Radboud Repository

Radboud University Nijmegen

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/199362

Please be advised that this information was generated on 2019-06-02 and may be subject to change.

Environmental Toxicology and Pharmacology 64 (2018) 48–59

Contents lists available at ScienceDirect

Environmental Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/etap

Time-Integrative Passive sampling combined with TOxicity Profiling (TIPTOP): an effect-based strategy for cost-effective chemical water quality assessment

Timo Hamers[a,*], Jessica Legradi[a], Nick Zwart[a], Foppe Smedes[b,c], Jasperien de Weert[b], Evert-Jan van den Brandhof[d], Dik van de Meent[d,e], Dick de Zwart[d,f]

[a] *Department Environment & Health (E&H), Vrije Universiteit Amsterdam, The Netherlands*
[b] *Deltares, Utrecht, The Netherlands*
[c] *Recetox – Masaryk University, Brno, Czech Republic*
[d] *National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands*
[e] *Radboud University Nijmegen, The Netherlands*
[f] *DdZ Ecotox, Odijk, The Netherlands*

ARTICLE INFO

Keywords:
Speedisk
silicone rubber
in vitro bioassay
in vivo bioassay
species-sensitivity distribution (SSD)
multiple substance potentially affected fraction of species (msPAF)
priority compounds
River Meuse
River Rhine
wastewater treatment plant (WWTP)
toxic pressure

ABSTRACT

This study aimed at demonstrating that effect-based monitoring with passive sampling followed by toxicity profiling is more protective and cost-effective than the current chemical water quality assessment strategy consisting of compound-by-compound chemical analysis of selected substances in grab samples. Passive samplers were deployed in the Dutch river delta and in WWTP effluents. Their extracts were tested in a battery of bioassays and chemically analyzed to obtain toxicity and chemical profiles, respectively. Chemical concentrations in water were retrieved from publicly available databases. Seven different strategies were used to interpret the chemical and toxicity profiles in terms of ecological risk. They all indicated that the river sampling locations were relatively clean. Chemical-based monitoring resulted for many substances in measurements below detection limit and could only explain $<$ 20% of the observed *in vitro* toxicity. Effect-based monitoring yielded more informative conclusions as it allowed for ranking the sampling sites and for estimating a margin-of-exposure towards chronic effect ranges. Effect-based monitoring was also cheaper and more cost-effective (i.e. yielding more information per euro spent). Based on its identified strengths, weaknesses, opportunities, and threats (SWOT), a future strategy for effect-based monitoring has been proposed.

1. Introduction

Within the European Water Framework Directive (WFD) legislation, “good surface water chemical status” is defined as a status in which concentrations of pollutants do not exceed predefined environmental quality standards (EQS) (European Commission, 2008, 2000). For surveillance monitoring purposes, EQS values have been defined for a set for 45 priority substances defined within the WFD and another eight substances for which legislation was in place before the WFD was introduced in 2000 (European Commission, 2013). However, a chemical water quality assessment based on compound-by-compound comparisons between the concentration of an individual pollutant and its EQS ignores (1) the contribution by non-analyzed and often unknown substances to the toxic potency present in the water, and (2) the combined effects of the different chemicals locally present in the environmental cocktail of contaminants. In addition, the time-point grab sampling method as currently applied within the WFD surveillance monitoring may miss incidental peak concentrations.

The general concept for ecological risk assessment is based on a three-way (TRIAD) approach, consisting of three lines of evidence, i.e. chemistry, ecology, and toxicology (Chapman et al., 1997). Within the WFD, however, only two lines of evidence are covered by the surveillance monitoring programs, i.e. chemistry and ecology, whereas the toxicity line of evidence is not represented. The TIPTOP (**T**ime-**I**ntegrative **P**assive sampling combined with **TO**xicity **P**rofiling) study described in this paper focuses particularly on this latter line of evidence, by exploring an alternative strategy for ecological risk assessment purposes.

The TIPTOP strategy consists of applying a combination of time-integrative passive sampling followed by toxicity profiling to

* Corresponding author at: Vrije Universiteit Amsterdam, Dept. Environment & Health, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands.
E-mail address: timo.hamers@vu.nl (T. Hamers).

https://doi.org/10.1016/j.etap.2018.09.005
Received 22 March 2018; Received in revised form 27 August 2018; Accepted 14 September 2018
Available online 20 September 2018
1382-6689/ © 2018 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/BY/4.0/).

demonstrate compliance with regulatory requirements for "good ecological water quality”. Throughout the manuscript, the terms passive sampling refers to a sampling technique based on free flow of a substance from the water to a receiving sampler due to differences between chemical potentials of the substance in water and the sampler (Vrana et al., 2005). Similarly, toxicity profiling refers to the testing of the sampler extract for its in vitro or in vivo activity towards a battery of biological endpoints resulting in a toxicological ‘fingerprint’ of the complex mixture of substances present in the sampler (Hamers et al., 2013). Although a combined strategy of passive sampling followed by effect-based bioassay measurements has been proposed in the past (European Commission, 2014; Hamers et al., 2013, 2010; Sabaliunas et al., 1998), the combination of both techniques has not been sufficiently scrutinized yet to serve as a serious alternative to the costly chemical monitoring that is currently required by the WFD.

The goal of the TIPTOP study described in the present paper was to demonstrate that an effect-based monitoring strategy using passive sampling followed by toxicity profiling is a practical method to determine surface quality chemical status. On beforehand, we hypothesized that this strategy is more cost-effective and leads to a more protective risk assessment than the current EU WFD strategy based on the chemical analysis of a continuously expanding suite of individual substances.

Within the TIPTOP study, we assessed the integrated toxic potency of passively sampled complex mixtures from well-studied surface water sampling sites, making use of state-of-the-art knowledge and technologies. To guarantee collection of a wide range of compounds, passive sampling was performed applying two types of samplers, i.e. partition-based passive samplers and adsorption-based passive samplers. Toxicity profiling was performed by testing the passive sampler extracts in a test battery consisting of *in vitro* and *in vivo* bioassays. To allow comparison to current water quality assessment, passive sampler extracts were also chemically analyzed for a selected set of target compounds and for the total molar sum of accumulated compounds. In addition, water concentrations at the sampling sites were retrieved from publicly available databases and reports.

The obtained chemical and toxicological profiles of the passive sampler extracts (presented in the Supplementary Material section S2) were interpreted in seven different ways to make a chemical water quality assessment. In general, the bioassays demonstrated low toxicity in the passive sampler extracts, which corresponded to low pollutant levels determined by chemical analysis. Consequently, differences in protectiveness between effect-based and chemical-based risk assessment were not demonstrated. Nevertheless, the TIPTOP study pointed out that effect-based risk assessment yielded a margin-of-exposure to concentration levels where effects were to be expected in the field. Therefore, it was more informative than chemical-based risk assessment, based on many measurements below limit of detection (LOD). In addition, effect-based risk assessment based on passive sampling followed by toxicity profiling was more cost-effective, not only because it yielded more informative conclusions, but also because cost estimates are lower due to the fact that less samples are required when monitoring time-integrative samples rather than grab samples. Finally, strengths, weaknesses, opportunities, and threats for the combined approach of passive sampling followed by toxicity profiling are discussed in a SWOT analysis, and a strategy has been proposed for applying this approach in future chemical risk assessment.

2. Methods

2.1. Sampling locations

Passive samplers were deployed at six river sites in the Dutch delta, which were well-characterized as WFD surveillance monitoring locations, and in the effluent stream of two urban wastewater treatment plants (WWTPs). The two WWTPs are monitored in four-year cycles for priority pollutant emissions within the framework of the European Pollutant Release Transfer Register (E-PRTR). Four-weekly grab samples from the six WFD sites are analyzed in monitoring programs dating back in some cases to 1977 by Rijkswaterstaat (RWS), a part of the Dutch ministry of Infrastructure and Water Management. The selected WFD sites are mainly located in a gradient of the River Meuse from the Belgian border to the Dutch Delta (see Figure S1). Names and abbreviating codes of the eight sampling locations together with their corresponding WFD water body monitoring framework is provided in Table 1.

2.2. Passive sampling

Silicone rubber sheets were chosen as partition-based passive samplers based on their superior permeability (Rusina et al., 2010, 2007) and robustness as experienced by several research groups (Emelogu et al., 2013; Smedes et al., 2007). Two silicone rubber samplers were deployed at each sampling site, i.e. one sampler for toxicity profiling and one sampler for chemical analysis. Prior to deployment, the silicone rubber sampler meant for chemical analysis was spiked with performance reference compounds (PRCs), which allowed assessment of the sampling volume of the sampler (see below). Silicone rubber sheets intended for toxicity profiling were not spiked with PRCs. Speedisks containing styrene divinylbenzene sorbent were chosen as adsorption-

Table 1
Overview of the eight TIPTOP sampling locations.

Location	Code	WFD water body/ WWTP	Monitoring framework	Comments
Eijsden	EIJSDPTN	Bovenmaas (Upper Meuse)	WFD	River Meuse enters The Netherlands Very well studied site
Keizersveer	KEIZVR	Bergse Maas	WFD	River Meuse, downstream from Eijsden
Bovensluis	BOVSS	Haringvliet-East	WFD	River Meuse mixed with River Rhine Beginning of the Dutch Delta – Hollands Diep
Steenbergen	STEENBGN	Volkerak	WFD	Dutch delta, downstream from Hollands Diep
Haringvlietsluis	HARVSS	Haringvliet-West	WFD	Dutch delta, downstream from Volkerak
Lobith	LOBPTN	Bovenrijn (Upper Rhine)	WFD	River Rhine enters The Netherlands Very well studied site
Kralingseveer	KRALSVR	WWTP	E-PRTR	Urban WWTP 10% industrial wastewater in influent
Amersfoort	AMFT	WWTP	E-PRTR	Urban WWTP 25% industrial wastewater in influent

based passive samplers for the uptake of more polar compounds, as pesticides, herbicides and fungicides, which are not sampled with partition-based materials with high affinity for hydrophobic compounds (Emelogu et al., 2013). Speedisks are robust samplers having the advantage over other absorption-based samplers that sorbent material is held by glass fiber filters without sorption (Smedes et al., 2011). Speedisks were not spiked with PRCs for reasons explained in 2.3. Technical details about preparation, deployment, and processing of the samplers are described in the Supplemental Material (sections S1.2 and S1.3).

