover,
955 the yearly variations of bioassay responses identified the C1 and C3 water and sediment samples as the
956 most toxic, whereas C4 samples only induced limited effects. These observations are in good

957 agreement with the conclusions drawn from bioavailability and biomarker LOEs (see Sections 3.2 and
958 3.3). Moreover, the independent exposure of model organisms/cell lines to water column or sediment
959 samples gave new insights into environmental contamination along the Seine River. The main part of
960 the toxic effects we noted appeared to be associated with the sediment compartment since the hazard
961 class indices in sediment were systematically higher than in water samples (Table 4).

962 Our results clearly show that the sediment compartment should be integrated into environmental
963 quality assessment procedures as a potential non-negligible source of toxicity for aquatic organisms.
964 Such procedures only based on water column analysis could underestimate the (eco)toxicological risk
965 for ecosystems.

966 The accuracy and environmental significance of toxic responses based on laboratory bioassays (and of
967 the resulting hazard assessment) could be improved by working on the reference and threshold values.
968 Many bioassays, such as *in vitro* tests, presently use ‘too clean to be real’ standards as negative
969 controls (*e.g.*, extraction blanks, ultra-pure water, etc.). The establishment of ‘truly environmental’
970 negative controls should be closely investigated to make these bioassay responses more realistic and
971 relevant in an environmental context. The calculation of reference and threshold values from the
972 analysis of environmentally ‘clean’ water and sediment samples could be a valuable alternative to
973 characterize the basal levels and variations of very specific responses such as hormone-mimetic,
974 genotoxicity, and cytotoxicity endpoints. We applied this kind of approach in some bioassays such as
975 the MELA (and the MELAc), with drinking water and Yville-sur-Seine sediment (a pristine site in the
976 vicinity of the Seine River axis) as negative controls for water and sediment samples, respectively.
977 The integration of other reference matrices from various geographical localizations could also make it
978 easier to understand response variability by taking into account the natural diversity and heterogeneity
979 of the reference environments.

980 Such a strategy would undoubtedly improve the accuracy and the relevance of environment quality
981 assessment using laboratory bioassays in large-scale studies as well as in more geographically-
982 restricted contexts with locally contrasted areas.

983

984 3.5. WOE integration

985 The results of each LOE were integrated into a global WOE index, and a hazard class was attributed to
986 each site and campaign (Fig. 5). Moreover, the contribution of each HQ index to the WOE value was
987 calculated, and is presented in Fig. 6.

988 WOE levels clearly reflected the anthropogenic gradient along the Seine River, with values increasing
989 from upstream to downstream of Paris (Fig. 5). They were systematically lower at Marnay (15% to

990 30%), intermediate at Bougival (38% to 64%) and the highest at Triel (39% to 70%) for all campaigns
991 and for annual average values. The only exception was the spring (C2) sampling period for which the
992 highest WOE index was noted at Bougival (Fig. 5). The resulting hazard classes varied from 'slight' to
993 'moderate' at Marnay, and from 'moderate' to 'major' at the two stations located downstream of Paris,
994 and the WOE index of Triel C1 was very close to the 'severe' hazard level. The year-round hazard was
995 assessed using the annual average LOE data. It summarized the overall hazard class as 'moderate' at
996 Marnay and 'major' at Bougival and Triel (Fig. 5).

997 Additionally, the seasonal variations identified when examining the results of each LOE were also
998 reflected in the WOE levels. The lowest value was consistently recorded for the winter (C4) campaign,
999 at each site. In contrast, the fall (C1) and summer (C3) campaigns had previously been identified as
1000 the most impacted ones, and coherently resulted in the highest WOE levels (Fig. 5). The only
1001 exception was the Bougival C2 sampling point; however, it should be borne in mind that the hazard
1002 assessment during the spring campaign was only based on bioaccumulation and water contamination
1003 data. For this reason, contributions of HQ indices are not discussed below because they are biased by
1004 the missing C2 campaign data.

1005 The contribution profiles from Marnay revealed that chemical contamination in the abiotic
1006 compartment was the main component of the WOE index, with a global contribution ($\text{ChemHQ}_{\text{water}}$
1007 plus $\text{ChemHQ}_{\text{sed}}$ contribution) around 80% for each sampling period and the annual average value
1008 (Fig. 6a). At the upstream site, the greatest risk was attributable to the contaminants analyzed in the
1009 sediment: the contributions of the $\text{ChemHQ}_{\text{sed}}$ indices reached 60% (vs. around 20% for the
1010 $\text{ChemHQ}_{\text{water}}$ indices). Similar observations were made for the downstream stations during the winter
1011 (C4) campaign. The global contribution of chemicals in the abiotic compartment also represented
1012 around 80% for Bougival and Triel during the C4 campaign; however, the contributions of water
1013 column and sediment contamination were almost the same, indicating that the chemical hazard was
1014 governed by contaminants analyzed in the water column as well as in the sediment at both downstream
1015 sites (Fig. 6b and 6c). The cumulative contributions of ChemHQs were also substantial at the
1016 downstream stations during the C1 and C3 campaigns, and the annual average values ranged between
1017 48% and 62% (Fig. 6b and 6c). However, biological effects recorded *in situ* (BiomHQs) and under
1018 laboratory conditions (ToxHQs) also contributed to the calculated WOE value in a non-negligible
1019 way. Biomarker responses thus accounted for 11% to 17% of the WOE indices, and bioassays
1020 contributed between 16% to 36% (Fig. 6b and 6c). Among the bioassays, the ones using sediment as a
1021 test phase usually yielded the highest contributions, suggesting again that the sediment compartment
1022 represents a noticeable hazard in terms of both contamination levels and biological effects.

1023 The increase of the biological effects noted at some particular sites/sampling dates, such as Triel
1024 C1/C3, could be attributed to seasonal variations in the contamination levels, as revealed by LOE#1

1025 results. The concentrations of some contaminants such as PFOS and pesticides are clearly influenced
1026 by seasonal factors such as hydrological conditions (Tamtam et al. 2008; Labadie and Chevreuil 2011)
1027 and/or by the seasonal use of certain chemicals such as pesticides, including metazachlor (Passeport et
1028 al. 2013).

1029 Moreover, the levels of metal elements in gammarids revealed that bioaccumulation was influenced
1030 not only by the contamination levels but also by seasonal variables like temperature, especially
1031 concerning essential elements (Lebrun et al. 2015). A more accurate characterization and
1032 understanding of these variations would make it possible to refine the reference values and/or to define
1033 specific thresholds according to the substance and the influence of confounding factors on its
1034 accumulation levels in organisms. For instance, determining whether the temperature influenced
1035 bioaccumulation levels by modulating metabolic rates or contaminant bioavailability would be of great
1036 interest. These adjustments could improve the conclusions of the WOE procedure in a relevant and
1037 reliable way, by monitoring the impact of external abiotic factors likely to modulate the time-course of
1038 accumulation by exposed organisms, especially in close but contrasted areas.

1039 Overall, the conclusions on the hazard represented by ‘chemical’ and ‘biological’ LOEs (*i.e.*, LOEs#1-
1040 2 vs. LOEs#3-4) are relatively coherent, although a shift in hazard severity was evidenced between the
1041 two types of LOEs. In fact, the hazard level is generally lower with ‘biological’ LOEs than with
1042 ‘chemical’ LOEs (Fig. 5). The table at the bottom of Fig. 5 also reveals a few exceptions (*e.g.*, Triel
1043 C4) for which the conclusion of the chemical and biological LOEs were not fully consistent,
1044 suggesting that the two approaches are complementary. These observations also emphasize the
1045 usefulness of WOE integration, as the class-based hazard ranking of the sites differed when
1046 considering the results of each LOE independently, especially at the downstream sites. For instance,
1047 when referring to the annual hazard class associated to ChemHQ_{sed} values, we failed to discriminate
1048 among the three sites because the environmental hazard was evaluated at the highest level (‘severe’) in
1049 all cases (Table 1), indicating that all sites represented the maximal hazard level in terms of sediment
1050 contamination. In contrast, the annual average hazard classes associated to sediment bioassays
1051 (ToxHQ_{sed}) identified a hazard level increasing from ‘slight’ at Marnay to ‘moderate’ and ‘major’ at
1052 Bougival and Triel, respectively (Table 4). This suggests that only Triel was faced with a high level of
1053 environmental hazard. However, only Bougival was faced with a high (‘major’) hazard level according
1054 to annual average bioavailability measurements (Table 2), while both downstream stations were
1055 classified as particularly impacted by biomarker responses (‘major’ and ‘severe’ hazard at Bougival
1056 and Triel, respectively, based on annual average data; Table 3). The environmental diagnosis of the
1057 three sites would have been substantially different if based on one or another LOE output, under- or
1058 over-estimating the environmental risk at each site depending on the LOE. This would also have led to
1059 contradictory conclusions on the impacts of the Oise River inputs between the two downstream
1060 stations. These two latter aspects would complexify the decision-making process for environmental

1061 managers. The solution lies in final WOE integration, which compiles the results from each LOE into
1062 a global hazard index associated to an integrative hazard class translating the overall environmental
1063 risk at each site.

1064 As a result, it seems relevant to further analyze the environmental hazard in the studied area, basing
1065 the diagnosis on various aspects including contamination levels, bioavailability, and biological
1066 responses. Although the biological effects recorded in the present study can be considered as quite
1067 limited in comparison to contamination levels, remobilization of contaminants (especially those
1068 trapped in sediment) and/or variations of some controlling factors (*e.g.*, temperature, flow rates, or
1069 physico-chemistry of the surrounding environment) would result in a strong increase in bioavailable
1070 contamination levels, and this in turn could induce more severe biological impacts.

