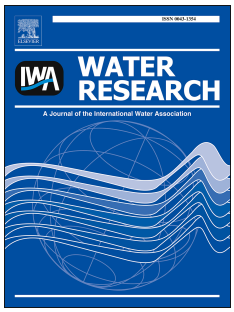
Accepted Manuscript



Effect-based nationwide surface water quality assessment to identify ecotoxicological risks

M.L. De Baat, M.H.S. Kraak, R. Van der Oost, P. De Voogt, P.F.M. Verdonschot

PII: S0043-1354(19)30426-9

DOI: https://doi.org/10.1016/j.watres.2019.05.040

Reference: WR 14689

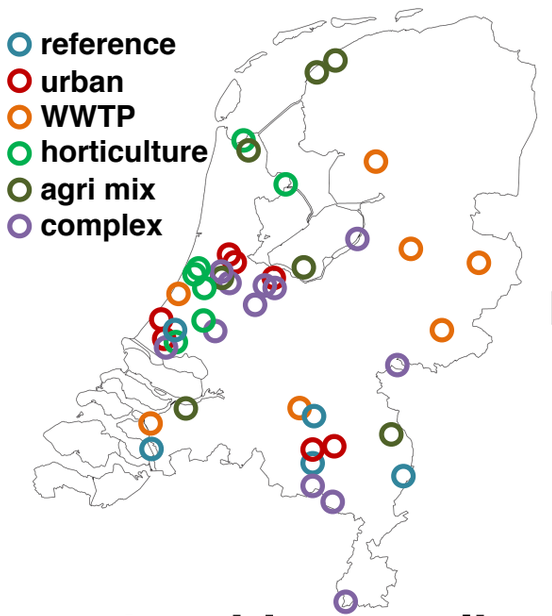
To appear in: *Water Research*

Received Date: 5 April 2019

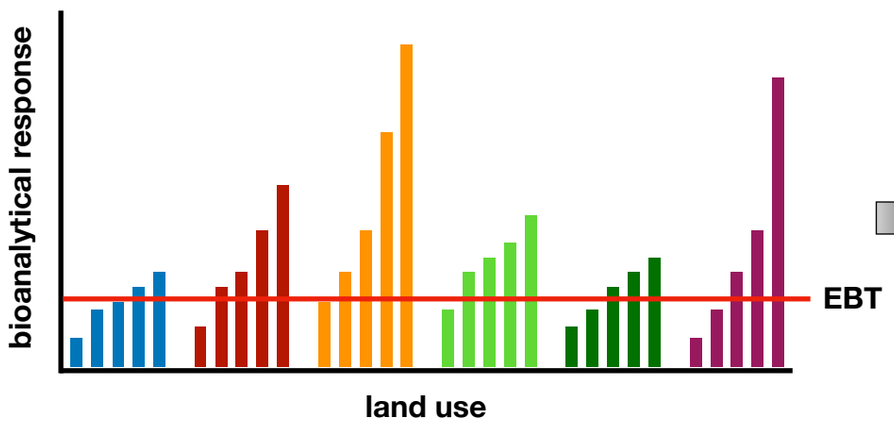
Revised Date: 8 May 2019

Accepted Date: 11 May 2019

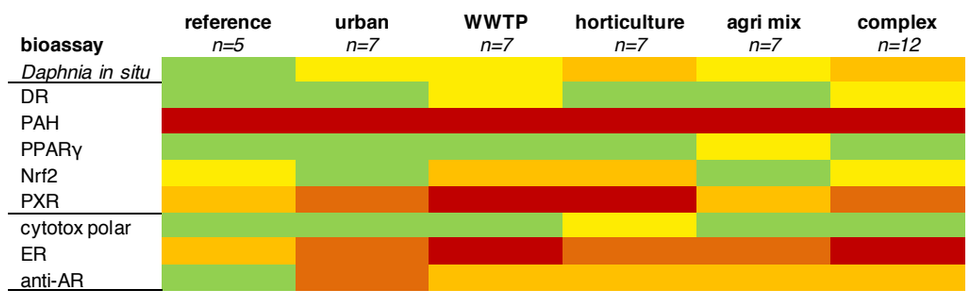
Please cite this article as: De Baat, M.L., Kraak, M.H.S., Van der Oost, R., De Voogt, P., Verdonschot, P.F.M., Effect-based nationwide surface water quality assessment to identify ecotoxicological risks, *Water Research* (2019), doi: https://doi.org/10.1016/j.watres.2019.05.040.



countrywide sampling



bioanalytical screening



ecotoxicological risk assessment

Effect-based nationwide surface water quality assessment to identify ecotoxicological risks

De Baat, M.L.,[1*] Kraak, M.H.S.,[1] Van der Oost, R.,[2] De Voogt, P.,[1,3] Verdonschot, P.F.M.[1,4]

1. Department of Freshwater and Marine Ecology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, The Netherlands
2. Department of Technology, Research and Engineering, Waternet Institute for the Urban Water Cycle, Amsterdam, The Netherlands
3. KWR Watercycle Research Institute, Nieuwegein, The Netherlands
4. Wageningen Environmental Research, Wageningen UR, The Netherlands

* Corresponding author. Department of Freshwater and Marine Ecology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Science Park 904, 1098 XH, Amsterdam, The Netherlands
email address: M.L.deBaat@uva.nl

Keywords: micropollutants, bioassay battery, passive sampling, Water Framework Directive, water monitoring

Abstract

A large portion of the toxic effects observed in surface waters cannot be attributed to compounds regularly measured by water authorities. Hence, there is an urgent need for an effect-based monitoring strategy that employs bioassays to identify environmental risks. The aim of the present study was to perform an effect-based nationwide water quality assessment to identify ecotoxicological risks in a wide variety of surface waters. At 45 locations silicone rubbers and polar organic chemical integrative samplers were exposed to surface water for 6 weeks. Alongside the passive samplers an *in-situ* daphnid test was performed. Subsequent to field exposure, accumulated compounds were extracted from the passive samplers after which a battery of *in vivo* and *in vitro* bioassays was exposed to the extracts. The bioassay battery was selected such that it could identify the risks posed by a wide range of chemical pollutants and their transformation products, while simultaneously allowing for targeted identification of groups of compounds that cause specific effects.

Bioassay responses were compared to effect-based trigger values to identify potential ecotoxicological risks at the investigated locations. Responses were observed in all bioassays, and trigger values were exceeded in 9 out of the 21 applied assays, allowing for ranking of the investigated locations based on ecotoxicological risks. No relationship between land use and the identification of ecotoxicological risks was observed. Based on the results, considerations regarding future improvements of effect-based monitoring are given. It is concluded that effect-based water quality assessment allowed prioritization of sites based on ecotoxicological risks, identified the presence of hazardous compounds regardless of being listed as priority substances, and meanwhile could prevent costly chemical analysis at sites with low ecotoxicological risks.

1. Introduction

According to the European Union (EU) Water Framework Directive (WFD)(The European Parliament and the Council of the European Union, 2013), chemical water quality is determined by monitoring surface waters for the presence of 45 (groups of) priority substances. However, the use of many of these compounds is restricted or banned and, as a result, concentrations of priority substances in European waters are decreasing (Altenburger et al., 2015; Fliedner et al., 2016). Simultaneously, industries have switched to a plethora of alternative compounds, which may enter the aquatic environment. Hence, the priority substances list is outdated, as the selected compounds are frequently absent nowadays, while the compounds present are not listed as priority substances (Busch et al., 2016; Schriks et al., 2010a; Schwarzenbach et al., 2006). Consequently, when toxic effects are observed in surface waters, these can often not be attributed to compounds measured by water authorities (Altenburger et al., 2015; Neale et al., 2015). Risks of pollutants to freshwater ecosystems are thus caused by mixtures of a myriad of (un)known, unregulated and unmonitored compounds (Daughton, 2005). Understanding of these risks requires a

paradigm shift, that allows for new holistic monitoring methods that do not solely depend on chemical analysis of priority substances, but contrastingly consider biological effects first (Hamers et al., 2018; Leusch et al., 2014; van der Oost et al., 2017a/b). Therefore, there is a need for an effect-based monitoring strategy that employs bioassays to identify environmental risk (Brack et al., 2017; Wernersson et al., 2015).

Bioassay responses to surface water samples are caused by the combined action of mixtures of all bioavailable (un)known compounds and their metabolites present, thereby overcoming the limitations posed by chemical analysis of a limited number of target compounds (Brack et al., 2017; Neale et al., 2015). Indeed, the applicability and reproducibility of a battery of bioassays to identify ecotoxicity in regular water quality monitoring has been shown in recent years (Blackwell et al., 2019; Di Paolo et al., 2016; Hamers et al., 2018; Jia et al., 2015; Leusch et al., 2014; Novák et al., 2018; van der Oost et al., 2017a). The ecotoxicity profiles of the surface water samples that are generated by such a bioassay battery allow for calculation and ranking of a cumulative ecotoxicological risk for the selected locations. Subsequently, at locations where risks are identified, it becomes relevant to investigate the drivers of the observed effects. The aim of the present study was therefore to identify ecotoxicological risks in an effect-based nationwide water quality assessment in a wide variety of surface waters in The Netherlands.

The success of effect-monitoring relies largely on the ease of use, endpoint specificity and scale of the used bioassays, as well as on the ability to interpret the measured responses. To ensure sensitivity to a wide range of potential stressors, while still providing specific endpoint sensitivity, the present study employed a previously successfully implemented bioassay battery including *in situ* whole organism assays as well as laboratory based whole organism *in vivo* and mechanism specific *in vitro* assays (van der Oost et al., 2017a). Adverse effects in the whole organism assays point to general toxic pressure and represent a high ecological relevance. *In vitro* or small-scale *in vivo* assays that target highly specific molecular initiating events allow for focused identification and subsequent confirmation of (groups of) toxic compounds (Brack et al., 2016; Escher et al., 2018; Neale

et al., 2017). The identification of ecotoxicological risks from bioassay battery responses follows from the comparison of bioanalytical signals to previously determined thresholds, defined as effect-based trigger values (EBT), that differentiate between acceptable and poor water quality (Tang et al., 2013). Recently van der Oost et al. (2017b) and Escher et al. (2018) derived EBTs for a variety of bioassays commonly applied in surface water quality assessment.