2.3. Chemical analysis

Extracts from passive samplers were chemically analyzed by GC-MS (MS) or LC-MS for a selected set of target compounds as well as the residual amounts of PRCs present in the silicone rubber extracts. In addition, the total molar sum of chemical substances accumulated in/on the silicone rubber and Speedisk samplers was quantified using vapor pressure osmometry, as described by Verhaar et al. (1995). Further details about sampler extraction and analytical methods are provided in the Supplementary Material (sections S1.3, S1.4 and S1.7).

Substance concentrations in the silicone rubber sampler were transformed into concentrations in the aqueous phase based on the sampling rate of the passive sampler (Rusina et al., 2010) and the retained fraction of PRCs detected on the sampler after deployment. The underlying method for this transformation (Booij and Smedes, 2010) has been described in detail in the Supplementary Material (section S1.5). Speedisk samplers were not spiked with PRCs because the adsorption samplers do not release compounds to the water phase, given their linear uptake characteristics in time. Therefore, Speedisk concentrations were transformed into aqueous phase concentration using Speedisk sampling rates based on those compounds that were detected both in the silicone rubber and the Speedisk samplers (see Supplementary Material section S1.5).

For the WFD locations, additional water concentrations of 205 target-analyzed organic contaminants determined from 1977 to 2014 (including the TIPTOP sampling period) were retrieved from the publicly available DONAR database (http://live.waterbase.nl/) by making a request to Helpdesk Water. For the two WWTP treatment plants, effluent concentrations of 318 organic contaminants were retrieved from publicly available reports (Baltussen, 2015, 2013, 2010) on the E-PRTR monitoring programs held in 2015, 2011, and 2007.

2.4. Toxicity Profiling

Extracts from passive samplers were tested in a battery consisting of five *in vitro* bioassays representing different mechanisms of action and seven small volume *in vivo* bioassays with species representing different trophic levels (Table 2). Further description of the methods with corresponding protocols and references is provided in the Supplementary Material.

In vitro bioassay responses were expressed as equivalent concentrations in the passive samplers of a reference compound typical for the observed specific response, e.g. pg estradiol-equivalents per gram silicone rubber or per Speedisk sampler for estrogenic compounds. *In vivo* bioassay responses were expressed as EC_{50} values in the test, i.e. gram silicone rubber or number or Speedisk samplers per liter test medium. To transform a measured bioassay response towards a passive sampler extract into a bioassay response towards the corresponding water phase, the measured bioassay response was corrected for - i.e. divided by (*in vitro*) or multiplied with (*in vivo*) - its sampled water volume (V) expressed in liter water per gram silicone rubber or per Speedisk sampler. For this transformation we need a single estimate per sample of the sampled volume of the original water. For the partitioning-based silicone rubber, however, it is in principle impossible to determine a single V value for all substances (and their corresponding combined toxicity) present in a silicone rubber passive sampler extract. In fact, V is a substance-specific measure as it depends on the partitioning of the compound between the silicone rubber and the water phase. To acknowledge differences in V values for different substances present in the silicone rubber sampler, a concentration-weighted average (V_{cwa}) value was determined to calculate bioassay responses to passive sampler extracts into bioassay responses to the corresponding water phase. V_{cwa} values were determined by weighing all V values for the individual compounds detected in the silicone rubber sampler by their concentration in the sampler relative to the total concentration of all detected compounds in the sampler. For the Speedisk sampler, the sampled water volume was considered to be equal for all different substances, because the sampler is supposed to act as in infinite sink, i.e. uptake is linear in time and the sampler does not reach equilibrium.

Table 2
Battery of *in vitro* and *in vivo* bioassays used for TIPTOP toxicity profiling.

Name	*In vitro/vivo*	Specific/General	Endpoint
DR-LUC (H4L1.1c4)	*In vitro*	Specific	Dioxinlike or PAH-like activity through arylhydrocarbon receptor (AhR) activation[a]
ER-LUC (VM7Luc4E2)	*In vitro*	Specific	Estrogenic activity through estrogen receptor (ER) activation
AR-EcoScreen	*In vitro*	Specific	Androgenic activity through androgen receptor (AR) activation Anti-androgenic activity through AR inactivation in the presence of 5-alpha-dihydrotestosterone (DHT)
TTR-binding	*In vitro*	Specific	Displacement of thyroid hormone precursor thyroxine (T4) from its plasma transport protein transthyretin (TTR)
Ames II	*In vitro*[b]	Specific	Mutagenic activity in TA98 strain with and without metabolic activation
A. fischeri[c]	*In vivo*	General	Reduced bioluminescence of *Allivibrio fischeri* bacteria indicative for decreased respiration
Microtox	*In vivo*	General	Reduced bioluminescence of *Allivibrio fischeri* bacteria indicative for decreased respiration
Algae PAM	*In vivo*	General	Reduced photosynthetic efficiency of green alga *Pseudokirchneriella subcapitata*
Thamnotoxkit F™	*In vivo*	General	Juvenile mortality (24 h) of crustacean *Thamnocephalus platyurus*
Daphnia magna acute toxicity test	*In vivo*	General	Juvenile immobility (24/48 h) of crustacean *Daphnia magna*
Rotoxkit F™	*In vivo*	General	Juvenile mortality (24 h) of rotifer *Brachionus calyciflorus*
QFET	*In vivo*	General	Zebrafish *Danio rerio* embryo toxicity (24/48/72/96/120 h)

[a] The type of response measured by the DR-LUC assay depends on the sample clean-up, i.e. with (dioxinlike) or without (PAH-like) sulfuric acid-silica clean-up.
[b] The Ames II assay is strictly an *in vivo* bioassay because it makes use of intact bacteria. Within TIPTOP, however, it is further considered as an *in vitro* bioassay, based on its small volume performance and its more specific endpoint (i.e. mutagenicity) compared to the general endpoints survival, growth, or reproduction determined in other *in vivo* bioassays.
[c] Although the *A. fischeri* and Microtox bioassay measure the same endpoint, they were regarded as two separate bioassays, since they used different protocols and were performed by different project partners.

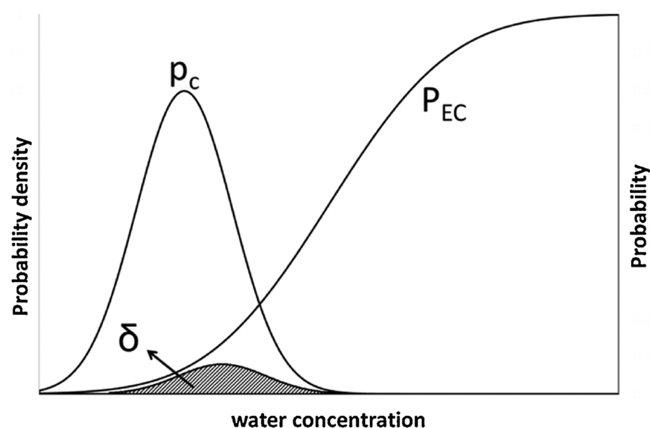


Fig. 1. Derivation of toxic pressure according to the Van Straalen-Aldenberg convolution (Aldenberg et al., 2002; Van Straalen, 2002). The toxic pressure is the product of the probability that a concentration value occurs in a water system (p_c) and the frequency by which a water concentration is higher than a critical effect concentration of an aquatic species (P_{EC}). Toxic pressure (i.e. the probability δ of having concentrations that exceed a critical effect concentrations) is obtained by integration of this probability product over all possible values of the water concentrations, which is represented by the shaded area. The cumulative distribution of critical effect concentrations (P_{EC}) is the same as the species sensitivity distribution (SSD).

Consequently all compounds have the same V value (and thereby the same V_{cwa} value) which is further used to calculate observed toxic potency in the passive sampler into waterborne toxic potency. For a more detailed description of V calculations we refer to the Supplementary Material (sections S1.5 and S3.1), where V values for Speedisk samplers are provided in Table S1, and V_{cwa} values for silicone rubber samplers in Table S7.

2.5. Toxic pressure calculations based on water concentrations

Toxic pressure has been defined here as the probability that a concentration in the field exceeds the critical effect concentration of a species as determined in a laboratory (e.g. acute EC50, chronic NOEC) (Van Straalen, 2002). As both the actual water concentration and the critical effect concentration are part of a statistical distribution (Fig. 1), the ultimate toxic pressure is a product of two probabilities, i.e. the probability for a water concentration to actually occur and the frequency by which that same water concentration exceeds the critical effect concentration for any aquatic species (i.e. a species sensitivity distributions or SSD). The resulting toxic pressure can be considered as the fraction of species expected to be potentially affected by that environmental concentration (Aldenberg et al., 2002).

Toxic pressure calculations were based on sample-specific concentrations of individual compounds in the water phase and on compound-specific *in vivo* aquatic toxicity data determined in laboratory experiments. The integrated toxic pressure of the complex mixture of substances in the field was expressed as the multiple substance Potentially Affected Fraction of species (msPAF), which was calculated according to the mixed model approach (De Zwart and Posthuma, 2005). In this approach, mixture toxic pressure of substances with the same Toxic Mode of Action (TMoA) was calculated by means of concentration addition, followed by aggregation across modes of action by means of response addition (Bliss, 1939). *In vivo* laboratory toxicity data were retrieved from a database that was originally described by De Zwart (2002) and has further been extended within the scope of three recent, closely related projects (ESF8, SOLUTIONS and eco-TTC). This database currently comprise 93197 records, covering 4411 substances, 2207 species, 24462 chronic NOEC values, and 46500 acute EC50 values (STOWA, 2016).