1071 To avoid such critical situations for the health of aquatic ecosystems, efforts should focus on
1072 decontamination and remediation procedures in the most impacted sites (Bougival and Triel) before
1073 more adverse effects occur at the population level, as suggested by some biological responses
1074 evidenced in the present study (*e.g.*, acute toxicity, altered microbial communities, reproductive
1075 impairments etc.).

1076

1077 *3.6. Bacterial communities*

1078 The bacterial community composition of the water column was studied using high throughput
1079 sequencing of bacterial 16S-rRNA genes during three sampling campaigns (C1, C3, and C4).
1080 Unfortunately, due to a technical problem, the data from the Triel sampling station were not available
1081 for the winter campaign (C4). The dissimilarity in bacterial community structure (number and relative
1082 abundance of different OTU_{0.03}) among the different samples are presented as a three-dimension
1083 NMDS graph (Fig. 7). This figure clearly shows that samples from Marnay were grouped together,
1084 while obviously separated from the Bougival and Triel samples that were close (except for the
1085 December 2012 Bougival sample). Similar differences were previously observed on ARISA
1086 (Automated Ribosomal Intergenic Spacer Analysis) profiles of river biofilms collected from the same
1087 sampling sites (Fechner et al. 2012). This suggests that the bacterial communities of the water samples
1088 collected upstream of Paris were different from those sampled downstream. These observations on
1089 bacterial communities are also in good agreement with the global hazard assessment illustrated by
1090 WOE indices, which were the lowest and varied between 15% to 30% for Marnay. In contrast, the
1091 integrative HQ calculated for the two downstream stations covered the same range of higher values
1092 (54% to 70%), except for the C4 campaign. For this campaign, WOE indices from Bougival were
1093 close to the range of values from Marnay; in parallel, bacterial community compositions were similar
1094 in the two stations during the winter campaign. Globally, this study shows that responses at the

1095 bacterial community level reflect the global disturbance of the environment particularly well. Results
1096 from river biofilms collected during the same campaigns also corroborate this observation (Faburé et
1097 al. 2015). The use of microbial communities in an ERA context might be very powerful in the future;
1098 however the conclusions from such an approach (*i.e.*, the classification of sites in relation to one
1099 another) can only be relative since they are drawn from inter-site qualitative comparisons, as reference
1100 levels are nowadays still lacking.

1101

1102 **4. Conclusion**

1103 The WOE approach applied in the present study proved efficient and relevant in terms of both global
1104 environmental hazard diagnosis and seasonality analysis. The procedure was particularly improved
1105 using external reference levels integrating natural variations of responses and confounding factors,
1106 especially in LOEs#2 and #3. This improved the reliability of WOE integration results, which better
1107 reflected the level of disturbance of organisms at each sampling time, without any interference related
1108 to acclimation or adaptation mechanisms likely to occur in chronically exposed populations. The
1109 establishment of reference values and thresholds from numerous studies conducted at the national
1110 scale also eliminated the need for a reference site in the study area, which could be very problematic in
1111 large rivers subjected to multiple and diffuse pressures. Our results reveal that at the upstream site,
1112 generally used as a relative reference or control site in previous investigations in the area, the low
1113 contamination levels nonetheless resulted in low but significant biological effects.

1114 This approach should be pursued and further developed at larger spatial scales. Bioaccumulation and
1115 biological responses to pollutants as well as baseline levels may be modulated and altered by long-
1116 term variations and trends in some key endpoints, *e.g.*, growth and reproduction, themselves governed
1117 by global factors and large-scale processes (*e.g.*, climate trends and changes, oceanographic cycles,
1118 etc.) (Garmendia et al. 2015). The in-depth characterization of the baseline levels and relevant effect
1119 thresholds for such environmentally relevant endpoints is thus a challenge to ensure their relevance
1120 within ERA purposes.

1121 Another strength of the present work lies in the use of gammarids from the same population for
1122 bioaccumulation measurements and biomarker analyses. A direct and strong connection was thus
1123 established between bioaccumulation levels and biological responses, strengthening the conclusions
1124 from the LOEs based on these data. Moreover, the use of these amphipods is entirely appropriate and
1125 relevant in the context of ecological/ecotoxicological field studies. Gammarids are widespread in
1126 European freshwaters, and are key actors in the functioning of these ecosystems as litter degraders and
1127 as a food source for fish and amphibian species. As a result, multiple biomarkers and bioassays using
1128 gammarids are available for field-testing of contaminant impacts. Moreover, modelling developments

1129 quantifying the natural variability of these markers in relation to abiotic factors enhance the reliability
1130 of the *in situ* methodology and allow for its implementation at large spatial and temporal scales in
1131 monitoring programs (Coulaud et al. 2011; Chaumot et al. 2015).

1132 The assessment of environmental quality was also improved by integrating water column and
1133 sediment analyses. In the water column, contamination (according to the selected compounds analyzed
1134 in this matrix) and toxicity (assessed by bioassays) remained relatively limited, but in the sediment
1135 variable stocks of pollutants accumulated and locally reached very high levels, with various impacts
1136 on laboratory-exposed organisms. These compounds are likely (at least partially) involved in some of
1137 the biological responses detected in gammarids through biomarker analyses, since they were
1138 significantly accumulated by exposed gammarids. This also proves that although we did not trace all
1139 contaminants in the water, a fraction of these contaminants is bioavailable to organisms through the
1140 water column. Overall observations suggest that a non-negligible ecological risk in the area could
1141 threaten benthic biota as well as pelagic organisms through the release and/or remobilization of the
1142 sediment-bound chemicals into the water column.

1143 Other investigations were performed within the framework of the PIREN-Seine program (*e.g.*,
1144 bacterial community analyses, bioaccumulation of metal and organic compounds, metal tolerance
1145 acquisition in biofilms), but were not integrated into the WOE model because reference levels and
1146 thresholds still remain difficult to set. However, it should be mentioned that these approaches gave
1147 similar results, clearly differentiating the upstream (Marnay) station from the downstream (Bougival
1148 and Triel) sites, in relation to the contamination gradient. There is no doubt that these experiments
1149 have to be further developed, but even in their current state they can be valuable tools with successful
1150 *in situ* deployment in a biomonitoring context, and it would be relevant to integrate them into an ERA
1151 procedure as they provide information at the community level.

1152 Finally, the WOE approach applied in the present study was based on the integration of each response
1153 into a global hazard index in a similar way, including pseudo normalization of the data. Thus the
1154 results from various sampling times remained comparable and were reported on a common grid of
1155 hazard classification. Such a procedure represents an advantageous and practical tool in the diagnosis
1156 of environmental hazard because it yields relevant information classifying the most problematic
1157 substances and effects, and gradually identifies the most impacted sites (comparing HQ values) with
1158 an associated hazard level. The most remarkable strong point of the approach lies in the ability of the
1159 model to integrate a large amount of endpoints characterizing various aspects of the environmental
1160 risk and to generate very 'simple' and 'comprehensible' integrative outputs from this large dataset,
1161 *i.e.*, the WOE index and the relative hazard class.

1162 This WOE model may be very helpful for environment managers in decision-making processes to plan
1163 remediation procedures and/or actions to reduce emissions and/or uses of problematic substances.

1164 Combined with the available biomonitoring tools used in the present study, this approach could also be
1165 implemented on a long-term basis to monitor the potential improvement of environmental quality
1166 following environmental management measures.

1167

1168 **References**

1169 Ayrault S, Roy-Barman M, Le Cloarec MF, Priadi CR, Bonté P, Göpel C (2012) Lead contamination
1170 of the Seine River, France: Geochemical implications of a historical perspective. *Chemosphere*
1171 87(8):902-910. doi: 10.1016/j.chemosphere.2012.01.043

1172 AZUR Environmental (1998) Microtox basic solid-phase test (Basic SPT). Carlsbad, CA, USA

1173 Bado-Nilles A, Jolly S, Porcher JM, Palluel O, Geffard A, Gagnaire B, Betoulle S, Sanchez W (2015)
1174 Applications in environmental risk assessment of leucocyte apoptosis, necrosis and respiratory
1175 burst analysis on the European bullhead, *Cottus* sp. *Environ Pollut* 184:9-17. doi:
1176 10.1016/j.envpol.2013.07.049

1177 Balaguer P, François F, Comunale F, Fenet H, Boussioux A-M, Pons M, Nicolas J-C, Casellas C
1178 (1999) Reporter cell lines to study the estrogenic effects of xenoestrogens. *Sci Total Environ*
1179 233(1-3):47-56. doi: 10.1016/s0048-9697(99)00178-3

1180 Barhoumi B, Clérandeau C, Landi L, Pichon A, Le Bihanic F, Poirier D, Anschutz P, Budzinski H,
1181 Driss MR, Cachot J (2016) Assessing the toxicity of sediments using the Medaka Embryo-
1182 Larval Assay and two other bioassays. *Environ Toxicol Chem (In press)*. doi: 10.1002/etc.3388

1183 Barjhoux I, Baudrimont M, Morin B, Landi L, Gonzalez P, Cachot J (2012) Effects of copper and
1184 cadmium spiked-sediments on embryonic development of Japanese medaka (*Oryzias latipes*).
1185 *Ecotoxicol Environ Saf* 79:272-282. doi: 10.1016/j.ecoenv.2012.01.011

1186 Bebianno MJ, Pereira CG, Rey F, Cravo A, Duarte D, D'Errico G, Regoli F (2015) Integrated
1187 approach to assess ecosystem health in harbor areas. *Sci Total Environ* 514:92-107. doi:
1188 10.1016/j.scitotenv.2015.01.050

1189 Belles A, Pardon P, Budzinski H (2014) Development of an adapted version of polar organic chemical
1190 integrative samplers (POCIS-Nylon). *Anal Bioanal Chem* 406(4):1099-1110. doi:
1191 10.1007/s00216-013-7286-2

1192 Benedetti M, Ciaprini F, Piva F, Onorati F, Fattorini D, Notti A, Ausili A, Regoli F (2012) A
1193 multidisciplinary weight of evidence approach for classifying polluted sediments: Integrating

