Hazard/Risk Assessment

Comparison of Substance-Based and Whole-Effluent Toxicity of Produced Water Discharges from Norwegian Offshore Oil and Gas Installations

Pepijn de Vries,a,* Robbert G. Jak,a and Tone K. Frostb

aWageningen Marine Research, Wageningen Universiteit en Research, Wageningen, The Netherlands
bEquinor Research Centre, Ranheim, Norway

Abstract: When assessing the environmental risks of offshore produced water discharges, it is key to properly assess the toxicity of this complex mixture. Toxicity can be assessed either through the application of whole-effluent toxicity (WET) testing or based on its substance-based chemical composition or both. In the present study, the toxicity assessed based on WET and substance-based was compared for 25 offshore produced water effluents collected for the Norwegian implementation of the Oslo–Paris convention risk-based assessment program. The objectives were, firstly, to examine the concurrence between toxicity estimates derived from these two lines of evidence; and, secondly, to evaluate whether toxicity of produced water discharges predicted from substance-based data is adequately addressed in comparison with ground truth reflected by WET. For both approaches, 50% hazardous concentrations (HC50s) were calculated. For at least 80% of the effluents the HC50s for the two approaches differed by less than a factor of 5. Differences found between the two approaches can be attributed to the uncertainty in the estimation of the concentration of production chemicals that strongly influences the substance-based estimated toxicity. By evaluating effluents on a case-by-case basis, additional causes were hypothesized. Risk management will particularly benefit from the strength of risk endpoints from both approaches by monitoring them periodically in conjunction over time. This way (in)consistencies in trends of both indicators can be evaluated and addressed. Environ Toxicol Chem 2022;41:2285–2304. © 2022 The Authors. Environmental Toxicology and Chemistry published by Wiley Periodicals LLC on behalf of SETAC.

Keywords: Aquatic toxicology; Whole-effluent toxicity; Species sensitivity distributions; Produced water; Hazardous concentration; Multisubstance potentially affected fraction

INTRODUCTION

Produced water and associated regulations

Produced water is a waste stream generated during production and processing of offshore oil and gas (Clark & Veil, 2009; dos Santos et al., 2014). It is a mixture containing residues from formation water and possibly injection water (Beyer et al., 2020; Neff et al., 2011; Scholten et al., 2000; Sørensen et al., 2019). It can also carry production chemicals which are added to facilitate the operational process. The differences in volume as well as the characteristics of produced water depend on many parameters including age of field, geographic location, reservoir type, and production technologies.

Traditionally, produced water discharges are managed by globally applied oil in water limits (Offshore Industry Committee, 2008; Oslo–Paris convention [OSPAR], 2014b; Smit et al., 2020; Zheng et al., 2016). However, risk-based approaches more specifically address environmental concerns of specific substances rather than these general oil in water limits. In Norway, this is addressed in the zero-harmful discharge program (Smit et al., 2011) and later by the OSPAR for the northeast Atlantic (OSPAR, 2014b).

The overall goal of the OSPAR risk-based approach is to assure and demonstrate that measures adequately reduce the environmental risk of produced water discharges (OSPAR, 2014b). Member states are allowed to assess the risk of produced water discharges using either data from whole-effluent toxicity (WET) tests, substance-based composition...
derived from produced water chemical characterization (also referred to as the substance-based approach), or a combination of the two approaches. How these first two lines of evidence should be combined in an overall risk management is not explicitly addressed by the guidelines.

Normally, the risk-based approach applies extrapolation to translate estimated toxicity into a protective environmental threshold value. In the present study, such extrapolation factors are omitted. This allows for a more straightforward comparison of toxicity without the introduction of uncertain factors that could hamper such comparison.

**Produced water composition and characterization**

Produced water contains oil-related constituents such as linear and cyclic alkanes and alkenes (aliphatic hydrocarbons), mono- and polyaromatic hydrocarbons, alkylphenols, and organic acids. It can also contain salts and heavy metals originating from formation water. The majority of aromatic hydrocarbons and aliphatic hydrocarbons (e.g., polycyclic aromatic hydrocarbon [PAHs] and alkylphenols) are present in the dispersed oil because of their low water solubility and hydrophobic properties, while the dissolved phase is dominated by monoaromatics (benzene, toluene, ethylbenzene, and xylene [BTEX]), phenols, and organic acids (Dórea et al., 2007; Neff et al., 2011; Røe Utvik, 1999; Zheng et al., 2016).

A wide variety of production chemicals are used to facilitate the production process including antifoams, biocides (to prevent fouling), corrosion inhibitors, and demulsifiers (Johnsen et al., 2004; Kelland, 2009, 2014; Lee & Neff, 2011; McCormack et al., 2001, 2002; Neff et al., 2011; Sørensen et al., 2019; Zheng et al., 2016). The chemical constituents and associated toxicity in these products are highly variable (Henderson et al., 1999).

Common organic and inorganic substances in produced waters and some commonly used production chemicals can be identified and quantified routinely by gas chromatography coupled to mass spectrometry (GC-MS) and liquid chromatography coupled to MS (LC-MS) or other routine techniques (Brown et al., 1992; Harman et al., 2011; Lee & Neff, 2011; Oetjen et al., 2017; Reed & Johnsen, 2012; Røe Utvik, 1999; Tibbetts et al., 1992).

However, produced waters form complex mixtures also containing a large number of lesser-known or unknown organic compounds. Identification and quantification of these substances are challenging, requiring more advanced methods of analysis (e.g., GC×GC-MS and higher-resolution MS [He et al., 2017; Melbye et al., 2009; Oetjen et al., 2017; Rowland et al., 2011; Thurman et al., 2017; Wang et al., 2013; Wilde et al., 2015]). Such techniques allow significantly higher chromatographic resolution and identification of polar substances but are not applied routinely.

**WET testing**

Direct toxicity or WET assessment of the effluent provides a direct measure of toxicity of mixtures including combined effects of multiple substances, both those known from chemical characterization and those of unknown identity, and accounts for interaction between substances (de Vlamming et al., 2000). For example, in the United States, WET assessments of onshore discharges to freshwater environments have been conducted since at least 1985 (US Environmental Protection Agency [USEPA], 1985, 1993, 1994, 1995) and later in Europe (see Maltby et al., 2000; Manusadzianas et al., 2003; Neamtu et al., 2009; Pessala et al., 2004; Tonkes et al., 1999). Also, WET assessments have been widely used for compliance monitoring of industry and sewage effluents to fresh- and marine waters (mainly coastal), for example, in the United States and Canada under the Clean Water Act (see Denton et al., 2000; de Vlamming et al., 2000). Furthermore, WET assessment has been applied to monitor effluent quality and to derive site-specific guidelines in Australia and New Zealand in both fresh- and marine waters (Adams et al., 2008; Van Dam & Chapman, 2001). In Australia and the Gulf of Mexico, WET testing is even a permit requirement for operations.

Corresponding practical guidance documents on WET assessment have also been developed for the marine environment in Europe (see Roex et al., 2007). However, experience with WET assessment of offshore effluents is limited (Girling, 1989; Somerville et al., 1987; Stremgren et al., 1995). A practical protocol was developed by Roex (2012) for whole-effluent assessment of produced water discharges from the offshore platforms.

Recently, substance- and WET-based risks of produced water discharges were determined and compared for 12 Australian offshore platforms by Parkerton et al. (2018). In 2009, a study was performed in OSPAR regional waters (De Vries & Karman, 2011). Karman and Smit (2019) evaluated the results from this pilot (complemented with more recent data) for produced water discharges in the Dutch sector focusing on WET results only. All of these studies focused on the risk (rather than hazard or toxicity) posed by produced water discharges, based on either WET testing or substance characterization or both.

Within OSPAR (Ferdinand et al., 2007), it is generally recognized that application of WET has an added value compared with the traditional substance-specific assessment, especially for effluents with complex composition. The substance-based approach utilizes information based on the chemical characterization of the effluent and is therefore limited to the known substances (characterized) in the effluent (de Vlamming et al., 2000) and ignores toxicant interactions (synergism/antagonism). Therewith, this approach enables identification and ranking of considered substances with highest potential toxicity to some extent. To link substances causally to the observed toxicity, further studies using iterative toxicity identification evaluation are required (Van Dam & Chapman, 2001). However, this is beyond the scope of the present study, and it would present additional challenges for samples that are degradable.

**Objectives**

The primary objective of the present study was to examine the concurrence between toxicity measured with WET and toxicity predicted with an substance-based approach. Also, we
evaluated whether the toxicity of produced water discharges predicted based on the multiple single substance data is adequately addressed in comparison with WET—this while assuming the WET represents the true toxicity of produced water effluent in the present study. To achieve a better understanding of the similarities and dissimilarities between the two approaches, an assessment of the toxicity estimates was performed in the present study, using acute WET and substance-based toxicity information directly without applying a precautionary assessment factor. Furthermore, the present study evaluated how well this toxicity can be estimated from the substance-based approach, to what extent differences between the two approaches can be explained, and whether certain produced water constituents serve as contributors to toxicity.

The composition of the individual naturally occurring substances (NOS) originates from the biannual sampling and chemical characterization of the produced water effluent. In the present study, these are used together with estimated discharge concentrations of production chemicals. Both were collected following Norwegian risk-based approach protocols. Insight into possible differences in the toxicity estimates resulting from the two lines of evidence is crucial because it forms the basis of risk-based assessment and management of produced water discharges. Therefore, it will help to improve the quality of the input data and refinement of the substance-based risk assessment undertaken by the operators. Moreover, it will provide insight on how to treat both lines of evidence in the context of environmental risk management in general.

METHODS

All produced water sampling and analyses were performed in parallel within the standard assessment program following OSPAR risk-based approach guidelines (OSPAR, 2014b; Roex, 2012). Therefore, the sampling and test protocols used in the present study were based on the recommendations in these guidelines.

Selection of and sampling at installations plus sample conservation, transport, and pretreatment

All installations with produced water effluents on the Norwegian continental shelf with an environmental impact factor (EIF) value ≥10 were included in the present study. The EIF is used as an environmental risk indicator for the management of produced water discharges on the Norwegian continental shelf (Smit et al., 2011). This resulted in a selection of 25 installations (and thus 25 effluents), of which 14 installations were operated by Equinor, the major Norwegian Oil and Gas producer on the Norwegian continental shelf (Figure 1). The remaining 11 installations were operated by five separate companies. Only one of the selected installations (Ula) produced solely gas, whereas one (Norne) produced a mixture of both oil and gas. All other platforms primarily produced oil (Supporting Information, Table S1). Sampling took place over a period between August 31, 2015, and March 6, 2017 (Supporting Information, Table S1).

WET sampling and sample pretreatment

Samples representative of the produced water being discharged were collected from the 25 offshore installations for WET testing and chemical characterization at the same event. The WET sampling protocols were followed (i.e., samples were not acidified for conservation; were kept at 0 °C to 5 °C and stored in darkness, where they are normally stable for 24–48 h; and the time between sampling offshore and start-up of toxicity testing in the laboratory [onshore] was mostly within 24 h for most effluents and did not exceed 48 h; see Supporting Information, Table S1), per Roex (2012). Samples for WET testing were collected in inert containers, rinsed with demineralized water, and then at least twice with the produced water (Roex, 2012). Containers were filled to the brim to avoid degradation by air. Additional physical and chemical parameters were measured in the produced water samples on arrival in the laboratory, prior to toxicity testing of the effluent. These included pH, ammonium, nitrite, conductivity, salinity, oxygen,
sulfide, and suspended particles. If any of these parameters were outside their optimal range in the assay, the sample was adjusted according to recommended ranges originating from Postma et al. (2002) prior to testing. This is because effects caused by exceedance of the optimal range for any of these parameters in each assay can have influence on the toxicity measured. Samples were adjusted before testing to maintain the oxygen content (by aeration), pH, and salinity (by adding either deionized water or NaCl) such that they were within the optimum range for each test protocol and species. Before any of these adjustments were made to samples that required them, the samples were first diluted (50%–70%) with clean filtered seawater to reach the preferred testing conditions (Supporting Information, Table S7). Produced water (effect) concentrations were always expressed as dilution percentages relative to its original volume prior to the dilution and adjustments.