2.6. Interpretation strategies

Based on the chemical and toxicity profiles obtained within TIPTOP, seven different strategies for water quality assessment were applied:

1. **Benchmarking of *in vitro* and *in vivo* toxicity profiles to WWTP effluents.** Similarity and dissimilarity in waterborne toxicity profiles (i.e. after correction for V_{cwa}) obtained for the eight sampling sites were evaluated by a multivariate hierarchical clustering method on z-score normalized bioassay results (between group linkage based on squared Eucledian distances). The classified toxicity profiles were used to benchmark the different river sampling sites against the expectedly more polluted WWTP sampling sites.
2. **WFD-like comparison between concentrations in the water and environmental quality standards.** The classical approach is used of comparing concentrations of target-analyzed substances compound-by-compound to critical threshold values. For this purpose, water concentrations collected for the river sites and the WWTP sites from publicly available databases and reports (see above) were straightforwardly compared to the Annual Average Environmental Quality Standard (AA-EQS) values defined within the WFD (European Commission, 2013).
3. **Comparison of *in vitro* toxic potencies to mechanism-specific trigger values.** *In vitro* bioassay responses are indicative for the presence of compounds with a specific mechanism of action, but are difficult to interpret in terms of ecological risk, because this requires an additional *in vitro* to *in vivo* extrapolation step. To enable the use of *in vitro* bioassays for ecological risk assessment, the observed *in vitro* toxic potencies of the passive sampler extracts were compared to preliminary trigger values that were very recently proposed (Escher et al., 2018; Van der Oost et al., 2017). Trigger values correspond to bioassay response levels at which chemicals are not expected to cause adverse effects at higher levels of biological organization. Derivation of such trigger values is a very topical research theme, and is still under debate.
4. **Calculation of toxic pressure based on concentrations in the water phase**. Water concentrations determined in water grab samples collected at the sampling sites were calculated into toxic pressure estimates using the mixed model approach (De Zwart and Posthuma, 2005). Alternatively, the concentrations determined in the silicone rubber and Speedisk samplers (expressed as μg/g SR or μg/SD, respectively) were used to calculate time-averaged concentrations in the passively sampled water phase. This was done by multiplying the concentrations in the passive samplers by the corresponding sampled volumes (V, expressed in L/g SR or L/SD, respectively), as derived in section S1.5 (Derivation of aqueous concentrations). For each compound i collected in the partitioning-based silicone rubber samplers, V_i depends on site-specific turbulence, as well as compound$_i$ specific absorption rates. The sampled water volumes for Speedisk adsorption samplers are only depending on site specific turbulence. The obtained time-weighted water concentrations were calculated into an overall toxic pressure estimate similarly as described above for the grab samples.
5. **Calculation of toxic pressure and HC5 based on the *in vivo* toxicity profile.** For each sample, the seven acute EC50 values (expressed as g SR/L or SD/L in the test medium) determined for the *in vivo* bioassays (Table 2) were multiplied by the concentration weighted sample volume V_{cwa} (see 2.4) to obtain the median effective concentration factor ($EC50_f$). Subsequently, these seven $EC50_f$ values were used to derive a sample-specific SSD based on

critical concentration factors of that particular water sample. From this SSD, the sample-specific toxic pressure was estimated as the msPAF value corresponding to a concentration factor of 1 (i.e. non-concentrated water). Alternatively, the SSD was used to determine the concentration factor corresponding to msPAF = 5%. Since the 5th percentile of a substance-specific SSD is generally accepted as the hazardous concentrations for that particular substance (European Commission, 2003), the obtained Hazardous Concentration factor 5% (HC_f5) value can be regarded as the margin of exposure of the toxic pressure in the field towards an accepted tolerance limit.

6 **Calculation of minimum toxic pressure based on molar concentrations in the passive sampler extracts.** Apart from specific TMoA that compounds have or have not, all organic compounds have a minimum toxicity, also known as narcosis or baseline toxicity. This is considered as a result of disturbance of biological membrane integrity by the partitioning of non-specifically acting neutral organics into membrane lipids (Escher and Hermens, 2002). Effects are reversible and occur if a certain threshold concentration level in the target body tissue is exceeded, i.e. the critical body burden (CBB). The total molar concentration in the passive sampler, as determined by the vapor pressure osmometry method, is considered to be dominated by compounds with a non-specific mode of action. Therefore, this total molar concentration was interpolated in SSDs of critical body burdens reported for narcotic compounds in the literature (Kipka and Di Toro, 2009; Verhaar et al., 1995). The resulting msPAF is regarded as an estimate of the baseline toxic pressure by narcotic substances.

7 **Calculation of TMoA-specific toxic pressure based on the*in vitro*toxicity profile.** The *in vitro* bioassay results were further used to determine the toxic pressure that might be attributed to a specific TMoA. For this purpose, a selection of chronic *in vivo* NOEC values was made for those compounds that were known to act through the same specific TMoA as determined in the *in vitro* bioassays. For each compound i, NOEC values were transformed into bioanalytical-equivalent (BEQ) concentrations of a reference compound indicative for the TMoA ($NOEC_{BEQ}$). This was done according to $NOEC_{BEQ} = NOEC_i \times REP_i$, with REP_i being the relative potency factor of compound i in the *in vitro* bioassay relative to this reference compound, i.e. $REP_i = EC50_{reference\ compound}/EC50_i$. EC50 values were retrieved from the ToxCast database (https://www.epa.gov/chemical-research/toxicity-forecasting). For each species, the obtained $NOEC_{BEQ}$ values for the different compounds were averaged in order to construct TMoA-specific SSDs. Finally, the TMoA-specific potency of the passive sampler extract measured *in vitro* bioassays (expressed as μg BEQ/g SR or μg BEQ/SD) was multiplied with the concentration weighted sample volume V_{cwa} (see 2.4) to obtain a waterborne TMoA specific potency (μg BEQ/L). This was interpolated into the TMoA-specific SSD. The obtained msPAF value was interpreted as the toxic pressure caused by that specific TMoA.

3. Results and Discussion

This chapter describes the outcomes of the seven different interpretation strategies in separate sections. The underlying bioassay results, i.e. equivalent concentrations of reference compound per gram silicone rubber or per Speedisk column (*in vitro*), or EC50 values expressed as gram silicone rubber or number or Speedisk samplers per liter test medium (*in vivo*), are not presented here, but are provided in the Supporting Information (section S3.2; Table S8).

3.1. Benchmarking toxicity profiles at river sites to WWTP sites

Benchmarking of the combined *in vivo* and *in vitro* toxicity profiles observed in river sites with WWTP toxicity profiles indicated similarity between the toxicity profiles from river sites Eijsden, Bovensluis, Haringvlietsluis, and Lobith (Fig. 2). The clustering analysis indicated that Speedisk and silicone rubber samplers from river site Keizersveer yielded higher toxicity profiles that differed from the other river sampling sites, similarly as observed for the WWTP samplers. Moreover, the two WWTP samples did not cluster together, indicating differences in toxicity profiles, probably due to differences in influent origin, i.e. with substantial (WWTP Amersfoort) or small (WWTP Kralingseveer) industrial contribution. Finally, higher toxicity was observed for the Speedisk sampler extracts collecting relatively polar substances than for the silicone rubber samplers collecting more lipophilic substances.

3.2. Comparison of target-analyzed concentrations to EQS values

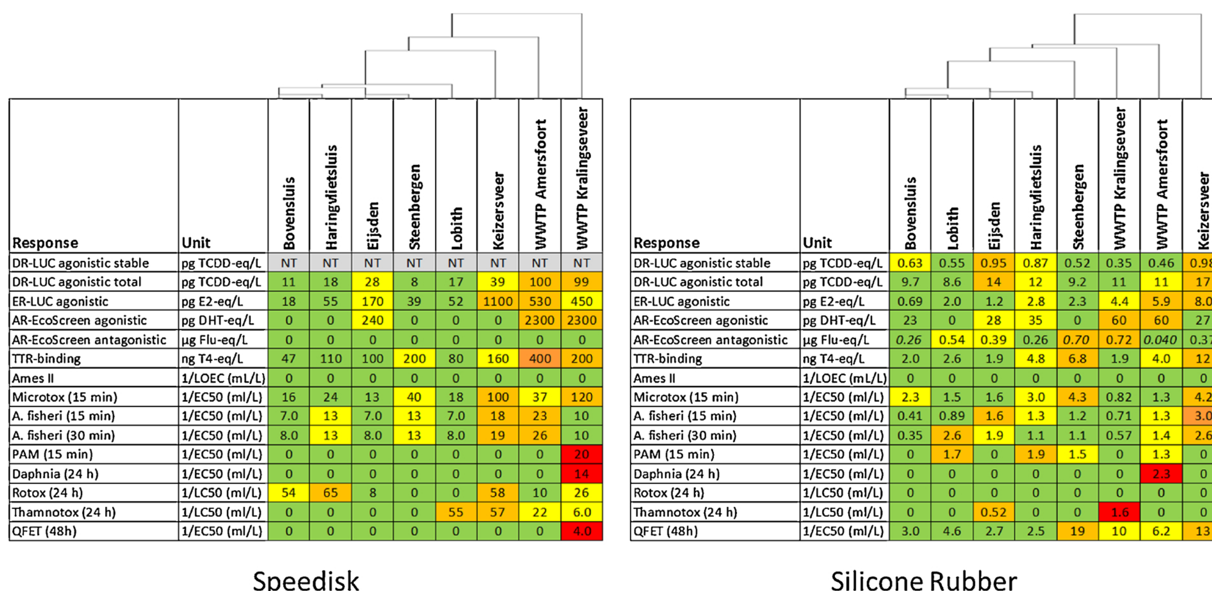
Twenty-nine out of the 205 organic substances measured by RWS in surface water at the river site locations have been defined as priority substances for which the measured concentrations could be compared to annual average environmental quality standard (AA-EQS) values. Out of the 2181 four-weekly observations made for these substances during 2014, AA-EQS exceedances were found in 77 (3.5%) of the cases. For the periods coinciding with the TIPTOP passive sampling campaign 7 exceedances were found out of 249 observations (i.e. 2.8%). In the WWTP effluents a total of 318 organic substances were measured within the E-PRTR monitoring program of which 18 correspond to an AA-EQS. In none of the cases, the measured concentrations in WWTP exceeded the AA-EQS values. All AA-EQS used in the present study are presented in the Supporting Information (Table S8).