- 1194 sediment chemistry, bioavailability, biomarkers responses and bioassays. Environment
1195 International 38(1):17-28. doi: 10.1016/j.envint.2011.08.003
- 1196 Bennett J, Cabbage J (1992) Review and Evaluation of Microtox Test for Freshwater Sediments.
1197 Washington State Department of Ecology, Washington, 28 pp
- 1198 Besse JP, Coquery M, Lopes C, Chaumot A, Budzinski H, Labadie P, Geffard O (2013) Caged
1199 *Gammarus fossarum* (Crustacea) as a robust tool for the characterization of bioavailable
1200 contamination levels in continental waters: Towards the determination of threshold values.
1201 Water Res 47(2):650-660. doi: 10.1016/j.watres.2012.10.024
- 1202 Billen G, Garnier J, Mouchel J-M, Silvestre M (2007) The Seine system: Introduction to a
1203 multidisciplinary approach of the functioning of a regional river system. Sci Total Environ
1204 375(1-3):1-12. doi: 10.1016/j.scitotenv.2006.12.001
- 1205 Blanchard M, Teil MJ, Guigon E, Larcher-Tiphagne K, Ollivon D, Garban B, Chevreuil M (2007)
1206 Persistent toxic substance inputs to the river Seine basin (France) via atmospheric deposition
1207 and urban sludge application. Sci Total Environ 375(1-3):232-243. doi:
1208 10.1016/j.scitotenv.2006.12.012
- 1209 Bols NC, Brubacher JL, Ganassin RC, Lee LEJ (2001) Ecotoxicology and innate immunity in fish.
1210 Dev Comp Immunol 25(8-9):853-873. doi: 10.1016/s0145-305x(01)00040-4
- 1211 Brouwer H, Murphy T, McArdle L (1990) A sediment-contact bioassay with *Photobacterium*
1212 *phosphoreum*. Environ Toxicol Chem 9(11):1353-1358
- 1213 Budzinski H, Letellier M, Thompson S, LeMenach K, Garrigues P (2000) Combined protocol for the
1214 analysis of polycyclic aromatic hydrocarbons (PAHs) and polychlorobiphenyls (PCBs) from
1215 sediments using focussed microwave assisted (FMW) extraction at atmospheric pressure. Fresen
1216 J Anal Chem 367(2):165-171
- 1217 Chapman PM, McDonald BG, Lawrence GS (2002) Weight-of-evidence issues and frameworks for
1218 sediment quality (and other) assessments. Hum Ecol Risk Assess 8(7):1489-1515. doi:
1219 10.1080/20028091057457
- 1220 Chapman PM, Wang F, Janssen CR, Goulet RR, Kamunde CN (2003) Conducting ecological risk
1221 assessments of inorganic metals and metalloids: Current status. Hum Ecol Risk Assess 9(4):641-
1222 697

- 1223 Chapman PM (2007) Determining when contamination is pollution - Weight of evidence
1224 determinations for sediments and effluents. *Environment International* 33(4):492-501. doi:
1225 10.1016/j.envint.2006.09.001
- 1226 Charron L, Geffard O, Chaumot A, Coulaud R, Quéau H, Geffard A, Dedourge-Geffard O (2013)
1227 Effect of water quality and confounding factors on digestive enzyme activities in *Gammarus*
1228 *fossarum*. *Environ Sci Pollut Res* 20(12):9044 - 9056. doi: 10.1007/s11356-013-1921-5
- 1229 Chaumot A, Geffard O, Armengaud J, Maltby L (2015) Chapter 11 - Gammarids as Reference Species
1230 for Freshwater Monitoring. In: Mouneyrac C, Amiard-Triquet J-C, Amiard C (Eds) *Aquatic*
1231 *Ecotoxicology: Advancing Tools for Dealing with Emerging Risks*. Academic Press, pp. 253-
1232 280
- 1233 Coulaud R, Geffard O, Xuereb B, Lacaze E, Quéau H, Garric J, Charles S, Chaumot A (2011) *In situ*
1234 feeding assay with *Gammarus fossarum* (Crustacea): Modelling the influence of confounding
1235 factors to improve water quality biomonitoring. *Water Res* 45(19):6417-6429. doi:
1236 10.1016/j.watres.2011.09.035
- 1237 Dagnino A, Sforzini S, Dondero F, Fenoglio S, Bona E, Jensen J, Viarengo A (2008) A "Weight-of-
1238 Evidence" approach for the integration of environmental "Triad" data to assess ecological risk
1239 and biological vulnerability. *Integr Environ Assess Manag* 4(3):314-326. doi:
1240 10.1897/ieam_2007-067.1
- 1241 Dedourge-Geffard O, Charron L, Hofbauer C, Gaillet V, Palais F, Lacaze E, Geffard A, Geffard O
1242 (2013) Temporal patterns of digestive enzyme activities and feeding rate in gammarids
1243 (*Gammarus fossarum*) exposed to inland polluted waters. *Ecotoxicol Environ Saf* 97:139-146.
1244 doi: 10.1016/j.ecoenv.2013.07.016
- 1245 Dufour V, Cruz J, Chambolle M, Granger D, Le Menach K, Pardon P, Litrico X, Budzinski H (2015)
1246 Study of water contamination by pesticides, comparison of grab sampling and passive sampling
1247 – Application to natural water and an industrialzone sewage near Bordeaux. *Revue des Sciences*
1248 *de l'Eau* 28(3):223-228. doi: 10.7202/1034011ar
- 1249 Dulio V, Andres S (2014) Recommandations du CEP auprès du MEDDE pour la sélection des
1250 Substances Pertinentes à Surveiller dans les milieux aquatiques pour le second cycle de la DCE
1251 (2016-2021). Rapport AQUAREF 2013, Final document, pp. 102.
1252 [http://www.onema.fr/IMG/pdf/Aquaref2013_F1a_INERIS_Rapport%20CEP_subst%20pertinen](http://www.onema.fr/IMG/pdf/Aquaref2013_F1a_INERIS_Rapport%20CEP_subst%20pertinentes_VF.pdf)
1253 [tes_VF.pdf](http://www.onema.fr/IMG/pdf/Aquaref2013_F1a_INERIS_Rapport%20CEP_subst%20pertinentes_VF.pdf) (In French)

- 1254 Ellman GL, Courtney KD, Andres Jr V, Featherstone RM (1961) A new and rapid colorimetric
1255 determination of acetylcholinesterase activity. *Biochem Pharmacol* 7(2):88-90
- 1256 Escartín E, Porte C (1996) Acetylcholinesterase inhibition in the crayfish *Procambarus clarkii*
1257 exposed to fenitrothion. *Ecotoxicol Environ Saf* 34(2):160-164. doi: 10.1006/eesa.1996.0058
- 1258 European Commission (EC) (2000) Directive 2000/60/EC of the European parliament and of the
1259 Council of 23 October 2000 establishing a framework for Community action in the field of
1260 water policy. *Official Journal of the European Union*, pp. 72
- 1261 European Commission (EC) (2013) Directive 2013/39/EU of the European parliament and of the
1262 Council of 12 August 2013 amending Directives 2000/60/EC and 2008/105/EC as regards
1263 priority substances in the field of water policy. *Official Journal of the European Union*, pp. 17
- 1264 European Union (EU) (2005) Common implementation strategy for the water framework directive:
1265 environmental quality standards (EQS), substance data sheet, priority substance No.28:
1266 Polycyclic Aromatic Compounds, Brussels, pp. 30
- 1267 Faburé J, Dufour M, Autret A, Uher E, Fechner LC (2015) Impact of an urban multi-metal
1268 contamination gradient: Metal bioaccumulation and tolerance of river biofilms collected in
1269 different seasons. *Aquat Toxicol* 159:276-289. doi: 10.1016/j.aquatox.2014.12.014
- 1270 Fechner LC, Gourlay-Francé C, Bourgeault A, Tusseau-Vuillemin MH (2012) Diffuse urban pollution
1271 increases metal tolerance of natural heterotrophic biofilms. *Environ Pollut* 162:311-318. doi:
1272 10.1016/j.envpol.2011.11.033
- 1273 French Ministry of Ecology Energy Sustainable Development and Planning (MEDAD) (2006) Arrêté
1274 du 9 août 2006 relatif aux niveaux à prendre en compte lors d'une analyse de rejets dans les eaux
1275 de surface ou de sédiments marins, estuariens ou extraits de cours d'eau ou canaux relevant
1276 respectivement des rubriques 2.2.3.0, 4.1.3.0 et 3.2.1.0 de la nomenclature annexée à l'article R.
1277 214-1 du code de l'environnement. *Journal Officiel de la République Française - Texte* 15 sur
1278 48 - pp. 3 (In French)
- 1279 French Ministry of Ecology Energy Sustainable Development and Planning (MEDAD) (2007)
1280 Circulaire n°2007/23 du 7 mai 2007 définissant les « normes de qualité environnementale
1281 provisoires (NQE_p) » des 41 substances impliquées dans l'évaluation de l'état chimique des
1282 masses d'eau ainsi que des substances pertinentes du programme national de réduction des
1283 substances dangereuses dans l'eau. *Bulletin officiel du MEDAD* 2007/15 - Texte 17/36 - pp. 13
1284 (In French)