An additional limitation of the present chemical water quality assessment is that grab sampling is commonly used for surface water sample collection. Yet, concentrations of compounds typically vary over time and therefore grab sampling only provides a snapshot of the chemical make-up of a water body (Jones et al., 2015). Passive sampling can overcome these limitations by exposing a sorbent to the target environment for several weeks to months, accumulating compounds from the water over time (Vrana et al., 2005). In this way, passive sampling integrates fluctuations in compound concentrations in time, and simultaneously enriches surface water samples to an extent that (bio)analytical detection limits become very low. Current limitations of passive sampling in water quality assessment are the compound selectivity of the receiving phase and the challenge of precisely determining the sampled volume of water (Roll and Halden, 2016). Nonetheless, the advantages of passive sampling compared to grab sampling outweigh its disadvantages, and passive sampling is increasingly applied as a valuable tool in the monitoring of environmental contaminants. Hence, the combination of passive sampling and effect-monitoring allows for time-integrated and reliable surface water quality assessment, that considers effects of all sampled (un)known compounds, regardless of priority lists.

In the present study, silicone rubber (SR) and polar organic chemical integrative sampler (POCIS) passive samplers were applied at 45 surface water locations. Alongside the passive samplers an *in-situ* daphnid test was performed. Subsequent to field exposure, accumulated compounds were extracted from the passive samplers after which a battery of *in vivo* and *in vitro* bioassays was exposed to the extracts. Bioassay responses were compared to effect-based trigger values to identify potential ecotoxicological risks at the

investigated locations. Finally, responses were related to surrounding land use, water body morphology and WFD ecological water quality assessment scores.

2. Material & Methods

2.1. Sampling sites

Sampling sites were selected in collaboration with 12 Dutch waterboards and the Dutch national water authority. Sites were classified based on the major surrounding land use or potential source of pollution. When classification of a location was not possible due to the diffuse or variable nature of contamination, it was assigned to the category “complex”. This resulted in the classification of 45 surface water locations into six categories (Figure S1): reference (n = 5), urban (n = 7), wastewater treatment plant effluent impacted (WWTP; n = 7), horticulture (n = 7), mixed agriculture (agri mix; n = 7) and complex (n = 12).

2.2. Deployment, extraction and estimating sampled volumes of passive samplers

2.2.1. Passive sampling devices

Silicone rubber (SR) sheets, with a weight of 20 g per set of six sheets, spiked with performance reference compounds (PRCs), were obtained from Deltares (Utrecht, The Netherlands) and applied for the sampling of nonpolar compounds.

POCIS, containing 0.22 g Oasis HLB sorbent, were obtained from Exposmeter (Tavelsjö, Sweden) and applied for the sampling of compounds in the more polar range. No sampler pre-treatment was required, and the samplers were transported to the study sites in their original airtight packaging.

2.2.2. Field deployment of passive samplers

SR sheets and POCIS were deployed simultaneously at each sampling location in cages to attach and protect the passive samplers during the exposure period. Cages were secured to the bottom or to the embankment to avoid loss of samplers and to ensure

136 permanent inundation. Per location, six SR sheets and four POCIS were exposed for a
137 period of six wk. After exposure, the samplers were transported to the laboratory and stored
138 at –20 °C until extraction.

139 *2.2.3. Extraction of silicone rubber*

140 SR sheets were cut into small pieces and put in precleaned thimbles of a Tecator
141 Soxtec Avanti 2050 extraction system. Extractions were performed in 80 mL of a
142 MeOH:acetonitrile (1:2 v/v) mixture with boiling stones. The extraction program was as
143 follows: 120 min boiling at 180 ˚C, 30 min rinsing, 5 min recovery, and 1 min drying. Cooled
144 extracts were filtered over glass fiber filters and collected in 250 mL glass bottles. Extraction
145 jars were rinsed twice with 10 mL of extraction mixture. Extracts were evaporated by
146 TurboVap II Zymark at 45 ˚C to approximately 5 mL, transferred quantitatively (rinsed twice
147 with 5 mL extraction mixture) to 15 mL conical tubes, evaporated under nitrogen, and finally
148 the end volumes were filled up to exactly 10 mL with extraction mixture.

149 *2.2.4. Extraction of POCIS*

150 To enable elution, the sorbent between the POCIS membranes was transferred
151 quantitatively into an empty solid phase extraction (SPE) column with a polyethylene frit.
152 Columns were dried under vacuum extraction, followed by centrifugation (2000 rpm, 15 min)
153 under nitrogen flow. Dry SPE columns were eluted three times with 3 mL of acetone, with 5
154 min equilibration time between elutions. Eluates were collected in 10 mL conical tubes, and
155 the end volumes were filled up to exactly 10 mL with acetone.

156 *2.2.5. Estimation of sampled water volumes*

157 *2.2.5.1. Silicone rubber*

158 SR sheets were spiked with PRCs with a wide hydrophobicity range (biphenyl D10
159 and polychlorinated biphenyl (PCB) congeners 1, 2, 3, 10, 14, 21, 30, 50, 55, 78, 104, 145,
160 and 204) that do not occur in Dutch surface waters. For PRC chemical analysis, SR extracts
161 were transferred to petroleum ether by adding 2 mL extract to 40 mL petroleum ether and
162 concentrated with Kuderna Danish at 80 ˚C. The petroleum ether extract was cleaned up
163 with aluminium oxide and silica gel column chromatography. The cleaned extract was

164 evaporated to exactly 2 mL and analysed with an Agilent 7890 Triple Quadrupole gas
165 chromatography mass spectrometer (GC-MS/MS) equipped with Edwards pump.
166 Quantification was performed using an external calibration series of 6 concentrations. The
167 rate of PRC dissipation was used to calculate the exchange rates (R_S values, in L/day) of the
168 samplers (Booij and Smedes, 2010). Subsequently, 50% of this calculated R_S was used as a
169 provisional estimation of the average extracted water volume per day as described by van
170 der Oost et al. (2017a).

171 *2.2.5.2. POCIS*

172 While standardized protocols for the determination of sampled volume of passive
173 samplers have been described for nonpolar samplers, no such consensus has yet been
174 reached for polar passive samplers (Harman et al., 2011). This is partly due to the different
175 nature of polar and nonpolar passive samplers, and the processes that hence dictate the
176 uptake of polar compounds in passive samplers (Harman et al., 2012). Sampling rates for
177 polar compound uptake in POCIS in stagnant to near stagnant water have been reported in
178 the range from 0.001 to 2.46 L/day, with an average sampling rate of 0.18 L/day (Harman et
179 al., 2012). Hence, to compare bioassay effects between sites, in the present study the same
180 estimated average sampled volume of 0.18 L/day was applied to determine the
181 concentration factor of all field deployed POCIS.

182 **2.3. Bioassay battery**

183 Whole organism bioassays and antibiotics WaterSCAN assays were performed at
184 the Waterproef Laboratory (Edam, The Netherlands). *In vitro* CALUX assays were
185 performed at the BioDetection Systems laboratories (Amsterdam, The Netherlands). Passive
186 sampling extracts were converted to other solvents before exposure in the bioassays. More
187 information on bioassay analytical details and solvent transfer is given in the supplementary
188 information (pages S2-5). An overview of the employed bioassays, their endpoints, and their
189 respective units of effect expression and EBTs is given in Table 1.

190 **2.3.1. *Daphnia in situ* exposure**

Daphnids were exposed to the surface water at 33 of the 45 study sites in glass jars. Field exposure was carried out during the first or second week of the passive sampler deployment. The survival of the *in situ* exposed daphnids was recorded after 1 wk of exposure. An observed mortality of 20% was used as trigger value for potential ecological effects (van der Oost et al., 2017b).

2.3.2. Whole organism bioassays

For the whole organism bioassays, SR passive sampler extracts were subjected to three bioassays. As these whole organism bioassays have no specific target compound group, toxicity is expressed as toxic units (TU), rather than reference compound equivalents. Herein, one TU represents the dilution at which the extract causes 50% effect for the respective endpoint of the test (EC_{50}).

2.3.2.1 Daphnia 48 h immobilization

The *Daphnia* 48 h immobilization assay was performed according to the Organisation for Economic Co-operation and Development (OECD) standard 202 (OECD Environmental Health and Safety, 2004), with reduced test volumes. EC_{50} values (volume percentage) were determined by nonlinear regression analysis with a log-logistic model by the statistical program SPSS (IBM Analytics). Bioassays were considered valid if >90% of the daphnids in the negative controls were mobile at the end of the test.

2.3.2.2. Algatox

The inhibition of algal growth was determined according to OECD standard 201 (OECD Environmental Health and Safety, 2006), with reduced test volumes, based on Peterson et al. (2005). After 72 h, exponential algal growth curves were determined to assess the percentage of growth inhibition compared to controls. Algal growth rate in the controls was required to reach 0.92/d, according to the OECD standard. The EC_{50} values were calculated using sigmoidal dose–response curves with variable slopes.

2.3.2.3. Microtox

The Microtox® test was performed by exposing the bioluminescent marine bacterium *Aliivibrio fischeri* to a dilution range of the passive sampler extracts. Toxicity was determined by quantifying inhibition of the luminescence produced by *A. fischeri* exposed to the extracts after 5, 15, and 30 min of exposure. Microtox Omni software (version 1.18) was used for determination of the TU values.

2.3.3. CALUX assays

Passive sampler extracts were analysed by a panel of in vitro CALUX® bioassays. Specific CALUX assays were performed on either polar or non-polar extracts, as suggested by van der Oost et al. (2017a). SR extracts were subjected to DR, PAH, PPARγ, Nrf2, PXR and p53 (with and without S9 metabolism) assays and POCIS extracts were subjected to ERα, anti-AR and GR assays, according to previously described protocols (Hamers et al., 2006; Murk et al., 1996; Sonneveld et al., 2004; Van Der Linden et al., 2008). The DR CALUX assay was performed without a sulfuric acid clean up step to eliminate PAHs and isolate the dioxins and dioxin-like polychlorinated biphenyls. To rule out confounding influences, cells were also monitored for cytotoxicity, which resulted in additional data for cytotoxicity caused by both polar and non-polar passive sampler extracts. The effects of the extracts were expressed as bioanalytical equivalents (BEQs) of the reference compounds (Table 1).