Only three compounds were responsible for the 77 EQS exceedances in river water, i.e. benzo[a]pyrene (25), fluoranthene (36), and tributyltin (16). For the PAHs, most exceedances were observed at locations Eijsden and Lobith, i.e. in 9 out of 13 samples per site for benzo[a]pyrene and in all 13 samples per site for fluoranthene. Since the freshwater EQS values for both PAHs are based on human-health via consumption of fishery products (QS_{biota}) (Sub-Group on Review of the Priority Substances List, 2011a, 2011b), EQS exceedance does not imply a risk for the aquatic organisms. In fact, because all measured concentrations were below their respective quality standards for aquatic species ($QS_{freshwater}$). For tributyltin, EQS exceedances were mainly observed at locations Bovensluis and Steenbergen (i.e. in 5 out of 13 samplers per site), with a maximum exceedance factor of five. Since this EQS is based on the HC5 concentration derived from an SSD with aquatic organisms (Anonymous, 2005), these exceedances indicate that tributyltin at these locations may cause toxic pressure levels > 5%.

3.3. Comparison of in vitro bioassay results to effect-based trigger values

Effect Based Trigger (EBT) values are meant to distinguish acceptable from poor water quality where EBT exceedance should be regarded as a trigger for further testing. The method for deriving EBT values, however, is still under development and debate. Two recent studies have proposed different methodologies for EBT derivation using a combination of laboratory aquatic toxicity data and field observations as point of departure (Van der Oost et al., 2017), or the EU environmental quality standards (EQS) for individual compounds (Escher et al., 2018). Overall, for four of the TMoA-specific *in vitro* bioassays applied in the present study, EBT values were available to which the bioassay responses were compared (Table 3), i.e. for PAH-like AhR agonistic, estrogenic, anti-androgenic, and TTR-binding acitivity.

In general, trigger values were only exceeded by Speedisk extracts (and not by silicone rubber extracts), except for the EBT-value of 6.36 ng BaPEq/L derived for the DR-LUC assay (Escher et al., 2018). This EBT value, however, is most probably too conservative, as it is > 20 times lower than the alternative EBT-value of 150 ng BaPEq/L (Van der Oost et al., 2017), which is driven by field observations.



Speedisk

Response	Unit	Bovensluis	Haringvlietsluis	Eijsden	Steenbergen	Lobith	Keizersveer	WWTP Amersfoort	WWTP Kralingseveer
DR-LUC agonistic stable	pg TCDD-eq/L	NT	NT	NT	NT	NT	NT	NT	NT
DR-LUC agonistic total	pg TCDD-eq/L	11	18	28	8	17	39	100	99
ER-LUC agonistic	pg E2-eq/L	18	55	170	39	52	1100	530	450
AR-EcoScreen agonistic	pg DHT-eq/L	0	0	240	0	0	0	2300	2300
AR-EcoScreen antagonistic	µg Flu-eq/L	0	0	0	0	0	0	0	0
TTR-binding	ng T4-eq/L	47	110	100	200	80	160	400	200
Ames II	1/LOEC (mL/L)	0	0	0	0	0	0	0	0
Microtox (15 min)	1/EC50 (ml/L)	16	24	13	40	18	100	37	120
A. fisheri (15 min)	1/EC50 (ml/L)	7.0	13	7.0	13	7.0	18	23	10
A. fisheri (30 min)	1/EC50 (ml/L)	8.0	13	8.0	13	8.0	19	26	10
PAM (15 min)	1/EC50 (ml/L)	0	0	0	0	0	0	0	20
Daphnia (24 h)	1/EC50 (ml/L)	0	0	0	0	0	0	0	14
Rotox (24 h)	1/LC50 (ml/L)	54	65	8	0	0	58	10	26
Thamnotox (24 h)	1/LC50 (ml/L)	0	0	0	0	55	57	22	6.0
QFET (48h)	1/EC50 (ml/L)	0	0	0	0	0	0	0	4.0

Silicone Rubber

Response	Unit	Bovensluis	Lobith	Eijsden	Haringvlietsluis	Steenbergen	WWTP Kralingseveer	WWTP Amersfoort	Keizersveer
DR-LUC agonistic stable	pg TCDD-eq/L	0.63	0.55	0.95	0.87	0.52	0.35	0.46	0.98
DR-LUC agonistic total	pg TCDD-eq/L	9.7	8.6	14	12	9.2	11	11	17
ER-LUC agonistic	pg E2-eq/L	0.69	2.0	1.2	2.8	2.3	4.4	5.9	8.0
AR-EcoScreen agonistic	pg DHT-eq/L	23	0	28	35	0	60	60	27
AR-EcoScreen antagonistic	µg Flu-eq/L	*0.26*	0.54	0.39	0.26	*0.70*	0.72	*0.040*	0.37
TTR-binding	ng T4-eq/L	2.0	2.6	1.9	4.8	6.8	1.9	4.0	12
Ames II	1/LOEC (mL/L)	0	0	0	0	0	0	0	0
Microtox (15 min)	1/EC50 (ml/L)	2.3	1.5	1.6	3.0	4.3	0.82	1.3	4.2
A. fisheri (15 min)	1/EC50 (ml/L)	0.41	0.89	1.6	1.3	1.2	0.71	1.3	3.0
A. fisheri (30 min)	1/EC50 (ml/L)	0.35	2.6	1.9	1.1	1.1	0.57	1.4	2.6
PAM (15 min)	1/EC50 (ml/L)	0	1.7	0	1.9	1.5	0	1.3	0
Daphnia (24 h)	1/EC50 (ml/L)	0	0	0	0	0	0	2.3	0
Rotox (24 h)	1/LC50 (ml/L)	0	0	0	0	0	0	0	0
Thamnotox (24 h)	1/LC50 (ml/L)	0	0	0.52	0	0	1.6	0	0
QFET (48h)	1/EC50 (ml/L)	3.0	4.6	2.7	2.5	19	10	6.2	13

Fig. 2. Hierarchical clustering of the combined *in vitro* and *in vivo* toxicity profiles of Speedisk (left) and silicone rubber (right) passive sampler extracts (expressed per liter water) per sampling site. For each bioassay response (R), the different colors indicate to which quartile (Qx) of the distribution of responses the observation belongs (Hamers et al., 2015). Green: R ≤ Q2; yellow: Q2 < R ≤ Q3; orange: Q3 < R < Q3 + 1.5xIQR; dark orange (outliers): Q3 + 1.5xIQR < R ≤ Q3 + 3xIQR; red (extremes): Q3 + 3xIQR < R, with IQR meaning interquartile range. Exception to this classification is made for responses below limit of quantification (0), which are colored green. AR-antagonistic responses in italics indicate that cytotoxicity could not be excluded. NT means not tested. Results from Ames II assay and *in vivo* bioassays has been expressed as reciprocal values of LOEC, EC50, or LC50, to allow similar interpretation as the *in vitro* bioequivalent concentrations, i.e. higher values indicate higher toxicity.

Escher et al (2018) also indicated that comparability between both EBT values could be obtained if more compounds with EQS values were tested in the DR-LUC bioassay than currently included in the EBT value (n = 4). Altogether, Speedisk extracts from the WWTP locations exceeded the EBT values for most bioassays. For the river sites, EBT exceedance was observed for location Keizersveer in the ER-LUC bioassay, indicating that ecological risk at this site cannot be excluded. The observed EBT exceedance for the TTR-binding assay at most locations should be regarded with caution, because the preliminary EBT of 60 ng T4-Eq/L is also based on few compounds (n = 5) (Escher et al., 2018).

A similar comparison was made for the *in vivo* bioassay results and their corresponding EBT values. This exercise was hampered, however, by the fact that for many samples toxicity was too low to be quantifiable. For those bioassay results with quantifiable results, only one EBT exceedance was observed, i.e. for WWTP Kralingseveer in the PAM bioassay (see Supplementary Information Table S9).

3.4. Toxic pressure estimates based on targeted chemical analyses

Toxic pressure estimates for the river sites were based on the

Table 3
Effect-based trigger (EBT) value exceedance factors for AhR agonistic, ER agonistic, AR-antagonistic, and TTR-binding TMoA in Speedisk (SD) and silicone rubber (SR) passive samplers.