- 1285 French Ministry of Ecology Energy Sustainable Development and Planning (MEDAD) (2015) Arrêté
1286 du 27 juillet 2015 modifiant l'arrêté du 25 janvier 2010 relatif aux méthodes et critères
1287 d'évaluation de l'état écologique, de l'état chimique et du potentiel écologique des eaux de
1288 surface pris en application des articles R. 212-10, R. 212-11 et R. 212-18 du code de
1289 l'environnement. Journal Officiel de la République Française - Texte 4 sur 125 - pp. 95 (In
1290 French)
- 1291 García-Armisen T, Inceoğlu Ö, Ouattara NK, Anzil A, Verbanck MA, Brion N, Servais P (2014)
1292 Seasonal variations and resilience of bacterial communities in a sewage polluted urban river.
1293 PLoS ONE 9(3):art. no. e92579. doi: 10.1371/journal.pone.0092579
- 1294 Garmendia L, Izagirre U, Soto M, Lermen D, Koschorreck J (2015) Combining chemical and
1295 biological endpoints, a major challenge for twenty-first century's environmental specimen
1296 banks. Environ Sci Pollut Res 22(3):1631-1634. doi: 10.1007/s11356-014-2925-5
- 1297 Geffard O, Xuereb B, Chaumot A, Geffard A, Biagianti S, Noël C, Abbaci K, Garric J, Charmantier
1298 G, Charmantier-Daures M (2010) Ovarian cycle and embryonic development in *Gammarus*
1299 *fossarum*: Application for reproductive toxicity assessment. Environ Toxicol Chem
1300 29(10):2249-2259. doi: 10.1002/etc.268
- 1301 Helmstetter MF, Alden RWI (1995) Toxic responses of Japanese medaka (*Oryzias latipes*) eggs
1302 following topical and immersion exposures to pentachlorophenol. Aquat Toxicol 32(1):15-29.
1303 doi: 10.1016/0166-445x(94)00059-y
- 1304 ISO (1999) Determination of the inhibitory effect of water samples on the light emission of *Vibrio*
1305 *fischeri* (luminescent bacteria test). Part 3: method using freeze dried bacteria. European
1306 Standard ISO 11348-3:1998. European Committee for Standardization, Bruxelles
- 1307 Jugan ML, Lévy-Bimbot M, Pomérance M, Tamisier-Karolak S, Blondeau JP, Lévi Y (2007) A new
1308 bioluminescent cellular assay to measure the transcriptional effects of chemicals that modulate
1309 the alpha-1 thyroid hormone receptor. Toxicol In Vitro 21(6):1197-1205. doi:
1310 10.1016/j.tiv.2007.03.020
- 1311 Jugan ML, Oziol L, Bimbot M, Huteau V, Tamisier-Karolak S, Blondeau JP, Lévi Y (2009) *In vitro*
1312 assessment of thyroid and estrogenic endocrine disruptors in wastewater treatment plants, rivers
1313 and drinking water supplies in the greater Paris area (France). Sci Total Environ 407(11):3579-
1314 3587. doi: 10.1016/j.scitotenv.2009.01.027
- 1315 Kinani S, Bouchonnet S, Creusot N, Bourcier S, Balaguer P, Porcher JM, Aït-Aïssa S (2010)
1316 Bioanalytical characterisation of multiple endocrine- and dioxin-like activities in sediments

- 1317 from reference and impacted small rivers. *Environ Pollut* 158(1):74-83. doi:
1318 10.1016/j.envpol.2009.07.041
- 1319 Labadie P, Chevreuril M (2011) Biogeochemical dynamics of perfluorinated alkyl acids and sulfonates
1320 in the River Seine (Paris, France) under contrasting hydrological conditions. *Environ Pollut*
1321 159(12):3634-3639. doi: 10.1016/j.envpol.2011.07.028
- 1322 Le Cloarec MF, Bonte PH, Lestel L, Lefèvre I, Ayrault S (2011) Sedimentary record of metal
1323 contamination in the Seine River during the last century. *Phys Chem Earth* 36(12):515-529. doi:
1324 10.1016/j.pce.2009.02.003
- 1325 Le Pape P, Ayrault S, Quantin C (2012) Trace element behavior and partition versus urbanization
1326 gradient in an urban river (Orge River, France). *J Hydrol* 472-473:99-110. doi:
1327 10.1016/j.jhydrol.2012.09.042
- 1328 Lebrun JD, Uher E, Tusseau-Vuillemin MH, Gourlay-Francé C (2014) Essential metal contents in
1329 indigenous gammarids related to exposure levels at the river basin scale: Metal-dependent
1330 models of bioaccumulation and geochemical correlations. *Sci Total Environ* 466-467:100-108.
1331 doi: 10.1016/j.scitotenv.2013.07.003
- 1332 Lebrun JD, Geffard O, Urien N, François A, Uher E, Fechner LC (2015) Seasonal variability and
1333 inter-species comparison of metal bioaccumulation in caged gammarids under urban diffuse
1334 contamination gradient: Implications for biomonitoring investigations. *Sci Total Environ*
1335 511:501-508. doi: 10.1016/j.scitotenv.2014.12.078
- 1336 MacDonald DD, Ingersoll CG, Berger TA (2000) Development and evaluation of consensus-based
1337 sediment quality guidelines for freshwater ecosystems. *Arch Environ Contam Toxicol* 39(1):20-
1338 31. doi: 10.1007/s002440010075
- 1339 Mersch-Sundermann V, Mochayedi S, Kevekordes S (1992) Genotoxicity of polycyclic aromatic
1340 hydrocarbons in *Escherichia coli* PQ37. *Muta Res-Genet Tox* 278(1):1-9. doi: 10.1016/0165-
1341 1218(92)90279-9
- 1342 Michel C, Bourgeault A, Gourlay-Francé C, Palais F, Geffard A, Vincent-Hubert F (2013) Seasonal
1343 and PAH impact on DNA strand-break levels in gills of transplanted zebra mussels. *Ecotoxicol*
1344 *Environ Saf* 92:18-26. doi: 10.1016/j.ecoenv.2013.01.018
- 1345 Munoz G, Giraudel JL, Botta F, Lestremau F, Dévier MH, Budzinski H, Labadie P (2015) Spatial
1346 distribution and partitioning behavior of selected poly- and perfluoroalkyl substances in

- 1347 freshwater ecosystems: A French nationwide survey. *Sci Total Environ* 517:48-56. doi:
1348 10.1016/j.scitotenv.2015.02.043
- 1349 Nouira T, Risso C, Chouba L, Budzinski H, Boussetta H (2013) Polychlorinated biphenyls (PCBs) and
1350 Polybrominated Diphenyl Ethers (PBDEs) in surface sediments from Monastir Bay (Tunisia,
1351 Central Mediterranean): Occurrence, distribution and seasonal variations. *Chemosphere*
1352 93(3):487-493. doi: 10.1016/j.chemosphere.2013.06.017
- 1353 Owen R, Buxton L, Sarkis S, Toaspern M, Knap A, Depledge M (2002) An evaluation of hemolymph
1354 cholinesterase activities in the tropical scallop, *Euvola (Pecten) ziczac*, for the rapid assessment
1355 of pesticide exposure. *Mar Pollut Bull* 44(10):1010-1017. doi: 10.1016/s0025-326x(02)00139-x
- 1356 Passeport E, Tournebize J, Chaumont C, Guenne A, Coquet Y (2013) Pesticide contamination
1357 interception strategy and removal efficiency in forest buffer and artificial wetland in a tile-
1358 drained agricultural watershed. *Chemosphere* 91(9):1289-1296. doi:
1359 10.1016/j.chemosphere.2013.02.053
- 1360 Piva F, Ciaprini F, Onorati F, Benedetti M, Fattorini D, Ausili A, Regoli F (2011) Assessing sediment
1361 hazard through a weight of evidence approach with bioindicator organisms: A practical model to
1362 elaborate data from sediment chemistry, bioavailability, biomarkers and ecotoxicological
1363 bioassays. *Chemosphere* 83(4):475-485. doi: 10.1016/j.chemosphere.2010.12.064
- 1364 Priadi C, Bourgeault A, Ayrault S, Gourlay-France C, Tusseau-Vuillemin M-H, Bonte P, Mouchel J-
1365 M (2011) Spatio-temporal variability of solid, total dissolved and labile metal: passive vs.
1366 discrete sampling evaluation in river metal monitoring. *J Environ Monit* 13(5):1470-1479. doi:
1367 10.1039/c0em00713g
- 1368 Quillardet P, Hofnung M (1985) The SOS Chromotest, a colorimetric bacterial assay for genotoxins:
1369 procedures. *Mutation Research - Environmental Mutagenesis and Related Subjects* 147(3):65-
1370 78. doi: 10.1016/0165-1161(85)90020-2
- 1371 Tamtam F, Mercier F, Le Bot B, Eurin J, Tuc Dinh Q, Clément M, Chevreuil M (2008) Occurrence
1372 and fate of antibiotics in the Seine River in various hydrological conditions. *Sci Total Environ*
1373 393(1):84-95. doi: 10.1016/j.scitotenv.2007.12.009
- 1374 Teil M-J, Tlili K, Blanchard M, Labadie P, Alliot F, Chevreuil M (2014) Polychlorinated biphenyls,
1375 polybrominated diphenyl ethers, and phthalates in roach from the Seine River basin (France):
1376 Impact of densely urbanized areas. *Arch Environ Contam Toxicol* 66(1):41-57. doi:
1377 10.1007/s00244-013-9955-8

- 1378 Thévenot DR, Moilleron R, Lestel L, Gromaire M-C, Rocher V, Cambier P, Bonté P, Colin J-L, de
1379 Pontevès C, Meybeck M (2007) Critical budget of metal sources and pathways in the Seine
1380 River basin (1994-2003) for Cd, Cr, Cu, Hg, Ni, Pb and Zn. *Sci Total Environ* 375(1-3):180-
1381 203. doi: 10.1016/j.scitotenv.2006.12.008
- 1382 Vicquelin L, Leray-Forget J, Peluhet L, LeMenach K, Deflandre B, Anschutz P, Etcheber H, Morin B,
1383 Budzinski H, Cachot J (2011) A new spiked sediment assay using embryos of the Japanese
1384 medaka specifically designed for a reliable toxicity assessment of hydrophobic chemicals.
1385 *Aquat Toxicol* 105(3-4):235-245. doi: 10.1016/j.aquatox.2011.06.011
- 1386 Wilson VS, Bobseine K, Lambright CR, Gray Jr LE (2002) A novel cell line, MDA-kb2, that stably
1387 expresses an androgen- and glucocorticoid-responsive reporter for the detection of hormone
1388 receptor agonists and antagonists. *Toxicol Sci* 66(1):69-81. doi: 10.1093/toxsci/66.1.69
- 1389 Xuereb B, Chaumot A, Mons R, Garric J, Geffard O (2009) Acetylcholinesterase activity in
1390 *Gammarus fossarum* (Crustacea Amphipoda): Intrinsic variability, reference levels, and a
1391 reliable tool for field surveys. *Aquat Toxicol* 93(4):225-233. doi:
1392 10.1016/j.aquatox.2009.05.006
- 1393
- 1394
- 1395