2.3.4. Antibiotics activity assay

Activities of 5 classes of antibiotics in the POCIS extracts were determined with the WaterSCAN assay, obtained from RIKILT (Wageningen, The Netherlands). The test system comprised 5 plates (details outlined in Pikkemaat et al., 2008): tetracyclines (T), quinolones (Q), β-lactams and macrolides (B+M), sulphonamides (S), and aminoglycosides (A). After incubation of the test plates, antibiotic activities were estimated and expressed as BEQ concentrations of the reference compounds (Table 1).

2.4. Data analysis

Bioassay effects were expressed as BEQ/L by using the estimated sampled water volumes of the respective passive samplers to determine the concentration factor of the used extracts. Subsequently, bioassay effects were compared to previously defined EBTs. EBTs from Escher et al. (2018) were utilized when available, and when the used reference compounds matched those applied in the current study. This was the case for the PAH, anti-AR and ER CALUX assays. For all other applied bioassays, EBTs from van der Oost et al. (2017b) were used (Table 1).

Average numbers of EBT exceedances per land use category were tested for equality of variances using a F-test, and subsequently differences between land use were tested for significance using a Two-sample T-test assuming equal variances ($\alpha = 0.05$). Statistical analyses were performed in Excel for Mac version 16 (Microsoft).

Multivariate analysis was applied to gain insight in the relationship between the surrounding land use, water type and ecological water quality and the bioanalytical responses. Only the tests that showed a response above the respective EBT were included. The total bioanalytical dataset consisted of 9 responding bioassays and 45 locations. Alongside this response matrix, two location variables were included in the multivariate analysis: A measure of ecological quality, expressed as WFD ecological quality assessment scores for macrofauna (EQR mafa), obtained from the Dutch waterboards, and the water type of the locations, expressed as ditch, pond or lake for lentic waters, and stream, channel or river for lotic waters. Missing values in the dataset were substituted with the average response value for each bioassay, to minimise their effect on multivariate analysis outcome. Bioassay responses were transformed to a logarithmic scale and the resulting dataset was ordinated by redundancy analyses (RDA) in CANOCO 4.2 for Windows (Ter Braak, 1990, 1988). The data analyses are fully described by Verdonschot and Ter Braak (1994). An unrestricted permutation test was used to test the validity of the total ordination as described by Ter Braak (1990) and Verdonschot and Ter Braak (1994).

3. Results

3.1. Bioassay battery responses to passive sampler extracts

Passive samplers for polar and non-polar compounds were successfully exposed at 45 surface water locations. During extraction, POCIS extracts were lost for three sampling locations, resulting in an incomplete dataset for these locations. Therefore, these locations were excluded from the comparison of EBT exceedances per location. All bioassays met their respective validity criteria. Responses were observed in all 21 bioassays, but for each bioassay there were clear differences in the strength of the responses between the locations. A representative example of the 21 bioassays is given in Figure 1, which depicts the estrogen receptor (ER) CALUX responses to the POCIS extracts. ER responses were observed at all but one location, with only a non-detect at one of the reference locations. The intensity of the response was highly variable for the different locations per land use, with the highest response at one of the urban locations. On average, the highest responses were observed at urban (0.40 ng EEQ/L), complex (0.38 EEQ/L) and WWTP (0.36 EEQ/L) locations, while the lowest responses were observed at the reference locations (0.13 EEQ/L). This is also reflected by the percentage of EBT exceedances per land use category, where EBT exceedance for the ER CALUX assay was observed at the majority of urban (71%), complex (89%) and WWTP (86%) locations, while the EBT was exceeded at 40% of the reference locations. Responses in the other 20 bioassays are listed in Table S1. This information was subsequently used to calculate the number of EBT exceedances per location.

3.2. Effect-based trigger value exceedances per location

All locations caused the exceedance of at least one EBT in the bioassay battery. The sum of EBT exceedances per location and the resulting average number of EBT exceedances per land use category are depicted in Figure 2. The variation between

297 locations within a land use category was largest for horticultural locations, while urban
298 locations showed the most consistent number of responses above the EBT per location. The
299 lowest average number of EBT exceedances was observed at reference locations (2), and
300 the highest number of EBT exceedances were observed at urban (3), WWTP (3.9), complex
301 (3.4) and horticulture (3.4) locations, including one location with seven EBT exceedances in
302 the latter. However, only at WWTP locations the average number of EBT exceedances was
303 significantly higher than at the reference ($p < 0.01$) and mixed agriculture ($p < 0.05$)
304 locations. The sum of EBT exceedances per site allowed for the ranking of sites based on
305 ecotoxicological risk, where the sites with the highest number of EBT exceedances are
306 assumed to be at the highest risk of surface water pollution (Table S1). EBT exceedances
307 were observed for 9 out of the 21 applied bioassays: For the *in situ Daphnia* test and for 8 *in*
308 *vitro* CALUX assays performed with both non-polar and polar extracts.

309

310 **3.3. Ecotoxicological risk identification**

311 Next, a heat map was constructed that visualizes the percentage of the investigated
312 locations with EBT exceedance per land use category (Figure 3). Interestingly, the EBT for
313 the PAH CALUX assay was exceeded at all the investigated locations, and hence this assay
314 did not allow for differentiation in ecotoxicological effect identification between locations or
315 land uses. Reference locations showed the lowest percentage of EBT exceedances,
316 however ecotoxicological risk was not completely absent, with responses in PAH, Nrf2, PXR
317 and ER CALUX tests. At urban locations, ecotoxicological risks were driven most strongly by
318 PXR, ER and anti-AR activity. At WWTP locations, the most profound contribution to
319 ecotoxicological risks was caused by PXR and ER activity. A less frequent, but nonetheless
320 substantial contribution to ecotoxicological risks was observed for Nrf2 and anti-AR activity.
321 Similarly, for horticultural locations, ecotoxicological risks were most frequently caused by
322 PXR and ER, with a contribution of ER and anti-AR activity. However, risk indication was

also frequently observed in the *Daphnia in situ* test, and polar extracts from horticultural locations were the only samples to cause cytotoxicity above the EBT. For mixed agricultural locations, risks were most frequently caused by ER activity, with less frequent contributions of PXR and anti-AR. Notably, mixed agricultural locations were the only ones to cause EBT exceeding PPARγ activity. Complex locations also showed EBT exceedances most frequently for PXR and ER, with less frequent exceedances for anti-AR activity and the *Daphnia in situ* test. DR activity above the EBT was observed only at WWTP and complex locations.

3.4. Response frequency per bioassay

Responses of all bioassays were summarized to gain insight into which assays responded most frequently to the passive sampler extracts and were hence the main determinants of the detection of ecotoxicological risks (Figure 4). Bioassay signals were categorized as no response, or a response below or above the EBT of that test. The frequency of effect detection in the bioassays ranged from largely no response (96% of locations) for the Algatox assay, to responses above the EBT at all locations for the PAH CALUX assay. The PAH, PXR, Nrf2 and DR CALUX assays showed a response at all locations, however with a varying frequency of responses above the EBT, with the most striking result for the PAH CALUX assay, for which the EBT was exceeded at all the investigated locations. Nine out of the battery of 21 bioassays showed responses above the EBT (Figure 3 & 4). The other 12 assays gave no response above their EBT. Out of these, nine showed no bioanalytical response at all at more than 50% of the investigated locations. These were the GR CALUX and the five antibiotics assays which were exposed to the polar passive sampler extracts, and the whole organism Daphniatox and Algatox and the *in vitro* p53 CALUX assay with S9 metabolism exposed to the non-polar extracts.

3.5. Multivariate analysis

The ordination result of the RDA with land use as explaining variable is presented as a correlation biplot of bioassay responses, land use, and environmental quality scores (Figure S1). The RDA revealed no significant variables in the dataset. Hence, land use, water type and the ecological quality score did not explain the variation observed in the bioassay battery responses.

4. Discussion

4.1. Effect-based identification of ecotoxicological risks

In the present study, an effect-based nationwide water quality assessment to identify ecotoxicological risks was performed. Effects were observed in all bioassays, and EBT exceedances were observed for 9 out of the 21 bioassays. The sum of EBT exceedances per site allowed for the ranking of sites based on ecotoxicological risk, rather than on the presence of a limited number of target compounds (Hamers et al., 2018), which can be considered as a proof of principle of effect-based water quality assessment. Subsequently, at locations where risks were identified, it becomes relevant to investigate the drivers of the observed effects.

The bioassays that showed responses above EBTs in the present study, and hence allowed the identification of ecotoxicological risks, were the DR, PAH, PPARγ, Nrf2 and PXR CALUX assays for non-polar extracts, the ER, anti-AR and cytotoxicity CALUX assays for polar extracts, and the *in situ Daphnia* assay. This is partly in line with previous findings by Escher et al. (2014) and van der Oost et al. (2017a), that identified high responses of *in vitro* assays for, amongst others, PAH, Nrf2, PXR and ER and anti-AR activity in surface water. Following from the observed CALUX responses, in the present study, risks were caused by both polar and non-polar organic extracts. Several of these tests indicated risks at the

majority of the studied locations. Most notably the PAH CALUX, which indicated ecotoxicological risks of polycyclic aromatic hydrocarbons (PAH) at all sites. This can in part be explained by the atmospheric origin of PAH loading to aquatic systems, causing the presence of PAHs even at locations with very limited anthropogenic pollution (Manoli and Samara, 1999). Interestingly, however, this was not the case for dioxins, which also partly find their way to the aquatic environment through atmospheric deposition (Kulkarni et al., 2008). In the present study, risk of dioxins was only observed for WWTP and complex locations, and infrequently at both. As for both groups of compounds the ultimate environmental sink is the sediment, which was not examined in the present study, this difference may be explained by the current emissions, which, in Europe, are more strongly regulated for dioxins and more common for PAHs (Kulkarni et al., 2008; Manoli and Samara, 1999).

The PAH and DR CALUX assays both target aryl hydrocarbon receptor binding, yet after different exposure times (4 vs. 24 h respectively), which affects the *in vitro* metabolization of PAHs (Pieterse et al., 2013). Since in the present study, water extracts subjected to the DR CALUX assay were not treated with a sulfuric acid clean up step to eliminate PAHs and isolate the dioxins and dioxin-like polychlorinated biphenyls, responses in the DR CALUX assay may be caused by stable PAHs that were not metabolized during the 24 h exposure. Thus, had the extracts been cleaned up with sulfuric acid, the three samples that showed EBT exceedance in the DR assay may well have lost their activity due to destruction of stable PAHs. This strengthens the observation that ecotoxicological risks in the investigated surface waters are much more common for PAHs than for dioxins.