	DR-LUC AhR agonistic[b]				ER-LUC ER agonistic				AR-EcoScreen AR antagonistic		TTR-binding	
EBT (ng RC-Eq/L)[a]	150		6.36		0.5		0.37[c]		25000	3280	60	
Reference[d]	VdO		E		VdO		A		VdO	E	E	
Location	SD	SR	SD	SR	SD	SR	SD	SR	SR[e]	SR[e]	SD	SR
Eijsden	0.61	0.30	14	7.2	0.34	0.00	0.46	0.00	0.02	0.12	1.7	0.03
Keizersveer	0.82	0.36	19	8.4	2.2	0.02	3.0	0.02	0.01	0.11	2.6	0.21
Bovensluis	0.24	0.21	5.6	4.9	0.04	0.00	0.05	0.00	0.01	0.08	0.78	0.03
Steenbergen	0.17	0.20	4.0	4.7	0.08	0.00	0.10	0.01	0.03	0.21	3.3	0.11
Haringvlietsluis	0.38	0.25	8.9	5.9	0.11	0.01	0.15	0.01	0.01	0.08	1.8	0.08
Lobith	0.35	0.18	8.4	4.3	0.10	0.00	0.14	0.01	0.02	0.17	1.3	0.04
WWTP Kralingseveer	2.1	0.23	50	5.3	0.89	0.01	1.2	0.01	0.03	0.22	3.5	0.03
WWTP Amersfoort	2.2	0.23	52	5.5	1.1	0.01	1.4	0.02	0.00	0.01	6.6	0.07

[a]EBT values are expressed in ng reference compound (RC) per liter water, with RC is BaP for DR-LUC, E2 for ER-LUC, Flu for AR-EcoScreen, and T4 for TTR-binding assay.
[b]The observed bioassay responses DR-LUC responses for total extracts in pg TCDD-eq/L were calculated into ng BaP-eq/L using a molar relative potency of 2.45*10-4 for BaP compared to TCDD, derived from Neale et al., 2017.
[c]Median out of five EBT values specifically derived for five cell-based bioassays, not including the ER-LUC assay performed in this paper.
[d]References VdO and E refer to Van der Oost et al., 2017 and Escher et al., 2018, respectively.
[e]Anti-androgenic responses to Speedisk extracts were all below the limit of detection.

Table 4
Toxic pressure calculations expressed as fraction of species potentially affected by multiple substances (msPAF) based on target-analyzed compounds in water and in passive samplers. Number of compounds exceeding the limit of detection are indicated by n.

Location	Water[a]		Speedisk		Silicone rubber	
	msPAF (%)	n	msPAF (%)	n	msPAF (%)	N
Eijsden	0.3	25	0.2	68	0.3	69
Keizersveer	0.0	22	0.1	40	0.1	74
Bovensluis	0.0	21	0.1	40	0.0	70
Steenbergen	0.0	12	0.1	49	0.1	69
Haringvlietsluis	0.0	20	0.1	39	0.0	68
Lobith	0.0	20	0.3	45	0.1	70
WWTP Kralingseveer	0.2/0.15	10/14	0.3	88	0.1	69
WWTP Amersfoort	0.1/0.2	12/15	0.3	67	0.1	74

[a] msPAF calculations based on targeted water analysis were based on average concentrations of four-weekly grab samples collected at the river locations in parallel to the passive sampling campaign and of 5 to 6 grab samples collected at the WWTPs during two different sampling campaigns in 2012-2013 (first value; (Baltussen, 2013)) and 2015 (second value; (Baltussen, 2015)).

Table 5
Hazardous concentration factors for 5% of the exposed species (HCf5) as derived from species-sensitivity distributions based on seven acute median effective concentration factors ($EC50_f$) obtained by testing the passive sampler extracts in seven *in vivo* bioassays.

Location	Based on observed acute $EC50_f$ values		Based on extrapolated chronic NOEC values	
	HCf5 Speedisk	HCf5 silicone rubber	HCf5 Speedisk	HCf5 silicone rubber
Eijsden	55	380	5.5	38
Keizersveer	7.7	120	0.77	12
Bovensluis	19	450	1.9	45
Steenbergen	17	160	1.7	16
Haringvlietsluis	14	270	1.4	27
Lobith	17	230	1.7	23
WWTP Kralingseveer	10	530	1.0	53
WWTP Amersfoort	24	190	2.4	19

average concentrations of the 205 organic compounds that were analyzed by RWS in grab river water samples collected during the passive sampling period. For the WWTP locations, toxic pressure estimates were based on the concentrations of the 318 organic compounds that were analyzed within the E-PRTR monitoring programs in WWTP effluent grab samples collected in 2013 or 2015. In addition, toxic pressure was calculated for each passive sampler based on the water concentration corresponding to the organic compounds analyzed in the Speedisk samplers (n = 171) and silicone rubber samplers (n = 181).

All toxic pressure calculations based on targeted chemical analyses pointed out that acute EC50 values were exceeded for at maximum 0.3% of the species, regardless of the sampling location (Table 4). Literally, an msPAF value of 0.3% means that only three out of 1000 exposed species experience an effective concentration above their acute EC50. It is not likely that the impact of such a low toxic pressure can be verified with ecological field observations (species census). At all sites, the calculated toxic pressure estimates could be attributed to a very limited set of pesticides ($n \leq 4$, in most cases n = 1). An overview of the responsible compounds per sample location is provided in Table S10 for water samples and Table S11 for passive samplers.

For two reasons, overall toxic pressure estimates were based on acute EC50 values rather than chronic NOEC values that are generally adopted to derive EQS values: i) empirical experience (De Zwart et al., 2009; Posthuma and De Zwart, 2012, 2006) learned us that acute EC50 exceedances correspond one-to-one to actual impacts on biodiversity observed in the field, and ii) due to higher availability of acute toxicity data, which are easier and cheaper to produce than chronic toxicity data, the acute SSDs underlying the msPAF calculations are generally derived from data on a more diverse set of species and are therefore more reliable than chronic SSDs.

To place the toxic pressure $\leq$0.3% findings into EQS perspective, the same strategy was used to calculate the overall toxic pressure of a hypothetical surface water with a toxicant loading equivalent to the AA-EQS concentrations for priority substances also used in section 3.2 (see Table S8), assuming that all chemicals are fully bioavailable. In this hypothetical waterbody, 4.7% of a generalized species assemblage would be exposed above their respective acute EC50. Given the one-to-one relationship between EC50 exceedance and loss of biodiversity, this result means that biodiversity would be reduced by almost 5%. Applied to NOEC, which is more in line with the methods applied in the setting of standard values (European Commission, 2011), the same 48 concentrations together would cause exceedance of chronic NOECs of 33.3% of the exposed species. On average, each of the 48 priority substances would cause NOEC exceedance of only 0.69%, indicating that, on average, current WFD-EQS are set well below the HC5-levels.

3.5. Toxic pressure estimates based on in vivo bioassay results

Based on the SSD of measured acute $EC50_f$ values (see Supplemental Material Figures S12 and S13 and Table S5 for the underlying $EC50_f$ values), the toxic pressure in non-concentrated water (i.e. corresponding to a concentration factor of 1) was assessed to be 0.0%, both for Speedisk and silicone rubber samplers. Using the same SSDs, the concentration factor of the water was derived for which 5% of the species is estimated to be exposed above acute EC50 value. This hazardous concentration factor (HCf5), which can be regarded as a margin-of-exposure towards acute effect concentrations, ranged from 8 to 55 for the Speedisk samplers and was > 100 for the silicone rubber samplers (Table 5). The SSDs of measured acute $EC50_f$ values were further extrapolated into SSDs for chronic NOECs by shifting the acute SSDs by a factor of 10 to the left (see Figures S12 and S13). Consequently, chronic toxic pressure estimates in non-concentrated water ranged for Speedisks from 0.0 to 4.9% with the exception of one of the river site locations Keizersveer (11.9%). The corresponding estimates for the margin-of-exposure towards chronic effects ranged from 1 to 5.5, with the exception again of Keizersveer, for which the margin-of-exposure was estimated as < 1 (i.e. HCf5 = 0.77; Table 5). For silicone rubber samplers, chronic toxic pressure estimates were all 0.0% with margins-of-exposure to chronic effects ranging from 12 to 53 (Table 5).

It should be realized that in many cases (indicated by a " > "sign in Table S5), the $EC50_f$ could not be quantified, because the passive sampling extracts were simply not concentrated enough. In order to include unquantifiable toxicity observations in the overall toxic pressure calculations, the highest test concentration multiplied by a factor of two was included as $EC50_f$ in the SSD calculation. This may however represent an overestimation of toxicity.

3.6. Minimum toxic pressure estimates based on total molar concentration

Under the assumption that the molar loading of the silicone rubber passive sampler material is equivalent to the molar loading of lipids in exposed organisms, minimum toxic pressure can be estimated by interpolation of the total molar concentration in the passive samplers determined by vapor pressure osmometry into SSDs for minimum toxicity. Verhaar et al. (1995) provided critical body residue levels for minimum toxicity compounds at three different levels of effect, which

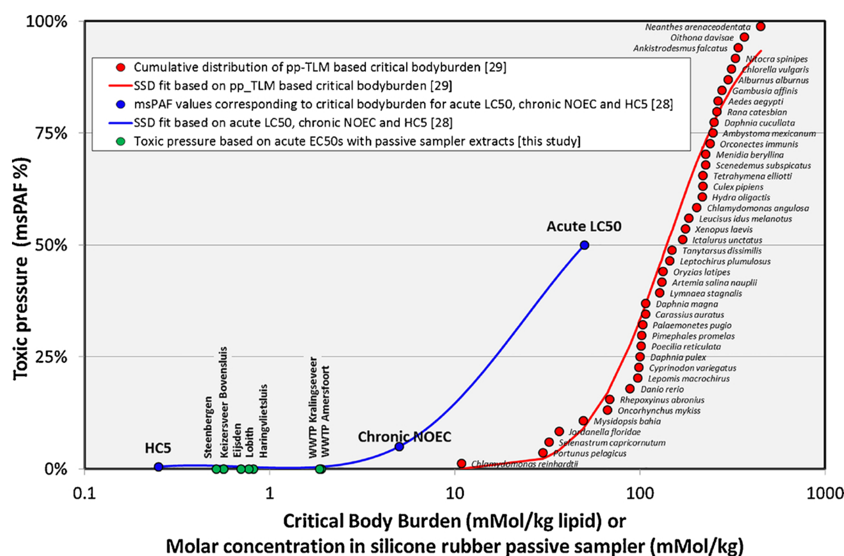


Fig. 3. SSDs for baseline minimum toxicity of non-polar narcotic compounds based on Verhaar et al. (1995) and Kipka and Di Toro (2009), together with the molar loading and corresponding toxic pressure (see 3.5) of the silicone rubber extracts (green dots).