1396 **Figure captions**

1397 **Fig. 1** Map of the three sampling stations (Marnay, Bougival, and Triel) along the Seine River (in the
1398 north of France) adapted from Faburé et al. (2015). Arrows indicate the water flow direction. The city
1399 of Paris is represented in dark grey, and the densely urbanized area surrounding the French capital is
1400 colored in light grey. Source (France map): www.histgeo.ac-aix-marseille.fr, © Daniel Dalet

1401

1402 **Fig. 2** Contribution of each class of chemicals to ChemHQ_{water} values (a), and ChemHQ_{sed} values
1403 using reference values for individual substances (b), or for each class of compounds (c). See Table 1
1404 footnote for details. PESTs, pesticides; AKPs, alkylphenols, PFAS, perfluoroalkyl substance (PFOS);
1405 MEs, metal elements; PAHs, polycyclic aromatic hydrocarbons, PCBs, polychlorinated biphenyls;
1406 PBDEs, polybromodiphenyl ethers; OCPs, organochlorine pesticides; CX, Xth campaign; AA, annual
1407 average value (mean of concentrations from the C1 to C4 campaigns)

1408

1409 **Fig. 3** Contribution of each class of chemicals to BioavHQ values (a), and of each class of biomarkers
1410 to BiomHQ values (b). Only the contributions of chemicals to the annual average (AA) BioavHQ
1411 values calculated for each sampling site are shown. PAHs, polycyclic aromatic hydrocarbons, PCBs,
1412 polychlorinated biphenyls; PBDEs, polybromodiphenyl ethers; OCPs, organochlorine pesticides; MEs,
1413 metal elements. Within LOE#3 (BiomHQ values), neurotoxicity was investigated by assessing AChE
1414 activity. Energy acquisition markers included cellulase, trypsin, and amylase enzymatic activities.
1415 Survival was assessed using mortality rates. Feeding rate measurements were used to track feeding
1416 behavior. Reproduction was studied through molt delay and the number of embryos and oocytes *per*
1417 female. Please note that reproduction markers and cellulase activity were not investigated in the C4
1418 campaign. CX, Xth campaign; AA, annual average value (mean of effects calculated from the C1 to C4
1419 campaigns); N/A, not applicable (as BiomHQ was equal to 0 for Marnay C4)

1420

1421 **Fig. 4** Contribution of each class of bioassay endpoints to ToxHQ_{water} (a) and ToxHQ_{sed} (b) values.
1422 Survival endpoints included embryonic and larval viability (MELA). Development was characterized
1423 by recording hatching success and time to hatch (MELA). The percentage of abnormal larvae was
1424 selected to illustrate teratogenicity (MELA). Biometric measurements of larvae at hatching, including
1425 total body length and head size, were used to evaluate *in ovo* growth (MELA). ER, TR, AR/GR, and
1426 Anti-AR induction factors were gathered to study endocrine-disrupting potency (ED *in vitro*
1427 bioassays). Cytotoxicity and genotoxicity potencies of sediment elutriates were evaluated using *in*

1428 *vitro* bioassays, using Microtox and SOS Chromotest procedures, respectively. CX, Xth campaign;
1429 AA, annual average value (mean of effects E(i) calculated from the C1 to C4 campaigns)

1430

1431 **Fig. 5** Weight of Evidence indices (WOEs) and associated hazard classes integrating the results of
1432 each LOE calculated for the three stations during the four sampling campaigns (C1 to C4), and annual
1433 average (AA) values. The hazard class attributed to each LOE hazard quotient (HQ) is summarized in
1434 the table below. Please note that in the C2 campaign, only ChemHQ_{water} and BioavHQ were evaluated

1435

1436 **Fig. 6** Contribution of each LOE hazard quotient (HQ) to the global WOE values calculated for the
1437 Marnay (a), Bougival (b), and Triel (c) stations. CX, Xth campaign; AA, annual average value. Please
1438 note that as only ChemHQ_{water} and BioavHQ were evaluated in the C2 campaign, details on the
1439 contributions for that campaign are not presented

1440

1441 **Fig. 7** 3D-Non-metric Multi-Dimensional Scaling graphical representation of the dissimilarity
1442 between bacterial communities of the different samples. Mar, Marnay; Bou, Bougival; Tri, Triel;
1443 Sept11, C1 campaign; Jul12, C3 campaign; Dec12, C4 campaign

1444

1445 **Supplementary Material captions (ESM_1.pdf)**

1446

1447 **Supplementary Tables**

1448

1449 **Table S1** List of chemical parameters included into the WOE approach for the water column
1450 (LOE#1), and associated weightings, CAS numbers and selected reference values

1451

1452 **Table S2** List of chemical parameters included into the WOE approach for the sediment (LOE#1), and
1453 associated weightings, CAS numbers and selected reference values

1454

1455 **Table S3** List of chemical parameters included into the WOE approach for bioavailability (LOE#2),
1456 selected reference values and hazard classes attributed to $RTR_w(i)$ values

1457

1458 **Table S4** Reference values (Ref.) and thresholds (Th(%)) assigned to biomarker responses (LOE#3)
1459 analyzed in the study

1460

1461 **Table S5** Details of the hazard classes attributed to the effect $E(i)$ values for biomarker responses
1462 (LOE#3)

1463

1464 **Table S6** Assigned weightings for the biomarker line of evidence (LOE#3) according to Piva et al.
1465 (2011)

1466

1467 **Table S7** Reference values (Ref.) and thresholds (Th(%)) assigned to each endpoint for *in vivo*
1468 (MELA) and *in vitro* (ED, Microtox, SOS Chromotest) bioassays analyzed in the study (LOE#4)

1469

1470 **Table S8** Details of the weighting calculations according to endpoint, matrix and exposure time,
1471 adapted from Piva et al. (2011)

1472

1473 **Table S9** Concentrations of organic micropollutants and metal elements measured in the water column
1474 (<0.7 µm) at the three stations during each sampling campaign, and integrated into the Weight-of-
1475 Evidence approach (LOE#1)

1476

1477 **Table S10** Concentrations of organic micropollutants and metal elements measured in sediment
1478 (<2 mm) at the three stations during each sampling campaign, and integrated into the Weight-of-
1479 Evidence approach (LOE#1)

1480

1481 **Table S11** Main physico-chemical characteristics and organic contamination levels in sediments
1482 sampled along the Seine River

1483

1484 **Table S12** Bioaccumulation levels of organic micropollutants and metal elements measured in
1485 gammarids at the three stations during each sampling campaign, and integrated into the Weight-of-
1486 Evidence approach (LOE#2)

1487

1488 **Table S13** Biomarker responses in gammarids at the three stations during each sampling campaign,
1489 integrated into the Weight-of-Evidence approach (LOE#3)

1490

1491 **Table S14** Bioassay endpoint values observed at the three stations during each sampling campaign,
1492 and integrated into the Weight-of-Evidence approach (LOE#4)

1493

1494 **Supplementary Figures**

1495

1496 **Fig. S1** Details on the calculation procedure implemented within LOE#1 (sediment and water column
1497 chemistry), adapted from Piva et al. (2011)

1498

1499 **Fig. S2** Details on the calculation procedure implemented within LOE#2 (bioavailability), adapted
1500 from Piva et al. (2011)

1501

1502 **Fig. S3** Details on the calculation procedure implemented within LOE#3 (biomarkers), adapted from
1503 Piva et al. (2011)

1504

1505 **Fig. S4** Details on the calculation procedure implemented within LOE#4 (bioassays), adapted from
1506 Piva et al. (2011)

1507

1508 **Fig. S5** Details on the calculation procedure implemented within WOE integration, adapted from Piva
1509 et al. (2011)