Besides the ubiquity of PAHs in surface waters, the detection of ecotoxicological risk also depends on the EBT value used for this specific test. In the case of the PAH CALUX assay, this EBT value was obtained from the study by Escher et al. (2018), in which EBTs were derived by read across from existing EU WFD environmental quality standards (EQS). This resulted in an EBT value of 6.21 ng benzo(a)pyrene equivalents (BEQ) per litre, which

401 is substantially lower than the EBT of 150 ng BEQ/L derived by van der Oost et al. (2017b).
402 Had we applied the latter, the resulting detection of ecotoxicological risk caused by toxic
403 PAH concentrations in surface water would have been markedly less dramatic, and would
404 have resulted in an EBT exceedance at only a single location. However, the study by Escher
405 et al. (2018) based their EBT on existing EQS values, about which a European wide
406 consensus exists, and which reliably indicates ecotoxicological risks to aquatic communities.
407 Hence, the dramatic EBT exceedance observed here may identify a serious risk posed by
408 PAHs in the majority of waterbodies, even at locations with very few other anthropogenic
409 pollution sources. Nonetheless, the profound influence of the value of the EBT for each
410 bioassay on the detection of ecotoxicological risks should not be underestimated. This
411 underlines the need for a standard procedure and consensus on EBT derivation and values
412 for the successful application of effect-based monitoring strategies in water quality
413 assessment.

414 **4.2. Identification of location and land use specific ecotoxicological risks**

415 Although several unique responses for the different land use and bioassay combinations
416 were observed, no land use specific responses or patterns became apparent, and only small
417 differences in EBT exceedances between land use types were found. This observation was
418 corroborated by the outcome of the multivariate analysis, which revealed no significant effect
419 of land use on the bioassay battery responses. The selected locations appear to suffer from
420 the presence of complex mixtures of micropollutants, frequently caused by the same drivers.
421 Hence, to identify pollution source specific drivers of ecotoxicological risks, in future
422 research locations should be selected that better represent a single pollution source and that
423 are more morphologically and biogeochemically similar to exclude confounding effects. Yet
424 at the same time, these findings also confirm the complex nature of surface water pollution
425 in large river deltas. This raises the question if categorizing sites into land use types is
426 appropriate at all, and if alternatively, sampling sites may better be considered independent
427 stochastic draws of diffuse pollution covering the industrialized world. When applying that

paradigm, in the present study, several discriminating bioassays allowed for the identification of locations at risk from chemical stressors, and for the ranking and subsequent prioritization of the locations that were at the highest risk from micropollutants.

4.3. Considerations for improved effect-based monitoring

EBT exceedances were observed for 9 out of the applied 21 bioassays, indicating that 12 bioassays were less effective in elucidating ecotoxicological risks at the studied locations. The bioassays that were not discriminating for ecotoxicological risks were the *in vivo* whole organism bioassays, the antibiotics assays, and the p53 (with and without S9 metabolism), GR and cytotoxicity (for non-polar compounds) CALUX assays.

4.3.1. p53 and cytotoxicity CALUX assays

The p53 and cytotoxicity CALUX assays indicate risks at a high organisational level caused by all compounds in a water sample (Escher et al., 2018; Van der Linden et al., 2014; van der Oost et al., 2017b). Hence, signals above the EBT in these assays would imply far stretching ecological effects in the field (Maltby, 1999). Therefore, although these tests did not respond frequently or severely to surface water passive sampler extracts in the present study, the inclusion of such tests in future bioassay batteries is recommended given their ecological relevance. Yet, the inclusion of S9 metabolism in the p53 test can be debated. The S9 metabolism in this assay can elucidate the enzymatic activation of mutagenicity in the sample. However, given the time integrative nature of passive sampling (six weeks in the present study), metabolism and activation of more toxic or persistent metabolites is expected to occur in the field rather than in the laboratory, and the added value of *in vitro* metabolization is negligible. This was also illustrated by the much lower p53 test response after S9 metabolism in the present study. Hence, the p53 assay without S9 metabolism should be sufficient to assess mutagenicity of surface water samples in monitoring strategies that apply passive sampling techniques.

4.3.2. Whole organism bioassays

For the whole organism bioassays, it can be argued that more sensitive alternatives should be developed and applied. For example, the Algatox assay showed no response to surface water extracts from approximately 95% of the locations in the present study. This is unexpected, as herbicides, that are the major target compound group of this bioassay, are the most frequently detected pesticide group in European surface waters (Booij et al., 2015; Schreiner et al., 2016). Recent work has shown that fluorescence based algal bioassays are efficient and effective in the assessment of toxicity to primary producers in regionwide screening efforts (de Baat et al., 2018; Novák et al., 2018; Sjollema et al., 2014). Hence, in the future, replacement of the Algatox assay with fluorescence based algal bioassays may result in more effective assessment of risks to primary producers in effect-based monitoring.

Finally, the applicability of the *in situ Daphnia* assay in micropollutant effect monitoring should be questioned. Although it was responsive and discriminating in the present and previous studies (van der Oost et al., 2017a), it is nearly impossible to determine the contribution of micropollutants to the observed mortality. Exposure of daphnids in the field for seven days gives rise to a multitude of confounding factors including oxygen dynamics, food availability, pH, salinity and temperature, and unless the effects of these on daphnid mortality can be fully excluded, the outcome of the test cannot be considered indicative of micropollutant risk in surface water. Nonetheless, the added value of *in situ* or active biomonitoring approaches in water quality assessment strategies should not be underestimated, as they represent the most realistic exposure scenario available in the effect-based toolbox. Recently, promising strategies to differentiate between the effects of chemical exposure and confounding factors in active biomonitoring with invertebrates were described (e.g. Brettschneider *et al*. 2019).

4.3.3. Antibiotics and GR CALUX assays

The antibiotics and GR CALUX assays target specific groups of compounds, and their inclusion in bioassay batteries is only justified when there is an assumable occurrence and

480 risk of these groups of compounds. Glucocorticoids mainly find their way into surface waters
481 through industrial and hospital effluents. Glucocorticoid concentrations in such effluents are
482 high, but decrease substantially after wastewater treatment (Schriks et al., 2010b; Van Der
483 Linden et al., 2008). Hence, application of the GR CALUX assay in surface water monitoring
484 is only marginally relevant, as the risk of glucocorticoids in surface waters is expected to be
485 negligible. Therefore, this test can be omitted in future bioassay batteries to save costs, or
486 be replaced with a more relevant endpoint to surface water toxicity like the anti-PR CALUX
487 assay, for which a recently defined EBT value is available (Escher et al., 2018).

488 Contrasting to glucocorticoids, antibiotics are ubiquitous in NW European surface waters.
489 They reach surface waters through diffuse input from the general public and the agri-food
490 sector, giving rise to surface water concentrations that are expected to cause risks to
491 bacteria, fungi and microalgae (Hernando et al., 2006; Kümmerer, 2009; Zhou et al., 2019).
492 Hence, risks of antibiotics in the here tested surface waters is assumable, and the lack of
493 effects above the EBT for the antibiotics assays in the present study is therefore
494 unexpected. Given the low responsiveness of the here applied WaterSCAN antibiotics assay
495 (Pikkemaat et al., 2008) and the ubiquity of antibiotics in surface waters, there is evidently a
496 need for a more sensitive detection method for antibiotics residues. A potentially suitable
497 alternative is the use of whole-cell based biosensors, in which, similar to CALUX assays,
498 receptor binding mediated bioluminescence detects antibiotics activity at a sublethal level
499 (Virolainen and Karp, 2014). However, this method is yet to be applied as bioanalytical tool
500 in surface water antibiotics screening.

501 **4.3.4. Sediments and metals**

502 As of yet, successful effect-based monitoring efforts have been focused mainly on
503 pollution of surface waters by organic compounds (Altenburger et al., 2019; Hamers et al.,
504 2018; van der Oost et al., 2017b), while relatively little attention has been given to the
505 inclusion of sediments as a relevant source of impaired ecological surface water quality.

Sediments are the largest chemical repositories on earth where harmful compounds accumulate, thereby representing a significant threat to the health of aquatic ecosystems (Burton, 2013). Despite their relevant role in aquatic ecosystem health, sediments are often overlooked and understudied in regular water quality assessment strategies like the WFD (Borja et al., 2004).

Metal pollution is another relevant source of impaired ecological surface water quality that is currently largely overlooked in effect-based monitoring efforts. Metal pollution can have severe detrimental effects on water quality owing to its toxicity, frequency and abundance (Armitage et al., 2007; Sin et al., 2001). Only very rarely have the effects of metal pollution on aquatic ecosystems been studied using a combination of passive sampling and bioanalytical tools (Roig et al., 2011). Given the relevance of these pollution sources to aquatic ecosystem health, the development of integrative strategies that include the effect-based assessment of metal pollution as well as sediment quality would be a valuable addition to future research efforts.

4.4. Conclusions

Scientists and water authorities together are faced with the challenge of the increasing complexity of pollution in surface waters, and how to make the impact of this pollution on aquatic ecosystems measurable. Traditional chemical target analysis of a limited selection of pollutants has lost its relevance. Fortunately, the current availability and future development of a wide variety of alternative tools, in the form of effect-based methods, allows for a holistic interpretation of the harmful effects of all chemicals present in surface waters without individual identification of the causing compounds. It is likely that the debate on the most efficient and effective combination of effect-based tools in bioassay batteries, a conclusive approach to EBT derivation, as well as the regionwide implementation of the resulting monitoring strategies, will be ongoing for some time to come. Yet, at present, there is no

practical limitation to the application of effect-based water quality assessment methods in regular water quality monitoring at a region- or nationwide scale.

Acknowledgements

This research was part of the Smart Monitoring project (443.324), funded by the Foundation for applied water research (STOWA), The Netherlands.