were used to construct an SSD, i.e. a) the median acute LC50 at 50 mMol/kg lipid invoking an acute toxic pressure of 50% (msPAF units); b) the median chronic NOEC, also defined as the HC50 (hazard concentration for 50% of species) at 0.5 mMol/kg lipid corresponding to a 10-fold lower acute toxic pressure of 5%; c) the 5th percentile of the chronic NOEC distribution also defined as HC5 (hazard concentration for 5% of species) at 0.25 mMol/kg lipid with again a 10-fold lower estimated acute toxic pressure of 0.5%. Kipka and Di Toro (2009) have published a minimum toxicity SSD based on critical body burdens derived from a poly parameter target lipid model (pp-TLM) calibrated against acute LC50 values for 42 different species. This model yielded a median (HC50) and 5% (HC5) acute toxic pressure at 140 and 39 mMol/kg lipid, respectively. Interpolation of the molar concentrations of the silicone rubber passive sampler in both SSDs yielded acute toxic pressure estimates of 0.0%, corresponding to the low toxic pressure estimates based on the *in vivo* results (see 3.5). Apparently, molar loadings at the sampling sites are far lower than critical body burdens, and therefore too low to invoke any significant effect.

At equilibrium, Jahnke et al. (2008) have demonstrated that the lipid based molar concentration of non-polar compounds is approximately a factor of 30 higher than the concentration in silicone rubber. If the silicone rubber sampler had only sampled non-polar substances, this finding implies that the green dots in Fig. 3 should be shifted by a factor 30 to the right. Even then, the observed toxic pressure estimates (see 3.5) fit perfectly in the pp-TLM based SSD, but the LC50-NOEC-HC5 based SSD seems to overestimate the observed toxic pressure. This is most likely caused by the fact that the silicone rubber sampler contains not only non-polar minimum toxicity, but also polar minimum toxicity, which is accounted for in the pp-TLM based SSD.

The estimated negligible minimum toxic pressure based on the vapor pressure osmometry was confirmed by an alternative method, which was based on calculating the concentrations in the water phase corresponding to the target-analyzed concentrations in the passive samplers into waterborne chemical activities. Minimum toxic pressure was estimated to be negligibly small (0.6-0.7%) according to similar methodology as shown in Fig. 1 (see Supporting Information section S3.5).

3.7. TMoA-specific toxic pressure estimates based on in vitro bioassay results

In this final strategy, SSDs were constructed based on chronic *in vivo* NOECs for all compounds that have *in vitro* AhR-agonistic, ER-agonistic, or AR-antagonistic potency in the H4L1.1c4, the BG1Luc4E2, or the mda-kb2 reporter gene assays, respectively, according to the ToxCast database. Before SSD calculation, *in vivo* NOEC values of all individual compounds were first calculated into a bioanalytical equivalent concentration of the reference compounds 2,3,7,8-tetrachlorodibenzo-[p]-dioxin (TCDD), 17ß-estradiol (E2), or flutamide (Flu), respectively, by multiplication with their compound-specific REP factor. For each TMoA, $NOEC_{BEQ}$ values from all compounds were averaged per species yielding a single average $NOEC_{BEQ}$ value per species. Finally the TMoA-specific SSD was constructed as the cumulative distribution function of $NOEC_{BEQ}$ values for all species (see Figure S15). Because this strategy assumes that all included compounds act as AhR-agonists, estrogens, androgens, or anti-androgens as their primary TMoA, only compounds with REP factors > 0.001 (i.e. being at least 1000x less potent than the bioassay specific reference compound) were included. This criterion prevents that very toxic compounds (i.e. low NOEC) that primarily act through a different primary TMoA but also have weak activity in the *in vitro* bioassay (i.e. low REP value) cause a shift of the SSD to the left. This may especially be the case for estrogenic and anti-androgenic potencies, which are exerted by a wide variety of compounds.

TMoA-specific toxic pressures were estimated by interpolation of the waterborne *in vitro* toxic potencies in the TMoA-specific SSDs, and ranged from 0.1% to 31% depending on the TMoA, sampling location and sampler type considered (Table 6). For the river sampling sites, ER- and AR-agonistic toxic pressures were < 5% except for ER-agonism in Keizersveer, confirming the results obtained by site-by-site comparison to effect-based trigger values (see Table 3). For AhR-agonistic (i.e. PAH-like) activity, toxic pressure estimates for the river sites were ~5%, for both passive sampler types, whereas the toxic pressure for the Speedisk extracts of the WWTP locations were clearly higher. These findings suggest indeed that the ultimate EBT value for PAH-like compounds has a value in between the conservative and the less conservative EBT estimate (i.e. 6.36 (Escher et al., 2018) and 150 ng BaP-Eq/L (Van der Oost et al., 2017); see 3.3).

Relatively high toxic pressure (> 10%) was estimated for the AR-antagonsitic potency collected by the silicone rubber samplers. In fact, bioassay responses (Flu-Eq/L) for the river sites exceeded the HC5 by a factor 7.8 to 112. Similar observations were made by Van der Oost et al. (2017), who found an exceedance of the anti-androgenic HC5 by a factor of 35 at background locations. Consequently, the authors decided to establish their EBT value (see section 3.3) on anti-androgenicity

Table 6
TMoA-specific toxic pressure (expressed as msPAF in %) calculated by interpolation of the waterborne toxic potency in its respective TMoA-specific SSD (ago: agonistic; anta: antagonistic; SD: Speedisk; SR: silicone rubber).

Location	AhR-ago		ER-ago		AR-ago		AR-anta[a]	
	SD	SR	SD	SR	SD	SR	SR[b]	SR[b]
Eijsden	7.0	5.7	3.1	0.16	1.2	0.32	15	26
Keizersveer	7.7	6.0	7.1	0.56		0.32	14	26
Bovensluis	5.2	5.0	0.9	0.11		0.28	12	23
Steenbergen	4.7	4.9	1.4	0.25			18	31
Haringvlietsluis	6.1	5.3	1.7	0.29		0.38	12	23
Lobith	6.0	4.8	1.6	0.22			16	29
WWTP Kralingseveer	10.1	5.2	4.8	0.38	3.8	0.53	18	31
WWTP Amersfoort	10.2	5.2	5.2	0.46	3.9	0.53	5.7	12

[a]Two different SSDs were constructed for anti-androgenicity because ToxCast reported for each compound two IC50 results in the mda-kb2 bioassay, i.e. based on the capacity to antagonize activation of the reporter construct by an EC100 = 10nM concentration of AR-agonist R1881 (left) or by an EC65 = 0.5 nM concentration of R1881 (right).
[b]Anti-androgenic responses to Speedisk extracts were all below the limit of detection.

levels found at background locations, rather than on the much more conservative SSD-based HC5. This is reflected by the different outcomes regarding anti-androgenicity in the EBT-based strategy (Table 3) and the SSD-based strategy (Table 6). This difference can most likely be attributed to the fact that the criterion REP > 0.001 for including compounds having anti-androgenicity as their primary TMoA may not be sufficiently stringent. This is due to the fact that the reference compound flutamide has relatively weak anti-androgenic potencies (IC50 values ranging from 5 to 30 μM), implying that other weak anti-androgenic compounds that most likely have other primary TMoA's are also included in this exercise. For estrogenic compounds, this is not the case, given the very low EC50 value (80 pM) of the estrogenic reference compound E2.

4. Conclusions

4.1. Hypothesis testing

One of the hypotheses tested in the TIPTOP study was that the newly proposed water quality monitoring method, *viz* passive sampling combined with toxicity profiling, is more protective than the traditional compound-by-compound comparison to water quality standards. The low toxicity observed for the passive sampler extracts in both the *in vitro* and *in vivo* bioassays hampered the evaluation of this hypothesis. Results from both chemical analysis of a limited set of target compounds and from *in vitro* and *in vivo* bioassays gave very low responses, indicating that the selected sampling sites were actually too clean to demonstrate the added value of effect-based measurements. Since both types of measurements, including all different interpretations in terms of ecological risk assessment strategies described in the Results section, indicated little to negligible risk from chemical substances, the ultimate conclusion can be that effect-based measurements gave consistent results, with no false positives when compared to the corresponding chemical analyses.

Moreover, the different interpretation strategies clearly illustrated that effect-based measurements yield more informative conclusions in terms of ecological risk assessment. In contrast to a chemical-analytical approach in which many substances are often found to be below the limit of detection, effect-based measurements account for all compounds in the mixture that contribute to the effect. Actually, the observed *in vitro* estrogenic and anti-androgenic potencies in the river samples could be explained at maximum for 15% and 17%, respectively (see Supplementary Material section S3.8), indicating that the observed toxicity – albeit very small – should be attributed to other compounds than the chemically analyzed compounds.

Furthermore, effect-based measurements enable the assessment of a margin-of-exposure towards concentrations where effects in the field may be expected. Estimation of such a margin-of-exposure is more informative than results from chemical analyses of dozens of compounds reported as < LOD. It may for instance be used to prioritize sampling sites for further investigation, such as the river sampling location Keizersveer, which was indicated by many different risk assessment strategies as the site with most deviating and highest toxicity profiles compared to the other river sampling locations.

As a consequence of the fact that an effect-based strategy yields more information per euro spent than chemical analyses, it can also be considered as more cost-effective. Moreover, annual costs for routine chemical analysis of determining 45 priority substances in 12 monthly grab samples is estimated to be 40 k€ per location. For the 8 silicone rubber and 8 Speedisk passive samplers described in the current study, sample collection (12 k€), chemical analysis (20 k€), and effect-based testing using the full battery of *in vitro* (45 k€) and *in vivo* bioassays (45 k€) and data analysis (12 k€) boiled down to 8.5 k€ per sample. In other words, a monitoring strategy based on a bi-annual sampling campaign with two types of passive samplers followed by toxicity profiling using the battery of *in vivo* and *in vitro* bioassays described in this study requires a budget of 34 k€ per location. This budget can further considerably be reduced by decreasing the number of substances to be chemically analyzed (i.e. to only performance reference chemicals for the silicone rubber sheets and a selected set of chemicals expected to be present in both samplers), by making a smart selection of bioassays to be performed (e.g. based on their responsiveness to the river sample extracts), and by making use of routine laboratories and automated data analysis formats to increase the throughput. Finally, it is expected that risk assessment approaches based on chemical analyses will only expand to monitoring an indefinitely large suite of chemicals. In the hypothetical case that the full set of chemicals should be monitored, it can be argued from a theoretical point of view that the number of mechanisms of action covered by these substances should always be lower than the number of substances, implying that effect-based monitoring is more cost-effective by definition. Based on these arguments, we conclude that the TIPTOP study confirms that a combination of passive sampling and toxicity profiling is a more cost-effective monitoring strategy than compound-by-compound chemical analyses.

4.2. SWOT analysis

Based on the experience gained in the present study, the following analysis was made regarding the strengths, weaknesses, opportunities, and threats (SWOT) of a combined approach of passive sampling followed by toxicity profiling.

	POSITIVE	NEGATIVE
INTERNAL	**Strengths** • reduction in costs • time-integrative sampling reduces uncertainty about missed substances present for short periods of exposure • reduced uncertainty about missed pollution episodes • endpoint evaluation closer to WFD-aim of good ecological status • easier interpretation in ecologically relevant terms • results are more informative • possible use as triage method to determine hot spots for in depth study	**Weaknesses** • misfit with current substance oriented legislation • difficult to attribute observed effects to underlying causation, polluting processes and sources of pollution • need for secluded sampling stations • need for further research into methods for sample preparation, extraction, and determination of molar concentration • possibility that peak exposures are averaged out over a long sampling period • theoretical impossibility to exactly translate the observed toxicity in the passive sampler to a corresponding toxicity in the water
EXTERNAL	**Opportunities** • no increase in analytical work load despite expected future increase in pollution diversity • design and setting of effect oriented EQS procedures are less demanding and do not need regular update • biological triage leaves more money available to in depth study of local hot spots and may drive the change to effect-based water quality assessment	**Threats** • changing to effect oriented legislation may take a long time • the proof of concept requires temporary simultaneous application of old and new system (i.e. temporary double costs) • water quality authorities require more bio-analytical and ecotoxicological personnel, which is less available • technologically oriented risk assessors tend to trust the outcome of chemical analyses better than the outcome of biological test systems

4.3. Strategy for future monitoring

The SWOT analysis clearly points out that implementation of effect-based measurements – despite their more informative conclusions in terms of ecological risk assessment and higher cost-efficiency - requires a paradigm shift from substance-based to effect-based risk assessment and legislation. Although such a shift requires an investment in effort, time, and money, it will ultimately be inevitable given the continuously expanding suite of compounds that may ultimately pose a threat to our water systems. Therefore, a strategy was proposed to determine water quality using passive sampler extracts, based on the TIPTOP experiences. The proposed strategy consists of four steps, i.e.

1. Determine narcotic toxic pressure based on total molar sum in passive samplers;
2. Determine generic toxic pressure based on *in vivo* bioassay results;
3. Determine TMoA-specific toxic pressure based on mechanistic *in vitro* bioassay results;
4. Compare toxic pressure estimates to maximum acceptable toxic pressure level.

Ad 1: Narcotic toxic pressure estimates

Baseline toxic pressure can be estimated by interpolation of the molar sum in the pp-TLM based SSD, as demonstrated in 3.6. In this study, the total molar sum was determined by vapor pressure osmometry. Alternatively, a minimal separation GC-MS method was proposed by Van Loon et al. (1996) to determine the total molar sum in surface water and drinking water extracts, which have the advantage of lower detection limits, less interference with lipids and other disturbing compounds, and a shorter run time.

Ad 2: Generic toxic pressure estimates

Generic toxic pressure can be determined by constructing sample-specific SSDs based on median effective concentration factors ($EC50_f$) determined in *in vivo* bioassays, similar as in 3.5. Within the current study, SSDs were based on the median (μ) and standard deviation (σ) of seven (log-transformed) $EC50_f$ values obtained in seven different *in vivo* bioassays. It is foreseen, however, that the recently described approach for toxic pressure assessment of data-poor chemical mixtures based on hierarchical Bayesian inference (Oldenkamp et al., 2015) can also be used for a confident μ and σ assessment, resulting in a confident assessment of generic toxic pressure, based on a limited number of *in vivo* bioassay results.

Generic toxic pressure estimates are currently based on substance concentrations in the water phase, which are calculated into msPAF values using SSDs, as shown in 3.4. Since toxic pressure calculations are basically the product of two probabilities (Fig. 1), it requires two probability distribution functions with known parameters, i.e. 1) a log-normal probability density function for the possible concentrations in the water and 2) a log-normal cumulative distribution function for the critical effect concentrations, each with realistic estimates of mean and standard deviation. As described by Van de Meent et al. (submitted), generic toxic pressure calculation can be applied to single chemicals, or to mixtures of chemicals with the similar variances of interspecies sensitivities (width of SSD), even under great uncertainty. This means that generic toxic pressure calculation as in Fig. 1 is a suitable method to assess the mixture toxic pressure in natural waters from the results of passive sampling, as done in the TIPTOP study, where means and standard deviations of exposure concentrations and critical effect concentrations of unknown sets of chemicals can often be estimated well enough for the purpose of assessment of the probability that one exceeds the other in the water system under study. Most importantly, this method does not require costly chemical analysis of all toxicants

potentially present in a water system. In addition, the method can be applied to mixtures of specifically-acting toxicants (e.g. estrogens) of which neither chemical structure nor concentrations are known, but of which the toxic effect can be assessed by means of a suitable bio-assay. Given the expected problems, however, regarding the acceptance of effect-based monitoring strategies (see section 4.2), it is not foreseen that chemical water quality assessment based on measuring molar concentrations will be accepted in the short term.

Ad 3: TMoA-specific toxic pressure estimate

Given the fact that a specific TMoA is usually exerted at relatively low concentrations by a relatively small subset of substances with often unknown identity, TMoA specific toxic pressure can best be determined using specific *in vitro* bioassays. Targeted chemical analysis is very likely to miss many of the unknown compounds working through a specific TMoA, whereas the molar concentration is considered to be a too generic measure of exposure. TMoA-specific toxic pressure can be estimated by interpolation of the observed *in vitro* bioassay responses (expressed in terms of BEQ) into an SSD based on chronic NOECs, which are converted into BEQ concentrations prior to SSD construction, similar as in 3.7. NOECs should be selected more strictly for inclusion in the SSD than done in 3.7, because the SSD should only consist of NOECs reflecting an adverse outcome that is most likely caused by the same TMoA as measured in the *in vitro* bioassay. Finally, TMoA specific SSDs can also be used to determine HC5 values (in terms of BEQ), which can serve as a basis for the derivation of trigger values for the mechanistic *in vitro* bioassays.

Ad 4: Comparison to a maximum acceptable toxic pressure level

So far, criteria for acceptability of toxic pressure have not been set explicitly, other than by formulating environmental quality criteria for a named set of substances, which should not be exceeded. It is recommended that the WFD formulates overall criteria for testing ‘good ecotoxicological water quality’. Lacking this, the generic acute toxic pressure of an hypothetical water containing all test substances at their respective EQS, which has been calculated to amount to an exceedance probability of almost 5% (see section 3.4), could be used. It seems reasonable to use the same 5% criterion for the exceedance probability of critical bodyburdens regarding the narcotic compounds and for the exceedance probability of chronic NOECs regarding compounds with a specific TMoA.

Acknowledgements

The TIPTOP project was funded by a Long-Range Initiative grant (LRI-ECO23) from the European Chemical Industry Council (CEFIC). The connected Research Liaison Team is acknowledged for providing feedback and guidance during the progress of the project. Nick Zwart was funded by the Dutch Technology Foundation (STW; project 12396). Marcel Kotte (Rijkswaterstaat), Manda Milosevic-Bilic (Hoogheemraadschap van Schieland en de Krimpenerwaard), and Jan Wout Koelewijn (Waterschap Vallei en Veluwe) are acknowledged for providing access to the passive sampling locations at the river sites, WWTP Kralingseveer, and WWTP Amersfoort, respectively. Joop Baltussen (BACO Adviesbureau BV) and Ed Steenbergen (Waterschap Rivierenland) are acknowledged for sharing the E-PRTR monitoring results of 2015 prior to reporting. We thank Ad Ragas and Rik Oldenkamp (Radboud University Nijmegen) for very constructive discussions regarding the monitoring strategy.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:10.1016/j.etap.2018.09.005.

References

Aldenberg, T., Jaworska, J.S., Traas, T.P., 2002. Normal Species Sensitivity Distributions and Probabilistic Ecological Risk Assessment. In: Posthuma, L., Suter IIG., Traas, T.P. (Eds.), Species Sensitivity Distributions in Ecotoxicology. Lewis Publishers CRC Press Taylor & Francis, Boca Raton FL, USA, pp. 49–102.

Anonymous, 2005. Environmental Quality Standards (EQS) Substance Data Sheet Priority Substance No. 30 Tributyltin compounds (TBT-ion) CAS-No. 688-73-3 (36643-28-4). Brussels.