Figures

Fig. 1

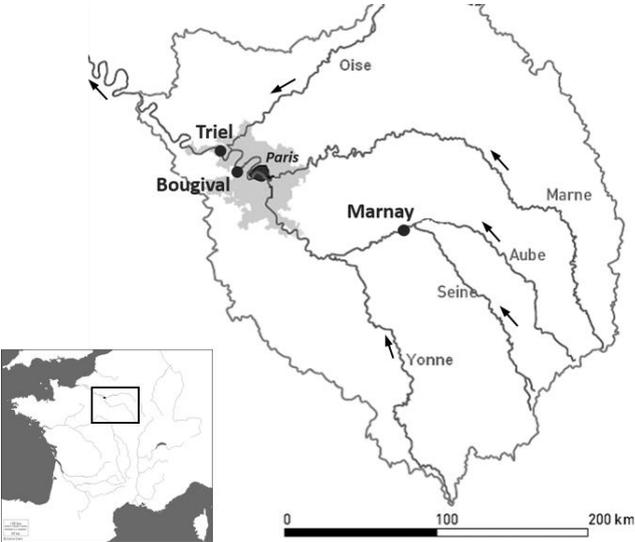
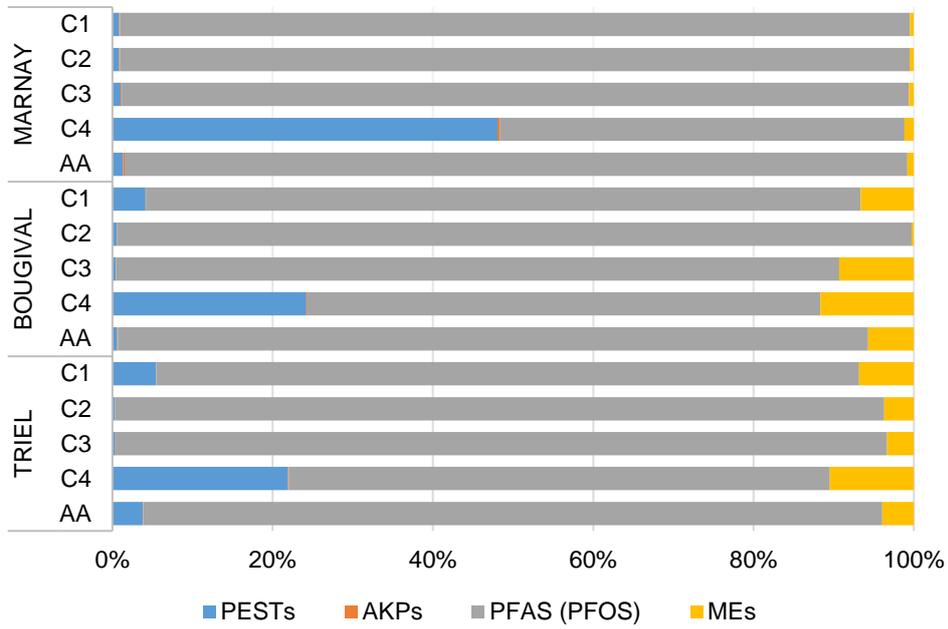
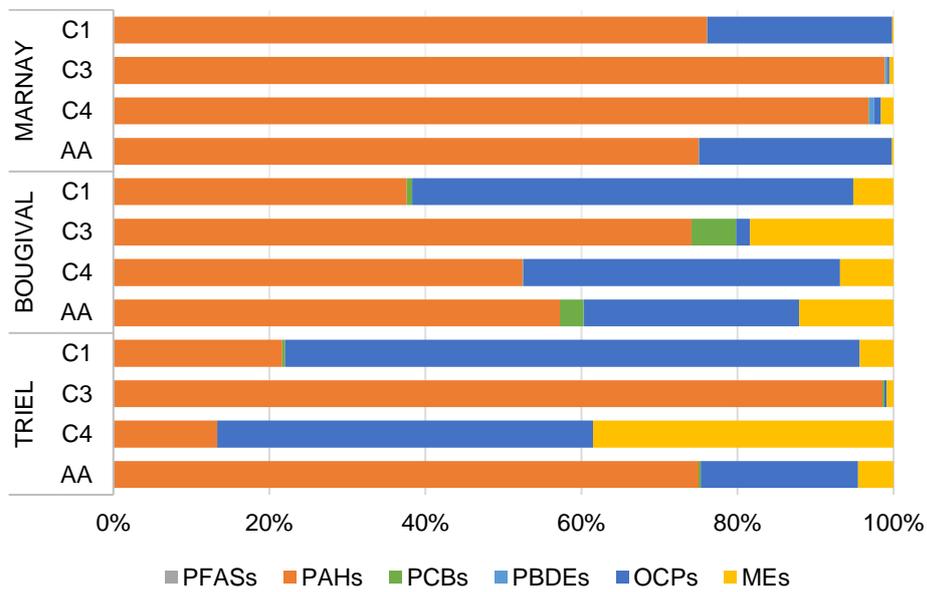


Fig. 2

(a)



(b)



(c)

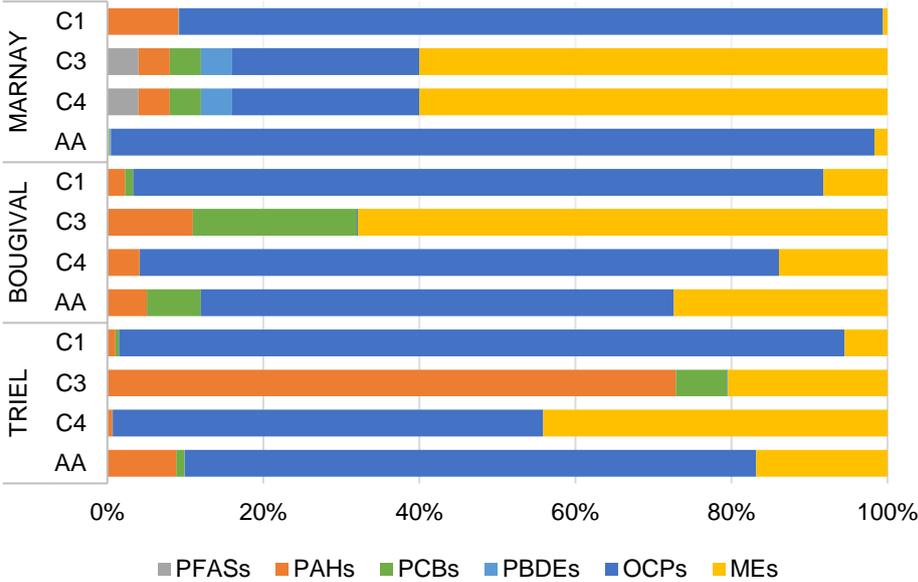
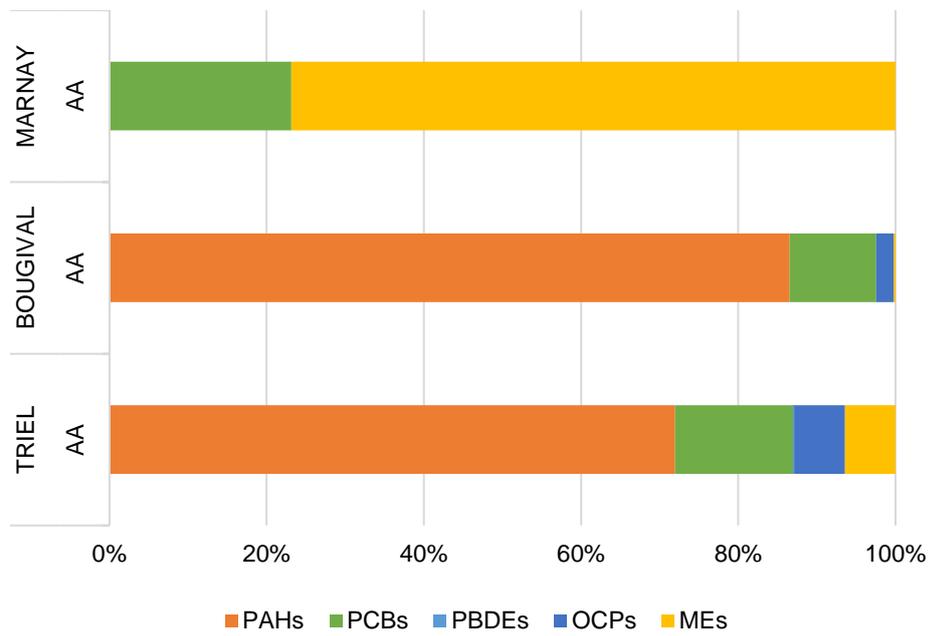


Fig. 3

(a)



(b)

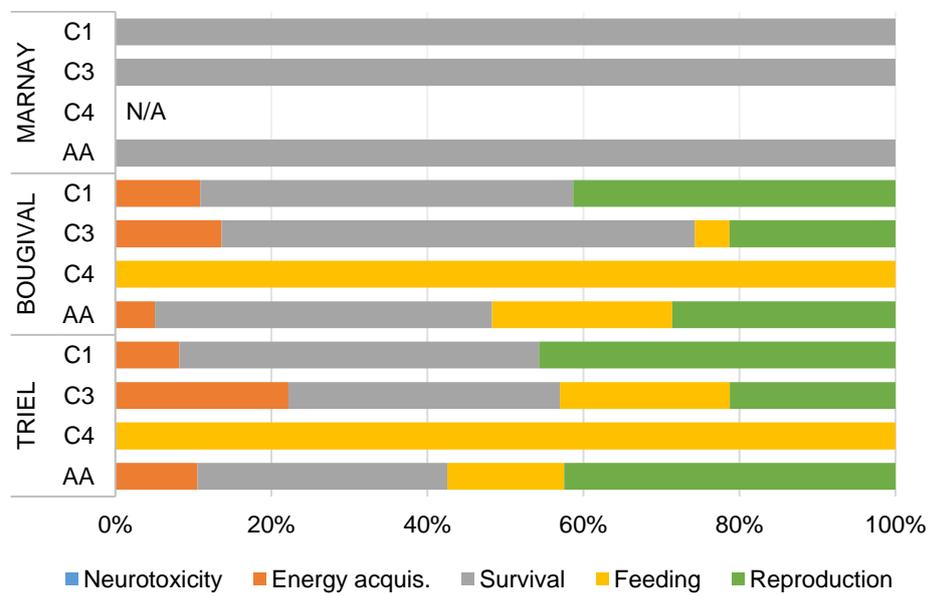
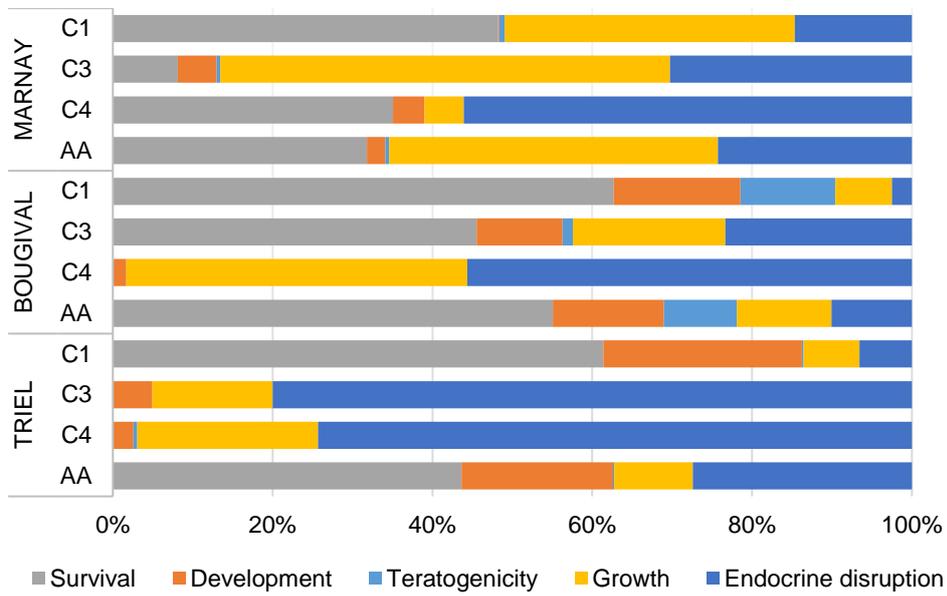