References

Altenburger, R., Ait-Aissa, S., Antczak, P., Backhaus, T., Barceló, D., Seiler, T.-B., Brion, F., Busch, W., Chipman, K., de Alda, M.L., de Aragão Umbuzeiro, G., Escher, B.I., Falciani, F., Faust, M., Focks, A., Hilscherova, K., Hollender, J., Hollert, H., Jäger, F., Jahnke, A., Kortenkamp, A., Krauss, M., Lemkine, G.F., Munthe, J., Neumann, S., Schymanski, E.L., Scrimshaw, M., Segner, H., Slobodnik, J., Smedes, F., Kughathas, S., Teodorovic, I., Tindall, A.J., Tollefsen, K.E., Walz, K.-H., Williams, T.D., Van den Brink, P.J., van Gils, J., Vrana, B., Zhang, X., Brack, W., 2015. Future water quality monitoring — Adapting tools to deal with mixtures of pollutants in water resource management. Sci. Total Environ. 512, 540–551. https://doi.org/10.1016/j.scitotenv.2014.12.057

Altenburger, R., Brack, W., Burgess, R.M., Busch, W., Escher, B.I., Focks, A., Mark Hewitt, L., Jacobsen, B.N., de Alda, M.L., Ait-Aissa, S., Backhaus, T., Ginebreda, A., Hilscherová, K., Hollender, J., Hollert, H., Neale, P.A., Schulze, T., Schymanski, E.L., Teodorovic, I., Tindall, A.J., de Aragão Umbuzeiro, G., Vrana, B., Zonja, B., Krauss, M., 2019. Future water quality monitoring: improving the balance between exposure and toxicity assessments of real-world pollutant mixtures. Environ. Sci. Eur. 31, 12. https://doi.org/10.1186/s12302-019-0193-1

556 Armitage, P.D., Bowes, M.J., Vincent, H.M., 2007. Long-term changes in macroinvertebrate
557 communities of a heavy metal polluted stream: The river Nent (Cumbria, UK) after 28
558 years. River Res. Appl. 23, 997–1015. https://doi.org/10.1002/rra.1022

559 Blackwell, B.R., Ankley, G.T., Bradley, P.M., Houck, K.A., Makarov, S.S., Medvedev, A. V.,
560 Swintek, J., Villeneuve, D.L., 2019. Potential Toxicity of Complex Mixtures in Surface
561 Waters from a Nationwide Survey of United States Streams: Identifying in Vitro
562 Bioactivities and Causative Chemicals. Environ. Sci. Technol. 53, 973–983.
563 https://doi.org/10.1021/acs.est.8b05304

564 Booij, K., Smedes, F., 2010. An improved method for estimating in situ sampling rates of
565 nonpolar passive samplers. Environ. Sci. Technol. 44, 6789–6794.
566 https://doi.org/10.1021/es101321v

567 Booij, P., Sjollema, S.B., van der Geest, H.G., Leonards, P.E.G., Lamoree, M.H., de Voogt,
568 W.P., Admiraal, W., Laane, R.W.P.M., Vethaak, A.D., 2015. Toxic pressure of
569 herbicides on microalgae in Dutch estuarine and coastal waters. J. Sea Res. 102, 48–
570 56. https://doi.org/10.1016/j.seares.2015.05.001

571 Borja, A., Valencia, V., Franco, J., Muxika, I., Bald, J., Belzunce, M.J., Solaun, O., 2004. The
572 water framework directive: Water alone, or in association with sediment and biota, in
573 determining quality standards? Mar. Pollut. Bull.
574 https://doi.org/10.1016/j.marpolbul.2004.04.008

575 Brack, W., Ait-Aissa, S., Burgess, R.M., Busch, W., Creusot, N., Di Paolo, C., Escher, B.I.,
576 Mark Hewitt, L., Hilscherova, K., Hollender, J., Hollert, H., Jonker, W., Kool, J.,
577 Lamoree, M., Muschket, M., Neumann, S., Rostkowski, P., Ruttkies, C., Schollee, J.,
578 Schymanski, E.L., Schulze, T., Seiler, T.B., Tindall, A.J., De Aragão Umbuzeiro, G.,
579 Vrana, B., Krauss, M., 2016. Effect-directed analysis supporting monitoring of aquatic
580 environments - An in-depth overview. Sci. Total Environ. 544, 1073–1118.
581 https://doi.org/10.1016/j.scitotenv.2015.11.102

582 Brack, W., Dulio, V., Ågerstrand, M., Allan, I., Altenburger, R., Brinkmann, M., Bunke, D.,
583 Burgess, R.M., Cousins, I., Escher, B.I., Hernández, F.J., Hewitt, L.M., Hilscherová, K.,
584 Hollender, J., Hollert, H., Kase, R., Klauer, B., Lindim, C., Herráez, D.L., Miège, C.,
585 Munthe, J., O'Toole, S., Posthuma, L., Rüdel, H., Schäfer, R.B., Sengl, M., Smedes, F.,
586 van de Meent, D., van den Brink, P.J., van Gils, J., van Wezel, A.P., Vethaak, A.D.,
587 Vermeirssen, E., von der Ohe, P.C., Vrana, B., 2017. Towards the review of the
588 European Union Water Framework Directive: Recommendations for more efficient
589 assessment and management of chemical contamination in European surface water
590 resources. Sci. Total Environ. 576, 720–737.
591 https://doi.org/10.1016/j.scitotenv.2016.10.104

592 Brettschneider, D.J., Misovic, A., Schulte-Oehlmann, U., Oetken, M., Oehlmann, J., 2019.
593 Detection of chemically induced ecotoxicological effects in rivers of the Nidda
594 catchment (Hessen, Germany) and development of an ecotoxicological, Water
595 Framework Directive–compliant assessment system. Environ. Sci. Eur. 31, 7.
596 https://doi.org/10.1186/s12302-019-0190-4

597 Burton, G.A., 2013. Assessing sediment toxicity: Past, present, and future. Environ. Toxicol.
598 Chem. 32, 1438–1440. https://doi.org/10.1002/etc.2250

599 Busch, W., Schmidt, S., Kühne, R., Schulze, T., Krauss, M., Altenburger, R., 2016.
600 Micropollutants in European rivers: A mode of action survey to support the development
601 of effect-based tools for water monitoring. Environ. Toxicol. Chem. 35, 1887–1899.
602 https://doi.org/10.1002/etc.3460

603 Daughton, C.G., 2005. "Emerging" chemicals as pollutants in the environment: A 21st
604 century perspective. Renew. Resour. J. 23, 6–23.

605 de Baat, M.L., Bas, D.A., van Beusekom, S.A.M., Droge, S.T.J., van der Meer, F., de Vries,
606 M., Verdonschot, P.F.M., Kraak, M.H.S., 2018. Nationwide screening of surface water
607 toxicity to algae. Sci. Total Environ. 645, 780–787.

https://doi.org/10.1016/j.scitotenv.2018.07.214

Di Paolo, C., Ottermanns, R., Keiter, S., Ait-Aissa, S., Bluhm, K., Brack, W., Breitholtz, M., Buchinger, S., Carere, M., Chalon, C., Cousin, X., Dulio, V., Escher, B.I., Hamers, T., Hilscherová, K., Jarque, S., Jonas, A., Maillot-Marechal, E., Marneffe, Y., Nguyen, M.T., Pandard, P., Schifferli, A., Schulze, T., Seidensticker, S., Seiler, T.-B., Tang, J., van der Oost, R., Vermeirssen, E., Zounková, R., Zwart, N., Hollert, H., 2016. Bioassay battery interlaboratory investigation of emerging contaminants in spiked water extracts – Towards the implementation of bioanalytical monitoring tools in water quality assessment and monitoring. Water Res. 104, 473–484. https://doi.org/10.1016/j.watres.2016.08.018

Escher, B.I., Allinson, M., Altenburger, R., Bain, P.A., Balaguer, P., Busch, W., Crago, J., Denslow, N.D., Dopp, E., Hilscherova, K., Humpage, A.R., Kumar, A., Grimaldi, M., Jayasinghe, B.S., Jarosova, B., Jia, A., Makarov, S., Maruya, K.A., Medvedev, A., Mehinto, A.C., Mendez, J.E., Poulsen, A., Prochazka, E., Richard, J., Schifferli, A., Schlenk, D., Scholz, S., Shiraishi, F., Snyder, S., Su, G., Tang, J.Y.M., Burg, B. Van Der, Linden, S.C.V. Der, Werner, I., Westerheide, S.D., Wong, C.K.C., Yang, M., Yeung, B.H.Y., Zhang, X., Leusch, F.D.L., 2014. Benchmarking organic micropollutants in wastewater, recycled water and drinking water with in vitro bioassays. Environ. Sci. Technol. 48, 1940–1956. https://doi.org/10.1021/es403899t

Escher, B.I., Aït-Aïssa, S., Behnisch, P.A., Brack, W., Brion, F., Brouwer, A., Buchinger, S., Crawford, S.E., Du Pasquier, D., Hamers, T., Hettwer, K., Hilscherová, K., Hollert, H., Kase, R., Kienle, C., Tindall, A.J., Tuerk, J., van der Oost, R., Vermeirssen, E., Neale, P.A., 2018. Effect-based trigger values for in vitro and in vivo bioassays performed on surface water extracts supporting the environmental quality standards (EQS) of the European Water Framework Directive. Sci. Total Environ. 628–629, 748–765. https://doi.org/10.1016/j.scitotenv.2018.01.340

634 Fliedner, Lohmann, N., Rüdel, H., Teubner, D., Wellmitz, J., Koschorreck, J., 2016. Current
635 levels and trends of selected EU Water Framework Directive priority substances in
636 freshwater fish from the German environmental specimen bank. Environ. Pollut. 216,
637 866–876. https://doi.org/10.1016/j.envpol.2016.06.060

638 Hamers, T., Kamstra, J.H., Sonneveld, E., Murk, A.J., Kester, M.H.A., Andersson, P.L.,
639 Legler, J., Brouwer, A., 2006. In Vitro Profiling of the Endocrine-Disrupting Potency of
640 Brominated Flame Retardants. Toxicol. Sci. 92, 157–173.
641 https://doi.org/10.1093/toxsci/kfj187

642 Hamers, T., Legradi, J., Zwart, N., Smedes, F., de Weert, J., van den Brandhof, E.-J., van
643 de Meent, D., de Zwart, D., 2018. Time-Integrative Passive sampling combined with
644 TOxicity Profiling (TIPTOP): an effect-based strategy for cost-effective chemical water
645 quality assessment. Environ. Toxicol. Pharmacol. 64, 48–59.
646 https://doi.org/10.1016/J.ETAP.2018.09.005