Chemical characterization of effluents

Samples for chemical characterization of NOS were collected, in parallel to the samples taken for WET testing, in accordance with the Norwegian Oil and Gas Recommended Procedures and Methods for Sampling and Chemical Analysis of Natural Occurring Substances (Norwegian Oil and Gas, 2013).

WET

On arrival at the laboratory, each of the 25 produced water samples was tested for acute toxicity with three marine species (bacteria Vibrio fischeri, algae Skeletonema costatum, and copepods Acartia tonsa). These test species and test standards are prescribed by OSPAR guidelines (OSPAR, 2014b). The results of WET testing are expressed in terms of produced water dilution that corresponds to a given degree of effect (e.g., 50% response) for the endpoint being examined. For example, when the 48-h 50% effect concentration (EC50) equals 5%, it indicates that a 20-fold dilution of the produced water sample causes a 50% response to the test species and endpoint studied after 48-h exposure to a produced water dilution series. For V. fischeri, the 15-min EC50 (luminescence inhibition) was determined using the Microtox® acute toxicity test kit applying International Organization for Standardization (ISO) NS-EN ISO 11348-3:2008 (ISO, 2019) standards. For S. costatum the 72-h EC50 (growth inhibition) was determined using ISO 10253:2006(E; ISO, 2006) standards. Finally, the 48-h 50% lethal concentration (LC50, referred to as EC50 in the remainder of the text for brevity) was determined for A. tonsa using ISO 14669-1999 (ISO, 1999) standards (see Supporting Information, Tables S6–S8).

Five to seven dilution treatments (nine in the bacteria test) ranging from 70% to 1% effluent were used for toxicity evaluation. Detailed descriptions of test setup and conditions (number and type of control treatments, number of replicates, pH, oxygen levels, temperature, test durations) can be found in the respective test protocol. A brief overview is shown in Supporting Information, Table S6. For evaluation of the quality of toxicity tests (and results), acceptance criteria for each test protocol are provided (Supporting Information, Table S8) and have been used in the present study.

The 50% hazardous concentration (HC50) was used as a common endpoint for the comparison of toxicity in the present study, which in the WET-based approach combines the WET data of all test species (HC50WET). The HC50 is expressed as the dilution percentage of the original produced water effluent at which up to 50% of the species are affected at up to their EC50, derived from single-species laboratory acute toxicity tests. A high HC50 therefore indicates a low toxicity, and a low HC50 indicates a high toxicity. For each sample, the HC50 together with 5% and 95% percentiles were calculated from the EC50 values (expressed as percentage) for the three tested species following Aldenberg et al. (2002; essentially HC50WET is the geometric mean EC50 for the tested species). For convenient comparison with substance-based results and other studies, WET per installation is also expressed in toxic units (TU) using TUWET = 100%/HC50WET. This is also done for each test species effect endpoint by using TUspecific = 100%/EC50.

Substance-based toxicity

Recommendations from the OSPAR risk-based approach guidelines were followed to select the NOS to be included (OSPAR, 2014a; 2014b). These NOS include the monoaromatic hydrocarbons BTEX, aliphatic hydrocarbons (dispersed oil), PAHs, naphthalenes, alkyl phenols, and relevant metals (As, Cu, Cr, Zn, Ni, Pb, Cd, and Hg). In addition to the guidelines, specific organic acids (which are not considered under standard OSPAR guidelines) have been included. A detailed list of selected NOS substances in the produced water effluent is included in Supporting Information, Table S2.

Concentrations of NOS were determined by chemical analysis in accordance with Norwegian Oil and Gas (2013) using accredited methods (good laboratory practice) for all individual constituents. For most of the organic NOS (including BTEX, [alkyl] phenols, naphthalenes, and PAHs), GC-MS or GC with a flame ionization detector was applied, while ion chromatography was also applied for organic acids. For metals, different analytical methods were applied: atomic absorption spectroscopy (AAS)–based methods (e.g., AAS, hydride generation AA/S high-resolution AAS, cold vapor AAS), inductively coupled plasma (ICP) atomic emission spectroscopy (e.g., high-resolution ICP-MS, dynamic reaction cell ICP-MS), and ICP-MS. A more detailed description of applied techniques is available from the guidelines (Norwegian Oil and Gas, 2013), as well as minimum requirements for the limit of quantification (LOQ) for each substance (see Supporting Information, Table S2). For the metals included in the present study, the normal LOQ ranged from 0.1 μg/L (Hg) to 15 μg/L (Zn) at the highest. For (alkyl) phenols, naphthalenes, and PAHs the maximum LOQ were 0.1 μg/L, 2 mg/L for organic acids, and 0.01–0.1 mg/L for BTEX.
Chemical constituents of production chemicals were not measured analytically. The confidential nature of many product formulations prohibits such analyses. Instead, their concentrations were estimated based on consumption amounts reported by the offshore industry following OSPAR (2020). For most organic production chemicals, estimates of discharged amounts of chemical constituents were based on the partitioning of the substances between water and n-octanol ($P_{OW}$) and, on the water cut, as a measure of the distribution of the chemical substances between oil and the water phase in produced water (Aas and others, 2002). The $P_{OW}$ data were provided by the suppliers of the product for all organic substances with the exception of surfactants, for which such extrapolation will not result in a reliable estimate (OSPAR, 2020). For these chemicals, reliable estimates may be generated from laboratory or field measurements. If not, precautionary assumptions are generally applied (e.g., 100% released to sea).

Discharge amounts of produced water and added production chemicals have been reported on a yearly basis in Norway, whereas their applied dosage varies. Some have been discharged continuously, whereas others have been intermittently. These reported chemical estimates may therefore not adequately reflect the actual discharge concentrations during sampling for WET and substance-based chemical characterization. Nevertheless, produced water discharge volumes have been monitored daily, and effort was made to obtain discharge consumption rates and estimations of discharge concentration of chemical constituents on daily and monthly bases for each installation. Therefore, four options for calculating the concentrations of production chemicals were applied in the present study: (1) daily estimates, based on the chemicals applied on the day of the sampling; (2) monthly estimates, based on the amount applied in the month of the sampling; (3) yearly average estimates; (4) disregarding production chemicals altogether. Note that the fourth option was not realistic (because it omits relevant constituents) and has only been included as a reference to be compared with the other options.

One group of production chemicals has been defined by OSPAR to pose little or no risk (PLONOR) to the environment (OSPAR, 2013; updated in 2021). Although these chemical substances have been considered nonhazardous, they may have contributed to the overall toxicity if discharged in large quantities and have thus been present in high concentrations in the effluent. The two most common PLONORs in this category were monoethylene glycol (MEG) and methanol and were accordingly included in the comparison. Because the chemical composition of the chemical products is confidential information (owned by the chemical suppliers), the product constituent identities were anonymized in the present study.

Because toxicity data availability was limited, some substances were grouped following the recommendations from OSPAR (2014a). Grouped substances are listed in Supporting Information, Table S3. For each substance, discharge concentrations were converted into micromoles per liter and then summed to obtain the total concentration of the group (assuming additivity of toxicity).

The calculation of the substance-based HC50 required information on the toxicity of each substance (group). Therefore, data collection focused on similar test species, endpoints, and duration for optimal comparability with the information collected for the WET-based HC50. For this purpose, similar acute EC50 values were collected from a number of toxicity databases, from which a specific subset of data was selected, which best reflect the test conditions and for each of the three species tested using the WET approach.

For the constituents of the production chemicals, EC50s were taken from the Harmonised Offshore Chemical Notiﬁcation Format (HOCNF) accessed from the NEMS chemicals database (https://www.n-ems.com/nems-chemicals). The HOCNF provided information on acute EC50 values for (standard) marine microalgae and crustacean species. It did not include information on bacterial toxicity, whereas bacterial toxicity was tested for the WET-based approach in the present study. Therefore, the marine bacteria (V. fischeri) toxicity was derived from the algal EC50 data that were available from the HOCNF, using interspecies correlation estimates (see Raimondo and Barron, 2020). For this purpose, data from Zhang et al. (2010) were used, which resulted in the relationship

\[
\ln \left( \frac{1}{EC50_{V.fischeri}} \right) = 0.4758 - 0.7016 \ln \left( \frac{1}{EC50_{algae}} \right) (r^2 = 0.61).
\]

For all NOS and the PLONORs, MEG and methanol (because no HOCNF was available for these production chemicals), acute algal and crustacean EC50 toxicity information was collected from the USEPA ECOTOX database (https://cfpub.epa.gov/ecotox/, release of December 15, 2015). Because this database does not contain bacterial EC50 data, this information was extracted from the ECETOC Aquatic Toxicity Database developed in 2003. When the latter database did not contain information on a required substance (group) the inter-species correlation estimate was used. Because data on PAHs are scarce and variable (De Vries and Jak, 2018), EC50 data from the databases above were supplemented with information from Verbruggen (2012).

The procedure outlined in Figure 2 was applied to calculate the substance-based HC50 as an overall toxicity indicator based on the discharge concentrations and the three geometric mean EC50 values for all substance groups, which conform with those of De Zwart and Posthuma (2005) and Smit and others (2008).

A final selection of EC50 data was made for each substance group by applying a set of criteria (Supporting Information, Table S4 and accompanying text). These criteria were based on a balance between similarity with WET test species and experimental conditions and data availability. This will result in a better representation of the tested assemblage (Dowse and others, 2013). Based on the final selection of EC50 values for each combination of substance group and species group, the geometric EC50 value was derived.

Then for each substance group $i$, the geometric mean value of these three species groups was taken and defined as $10^{X_{mi}}$ (labeled benchmark in Figure 2). For each substance group, the toxic unit (TU) is calculated by dividing the substance group effluent concentration ($C_{i}$) by this benchmark: $TU = C_{i}/10^{X_{mi}}$ (Step II in Figure 2). Each substance group has been assigned to a specific toxic mode of action (Step I in Figure 2, see also...
Table 1). For each toxic mode of action (labeled \( j \)), the toxic units were summed: \( TU_j = \sum_{i} TU_{ij} \) (Step III in Figure 2).

Next, it was assumed that each toxic mode of action acts independently, and the toxicity for each toxic mode of action could be represented by a generic species sensitivity distribution (SSD; Posthuma et al., 2002). For each toxic mode of action \( j \), it was assumed that \( TU_j \) is at the center of its SSD curve, and the steepness could be represented by a generic slope associated with the toxic mode of action (Table 2). The multisubstance PAF (msPAF) is then calculated from the individual PAFs. The dilution factor \( d \) is iteratively adjusted until msPAF = 0.5. In that case, HC50 = 100%/d. Roman numerals are used to label steps in the procedure and are explained in more detail in the main text.