Fig. 4

(a)



(b)

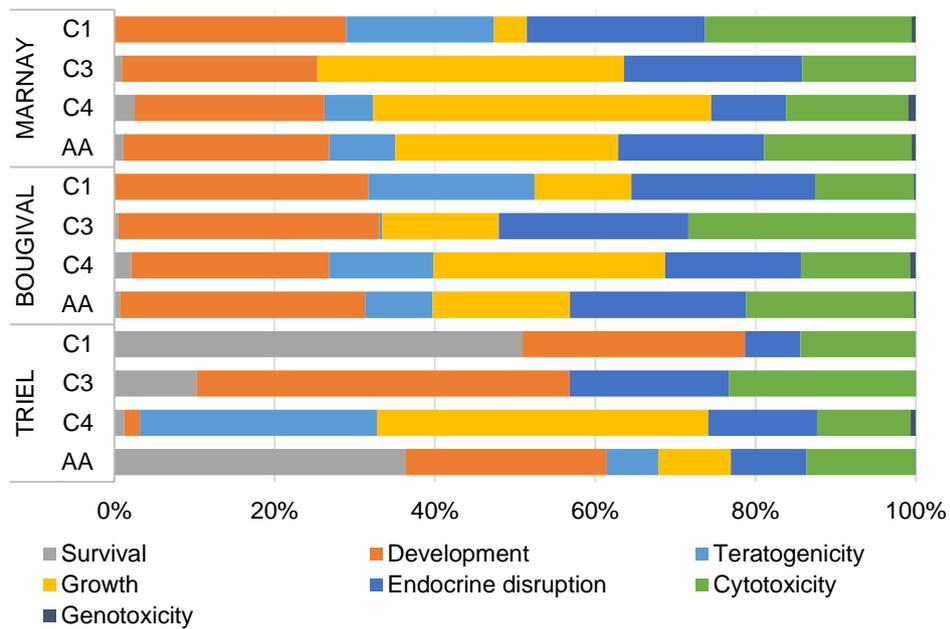
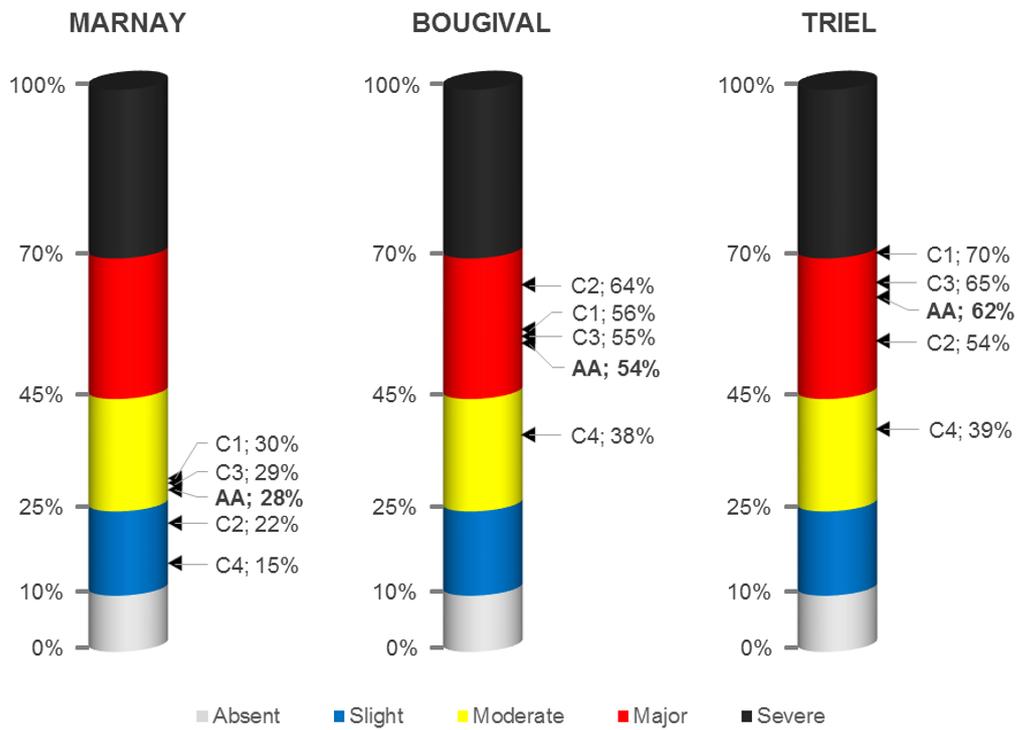


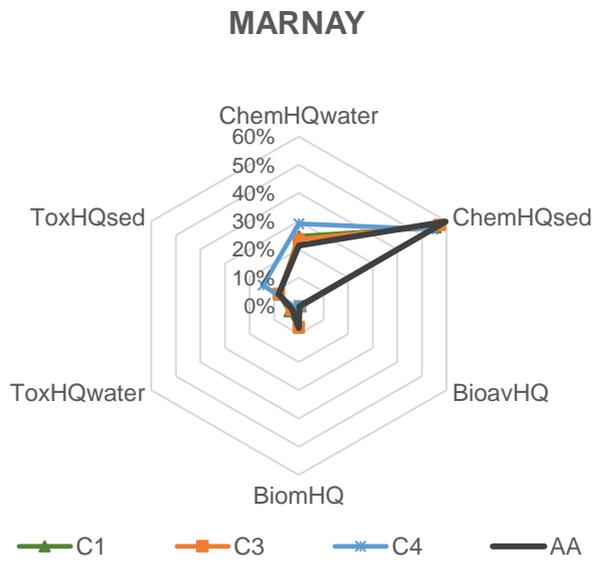
Fig. 5



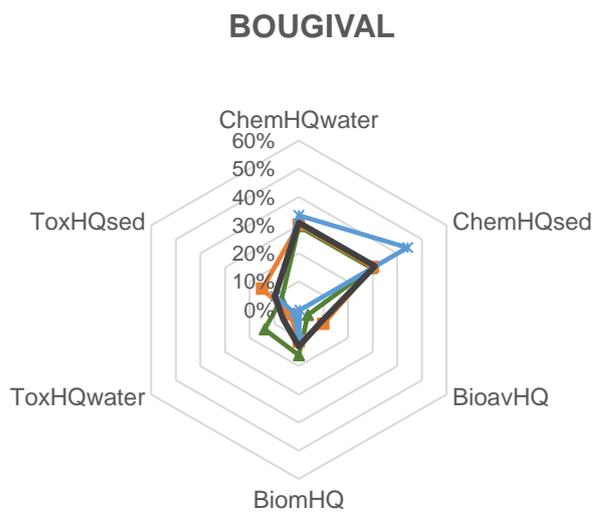
Summary of the hazard classes attributed to the HQs values within each LOE																
HQs	MAR_C1	MAR_C2	MAR_C3	MAR_C4	MAR_AA	BOU_C1	BOU_C2	BOU_C3	BOU_C4	BOU_AA	TRI_C1	TRI_C2	TRI_C3	TRI_C4	TRI_AA	
ChemHQ _{water}	Moderate	Moderate	Moderate	Moderate	Moderate	Severe	Severe	Severe	Major	Severe	Severe	Severe	Severe	Major	Severe	
ChemHQ _{sec}	Severe	N/A	Severe	Moderate	Severe	Severe	N/A	Severe	Severe	Severe	Severe	N/A	Severe	Severe	Severe	
BioavHQ	Slight	Slight	Slight	Slight	Slight	Major	Major	Major	Slight	Major	Major	Moderate	Moderate	Slight	Moderate	
BiomHQ	Moderate	N/A	Moderate	Slight	Moderate	Severe	N/A	Major	Major	Major	Severe	N/A	Severe	Major	Severe	
ToxHQ _{water}	Absent	N/A	Absent	Absent	Absent	Moderate	N/A	Absent	Absent	Slight	Major	N/A	Slight	Absent	Slight	
ToxHQ _{sec}	Slight	N/A	Slight	Slight	Slight	Slight	N/A	Moderate	Slight	Moderate	Severe	N/A	Severe	Slight	Major	

Fig. 6

(a)



(b)



(c)