647 Harman, C., Allan, I.J., Bäuerlein, P.S., 2011. The challenge of exposure correction for polar
648 passive samplers the PRC and the POCIS. Environ. Sci. Technol.
649 https://doi.org/10.1021/es2033789

650 Harman, C., Allan, I.J., Vermeirssen, E.L.M., 2012. Calibration and use of the polar organic
651 chemical integrative sampler-a critical review. Environ. Toxicol. Chem. 31, 2724–2738.
652 https://doi.org/10.1002/etc.2011

653 Hernando, M.D., Mezcua, M., Fernández-Alba, A.R., Barceló, D., 2006. Environmental risk
654 assessment of pharmaceutical residues in wastewater effluents, surface waters and
655 sediments. Talanta 69, 334–342. https://doi.org/10.1016/j.talanta.2005.09.037

656 Jia, A., Escher, B.I., Leusch, F.D.L., Tang, J.Y.M., Prochazka, E., Dong, B., Snyder, E.M.,
657 Snyder, S.A., 2015. In vitro bioassays to evaluate complex chemical mixtures in
658 recycled water. Water Res. 80, 1–11. https://doi.org/10.1016/j.watres.2015.05.020

659 Jones, L., Ronan, J., McHugh, B., McGovern, E., Regan, F., 2015. Emerging priority
660 substances in the aquatic environment: a role for passive sampling in supporting WFD
661 monitoring and compliance. Anal. Methods 7, 7976–7984.
662 https://doi.org/10.1039/c5ay01059d

663 Kulkarni, P.S., Crespo, J.G., Afonso, C.A.M., 2008. Dioxins sources and current remediation
664 technologies - A review. Environ. Int. 34, 139–153.
665 https://doi.org/10.1016/j.envint.2007.07.009

666 Kümmerer, K., 2009. Antibiotics in the aquatic environment - A review - Part I. Chemosphere
667 75, 417–434. https://doi.org/10.1016/j.chemosphere.2008.11.086

668 Leusch, F.D.L., Khan, S.J., Laingam, S., Prochazka, E., Froscio, S., Trinh, T., Chapman,
669 H.F., Humpage, A., 2014. Assessment of the application of bioanalytical tools as
670 surrogate measure of chemical contaminants in recycled water. Water Res. 49, 300–
671 315. https://doi.org/10.1016/j.watres.2013.11.030

672 Maltby, L., 1999. Studying Stress: The Importance of Organism-Level Responses. Ecol.
673 Appl. 9, 431. https://doi.org/10.2307/2641131

674 Manoli, E., Samara, C., 1999. Polycyclic aromatic hydrocarbons in natural waters: Sources,
675 occurrence and analysis. TrAC - Trends Anal. Chem. 18, 417–428.
676 https://doi.org/10.1016/S0165-9936(99)00111-9

677 Murk, A.J., Legler, J., Denison, M.S., Giesy, J.P., van de Guchte, C., Brouwer, A., 1996.
678 Chemical-Activated Luciferase Gene Expression (CALUX): A Novel in Vitro Bioassay
679 for Ah Receptor Active Compounds in Sediments and Pore Water. Fundam. Appl.
680 Toxicol. 33, 149–160. https://doi.org/10.1006/FAAT.1996.0152

681 Neale, P.A., Ait-Aissa, S., Brack, W., Creusot, N., Denison, M.S., Deutschmann, B.,
682 Hilscherová, K., Hollert, H., Krauss, M., Novák, J., Schulze, T., Seiler, T.B., Serra, H.,
683 Shao, Y., Escher, B.I., 2015. Linking in Vitro Effects and Detected Organic

684 Micropollutants in Surface Water Using Mixture-Toxicity Modeling. Environ. Sci.
685 Technol. 49, 14614–14624. https://doi.org/10.1021/acs.est.5b04083

686 Neale, P.A., Altenburger, R., Aït-Aïssa, S., Brion, F., Busch, W., de Aragão Umbuzeiro, G.,
687 Denison, M.S., Du Pasquier, D., Hilscherová, K., Hollert, H., Morales, D.A., Novák, J.,
688 Schlichting, R., Seiler, T.-B., Serra, H., Shao, Y., Tindall, A.J., Tollefsen, K.E., Williams,
689 T.D., Escher, B.I., 2017. Development of a bioanalytical test battery for water quality
690 monitoring: Fingerprinting identified micropollutants and their contribution to effects in
691 surface water. Water Res. 123, 734–750. https://doi.org/10.1016/j.watres.2017.07.016

692 Novák, J., Vrana, B., Rusina, T., Okonski, K., Grabic, R., Neale, P.A., Escher, B.I., Macová,
693 M., Ait-Aissa, S., Creusot, N., Allan, I., Hilscherová, K., 2018. Effect-based monitoring
694 of the Danube River using mobile passive sampling. Sci. Total Environ. 636, 1608–
695 1619. https://doi.org/10.1016/j.scitotenv.2018.02.201

696 OECD Environmental Health and Safety, 2006. OECD 201: Freshwater Alga and
697 Cyanobacteria, Growth Inhibition Test, in: OECD Guidelines for Testing of Chemicals.
698 https://doi.org/10.1787/9789264069923-en

699 OECD Environmental Health and Safety, 2004. OECD 202: Daphnia sp., Acute
700 Immobilisation Test, in: OECD Guidelines for Testing of Chemicals.
701 https://doi.org/10.1787/9789264069947-en

702 Peterson, H.G., Nyholm, N., Ruecker, N., 2005. Algal microplate toxicity test suitable for
703 heavy metals, in: Small-Scale Freshwater Toxicity Investigations: Volume 1 - Toxicity
704 Test Methods. Springer-Verlag, Berlin/Heidelberg, pp. 243–270.
705 https://doi.org/10.1007/1-4020-3120-3_7

706 Pieterse, B., Felzel, E., Winter, R., Van Der Burg, B., Brouwer, A., 2013. PAH-CALUX, an
707 optimized bioassay for AhR-mediated hazard identification of polycyclic aromatic
708 hydrocarbons (PAHs) as individual compounds and in complex mixtures. Environ. Sci.

709 Technol. 47, 11651–11659. https://doi.org/10.1021/es403810w

710 Pikkemaat, M.G., Dijk, S.O. v, Schouten, J., Rapallini, M., van Egmond, H.J., 2008. A new
711 microbial screening method for the detection of antimicrobial residues in slaughter
712 animals: The Nouws antibiotic test (NAT-screening). Food Control 19, 781–789.
713 https://doi.org/10.1016/j.foodcont.2007.08.002

714 Roig, N., Nadal, M., Sierra, J., Ginebreda, A., Schuhmacher, M., Domingo, J.L., 2011. Novel
715 approach for assessing heavy metal pollution and ecotoxicological status of rivers by
716 means of passive sampling methods. Environ. Int. 37, 671–677.
717 https://doi.org/10.1016/j.envint.2011.01.007

718 Roll, I.B., Halden, R.U., 2016. Critical review of factors governing data quality of integrative
719 samplers employed in environmental water monitoring. Water Res. 94, 200–207.
720 https://doi.org/10.1016/j.watres.2016.02.048

721 Schreiner, V.C., Szöcs, E., Bhowmik, A.K., Vijver, M.G., Schäfer, R.B., 2016. Pesticide
722 mixtures in streams of several European countries and the USA. Sci. Total Environ.
723 573, 680–689. https://doi.org/10.1016/j.scitotenv.2016.08.163

724 Schriks, M., Heringa, M.B., van der Kooi, M.M.E., de Voogt, P., van Wezel, A.P., 2010a.
725 Toxicological relevance of emerging contaminants for drinking water quality. Water
726 Res. 44, 461–476. https://doi.org/10.1016/j.watres.2009.08.023

727 Schriks, M., Van Leerdam, J.A., Van Der Linden, S.C., Van Der Burg, B., Van Wezel, A.P.,
728 De Voogt, P., 2010b. High-resolution mass spectrometric identification and
729 quantification of glucocorticoid compounds in various wastewaters in the Netherlands.
730 Environ. Sci. Technol. 44, 4766–4774. https://doi.org/10.1021/es100013x

731 Schwarzenbach, R.P., Escher, B.I., Fenner, K., Hofstetter, T.B., Johnson, C.A., von Gunten,
732 U., Wehrli, B., 2006. The challenge of micropollutants in aquatic systems. Science 313,
733 1072–7. https://doi.org/10.1126/science.1127291

734 Sin, S.N., Chua, H., Lo, W., Ng, L.M., 2001. Assessment of heavy metal cations in
735 sediments of Shing Mun River, Hong Kong. Environ. Int. 26, 297–301.
736 https://doi.org/10.1016/S0160-4120(01)00003-4

737 Sjollema, S.B., Martínezgarcía, G., Van Der Geest, H.G., Kraak, M.H.S., Booij, P., Vethaak,
738 A.D., Admiraal, W., 2014. Hazard and risk of herbicides for marine microalgae. Environ.
739 Pollut. 187, 106–111. https://doi.org/10.1016/j.envpol.2013.12.019

740 Sonneveld, E., Jansen, H.J., Riteco, J.A.C., Brouwer, A., van der Burg, B., 2004.
741 Development of Androgen- and Estrogen-Responsive Bioassays, Members of a Panel
742 of Human Cell Line-Based Highly Selective Steroid-Responsive Bioassays. Toxicol.
743 Sci. 83, 136–148. https://doi.org/10.1093/toxsci/kfi005

744 Tang, J.Y.M., McCarty, S., Glenn, E., Neale, P.A., Warne, M.S.J., Escher, B.I., 2013.
745 Mixture effects of organic micropollutants present in water: Towards the development of
746 effect-based water quality trigger values for baseline toxicity. Water Res. 47, 3300–
747 3314. https://doi.org/10.1016/j.watres.2013.03.011

748 Ter Braak, C.J.F., 1990. Update notes: CANOCO version 3.1., Agricultural Mathematics
749 Group, Wageningen. Agricultural Mathematics Group, Wageningen.

750 Ter Braak, C.J.F., 1988. CANOCO—A FORTRAN program for canonical community
751 ordination by [partial][detrended][canonical] correspondence analysis, principal
752 component analysis and redundancy analysis (version 2.1).