**TABLE 1:** Results of analyses of variance, where the response variable is the perpendicular distance of 50% hazardous concentration substance-based (HC50SB) and HC50WET data from the identity relation (y = x)

| Option | Parameter | df | Sum of squares | Explained variance (%) | Mean of squares | \( F \) | \( Pr(>F) \) |
|--------|-----------|----|----------------|------------------------|----------------|------|-----------|
| (1) Day | PC1 (PAH 2–3 rings; naphthalenes; PAH 4+ rings) | 1 | 0.34 | 12% | 0.344 | 4.9 | 0.041 |
| (1) Day | PC2 (arsenic; zinc; lead) | 1 | 0.05 | 2% | 0.054 | 0.8 | 0.392 |
| (1) Day | PC3 (phoenol C0–C3; BTEX; organic acids) | 1 | 0.26 | 9% | 0.257 | 3.6 | 0.073 |
| (1) Day | dg | 2 | 0.80 | 27% | 0.400 | 5.7 | 0.012 |
| (1) Day | Operator | 1 | 0.21 | 7% | 0.208 | 2.9 | 0.103 |
| (1) Day | Residuals | 18 | 1.27 | 43% | 0.071 | – | – |
| (2) Month | PC1 (PAH 2–3 rings; naphthalenes; PAH 4+ rings) | 1 | 0.07 | 3% | 0.069 | 1.9 | 0.185 |
| (2) Month | PC2 (phoenol C0–C3; BTEX; production chemicals) | 1 | 0.57 | 24% | 0.517 | 15.8 | 0.001 |
| (2) Month | PC3 (cadmium; arsenic; mercury) | 1 | 0.19 | 8% | 0.193 | 5.3 | 0.033 |
| (2) Month | dg | 2 | 0.67 | 28% | 0.333 | 9.2 | 0.002 |
| (2) Month | Operator | 1 | 0.25 | 10% | 0.248 | 6.8 | 0.018 |
| (2) Month | Residuals | 18 | 0.65 | 27% | 0.036 | – | – |
| (3) Year | PC1 (PAH 2–3 rings; naphthalenes; PAH 4+ rings) | 1 | 0.26 | 9% | 0.261 | 4.3 | 0.053 |
| (3) Year | PC2 (phoenol C0–C3; BTEX; production chemicals) | 1 | 0.66 | 23% | 0.661 | 10.9 | 0.004 |
| (3) Year | PC3 (cadmium; arsenic; mercury) | 1 | 0.21 | 7% | 0.211 | 3.4 | 0.082 |
| (3) Year | dg | 2 | 0.48 | 16% | 0.48 | 3.9 | 0.039 |
| (3) Year | Operator | 1 | 0.24 | 8% | 0.24 | 3.9 | 0.064 |
| (3) Year | Residuals | 18 | 1.10 | 37% | 0.061 | – | – |

Analyses were performed separately for three out of the four options for handling added chemicals. Each of the principal components is followled by the top three ranking substance groups that best reflect the principal component.

WET = whole-effluent toxicity; df = degrees of freedom; \( Pr(>F) \), probability of obtaining an F value, considered significant when <0.05 (text in bold face); PC = principal component; PAH = polycyclic aromatic hydrocarbon; BTEX = benzene, toluene, ethylbenzene, and xylene; dg = dominant substance group.
slope parameter, $S_{m_j}$. The values of $S_{m_j}$ for each toxic mode of action $j$ were taken from Harbers et al. (2006) or based on data from Crommentuijn et al. (1997) or the National Institute of Public Health and the Environment (2005) when not available from the former (Table 2). The SSD curve for each toxic mode of action was thus defined as follows (Step IV in Figure 2):

$$PAF_j(d) = \text{pnorm}\left(\log_{10}\left(\frac{\text{TU}_j}{d}\right), 0, S_{m_j}\right)$$

(1)

In Equation 1, $PAF_j$ has been defined as the SSD response defined as the potentially affected fraction of species for toxic mode of action $j$; pnorm has been defined as the cumulative normal distribution expressed as a function of the parameters quantile $\left(\log_{10}\left(\frac{\text{TU}_j}{d}\right)\right)$, mean (in this case 0), and standard deviation (in this case the slope parameter $S_{m_j}$); and $d$ has been defined as the dilution factor by which the effluent needs to be diluted to reach $PAF_j$. The multisubstance PAF (msPAF) of the entire effluent has been expressed as follows (Step V in Figure 2):

$$\text{msPAF}(d) = 1 - \prod_j (1 - PAF_j(d))$$

(2)

The HC50 of the effluent was defined as the dilution percentage of the effluent required to reach $msPAF(d) = 0.5$. The HC50 value could thus be obtained by modifying the value of $d$ in the equations above until $msPAF(d) = 0.5$ (where HC50 = 100%/d). This was achieved numerically with a one-dimensional root (zero) finding algorithm based on Brent (1973) implemented in R (R Foundation for Statistical Computing, 2013; Step VI in Figure 2).

**Comparison of HC50**

The substance-based and WET-based toxicity (HC50$_{SB}$ and HC50$_{WET}$, respectively) were initially compared visually by scatter plotting both values against each other for all installations. When the toxicity indicator for each approach would have been identical, both HC50$_{SB}$ and HC50$_{WET}$ were positioned on the diagonal line $y = x$. Therefore, for further analysis of each installation the signed perpendicular distance (on a double logarithmic scale) from this diagonal was calculated:

$$\text{dist} = \left|\frac{1}{\sqrt{2}} \log_{10} \frac{HC50_{SB}}{HC50_{WET}}\right|$$

(3)

To explain differences present in the two lines of evidence, dist was used as a response variable in a linear regression model:

$$\text{dist} = \beta_0 + \beta_1 \times \text{PC1} + \beta_2 \times \text{PC2} + \beta_3 \times \text{PC3} + \beta_4 \times \text{dg} \times \beta_5 \times \text{operator}$$

(4)

In Equation 4, $\beta_k$ was used to represent regression intercept and slope parameters that were fitted, and PC1, PC2, and PC3 represent the first three principal components, obtained from the TU$s calculated for each sample. For the principal component analyses, the TU values were log-transformed, then scaled and normalized (such that the mean value was 0 and the standard deviation 1). The principal components were calculated using the “prcomp” function from the stats package shipped with R (R Foundation for Statistical Computing, 2013). In essence, these principal components reflected differences in how contributions to toxicity in each sample are composed. Nonnumerical explanatory variables were dominant substance group $dg$ and operator. The first $dg$ represented the substance group that has a dominating contribution to toxicity (i.e., largest sum of TU values). The latter represented the company that operates the installation, where only Equinor and others (grouped) were distinguished. The other operators were pooled because otherwise there were insufficient samples per operator for statistical analyses.

| TMoA                | TMoA abbreviation | $S_{m_j}$ ($\log_{10}$) | Substance group(s) associated with TMoA | $S_{m_j}$ (based on data) from |
|--------------------|------------------|-------------------------|----------------------------------------|-------------------------------|
| Nonpolar narcosis  | NN               | 0.65                    | Aliphatic hydrocarbons; BTEX; MEG; methanol; naphtalene; organic acids; PAH (2–3 rings); PAH (4+ rings) | Harbers et al. (2006)         |
| Polar narcosis     | PN               | 0.58                    | Phenol (C0–C3); phenol (C4–C5); phenol (C6+) | Harbers et al. (2006)         |
| Unknown*           | UN               | 0.85                    | Production chemical (additives)        | Harbers et al. (2006)         |
| Zinc               | Zn               | 0.98                    | Zinc                                    | National Institute of Public Health and the Environment (RIVM; 2005) |
| Copper             | Cu               | 0.98                    | Copper                                  | RIVM (2005)                   |
| Nickel             | Ni               | 2.25                    | Nickel                                  | RIVM (2005)                   |
| Cadmium            | Cd               | 1.12                    | Cadmium                                 | RIVM (2005)                   |
| Lead               | Pb               | 0.98                    | Lead                                    | RIVM (2005)                   |
| Mercury            | Hg               | 0.83                    | Mercury                                 | RIVM (2005)                   |
| Arsenic            | As               | 0.98                    | Arsenic                                 | Crommentuijn et al. (1997)    |
| Chromium           | Cr               | 0.91                    | Chromium                                | Crommentuijn et al. (1997)    |

*The toxic mode of action labeled as “unknown” (UN) was assigned to all substances for which the toxic mode of action was not known (all production chemical additives except monoethylene glycol and methanol). The slope for this group (with an unknown toxic mode of action) was based on the median slope for all substances (with all toxic modes of action) as collected by Harbers et al. (2006).

TMoA = toxic mode of action; $S_{m_j}$ = slope parameter; BTEX = benzene, toluene, ethylbenzene, xylene; MEG = monoethylene glycol; PAH = polycyclic aromatic hydrocarbon.
An analysis of variance (ANOVA) was applied to this linear regression model to establish the amount of variance explained by the variables and their significance. The ANOVA was performed separately for each option (for handling production chemicals). This analysis was not performed for the option in which production chemicals were omitted.

Validity of the ANOVA was evaluated by performing Shapiro-Wilk tests for normality of the residuals. The association between nonnumerical variables was evaluated with χ² statistics and their homogeneity with Bartlett’s test (Bartlett & Fowler, 1937), all using R.

RESULTS

The present study focuses on the comparison of the toxicity endpoints (HC50₅₀ and HC50₆₇). Raw data underpinning these toxicity indicators are provided as Supporting Information. These include measured NOS discharge concentrations (Supporting Information), estimated discharge concentrations of production chemicals (Supporting Information), toxicity data (EC50) as selected for the toxicity calculations (Supporting Information, Table S5), toxicity (EC50) measured for the whole effluent (WET; Supporting Information, Table S9), and physical parameters of the effluent samples on arrival at the laboratory prior to WET tests (Supporting Information, Table S10). Toxicity data aggregated per substance group and biota are provided in the main text (Table 3). A summary of the concentrations of individual NOS in the effluents is given in Table 4.