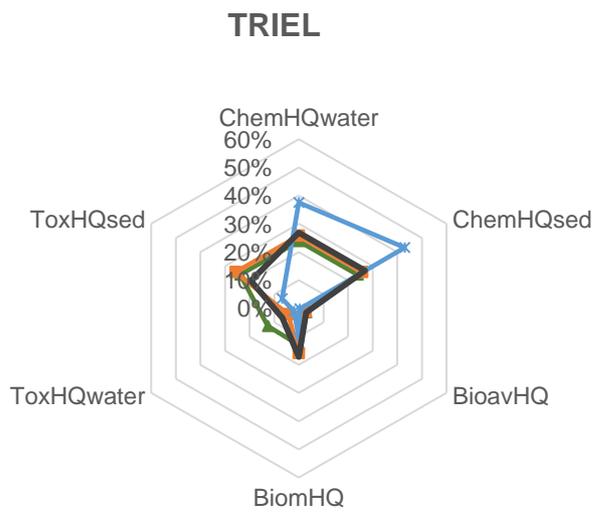


Fig. 7

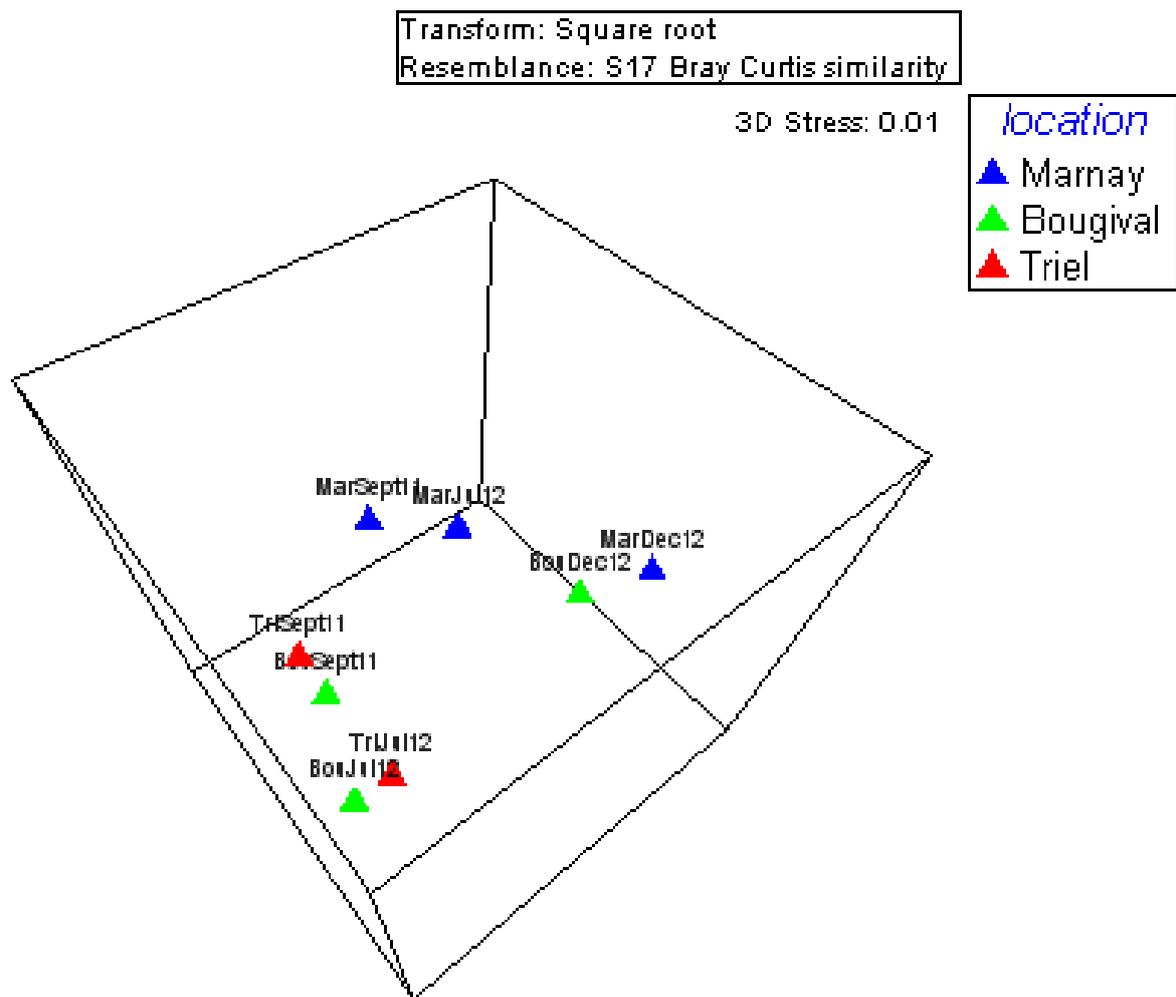


Table 1 Chemical Hazard Quotients calculated for the water column (ChemHQ_{water}) and for sediments (ChemHQ_{sed})

	Water column		Sediment			
	ChemHQ _{water}	Hazard class	ChemHQ _{sed} ^a	Hazard class	ChemHQ _{sed} ^b	Hazard class
MAR_C1	5.8	Moderate	92.6	Severe	24.3	Severe
MAR_C2	5.6	Moderate	N/A	N/A	N/A	N/A
MAR_C3	5.3	Moderate	17.1	Severe	0.3	Absent
MAR_C4	3.4	Moderate	6.2	Moderate	0.3	Absent
MAR_AA	4.7	Moderate	41.2	Severe	10.4	Major
BOU_C1	33.3	Severe	318	Severe	200	Severe
BOU_C2	14.9	Severe	N/A	N/A	N/A	N/A
BOU_C3	25.1	Severe	427	Severe	116	Severe
BOU_C4	9.9	Major	150	Severe	74.3	Severe
BOU_AA	19.6	Severe	301	Severe	133	Severe
TRI_C1	52.2	Severe	866	Severe	681	Severe
TRI_C2	29.1	Severe	N/A	N/A	N/A	N/A
TRI_C3	32.2	Severe	2,756	Severe	117	Severe
TRI_C4	11.3	Major	296	Severe	258	Severe
TRI_AA	30.5	Severe	1,309	Severe	355	Severe

a, Hazard quotient calculated using reference value (and measured concentration) for individual substance when available (excepted for ΣPBCs, ΣDDT, ΣDDE and ΣDDD; see ESM Table S2 for definition)

b, Hazard quotient calculated using reference value (and measured concentration) for a class of substances when available (for ΣDDTs, ΣPBDEs and ΣPAHs; see ESM Table S2 for definition)

N/A, not applicable; MAR, Marnay; BOU, Bougival; TRI, Triel, CX, Xth campaign; AA, annual average (mean of concentrations for C1 to C4 campaigns) value

Table 2 Bioavailability Hazard Quotients (BioavHQs) calculated using bioaccumulation levels in *Gammarus fossarum* following a 7 days-*in situ* exposure

	BioavHQ	Hazard class index^a	Hazard class
MAR_C1	5.8	147%	Slight
MAR_C2	6.0	135%	Slight
MAR_C3	13.4	165%	Slight
MAR_C4	1.8	150%	Slight
MAR_AA	6.5	129%	Slight
BOU_C1	58.6	924%	Major
BOU_C2	189	2006%	Major
BOU_C3	262	2300%	Major
BOU_C4	1.8	150%	Slight
BOU_AA	233	2012%	Major
TRI_C1	66.0	924%	Major
TRI_C2	64.7	665%	Moderate
TRI_C3	75.7	788%	Moderate
TRI_C4	1.8	150%	Slight
TRI_AA	77.4	829%	Moderate

a, Index used to attribute the hazard class (see Section 2.7.2 for details on calculation).

MAR, Marnay; BOU, Bougival; TRI, Triel, CX, Xth campaign; AA, annual average (mean of the measurements for C1 to C4 campaigns) value

Table 3 Biomarker Hazard Quotients (BiomHQs) calculated using biomarker responses in *Gammarus fossarum* following *in situ* exposure along the Seine River

	BiomHQ	Hazard class index^a	Hazard class
MAR_C1	26.4	154%	Moderate
MAR_C3	9.3	151%	Moderate
MAR_C4	0.0	70%	Slight
MAR_AA	11.9	151%	Moderate
BOU_C1	37.4	401%	Severe
BOU_C3	17.8	298%	Major
BOU_C4	12.5	216%	Major
BOU_AA	22.2	327%	Major
TRI_C1	51.2	401%	Severe
TRI_C3	26.9	446%	Severe
TRI_C4	9.4	216%	Major
TRI_AA	34.2	460%	Severe

a, Index used to attribute the hazard class (see Section 2.7.3 for details on calculation).

MAR, Marnay; BOU, Bougival; TRI, Triel, CX, Xth campaign; AA, annual average (estimated from the mean of the effects E(i) calculated for C1 to C4 campaigns) value

Table 4 Bioassay Hazard Quotients calculated for the water column (ToxHQ_{water}) and for sediment (ToxHQ_{sed})

	Water column			Sediment		
	ToxHQ _{water}	Hazard class index ^a	Hazard class	ToxHQ _{sed}	Hazard class index ^a	Hazard class
MAR_C1	4.8	0.6	Absent	16.2	1.1	Slight
MAR_C3	3.4	0.4	Absent	16.5	1.1	Slight
MAR_C4	0.8	0.1	Absent	14.9	1.0	Slight
MAR_AA	3.0	0.3	Absent	15.9	1.1	Slight
BOU_C1	32.7	3.8	Moderate	28.9	1.9	Slight
BOU_C3	7.8	0.9	Absent	57.6	3.9	Moderate
BOU_C4	3.2	0.4	Absent	25.0	1.7	Slight
BOU_AA	14.5	1.7	Slight	37.2	2.5	Moderate
TRI_C1	36.5	4.2	Major	113	10	Severe
TRI_C3	11.6	1.3	Slight	68.6	8.5	Severe
TRI_C4	3.2	0.4	Absent	18.2	1.3	Slight
TRI_AA	17.1	2.0	Slight	84.1	5.8	Major

a, index used to attribute the hazard class (see Section 2.7.4 for details on calculation).

MAR, Marnay; BOU, Bougival; TRI, Triel, CX, Xth campaign; AA, annual average (estimated from the mean of the effects E(i) calculated for C1 to C4 campaigns) value