753 The European Parliament and the Council of the European Union, 2013. Directives of 12
754 August 2013 amending Directives 2000/60/EC and 2008/105/EC as regards priority
755 substances in the field of water policy, Official Journal of the European Union.
756 https://doi.org/http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32013L0039

757 Van Der Linden, S.C., Heringa, M.B., Man, H.Y., Sonneveld, E., Puijker, L.M., Brouwer, A.,
758 Van Der Burg, B., 2008. Detection of multiple hormonal activities in wastewater

759 effluents and surface water, using a panel of steroid receptor CALUX bioassays.
760 Environ. Sci. Technol. 42, 5814–5820. https://doi.org/10.1021/es702897y

761 Van der Linden, S.C., von Bergh, A.R.M., van Vught-Lussenburg, B.M.A., Jonker, L.R.A.,
762 Teunis, M., Krul, C.A.M., van der Burg, B., 2014. Development of a panel of high-
763 throughput reporter-gene assays to detect genotoxicity and oxidative stress. Mutat.
764 Res. - Genet. Toxicol. Environ. Mutagen. 760, 23–32.
765 https://doi.org/10.1016/j.mrgentox.2013.09.009

766 van der Oost, R., Sileno, G., Janse, T., Nguyen, M.T., Besselink, H., Brouwer, A., 2017a.
767 SIMONI (Smart Integrated Monitoring) as a novel bioanalytical strategy for water quality
768 assessment: Part II-field feasibility survey. Environ. Toxicol. Chem. 36, 2400–2416.
769 https://doi.org/10.1002/etc.3837

770 van der Oost, R., Sileno, G., Suárez-Muñoz, M., Nguyen, M.T., Besselink, H., Brouwer, A.,
771 2017b. SIMONI (Smart integrated monitoring) as a novel bioanalytical strategy for water
772 quality assessment: Part I–model design and effect-based trigger values. Environ.
773 Toxicol. Chem. 36, 2385–2399. https://doi.org/10.1002/etc.3836

774 Verdonschot, P.F.M., Ter Braak, C.J.F., 1994. An experimental manipulation of oligochaete
775 communities in mesocosms treated with chlorpyrifos or nutrient additions: multivariate
776 analyses with Monte Carlo permutation tests. Hydrobiologia 278, 251–266.
777 https://doi.org/10.1007/BF00142333

778 Virolainen, N., Karp, M., 2014. Biosensors, antibiotics and food, in: Bioluminescence:
779 Fundamentals and Applications in Biotechnology. Springer, Berlin, Heidelberg, pp.
780 153–185. https://doi.org/10.1007/978-3-662-43619-6_5

781 Vrana, B., Allan, I.J., Greenwood, R., Mills, G.A., Dominiak, E., Svensson, K., Knutsson, J.,
782 Morrison, G., 2005. Passive sampling techniques for monitoring pollutants in water.
783 TrAC Trends Anal. Chem. 24, 845–868. https://doi.org/10.1016/j.trac.2005.06.006

784 Wernersson, A.-S., Carere, M., Maggi, C., Tusil, P., Soldan, P., James, A., Sanchez, W.,
785 Dulio, V., Broeg, K., Reifferscheid, G., Buchinger, S., Maas, H., Van Der Grinten, E.,
786 O'Toole, S., Ausili, A., Manfra, L., Marziali, L., Polesello, S., Lacchetti, I., Mancini, L.,
787 Lilja, K., Linderoth, M., Lundeberg, T., Fjällborg, B., Porsbring, T., Larsson, D.J.,
788 Bengtsson-Palme, J., Förlin, L., Kienle, C., Kunz, P., Vermeirssen, E., Werner, I.,
789 Robinson, C.D., Lyons, B., Katsiadaki, I., Whalley, C., den Haan, K., Messiaen, M.,
790 Clayton, H., Lettieri, T., Carvalho, R.N., Gawlik, B.M., Hollert, H., Di Paolo, C., Brack,
791 W., Kammann, U., Kase, R., 2015. The European technical report on aquatic effect-
792 based monitoring tools under the water framework directive. Environ. Sci. Eur. 27, 7.
793 https://doi.org/10.1186/s12302-015-0039-4

794 Zhou, S., Di Paolo, C., Wu, X., Shao, Y., Seiler, T.-B., Hollert, H., 2019. Optimization of
795 screening-level risk assessment and priority selection of emerging pollutants – The
796 case of pharmaceuticals in European surface waters. Environ. Int. 128, 1–10.
797 https://doi.org/10.1016/j.envint.2019.04.034

Table 1 Bioassay battery applied to assess toxicity at 45 surface water locations in The Netherlands. Effect-based trigger values (EBT) were previously defined by Escher et al., 2018 (PAH, anti-AR and ER CALUX) and Van der Oost et al., 2017b.

	Bioassay	**Endpoint**	**Reference compound**	**EBT**	**Unit**
in situ	Daphnia in situ	Mortality	n/a	20	% mortality
in vivo non-polar	Daphniatox	Mortality	n/a	0.05	TU
	Algatox	Algal growth inhibition	n/a	0.05	TU
	Microtox	Luminescence inhibition	n/a	0.05	TU
in vitro CALUX non-polar	cytotox nonpolar	Cytotoxicity	n/a	0.05	TU
	DR	Dioxin (-like) activity	2,3,7,8-TCDD	50	pg TEQ/L
	PAH	PAH activity	benzo(a)pyrene	6.21	ng BEQ/L
	PPARγ	Lipid metabolism inhibition	rosiglitazone	10	ng REQ/L
	Nrf2	Oxidative stress	curcumin	10	µg CEQ/L
	PXR	Toxic compound metabolism	nicardipine	3	µg NEQ/L
	p53 -S9	Genotoxicity	actinomycin D	0.005	ng AEQ/L
	p53 +S9	Genotoxicity (after metabolism)	actinomycin D	0.005	µg CEQ/L
in vitro CALUX polar	cytotox polar	Cytotoxicity	n/a	0.05	TU
	ER	Estrogenic activity	17ß-estradiol	0.1	ng EEQ/L
	anti-AR	Antiandrogenic activity	flutamide	14.4	µg FEQ/L
	GR	Glucocorticoid activity	dexamethasone	100	ng DEQ/L
in vitro antibiotics polar	T	Bacterial growth inhibition (Tetracyclines)	oxytetracycline	250	ng OEQ/L
	Q	Bacterial growth inhibition (Quinolones)	flumequine	100	ng FEQ/L
	B+M	Bacterial growth inhibition (β-lactams and Macrolides)	penicillin G	50	ng PEQ/L
	S	Bacterial growth inhibition (Sulfonamides)	sulfamethoxazole	100	ng SEQ/L
	A	Bacterial growth inhibition (Aminoglycosides)	neomycin	500	ng NEQ/L

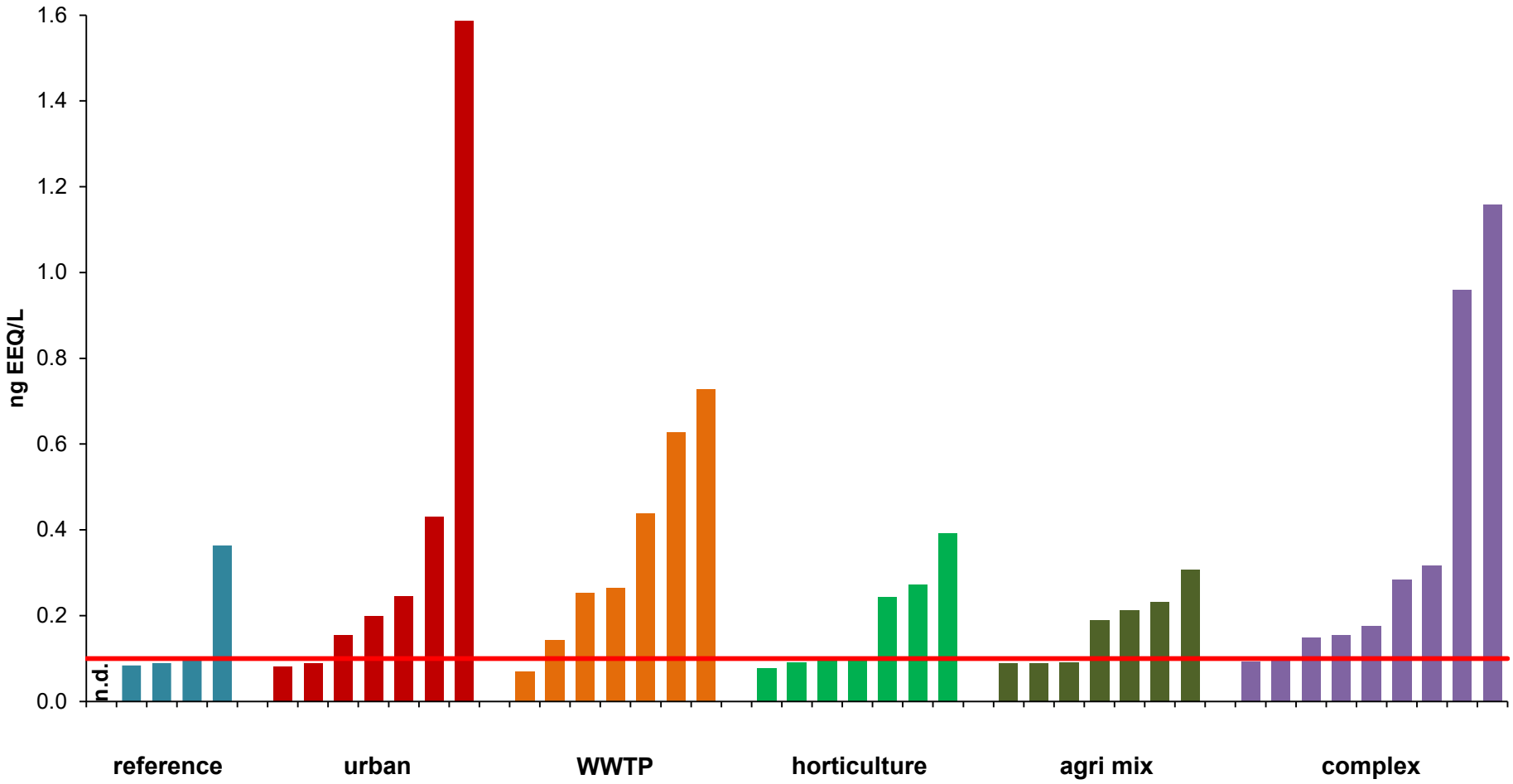


Figure 1 Estrogen receptor chemically activated luciferase expression (ER CALUX) assay responses to POCIS passive sampler extracts from 42 surface water locations with different surrounding land uses expressed as 17β-estradiol equivalents (ng EEQ/L). The red line indicates the effect-based trigger value (0.1 ng EEQ/L). WWTP = wastewater treatment plant and agri mix = mixed agriculture, n.d. = no detected bioanalytical response. **PRINT IN COLOR**