The algae S. costatum was found to be the most sensitive species in the WET tests on Norwegian continental shelf, for 68% of the installations, whereas for the crustacean A. tonsa this was 20% and for the bacterium V. fischeri this was 12%. Figure 3 and Supporting Information, Table S12 show TU values for each of the four options for addressing production chemicals per installation (n = 25) stacked for each substance group (i). The TUₚₑₛ for each test species across all effluents is also shown and ranged from 1.5 to >100 (geometric mean of 15). The S. costatum EC50 of four effluents was <1% (TUₚₑₛ > 100). For effluents with EC50s <1%, the value was divided by a correction factor of 2 (giving EC50 = 0.5% and TUₚₑₛ = 200), used further in estimation of toxic units. Using

### Table 3: Toxicity data per substance group

| Substance group          | Algae: log₁₀ EC50 | Crustacea: log₁₀ EC50 | Bacteria: log₁₀ EC50 |
|--------------------------|-------------------|-----------------------|----------------------|
|                          | Log₁₀ (µmol/L ± SD (n)) | Indicative (µg/L) | Log₁₀ (µmol/L ± SD (n)) | Indicative (µg/L) | Log₁₀ (µmol/L ± SD (n)) | Indicative (µg/L) |
| Aliphatic hydrocarbons    |                   |                      |                      |                   |                      |                      |
| BTEX                     | 1.81 ± 0.10 (4)   | 3.77                  | 3.02 ± 0.55 (10)    | 4.98              | 2.45 ± 0.78 (2)      | 4.41                  |
| Naphthalenes             | 1.19 (1)          | 3.29                  | 1.43 ± 0.21 (5)     | 3.54              | 1.04 ± 0.25 (7)      | 3.15                  |
| Organic acid             | 3.09 ± 0.00 (2)   | 4.87                  | 3.49 ± 0.70 (31)    | 5.27              | 2.24 ± 0.04 (6)      | 4.02                  |
| PAH 2–3 rings            | 0.787 ± 1.48 (9)  | 3.01                  | 0.686 ± 0.31 (5)    | 2.91              | 0.639 ± 0.69 (38)    | 2.86                  |
| PAH 4+ rings             | 2.03 ± 4.46 (2)   | 4.34                  | −0.267 ± 0.27 (7)   | 2.04              | 1.59 ± 0.87 (18)     | 3.90                  |
| Phenol C0–C3             | 2.72 ± 0.00 (2)   | 4.69                  | 2.6 ± 0.38 (6)      | 4.57              | 2.43 ± 0.15 (3)      | 4.40                  |
| Phenol C4–C5             | 1.2 (1)           | 3.38                  | 2.17 (1)            | 4.35              | 1.05                 | 3.23                  |
| Phenol C6+               | −0.168 (1)        | 2.15                  | 0.309 (1)           | 2.62              | 0.0885               | 2.40                  |
| Monoethyleneglycol       | 5.68 (1)          | 7.47                  | 5.67 ± 0.32 (31)    | 7.46              | 6.26 (1)             | 8.05                  |
| Methanol                 | 5.49 (1)          | 7.00                  | 5.57 (2)            | 7.08              | 6.11 ± 0.21 (2)      | 7.62                  |
| Arsenic                  | 1.25 ± 1.12 (6)   | 3.12                  | 0.83 (2)            | 2.71              | 1.09                 | 2.96                  |
| Cadmium                  | 0.108 (1)         | 2.16                  | 0.102 ± 0.34 (9)    | 2.15              | 2.32 (1)             | 4.37                  |
| Chromium                 | 1.06 ± 0.41 (5)   | 2.78                  | 2.42 ± 0.13 (6)     | 4.14              | 2.75 ± 0.47 (2)      | 4.47                  |
| Copper                   | 0.591 (1)         | 2.39                  | 0.233 (1)           | 2.04              | 1.17 (1)             | 2.97                  |
| Lead                     | −1.03 (1)         | 1.29                  | 0.508 (2)           | 2.82              | −0.17 (1)            | 2.14                  |
| Mercury                  | −0.413 ± 0.32 (10)| 1.89                  | −1.12 ± 0.09 (9)    | 1.18              | −0.461 ± 0.34 (2)    | 1.84                  |
| Nickel                   | 1.61 ± 0.83 (10)  | 3.38                  | 2.01 (1)            | 3.78              | 1.34                 | 3.11                  |
| Zinc                     | 0.337 (1)         | 2.15                  | 1.45 (1)            | 3.27              | 2.58 (1)             | 4.40                  |

Numbers are mean log₁₀-transformed 50% effect concentrations (EC50s) expressed in micromoles per liter. The plus–minus sign shows the standard deviation of the log₁₀-transformed EC50 values (only shown when n > 1). Between parentheses is the number of records on which the mean is based. This number is missing when the EC50 is based on an interspecies correlation estimate. These toxicity values are used to calculate the toxic units. The geometric mean EC50 is also provided indicatively in log₁₀-transformed micrograms per liter, for easier comparison with other studies.

- Because molar mass varies within substance groups, the mass of the largest fraction is used as a representative to convert micromoles into micrograms.
- Toluene is used as a representative to convert micromoles into micrograms.
- Naphthalene is used as a representative to convert micromoles into micrograms.
- Acetic acid is used as a representative to convert micromoles into micrograms.
- Phenanthrene is used as a representative to convert micromoles into micrograms.
- Fluoranthene and pyrene are used as a representative to convert micromoles into micrograms.
- Phenol is used as a representative to convert micromoles into micrograms.
- Butyl phenol is used as a representative to convert micromoles into micrograms.
- Octyl phenol is used as a representative to convert micromoles into micrograms.
- BTEX = benzene, toluene, ethylbenzene, xylene; PAH = polycyclic aromatic hydrocarbon.
TABLE 4: Geometric mean discharge concentrations of naturally occurring substances measured in all samples (platforms)

| Substance                              | Substance group        | Discharge concentrations (µg/L) |
|----------------------------------------|------------------------|---------------------------------|
| Aliphatic hydrocarbons                 | Aliphatic HCs          | 1.2E+04 (2.8E+03–1.8E+05)       |
| Benzene                                | BTEX                   | 4.7E+03 (5.0E+02–1.7E+04)       |
| Toluene                                | BTEX                   | 4.4E+03 (1.2E+03–1.4E+04)       |
| Ethylbenzene                           | BTEX                   | 3.0E+02 (7.3E+01–6.0E+02)       |
| Sum xylene                             | BTEX                   | 1.8E+03 (4.4E+02–6.3E+03)       |
| Naphthalene (and alkyl homologs)       | Naphthalenes           | 8.9E+02 (2.7E+02–3.3E+03)       |
| Formic acid                            | Organic acid           | 1.7E+03 (1.0E+03–1.6E+04)       |
| Acetic acid                            | Organic acid           | 8.0E+04 (7.3E+03–7.0E+05)       |
| Propionic acid                         | Organic acid           | 8.8E+03 (1.0E+03–8.0E+04)       |
| Butanoic acid                          | Organic acid           | 2.1E+03 (1.0E+03–1.4E+04)       |
| Pentanoic acid                         | Organic acid           | 1.1E+03 (1.0E+03–3.0E+03)       |
| Hexanoic acid                          | Organic acid           | 1.0E+03 (1.0E+03–1.0E+03)       |
| Acenaphthene                           | PAH 2–3 rings          | 1.2E+00 (3.0E–02–6.6E+00)       |
| Acenaphthylene                         | PAH 2–3 rings          | 8.3E–01 (1.4E–01–3.0E+00)       |
| Anthracene                             | PAH 2–3 rings          | 1.8E–01 (1.0E–02–1.0E+00)       |
| Dibenzothiophene and alkyl homologs    | PAH 2–3 rings          | 2.2E+01 (4.2E+00–1.1E+02)       |
| Fluorene                               | PAH 2–3 rings          | 9.6E+00 (1.5E+00–4.3E+01)       |
| Phenanthrene (and alkyl homologs)      | PAH 2–3 rings          | 6.6E+01 (1.8E+01–4.6E+02)       |
| Benz[a]anthracene                      | PAH 4+ rings           | 8.0E–02 (1.0E–02–5.5E–01)       |
| Benzo[a]pyrene                         | PAH 4+ rings           | 2.2E–02 (5.0E–03–2.7E–01)       |
| Benzo[ghi]perylene                     | PAH 4+ rings           | 3.5E–02 (5.0E–03–6.5E–01)       |
| Benzo[b]fluoranthene                   | PAH 4+ rings           | 9.1E–02 (1.0E–02–9.9E–01)       |
| Benzo[k]fluoranthene                   | PAH 4+ rings           | 9.9E–03 (5.0E–03–7.3E–02)       |
| Chrysene                               | PAH 4+ rings           | 3.9E–01 (4.0E–02–2.4E+00)       |
| Dibenzo[a,h]anthracene                 | PAH 4+ rings           | 1.2E–02 (5.0E–03–2.5E–01)       |
| Fluoranthene                           | PAH 4+ rings           | 1.9E–01 (5.0E–03–1.7E+00)       |
| Indeno[1,2,3-cd]pyrene                  | PAH 4+ rings           | 8.8E–03 (5.0E–03–1.0E+01)       |
| Pyrene                                 | PAH 4+ rings           | 3.5E–01 (4.7E–02–3.7E+00)       |
| Phenol (C1–C3 alkyl phenols)           | Phenol C0–C3           | 2.4E+03 (2.0E+02–3.8E+04)       |
| Butylphenol (and other C4 alkyl phenols) | Phenol C4–C5         | 5.0E+01 (1.2E+01–2.0E+02)       |
| Pentylphenol (and other C5 alkyl phenols) | Phenol C4–C5        | 1.7E+01 (9.0E–01–9.6E+01)       |
| Octylphenol (and C6–C8 alkyl phenols)  | Phenol C6+             | 8.7E–01 (1.2E–01–4.6E+00)       |
| Nonylphenol (and other C9 alkyl phenols)| Phenol C6+            | 4.7E–02 (1.0E–02–1.9E–01)       |
| Arsenic                                | Arsenic                | 7.4E–01 (2.6E–02–1.1E+02)       |
| Cadmium                                | Cadmium                | 2.9E–02 (5.0E–03–6.0E–01)       |
| Chromium                               | Chromium               | 8.1E–01 (2.0E–01–1.1E+01)       |
| Copper                                 | Copper                 | 1.2E+00 (1.3E–01–1.9E+01)       |
| Lead                                   | Lead                   | 1.4E–01 (2.0E–02–2.2E+01)       |
| Mercury                                | Mercury                | 8.0E–03 (2.6E–05–1.8E–01)       |
| Nickel                                 | Nickel                 | 8.5E–01 (8.8E–02–5.4E+00)       |
| Zinc                                   | Zinc                   | 3.3E+00 (4.1E+01–5.1E+02)       |

Values in parentheses indicate minimum to maximum range. Individual measurements of each effluent are available as Supporting Information. PAH = polycyclic aromatic hydrocarbon.

this correction and combining the data from all tested species, the TU_{WET} per installation ranged from 3 up to 80 with a geometric mean of 15 (Supporting Information, Table S9; Figure 3). Note that toxic units are not the ultimate toxicity indicators used in the present study. For that purpose, the HC50 is used. Because these are aggregated per toxic mode of action, it is no longer possible to distinguish between the individual chemical constituents.

Figure 3 confirms that production chemicals form a major component in the toxicity estimate of produced water discharges in most produced water samples. Organic acids also contribute considerably to the acute toxicity of produced water discharges. At first glance of Figure 3 the option using monthly estimates of production chemicals nicely matches with the toxicity measured in the WET tests, resulting in similar patterns. There are exceptions, for instance, platform Valhall where the production chemical concentrations seem to be overestimated (assuming that all other constituents are assessed accurately).

In most cases, TU_{WET} > \sum TU (72%, 68%, 64%, and 96% of all installations for each respective option 1–4; see Supporting Information, Tables S9 and S12).

Note that TU and TU_{WET} are not the definitive toxicity indicators in the present study, which are HC50_{SB} and HC50_{WET}. Calculated HC50_{SB} and HC50_{WET} for each of the 25 platforms (sampled effluents) are depicted in Figure 4 (Supporting Information, Table S13). Note that the HC50 represents the dilution percentage where at least 50% of the species are exposed at or below their EC50. For example, when the HC50 = 5%, this means that the effluent needs to be diluted to at least 5% (i.e., a 20-fold dilution of the produced water sample) to affect less than half of the relevant species at or below the selected effect threshold (EC50 in the present study). Furthermore, the higher the HC50 value, the fewer species are potentially affected and the lower the mean toxicity. Each panel in Figure 4 shows the results for the four different options applied for estimating production chemical discharge.
concentrations in the substance-based approach. Note that the values for $HC50_{WET}$ (on the vertical axis) are identical in each panel; only the $HC50_{SB}$ values vary with each of these options.

Figure 4 visualizes how well the substance-based toxicity matches with the WET-based toxicity. For that purpose, there are two aspects that need to be considered: the distance from the diagonal line ($y = x$), which indicates systematic differences, and the magnitude of scattering (i.e., correlation), which reveals random differences. Figure 4 shows that when production chemicals are included (options 1–3) in the analysis, the ratio for the substance-based HC50 over the WET-based HC50 is, with the exception of one effluent (Balder; using a daily estimate of production chemicals), within an order of magnitude and in most cases even within half an order of magnitude (see also Supporting Information, Table S13). However, the confidence intervals of the $HC50_{WET}$ (and thus the ratios) may stretch over several orders of magnitude.

The ratio of $HC50_{SB}$ over $HC50_{WET}$ is visualized in Figure 5, which shows the cumulative percentage of platforms (samples) where its effluent is below a specific ratio. If the substance-based and WET-based toxicity values were similar, a median ratio of 1 would be expected (50% of the platforms are below and 50% of the platforms are above the ratio of 1). Figure 5 shows that when production chemicals are disregarded as constituents of the effluent (option 4) in the substance-based approach, the ratio shifts to the right. In that case, the median ratio of the $HC50_{SB}$ over $HC50_{WET}$ is well above 1 (~2.7), and thus the substance-based toxicity is generally less conservative than the WET-based toxicity ($HC50_{SB} > HC50_{WET}$) for 96% of the installations (Table 5), as can be expected when omitting relevant substances from the toxicity estimate. Note that this option was included as a reference only. For the other estimate options that do include production chemicals as constituents of the effluent, the median ratio is closer to, but still above, 1 (median between 1.3 and 1.4), where between 64% and 72% of the installations have $HC50_{SB} > HC50_{WET}$. This indicates that WET-based toxicity is more conservative than substance-based toxicity in most cases. Results from Figure 5 are further summarized in Table 5, where it is shown that (depending on the method for estimating production chemical discharge concentrations, excluding the reference option 4), at least 80% of the samples the HC50 of both approaches differs by a factor 5 or less. For at least 96% of the samples this difference is a factor of 10 or less (Table 5). These factors are selected to objectively qualify the differences between the substance-based and WET-based HC50.

It is also notable in Figure 5 that the distribution of ratios of $HC50_{SB}$ over $HC50_{WET}$ is skewed (Pearson’s moments of skewness are 0.9, −1.6, −0.9, and −0.8, respectively, for the option without production chemicals and the daily, monthly, and annual production chemical concentration estimates). For platforms where the ratio of $HC50_{SB}$ over $HC50_{WET}$ is less than 1 (indicating that substance-based toxicity is more conservative than the WET-based toxicity [Supporting Information, Table S14]), the ratio decreases more rapidly, differences between the three options become more apparent (larger), and a second inflection point appears in the S-shaped curves in Figure 4. Note that these platforms (at the left-hand side of the distribution in Figure 5) all have these production chemicals as the dominant contributor to the toxicity of the effluent, whereas on the right-hand tail of the distributions toxicity is more frequently dominated by organic acids.
Finally, the results from the ANOVAs (Table 1) show how the difference between HC50S_B and HC50WE_T (expressed as the perpendicular distance from the line $x = y$: dist from Equation 3) is explained by several variables. Positive distance values (dist $>0$) indicate values above the line $x = y$, and thus HC50SB > HC50WET (a lower substance-based toxicity compared with the WET-based toxicity), whereas negative values (dist $<0$) indicate the opposite (a lower WET-based toxicity compared with the substance-based toxicity). Because the distance is with both axes on a 10 base logarithmic scale, an absolute value of 1 indicates a difference of an order of magnitude (i.e., a factor of 10 on a linear scale).

When focusing only on the terms that significantly explain variance (probability of obtaining an $F$ value, Pr[$>F$] $< 0.05$) in the distance between the two approaches, the second option (using monthly production chemical concentration estimates) explains most variance (70%). This is opposed to option 1 and option 3, which both only explain 39% of the variance. The generally large portion of unexplained variance suggests a weak correlation between HC50S_B and HC50WET, as can also be seen in Figure 4 from the high amount of scattering.

The principal components (PC1–PC3, see also Supporting Information, Figure S2) reflect the chemical characterization of the effluent. Significant effects of these principal components in the ANOVA indicate that they are partly responsible for differences observed between the WET-based and substance-based toxicity. This in turn suggests that either WET or the substance-based approach does not adequately address the contribution of the substances represented by the principal components to the toxicity. Unfortunately, this cannot be pinned down to specific substances because substance concentrations in the effluents are correlated. Fortunately, each principal component does reflect a selection of substances. Table 1 shows which substance groups are reflected by which...
principal component. For option 2 (for which most variance is explained significantly, \( Pr[F] < 0.05 \)), the smaller phenols, BTEXs, production chemicals, and heavy metals seem to be responsible for the differences (Table 1); this is also the case for option 3. For the other options (1 and 3), PAHs and naphthalenes are also responsible (Table 1) but to a lesser degree.

The substance group that has a dominating contribution to the substance-based toxicity \((dg)\) also explains a significant part (16%–28%) of the variation for each of the three analyzed options (Table 1). It is obvious that the substance group that has a major contribution to toxicity (e.g., production chemicals) will have a considerable effect on the comparison of the two approaches. Finally, the explanatory variable operator significantly explained the difference between substance-based and WET-based toxicity for option 2 (explaining 10% of the variation).

**DISCUSSION**

In the present study, samples of discharged produced water from 25 Norwegian installations were collected, for which the toxicity was determined using two lines of evidence: one based on the substance composition (substance-based) and one based on the toxicity of the whole effluent (WET). Toxicity was in both cases expressed as HC50, and differences were compared using ANOVAs for different options of including production chemicals. There are several statistical, practical, and

![Figure 5](image)

**FIGURE 5:** The ratio of substance-based (SB) over whole-effluent toxicity (WET)-based 50% hazardous concentration (HC50; \( \frac{HC50_{SB}}{HC50_{WET}} \)) is sorted for each installation (represented by a marker) from low to high on the x-axis. The distance used in analyses of variance can be obtained from this ratio (on the x-axis) using Equation 3 in the main text. The cumulative percentage of installations is shown on the y-axis. If both substance-based and WET-based hazards are identical indicators, the median ratio would be 1 (i.e., 50% of the platforms would have a ratio of 1 or more). Different colors show the different options for handling chemical additive concentration estimates. Marker shapes (regardless of color) indicate which substance group dominates the substance-based hazard for the sample represented by the marker. The distinction of dominating substance group is not meaningful for the option where production chemicals are not included and therefore is not depicted with markers. The options that include the production chemicals show an additional inflection point in the curve with the samples where production chemicals dominate the hazard concentrated in the left-hand tail.

| Option | Effluents where \( HC50_{SB} \geq 0.5 \) and \( \leq 2 \) | Effluents where \( HC50_{SB} \geq 0.5 \) and \( > 5 \) | Effluents where \( HC50_{SB} \geq 0.1 \) and \( \leq 10 \) | Effluents where \( HC50_{SB} > HC50_{WET} \) |
|--------|-----------------|-----------------|-----------------|-----------------|
| (1) Day| 52% | 88% | 96% | 72% |
| (2) Month| 48% | 84% | 96% | 64% |
| (3) Year| 52% | 80% | 100% | 64% |
| (4) Excluded| 32% | 72% | 92% | 96% |

Results are shown for each of the 4 options for estimating the production chemical concentrations. HC50 = 50% hazardous concentrations; SB = substance-based; WET = whole-effluent toxicity.
methodological issues that need to be considered when interpreting the results. These are discussed below.

**Statistical considerations**

**Diagnosis of analyses.** The ANOVA demonstrated that large parts of the variation remain unexplained (27%–43%). In the following section, potential causes for the remaining variance are hypothesized. In the remainder of this section, the validity of the ANOVA model is discussed.

The representativity of the selected samples (installations) is important for the interpretation of the results. As stated in the methodology, all installations on the Norwegian continental shelf with an EIF $\geq$10 were selected, which means that the samples are representative for the higher risk installations in that area. In other areas, the age and composition of oil and gas fields (and therefore their produced water) may differ. This bias should be considered when the present study findings are extrapolated to other areas.

For the ANOVAs, it is important that the boundary conditions for such an analysis are met. One of the assumptions in an ANOVA is that the explanatory variables are independent (i.e., not correlated or associated). This is one of the reasons the substance composition is summarized using a principal component analysis. The major principal components are by definition not correlated.

Although conceptually the dominant substance group (dg from Equation 4) is dependent on the chemical composition (which is reflected by the principal components PC1, PC2, and PC3 from Equation 4), association was mostly not found between these explanatory variables (Supporting Information, Table S16). This is also the case for the explanatory variable operator, which has been found to be associated with some of the major principal components (Supporting Information, Table S17). This suggests that the composition of the effluent (or production chemical concentration estimates) may be different for the different operators. However, no association has been found between the variables operator and dominating substance group (dg from Equation 4; Supporting Information, Table S15).

In the ANOVA, residuals (deviation from the response variable that cannot be derived from the explanatory variables) are assumed to be normally distributed, as yet another boundary condition. There is no evidence for violation of this assumption because for none of the three ANOVAs (Table 1) could the null hypothesis of the Shapiro-Wilk test for normality be rejected (Supporting Information, Table S14). Another assumption in ANOVAs is that the variation in factorial variables is equivalent. However, after applying Bartlett’s test for equal variance, it is concluded that variance is not equivalent for all variables (see Supporting Information, Table S18 and Figures S1 and S2) as one of the boundary conditions of an ANOVA. Because not all assumptions for the ANOVAs are met, the results from the analyses should be interpreted with caution.

**Interpretation of analyses.** Keeping the caution above in mind, all produced water effluents on the Norwegian continental shelf (with EIF $\geq$10) are included (and the samples are representative for that selection). Therefore, differences in variance (although complicating the ANOVA) do indicate differences between groups of samples. The most notable difference in variance is seen for the samples in which production chemicals dominate the toxicity ($n = 13$, $n = 15$, $n = 17$ out of 25 samples for options 1, 2, and 3, respectively), where the variance is larger than in the other samples (Supporting Information, Figure S4 and Table S18). This suggests that the error in toxicity estimates is larger in the groups with greater variance, which makes sense because production chemical discharge concentrations are estimated based on usage and are not measured.

Overall, the production chemicals dominate the toxicity, as shown in Figure 3. This is as expected because installations included in the present study are those that have the highest risk potential (EIF $\geq$10) on the Norwegian continental shelf, are represented by oil-producing installations (except one), and are dominated by installations that are late in their production phase and therefore produce larger volumes of produced water. Therefore, most of them also consume and discharge larger amounts of production chemicals. There are uncertainties in estimations of discharge concentrations of production chemicals, which is also reflected by the different options analyzed in the present study. In particular, released fractions of surface-active chemicals (surfactants) cannot be estimated reliably. Conservative estimates are therefore applied instead. For several installations this has probably resulted in an overestimation of annual discharge concentrations (option 3) and consequently in an overestimation of the overall toxicity (and the contribution to toxicity) of produced water effluents. This can explain why relatively many effluents (ranging from 13 to 17 out of 25 depending on the analysis option) have production chemicals dominating the toxicity. It can also explain why substance based toxicity more frequently overestimates the toxicity for those effluents when compared to WET-based toxicity.

This is further supported by the observation that when production chemicals are omitted the distance between HC50$_{B}$ and HC50$_{WET}$ is skewed to the right, which, as expected, indicates that the substance-based toxicity is underestimated (Figure 5). In contrast, for each of the other options the same results are skewed to the left (Figure 5). This suggests that each of these options overestimates the toxicity caused by these chemicals.

The pooled operators (marked other) show more variance in their substance-based toxicity estimates compared to WET toxicity (Supporting Information, Table S18 and Figure S3, particularly options 2 and 3) when compared to the Equinor installations. Also, in the pooled group substance-based toxicity is more frequently higher than the WET-based toxicity compared to the Equinor installations (Table 1; Supporting Information, Figure S3). Assuming that the composition of NOS in both groups is comparable, these results suggest that these differences are caused by uncertainties in production chemical concentration estimation and variable precautionary assumptions among operators. However, it should be noted that there might be unknown/latent variables associated with the operators.
Interestingly, in Figure 4, it is shown that for option 4 where production chemicals are omitted constituents of the effluent in the prediction of substance-based toxicity, the WET-based toxicity is greater than the substance-based toxicity for all effluents (HC50sub > HC50WET), except one (Valhall). In the case of Valhall, the toxicity estimate is dominated by organic acids in option 4 (Figure 3). In fact, it has the highest contribution of organic acids to substance-based toxicity in comparison to all other effluents (Figure 3). Because especially the short-chained carboxylic acids are volatile, the toxicity may have changed after aeration in the WET test. This could also be the case for other effluents.

**Practical and methodological considerations**

All samples were aerated before WET testing (Supporting Information, Table S10). As postulated above for Valhall, aeration could have removed volatile toxicants (e.g., BTEX, phenol) from and stimulate degradation of substances in the mixture, potentially underestimating toxicity in the WET tests. Other adjustment can affect bioavailability of the toxicants (e.g., pH). Therefore, sample pretreatment and adjustment prior to WET testing were kept at a minimum to make deviations of the samples to the chemical analysis as small as possible. Nonetheless, all samples needed adjustment to some extent (Supporting Information, Table S10). The methodological challenges in the WET assessment may have influenced the measured toxicity in both directions, dependent on the test species and the effluent composition.

In addition, boundary conditions (acceptance criteria) set in their respective toxicity testing protocols are not always met (Supporting Information, Table S11 and Figure S1). For 19 effluents (76%), it was observed that the acceptance criteria were not met for one or more test, where for 16 of the effluents (64%) this potentially affected the test endpoint for one or two test species (Supporting Information, Table S11). For those tests, WET may have been overestimated. However, because this did not have an effect on the statistical analyses, all samples were included in the present study (Supporting Information, Figure S1). Despite the lack of statistical significance, failure to meet acceptance criteria should be considered when evaluating effluent toxicity case by case for each effluent.

Only three species were tested for each sample, generating a limited representativeness for estimation of WET toxicity in the present study. A larger number of species would improve the certainty of WET toxicity estimates. When working with percentile hazardous concentrations (HCx), this number (and the inclusion of bacteria) is generally not sufficient for conservative hazard assessment (see Wheeler et al., 2002). However, conservative assessment was not the objective in the present study.

There are also uncertainties associated with the substance-based approach. The approach characterizes only known substances, commonly associated with produced water, but is not exhaustive. This means that unknown constituents possibly present in the produced water effluent are ignored, and therefore have not been evaluated in the present study. These substances potentially contribute to the overall toxicity of the effluent and can result in underestimation of the substance-based toxicity.

Naphthenic acids are gaining attention in that respect and are potentially relevant. They are classified as carboxylic monooacids containing a cyclo-aliphatic structure. Nearly all crude oils contain some naphthenic acids. They are highly soluble in water, much more so than other organics in produced water. In addition, naphthenic acids biodegrade slower in comparison with other organic acids. In the past, naphthenic acids have been isolated from the effluent of installation Heidrun and were measured at concentrations of roughly 50 mg/L (Mediaas et al., 2003). Similar levels have been found in produced water more recently as well (Bertheussen et al., 2021; Samanipour et al., 2020). Toxicity to aquatic species justified considering naphthenic acids as a produced water constituent of concern (Jones et al., 2011; Thomas et al., 2009). Extracts of naphthenic acid from produced water were shown to have an EC50 of 23 mg/L for S. costatum and 6.4 mg/L for A. tonsa (Frost et al., 2003). Assuming that naphthenic acid concentrations in the present study are in the same range as those measured in the past (Bertheussen et al., 2021; Mediaas et al., 2003; Samanipour et al., 2020), they could have contributed to the toxicity observed in the present WET tests. Recently, a standardized method for routine chemical characterization of the total concentration of naphthenic acids (as a group) in produced water effluents was developed (Bertheussen et al., 2021) and is currently implemented in the biannual analysis of produced water effluents on the Norwegian continental shelf. This was not yet the case in the present study and was therefore not addressed in the substance-based approach.

Sørensen et al. (2019) demonstrated using high-resolution analytical techniques in combination with acute toxicity testing that for some produced water effluents on the Norwegian continental shelf potentially “unknown” contributors to produced water toxicity toward, for example, A. tonsa could be attributed mainly to components in the polar substance fraction, whereas the toxicity of other produced water extracts was attributed mainly to the nonpolar substance fractions. A range of previously uncharacterized produced water polar and non-polar constituents, including those likely derived from production chemicals, such as trithiolane, imadazolines, and “quats,” were tentatively identified using advanced analytical techniques, including GC × GC-MS and high-resolution LC-MS.

Another concern is the way toxicity information was used to estimate the substance-based toxicity of produced water. For some combinations of species and substances, no toxicity data are available. For this purpose, substances were grouped into chemically similar groups. For these groups, acute toxicity of the substances can be highly variable, as is the case for PAHs and the aliphatic hydrocarbons, where toxicity is associated with carbon chain length and structure (see Redman et al., 2012). Other uncertainties may originate from the methodological process and associated assumptions, such as the classification of substances to specific toxic mode of actions and their associated parameters (Table 2).
In addition, even after grouping of substances, surrogate test species had to be selected because toxicity data for the preferred species, that is, those used in the WET tests, were not available (Supporting Information, Tables S3–S5). When it is assumed that the sensitivity of these surrogates deviates randomly in both directions compared to the target species sensitivity, it should only affect the sensitivity of the analyses but not the outcome. Unfortunately, in reality this error is not random. Hendriks et al. (2013) showed that in data-deficient situations toxicity is most likely underestimated. Although this effect is less than an order of magnitude, it contributes to discrepancies between the WET and substance-based results.

**Comparison with literature**

Karman and Smit (2019) studied the risk of produced water discharges from Dutch installations based on WET. Tested species as well as measured WET correspond well with the present study. The EC50 of the most sensitive species ranged from 1.4% to 16.9% of the effluent in Karman and Smit (2019), where bacteria, algae, and crustacea were tested. In the present study, EC50 of the most sensitive species ranged from <1.0% to 15%. In the present study, algae were found to be the most sensitive species in 68% of the samples. This was only the case in 16% of the samples analyzed by Karman and Smit (2019). Possibly this is attributable to a different composition of their samples: They collected samples from Dutch installations which mainly produce gas, whereas the present study focuses on Norwegian installations which mainly produce oil. In addition, they tested a generally less sensitive alga, Phaeodactylum tricornutum, compared to the algae (S. costatum) tested in the present study (USEPA, 2022).

Parkerton et al. (2018) address environmental risk of produced water discharges of 12 offshore installations in Australia. In contrast to Karman and Smit (2019), Parkerton et al. (2018) not only calculated risk based on WET but also assessed substance-based risk. The latter was done by a prediction of msPAF based on substances targeted for analytical characterization. Furthermore, their study also omitted precautionary assessment factors in the evaluation of the risk. In that respect, the study by Parkerton et al. (2018) is relevant to the present study. Nevertheless, there are several (methodological) dissimilarities between these studies, which are summarized in Table 6.

An important difference is that Parkerton et al. (2018) determined the environmental concentrations by dilution modeling (ignoring loss properties) and not discharge concentrations as in the present study. Furthermore, they used (sub)chronic toxicity data of specific test species (EC10) from their WET tests and compared those to acute EC50/LC50 of nonspecific test species in their substance-based approach. Despite these differences, our findings are discussed in relation to the findings of Parkerton et al. (2018).

Parkerton et al. (2018) identified sulfide as an important constituent of produced water discharges. The sulfide concentrations in the effluents of the present study were measured at arrival at the onshore laboratories, using a test kit with a high LOQ (10 mg/L). Although all the present samples had levels below this LOQ (Supporting Information, Table S10), the log mean acute EC50 reported by Parkerton et al. (2018) is considerably lower (0.02 mg/L) than this LOQ. The role of sulfides in the present study is therefore inconclusive. Parkerton et al. (2018) did report a negative correlation between sulfide and aliphatic hydrocarbon levels, which makes it difficult to derive a causal relationship between either constituent and their contribution to the overall toxicity. This is why in the present study principal component analyses were used to eradicate such correlations between constituents.

**TABLE 6: Summarizing overview of methodological differences between the present study and Parkerton et al. (2018)**

| Aspect                          | Present study                                      | Parkerton et al. (2018)                        |
|--------------------------------|---------------------------------------------------|-----------------------------------------------|
| Samples                        | Norwegian continental shelf; 25 installations:    | Bass Strait, Australia; 12 installations: type of producers not indicated |
|                                | mainly oil producers                               | Based on PETROTOX model for nonspecific species using fractions of hydrocarbons |
| Hydrocarbon toxicity           | Geometric mean of aliphatic hydrocarbon representatives for target species | Acute toxicity for nonspecific species          |
| Toxicity in SB approach        | Acute toxicity for target species similar to those tested for WET | Five (sub)chronic and one acute test (EC10 and EC50 available) |
| WET testing                    | Three acute tests (EC50)                          | Mean of five (sub)chronic tests (EC10)         |
| Estimation of WET-based TU     | Geometric mean of three acute tests (EC50)        | Measured limited specific constituents, ignored the remainder |
| Production chemical concentrations | Estimated all constituents from reported usage    | Not included                                   |
| Organic acids                  | Included                                           | Included                                       |
| Ammonia                        | Measured ammonium but not included                | Included                                       |
| Cyanide                        | Not measured                                       | Included                                       |
| Sulfide                        | Measured below LOQ (10 mg/L), not included         | Included                                       |
| Endpoint of comparison         | Toxicity: HC50 (acute) neither assessment factors nor predicted discharge dilutions are considered | Risk: msPAF (<0.05) based on SSDs using acute toxicity data in the SB approach and chronic data from WET testing, encompassing predicted environmental dilution, ignoring loss processes (e.g., volatilization, oxidation and [bio]degradation) on discharge |

SB = substance-based; WET = whole-effluent toxicity; EC50/EC10 = 50% and 10% effective concentrations; TU = toxic unit; LOQ = limit of quantification; HC50 = 50% hazardous concentration; msPAF = multisubstance potentially affected fraction; SSD = species sensitivity distribution.
Similarly, ammonia was not identified in the present study as a constituent of concern at the time of sampling, whereas Parkerton et al. (2018) did. However, its contribution to the risk appeared to be minor in their study. In the present study, ammonium levels were measured in the effluents in the range 6.5–70 mg/L (Supporting Information, Table S10). When considered indicative for the levels of ammonia these are also not relevant in the present study and were therefore not included. Parkerton et al. (2018) also identified cyanide as a substance of concern but did not find a significant contribution of this substance to the risk.

Although not highly toxic (Table 3), short-chained organic acids (formic acid to hexanoic acid) were present in high concentrations in the produced water mixtures (Table 4); and the reported range for the sum of this group was 0.012–0.80 g/L. This resulted in a relatively high contribution of these constituents to the acute toxicity in the present study (Figure 3). Organic acids were not addressed by Parkerton et al. (2018) and are also not included in the OSPAR risk-based approach guidelines (OSPAR, 2014a, 2014b). In the field, they are expected to (bio)degrade and/or evaporate rapidly but will cause a response in acute WET tests at the prevailing condition with, for example, short exposure durations. All this being said, the toxic units shown in Figure 3 do allow for a comparison with the toxic units reported by Parkerton et al. (2018). In the present study, the TUspec value across all acute WET tests (75 in total) ranged from 1.5 to >100 (the 50% effect level was beyond the lowest test concentration of 1% effluent), with a median value of 16 (Supporting Information, Table S9). Parkerton et al. (2018) reported toxic units for their acute tests ranging from 1.0 to 62.5, with a median value of 5.4. In the present study, the geometric means of toxic unit were used to reflect the hazard of each effluent and ranged from 3 to 80, with a median of 14. In Parkerton et al. (2018), the geometric mean toxic unit for each effluent ranged from 2.9 to 10.3, with a median of 4.6, based on their acute data. Effluents from the present study were found to have a higher toxicity than those tested by Parkerton et al. (2018). This difference between the studies cannot be explained with the available data.

Parkerton et al. (2018) calculated substance-based toxic units that are considerably more conservative than the toxic units they measured in WET tests, despite testing sensitive species in (sub)chronic WET tests and basing the substance-based toxic unit on acute toxicity. In the present study, these are more in the same range.

The most striking difference between the present study and that of Parkerton et al. (2018) is the contribution of production chemicals. They included a limited set of production chemicals which were characterized chemically but eventually assessed their role to be negligible. In contrast, the present study identified them as a major contributor. The methodological differences identified earlier (Table 6) could be the cause of this discrepancy. Production chemicals, included differently in the two studies, have historically been constituents of concern on the Norwegian continental shelf (Smit et al., 2011); the present results are therefore in line with expectations. However, without actual measurements of production chemicals to support our observations, this remains uncertain.

**General considerations**

As indicated, the present study does not address exposure levels (i.e., dispersion of the discharge) and therefore does not quantify risk. Furthermore, precautionary assessment factors (i.e., to translate the laboratory-scale experimental toxicity data to field conditions) are omitted in the present study. This was done because such factors may differ per substance and approach (WET vs. substance-based). Introducing such variable factors would obscure true differences between the WET-based and substance-based approaches and conflict with the present objectives. Of course, when applied in risk management, the precautionary principle should be honored to account for uncertainty in risk assessment. When both WET and substance-based approaches are applied conjointly in risk assessment, one should be aware of differences in precautionary assumptions between the two approaches.

Exclusion of precautionary adjustment in the present study allowed direct comparison between the toxicity of the WET and substance-based approaches for the 25 offshore effluents on the Norwegian continental shelf. For at least 80% and 48% of the effluents the WET-based and substance-based HC50 values differ by less than a factor of 5 and 2, respectively (Table 5). The toxicity estimate of the WET and substance-based approaches (ratio HC50sub over HC50WET close to 1) is nearly identical for four samples (Ekofisk M, Grane, Snoerre A, and Statfjord C), when only the options that address production chemicals are considered (options 1–3; Supporting Information, Table S13), while a ratio of the HC50sub over HC50WET of >1 (WET-based toxicity greater than substance-based toxicity) occurred for at least 16 (64%) effluents and a ratio <1 (substance-based toxicity greater than substance-based toxicity) for at most 9 (36%) of the 25 effluents (Table 5). This difference between substance-based and WET-based HC50 values was significantly affected by the production chemical discharge data provided by the different operators but could also be affected by any of the factors described above. Overall, added production chemicals have a dominating contribution to toxicity estimates for most effluents where the HC50sub over HC50WET ratio was <1. Uncertainties in estimation (overestimation) of added production chemical concentrations in produced water effluents informing the substance-based approach may explain this and require further attention when improving the toxicity, hazard, and risk assessment of produced water discharges of Norwegian installations. Furthermore, the comparison of toxicity indicators also tells us whether all information included based on substances (substance-based hazard) adequately describes their contribution to the actual toxicity of the effluent. The present study has provided insight into significant data gaps and limitations, for example, in the quality of data, missing data, and indications of the potential presence of “unknown” substances and interactions being overlooked in the produced water
effluent that deserve more attention. The knowledge obtained from the present study can refine current environmental risk-assessment practice and improve interpretation of corresponding risk-level estimations.

Based on these discrepancies, it is recommended that OSPAR guidelines need to be more explicit on why and how the two lines of evidence need to be used. For instance, when determining acceptable risk levels of produced water discharges, precautionary assumptions are required; and prioritizing chemical groups or installations for mitigating measures might need more realistic risk estimates.

More generally, acute WET assessed with only three marine species does not adequately reflect long-term (chronic) effects to the full ecosystem’s species assemblage. The substance-based approach does not adequately address unknown constituents or interactions. Neither approach can therefore be seen as ground truth. This was also recognized by Smit et al. (2020). Ideally, both approaches are applied periodically, such that (preferably downward) trends in the toxicity, hazard, and risk indicators over time can be observed. The evidence will be more compelling when the two indicators show similar trends over time. When the WET-based and substance-based approaches show contrasting trends, information from both lines of evidence (effluent composition, predicted toxicity, and measured WET) can be used to postulate and investigate potential causes.

CONCLUSIONS AND RECOMMENDATIONS

In general, the present study shows that the substance-based approach does not allow for a highly precise estimate of the toxicity in comparison with ground truth reflected by WET. The acute toxicity of the WET and substance-based approaches is nearly identical for four effluents, while for eight installations the substance-based toxicity is overestimated, most likely because of inaccurate calculation of discharge concentrations for production chemicals. For the remainder of the installations (13), the substance-based toxicity is underestimated, most likely caused by the incomplete chemical characterization of the effluent. For instance, the presence of “unknown” substances might potentially contribute to the overall toxicity (e.g., naphthenic acids). However, the differences are generally less than a factor of 5. The gap between the two methods is not expected to be bridged because of inherent differences between the approaches, although there is still room for improvement, especially when it concerns their practical application and interpretation in a legislative context, where differences between the approaches should be addressed more explicitly. We advise applying both approaches periodically and comparing trends over time between them.

A more accurate prediction of the concentrations of production chemicals in produced water effluents will certainly strengthen the substance-based approach because this group frequently has a dominating contribution to the toxicity of the produced water discharge, and concentrations are estimated from usage and conservative assumptions (e.g., for surfactants).

It is also shown that operators have an effect on how well substance-based toxicity estimates match with WET-based toxicity. Because this is most likely caused by differences in the way operators estimate discharge concentrations of the production chemicals, this process should be further harmonized. A concrete improvement identified in the present study results from the fact that the difference between the WET-based and substance-based toxicity was smallest when using the monthly concentration estimates for the production chemicals (70% of the variance could be significantly explained). However, when implemented in regulations, one may not be interested in monthly snapshots of risks in time but rather cumulative risks reflected by yearly averages. In addition, the role of production chemicals may change when other shortcomings (e.g., addressing unknown constituents and interactions) are also addressed. Such intents and best practices should also be made explicitly part of the OSPAR guidelines because they are currently ambiguous in that respect.

Supporting Information—The Supporting Information is available on the Wiley Online Library at https://doi.org/10.1002/etc.5414.

Acknowledgment—Norsk Olje & Gass (Norwegian Oil and Gas Association) is acknowledged for funding the research presented in the present study. The member companies for delivery of input data used in the analysis InteRTEK Westlab and Baker Hughes are acknowledged for conducting WET tests; Sintef MOLAB is acknowledged for chemical characterization of water samples. We gratefully acknowledge U. Musdalslien (Equinor) for her support in collecting and quality assurance of input data used in the analysis from all operators on the Norwegian continental shelf. Four anonymous reviewers are acknowledged for their constructive feedback which helped improve the quality of this article.

Author Contributions Statement—Pepijn de Vries: Conceptualization; Data curation; Formal analysis; Methodology; Visualization; Writing—original draft; Writing—review & editing. Robbert Jak: Conceptualization; Data curation; Funding acquisition; Writing—review & editing. Tone Frost: Conceptualization; Data curation; Resources; Writing—original draft; Writing—review & editing.

Data Availability Statement—Data, associated metadata, and calculation tools are available as supporting information.

REFERENCES

Aas, N., Knudsen, B., Sæten, J. O., & Nordstad, E. (2002, March 20–22). Mass balance of production chemicals [Conference presentation]. SPE International Conference on Health, Safety and Environment in Oil and Gas Exploration and Production. Kuala Lumpur, Malaysia.

Adams, M. S., Stauber, J. L., Binet, M. T., Molloy, R., & Gregory, D. (2008). Toxicity of a secondary-treated sewage effluent to marine biota in Bass Strait, Australia: Development of action trigger values for a toxicity monitoring program. Marine Pollution Bulletin, 57(6), 587–598. https://doi.org/10.1016/j.marpolbul.2007.12.012
Aldenberg, T., Jaworska, J. S., & Traas, T. P. (2002). Normal species sensitivity distributions and probabilistic ecological risk assessment. In L. Posthumus, G. W. Suter, & T. P. Traas (Eds.), Species sensitivity distributions in ecotoxicology (pp. 49–102). CRC.

Bartlett, M. S., & Fowler, R. H. (1937). Properties of sufficiency and statistical tests. Proceedings of the Royal Society of London. Series A—Mathematical, Physical and Engineering Sciences, 160(901), 268–282. https://doi.org/10.1098/rspa.1937.0109

Bertheussen, A., Kvalvåg, Ø., & Havre, T. E. (2021). Quantification of naphthenic acids in produced water (NOG-0324-PR-001 Rev 02). Worley Origo Process.

Beyer, J., Goksøyr, A., Hermann, D. Ø., & Klungsøyr, J. (2020). Environmental effects of offshore produced water discharges: A review focused on the Norwegian continental shelf. Marine Environmental Research, 163, Article 105155. https://doi.org/10.1016/j.marenvres.2020.105155

Brent, R. P. (1973). Algorithms for minimum without derivatives. Prentice-Hall.

Brown, J. S., Sauer, T. C., Wade, M. J., & Neff, J. M. (1992). Chemical and toxicological characterization of produced water fresh extracts. In J. P. Ray & F. R. Engelhard (Eds.), Produced water: Technological/environmental issues and solutions (pp. 113–131). Springer.

Clark, C. E., & Veil, J. A. (2009). Produced water volumes and management practices in the United States (ANL/EVS/R-09/1). US Department of Energy.

Crommentuijn, T., Polder, M. D., & Van de Plassche, E. J. (1997). Maximum permissible concentrations and negligible concentrations for metals, taking background concentrations into account (601501001). National Institute of Public Health and the Environment.

Denton, D., Fox, J., Fulk, F., Greenwald, K., Narvaez, M., Norberg-Bohmer, T., & Klungsøyr, J. (2020). Environmental effects of offshore produced water discharges: A review focused on the Norwegian continental shelf. Marine Environmental Research, 163, Article 105155. https://doi.org/10.1016/j.marenvres.2020.105155

Brent, R. P. (1973). Algorithms for minimization without derivatives. Prentice-Hall.

Brown, J. S., Sauer, T. C., Wade, M. J., & Neff, J. M. (1992). Chemical and toxicological characterization of produced water fresh extracts. In J. P. Ray & F. R. Engelhard (Eds.), Produced water: Technological/environmental issues and solutions (pp. 113–131). Springer.

Clark, C. E., & Veil, J. A. (2009). Produced water volumes and management practices in the United States (ANL/EVS/R-09/1). US Department of Energy.

Crommentuijn, T., Polder, M. D., & Van de Plassche, E. J. (1997). Maximum permissible concentrations and negligible concentrations for metals, taking background concentrations into account (601501001). National Institute of Public Health and the Environment.

Denton, D., Fox, J., Fulk, F., Greenwald, K., Narvaez, M., Norberg-Bohmer, T., & Klungsøyr, J. (2020). Environmental effects of offshore produced water discharges: A review focused on the Norwegian continental shelf. Marine Environmental Research, 163, Article 105155. https://doi.org/10.1016/j.marenvres.2020.105155

Brent, R. P. (1973). Algorithms for minimization without derivatives. Prentice-Hall.

Brown, J. S., Sauer, T. C., Wade, M. J., & Neff, J. M. (1992). Chemical and toxicological characterization of produced water fresh extracts. In J. P. Ray & F. R. Engelhard (Eds.), Produced water: Technological/environmental issues and solutions (pp. 113–131). Springer.

Clark, C. E., & Veil, J. A. (2009). Produced water volumes and management practices in the United States (ANL/EVS/R-09/1). US Department of Energy.

Crommentuijn, T., Polder, M. D., & Van de Plassche, E. J. (1997). Maximum permissible concentrations and negligible concentrations for metals, taking background concentrations into account (601501001). National Institute of Public Health and the Environment.

Denton, D., Fox, J., Fulk, F., Greenwald, K., Narvaez, M., Norberg-Bohmer, T., & Klungsøyr, J. (2020). Environmental effects of offshore produced water discharges: A review focused on the Norwegian continental shelf. Marine Environmental Research, 163, Article 105155. https://doi.org/10.1016/j.marenvres.2020.105155

Brent, R. P. (1973). Algorithms for minimization without derivatives. Prentice-Hall.

Brown, J. S., Sauer, T. C., Wade, M. J., & Neff, J. M. (1992). Chemical and toxicological characterization of produced water fresh extracts. In J. P. Ray & F. R. Engelhard (Eds.), Produced water: Technological/environmental issues and solutions (pp. 113–131). Springer.

Clark, C. E., & Veil, J. A. (2009). Produced water volumes and management practices in the United States (ANL/EVS/R-09/1). US Department of Energy.

Crommentuijn, T., Polder, M. D., & Van de Plassche, E. J. (1997). Maximum permissible concentrations and negligible concentrations for metals, taking background concentrations into account (601501001). National Institute of Public Health and the Environment.
Produced water toxicity vs. substance based estimates—Environmental Toxicology and Chemistry, 2022;41:2285–2304

Melbye, A. G., Brakstad, O. G., Hokstad, J. N., Gregersen, I. K., Hansen, B. H., Booth, A. M., & Tollefsen, K. E. (2009). Chemical and toxicological characterization of an unresolved complex mixture-rich biodegraded crude oil. Environmental Toxicology and Chemistry, 28(5), 1815–1824. https://doi.org/10.1897/08-545.1

National Institute of Public Health and the Environment. (2005). e-Tox database.

Neff, J., Lee, K., & DeBois, E. M. (2011). Produced water: Overview of composition, fate, and effects. In K. Lee & J. Neff (Eds.), Produced water, environmental risks and advances in mitigation technologies. Springer.

Produced water 2: Environmental issues and solutions in produced water. Environment and Health Chemistry, 1(1), 76–85. https://doi.org/10.1016/s11356-09-0101-0

Neff, J., Lee, K., & DeBois, E. M. (2011). Produced water: Overview of composition, fate, and effects. In K. Lee & J. Neff (Eds.), Produced water, environmental risks and advances in mitigation technologies. Springer.

Norwegian Oil and Gas. (2013). Recommended guidelines for sampling and analysis of produced water. https://www.norskeljepoggass.no/contentassets/799e40a4a08f42dd2c2cd1flf948a8bdf8835-norwegian-oil-and-gas-recommended-guidelines-for-sampling-and-analysis-of-produced-water-rev-1-11.12.2012.pdf

Oetjen, K., Giddings, C. G. S., McLaughlin, M., Nell, M., Blotveugel, J., Helbling, D. E., & Higgins, C. P. (2017). Emerging analytical methods for the characterization and quantification of organic contaminants in flowback and produced water. Trends in Environmental Analytical Chemistry, 15, 12–23. https://doi.org/10.1016/j.teac.2017.07.002

Offshore Industry Committee. (2008). Meeting of the Offshore Industry Committee (OIC): Summary record (OIC 08/13/1–E). OSPAR Commission.

OSPAR. (2013). OSPAR list of substance used and discharged offshore which are considered to pose little or no risk to the environment (PLONOR) (OSPAR Agreement 2013–06). https://www.ospar.org/documents/32939

OSPAR. (2014a). Background document for establishment of a list of predicted no effect concentrations (PNECs) for naturally occurring substances in produced water. PNECs list for use in the risk based approach to the management of produced water discharged from offshore installations (OSPAR Recommendation 2012/25).

OSPAR. (2014b). OSPAR 12/221, annex 19. OSPAR guidelines in support of recommendation 2012/25 for a risk-based approach to the management of offshore produced water discharges from offshore installations (OSPAR Agreement 2012–7).

OSPAR. (2020). OSPAR guidelines for completing the harmonised offshore chemical notification format (HOCNF) (OSPAR Agreement 2012/05).

Parkerton, T. F., Bok, M., Ireland, A. W., & Prosser, C. M. (2018). An evaluation of cumulative risks from offshore produced water discharges in the Bass Strait. Marine Pollution Bulletin, 126, 610–621. https://doi.org/10.1016/j.marpolbul.2017.10.003

Pessala, P., Schultz, E., Nakari, T., Joutti, A., & Herve, S. (2004). Evaluation of an unresolved complex mixture (UCM) of produced water-revisited. Chemosphere, 55(8), 1321–1328. https://doi.org/10.1016/j.chemosphere.2004.06.002

Smit, M. G. D., Hayes, S., Pelz, O., Arnold, R., & Brown, W. (2020, July 24). New produced water risk-based approach guidance [Virtual conference presentation]. SPE International Conference and Exhibition on Health, Safety, Environment, and Sustainability.

Smit, M. G. D., Jak, R. G., Rye, H., Frost, T. K., Singsaas, I., & Karman, C. C. (2008). Assessment of environmental risks from toxic and nontoxic stressors; a proposed concept for a risk-based management tool for offshore drilling discharges. Integrated Environmental Assessment and Management, 4(2), 177–183. https://doi.org/10.1897/iaem.2007-036.1

Somerville, H. J., Bennett, D., Daventry, N. J., Holt, M. S., Lynes, A. C., Mahieu, A., & Wilkinson, T. G. (1987). Environmental effect of produced water from North Sea oil operations. Marine Pollution Bulletin, 18(10), 549–558. https://doi.org/10.1016/0025-326X(87)90359-9

Sørensen, L., McCormack, P., Aitkin, D., Robson, W. J., Booth, A. M., Faksness, L.-G., & Starseth, T. R. (2019). Establishing a link between composition and toxicity of offshore produced waters using comprehensive analysis techniques—A way forward for discharge monitoring. Science of the Total Environment, 694, Article 133682. https://doi.org/10.1016/j.scitotenv.2019.133682

Thurman, E. M., Ferrer, I., Rosenblum, J., Linden, K., & Ryan, J. N. (2017). Identification of polypropylene glycols and polyethylene glycol carboxylates in flowback and produced water from hydraulic fracturing. Journal of Hazardous Materials, 323, 11–17. https://doi.org/10.1016/j.jhazmat.2016.02.041

Tibbetts, P. J. C., Buchanan, I. T., Gawel, L. J., & Large, R. (1992). A comprehensive determination of produced water composition. In J. P. Ray & F. R. Engelhardt (Eds.), Produced water: Technological/environmental aspects. CRC.

Tollefsen, K. E., Geerds, G., & Gjelsvik, A. (1999). Assessment of complex industrial effluents in The Netherlands using a whole effluent toxicity (WET) approach. Water Science and Technology, 39(10–11), 55–61. https://doi.org/10.2166/wst.1999.0630

US Environmental Protection Agency. (1985). Methods for measuring the acute toxicity of effluents to freshwater and marine organisms (EPA600/8-85/013).

US Environmental Protection Agency. (1993). Methods for measuring the acute toxicity of effluents and receiving waters to freshwater and marine organisms (EPA600/4-90027P).
US Environmental Protection Agency. (1994). Short-term methods for estimating the chronic toxicity of effluents and receiving waters to marine and estuarine organisms (EPA-600-4-01-003).

US Environmental Protection Agency. (1995). National policy regarding whole effluent toxicity enforcement.

US Environmental Protection Agency. (2022). Interspecies correlation estimation—WebICE calculator. https://www3.epa.gov/webice/iceCalculate.html?Surrogate=Phaeodactylum+tricornutum&Predicted=Skeletonema+costatum

Van Dam, R., & Chapman, J. (2001). Direct toxicity assessment (DTA) for water quality guidelines in Australia and New Zealand. Australasian Journal of Ecotoxicology, 7, 175–198.

Verbruggen, E. M. J. (2012). Environmental risk limits for polycyclic aromatic hydrocarbons (PAHs); for direct aquatic, benthic, and terrestrial toxicity (607711007/2012). National Institute of Public Health and the Environment.

Wang, B., Wan, Y., Gao, Y., Yang, M., & Hu, J. (2013). Determination and characterization of oxy-naphthenic acids in oilfield wastewater. Environmental Science & Technology, 47(16), 9545–9554. https://doi.org/10.1021/es401850h

Wheeler, J. R., Grist, E. P. M., Leung, K. M. Y., Morritt, D., & Crane, M. (2002). Species sensitivity distributions: Data and model choice. Marine Pollution Bulletin, 45(1), 192–202. https://doi.org/10.1016/S0025-326X(01)00327-7

Wilde, M. J., West, C. E., Scarlett, A. G., Jones, D., Frank, R. A., Hewitt, L. M., & Rowland, S. J. (2015). Bicyclic naphthenic acids in oil sands process water: Identification by comprehensive multidimensional gas chromatography-mass spectrometry. Journal of Chromatography A, 1378, 74–87. https://doi.org/10.1016/j.chroma.2014.12.008

Zhang, X. J., Qin, H. W., Su, L. M., Qin, W. C., Zou, M. Y., Sheng, L. X., & Abraham, M. H. (2010). Interspecies correlations of toxicity to eight aquatic organisms: Theoretical considerations. Science of the Total Environment, 408(20), 4549–4555. https://doi.org/10.1016/j.scitotenv.2010.07.022

Zheng, J., Chen, B., Thanyamanta, W., Hawboldt, K., Zhang, B., & Liu, B. (2016). Offshore produced water management: A review of current practice and challenges in harsh/Arcic environments. Marine Pollution Bulletin, 104(1), 7–19. https://doi.org/10.1016/j.marpolbul.2016.01.004