Baltussen, J., 2015. Watergerelateerde emissies vanuit rwzi’ s in het kader van de PRTR (jaar 2015). Amersfoort, The Netherlands.

Baltussen, J., 2013. Watergerelateerde emissies vanuit rwzi’ s in het kader van de iPRTR. Amersfoort, The Netherlands.

Baltussen, J., 2010. Emissie onderzoek op een zestal rwzi’ s in het kader van de E-PRTR. Amersfoort, The Netherlands.

Bliss, C.I., 1939. The toxicity of poisons applied jointly. Ann. Appl. Biol. 26, 585–615.

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

Chapman, P.M., Anderson, B., Carr, S., Engle, V., Green, R., Hameed, J., Harmon, M., Haverland, P., Hyland, J., Ingersoll, C., Long, E., Rodgers Jr, J., Salazar, M., Sibley, P.K., Smith, P.J., Swartz, R.C., Thompson, B., Windom, H., 1997. General guidelines for using the sediment quality triad. Mar. Pollut. Bull. 34, 368–372.

De Zwart, D., 2002. Observed regularities in SSDs for aquatic species. Species Sensitivity Distributions in Ecotoxicology. pp. 133–154 Boca Raton FL, USA.

De Zwart, D., Posthuma, L., 2005. Complex mixture toxicity for single and multiple species. Environ. Toxicol. Chem. 24, 2665. https://doi.org/10.1897/04-639R.1.

De Zwart, D., Posthuma, L., Gevrey, M., Von der Ohe, P.C., De Deckere, E., 2009. Diagnosis of Ecosystem Impairment in a Multiple-Stress Context — How to Formulate Effective River Basin Management Plans. Integr. Environ. Assess. Manag. 5, 38–49.

Emelogu, E.S., Pollard, P., Robinson, C.D., Smedes, F., Webster, L., Oliver, I.W., McKenzie, C., Seiler, T.B., Hollert, H., Moffat, C.F., 2013. Investigating the significance of dissolved organic contaminants in aquatic environments: Coupling passive sampling with in vitro bioassays. Chemosphere 90, 210–219. https://doi.org/10.1016/j.chemosphere.2012.06.041.

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., 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.

Escher, B.I., Hermens, J.L.M., 2002. Critical Review Modes of Action in Ecotoxicology : Their Role in Body Burdens, Species Sensitivity, QSARs, and Mixture Effects. Environ. Sci. Technol. 36, 4201–4217. https://doi.org/10.1021/es015848h.

European Commission, 2014. TECHNICAL REPORT ON AQUATIC EFFECT-BASED MONITORING TOOLS (No. 2014-077). https://doi.org/10.2779/7260.

European Commission, 2013. Directives of 12 August 2013 amending Directives 2000/60/EC and 2008/105/EC as regards priority substances in the field of water policy. Off. J. Eur. Union L226, 1–17. https://doi.org/http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32013L0039.

European Commission, 2011. Technical Guidance For Deriving Environmental Quality Standards - Common Implementation Strategy for the Water Framework Directive (2000 / 60 / EC) Guidance Document No. 27.

European Commission, 2008. Directive 2008/105/EC of the European Parliament and of the Council of 16 December 2008 on environmental quality standards in the field of water policy, amending and subsequently repealing Council Directives 82/176/EEC, 83/513/EEC, 84/156/EEC, 84/491/ECC. Off. J. Eur. Union L348, 84–97.

European Commission, 2003. Technical guidance document on risk assessment - Part II (No. EUR 20418 EN/2). European Chemicals Bureau, Ispra, Italy.

European Commission, 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. Off. J. Eur. Comm L327, 1–72.

Hamers, T., Kamstra, J.H., van Gils, J., Kotte, M.C., van Hattum, A.G.M., 2015. The influence of extreme river discharge conditions on the quality of suspended particulate matter in Rivers Meuse and Rhine (The Netherlands). Environ. Res. 143, 241–255. https://doi.org/10.1016/j.envres.2015.10.019.

Hamers, T., Legler, J., Blaha, L., Hylland, K., Marigomez, I., Schipper, C.A., Segner, H., Vethaak, A.D., Witters, H., de Zwart, D., Leonards, P.E.G., 2013. Expert opinion on toxicity profiling-report from a NORMAN expert group meeting. Integr. Environ. Assess. Manag. 9, 185–191. https://doi.org/10.1002/ieam.1395.

Hamers, T., Leonards, P.E.G., Legler, J., Dick Vethaak, A., Schipper, C.A., 2010. Toxicity profiling: An integrated effect-based tool for site-specific sediment quality assessment. Integr. Environ. Assess. Manag. 6, 761–773. https://doi.org/10.1002/ieam.75.

Jahnke, A., McLachlan, M.S., Mayer, P., 2008. Equilibrium sampling: Partitioning of organochlorine compounds from lipids into polydimethylsiloxane. Chemosphere 73, 1575–1581. https://doi.org/10.1016/j.chemosphere.2008.09.017.

Kipka, U., Di Toro, D.M., 2009. Technical basis for polar and nonpolar narcotic chemicals and polycyclic aromatic hydrocarbon criteria. III. A polyparameter model for target lipid partitioning. Environ. Toxicol. Chem. 28, 1429–1438. https://doi.org/10.1897/08-364.1.

Neale, P.A., Altenburger, R., Aït-aïssa, S., Brion, F., Busch, W., De Aragao Umbuzeiro, G., Denison, M.S., Du Pasquier, D., Hilscherová, K., Hollert, H., Morales, D.A., Novák, J., Schlichting, R., Seiler, T.-B., Serra, H., Shao, Y., Tindall, A.J., Tollefsen, K.E.,

Williams, T.D., Escher, B.I., 2017. Development of a bioanalytical test battery for water quality monitoring: Fingerprinting identified micropollutants and their contribution to effects in surface water. Water Res. 123, 734–750. https://doi.org/10.1016/j.watres.2017.07.016.

Oldenkamp, R., Hendriks, H.W.M., Van De Meent, D., Ragas, A.M.J., 2015. Hierarchical Bayesian Approach to Reduce Uncertainty in the Aquatic Effect Assessment of Realistic Chemical Mixtures. Environ. Sci. Technol. 49, 10457–10465. https://doi.org/10.1021/acs.est.5b02651.

Posthuma, L., De Zwart, D., 2012. Predicted mixture toxic pressure relates to observed fraction of benthic macrofauna impacted by contaminant mixtures. Environ. Toxicol. Chem. 31, 2175–2188. https://doi.org/10.1002/etc.1923.

Posthuma, L., De Zwart, D., 2006. Predicted effects of toxicant mixtures are confirmed by changes in fish species assemblages in Ohio, USA. Rivers. Environ. Toxicol. Chem. 25, 1094–1105.

Rusina, T.P., Smedes, F., Klanova, J., Booij, K., Holoubek, I., 2007. Polymer selection for passive sampling: A comparison of critical properties. Chemosphere 68, 1344–1351. https://doi.org/10.1016/j.chemosphere.2007.01.025.

Rusina, T.P., Smedes, F., Koblizkova, M., Klanova, J., 2010. Calibration of silicone rubber passive samplers: Experimental and modeled relations between sampling rate and compound properties. Environ. Sci. Technol. 44, 362–367. https://doi.org/10.1021/es900938r.

Sabaliunas, D., Lazutka, J., Sabaliuniene, I., Sodergren, A., 1998. Use of semipermeable membrane devices for studying effects of organic pollutants: Comparison of pesticide uptake by semipermeable membrane devices and mussels. Environ. Toxicol. Chem. 17, 1815–1824. https://doi.org/10.1897/1551-5028(1998)017<1815:uosmdf>2.3.co;2.

Smedes, F., Beetlje, H., Jonker, M., Kotte, M., 2011. Sorption of various polar chemicals to different sorbents and membrane materials applied in passive sampling. SETAC Europe 23rd Annual Meeting Glasgow. pp. 308–309.

Smedes, F., Zande, T., Van Der, Roose, P., Davies, I.M., 2007. ICES Passive sampling trial survey for water and sediment (PSTS) 2006 – 2007. Part 3: Preliminary interpretation of field data. Proceedings ICES Annual Science Conference. pp. 1–29.

STOWA, 2016. Ecologische Sleutelfactor Toxiciteit. Hoofdrapport, deelrapporten en rekentools. https://www.stowa.nl/publicaties/ecologische-sleutelfactor-toxiciteit-hoofdrapport-deelrapporten-en-rekentools.

Sub-Group on Review of the Priority Substances List, 2011a. PAH-5-6-rings EQS dossier 2011.

Sub-Group on Review of the Priority Substances List, 2011b. Fluoranthene EQS dossier.

Van de Meent, D., De Zwart, D., Posthuma, L., 2018. 2018. Screening level estimation of release rates of currently used chemicals to the environment. Environ. Chem. Toxicol submitted.

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

Van Loon, W.M., Wijnker, F.G., Verwoerd, M.E., Hermens, J.L., 1996. Quantitative determination of total molar concentrations of bioaccumulatable organic micropollutants in water using C18 empore disk and molar detection techniques. Anal. Chem. 68, 2916–2926. https://doi.org/10.1021/ac951136w.

Van Straalen, N.M., 2002. Theory of Ecological Risk Assessment Based on Species Sensitivity Distribuions. In: Posthuma, L., Suter IIG., Traas, T.P. (Eds.), Species Sensitivity Distributions in Ecotoxicology. Lewis Publishers CRC Press Taylor & Francis, Boca Raton FL, USA, pp. 37–48.

Verhaar, H.J.M., Busser, F.J.M., Hermens, J.L.M., 1995. Surrogate Parameter for the Base-Line Toxicity Content of Contaminated Water - Simulating the Bioconcentration of Mixtures of Pollutants and Counting Molecules. Environ. Sci. Technol. 29, 726–734. https://doi.org/10.1021/es00003a021.

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