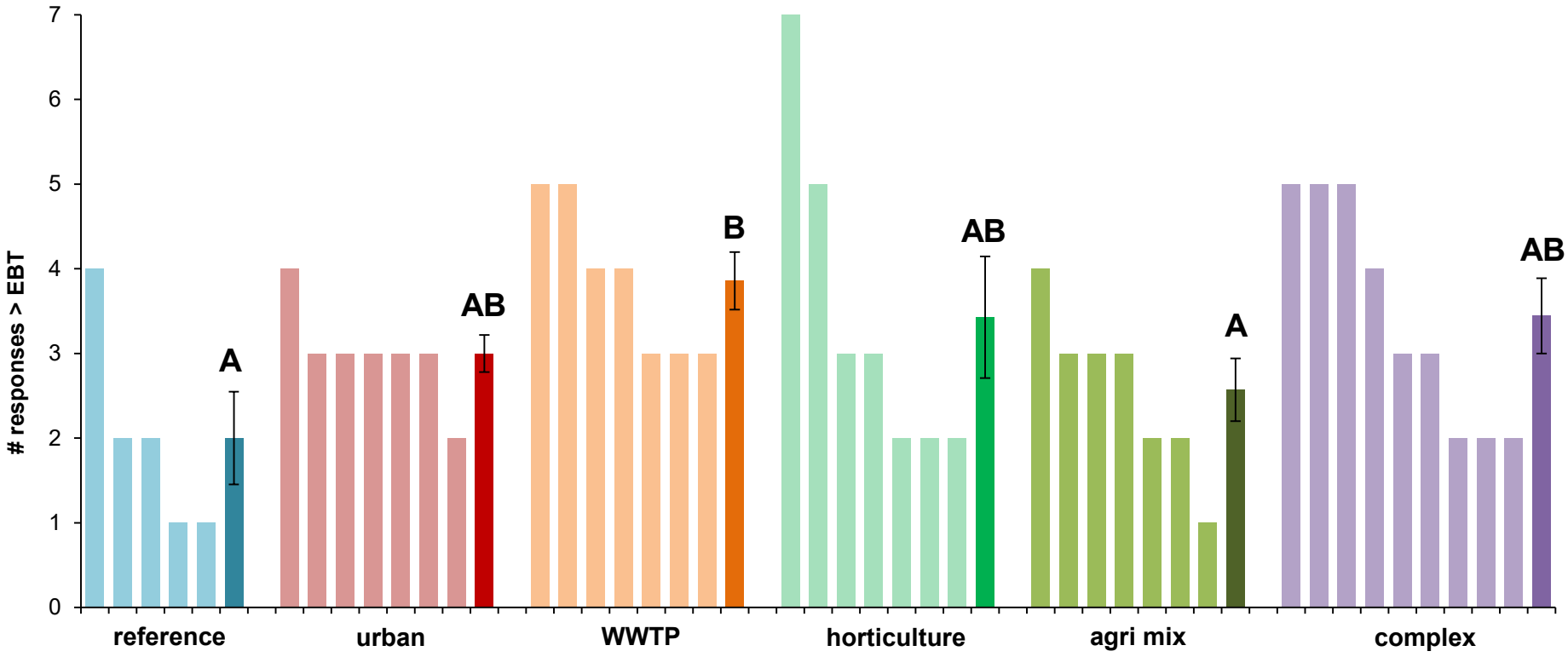


Figure 2 Number of effect-based trigger value (EBT) exceedances per location (light bars) and average number of exceedances per land use category (dark bars, ±SE) of a panel of 21 bioassays at 42 surface water locations grouped by surrounding land use. WWTP = wastewater treatment plant and agri mix = mixed agriculture. Statistical differences between land use averages are indicated with letters (Two-sample T-test assuming equal variances, α = 0.05) **PRINT IN COLOR**

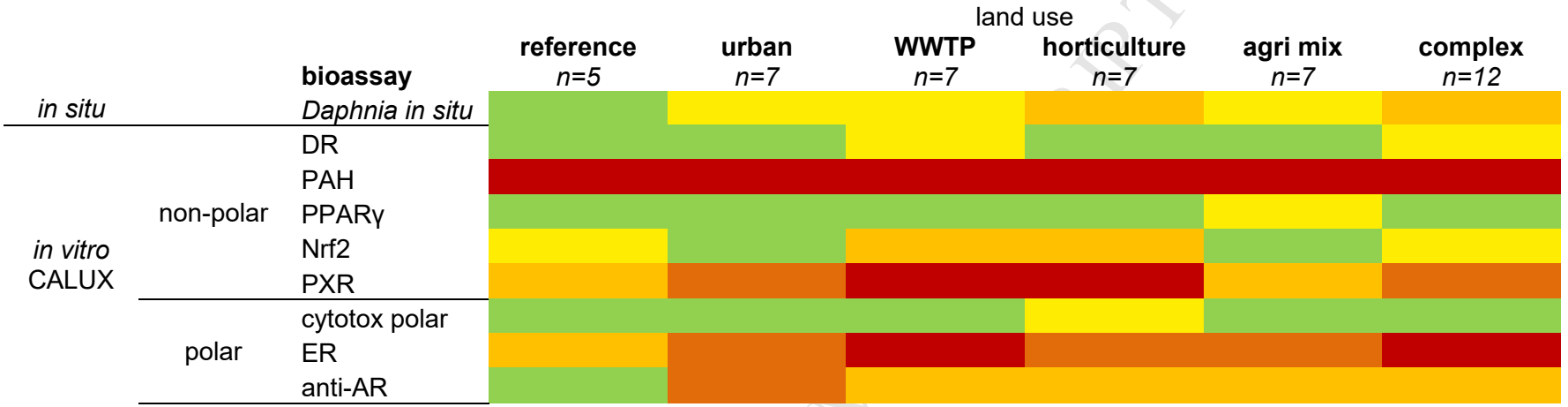


Figure 3 Heat map depicting responses in the 9 bioassays that gave a signal above the EBT for 45 surface water locations with different surrounding land uses. Colours indicate the percentage of the investigated locations with EBT exceedance per land use category: green = 0%, yellow = 0-25%, light orange = 25-50%, dark orange = 50-75% and red = 75-100% of locations. Deviating number of samples per land use for *Daphnia in situ*: reference = 4, urban = 5, WWTP = 6, horticulture = 2, agri mix = 6 and complex = 10. For complex land use, n = 9 for polar CALUX assays. **PRINT IN COLOR**

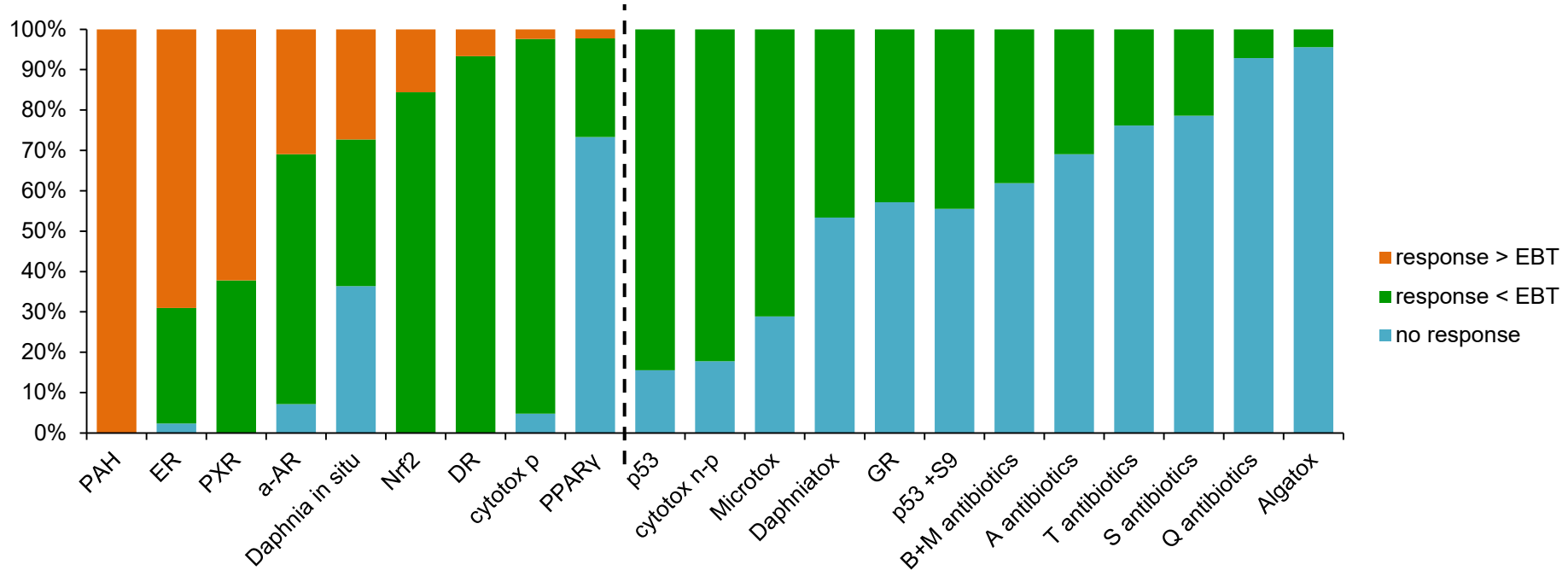


Figure 4 Frequency of responses of a panel of 21 bioassays to passive sampler extracts from 45 surface water locations. Colours indicate the bioassay responses and EBT exceedances at the percentage of study locations. The dashed line indicates the division between bioassays with and without EBT exceedance in the present study. **PRINT IN COLOR**

Highlights

- Responses were observed in all 21 bioassays in a nationwide campaign
- Effect-based trigger values were exceeded for 9 out of 21 bioassays
- Effect-based assessment allowed prioritization based on ecotoxicological risks
- Improvements of effect-based monitoring are proposed

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: